PROJECT THOR

Project Thor

Team Members:

Name	Email
Adonay Pichardo	apichardo2019@my.fit.edu
Jared Blanco	jblanco2018@my.fit.eud
Josh Temel	jtemel2018@my.fit.edu
Luke Bonenberger	lbonenberger2018@my.fit.edu

Faculty Advisor:

Name	Email
Sid Bhattacharyya	sbhattacharyya@fit.edu

Client:

Name	Email
Amitabh Nag	anag@fit.edu

Date(s) of Meeting(s) with Client:

Date	Topic
September 3, 2021	• Better understanding the dataset

Meetings with Client:

Up to the time of the writing of this document, there have been no meetings with the client to further discuss the project development process for this semester. However, there will be a meeting planned within the near future (1515 on 18JAN2022).

Goals & Motivation:

Motivation

Today's society has grown more and more dependent on the use of computers and this trend doesn't seem to be slowing down anytime soon. Along with this increased use of computers comes an increased importance for the data being transmitted and stored by these computers to be secure. Designing the methods used by computers to create encryption keys that are practically immune to attacks has proven to be a very complicated (if not impossible) challenge that computer scientists still struggle with today and random numbers lie at the root of the problem.

Computers primarily use pseudo-random numbers which are created using a mathematical formula that produces a deterministic, periodic sequence of numbers, which is completely dependent on the initial state or seed. However, today's computing systems are powerful enough to predict the random numbers being generated by PRNGs within a reasonable amount of time. This has placed an increased emphasis on finding a truly non-deterministic process which we believe may be found in the use of natural phenomena to inject true randomness.

Goal(s)

Our overall goal for this project is to develop a web application that allows users to generate encryption keys with high-entropy, truly random data that is gathered from the study of lightning strikes.

This semester we plan to soon finish out development and refinement process and move into the testing and verification phase. This will include unit testing, peer review testing, and bug fixing. Additionally, we also plan to begin the project's reporting phase in which we will begin developing documentation (write ups) that will be posted on the web application, create a poster for the showcase, and write up our results from all the testing.

Approach:

Key Features

Feature 1: Automated Data Transfer

The original data set will be provided to us via ASCII text files. Much of the data within this file is unusable due to its predictable nature. Only a few data types will be potentially useful for this project. Therefore, to make them more easily accessed and usable, the web application will automatically parse the relevant data from the ASCII files and add it into our working database from which the web application will be able to generate random numbers.

Feature 2: Random Number Generator

The web application, using the random data gathered in the database, will be able to algorithmically generate random numbers. These random numbers will be theoretically less predictable than those produced by pseudo-random number generators currently being used as the seed data being pulled from the database is considered to be random by nature.

Feature 3: Create Encryption Keys

Users will be able to use our web application to create cryptographic keys. These keys will be generated using the random number generator mentioned above and should be very secure due to the fact that they were developed using numbers and data sets that have a higher degree of entropy (randomness) than those typically employed by other modern cryptographic algorithms.

Feature 4: File Upload for Encryption

Users will be able to use our web application to upload files from their local machines to be encrypted using modern encryption algorithms. The specific algorithms to be used are yet to be decided, but the user will have the option of using at least one, and possible more using a drop down menu. Additionally, the specific file formats we will accept are pdf and txt, but may also expand to include others.

Novel Features

Feature 1: Collected Data

The original data set that will be collected, sorted, and transferred into our database is a novel feature to our project as it will have a degree of randomness inherent to it due to it being based directly from lightning. A natural phenomena that is highly abundant yet very unpredictable.

Technical Challenges

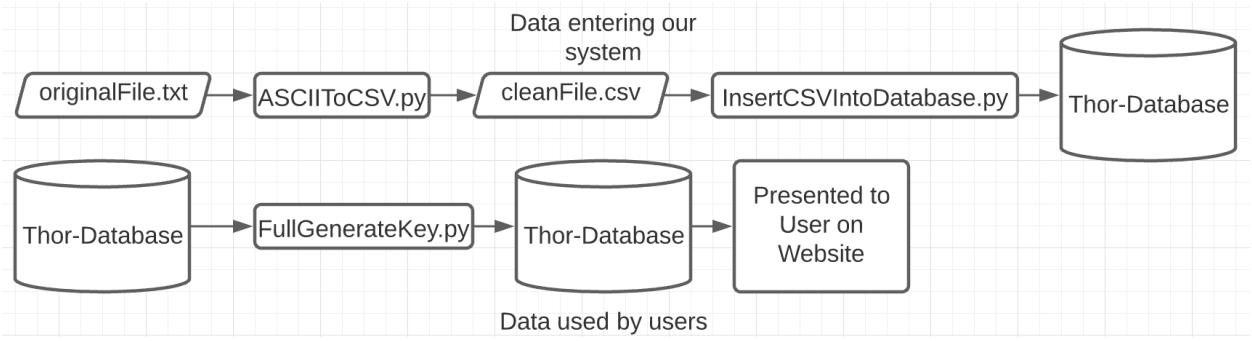
This project is relatively simple in the sense that there is not a lot to the deliverables (web application and database), but it does come with a lot of technical challenges for our team to work through. Many of these challenges will be found in data analysis. However, there are also some computer science challenges that will need to be met as well. These are listed below:

