Batch Analysis of Network Security Monitoring Data

Renu Tiwari, Jiang Xing Kun

January 2021

1 Duties Description performed by Project Team Members

1.1 Renu Tiwari (UYFIND)

- Data preprocessing on Jupyter Notebook: Initial exploration of data to understand the fields and get summary statistics and visualizations using Python.

- Streaming .csv data using kafka into "streaming" topic.

- Show .csv data on UI localhost after kafka stream.

- Use Spark Consumer to read data and do filtration.

- Created function to filter on basis of only TCP protocol accepted and export data to mysql and mongoDB.

- Explore different databases.

- Created configuration setup for MongoDB and migrated from MySQL to MongoDB database.

- Conversion of JavaInputDStream to JavaRDD.

- Code repository management on GitHub: creating repository, reviewing pull requests and access management.

- Participating in project team meetings.

- Prepared latex report, duties and presentation material.

1.2 Jiang Xing Kun (R0F1R5)

- Loading data into wire-shark and export to .csv data.

- Setting up Kafka Connect to read .csv data into 'topic1' stream

- Setting up Spark Kafka Stream Connection.

- Do filtration on basis of data length and export data to MySQL.

- Do setup for Flink Streaming.

- Prepare Dashboard.html for pie chart visualization.

- Pushing code contribution on GitHub.

- Participating in project team meetings,