Privacy-Preserving Fraud Detection via Cooperative Mobile Carriers with Improved Accuracy

Wenyan Yao, Na Ruan*, Feifan Yu, Weijia Jia, Haojin Zhu
Shanghai Jiao Tong University, China
Email: naruan@cs.sjtu.edu.cn

***Abstract*—With the explosive growth of users in mobile carrier, telecommunication fraud causes a serious loss to both of the users and carriers. The academia has an increasing interest in the issue of detecting and recognizing fraudster, and varies strategies have been proposed to prevent the attack and fraudulent activity. However, fraudsters are always inclined to hide their identity and perform the fraudulent activity through different mobile carriers, which makes the previous methods less effective in fraud detection. In this paper, we propose a novel strategy with a high accuracy and security through the cooperation among mobile carriers. We introduce the Latent Dirichlet Allocation (LDA) model to profile users in different carriers. In order to match the fraud accounts, we propose a strategy based on Maximum Mean Discrepancy (MMD) to analyze and compare the distribution of statistical samples. Meantime, during the cooperation of carriers, there is a risk of privacy disclosure. To deal with this weakness, we also demonstrate that our method can detect the fraudulent accounts without leaking the private records and data of user accounts based on the differential privacy.**

***Index Terms*—Fraud Detection, Cooperation Mobile Carriers, Privacy-Preserving, Improved Accuracy**

I. Introduction

With the explosive growth of users in mobile carrier, telecommunication fraud causes a serious loss to both of the users and carriers. In order to detect the fraud activity, many researchers proposed different strategies using machine learning, statistical model and other approach. A good example is Bolton R J., et al [9] described how statistical model can help mobile carriers detecting fraudsters. Weatherford M. [14] focused on neural networks to utilize historical records to generate the patterns of normal user for long-term. Moreover, some industry developed software to detect fraud, such as TransNexus, they developed a system called NexOSS to detect fraudsters who use VoIP network in 2015.

Especially, the academia has an increasing interest in the issue of detecting and recognizing fraudster in mobile carrier, and varies strategies have been proposed to prevent the attack and fraudulent activity. Becker R A., et al [15] introduced many detection strategy, such as Early Threshold-Based Alerting which utilizes historical data to find a threshold to distinguish normal account and fraud account, however, in the real scenarios there are too many users with different behaviors thus this method will predict a normal user as a fraudster mistakenly. They also described the Signature-Based Alerting whose basic idea is to profile the behavior of users in the mobile carrier, which needs a accurate and efficient profiling method. Furthermore. Yusoff M I M., et al [16] proposed a method using statistical model like Gaussian Mixed Model to profile users.

However, there are still some challenges in this area. Firstly, fraudsters are always inclined to hide their identity and perform the fraudulent activity through different mobile carriers, which makes the previous methods less effective in fraud detection. Like others, Olszewski D. [2] proposed a method based on Latent Dirichlet Allocation (LDA) to profile users, then built a automatic threshold to detect fraudster in only one mobile carrier, which can hardly detect the fraudster who hide in multiple mobile carriers. Secondly, in the mobile carrier, there is a large number of data need to be analyzed simultaneously, however, there is only a small number of fraudulent call sample which we can use to learn the behavior of fraudsters. For example, Henecka W., et al [1] proposed a fraud detection across multiple databases, but they only used one feature of call samples (destination) to profile the user, moreover their matching strategy only focused on the distance of two signatures, thus the accuracy of their model need to be improved. Thirdly, if we detect fraudster through cooperation among multiple mobile carriers, they need to exchange data. Therefore, in this process, the attacker can have the chance to get the private call record of individual users, which poses a serious threat to the privacy security of normal users.

To address the above challenges, we propose a novel strategy with a high accuracy and security through the cooperation among mobile carriers. The contribution of this work can be summarized as follows:

1) To thwart the sophisiticated fraudsters who can hide in the multiple mobile carriers, we propose a Cooperative Fraud Detection model based on the cooperation among multiple carriers.
2) To increase the detection accuracy, we propose an efficient and accurate profiling method to profile the behavior of mobile phone user, and we also propose a comprehensive and accurate matching method.
3) To prevent privacy from possible disclosure, we applicate differential privacy to insure that our method can detect the fraudulent accounts without leaking the private records and data of user accounts.

We also conduct a set of experiments and compare with previous work to evaluate the accuracy and the efficiency of the our detection model. We use the ROC (Receiver Operating Characteristic) curves and the value AUROC (Area Under Receiver Operating Characteristic) which is the common method used in the research of fraud to evaluate performace of our model. Moreover, we take the scale of data into consideration, using different parameters to show the influence of different features of datasets. The evaluations show that our detection model also has a high accuracy, effieciency and can prevent privacy from disclosure efficiently.

The remainder of this work is organized as follows. In Section II, we briefly introduce the relevant backgroud knowledge including Latent Dirichlet Allocation (LDA), Maximum Mean Discreapancy (MMD) and Differential Privacy. In Section III, we describe our Cooperative Fraud Detection Model and

*:corresponding author

978-1-5090-6599-8/17/$31.00 ©2017 IEEE

its application scenarios. In Section IV to Section VI, we completely introduce our approach including profiling module, matching module and privacy protection module. In Section VII, we show the evaluation of our work. A conclusion is drawn in Section VIII.

II. Preliminaries

In this section, we introduce the basics of Latent Dirichlet Allocation (LDA) model, Maximun Mean Discrepancy (MMD) and Differential Privacy.

A. Latent Dirichlet Allocation(LDA)

Latent Dirichlet Allocation (LDA) model first proposed by Blei D M., et al [10] to recongnize large-scale document collection or corpus. It is a generative probabilistic model of a corpus and a three-level Bayes model, the basic idea is that the document is a random mixture over latent topic, and each topic is characterized by a distribution over words. Compared to other method, LDA model can deal with large number of data efficiently, thus it is suit for our scenarios. The process of LDA to generate a document is defined as follows:

Definition 1. *For each document w in a corpus D:*
1.Choose $N \sim Poisson(\xi)$
2.Choose $\theta \sim Dir(\alpha)$
3.For each of the N word w_n: Choose a topic $z_n \sim Multinomial(\theta)$. Choose a word w_n from $p(w_n|z_n, \beta)$

Where the **w** denotes a document, w_n denotes the nth word in the document sequence, N denotes the number of words in a document, **z** denotes the topic. Fig.1 shows the graphical LDA model where the boxes are plates representing replicates, the outer plate represents documents, while the inner plate represents the repeated choice of topics and words within a document, and M denotes the number of documents in a topic.