1. Automating raw data entry into the database. (Completed)
2. Creating a random number generator using multiple parameters drawn from the data set. (In Progress)
3. Integrating a strong encryption algorithm that uses our random number generator. (In Research)
4. Learning how to use the chosen website backend option. (Completed)
5. Learning about integrating backend to database. (Completed)
6. Prevent recreation of previously calculated numbers. (In Progress)

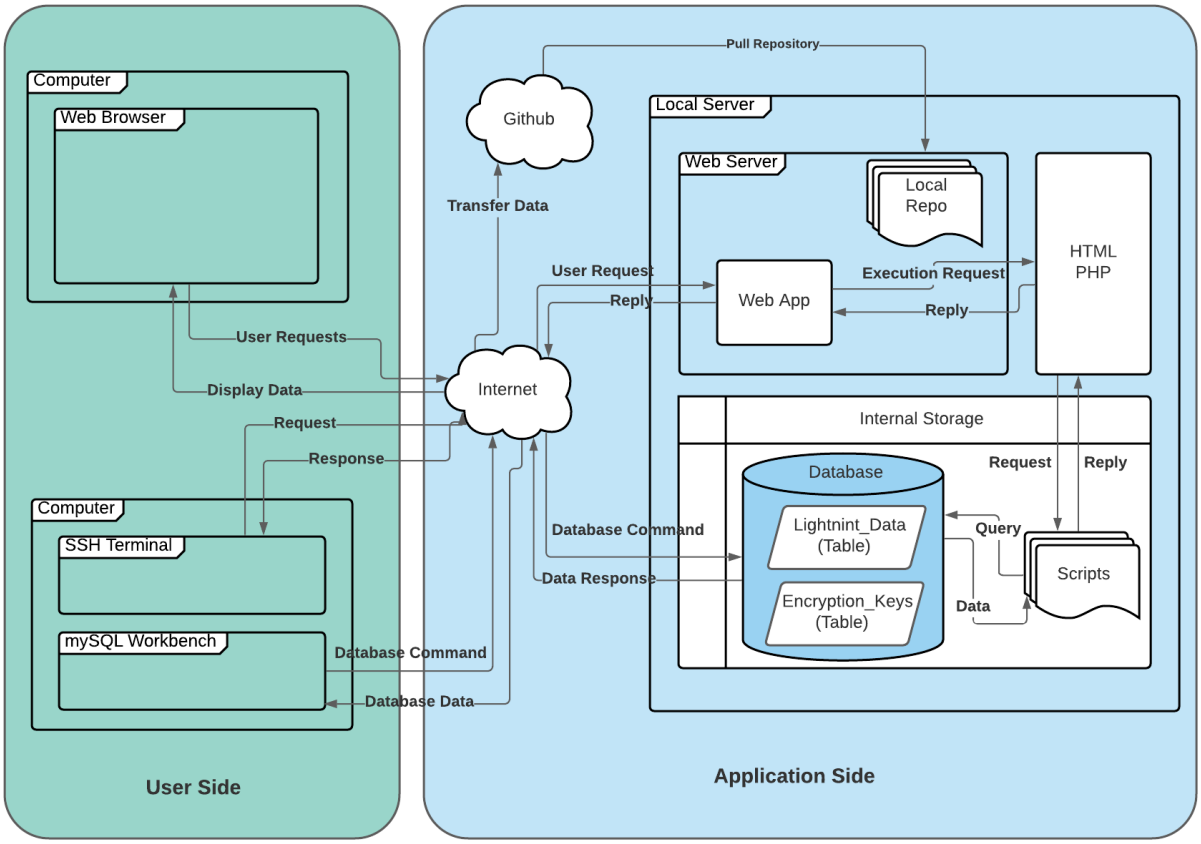
Although some are now completed, these challenges continue to adapt and come back around to again test the team. Additionally, with the generation of keys now being derived from equally weighted fields gathered from the automated lightning data import, the ability to generate an exponential amount of additional keys per lightning strike is now a viable option. This is possible by randomly generating keys by combining different fields from different lightning strikes all within a time period. The challenge with this algorithm is the lack of documentation on implementation, currently our groups is following a similar structure to the scientific method to derive this algorithm while utilizing various entropy calculations to support our findings

Design

Data Flow Diagram illustrates lightning data going through clean up, parsing, storage, use, and finally delivering.



