Anomalous Behavior Detection in Mobile Network

Mon Mon Ko, Mie Mie Su Thwin
University of Computer Studies, Mandalay, Myanmar Computer Emergency Response Team Ministry of Science and Technology, Yangon ,Myanmar
nannlaypyaenu@gmail.com, drmiemiesuthwin@mmcert.org.mm

Abstract

New security threats emerge against mobile devices as the devices' computing power and storage capabilities evolve. Preventive mechanisms like authentication, encryption alone are not sufficient to provide adequate security for a system. There is a definite need for Anomaly detection systems that will improve security on the mobile phone. In this work, we propose User Group Partition Algorithm and Behavior Pattern Matching Algorithm to extract anomalous calls from mobile call detail records effectively.

1. Introduction

Mobile devices have evolved and experienced a great success over the last few years. Such devices are capable of performing sophisticated tasks and communicate through various wireless interfaces[1]. However, along with their popularity, mobile devices face an everyday growing number of security threats. This is despite the variety of peripheral protection mechanisms proposed in the literature in recent years. Without doubt, authentication and access control methods can be used in many cases, but alone, they are not sufficient to offer integral protection against intrusions[2]. Overall, with the increasing risk of mobile malware, the theft or loss of mobile devices and the physical vulnerability, i.e. rewiring a circuit on the chip or using probing pins to monitor data flows to retrieve private keys or find flaws in the hardware components, designing a highly secure mobile device is still a very challenging task[3].

While more than four billion people enjoy their mobile devices using 2G/3G mobile networks, Kaspersky Lab has very recently identified 39 new mobile malware families (SMS trojans, iPhone malware, Android spyware) with 143 modifications[4]. According to a ScanSafe report malware volumes grew 300% in 2008, and it is noted that most of the legitimate web pages crawling on the Internet are not trustworthy or infected by different kinds of viruses[5]. Moreover, according to the UK Home Office, 69% of robberies include a mobile device. As a result, a need for more intelligent and sophisticated security controls such as Intrusion Detection Systems (IDSs) for mobile devices is necessary[6]. In general, there are two basic approaches in IDS to detect an intrusion: a) misuse based (also called signature-based), and b) anomaly based (also called behavior-based). Although misuse based IDS can immediately be employed to monitor the mobile environment, only an anomaly-based IDS is able to detect new, unforeseen vulnerabilities and variants of known attacks[7]. Anomaly-based intrusion detection profiles normal behavior and attempts to identify patterns of user activities that deviate from a predefined or dynamically updated profile[8,9]. Whilst much research has been devoted to IDS, in the context of anomaly detection, the exploration of what is defined as "normal" has been limited and several important problems remain unsolved[10,11].

In this paper we concentrate on anomaly-based IDS for mobile network. We use a data set generated from a database of real world mobile communication network information. The database provides the following information for each transaction (use of a service by a customer): the initiation time, the duration (in minutes), and the type of the service. From the database, we generated a data set with 13,280 examples.

1.1. Motivation

The motivation behind this approach is to seek proper as opposed to normal behavior. It also to overcome the drawbacks of existing approaches such as specification increasing, high false alarm rate, and to provide the solve options for users.

1.2. Purposes of the System

The purposes of this System are:

- To study anomaly detection in mobile network elaborately.
- To propose an effective detection system using GP algorithm and BPM algorithm and summarizing mobile user behavior.
- To eliminate the drawbacks of previous works in anomaly detection in mobile network.

- To develop the applicable anomaly detection system of mobile network effectively.
- To help for getting high accuracy in detection of anomalies in call detail records.
- To highlight the importance of anomaly detection process for mobile telecommunication network.
- To observe the usefulness of mobile user behavior patterns.
- To pinpoint the impacts of CDR features for mobile user behavior monitoring.

2. Literature Review

This chapter presents some knowledge and overview of approaches in the literature concerned with anomalous behavior detection in mobile phone system.

