

Telecom Call Data Record Analysis

Mr. Ritik SR Jha, Ms. Sonam Ram

B.Tech Student (CSE), RITEE, Raipur, Chhattisgarh, India

Abstract-Client needs to develop a POC (Proof of Concept) for one of its esteemed clients for Unified Call Data Record Analysis. The challenge is that the CDR (Call Data Record) is in raw format and does not have column names also present and it has 435 columns in total. Contains date and time information in a regular pattern like- "2019 06 20 03 27 17.906". contains some unwanted data. Information spread across multiple columns. We need to visualize the Analytics for Call data, Device data and Service data.

CDR Analysis and Insights is a real time call data records analysis project, in which earlier there was the raw data that had 16k+ Rows and 435 Columns and according to the requirement of the client, The CDR data is represented using a UI which contains three tabs, Call Analytics Tool , Device Analytics Tool and Service Analytics Tool which will give information about the calls, devices and services in the form graph, cards , tables and pie chart respectively, hence making it easy to Visualize and Analyze Data so that some insights can be made over the data

Keywords: Data Visualization, Data Analysis, Call Data Record, WireFrame, Python.

1. INTRODUCTION

Telecom Call Data Record (CDR) analysis has gained significant attention in the field of telecommunications and data analytics due to its potential for extracting valuable insights and supporting various applications such as network optimization, fraud detection, and customer behavior analysis. In this section, we review the existing literature on CDR analysis, highlighting key contributions and approaches employed by previous researchers.

We need to visualize the Analytics for Call data, Device data and Service data.

1. We should be able to run the Python script from the command line and it should open a Web Page in the browser with a Descent and Usable Interface.
2. User should have an option on the Web Page to close the opened browser after his work is complete, it

should also close/kill the script running in the background.

3. UI should have a Call Data Analytics (Interactive Visualisation) to generate a Line Graph based on the filtration of Start date and End date, Group and based on Hourly, Weekly and Daily. User should be able to select any Start Date and End Date, Group and Type of Graph and based on that we need to show the graph. We need to show important information like total calls, Incoming and Outgoing, Missed Calls, Max Duration etc in the form of Cards. Also show the records in the form of tabular data.

4. UI should have a Device Data Analytics and Service Data Analytics (Interactive Visualisation) to generate a Pie Chart.

2. RELATED WORK

In this project, We focused on Data Analysis and Data Visualization Skills, which included working with raw, unfiltered datasets and forming graphs and pie charts. We also got to know about various analytical tools provided to us by Python, libraries such as: import webbrowser, import dash, import re and many more.

Network Traffic Analysis

proposed a method to analyze CDRs to detect network congestion and improve quality of service by dynamically rerouting traffic. employed machine learning techniques to identify abnormal traffic patterns and predict network failures.

Customer Segmentation and Behavior Analysis

CDR analysis has been widely utilized for customer segmentation and behavior analysis, enabling telecom operators to gain a deeper understanding of their customer base. Researchers have employed clustering algorithms and data mining techniques to identify distinct customer segments based on calling patterns, usage preferences, and geographical locations.

Fraud Detection

CDR analysis plays a crucial role in fraud detection and prevention in the telecom industry. proposed a fraud detection framework that utilizes CDRs and machine learning algorithms to identify abnormal call traffic associated with SIM card cloning.

Privacy-Preserving Techniques

presented a privacy-preserving scheme for location-based services using CDRs, enabling spatial analysis without compromising user privacy.

3. PROPOSED METHODOLOGY

We would be going to use Agile Methodology. Release Early and Release Often

This is the way I would have approached to solve a data science project: Create a Backlog—> To Do—> Doing—> Done.

Create A User Story As a User type I want a goal for some reason.

Step 1 (Analysis Phase)

First step you need to do is read the document carefully and download the dataset. It is very important for any data science project to have a proper understanding of the dataset. (Domain Expert for Telecom) These are the columns which are required to be read for the project, Column numbering starts from 0

1 serviceProvider
4 group Information
5 direction
9 startTime
13 endTime
14 Miss Call Information
31 Group ID
120 userId
180 twoStageDialingDigits
146 relatedCallId
147 relatedCallIdReason
267 vpDialingfacResult
312 locationType
345 userAgent

Deliverable of this phase would be a CSV file with reduced columns required to meet the client expectation. You should be able to read a single record and understand it.

First create a csv with enhanced columns of missing

data Then we need to split the CSV into three CSV for Call, Device and Service related data.

Step 2 (Design Phase)

Now after understanding the client's expectation, try to create a Wireframe or rough sketch of the UI. How will the client actually use your software? Will there be Drop down, text-boxes, buttons etc,

Also what actions would happen on the Web Page Which Graphs/Chart should be used? How many pages should I design? This step will better consolidate the client's understanding. Deliverable of this phase would be a Wireframe document.

