To Analyse and Optimise the Issues in Call Drops and Enhance the Real Time (CDR) Processing

G. V. Ashok and P. Vasanthi Kumari

March 5, 2020

To Analyse and optimise the issues in call drops and enhance the real time (CDR) processing

Ashok GV
Dept. of Computer Science Engineering
Dayananda Sagar University,
Bangalore, Karnataka
ashokgvm@gmail.com

Dr. Vasanthi Kumari P
Associate Professor
Dept. of Computer Science Engineering
Dayananda Sagar University,
Bangalore, Karnataka
vasanthi-bca@dsu.edu.in

Abstract

Telecom dropped-call rate is the fraction of the calls which gets terminated before any of the consumers finishes the conversation and disconnects frequently. Call drop is a huge problem, despite warnings from the telecom regulator, the problem persists even after years. Causes of call drops happens with different factors such as Demand of wireless cellular connectivity, Signal strength, Network Configuration, Transmission Problems, Trans-receiver and receiver problem, Propagation factors, Irregular user behaviour, radio propagation for wireless communications, non-availability of sites due to acquisition problems, sealing of sites by local authorities due to fear of electro-magnetic fields from mobile radiations etc. Even inadequate infrastructure and overloaded networks aggravated the call drop issues whereas, setting up of more towers for better coverage would also directly affect the health of habitants around. In this paper, the overall condition of call drops has been reviewed and possible ways to minimize the spectacles of call drops using Linear Discriminant Analysis (LDA) is a dimensionality reduction technique. The goal of LDA is to project the features in higher dimensional space onto a lower-dimensional space in order to avoid the curse of dimensionality and also reduce resources and dimensional costs This paper gives to telecom service providers to improve their infrastructure to minimize the effect of call drop and provide quality services with the advanced technologies with precise algorithms.

Keywords: Call drops, Source, Destination, Event type, CDR, Liner Discriminant Analysis (LDA)

I INTRODUCTION

"Quality of Service" (QoS) is a big concern, Concerned over the worsening call drop situation in the world or country. The issue of call drops in mobile networks has now become a new burning topic in public discourse. Whether you have called another user or someone else has called you on your mobile phone, if the call is interrupted before the users decide to terminate the call. A dropped call directly affects the quality of service (QoS) that is expected to be maintained in a mobile wireless network. All big telecom networks have rampant call drops causing trouble for customers. A survey has shown that 56% of telco users face severe call connect and call drop problems.

II LITERATURE REVIEW

1. Mr. Nilesh R. Gode [1] Ms Jyoti Mali[2], Atharva College of Engineering, Mumbai University, published on International Journal of Scientific & Engineering Research, Volume 7, Issue 2, February-2016 ISSN 2229-5518.
   Paper published on "CALL DROP ISSUE'S OF MOBILE PHONES IN INDIA" and given details on details idea about how the "Call Drop

Issue" occur, what are the different factor responsible for causing this, how can we calculate the call Drop for 2G and 3G that data is completely calculated by TRAI with different service provider. The possible solution for consumer point of view and finally try to give a possible solution for reducing the Call Drop Issue.

2. José Antonio Iglesias, Agapito Ledezma, Araceli Sanchis and Plamen Angelov
Computer Science Department, Carlos III University of Madrid, Computing and Communications Department, Lancaster University, Lancaster LA14WA, UK, Published: 5 August 2017.

Paper published on "Real-Time Recognition of Calling Pattern and Behaviour of Mobile Phone Users through Anomaly Detection and Dynamically-Evolving Clustering".

In this paper explains competitive telecommunications market, the information that the mobile telecom operators can obtain by regularly analysing their massive stored call logs, is of great interest. Although the data that can be extracted nowadays from mobile phones have been enriched with much information, the data solely from the call logs can give us vital information about the customers. This information is usually related with the calling behaviour of their customers and it can be used to manage them. However, the analysis of these data is normally very complex because of the vast data stream to analyse. Thus, efficient data mining techniques need to be used for this purpose. In this paper, a novel approach to analyse call detail records (CDR) is proposed, with the main goal to extract and cluster different calling patterns or behaviours, and to detect outliers. The main novelty of this approach is that it works in real-time using an evolving and recursive framework.

III PROBLEM STATEMENT

1. Network Call Drops:

- Analysis of the call drops and report the analytics to recognize issue either with service provider's networks source or destination in each region and retain/prevent the existing subscribers.
- Which region in my network at the most dropped calls in the past hours, day, week and which of my customers were most affected? Are these customers profitable? Are they likely to churn?
- Which of the outages were due to handset problem, wireless coverage problems or switch problems ?

