LETTER

One-Class Naïve Bayesian Classifier for Toll Fraud Detection

Pilsung KANG[†a)], ***Member***

SUMMARY In this paper, a one-class Naïve Bayesian classifier (One-NB) for detecting toll frauds in a VoIP service is proposed. Since toll frauds occur irregularly and their patterns are too diverse to be generalized as one class, conventional binary-class classification is not effective for toll fraud detection. In addition, conventional novelty detection algorithms have struggled with optimizing their parameters to achieve a stable detection performance. In order to resolve the above limitations, the original Naïve Bayesian classifier is modified to handle the novelty detection problem. In addition, a genetic algorithm (GA) is employed to increase efficiency by selecting significant variables. In order to verify the performance of One-NB, comparative experiments using five well-known novelty detectors and three binary classifiers are conducted over real call data records (CDRs) provided by a Korean VoIP service company. The experimental results show that One-NB detects toll frauds more accurately than other novelty detectors and binary classifiers when the toll frauds rates are relatively low. In addition, The performance of One-NB is found to be more stable than the benchmark methods since no parameter optimization is required for One-NB.

key words: *one-class Naïve Bayesian classifier, toll fraud detection, genetic algorithm, novelty detection*

1. Introduction

Voice-over-IP (VoIP) is defined as a class of products that allow advanced communication services over data networks [1]. Recently, VoIP technology has been rapidly adopted by both consumers and enterprises because it can provide lower costs through equipment consolidation and help create new business models by offering greater flexibility and more features than traditional telephony services [2], [3]. However, VoIP technology is more vulnerable to potential misuses, such as denial of service or service abuse, due to the high-level complexity of its architecture, protocols, and implementation [3].

Among the potential misuses, toll fraud attack, which refers to unauthorized access and use of a VoIP network, is the most serious problem for the VoIP service provider because it utilizes the organization bandwidth and incurs a heavy cost [4]. According to a world-wide survey conducted by the Communications Fraud Control Association (CFCA), estimated losses due to fraud range from US$ 54.4 to 60 billion [4], and more than half of these losses are associated with toll fraud attacks. Conventional remedies for toll fraud attacks include firewall configurations and port protections. Although a set of pre-defined anti-fraud systems are being implemented, in practice they cannot detect all types of toll frauds. Therefore, it becomes crucial to identify the fraud as early as possible, even if this is done only after a call has been completed, for the sake of billing evidence or adding monitoring rules.

Until now, most anti-toll fraud systems have been heavily dependent on hardware or network protocol-level technologies, based on a set of pre-defined defensive rules generated by domain expert engineers. However, little attention has been paid to machine learning-based software-level approaches. In [5], generating a set of significant rules on the basis of an inductive rule mining algorithm for cellular phones achieved detection accuracy between 80% and 90% according to various alarm mechanisms. In [6], three machine learning-based algorithms were employed, i.e., C4.5, Naïve Bayesian classifier, and support vector machine (SVM) to detect abnormal traffics in a VoIP system operated in China. In [7], Latent Dirichlet Allocation (LDA) was utilized for user profiling to distinguish fraudulent from normal behaviors.

Although all the above studies achieved favorable detection performances, there are some important issues that need to be addressed. First, binary classification scheme has been commonly adopted for fraud detection; its performance can be guaranteed if and only if sufficient examples are provided for both normal and fraud calls. In general, however, not only do normal calls far outnumber fraud calls, but also the characteristics of the fraud calls are too diversified to be generalized as an identical class. Therefore, it is more practically appropriate to build the fraud detection model based on novelty detection scheme where only normal calls are used to describe the target class. However, most of the conventional novelty detection algorithms, as well as binary classification algorithms, have their own parameters to be optimized. The existence of parameters may cause the performance instability because different parameters can be selected under different experimental settings.

In order to overcome the limitations listed above, we propose a one-class Naïve Bayesian classifier (One-NB) for toll fraud detection. Unlike binary classification, One-NB is based on the novelty detection scheme so that it is not mandatory to generalize the frauds as only one class during the model construction. Furthermore, since there is no parameters to be optimized, One-NB can result in a stable performance because it always gives the identical prediction outcomes for the same test data.

Manuscript received July 1, 2013.
Manuscript revised September 30, 2013.