Architecture Diagram illustrates the current state of the system(s).



Evaluation

There are many ways to measure the success of our project. However, we are going to focus on two primary measures: one will be a series of metrics that determine the randomness of the numbers being generated and another will be a measure of the customer satisfaction based on how well our project met the requirements. Ultimately, the later will be determined by our final advisor grade, but in the testing phase of the project, acceptance testing will also be considered.

Progress Summary

Module/Feature	% Completed	To Do
Automated Data Transfer	100%	
Random Number Generator	90%	• Generate more

		permutations from existing dataset. Fix/refine algorithm (track already generated permutations so as to not duplicate). • Add data refinement algorithm to the algorithm.
Create Encryption Keys	95%	Currently only reading the first line from the database. Needs to be debugged to read a new line every time.
File Upload for Encryption	0%	Needs started.

Milestone 4 (Feb 14):

Itemized tasks:

A. Update Demos
B. Update Documentation
C. Create Problem Statement Documentation for Web Application
D. Create Solution Documentation for Web Application
E. Drafting Web Application Testing Documentation

Milestone 5 (Mar 21):

Itemized tasks:

A. Perform Requirements Verification Testing
B. Analyze Results of Web Application Testing
C. Perform Metrics Testing

Milestone 6 (April 18):

Itemized tasks:

A. Complete Final Showcase Poster
B. Complete Final Showcase Ebook Page
C. Complete Final Demo Update
D. Complete Final Documentation Update
E. Complete Documentation for Future Teams Interesting in Project

Task Matrix for Milestone 4

Task	Adonay	Jared	Josh	Luke
Update Demos	25%	25%	25%	25%
Update documentation	25%	25%	25%	25%
Create and add detailed explanation page to the problem section of the web app	0%	0%	50%	50%
Create draft of solution explanation for the solution section of the web app	0%	0%	50%	50%
Refine the data collaboration algorithm to ensure the production of non-duplicate keys with command safeguard	50%	50%	0%	0%

techniques				
Begin outlining a more detailed test plan / system for testing the functionality of the web application so that we can test ourselves and allow the class to find and report bugs.	50%	25%	25%	0%

Task Descriptions:

Task 1: Update Demos

Currently our group has live demos of our web application, key generator, and automated data transfer. At the end of this milestone, these demo will need to updated to include the latest progress made. Additionally, we hope to add another demo showing how our group is generating more data from our existing data.

Task 2: Update Documentation

Our project documentation had been treated as ‘live’ documents that are ever changing. As a result, they need to be updated to reflect the latest project design, test plan, system architecture, etc.

Task 3: Problem Documentation

One part of our web application deliverable is including an explanation section that explains to those who want to find out more about the project the problem being solved. This task’s goal is to create that documentation with a detailed description of the problem statement. It will then be published on the web application. This should expand on the existing problem statement published on the homepage and simple expand into more technical detail.

Task 4: Draft Solution Explanation

This is similar to task three, but the topic of the write up should be our groups solution to the problem. Again, it should include technical detail and go beyond our existing ‘elevator pitch’ solution published on the web application’s homepage.

Task 5: Data Generator Refinement

Now that the lightning fields are now equally weighted, the ability to randomly combine these fields within each other is now an effective way of generating these keys. This algorithm is similar to how linux combines multiple random factors to determine a random key. The current limitation with this implementation is a lack of testing on the possibility of duplicate numbers. In theory this has a 1/(1276^4) chance of occurring but this is dependent on the algorithm being successful in randomly choosing these keys.

Task 6: Draft Test Plan

Our group already has an existing test plan document that addresses our plan for completing the requirement testing for the project. However, we'd also like to open testing up to our class and peers in an attempt to find bugs in our web application. We've discussed making it a competition and creating a bug report template and reporting system. The goal of this task is to create a document that formalizes what we've discussed so that we can execute successfully.

Approval from Faculty Advisor

"I have discussed with the team and approve this project plan. I will evaluate the progress and assign a grade for each of the three milestones."

Signature: on file_____ Date: _____