IV PROPOSED METHODOLOGY

Over come up with above problem statement, Analyse call drops and report the analytics to recognize issue either with service provider's networks source or destination in each region and retain/prevent the existing subscribers and optimise the issues in call drops and enhance the real time (CDR) processing with security.

1. Process:

Have taken below two of business scenarios/use cases for the present research and able analyse results and taken for future implementation.

- Analyse call drop in each region
- Call duration in every hour

2. Created framework to generate CDR's

Every voice call and IP service in a telecommunications network generates usage records. These service usage records, Call Details Records (CDR) for a voice network and IP Data Records (IPDR) for IP networks, contain information about the call or session that is used for applications such as billing, service quality monitoring and fraud detection. Data record formats are controlled by standards, but unfortunately standards vary by industry.

Most service providers analyse CDRs through batch processing, usually once a day, but sometimes as infrequently as weekly or even monthly. Even where certain types of CDRs are processed more frequently, this is at best hourly.

It's difficult to get the real time data from any of the organization or service providers hence created the framework to generate the Real time CDR's by using IntelliJ IDE with Java programming language and feature files are written using cucumber. This framework can utilise to generate the lacs or crores of real time CDR data with in fraction of seconds.

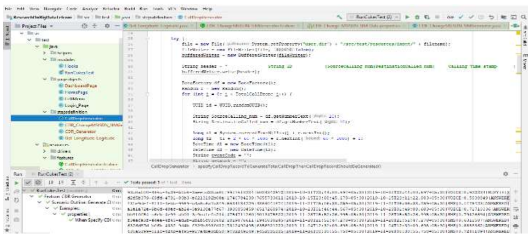

Fig 1: Framework to generate real time CDR data

Have taken examples of real time CDR's, which used to identify the network failures based on source, destination, Event codes and call results.

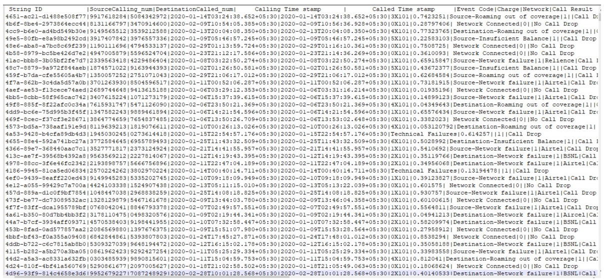

Fig 2: Example of CDR data

Get Latitude and Longitude based on the Location cities/places of call made which will helps to determine the location. Latitude and longitude used to identify caller's geographical area which is used to calculate the network failure area.

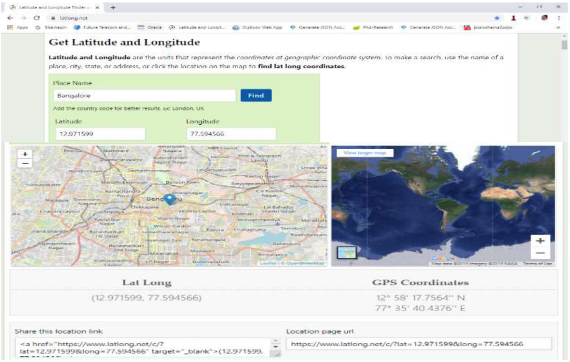

Fig 3: Get Location based on Latitude and longitude.

3. System Design flow for call drops

Subscriber data has been stored in Database with multiple tables and columns. Required data has been extracted to analysis from different column based table. These call drops are identified based on the event type and call results. There are few scenarios are identified call drops based on their source and destination network failures.

Below diagram represents the scenarios of call drop happens due to Source (S), Destination (D) network. Identified the success or call drops based on Event code.