[†]The author is with the IT Management Programme, Seoul National Univeristy of Science and Technology, 139–743, Republic of Korea.
a) E-mail: pskang@seoultech.ac.kr

DOI: 10.1587/transinf.E97.D.1353

2. One-Class Naïve Bayesian Classifier

In this paper, we set a more realistic scenario than that used in previous studies, in which fraud rates are uncertain but variable over different periods. If the fraud rate in one period is relatively high, an accurate binary classification model can be built. If the fraud rate is low in another period, on the other hand, it is sometime even impossible to train the binary classification algorithm. In order to secure robustness to uncertain fraud rates over different periods, novelty detection algorithms, also known as one-class classification algorithms, are more appropriate.

The Naïve Bayesian classifier was originally proposed for multi-class classification. When predicting the membership likelihood to the class C_i based on a set of explanatory variables $x_1, x_2, \ldots, x_d$, Bayes' rule [8] can be written as

$$P(C_i|x_1, x_2, \ldots, x_d) = \frac{P(x_1, x_2, \ldots, x_d|C_i) \times P(C_i)}{P(x_1, x_2, \ldots, x_d)}. \quad (1)$$

In the Naïve Bayesian classifier [9], it is assumed that the explanatory variables are statistically independent, and therefore Eq. (1) can be rewritten as

$$P(C_i|x_1, x_2, \ldots, x_d) = \frac{P(x_1|C_i) \times P(x_2|C_i) \times \cdots \times P(x_d|C_i) \times P(C_i)}{P(x_1, x_2, \ldots, x_d)}. \quad (2)$$

In toll fraud detection models, the fraud likelihood of a new call during a certain monitoring period should be computed by the Naïve Bayesian classifier:

$$\text{Likelihood} = P(\text{fraud}|x_1, x_2, \ldots, x_d, \text{period} = t + 1). \quad (3)$$

Assuming that the overall call patterns in two consecutive periods are consistent, we can rewrite Eq. (3) according to Bayes' rule:

$$\begin{aligned} &P(\text{fraud}|x_1, x_2, \ldots, x_d, \text{period} = t + 1) \quad (4)\\ &\propto P(\text{fraud}|x_1, x_2, \ldots, x_d, \text{period} = t)\\ &= \frac{P(x_1, \ldots, x_d|\text{fraud}, \text{period} = t) \times P(\text{fraud}|\text{period} = t)}{P(x_1, \ldots, x_d|\text{period} = t)}. \end{aligned}$$

Since $P(\text{fraud}|\text{period} = t)$ is identical during the test period, the One-NB computes the fraud likelihood in time period $t + 1$ as

$$\begin{aligned} \textit{Likelihood} &\propto \frac{P(x_1, \ldots, x_d|\text{fraud}, \text{period} = t)}{P(x_1, \ldots, x_d|\text{period} = t)}.\\ &= \prod_{k=1}^{d} \frac{P(x_k|\text{fraud}, \text{period} = t)}{P(x_k|\text{period} = t)}. \quad (5) \end{aligned}$$

3. Experiments

In this study, toll fraud detection models were built as follows. First, actual raw CDRs were collected by a major Korean VoIP service provider with the labeled fraudulent calls identified by expert engineers. Then, six novelty detection algorithms including One-NB were trained based on one week's data. In order to improve the detection performance, genetic algorithm was employed to determine significant variables. Finally, their toll fraud detection performances were compared using the data collected from the following week. Further, three well-known binary classification algorithms with GA variable selection were also tested to verify the detection performance of One-NB.

Table 1 The number of calls and fraud rate in each week.

Period	Total	Normal	Frauds	Fraud rate
Week 1	9,648	5,178	4,470	46.33%
Week 2	12,201	6,156	6,045	49.55%
Week 3	7,099	5,933	1,166	16.42%
Week 4	6,983	6,671	312	4.47%

Table 2 The explanatory variables used in this paper.

Variable Name	Class	Description
CALLER_ID	Nominal	Caller's phone number
AREA_CODE	Nominal	Receiver's country code
COMM_CODE	Nominal	Service product code
RATE_GROUP	Nominal	Charge rate class
DURATION	Numeric	Call time + waiting time
CHARGE	Numeric	Total amount of money being charged
BILL_YN	Binary	The result of billing request
CALLTYPE	Nominal	Call type classified by the server
CHARGECLASS	Nominal	Type of charge class
CALLINGLOCNUM	Nominal	Local number of the receiver
PREFIX	Nominal	Receiver's regional and country code
SSW_DURATION	Numeric	Actual call time
DISCONNREASON	Nominal	Code for disconnection
FAILCODE	Nominal	Code for call failure

3.1 Data Preparation & Variable Selection

A group of customers who had been consistently monitored by the VoIP service provider because of their historically frequent toll fraud attempts were selected in this experiment. Table 1 summarizes the number of total/normal/fraud calls made by the customers in June 2012. During the first two weeks, toll frauds were frequently attempted so that the fraud rate rose to as high as 50% of the total calls, whereas it decreased to less than 5% in the final week. Under this condition, more reliable experimental results can be obtained because we could compare the performance of One-NB with other novelty detectors and binary classifiers over different fraud rates.

