**Introduction**

Background and Challenge

Social networks are an integral part of life for most people. People make friends on Facebook and Twitter, find jobs on LinkedIn. Concerning Twitter, it is the second-largest social network, has more than 1 billion registered users and 330 million monthly active users. It proves that people rely heavily on social networks. They are very willing to share their daily life on those platforms. However, the existence of social networks inevitably forces people to face privacy issues [1]. For professionals, except for the usual information leakage, also have the risk of leaking the content of tweets. Due to hundreds of millions of people using social networking platforms, some people may think that they will not be discovered when unscrupulous talking about sensitive information related to the company, such as salary, revealing the idea of job-hopping, complaining about colleagues, company and employer. However, the recommendation system of social networks would constantly recommend users to workmates or friends of friends. So many ways can lead to leakage of improper comments. Employees who post sensitive information might face pressure from their companies if their inappropriate statements are discovered by companies. Some professionals use non-public social accounts to send posts. But if a post includes any name, location, or company can also be easily recognized by people around them. Therefore, posting sensitive information about the company on social platforms without being discovered needs to be studied.

In addition, people are increasingly suffering from employment pressure due to the COVID-19 pandemic, leaving many job seekers in a weaker position. From March to April 2020, the unemployment rate rose from 4.4% to 14.7%. Not only because of the depressed market, information inequality between job seekers and companies is also a crucial reason. Companies always want to select fewer demanding employees, such as accepting lower pay or working overtime without complaint. Corporations can sift through resumes, compare candidates and choose the best value for money. And job seekers are stuck waiting again and again. Even if he does land a job, he may encounter a poor working environment but doesn't realize it until he starts working. There is also information inequality in terms of salary, which job seekers are most interested in. Newly graduated students have no idea of the salary level of different cities, companies and different positions. So they may suffer loss in the salary aspect when looking for a job. Besides, people seldom talk about salary. If someone gets a disproportionate amount of money, they may not find out. Some efforts have been made to address this situation. In 2016, LinkedIn launched Salary Insights, which is a system that gathers salary information from members to provide salary insights to job seekers. In 2017, Kenthapadi proposed the LinkedIn compensation product, which helps people calculate their earning potential by collecting a large amount of data [3]. These are rewarding and innovative products. But at present, the coverage of these products is low and is not suitable for many regions and positions. Also, many job seekers are looking for information other than salaries, such as working environment and intensity. Therefore, it is an urgent need for a secure social platform, which allows people to talk freely about their careers without the risk of being discovered.

Solution

SafeChat is a Web-based social networking platform that detects sensitive information in real-time. In SafeChat, authenticated users can post sensitive information without fear of being discovered by their bosses. Users are required to fill out basic information and the company they belong to when signing up for the social networking platform. Unlike other social networks, SafeChat allows people to choose to post anonymously, allowing them to hide their identities while posting. In addition, people can choose to encrypt posts through the Base64 algorithm. If the user encrypts the content, he is also anonymous automatically. All encrypted information cannot be seen by other employees of the same company. When other people read this post, they can click the decryption button next to it, then navigate into the decryption interface. The decryption page contains decrypted contents of tweets, which is forbidden to copy. The page is also full of watermarks with the reading user's real name, which helps prevent people from taking screenshots or photos to spread it. When users publish information, the system will automatically detect sensitive information. If there is sensitive information, NER detection will be performed on it, with the main detection objects being name, organization, location and money. If the post contains sensitive information, but the NER tool does not detect identity information, the system prompts the user to remain anonymous. If sensitive information and entity information are both distinguished, the system prompts the user to encrypt it.

**Requirements**

Requirements gathering

Collect requirements through interviews at the beginning of the project. The purpose of the interview is to understand the behavioural characteristics and preferences of employees using social platforms. And their pain points in using social networks. All the questions should refine according to the interview purpose, and conservations should build around the work and social network. Users interviewed need to be screened in combination with SafeChat features, namely, people who like to use social networks and are already working or about to work. Divide users into core users and potential users. Core users are those who want to be active on workplace social platforms, and potential users are defined as the target user but not currently considered using SafeChat. The core users of this research object are two employees of Internet companies and an HR of an Internet company. Potential users are two graduates. Before the interview, different questions should be chosen for different types of interviewees, and each interviewer should answer about ten questions. During the interview, supplementary questions can be asked by adjusting the space flexibly. After the interview, collate all content for the requirements design of SafeChat and prioritize functions using Moscow. Interview questions are in the appendix.

Requirements List

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | Module | Description | MoSCoW | Implemented |
| 1 | Basic Function | User can register | M | Y |
| 2 | User can login | M | Y |
| 3 | System should encrypt users' passwords | S | Y |
| 4 | Users can upload profile pictures | S | Y |
| 5 | Users can modify personal information | M | Y |
| 6 | Send Posts | Users can send posts directly | M | Y |
| 7 | Users can insert emoticons in posts | C | Y |
| 8 | Users can send posts anonymously | M | Y |
| 9 | Users can send encrypted posts | M | Y |
| 10 | System must detect sensitive information in posts | M | Y |
| 11 | System must use NER tool detect posts | M | Y |
| 12 | System must suggest sending mode to users | M | Y |
| 13 | Read Posts | Users can view the avatar and name of the publisher | M | Y |
| 14 | Users can read all direct and anonymous posts | M | Y |
| 15 | Users can only view encrypted posts from employees of other companies | M | Y |
| 16 | User can like posts | S | Y |
| 17 | Decrypt Posts | Users can decrypt posts from employees of other companies | M | Y |
| 18 | System could place the watermark of the user's name and email on the decryption interface | C | Y |
| 19 | System must forbid users to copy content on the decryption page | S | Y |
| 20 | Comment Posts | Users can view post comments | M | Y |
| 21 | Users can comment on post | S | Y |
| 22 | Users can insert emojis into comments | C | Y |
| 23 | Users can make comment anonymously | S | Y |
| 24 | Users can encrypt their comments | W | N |
| 25 | Search Posts | User can find posts by entering keywords | S | Y |
| 26 | System should display the content and publisher name for non-anonymous posts | S | Y |
| 27 | System must display only the content of anonymous posts | M | Y |

**Implementation**

System Structure