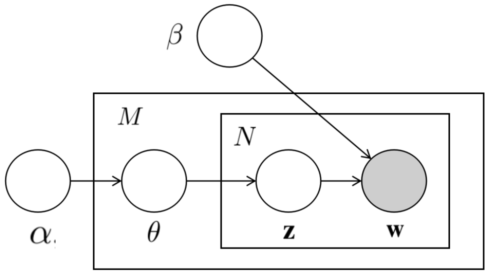


Fig. 1: LDA Model

B. Maximun Mean Discrepancy(MMD)

Arthur Gretton et al [4], proposed a framework for analyzing and comparing distributions, using statistical tests to determine if two samples are drawn from different distributions. Firstly, the problem is defined as follows:

Problem 1. *Let p_x and p_y be Borel probability of meatures defined on a domain $\mathcal{X}$. Given observations $X := \{x_1, \cdots, x_m\}$ and $Y := \{y_1, \cdots, y_n\}$, drawn independently and identically distributed (i.i.d.) from p_x and p_y, respectively, can we decide whether $p_x \neq p_y$?*

To solve this problem, we have the Lemma.1:

Lemma 1. *Let $(\mathcal{X}, d)$ be a metric space, and let p_x, p_y be two Borel probability measures defined on $\mathcal{X}$. Then $p_x = p_y$ if and only if $E_{x \sim p_x}(f(x)) = E_{y \sim p_x}(f(y))$ for all $f \in C(\mathcal{X})$, where $C(\mathcal{X})$ is the space of bounded continuous functions on $\mathcal{X}$.*

Then carriers can choose a rich and general function classses $\mathcal{F}$ to define MMD which can measure the disparity between two samples, the definition is:

Definition 2. *Let $\mathcal{F}$ be a class of functions $f : X \to \mathbb{R}$ and let p_x, p_y, X, Y be defined as above. We define the maximum mean discrepancy (MMD) as:*

$$MMD[\mathcal{F}, p_x, p_y] := sup_{f \in \mathcal{F}}(E_{x \sim p_x}[f(x)] - E_{y \sim p_y}[f(y)]) \tag{1}$$

C. Differential Privacy

Dalenius first articulated a desideratum that foreshadows for databases the notion of semantic security: access to a statistical database should not enable one to learn anything about an individual that could not be learned without access. However, achieving this type of privacy was proved to be impossible facing attackers with auxiliary information. Dwork C [5] proposed a new model, which was named of differential privacy, to ensure that, any given disclosure will be, within a small multicative factor, just as likely whether or not the individual participant in the database, i.e., the presence of an individuals data would not be the cause of information disclosure. The rigorous definition of differential privacy is showed in Definition 3.

Definition 3. *A randomlized function M gives $varepsilon$-differential privacy if for all data sets D_1 and D_2 differenting on at most one element, and all $S \subseteq Range(M)$, we have:*

$$Pr(M(D_1) \in S) \leq exp(\varepsilon) \times Pr(M(D_2) \in S) \tag{2}$$

Differential privacy allows users to interact with the database only by statistical queries. Compared to privacy methods based on publishing perturbed versions of the dataset, such as homomorphic encryption, differential privacy provides much higher efficiency, prevent the unsafety caused by disclosure of encryption keys, and performs better dealing with complicated and various real data and new kinds of attacks. Random noise whose magnitude is chosen as a function of L_1-sensivity is added to each query result to achieve differential privacy when result is numeric. L_1-sensitivity which is showed in Definition 4. is the largest change that a single participant could have on the output to the query function.

Definition 4. *For $f : D \to R_d$, the L_1-sensitivity of f is:*

$$\triangle f = max_{D_1, D_2} ||f(D_1) - f(D_2)|| \tag{3}$$

for all D_1, D_2 differing in at most one element.

III. Cooperative Fraud Detection Model And Attack Model

In this section, we present the application scenairos of our work, propose our cooperative fraud detection model, as well as the attack model including the different fraud forms and possible privacy attack.

A. Application Scenarios

In our work, to achieve the goal that detects the fraudsters accurately and efficiently, we propose a method based on the cooperation of multiple mobile carriers. Our application scenario is showed in Fig.2, where the box with imaginary line denotes the fraud accounts detected by our model in the carrier B, C, D. The application scenario consists of multiple moblie phone carriers, and one of them contains a known fraud list and a database with users data and the other carriers can utilize our model to analyze their data and possible fraud list to detect fraudsters in their own database even the hidden fraudsters.

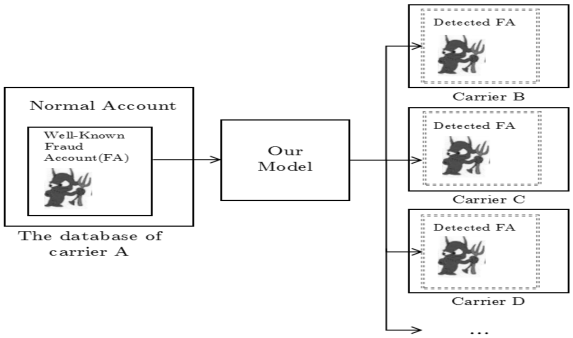


Fig. 2: Application Scenarios

B. Our Cooperative Fraud Detection model

In application senarios of the traditional method, the fraud detection is done in one mobile carrier, which we can assume that the work is done in one database which contains individual call data records (CDR). And the detection model use profiling method to profile the behavior and habit of individual account through the features of their CDRs which consists of e.g. destination, duration, type, cost and so on. Due to the trait of the fraud accounts that is always deviate from the normal user, using classification algorithm with an appropriate threshold can determine the fraudsters.

However, the experienced and sophisticated fraudsters can hide their behavior by changing their pattern of phone call, which made their strange and fraudulent behavior can not be caught by previous fraud detection approach, and it also takes high cost for mobile carrier to detect these fraud accounts. But, the same type of fraudsters always use the similar pattern to do the fraud activity, thus without losing generality we can assume that if a fraudster has accounts in both mobile carrier A and mobile carrier B, then the characteristics of these accounts should be alike. Based on this assumption, we build a cooperative fraud detection model between two carriers, and it is not hard to extend to multiple mobile carriers. The framework of our work is described as follows. Firstly, we extract the CDRs and other data from one carrier, we let it be carrier A. Then through the profiling module based on LDA model, we can profile the behavior of fraud accounts in the carrier A efficiently. After this, in the carrier B, using our match module which is based on MMD to determine the similarity of these accounts. Finally, we can screen out and detect the fraudulent accounts in the carrier B. Our model is showed in Fig.3.