The explanatory variables used in the experiment are shown in Table 2. Among 25 attributes in the raw CDRs, only 14 variables are taken into consideration after eliminating irrelevant or redundant variables. Since most of the remaining variables are nominal with a large number of cases, each case in a variable is transformed to a numeric value using Eq. (5) when computing the fraud likelihood. In order to improve the detection performance and model efficiency, genetic algorithm (GA) is employed to select significant variables. GA finds a set of pseudo-optimal input variables based on an evolutionary search method as fol-

lows. An initial population consisting of sufficient chromosomes is created, where each gene in a chromosome indicates whether the corresponding input variable is activated during the training. Then, novelty detectors or binary classifiers are trained using the activated variables in each chromosome. Chromosomes with higher detection ability survive and generate a new population through crossover and mutation. This procedure is repeated until an optimal set of variables can be obtained. In the experiment, the number of populations in each generation is set to 50, whereas the total number of generations is set to 100. The crossover rate and mutation rate are set to 0.5, and 0.05, respectively. Finally, the fitness function is set to the F1 measure as explained in the following section.

3.2 Benchmark Methods & Performance Measure

In order to verify the performance of One-NB, five well-known novelty detection algorithms and three binary classification algorithms are employed as benchmark methods. As benchmark novelty detectors, Gaussian density estimator (Gauss) [10], a mixture of Gaussians (MoG) [11], K-Means clustering (KMC) [12], The Parzen window density estimator (Parzen) [13], and k-nearest neighbor (k-NN) are used. Gauss assumes that the normal data are generated from an underlying Gaussian distribution. When the training CDRs are provided, Gauss estimates their mean and covariance, and the likelihood of a new call belonging to the normal CDRs is computed based on the estimated parameters: the higher the score, the lower fraudulence level of the new call. Since Gauss requires a strong assumption of unimodality, it is often violated in practice. By relaxing the unimodality assumption of Gauss, MoG allows the training data to have more than one modals with different parameters, each of which still follows the Gaussian distribution. KMC even relaxes the Gaussian assumption on each cluster in the MoG. Thus, KMC finds the K clusters that maximize the homogeneity among the instances in the same cluster and the heterogeneity among different clusters. Parzen is an extreme version of MoG in that the number of clusters equals the number of training instances. k-NN computes the novelty score of a new call based on the similarity between the test call and its k nearest neighbors [14]. The premise behind the k-NN is that if an instance is an outlier, it should be located far from other normal instances so that the distance to its nearest neighbors is larger than that of other normal instances. Among a number of variations in k-NN, the average distance is adopted as a novelty score in this study.

As benchmark binary classifiers, logistic regression (LR) [13], artificial neural networks (ANN) [9], and support vector machine (SVM) [15] are selected. LR is a *de facto* algorithm for binary classification. It is a simple linear classifier that is trained by maximizing the log-odds ratios of the two classes. ANN is one of the most widely adopted non-linear classification algorithm. It mimics the human brain's information processing by constructing a structure of networks that connect input variable and outcomes where a number of hidden layers or nodes are placed between them. SVM is a state-of-the art classifier that is based on the structured risk minimization SRM) principle and is able to provide the global optimum under a given parameter setting.

Table 3 Parameter search space for each novelty detection algorithm.

Algorithm	Parameter	Candidates
Gauss	None	-
MoG	Number of Modals	[2, 3, 4, ⋯, 10, 15, 20, 25, 30]
KMC	Number of clusters	[2, 3, 4, ⋯, 10, 15, 20, 25, 30]
Parzen	Kernel width	[0.01, 0.05, 0.1, 0.5, 1, 2, 3, 4, 5, 7, 10]
k-NN	Number of neighbors	[2, 3, 4, ⋯, 10, 15, 20, 25, 30]
One-NB	None	-
LR	None	-
ANN	Number of hidden nods	[1, 2, 3, ⋯, 10, 15, 20, 25, 30]
SVM	Kernel type	RBF-kernel
	Kernel width	$[2^{-5}, 2^{-4}, \cdots, 2^{4}, 2^{5}]$
	Cost	[0.1, 1, 2, ⋯, 10, 20, 30, 50, 100]

Since most of the algorithms have their own model parameters to be optimized as shown in Table 3, we used 10-fold cross validation with the training data set to select the best parameters and used them for model training and evaluation. Since the training data sets are randomly divided during 10-fold cross validation, different parameters can be selected as a result of different sampling. In order to investigate the performance stability, this procedure, i.e., parameter selection based on 10-fold cross validation and test the detector with the selected parameters, is repeated 30 times and the average and standard deviation of the detection performance are recorded.