The work in [12] proposed a prototype of a tool, based ona supervised Artificial Neural Network (ANN), to detectanomalous behaviour on mobile communications, suchas service fraud and Subscriber Identity Module (SIM)card cloning. The authors, based on their prototype, reportaccuracy of a 92.50% detection of fraudulent users and a92.5% correct classification of legitimate users. The workin [13] proposed the Bayes Decision Rule (BDR) towardsthe generation of mobility user profiles within the GlobalSystem for Mobile Communications (GSM) network. Byutilising their method the authors managed to achieve aTPR of 83.50%. One problem with this approach is theprivacy of the end-user's usage log files, which are exposedto the telecom carriers in order to detect mistrusted users,as explained in [14].

Hollm´en [15] has proposed fraud detection techniquesin mobile networks by means of user profiling andclassification. Specifically, the author used ANN andprobabilistic models to detect anomalous usage andachieved a TPR of 69%. However, the presented methodfor fraud detection is based on an available large databasewith billions of records. As a result, this method can beseen only as a specific user profiling problem in frauddetection. The authors in [16] used ANN to form shortand long-term statistical behavior profiles for GSM andUniversal Mobile Telecommunication Systems (UMTS)networks. They define two time spans over the call datarecords, i.e. a shorter sequence or Current BehaviorProfile (CBP) and a longer one or Behavior ProfileHistory (BPH). They also used the maximal entropyprinciple to create statistical profiles and Hellinger distanceto calculate the distance between CBP and BPH. If thisdistance is greater than some pre-determined threshold, analarm is raised.The authors in [14] discussed how ANN and othertools can be applied against frauds in first generation (1G)mobile networks. They also presented an on-line securitysystem for fraud detection of mobile phone operationsusing the RBF model. They have pointed out that it isvery hard to build a system capable of identifying anypossible fraud; however they managed a TPR of 97.50%.

Also, the authors in [17] proposed an on-line anomalydetection algorithm, based on Markov Model, where thekey distinguishing characteristic is the use of sequencesof network cell IDs traversed by a user. With this IDSthey attempted to address the problem of SIM cloningand MAC-address spoofing. Through their experimentalprocedure a TPR of 87.50% has been attained. The work in[3] proposed a mobility-based anomaly detection schemeto detect cloning attacks and cell phone losses. The authorsemployed several methods, such as high order Markovtechniques, the exponentially Weighted Moving AverageModel (WMAM) and the Shannon's entropy in orderto explore normal usage profile. The highest TPR theyachieved was 89%.

Recently, in [18] the authors presented a testedfor experimenting with anomaly detection algorithmsand demonstrated its properties using two unsupervisedanomaly detection methods, i.e. Self-Organizing Map(SOM) and clustering. They conclude that both methodsare suitable for network monitoring.

The work in [19] presented a behavioral detection framework for malwaretargeting mobile devices. Particularly, the framework generates a malicious behavior signature database by extracting the key behavior signatures from the mobilemalware. By using this scheme the authors tried to applya method called Temporal Logic of Causal Knowledge(TLCK) in order to address the challenge of behavioraldetection. This is managed by providing a compact "spatial-temporal" representation of program behavior.

To identify malicious behavior they used Support VectorMachine (SVM) classification to train a classifier fromboth normal and malicious data. Their evaluation on bothsimulated and real-world malware samples indicates that behavioral detection is able to identify current mobileviruses and worms with more than 96% accuracy. The authors in [20] proposed VirusMeter, a malware detectionsystem and cross-evaluated Linear Regression, ANN andDecision Trees, for their ability to detect anomalous behaviors on mobile devices. By monitoring powerconsumption on a mobile device and using ANN theyachieved TPR of 98.60%. However, VirusMeter detectioncan be affected because the precision of battery

powerindicators may vary significantly between different mobileOperating Systems (OS).

3.Research Methodology

The objective of this Research is to develop an anomalous behavior detection system for mobile network, should be able to effectively detect the anomalous call and messages in mobile network communication. In order to accomplish this objective, a study of existing approaches to anomalous behavior detection in mobile network is conducted. The problem of anomalous behavior detection in mobile network has been investigated by researchers and many kinds of anomalous behavior detection methods have been proposed. In general, there are two Approaches which can be observed.

The first Approach is Anomaly-based Detection. Most of the Intrusion Detection Methods use the Anomaly-based intrusion detection. Actually Anomaly detection relies on models of the intended behavior of users and applications and interprets deviations from this 'normal' behavior as evidence of malicious activity. The second Approach, Specification-based intrusion detection defines the precise expected behavior of the system for specific events like messaging and calling.