Source (S)	Destination (D)	Event Code
S (No Issue)	D (No Issue)	0x101
S (Network Failure)	D (No Issue)	1x101
S (No Issue)	D (Network Failure)	2x101
S (Roaming – Out of Coverage)	D (No Issue)	3x101
S (No Issue)	D (Roaming – Out of Coverage)	4x101
S (Call Drop – Insufficient Balance)	D (No Issue)	5x101
S (No Issue)	D (Roaming – Insufficient Balance)	6x101

Fig 4: Identify call drop based on Event code

Below diagram depicts call flows on event code based on success or call drops

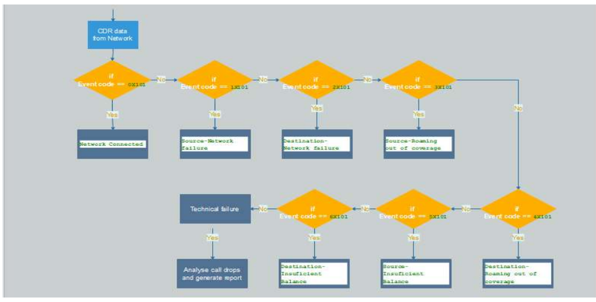

Fig 5: Flowchart diagram for call flows on event code based on success or call drops

4. Design and Experiment

The functional flow of real time application is as given below

- The data generated in Network Elements from service provider end is stored in Network File System (NFS)
- The NFS has set of files in text format.
- It is transferred to Local File System (LFS)
- Process multiple files using Java program with Apache Storm
- Load the data into HDFS
- Place analytics over Apache Spark and HDFS to run business use cases.

Fig 6: File Process System

Linear Discriminant Analysis (LDA)

Applied Linear Discriminant Analysis (LDA) algorithm to identify the call drop based on the regions and have got accurate results to identify the network failures.

Linear Discriminant Analysis (LDA) is a dimensionality reduction technique. It is used as a pre-processing step in Machine Learning and applications of pattern classification. The goal of LDA is to project the features in higher dimensional space onto a lower-dimensional space in order to avoid the curse of dimensionality and also reduce resources and dimensional costs.

LDA is a supervised classification technique that is considered a part of crafting competitive machine learning models. This category of dimensionality reduction is used in areas like image recognition and predictive analysis in marketing.

Below examples depicts how Liner Discriminant Analysis (LDA) used identify the call drops due to network failures on both source and destination. We have taken two set of data(x1, x2) which are successfully connected from the Source & Destination(x1) and call drops(x2) happens on both regions.

Class	X1 (Source)	X2 (Destination)	Status on Call drop
1	29.48	66.27	No Call drop
1	25.53	77.85	No Call drop
1	35.67	56.51	No Call drop
1	31.57	54.67	No Call drop
2	25.82	44.58	Call drop
2	21.56	62.22	Call drop
2	32.73	35.21	Call drop

Fig 7: Classification of data

We plot the features, we can see the data in linearly separable. We can draw a line to separate the two groups. The problem is to find the line and to rotate the feature in such a way to maximize the distance between the groups and to minimize distance with in group.

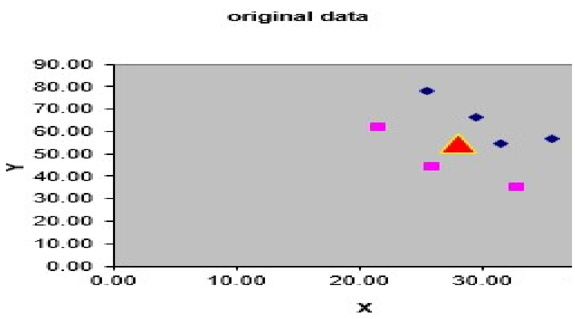

Fig 8: Original data

X = Features (or independent variables) of all data. Each row (dependent by K) represents one of the object, each column stands for one feature(Source and Destination).

Y= Group of the obect (or independent variables) of all data. Each row represents one of the object and it has only one column.

In our example

$$x = \begin{bmatrix} 29.48 & 66.27 \\ 25.53 & 77.85 \\ 35.67 & 56.51 \\ 31.57 & 54,67 \\ 25.82 & 44.58 \\ 21.56 & 62.22 \\ 32.73 & 35.21 \end{bmatrix} \quad y = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{bmatrix}$$

X_k = data of row ^{k}.

For example x_3= $[35.67 \quad 56.51]$ x_7=$[32.73 \quad 35.21]$

^{g} number of groups in y. In our example $^g = 2$

X_i = feature data for group ^{i}. Each row represents one object, each column stands for one feature.

We separate x into several groups based on the number of category in ^{y}

$$x1 = \begin{bmatrix} 29.48 & 66.27 \\ 25.53 & 77.85 \\ 35.67 & 56.51 \\ 31.57 & 54,67 \end{bmatrix} \qquad x2 = \begin{bmatrix} 25.82 & 44.58 \\ 21.56 & 62.22 \\ 32.73 & 35.21 \end{bmatrix}$$

μ_i = mean of feature in group I , which is avarage of x_i

$$\mu 1 = [30.56 \quad 63.83] \qquad \mu 2 = [26.71 \quad 47.34]$$

μ = global mean vector, that is mean of the whole data set.