The F1-measure is adopted as a performance measure. Four outcomes can result from detection models under a certain threshold: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). Recall is calculated as the correctly detected frauds over the total number of actual frauds and computed as TP/(TP+FN), whereas precision is calculated as the correctly detected frauds over the total number of calls identified as fraud by the model and computed as TP/(TP+FP). The F1-measure is the harmonic mean of recall and precision:

$$\text{F1-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}. \tag{6}$$

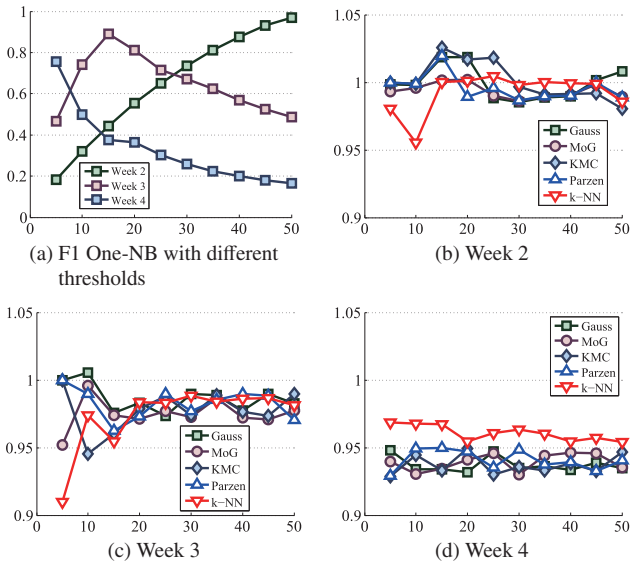
Since the F1-measure is dependent on the threshold setting, we vary the threshold from the top 5% novelty scores to the top 50% in increments of 5% in the case of novelty detection algorithms.

3.3 Experimental Results

Figure 1 (a) shows the F1-measure of One-NB for the three test periods for various threshold settings. Since all the novelty detectors present similar trends with One-NB, the relative F1 performance of the other five novelty detectors in each period is presented in Fig. 1 (b), (c), and (d). It was observed that the peak of the F1-measure was reported at the threshold that is closest to the actual toll fraud rate in the test period; the F1-measure was highest when the top 50%

Table 4 The performance of each novelty detector in each week (Values preceding ± is the average of 30 repetitions whereas values following ± is the standard deviation of each measure.).

Period	Measure	Variables	Gauss	MoG	KMC	Parzen	k-NN	One-NB
Week 2	Recall	All	0.9492	0.9371±0.0241	0.9117±0.0362	0.9105±0.0209	0.9317±0.0177	0.9515
		GA	**0.9811**	0.9613±0.0214	0.9579±0.0301	0.9613±0.0219	0.9578±0.0326	0.9734
	Precision	All	0.9405	0.9007±0.0514	0.9104±0.0415	0.9130±0.0332	0.9217±0.0287	0.9428
		GA	**0.9721**	0.9525±0.0275	0.9407±0.0317	0.9532±0.0141	0.9547±0.0187	0.9644
	F1	All	0.9448	0.9128±0.0351	0.9108±0.0357	0.9112±0.0217	0.9306±0.0207	0.9471
		GA	**0.9766**	0.9586±0.0256	0.9501±0.0179	0.9580±0.0206	0.9550±0.0281	0.9689
Week 3	Recall	All	0.6364	0.6185±0.0405	0.6108±0.0540	0.6558±0.0207	0.6627±0.0141	0.7168
		GA	0.8302	0.8210±0.0314	0.8028±0.0485	0.8174±0.0258	0.8078±0.0218	**0.8508**
	Precision	All	0.6967	0.6784±0.0442	0.6526±0.0335	0.7220±0.0208	0.7189±0.0258	0.7407
		GA	0.9089	0.9185±0.0350	0.9058±0.0288	0.9057±0.0278	0.8979±0.0117	**0.9315**
	F1	All	0.6652	0.6534±0.0388	0.6310±0.0447	0.6980±0.0236	0.6933±0.0186	0.7286
		GA	0.8678	0.8660±0.0410	0.8540±0.0207	0.8558±0.0267	0.8490±0.0174	**0.8893**
Week 4	Recall	All	0.6276	0.6014±0.0474	0.6108±0.0524	0.5453±0.0117	0.6348±0.0165	0.6469
		GA	0.7596	0.7530±0.0354	0.7427±0.0337	0.7508±0.0259	0.7818±0.0127	**0.8013**
	Precision	All	0.5504	0.4986±0.0554	0.5158±0.0544	0.4785±0.0216	0.5673±0.0225	0.5875
		GA	0.6791	0.6671±0.0374	0.6604±0.0405	0.6698±0.0258	0.6907±0.0155	**0.7163**
	F1	All	0.5865	0.5450±0.0521	0.5449±0.0538	0.5053±0.0198	0.5993±0.0208	0.6158
		GA	0.7171	0.7108±0.0366	0.7025±0.0385	0.7028±0.0227	0.7327±0.0138	**0.7564**