Some researcher proposed specification based intrusion prevention approach to detect unauthorized access to sensitive services, such as SMS, audio, and video services. However in such approach, the number of specifications increase if the number of attacks to be prevented increases.

Some researcher proposed Knowledge-based intrusion detection schemes apply the knowledge they have accumulated about specific attacks and system vulnerabilities. Using this knowledge database, any action that is not explicitly recognized as an attack is considered acceptable. Otherwise, an alarm is triggered by the system. There are many different characteristics of intrusion systems available in the marketplace. Expert systems are based upon knowledge based intrusion detection techniques. Each attack is identified by a set of rules. Rule-based languages are used for modeling the knowledge that experts have accumulated about attacks/frauds. Information regarding some intruders has also been added to these systems. A major drawback of knowledge-based intrusion systems is the difficulty in gathering the information on known attacks (which should be updated regularly), and developing a comprehensive set of rules that can be used to identify intrusive behavior.

Actually anomalous behavior detection for mobile phone network is needed to reduce the drawbacks of previous works in anomaly detection in mobile phone network. And a relatively lightweight and fast method is also needed to do the grouping and matching mobile phone call patterns. So in this System, the mobile phone call detail records are used as inputs to effectively detect the anomalous mobile phone call behavior.

3.1.Data Collection

In order to develop models of normal and abnormal behavior and to be able to assess the diagnostic accuracy of the models, call data exhibiting both kinds of behavior is needed. Gathering normal call data is relatively easy as this mode dominates the population, but collecting abnormal call data is more problematic. Abnormal call data is relatively rare and the data collection involving human labor is expensive. In addition, the processing and storing of data is subject to restrictions due to legislation on privacy of data.

We use a data set generated from a database of real world mobile communication network information.The database provides the following information for each transaction (use of a serviceby a customer): the initiation time, the duration (in minutes), and the type of the service. Fromthe database, we generated a data set with 13,280 examples.

3.2. Data Structure

According to our study Telephone call, SMS are focused on.

Every record of the Telephone call data file is composedof the following collected features. First, the featureNumber refers to the telephone number of the calleror the callee.The Timestampfeature represents thedate and time a telephone call took place. Next,the Transaction feature indicates the direction of a call,that is incoming or outgoing. The Duration featurerepresents the duration of a call in seconds.Eachrecord of the SMS data file in turn is composedof the following features. The Number feature refers to themobile number the particular message has sent or received.The Timestampfeature represents a date and is referred to the dateand time an SMS has been sent or received. Next, the Transaction feature indicates the direction of an SMS (incoming oroutgoing).

3.3. Proposed Approach

We have two key ideas for our proposed approach: a call detail record user assessed and how the anomalous call record has been detected. Based on these ideas, call detail record; there are three main steps for detecting the anomalous call record behavior. They are 1) grouping the normal mobile phone users according to their mobile phone usage, 2) matching the abnormal mobile phone user with grouped mobile user, 3) determining the new mobile phone transaction is normal or not.

We partition the mobile phone usage patterns based on the following facts:
1) Counts and Duration of International Call,
2) Counts and Duration of International Incoming Call,
3) Counts and Duration of International Outgoing Call,
4) Counts and Duration of Local Call,
5) Counts and Duration of Local Incoming Call,
6) Counts and Duration of Local Outgoing Call.

First of all, we proposed the cleaning algorithm to eliminate the negligible and missing transaction data for our Anomalous Behavior Detection in Mobile Network (ABDMN) system.

