**DOCUMENTATION OF WORK DONE BY PARNA GOSWAMI**

Here is the step-by-step process of how things are going for the malicious URL detection part of the OSINT project:

* Data for malicious URL is collected from three sources: XForce, Kaggle, Phish tank.
* All the collected URLs are distributed into two files: only\_malicious\_urls.csv and only\_benign\_urls.csv
* Then I made a bloom filter in python using the inbuilt bloom-filter module of python and pickled it using the pickle module of python.
* This .pkl file is then stored in the python folder of our android studio.
* Then I wrote the UI of my app and used Chaquopy to integrate the python file with java.
* Now, since our java code can read the data in the bloom filter i.e., basically interact efficiently with the filter, so we will do the detection of the presence or absence of a particular URL in our filter.
* So, we input a URL and check if it is present or not in the spam URL dataset. If yes, then we show the user that the given URL is malicious else, we show that the input URL is not malicious.
* If the detected URL is malicious, the we advise the user to go to the XForce link, provided below, for confirmation.
* There is an extra feature i.e., there is a “Read SMS” button which reads the last SMS and detects the presence of URL in it and checks whether it is malicious or not.
* Then, I worked for the delta calculation of the model. We are calculating the delta based on last updating time of the database.
* We are detecting any insertion, updation and deletion from the database and accordingly our .pkl file is getting updated because of the infinite loop which we are running.
* So, to push the delta to the device, all we need to do is reload our app and then we are into the new and updated bloom filter.
* All the work which is going on here, needs no internet and everything is happening locally, in our device.