C. Attack model

In this study, we consider the following attack and fraud models:

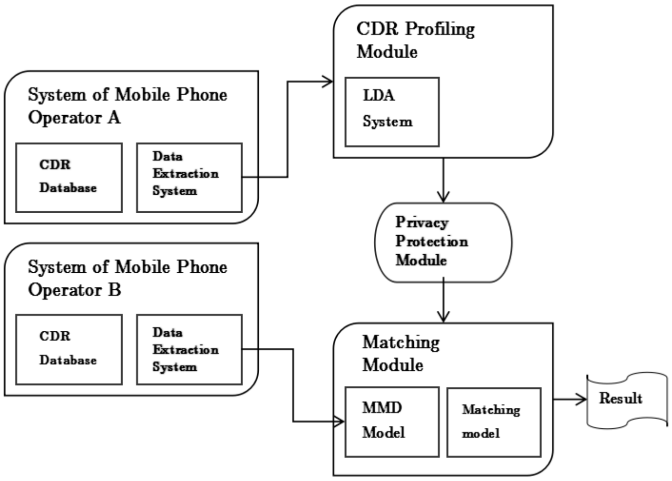


Fig. 3: Cooperative Fraud detection Model

- Conventional fraud: In the type of conventional fraud, fraudsters make phone call to very large range of normal individual accounts to allure normal user to produce extra fees or other fraud means. Thus in this kind of fraud, potential fraudulent accounts always have abnormal behavior and feature, e.g. very high rising or suspicious calling rate, large-scale destination. Because of this, using fraud detection based on profiling method can detect the fraudulent account accurately. However, mobile phone call fraud, differ greatly from credit fraud and insurance fraud, contains large number of data that needs to be analyzed. Moreover, there is only a small number of fraudulent call samples that can be used to reference. We propose a model also based on user profiling using LDA model, and introduced a kernel method to match profile which ensure our method can detect fraudsters accurately. Meantime, our method is based on cooperation of multiple mobile carrier which means we can solve the problem of lack of fraudulent call samples.
- Subscription fraud: In this type of fraud, fraudsters change their device and service with no intent to pay, e.g. fraudsters sign up a new line or a new account in a mobile carrier from original mobile carrier to continue fraud activity. In this scenarios, the traditional detection schemes which are based on classifying the behavior of accounts in only one mobile carrier can not detect the fraudulent account accurately and efficiently. We propose an accurate matching method based on MMD which make it can detect this kind of fraudsters efficiently.
- Privacy attacks: In the type of privacy attacks, attackers pose as different carriers who make statistical queries on an carriers database, and make queries to get information of an accounts specific call record. For instance, attackers make queries for the sum of the call duration of the first 100 and 101 call records of an account, to find out the duration of the 101th call record. Even if the carrier only answer the queries for statistics of all call records of an account, the attacker can repeatedly make such queries, and get the information of a new record with high probabilities when the answers change. Considering the necessarity of fraud detection in time, it is not practical to stop answering after an answer for a long time. The cooperative fraud detection model requires

large amount of call records to improve the accuracy, which makes only answering the queries for the first 100 call records inappropriate. We proposed a method based on differential privacy to prevent this kind of attacks.

IV. Profiling Module

In this section, we propose a profiling module based on Latent Dirichlet Allocation (LDA) to profile the behavior of accounts.

A. Notations

Here, we give notations of variables:
K : Number of latent class
ξ : The parameter of Poisson distribution
α : It is the parameter of prior Dirichlet distribution over the latent class
V : The number of features
β : It is $K \times V$ matrix, whose rows denote the parameters of the Multinomial distributions
a : Denote the feature vector.
N : Denote the number of iterations.
Γ : Denote the Gamma Function.
z : Denote the class
z_i : Denote the ith class

B. Using LDA model to profile user

The Signature Based Fraud Detection is an efficient method for mobile carrier to detect fraudulent account, however, the accuracy of this method relies on an efficient and accurate signature generating method. The main problem of profiling behavior of users is that how to utilize historical data to reflect the behavior pattern of the individual account, as well as to predict the following user behavior since we can distinguish different type of users only in this situation.

We use LDA model to profile users behavior in the mobile carrier. Due to this model is a generative probabilistic model for collections of discrete data. Its original goal is to find the short description of the members with in a group that enable efficient processing of large collections, meantime, preserving the essential statistical relationships which makes it is appropriate for profiling mobile users.

It is a three-level hierarchical Bayesian model. We assume accounts are represented as finite mixture over latent class, and the classes are represented by a distribution over multinomial. In our method, we defined 4 features which are calling type, duration, time and cost as multinomial. Thus, an account is represented by probabilities vector of the K class, and a class is represented by probabilities of the 4 features.

An account can be generated using following procedure. The hiden variables θ and z are edtimated using varational approximation. A k-dimensional Dirichlet random variable θ can take values in the $(k-1)$-simplex. A k-vector θ lies in the $(k-1)$-simplex if:

$$\theta_i \geq 0, \sum_{i=1}^{k} \theta_i = 1 \tag{4}$$

And has following probability density on this simplex:

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \theta_1^{\alpha_1 - 1} \cdots \theta_k^{\alpha_k - 1} \tag{5}$$

Algorithm 1 generating accounts data using LDA

```
Input: $\xi$, $\alpha$, $\beta$
Output: $a$
randomly draw the number of iterations $N \sim Poisson(\xi)$;
randomly draw the parameter for generating account from
  the class distribution $\theta \sim Dirichlet(\alpha)$;
for each of the $N$ multinomials $a_i$ do
    draw the class $z_i$, $z \sim Multinomial(\theta)$;
    draw the feature $a_i$ from $p(a|z_i, \beta)$ which is a multinomal probability distribution vector of features $a$ in the class $z_i$, which is in the row of the matrix-parameter $\beta$
end for
return $a$;
```

And the parameters α and β of this model can be estimated by using the EM algorithm(α and β maximize the (marginal) log likelihood of the data).