Algorithm CleanCDR
CleanCDR (CDR dataset)
for each CDR in dataset
begin
if the call duration <= 0
then delete the call detail record
end if
Update CDR dataset
end
end for

Later, we apply the user group partitioning algorithm to group the mobile phone users with similar mobile phone usage.
Definitions of Terms in the Proposed Algorithm
CDR = Call Detail Record
GP = Group Partitioning
MaxDur = Maximum Duration
MinDur = Minimum Duration
GCount = Group Count

Algorithm : Group Partitioning (GP)
Input: MaxDur, MinDur, GCount, CDR
Output: Group Data
Begin
Start with MaxDur, MinDur, GCount, CDR
For i= 1 to GCount
Dist = MaxDur /GCount
Group (i)Min = MinDur
Group(i)Max = Dist * i
MinDur = Group(i)Min
End for
For I = 1 to GCount
For each record in CDR
If during of record is between Group(i)
Insert into Group(i)
Endif
End for
Endfor
end

Then, we apply the behavior pattern matching algorithm to detect the anomalous mobile phone calls.
Definitions of Terms in the Proposed Algorithm
BPM = Behavior Pattern Matching (BPM)
GP = Group Partitioning
seq = Sequence
m = Basic Vector
smty = Similarity

Algorithm Behavior Pattern Matching (BPM)
Input: Group Data and Test Data
Output: Similarity
begin
Start with k group and k test data,
Select the length m of the sequence (seq)
seq = m * unit vector
Match(group data, test data)
Store maximum value as highest similarity measure
Store the vector containing the maximum values.
Repeat for all group – test combination
end
Algorithm: Match Function
Input: group data, test data
Output: maximum similarity
begin
smty = 0
for each position I, in the sequence length
if position I holds no of call info then
ifseq_test(i) = seq_group(i) AND seq_test(i) != 0 then smty+=1
record position,
elseif position I holds duration info then
ifseq_test(i) 1= 0 AND seq_test(i) <= (1+X) * seq_group(i) AND
seq_test(i) >= (1-X) * seq_group(i) then smty += 1
smty(seq_test(i),k)= max {smty (seq_test(i), seqj)}
end

System flow diagram for our proposed system is shown in figure 3.1.

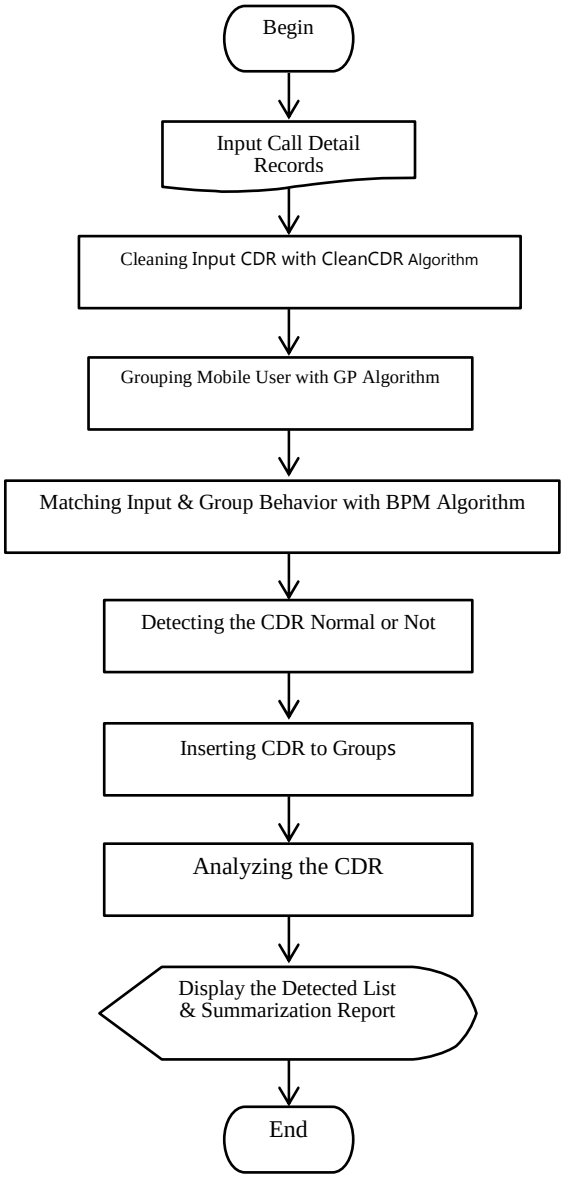

Figure 3.1 System Flow Diagram

4. Implementation of Proposed System

The main objective of this System is to propose an anomalous behavior detection system. The system proposes User Group Partition Algorithm and Behavior Pattern Matching Algorithm to extract anomalous calls from mobile call detail records effectively.

