Identifeye

Zak Fahey, Mitchell Haas, Brendan Fisher

Who are we?

- Zak Fahey - Computer Science
  faheyzt@mail.uc.edu
- Mitchell Haas - Computer Science
  hassm3@mail.uc.edu
- Brendan Fisher - Computer Science
  fisherb8@mail.uc.edu
- Raj Bhatnagar - Advisor
  bhatnark@ucmail.uc.edu

Purpose and Goal

- Changing your identity online can be as simple as flipping a switch.
- By using VPNs users can circumvent bans, or post as anyone online.
- People can also get into accounts who aren't supposed to access them.
- We use data analysis to identify users who are changing their online identity.
- Essentially, we want to map online identities to real people, even if the data is obscured.

Merits, and Novelty

- There is no third party plug in for identifying users with multiple accounts
- The data from the third party is anonymized to secure user information, and the data maintains its ability to be analyzed
- We worked with a third party to get real world video game data from their server
- This application solves problems faced by many communities such as video game servers, and online reviews. Anywhere where people have a reason to create multiple online accounts
- Identifying user's is a research field in computer science. It has merits both as a research field and multiple real world applications

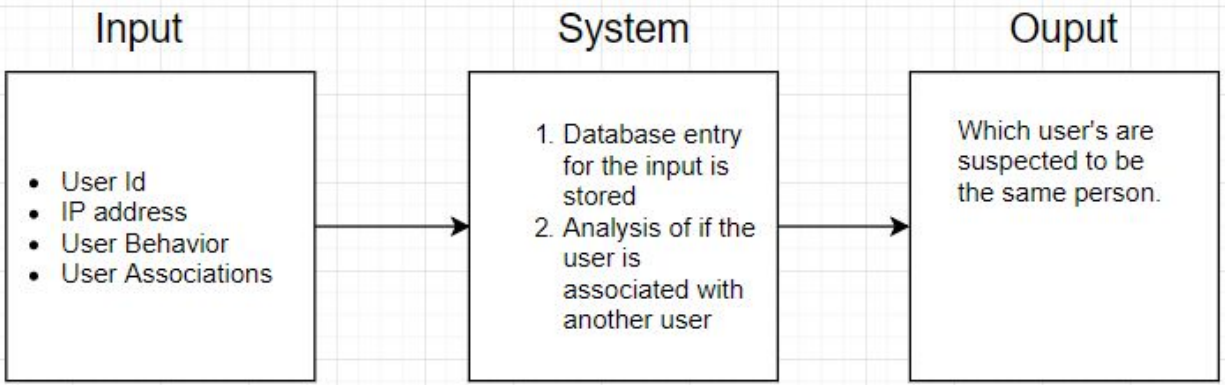
Broader Impacts

- Ban invasion in games, social media, and review sites is very common
- In video games having a way to keep to toxic members who are harassing the community with scams, or inappropriate actions/ language will help keep all ages in the community safe from negative activities
- Reviews with multiple accounts can artificially inflate a products review and perceived value which causes customers to waste money on a bad product when the review system is unreliable.
- Multiple accounts on social media can spread fake news or panic to community at large. finding these accounts before they can cause damage is essential to keep well informed.
- Identifying users with multiple accounts who are performing malicious activities can help keep these and other communities positive and safe