In this example $\mu = [28.909 \quad 56.758]$

X_i^0 = mean corrected data, that is features data for group ^{i}, x_i , minus the global mean vector μ

$$x1 = \begin{bmatrix} 0.569 & 9.510 \\ -3.379 & 21.092 \\ 6.761 & -0.248 \\ 2.661 & -2.088 \end{bmatrix} \quad x2 = \begin{bmatrix} -3.086 & -12.181 \\ -7.351 & 5.465 \\ 3.825 & -21.552 \end{bmatrix}$$

$$\mathbf{c}_i = \frac{\left(\mathbf{x}_i^o\right)^T \mathbf{x}_i^o}{n_i}$$

= covariance matrix of groups ^{i}

$$c1 = \begin{bmatrix} 16.132 & -18.271 \\ -18.271 & 134.936 \end{bmatrix} \quad c2 = \begin{bmatrix} 26.063 & -28.341 \\ -28.341 & 214.234 \end{bmatrix}$$

$$C(r,s) = \frac{1}{n} \sum_{i=1}^{g} n_i \cdot c_i(r,s)$$

= pooled within group covariance matrix. It is calculated for each entry (^{r,s}) in the matrix.

In our example

$\frac{4}{7}*16.132+\frac{3}{7}*26.063 = 20.388,$

$\frac{4}{7}*-18.271+\frac{3}{7}*-28.341 = -22.587$

$\frac{4}{7}*-18.271+\frac{3}{7}*-28.341 =-22.587,$

$\frac{4}{7}*134.936+\frac{3}{7}*214.234 = 168.921$

$$c = \begin{bmatrix} 20.388 & -22.587 \\ -22.587 & 168.921 \end{bmatrix}$$

The inverse of the pooled covariance matrix is

$$C^{-1} = \begin{bmatrix} 20.388 & -22.587 \\ -22.587 & 168.921 \end{bmatrix}$$

P = prior probability vector (each row represent prior probability of group ^{i}) that is

$$p_i = \frac{n_i}{N}$$

$$p = \begin{bmatrix} 0.571 \\ 0.429 \end{bmatrix} = \begin{bmatrix} \frac{4}{7} \\ \frac{3}{7} \end{bmatrix}$$

Discriminant function

$$f_i = \mu_i C^{-1} x_k^T - \tfrac{1}{2} \mu_i C^{-1} \mu_i^T + \ln(p_i)$$

We should assign object ^{k} to group ^{i} that has maximum f_i

Below diagram depicts classification of success and failures of call drops based on network region

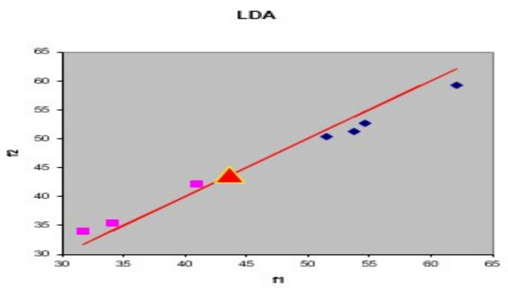

Fig 9: Result of classification for success and failures of call drops

LDA is considered to be a very simple and effective method, especially for classification techniques. Since it is simple and well understood, so it has a lot of extensions and variations:

- **Quadratic Discriminant Analysis (QDA)** – When there are multiple input variables, each of the class uses its own estimate of variance and covariance.
- **Flexible Discriminant Analysis (FDA)** – This technique is performed when a non-linear combination of inputs is used as splines.
- **Regularized Discriminant Analysis (RDA)** – It moderates the influence of various variables in LDA by regularizing the estimate of the covariance.

V RESULTS AND DISCUSSIONS

Have taken below two of business scenarios/use cases to present results and display the analytics on dashboard.

Use Cases

The use cases generated by business users are given below

- Analyze call drop in each region
- Call duration in every hour

Below graph represents the call connected and call drops happens every 5 seconds between the service providers.