4.1Phase 1: Grouping Mobile Phone User Groups

The goal of this Phase is to effectively classify the call record data based on the attributes of call detail record.The inputs of this Phase are Call Detail Records.The processes that effectively carry out for this Phase are: 1.Cleaning the Call Record Data, 2.Grouping the dataset based on Call Duration and number of calls.The output of this Phase is Call Detail Record Groups with similar behavior.

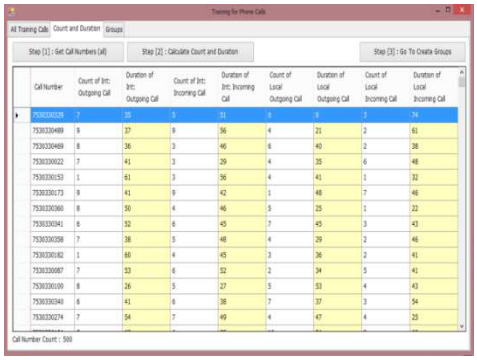

Figure 4.1. Calculating Count and Duration of Mobile Phone Calls

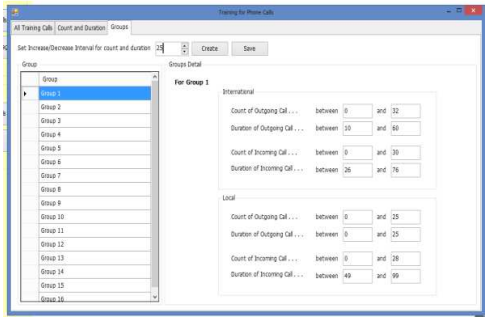

Figure 4.2. Grouping Mobile Phone Calls

4.2 Phase 2: Behavior Pattern Matching

The goal of this Phase is to effectively detect Sudden Change in Behavior of the mobile phone user.The inputs for this Phase are the Call Detail RecordGroups which are the output of Group Partitioning Phase, the first Phase of the System.The processes proposed for this Phase of the System are:Matching new input with Group behavior, Detect the call data if normal or abnormal, Analyzing the Call DetailRecords.The output of this Phase is detected anomalous mobile phone call patterns, the primary goal of the Proposed System.

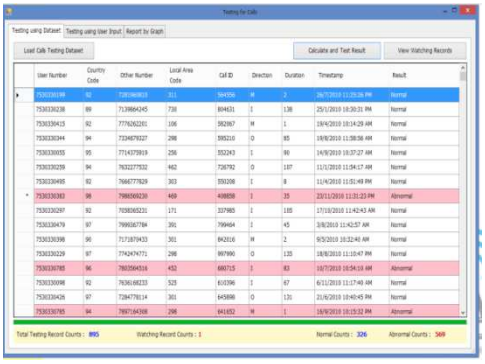

Figure 4.3. Abnormal Call Detail Record Expressed in Red Color

Then the system returns the user watching list of abnormal call detail records as shown in figure 4.3.

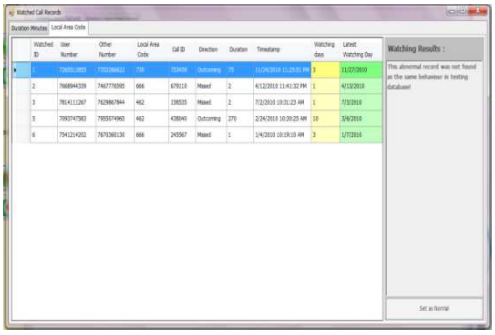

Figure 4.4. Watching List of Abnormal Call Detail Records

5. Conclusion

Call Detail Record Data contain abnormal call record, anomalies that could negatively affect the mobile phone users and network carrier.Detecting anomalous behavior call record is an interesting problem. By using the Proposed Group Partition Algorithm and Behavior Pattern Matching Algorithm, the sudden changed call behavior can be detected effectively.By Using the Summery CDR Algorithm, the summarization report of CDR can be effectively delivered. Detecting the anomalous call behavior can help for many research areas of mobile phone network.

References

A.H. Artail, M. Raydan, Device-aware desktop web pagetransformation for rendering on handhelds.Personal and Ubiquitous Computing 2005; 9(6): 368-380, DOI: 10.1007/s00779-005-0348-5.

