REAL-TIME FRAUD DETECTION IN TELECOMMUNICATION NETWORK USING CALL PATTERN ANALYSIS

Kehelwala Gamaralalage Dasun Chamara Kehelwala

(148223L)

Degree of Master of Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

December 2017

REAL-TIME FRAUD DETECTION IN TELECOMMUNICATION NETWORK USING CALL PATTERN ANALYSIS

Kehelwala Gamaralalage Dasun Chamara Kehelwala

(148223L)

Dissertation submitted in partial fulfillment of the requirements for the degree
Master of Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

December 2017

DECLARATION

Candidate:

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my report, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

……………….…………… ……………………….

K.G.D.C. Kehelwala Date

Supervisor:

The above candidate has carried out research for the Masters Dissertation under my supervision.

……………….…………… ..……………………

Dr. H.M.N. Dilum Bandara Date

Abstract

Telecommunication service providers are losing considerable percentage of their annual revenue due to fraudulent activities. Such activities also deteriorate customer experience. Therefore, real-time detection of such fraudulent activities is required to minimize the revenue loss and to preserve customer experience. Illegal termination of International calls (aka. SIMbox fraud) and extreme usage scenarios related to International revenue share fraud are two major fraudulent activities which make highest impact. While such activities can be detected by identifying behavioral and calling patterns of subscribers, they need to be detected in real time so that subscriber connections linked with an ongoing fraud activity can be terminated to minimize the impact of threat or revenue loss. Call Detail Records (CDRs) produced by telecommunication equipment contains attributes that are specific to a phone call or other communication transactions handled by the device could be used to detect behavioral and calling patterns of subscribers. However, traditional CDR analysis techniques do not facilitate time-sensitive monitoring and analytical requirements. Therefore, we propose a Complex Event Processing (CEP) based solution for the real-time identification of fraudulent and extreme usage subscriber patterns. We identified a rich set of features and set of call patterns, and then combined batch analytics with real-time analytics to increase the detection accuracy. We demonstrated the utility of the proposed solution using a real dataset from a service provider. The proposed solution achieved an accuracy of 99.9% with average latency of 16 call attempts per detection at input event rate of 230 events per second with modest hardware.

Keywords: Complex Event Processing, Data analytics, Call Detail Records, call patterns

ACKNOWLADGEMENT

My sincere gratitude goes to my family members for the continuous support and motivation given to make this thesis a success. I also express my heartfelt appreciation to Dr. Dilum Bandara, my supervisor, for the supervision, advice and valuable feedback given throughout to make this research a success. I also thank to Mr. Ruchira Yasaratne, Mr. Sampath Ilesinghe and Mr. Pradeep De Almeida of the Dialog Axiata PLC, for providing approvals to proceed this project by keeping trust on me. Last but not least I also thank my friends who supported me in this whole effort.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
ASCII	American Standard Code for Information Interchange
BAM	Business Activity Monitor
BI	Business Intelligence
BSC	Base Station Controller
CDR	Call Detail Record
CEP	Complex Event Processor
CLI	Calling Line Identification
CTR	Click-through Rate
CUP	Current User Profile
DAHP	Database-Active Human-Passive
DAS	Data Analytics Server
DBMS	Database Management System
DDOS	Distributed Denial of Service
DOS	Denial of Service
DSMS	Data Stream Management System
EDGE	Enhanced Data rates for GSM Evolution
FDT	Fraud Detection Tool
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GT	Global Title
HADP	Human-Active Database-Passive
HSPA	High Speed Packet Access
HTTP	Hypertext Transfer Protocol
IDD	International Direct Dialing
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
ISC	International Switching Center
ISDN	Integrated Services Digital Network
ISUP	ISDN User Part

LTE	Long Term Evolution
LKR	Sri Lankan Rupee
MCC	Mobile Country Code
MLP	Multi-Layer Perception
MNC	Mobile Network Code
MO	Mobile Originated
MSC	Mobile Switching Center
MSISDN	Mobile Station - ISDN
MT	Mobile Terminated
NFA	Non-Deterministic Finite Automata
NN	Neural Networks
OCS	Online Charging Node
PABX	Private Automatic Branch Exchange
QoS	Quality of Service
RFID	Radio Frequency Identification
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol
SMS	Short Message Service
SOM	Self-Organizing Map
SVM	Support Vector Machine
TDM	Time Division Multiplexing
TMSC	Tandem Mobile Switching Center
UPH	User Profile History
UTMS	Universal Mobile Telecommunications System
VLR	Visitor Location Register
VoIP	Voice over Internet Protocol