(a) F1 One-NB with different thresholds (b) Week 2 (c) Week 3 (d) Week 4

Fig. 1 The F1-measures (y-axis) with the top N% (x-axis) novelty scores as thresholds for each novelty detection algorithm.

novelty score was set to the threshold in week 2, in which the actual fraud rate was 49.55%. Similarly, the F1-measure was highest with the thresholds of the top 15% and 5% novelty scores for week 2 and week 3 where the actual fraud rates were 16.42% and 4.47%, respectively. Another observation is that toll fraud detection problem becomes more difficult when the actual fraud rate decreases. When toll fraud attempts were prevalent, as in week 2, most of them could be identified correctly, and false alarms and misses rarely occurred. When toll frauds were infrequently attempted, on the other hand, which is a more realistic situation, the highest F1-measures reported were between 0.7 and 0.9, depending on the test period and novelty detector. It is worth noting that since billions of calls are made in a week using the equipment of the VoIP service company, only a 1% increase in the F1-measure can significantly help the business increase its revenue.

Table 5 The number of variables selected by GA.

Detector	Week 2	Week 3	Week 4	Average
Gauss	4	**3**	**4**	3.67
MoG	6	4	**4**	5.67
KMC	4	7	6	5.67
Parzen	6	**3**	6	5.00
k-NN	6	4	**4**	4.67
One-NB	**2**	4	**4**	**3.33**

Table 4 shows the recall, precision, and F1-measure for all the novelty detectors in each test period when the threshold was set to the top 50%, 15%, and 5% for weeks 2, 3 and 4, respectively. In the Variables column, All means that the novelty detectors are trained on the basis of all variables whereas GA mean that they are trained on the basis of the selected variables by GA. It is observed that the variables selected by GA improved the detection performance of all novelty detectors and the performance improvement is more noticeable when the fraud rates are low. As mentioned above, the F1-measure was close to 1 in week 2. Among the novelty detectors, Gauss achieved a higher F1-measure (0.9766) than the others. In weeks 3 and 4, on the other hand, One-NB resulted in the highest Recall, Precision, and F1-Measure on average. In addition, One-NB resulted in the same performance since it does not have any model parameter, but other novelty detectors, except Gauss, even significant variations across the repetitions. It is worth noting that the performance of One-NB became noteworthy as the actual fraud rate decreased. It is of practical use that, for the entire customer base, the actual fraud rate should be lower than that of week 4. Consequently, One-NB would be more effective than other novelty detectors when deployed to general customers.

Table 5 summarizes the number of variables selected by GA for each novelty detector. One-NB used only between two (week 2) and four (week 3 and 4) variables, which resulted in it using the fewest variables on average,

Table 6 The performance comparison of One-NB with binary classification algorithms.

Period	Measure	LR	ANN	SVM	One-NB
Week 2	Recall	0.9902	0.9714±0.0237	**0.9929**±0.0074	0.9734
	Precision	0.9795	0.9600±0.0281	**0.9809**±0.0068	0.9644
	F1	0.9842	0.9624±0.0255	**0.9868**±0.0069	0.9811
Week 3	Recall	0.8701	0.8526±0.0504	0.8740±0.0124	**0.8508**
	Precision	0.7551	0.7636±0.0456	0.7617±0.0148	**0.9315**
	F1	0.8045	0.8188±0.0488	0.8200±0.0136	**0.8893**
Week 4	Recall	0.6442	0.6731±0.1047	0.6795±0.0258	**0.8013**
	Precision	0.7028	0.7192±0.1148	0.7260±0.0301	**0.7163**
	F1	0.6712	0.6950±0.1085	0.7026±0.0285	**0.7564**

3.33, followed Gauss with 3.67. If the fraud detection performance is comparable, models with fewer variables are easier to manage, and thus the efficiency can be improved.