Given the parameters α and β, the joint distribution of a latent class mixture θ, and z. The vector of V features a is given by:

$$p(\theta, z, a|\alpha, \beta) = p(\theta|\alpha) \prod_{i=1}^{K} p(z|\theta) p(a|z_i, \beta) \tag{6}$$

Then we can define the marginal distribution of an account in the mobile carrier as:

$$p(a|\alpha, \beta) = \int p(\theta|\alpha) (\prod_{i=1}^{N} \sum_{i=1}^{N} p(z_i|\theta) p(a_i|z_i, \beta)) \tag{7}$$

For each account, we can calculate the distribution as follows:

$$\begin{aligned} p(a_{LDA}) &= \int_{\triangle} p(a|\theta) p(\theta|c_n) \mathrm{d}\theta \\ &= \sum_{k=1}^{K} p(a_{LDA}|k) E_{p(\theta|c_n)}\{\theta_k\} \\ &\approx \sum_{k=1}^{K} p(a_{LDA}|k) E_{D(\theta|c_n)}\{\theta_k\} \\ &= \sum_{k=1}^{K} p(a_{LDA}|k) \frac{\gamma_{kn}}{\sum_{i=1}^{K} \gamma_{in}} \end{aligned} \tag{8}$$

where a_{LDA} represents an account, c_n denotes the phone calls of this account, γ_{in} denotes the variational free parameter. $E_{D(\theta|c_n)}\{\theta_k\}$ denotes expectation of discrete random variable θ_k with respect to $D(\theta|c_n)$. $E_p\{\theta_k\}$ denotes expectation of discrete random variable θ_k with respect to p.

V. Matching module

In this section, we propose a matching approach between two profiles of mobile phone accounts based on MMD. In our fraud detection model, a key challenge is that we need to distinguish even tiny difference between two profiles to achieve high detection accuracy, MMD can predict the similarity of two distribution efficiently and does not rely on any assumption of distribution. Meantime in our application scenarios, without losing generality, we can assume if we find that two samples generated by our profiling module from two mobile accounts are derived from the same distribution, we can predicate they are the same type of user. Because the same type of fraudsters always use the same or similar fraud pattern, and it can save cost.

A. Notions

Here, we list some variables and function used in this section:

P_i : the profile of the ith account in mobile carriers.

p_i : the distribution of P_i.

x_i : the ith samples of the profile P_x, in our method we use different time quantum such as 1/12/2016 to 5/12/2016 and 6/12/2016 to 10/12/2016.

y_i : the ith samples of the profile P_y.

$\mathscr{F}$: the function class of f.

$\mathscr{H}$: Reproducing Kernel Hilbert Space.

$\mathscr{X}$: Compact Metric Space.

k : the Gaussian Kernel Function.

x_c : tbe center of the kernel function.

σ : the width of kernel funtion which can control the influence range of kernel function.

x^* : the normalized features of the profile P_x.

$Fraud_A$: the fraud list of carrier$A$.

$Fraud_B$: the fraud list of carrier$B$.

$threshold$: the parameter of our model which control the tolerality of our model.

B. Our matching method

In our work, we generate the profile of individual account in mobile carriers. Thus for every accounts i we have the generating profile from the profiling module, P_i. Essentially, they are the statistical samples which are derived from typical distribution p_i. We have $P_i \sim p_i$, and they may be derived from the same distribution or not. If we want to detect the fraudsters in one moblie phone carriers, the efficient way is to find the alilke fraudsters in other moblie phone carriers. Thus carriers need to compare P_i and P_j to determine whether they are the same type of user, which means whether $p_i = p_j$.

if we choose two profile samples in the database such as P_x, P_y thus we have:

$$\begin{aligned} P_x &:= [x_1, x_2, \dots x_m] \\ P_y &:= [y_1, y_2, \dots y_n] \end{aligned} \tag{9}$$

Then we assume there is a unspecified function class $\mathscr{F}$, and the function in the $\mathscr{F}$ can help us to measure the disparity between p_i and p_j. According to Definition.2, we have the MMD of these samples as 1. And we can formulate a biased empirical estimate of the MMD is:

$$MMD_b[\mathcal{F}, p_x, p_y] := sup_{f \in \mathcal{F}}(\frac{1}{m}\sum_{i=1}^{m} f(x_i) - \frac{1}{n}\sum_{i=1}^{n} f(y_i)]) \tag{10}$$

In order to estimate the MMD, carriers need to find an apporiate function class which is rich enough to generally indentify whether $p_x = p_y$, and it is needed to be restrictive enough to provide useful finite sample estimates. If the class $\mathscr{F}$ is the unit ball in a Reproducing Kernel Hilbert Space (RKHS) $\mathscr{H}$, the empirical MMD can be computed efficiently. We have the Theorem.1:

Theorem 1. *Let $\mathscr{F}$ be a unit ball in a universal RKHS $\mathscr{H}$, defined on the compact metric space $\mathscr{X}$, with associated kernel $k(.,.)$. Then $MMD[\mathcal{F}, p_x, p_y] = 0$ IF and only if $p_x = p_y$.*

In order to exhibit the maximum discrepancy between two distributions, a witness function f is needed. In our model, the population $f(x)$ and its empirical estimate $\hat{f}(x)$ are:

$$\begin{aligned} f(x) &\propto \langle \phi(x), \mu[p_x], \mu[p_y] \rangle \\ &= E_{x' \sim p_x}[k(x, x')] - E_{y' \sim p_y}[k(x, y')] \end{aligned} \tag{11}$$

$$\begin{aligned} \hat{f}(x) &\propto \langle \phi(x), \mu[P_x], \mu[P_y] \rangle \\ &= \frac{1}{m}\sum_{i=1}^{m} k(x_i, x) - \frac{1}{n}\sum_{i=1}^{n} k(y_i, x) \end{aligned} \tag{12}$$

Where the $k(x_i, x)$ is a kernel function. In order to fomulate the accurate MMD between p_x and p_y, we need a comprehensive kernel function, in our model, we choose the Gaussian Radial Basis Function (RBF) Kernel, which is defiend as follows:

$$k(x, x_c) = Exp(\frac{-(||x - x_c||)^2}{(2\sigma)^2}) \tag{13}$$

In our model, to assure the accuracy of MMD, an appropriate kernel width σ is needed. However, if we set $\sigma = 0$ or set $\sigma \to \infty$ the empirical MMD is zero for any two distribution samples. Without lossing the generality, we set the σ to be the median distance among all point pairs in the P to avoid the extreme situation.

Another problem is that in the kernel function, the value of every dimension in the vector P should be in the same value range such as $[0, 1]$. However, in our application scenarios, the duration recorded as seconds is much bigger than other features, which will make other features lose their influence. Thus the Min-Max Normalization can be used to do the linear transformation, which will make the value maps to $[0, 1]$. The transformation is defined as follows:

$$x^* = \frac{x - min(x)}{max(x) - min(x)} \tag{14}$$

Finally, according to the (10), (12), (13), (14), we can formuate the Maximum Mean Discrepancy of any two profiles generated by our profiling module.