[1] G.W. Chow, A. Jones. A Framework for AnomalyDetection in OKL4-Linux Based Smartphones, Proceedings of the 6th Australian Information Security Management Conference, 2008.

[2] B. Sun ,Z. Chen , R. Wang , F. Yu , VCM. Leung.Towards adaptive anomaly detection in cellular mobile networks, Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC06), Vol 2, 2006; 666-670.

[3] B. Sun , Y. Xiao ,K. Wu . Intrusion Detection in CellularMobile Networks. Book chapter in Wireless Mobile Network Security. Springer, 2007; 183-210, ISBN: 0387280405.

[4] I. Naumann, G. Hogben,L. Fritsch , R. Benito ,R. Dean .Security Issues in the Context of Authentication Using Mobile Devices (Mobile eID), European Network and information Security Agency (ENISA), January. 2008.

[5] GSM. World Mobile. Market Data Summary (Q22009). Available at:http://www.gsmworld.com/newsroom/market-data/ market_data_summary.htm (Accessed 16 Feb. 2011).

[6] Mobile World Congress. Visit Kaspersky Lab atMobile World Congress 2009 in Barcelona. Availableat:http://www.kaspersky.com/news?id=207575745 (Accessed 16 Feb. 2011).

[7] M. Landesman. The World's Largest SecurityAnalysis of Real-World Web Traffic, AnnualGlobal Threat Report, ScanSafe STAT. Available at: http://www.scansafe.com/downloads/gtr/2009_AGTR.pdf (Accessed 16 Feb. 2011).

[8] B. Ray. Home Office discusses thief-proof phones. Available at:http://www.theregister.co.uk/2007/05/25/home_office_phone_crime(Accessed 16 Feb. 2011).

[9] C. Kruegel , F. Valeur, G. Vigna. Intrusion Detection and Correlation: Challenges and Solutions. Book chapter Computer security and Intrusion Detection, Springer, 2005.

[10] [11] KK. Singh. Hybrid Profiling Strategy for IntrusionDetection, Department of Computer Science University of British Columbia, 2004.

[11] Y. Moreau , H. Verrelst, J. Vandewalle. Detection ofobile phone fraud using supervised neural networks:A first prototype, Proceedings of the 7th internationalConference on Artificial Neural Networks (ICANN'97)1997; 1065-1070.

[12] D. Buschkes, R. Kesdogan, P. Reichl. How toIncreaseSecurity in Mobile Networks by Anomaly Detection,Proceedings of the Computer Security ApplicationsConference, Phoenix, December. 1998; 3-12.

[13] A. Boukerche,MSMA. Notare. Behavior-Based IntrusionDetection in Mobile Phone Systems. JournalOf Parallel and Distributed Computing 2002; 62(9):1476-1490.

[14] J. Hollm´en. User profiling and classification for frauddetection in mobile communications networks, PhDThesis, Helsinki University of Technology, 2000.

[15] P. Burge, J. Shawe-Tylor. An unsupervised neuralnetwork approach profiling the behavior of mobilephone users for use in fraud detection. Journal of Parallel and Distributed Computing 2001; 61(7): 915-925.

[16] [17]B. Sun, F. Yu ,K. Wu ,VCM. Leung .Mobilitybasedanomaly detection in cellular mobile networks,Proceedings of the ACM wireless security (WiSe '04),Philadelphia, PA, 2004; 61-69.

[17] P.Kumpulainen,K.Htnen.Anomaly Detection AlgorithmTest Bench for Mobile Network Management,Tampere University of Technology, 2008.

[18] A. Bose ,X. Hu ,KG. Shin ,T. Park. Behavioral Detectionof Malware on Mobile Handsets, Proceedings ofthe 6th international conference on Mobile systems,applications, and services (MobiSys '08), USA, June.2008.

[19] L. Liu ,G. Yan ,X. Zhang , S. Chen . VirusMeter:Preventing Your Cellphone from Spies, Proceedingsof the 12th International Symposium on RecentAdvances in Intrusion Detection, Lecture Notes InComputer Science, Springer-Verlag, 2009; 244-264,DOI: 10.1007/978-3-642-04342-0 13.