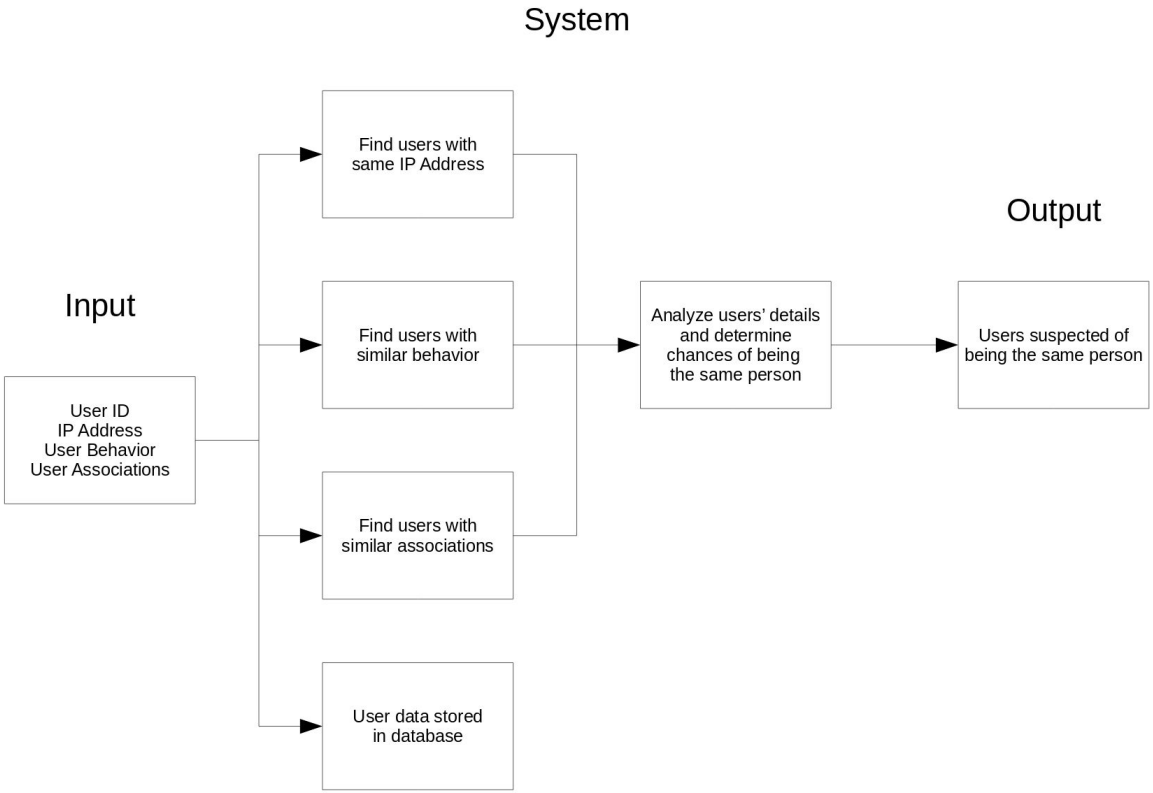
Computing and Software Design

- There are three main components to Identifeye
- The Anonymizer which is a script that takes in source data and remove any sensitive information by hashing it, while still maintaining the data's ability to be analyzed.
- The Interfacing library which is a C# library that is used to communicate with the analysis server.
- The Analysis server which is the main Python application that takes in the anonymized data, stores in a graphical data structures, runs the data analysis, and outputs the connection percentage between banned user accounts and other accounts in the data.

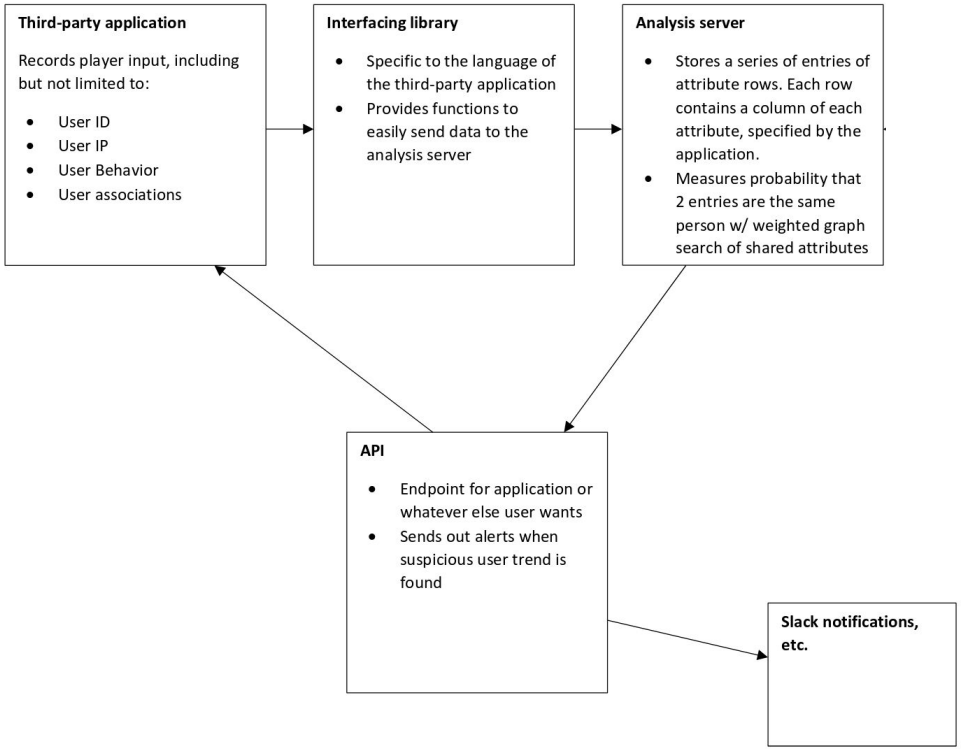
Design Diagram Level 0



Design Diagram Level 1



Design Diagram Level 2



Technologies

- There is a variety of technologies learned and used in the development of Identifeye. Below is a detailed list of some the technologies involved.
- Python is the language used for the large majority of the project it was used in the anonymizer and the analysis program.
- C# was the main language used for the interfacing library.
- GRPC was used in the interfacing library and is an open source remote procedure call framework.
- Just as import were the communication technologies used used such as GroupMe and Discord for group communication as well as Google Drive and Github for file sharing to keep the most up to date versions of project available to all group members.

Milestones

- There was a lot of significant milestones in the development of this project. Below is a list of the milestones from winter break to the senior design expo.
- Explore the viability of text analysis in the project. We found out we could not get the comments data from the real world data set so this idea was deemed non-viable. November 2018 - January 2019
- Develop the anonymizer program. January - March 2019
- Develop the interfacing library. January - March 2019
- Develop the core of the analysis server in Python. January - March 2019
- Test the application with data and verify that it works. March 2019
- Obtained real world data from the video game server. Mid April 2019
- Senior Design Poster developed. March - April 2019
- Finalize documentation for senior design. April 2019

Results

- We achieved a lot while developing this project, but perhaps the most valuable achievement was learning what it takes to design and development a project.
- Completed the planning stage for the project.
- Adjusted the scope of the project.
- Developed an anonymizer program.
- Developed an interfacing library.
- Developed an analysis program.
- Completed a senior design poster and expo.
- Obtained real world data from a video game server.
- Completed a variety of senior design documentation.

Challenges

- We faced a variety of challenges throughout this project. Listed below are some of the challenges we faced.
- Having a full load of classes, while developing a project that ideally would have multiple full time developers.
- Missing some masters level courses, and experience in machine learning which limits the scope of the project.
- Searching for jobs during winter break, and both semesters.
- Relying on a third party company to provide real world video game data that we final obtained after months of communication that started in the fall of 2018.
- Learning new technologies for programming and designing components of the project. As well as learning new programming languages.