Then carriers need to predict an account in carrier B whether matchs a fraud account in carrier A. For every fraud accounts in the carrier A, we determine the MMD with the every accounts in the carrier B, then we retain minimum MMD. If the minimum MMD is smaller than a value, we can predict that the account in the carrier B is a fraudster. The Algorithm 2 describes the pseudocode.

Algorithm 2 Predict the fraud account

```
Input: profile $P_i$ for every accounts, $Fraud_B$, $threshold$
Output: $Fraud_A$.
set the $minimum = \infty$
for each account $i$ in carrier$A$ do
    for each account $j$ in $Fraud_B$ do
        determine the $MMD(P_i, P_j)$ between two accounts $i$ and $j$
        if the $MMD(P_i, P_j)$ is lower than the $minimum$ of $i$ then
            update the $minimum$
        end if
    end for
    if the $minimum$ is lower than $threshold$ then
        add account $i$ into $Fraud_A$
    end if
end for
return $Fraud_A$;
```

VI. Privacy Protection Module

In this section, we propose a privacy protecting method based on differential privacy during carriers make information interaction.

A. Notions

Here, we list some variables and functions used in this section. Many of the variables and functions that have been used in last two sections are also used in this one, we do not repeat explaning them unless their meaning changes.

$G_k(S)$: the sum of k powers of all elements in set S, e.g. $G_2(a,b,c) = a^2 + b^2 + c^2$

$x_{i,j}$: the jth feature of x_i

$X_{i,j}$: the jth feature of X_i

$Y_{i,j}$: the jth feature of Y_i

B. The method based on differential privacy

As showed above, in order to compute the MMD between an account in carrier A and an account in carrier B, A and B need to show information about the profile of the accounts to each other, as the estimate of the witness function of MMD requires that. However, it is necessary for carriers to protect their users privacy, i.e., carrier A cannot get exactly the features of an accounts profile in carrier B. We hope the estimate of the witness function of MMD can be expressed as expression of statistics on an accounts profile. Thus, carrier A makes queries for statistics on the account in carrier B to compute the estimate of the witness function value on an accounts profile in opertaor A, and provides B with statistics which B needs to compute the estimate of the witness function value on that accounts profile in carrier B, and the MMD is calculated out without directly showing accounts profile to each other. Then we add noise to the query results to ensure that privacy attackers cannot get the atttibutes of an accounts specific call record, based on different privacy.

First we show that the estimate of the witness function of MMD can be approximately expressed as expression of statistics on the accountsprofile such as $G_k(x_i)$ and $G_k(y_i)$. We only need to show (12) can be expressed in this way. In fact, it is computed by carrier A, and A knows all x_i, thus only need to show that (12) can be estimated without the knowledge of accurate value of y_i. Let $Y_k = \frac{y_k}{2\sigma}$, $X_j = \frac{x_j}{2\sigma}$, and according to (13), we have:

$$f(\hat{x_j}) = \\ \frac{1}{m}\sum_{i=1}^{m} Exp((||X_i - X_j|||)^2) - \frac{1}{n}\sum_{k=1}^{n} Exp((||Y_k - X_j||)^2) \tag{15}$$

As we mentioned above, we set σ to be the median distance among all point pairs. Because A doesnt know the exact value of y_k, A regards all x_i as P. If the account in carrier A is the same kind of frauster as the fraud account in B that is compared with the account in A, the distance between y_i and x_j is in the range of distances between all other x_i and x_j with very high probablity. Therefore, in this case, for all Y_k, we have:

$$||Y_k - X_j|| \leq 1 \tag{16}$$

Consider series expansion, we have:

$$Exp(t) = \sum_{i=0}^{\infty} \frac{t^i}{i!} \tag{17}$$

Consider a function $r(t)$:

$$r(t) = \frac{Exp(t) - (1 + t + \frac{t^2}{2} + \frac{t^3}{6})}{Exp(t)} \tag{18}$$

It is easy to derive that $r'(t) = Exp(-t)\frac{t^3}{6} > 0,\ t > 0$. Thus when t is not larger than 1, the biggest $r(t)$ is $r(1)$, less than 2%. We use $1 + t + \frac{t^2}{2} + \frac{t^3}{6}$ as an approximate estimate of $Exp(t)$ to compute $\hat{f}(x_j)$. As shown above, the error of computing is less than 2% of the largest $k(y_k, x_j)$ with very high probablity. Comparing with the large difference between MMD of two different kind of accounts and MMD of two accounts of the same kind, this error is little, which causes little influence on the detecting of fraud accounts that are really the same kind of fraudsters as the fraud account in B, with statistics on whose profile MMD is computed.

We use K features of user in the mobile carriers, the kernel function can be transformed as follows:

$$\begin{aligned} & Exp((||Y_k - X_j||)^2) \\ & \approx 1 + (||Y_k - X_j||)^2 + \frac{(||Y_k - X_j||)^4}{2} + \frac{(||Y_k - X_j||)^6}{3} \\ & = 1 + \sum_{s=1}^{K}(Y_{k,s}^2 - 2Y_{k,s}X_{j,s} + X_{j,s}^2) + \sum_{s=1}^{K}\sum_{t=1}^{K} \\ & (Y_{k,s}^2 Y_{k,t}^2 + X_{j,s}^2 X_{j,t}^2 - 4Y_{k,s}^2 Y_{k,t} X_{j,t} - 4X_{j,s}^2 X_{j,t} Y_{k,t} + \\ & 2X_{j,s}^2 Y_{k,t}^2) + \sum_{s=1}^{K}\sum_{t=1}^{K}\sum_{r=1}^{K}(Y_{k,s}^2 Y_{k,t}^2 Y_{k,r}^2 - 2Y_{k,s}^2 Y_{k,t}^2 Y_{k,r} X_{j,r} \\ & + \cdots + 2Y_{k,s}^2 X_{j,t}^2 X_{j,r}^2) \end{aligned} \tag{19}$$

Therefore, $\hat{f}(x_j)$ can be computated given the values of $G_k(Y_{i,s})$ and other statistics on Y_i without using the exact value of Y_i. As mentioned above, attackers can get the value of $Y_{k,s}$ by querying for $\sum_{l=1}^{k-1} Y_{l,s}$ and $\sum_{l=1}^{k} Y_{l,s}$. We have to at least add noises to the results of these queries. As for other statistics such as $\sum_{k=1}^{n} Y_{k,s}^2 Y_{k,t}^2 Y_{k,r}^2$, attackers cant get the value directly in the same way. Combining the results of different queries and solving the equation group to get the information is out of our privacy attack model, and the computation complexity is high when n is large. Thus, we only add noises to the results of queries for $\sum_{l=1}^{k-1} Y_{l,s}$, which also improves the availability of estimation result of MMD in this way.