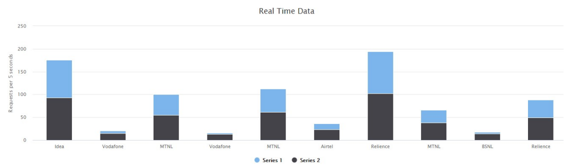

Fig 10: Data visualisation on call made each 5 seconds

Result of our computation are given below on LDA

Training Data, D		
class	X1 (Source)	X2 (Destination)
1	29.48	66.27
1	25.53	77.85
1	35.67	56.51
1	31.57	54.67
2	25.82	44.58
2	21.56	62.22
2	32.73	35.21
prediction	28.10	54.57
Average	28.909	56.758
Std dev	4.515	12.997

Mean Corrected Data		Discriminant function		Results
X1o	X2o	f1	f2	Classification
0.569	9.510	54.722	52.597	1
-3.379	21.092	53.697	51.279	1
6.761	-0.248	62.035	59.158	1
2.661	-2.088	51.557	50.376	1
-3.086	-12.181	31.770	34.051	2
-7.351	5.465	34.147	35.370	2
3.825	-21.552	40.965	42.186	2
-0.809	-2.191	43.676	43.721	2

Fig 11: Consolidated computation data on LDA

The discriminant function is our classification rules to assign the object into separate group. If we input the call drops that have source 28.10 and destination and 54.57, reveal that it does have more call drop in that region.

Transforming all data into discriminant function (f_1,f_2) we can draw the training data and the prediction data into new coordinate. The discriminant line is all data discriminant function X= {f_1,f_2} and Y = {f_1,f_2}.

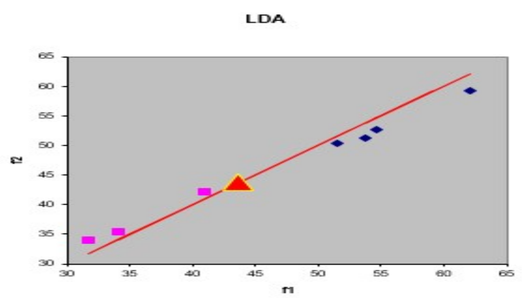

Fig 12: Result of classification for success and failures of call drops

VI CONCLUSION AND FUTURE WORK

World's fastest growing economy is plagued by poor infrastructure and overloaded networks to an extent that many callers are cut off even before they can finish a sentence. Issues such as call drops, network outage and connectivity gaps etc are being faced by the users' due to the absence of comprehensive mobile tower. The dropped-call rate is one of the key performance indicators (KPI) used by the network operators to assess the performance of their networks. It is assumed to have direct influence on the customer satisfaction with the service provided by the network and its operator. In mobile networks this is achieved by improving radio coverage, expanding the capacity of the network and optimising the performance of its elements, all of which may require considerable effort and significant investments on the part of the network operator. Call drops has been reviewed and possible ways to minimize the spectacles of call drops using Linear Discriminant Analysis (LDA) which is dimensionality reduction technique. As discussed in the results have shown the accuracy on results based on the LDA algorithms. This paper provides the reasons of call drop and basis for resolving the same. It provides a summarized view of the problem of call dropping, how it occurs and what are the methods which can be opt to improve the telecom infrastructure and minimize the call drop. Even we will have to consider the naturally not impact on the animals or birds if increases the more network towers. Still some challenges with no mitigation strategies. Our future research will concentrate on developing a more complete understanding of challenges associated with liner or non-liner algorithms using machine learning.

REFERENCES

[1] http://ieeexplore.ieee.org/

[2]. https://en.wikipedia.org/wiki/Big_data

[3] http://bigdata-madesimple.com/11-interesting-big-data-case-studies-in-telecom/

[4] http://www.economist.com

[5] Big Data Essentials Kindle Edition by Anil Maheshwari (June 2016)

[6] Handoff and Drop Call Probability: Case Study of Nigeria's Global System for Mobile Communication Sector, Rex NdubuisiAli, Scholars Journal of Engineering and Technology, 2015.

[7] Handoff Schemes and its Performance Analysis of Priority within a particular channel in wireless systems, Chandan Singh Ujarari, Arun Kumar,

International Journal for Research in Applied Sciences & Engineering Technology, May 2015

[8] Decreasing Call Blocking Rate by Using Optimization Technique Vinay Prakash Sriwastava Jalneesh Singh, Vinay Kumar Verma.

[9] The Computing Big Data Review 2015 summarizes the results of a comprehensive research programmer undertaken by Computing during the first quarter of 2015.

[10] http://www.informationweek.com/big-data/big-data-analytics/big-data-predictions-for-2016/d/d-id/1323671

[11] https://www.splunk.com/en_us/solutions/solution-areas/big-data.html