The performance comparison of One-NB with the benchmark binary classification algorithms are summarized in Table 6. As we expected, binary classification algorithms achieved higher detection performance than One-NB when there are sufficient fraud calls in the training data set (Week 2). However, when the fraud rate decreases, their performance also decreased so that F1 measure of those classifiers are lower than that of One-NB in Week 2 and Week 3. To make matters worse, ANN reports relatively high performance fluctuations as the fraud rate decreases. Since it usually suffers from the over-fitting problem, insufficient information on the minority class, i.e., fraud calls, might have led ANN to inappropriate classification boundary.

In summary, the proposed One-NB for toll fraud detection was found to be highly effective, especially when the fraud rate was low compared to both other novelty detection algorithms and binary classification algorithms. In addition, it can provide a more stable performance compared to benchmark algorithms because there is no parameter to be optimized.

4. Conclusion & Discussion

In this paper, the One-NB is proposed in order to detect toll frauds in a VoIP service. In One-NB, historical fraud occurrence was utilized to compute the fraud likelihood of the Bayes' rule, assuming that the input variables are statistically independent. The experimental results on a real data set confirmed that One-NB resulted in the highest but stable F1-measure when the fraud rates were relatively low, and these performances were achieved with the minimum set of input variables as compared to the benchmark methods.

Some further research directions are as follows. First, since the actual CDRs were collected only for a group of monitored customers, their fraud rate was much higher than the average fraud rate of the entire customers. Thus, One-NB should be tested for different customer groups to verify its general performance. Second, the test period was fixed to one week in the experiment; it would be more practical if the test period was adjusted dynamically to reflect the current toll fraud occurrence situation.

Acknowledgement

The work was supported by the research program funded by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science,and Technology (MEST) (No.2011-0021893).

References

[1] A. Keromytis, "A comprehensive survey of Voice over IP security research," IEEE Communications Survey & Tutorials, vol.14, no.2, pp.514–537, 2012.
[2] H. Abdelnur, R. State, and O. Festor, "Advanced fuzzing in the VoIP space," J. Computer Virology, vol.6, no.1, pp.57–64, 2010.
[3] A. Keromytis, "Voice-over-IP security: Research and practice," IEEE Security & Privacy, vol.8, no.2, pp.76–78, 2010.
[4] R. Dantu, S. Fahmy, H. Schulzrinne, and J. Cangussu, "Issues and challenges in securing VoIP," Computers & Security, vol.28, no.8, pp.743–753, 2009.
[5] T. Fawcett and F. Provost, "Combining data mining and machine learning for effective user profiling," Proc. Second International Conference on Knowledge Discovery and Data Mining (KDD'96), pp.8–13, Portland, OR, USA, 1996.
[6] Y. Liu, G. Li, W. Yu, J. Xu, and J. Yang, "Abnormal traffic detection in VoIP by using machine learning approaches," Proc. 2008 International Conference on Artificial Intelligence (ICAI'08), pp.295–300, Las Vegas, NV, USA, 2008.
[7] D. Olszewski, "A probabilistic approach to fraud detection in telecommunications," Knowledge-Based Systems, vol.26, pp.246–258, 2012.
[8] J. Stone, Bayes' Rule: A Tutorial Introduction to Bayesian Analysis, Sebtel Press, 2013.
[9] C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[10] V. Barnett and T. Lewis, Outliers in Statistical Data, Wiley & Sons, New York, NY, USA, 1994.
[11] G. McLachlan and D. Peel, Finite Mixture Models, Wiley & Sons, New York, NY, USA, 2000.
[12] D. Tax, One-Class Classification: Concept-Learning in the Absence of Counter-Examples, Ph.D. thesis, Delft University of Technology, June 2001.
[13] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, John Wiley & Sons, New York, NY, USA, 2001.
[14] P. Kang and S. Cho, "A hybrid novelty score and its use in keystroke dynamics-based user authentication," Pattern Recognit., vol.42, no.11, pp.3115–3127, 2009.
[15] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, vol.2, no.2, pp.121–167, 1998.