Lets consider the details of adding noises.

Theorem 2. *For a query $f : D \leq R^d$, the mechanism K_f that adds independently generated noise L with distribution*

$$Lap(0,\sigma) : Pr(L = x) = \frac{1}{2\sigma}Exp(-\frac{||x||}{2\sigma}) \tag{20}$$

it gives $\frac{\triangle f}{\sigma}$-differential privacy.

Its important to note that $\sum_{l=1}^{k} Y_{l,s}^q$ where $(q > 1)$ can be computed with the values of $\sum_{l=1}^{k} Y_{l,s}$ and some symmetric polynomials of $Y_{l,s}$. Since we need to ensure the availability of estimation result of MMD, we hope to add noises to as few results of queries as possible. We consider adding noises to the result of query for $\sum_{l=1}^{k} Y_{l,s}$, and computing the result of query for $\sum_{l=1}^{k} Y_{l,s}^2$ based on the perturbed result of $\sum_{l=1}^{k} Y_{l,s}$ and the real value of $\sum_{1 \leq l < m \leq k} Y_{l,s}Y_{m,s}$. Attackers can not get the value of $Y_{i,s}$ directly from the results of quests for $\sum_{1\leq l<m\leq k} Y_{l,s}Y_{m,s}$ and $\sum_{1\leq l<m\leq k-1} Y_{l,s}Y_{m,s}$. However, because we have:

$$Y_{k,s} = \frac{\sum_{1\leq l<m\leq k} Y_{l,s}Y_{m,s} - \sum_{1\leq l<m\leq k-1} Y_{l,s}Y_{m,s}}{\sum_{l=1}^{k-1} Y_{l,s}} \tag{21}$$

the numerator can be calculated out accurately, and though attackers can only get the perturbed value of denominator, it cannot be too far from the real value, or would result in large error in estimation of MMD. Thus, $Y_{k,s}$ calculated in this way is close to the real value of it.

Theorem 3. *Let M_i each provides ε-differential privacy. $M(M_1(D), M_2(D), , M_n(D))$ provides $\sum_{i=1}^{n} \varepsilon$-differential privacy.*

Therefore, we can add noises as follows. For the result of query for $\sum_{l=1}^{k} Y_l^q$ from carrier A, carrier B answers $\sum_{l=1}^{k} Y_l^q + noise$.

Since the estimation of MMD also requires B questing A for statistics. Both A and B adding noises to their results may make the noise in the estimate of the MMD superfluous. Thus we consider A and B contributing partial noises such that the aggregated noise guarantees different privacy.

Lemma 2. *Laplace distributed random variable $L \sim Lap(0, \sigma)$ can be simulated by the sum of $2n$ random variables as follows:*

$$L = \sum_{i=1}^{n} (G_i - H_i) \tag{22}$$

where G_i and H_i are independent Gamma distributed random variables with densities following the formula:

$$Pr(G_i = x) = Pr(H_i = x) = \frac{(\frac{1}{\alpha})^{\frac{1}{n}} x^{\frac{1}{n-1}} e^{-\frac{x}{\alpha}}}{\Gamma(\frac{1}{n})} \tag{23}$$

Γ *is the Gamma Function.*

According to Lemma 2, carrier A and carrier B can add Gamma noises to their results of queries, such that the aggregated noise in the estimate of MMD is Laplacian noise.

VII. Performance Evaluation

In this section, to evaluate the performance of our Cooperative Fraud Detection Model and the Privacy Protection Module, we conduct a set of experiments with different datasets and simulations using Mathematica, MATLAB, PYTHON and C++. In the following, we present details of our evaluations and show the results of these evaluations. We also compare the results with other detection methods and matching schemes to evaluate the performance of our work, as well as influence of different features on accuracy.

A. Evaluation Settings

1) Cooperative Fraud Detection Model: In our evaluation, we use original real-world data to generate a large number of data using different distribution and devided them into six groups which have the different data scale. We set six groups of experiments. We take scale of the dataset and number of accounts into consideration, thus in our experiments we will show the different performances of the different data scale. Meantime, in the process of simulation of CDRs, we will use the same distribution with different parameters for some accounts to evaluate influence of this factor. The details of the number of accounts, CDRs, etc. are in the Table I, where N denotes the number of experiments, Num_a denotes the number of accounts, Num_f denotes the number of fraud accounts, Num_c denotes the number of average CDRs in an account, Num_t denotes the number of types of fraud accounts, Num_s denotes the number of features of accounts.

TABLE I: The Data scale of Datasets

N	Num_a	Num_f	Num_c	Num_t	Num_s
1	1000	15	100	15	5
2	1000	30	100	15	5
3	2000	15	100	15	5
4	1000	15	200	15	5
5	1000	15	200	5	5
6	1000	15	200	15	3

TABLE II: The Number of features

N	duration	type	time	cost	dial or answer
1	✓	✓	✓	✓	✓
2	✓	✓	✓	✓	✓
3	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓
5	✓	✓	✓	✓	✓
6	✓	✓	×	×	✓

Moreover, we take number of features of an account in the mobile carrier into consideration, we set different features in the our experiment. And the features we used are in the Table II, where N denotes the group number of the experiments, ✓ denotes this experiment contains this feature, × denotes that this experiment does not contain this feature and we divide time into morning, noon, afternoon, evening, and midnight.

2) Privacy Protection Module: We use the data of 4th experiment to set the stimulation to evaluate the influence of noise on the MMD result. The parameters are showed in the Table.III. The median kernel width σ is 0.191, and the averagex_i of a fraud account is 0.32, and in our model, the value range of x_i is from 0.18 to 1 and y_k is from 0 to 1. The $noise$ which denotes $\frac{noise}{\sum y_k}$ is varied from 0 to 1.

TABLE III: Parameters of kernel and noise

σ	$x_{average}$	range of y_k	range of x_i	range of $noise$
0.191	0.32	0∼1	0.18∼1	0∼1

B. Evaluation Results

In this subsection, we will present AUROC of our evaluation firstly. Secondly, we will compare our work with previous work in ROC. Thirdly, we will discuss influence of different parameters. Finally, we will analyze the influence of noise to our model.