Step 3 (Research and Learning Phase)

Now after understanding the client's expectation, create a list of how to. This list will be a technical How to do List. For Example, How to Run the Script of Python from Command Line or Terminal ? How to open the Browser from Python with a specified URL ? How to kill the process when someone clicks the CLOSE button on the browser ? How to fill the drop down with values from the Pandas in python How to show a Chart/Graph in a webpage using Python ? Dash Google all these questions and read articles so that you now know how to technically solve that part of the problem Deliverable of this phase would be a lot of understanding notes from reading multiple articles/blogs and watching videos.

Step 4 (Development Phase)

Segregate the list of clients expectation into separate parts

- a) User Experience of Running the Project
- b) UI Expectation for taking input from the user
- c) UI Expectation for the output

Step 5 (Backend Development)

Create a Python script with the UI element separate to solve the core expectation from the client for Call, Device and Service Analysis. Take inputs from the user and create a custom python function which filters the required data from the dataset.

Step 6 (Frontend Development)

Now Create the UI for the clients Expectation. Then link to the UI with the backend functions created. Both the Frontend and Backend Development can go side by side also.

Step 7 (Integrated Testing)

Manual Testing in Development Environment Manual Testing in Production Environment Here you will feel the need for having the Logging and Exception handling.

Step 8 (UAT Phase - User Acceptance Test)

In this phase the full project is given to the client and installed on the client's machine for verification of clients expectations. The client's raised points need to be fixed and resubmitted to the client again for the confirmation. This is an iterative process and might take 2-3 iterations depending on the number of points raised.

Step 9 (Packaging for Final Production Delivery)

In this phase the source code and other deliverables are submitted to the client A project Document needs to

be submitted User Manual Document needs to be submitted with Source Code Documentation.

4. RESULTS AND DISCUSSION

With the help of the cleared data, a web UI is created displaying 3 tabs: Call Analytics Tool, Device Analytics Tool, Service Analytics Tool. The former tab will give the information about the calls, also the information according to the user's selection of data from the four dropdowns in a graphical representation (weekly, day wise hourly), also gives the information in the form of cards (in coming, outgoing, missed calls) and data table. Device Analytics as well as Service Analytics gives the information regarding the devices and services respectively in the form of pie charts.

Main UI Reference

CDR Analysis with Insights

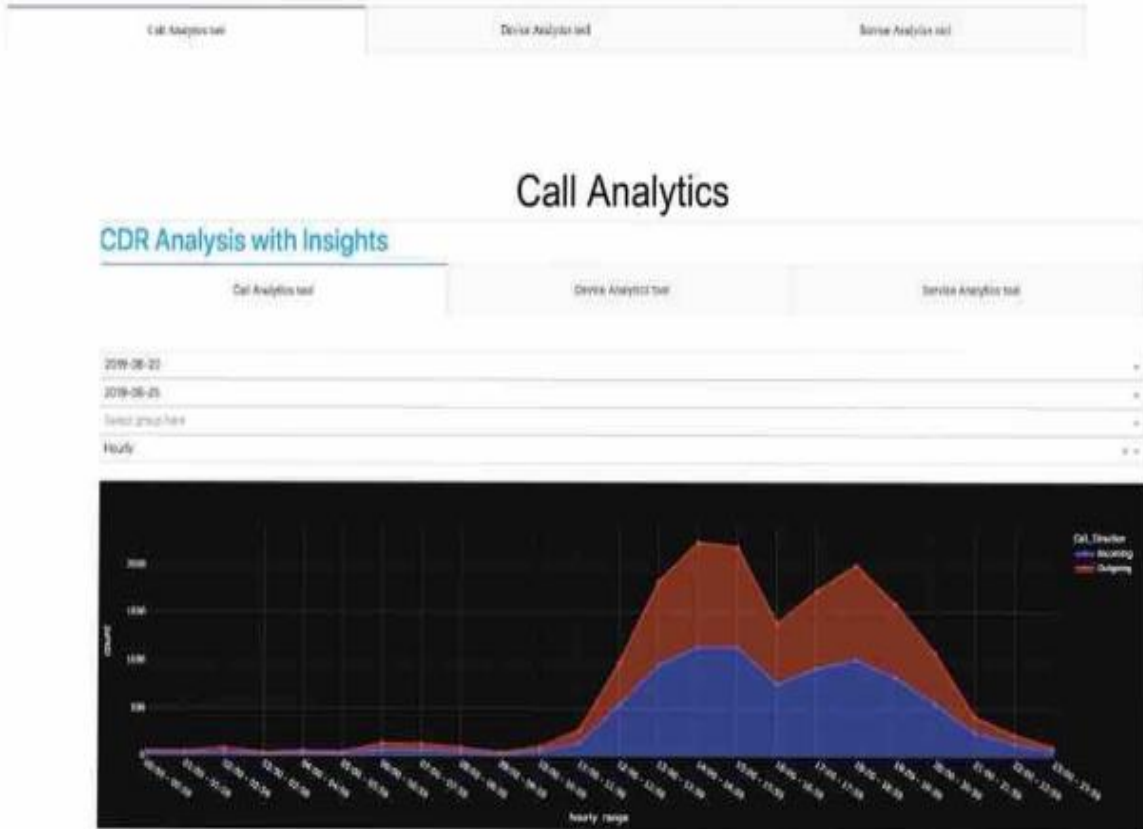
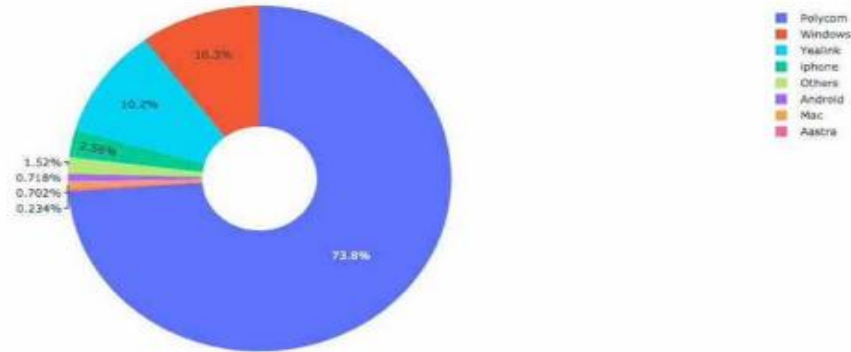