1) Cooperative Fraud Detection Model: Firstly we present AUROC value of six experiments to show the accuracy of our model, and the average value is higher than the previous work which means that our model can detect the fraud accurately. The AUROC values of six experiments are in Fig.4, the imaginary line denotes the $AUROC = 0.966$ and we can find that experiment 1 and 2 have the similar performance, the AUROC of experiment 4 and 5 are higher than experiment 1 and 2, but accuracy of experiment 3 and 6 are not very high which we will discuss later.

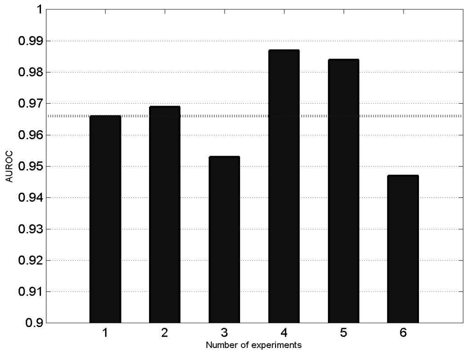


Fig. 4: AUROC

We present the ROC curves compared with the work of Henecka W., et al[1]. The ROC curves are showed in Fig.5. We can find that they used the different profiling model and matching model with us, and in Fig.5, when we detect 80% fraudsters, our matching model have the lower false rate, thus the accuracy of our model is higher. The first reason is that we use the LDA model which can generate the probalistic profile from the historical record with multiple features, however they only use the destination to profile user. Secondly, compared with the different matching model, our model based on MMD have the higher accuracy, because MMD method can determine the similarity without relying on the parameter of the distribution and the appropriate kernel width can increase the convergence rate of two samples, and it does not require density estimates as an intermediate step.

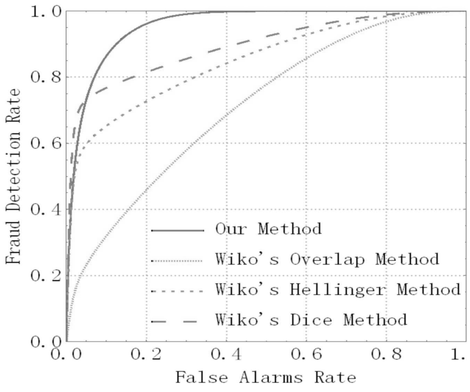


Fig. 5: ROC Curves of Our Model Compared with Wilko's

Thirdly, we compare our model with the previous work based on LDA in only one mobile carrier [2] using the 5th experiment, because the data scale of the 5th experiment is similar to theirs. Fig.6 shows the ROC curves. It shows that when we have the same detection rate, our model have lower false rate, when we detect whole of fraudsters, our model have the lower false rate than theirs. Moreover, AUROC of our model (0.987) is higher than theirs (0.967). However, the gap between two model is not very large, because they also use the LDA model which can utilize the historical record efficiently to profile the users in the mobile carriers. The reason of this gap is that in our application envrionment, there is a group of fraudstrs that hide in multiple carriers and they change their behavior which make they are hard to be detected by detection model based on one carrier. But, in our model, we utilize the cooperation of multiple carriers to deal with this kind of fraudsters.

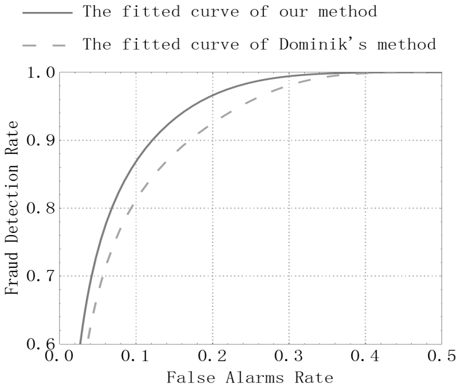


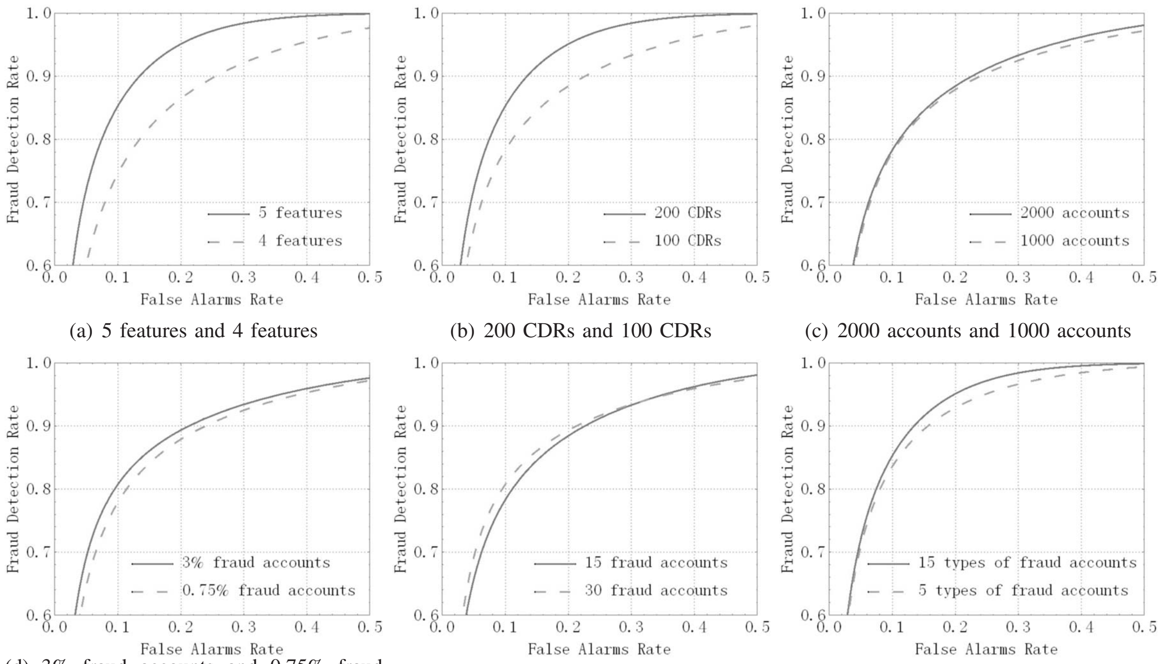
Fig. 6: ROC Curves of Our Model Compared with Dominik's

Finally, we evaluate the influence of the factors of data such as the number of accounts to our model through comparing the performance of two experiments which only have one different feature.

1) The AUROC of experiment1 (0.966) is lower than experiment2 (0.969) in Fig.7(e), however, the gap is very small which means the number of fraud accounts does not affect the performance of our model.
2) The AUROC of experiment3 (0.953) is lower than experiment1 (0.966), the reason is that the number of norm accounts is far more than fraud accounts thus more accounts in a data set increase the possibility that some normal accounts have the same behavior with the fraud accounts, which incur that we predict mistakenly some norm accounts as fraud accounts, which can find in Fig.7(c) and Fig.7(d).
3) The AUROC of experiment4 (0.987) is higher than experiment1 (0.966) in Fig.7(b), which means the higher number of CDRs for an account will have better performance, the possible reason is that more CDRs can help us profiling the user accurately.
4) The AUROC of experiment4(0.987) is higher than experiment6(0.947) in Fig.7(a), which means more features in an account can have the better accuracy, the possible reason is that more features can help our matching model comparing two account comprehensively.

Conclusively, more information of fraud users and normal users, as well as more features of accounts can help us detect fraudsters accurately. To understand the influence of these factors intuitively, Fig.7 shows the ROC curves comparing every two experiments which have only one different features.

2) Privacy Protection Module: In our model, we add noise to avoid attackers to get private CDR data, however, the noise also can have influence on the exact result of MMD. Thus, we did the stimulation to evaluate the influence of noise on the result of MMD. We draw the $noise$ from the Laplace distribution. The results are described in Fig.8, where $noise$ denotes $\frac{noise}{\sum y_k}$, y denotes the value of y_k, and the $FalseRate$ denotes the ratio that the noise affect the result of MMD. We can find that if the y is varied from 0 to 1, the stronger noise have the stonger influence on the MMD result. Therefore, carrier need to control the $\frac{noise}{\sum y_k}$ lower than 0.1 to insure that the result of MMD do not affect the accuracy of our detection model, meantime insure the attackers can not get the private data through query $\sum y_{k-1}$ and $\sum y_k$.



(a) 5 features and 4 features

(b) 200 CDRs and 100 CDRs

(c) 2000 accounts and 1000 accounts

(d) 3% fraud accounts and 0.75% fraud accounts

(e) 30 fraud accounts and 15 fraud accounts

(f) 15 types of fraud and 5 types of fraud

Fig. 7: ROC Curves with Different Data

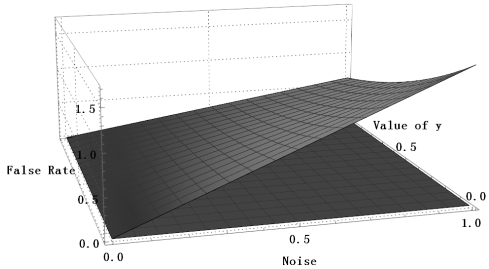


Fig. 8: Influence of Noise to MMD Result

VIII. Conclusion

In this paper, we proposed an efficient cooperative fraud detection among mobile carriers. We succeeded in protecting private data of users in mobile carriers during cooperation of multiple carriers. Through a set of comprehensive evaluations, we find that our detection have the improved accuracy, meantime, our model does not affect detection accuracy during applying privacy protection. The accuracy of our method is significantly higher than previous work, meantime, our method have higher efficiency.

Furthermore, firstly, our future research may be focused on applying our model in bigger dataset. Moreover, we plan to evaluate the efficiency of our privacy protection scheme quantificationally further. We will further study the influence of different noise generation method to our model. Finally, we will try to extend our model to general scenarios to solve other fraud problem.

Acknowledgements

This work is supported by Chinese National Research Fund (NSFC) Key Project No. 61532013; National China 973 Project No. 2015CB352401; Shanghai Scientific Innovation Act of STCSM No.15JC1402400 and 985 Project of Shanghai Jiao Tong University with No. WF220103001; National Science Foundation of China (NSFC) (No: 61672350)

References

[1] Henecka W, Roughan M. *Privacy-Preserving Fraud Detection Across Multiple Phone Record Databases*, IEEE Transactions on Dependable and Secure Computing, 2015, 12(6): 640-651.

[2] Olszewski D. *A probabilistic approach to fraud detection in telecommunications*, Knowledge-Based Systems, 2012, 26: 246-258.

[3] Xing D, Girolami M. *Employing Latent Dirichlet Allocation for fraud detection in telecommunications*, Pattern Recognition Letters, 2007, 28(13): 1727-1734.

[4] Gretton A, Borgwardt K M, Rasch M J, et al. *A kernel two-sample test*, Journal of Machine Learning Research, 2012, 13(Mar): 723-773.

[5] Dwork C. *Differential privacy: A survey of results*, International Conference on Theory and Applications of Models of Computation. Springer Berlin Heidelberg, 2008: 1-19.

[6] Tseng V S, Ying J C, Huang C W, et al. *FrauDetector: A Graph-Mining-based Framework for Fraudulent Phone Call Detection*, Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015: 2157-2166.

[7] Becker R A, Volinsky C, Wilks A R. *Fraud detection in telecommunications: History and lessons learned*, Technometrics, 2012.

[8] Girolami M, Kabn A. *Sequential activity profiling: latent Dirichlet allocation of Markov chains*, Data Mining and Knowledge Discovery, 2005, 10(3): 175-196.

[9] Bolton R J, Hand D J. *Statistical fraud detection: A review*, Statistical science, 2002: 235-249.

[10] Blei D M, Ng A Y, Jordan M I. *Latent dirichlet allocation*, Journal of machine Learning research, 2003, 3(Jan): 993-1022.

[11] Inan A, Kantarcioglu M, Ghinita G, et al. *Private record matching using differential privacy*, Proceedings of the 13th International Conference on Extending Database Technology. ACM, 2010: 123-134.

[12] Dwork C. *A firm foundation for private data analysis*, Communications of the ACM, 2011, 54(1): 86-95.

[13] Haeberlen A, Pierce B C, Narayan A. *Differential Privacy Under Fire*, USENIX Security Symposium. 2011.

[14] Weatherford M. *Mining for fraud*, IEEE Intelligent Systems, 2002, 17(4): 4-6.

[15] Becker R A, Volinsky C, Wilks A R. *Fraud detection in telecommunications: History and lessons learned*, Technometrics, 2012.

[16] Yusoff M I M, Mohamed I, Bakar M R A. *Fraud detection in telecommunication industry using Gaussian mixed model*, 2013 International Conference on Research and Innovation in Information Systems (ICRIIS). IEEE, 2013: 27-32.