Figure 1 : CALL ANALYTICS



Group	DeviceID	DeviceType	Incoming	Outgoing	Missed Calls	Total_milli_secs
1219551297	4864927148@vulp.vo.llnwd.net	PolycornVXX-VXX-400-0A/S-7.1.2205_00442307044	0	0	0	0
1219551297	4864927148@vulp.vo.llnwd.net	PolycornVXX-VXX-400-0A/S-4.7.1694_00442300864	0	0	0	0
1219551297	4864927148@vulp.vo.llnwd.net	Polycorn@vulldPolat12-022P_316-0A/6.0-2.11332_00442300864	49	0	0	14
12302153919	22642855484@vulp.vo.llnwd.net	bc-0e-v-0C-0se (22-4.3 108 IPhone)	0	2	0	13
12302153919	22642855484@vulp.vo.llnwd.net	Polycorn@vulldPolat12-022P_316-0A/4.0-2.11332_00442300864	0	0	0	0

Figure 2 : CALL ANALYTICAL TOOL
Device Analytics



Service Analytics

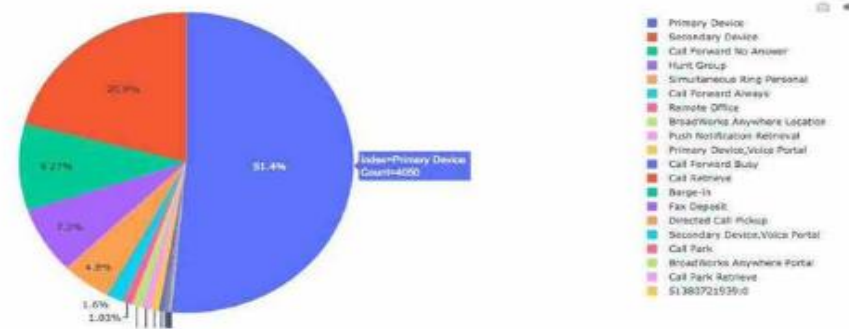


Figure 3: PIE CHART REPRESENTATION

5.CONCLUSION

This is a python based data analytics article and the UI was created using dash module and its framework. This Article has been a very interesting project that has made me learn many new concepts as well as boosted my confidence, skills and knowledge.

REFERENCE

- [1] <https://www.stackoverflow.com/>
- [2] <https://www.tutorialspoint.com/>
- [3] <https://www.google.com/>
- [4]Introduction to Plotly Dash Web Application Development Framework | Plotly Dash Tutorial Part - 1
- [5] Part 1. Layout | Dash for Python Documentation | Plotly
- [6]<https://www.datacamp.com/community/tutorials/learn-build-dash-python> (skip after Authentication Section)

7.BIOGRAPHY

RITIK SR JHA is pursuing Bachelor's of Technology in the Computer Science and Engineering Department at Raipur Institute of Technology, affiliated to Chhattisgarh Swami Vivekanand Technical University, Bhilai, Chhattisgarh, India. I am interested in Machine Learning, Big Data and Backend Development.