Non Technical Loses Detection
Experts Labels vs. Inspection Labels in the Learning Stage

Fernanda Rodríguez, Federico Lecumberry and Alicia Fernández
Instituto de Ingeniría Eléctrica, Facultad de Ingeniería, Universidad de la República,
J. Herrera y Reissig 565, 11300, Montevideo, Uruguay

Keywords: Electricity Fraud, Support Vector Machine, Optimum Path Forest, Unbalance Class Problem, Combining Classifier, UTE.

Abstract: Non-technical losses detection is a complex task, with high economic impact. The diversity and big number of consumption records, makes it very important to find an efficient automatic method for detection the largest number of frauds with the least amount of experts' hours involved in preprocessing and inspections. This article analyzes the performance of a strategy based on learning from expert labeling: suspect/no-suspect, with one using inspection labels: fraud/no-fraud. Results show that the proposed framework, suitable for imbalance problems, improves performance in terms of the $F_{measure}$ with inspection labels, avoiding hours of experts labeling.

1 INTRODUCTION

Improving non-technical loss detection is a huge challenge for electric companies. In Uruguay the national electric power utility (henceforth UTE) faces the problem by manually monitoring a group of customers. A group of experts inspect at the monthly consumption curve of each customer and indicates those with some kind of suspicious behavior. This set of customers, initially classified as suspects are then analyzed taking into account other factors (such as fraud history, electrical energy meter type, etc.). Finally a subset of customers is selected to be inspected by an UTE's employee, who confirms (or not) the irregularity. The procedure is illustrated in Figure 1. The procedure described before, has major drawbacks, mainly, the number of customers that can be manually controlled is small compared with the total amount of customer (around 500.000 only in Montevideo).

Several studies with a Pattern Recognition approach have addressed the detection of non-technical losses, both supervised or unsupervised. Leon et al. review the main research works found in the area between 1990 and 2008 (Leon et al., 2011). Here we present a brief review that builds on this work and wide it with new contributions published between 2008 and 2013. Several of these approaches consider unsupervised classification using different techniques such as fuzzy clustering (dos Angelos et al., 2011), neural networks (Markoc et al., 2011; Sforna, 2000), among others. Monedero et al. use regression based on the correlation between time and monthly consumption, looking for significant drops in consumption (Monedero et al., 2010). Then they make a second stage where suspicious customers are eliminated if the consumption of these depend on the economy of the moment or the year's season. Only major customers were inspected and 38% were detected as fraudulent. Similar results (40%) were obtained in (Filho et al., 2004) using a tree classifier and customers who had been inspected in the past year. In (Depuru et al., 2011) and (Yap et al., 2007) SVM is used. In the latter, Modified Genetic Algorithm is employed to find the best parameters of SVM. In (Yap et al., 2012), is compared the methods Back-Propagation Neural Network (BPNN), Online-sequential Extreme Learning Machine (OS-ELM) and SVM. Biscarri et al. (Biscarri et al., 2008) seek for outliers, Leon et al. (Leon et al., 2011) use Generalized Rule Induction and Di Martino et al. (Di Martino et al., 2012) combine CS-SVM classifiers, One class SVM, and C4.5 OPF using various features derived from the consumption. Different kinds of features are used among this works, for examples, consumption (Biscarri et al., 2008; Yap et al., 2007), contracted power and consumed ratio (Galvn et al., 1998), Wavelet transformation of the monthly con-

Rodríguez F., Lecumberry F. and Fernández A..
Non Technical Losses Detection - Experts Labels vs. Inspection Labels in the Learning Stage.
DOI: 10.5220/0004823506240628
In *Proceedings of the 3rd International Conference on Pattern Recognition Applications and Methods* (ICPRAM-2014), pages 624-628
ISBN: 978-989-758-018-5

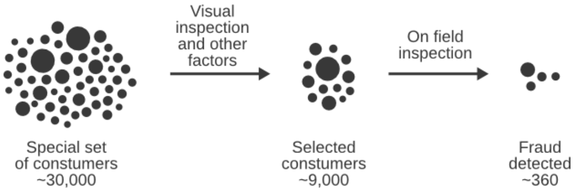

Figure 1: Manual fraud detection scheme.

sumption (Jiang et al., 2002), amount of inspections made to each client in one period and average power of the area where the customer resides (dos Angelos et al., 2011), among others.

On the other hand, Romero proposes (Romero, 2012) a method to estimate and reduce non-technical losses, such as advanced metering infrastructure, fraud deterrence prepayment systems, system remote connection and disconnection, etc. Lo et al. based on real-time measurements, design (Lo et al., 2012) an algorithm for distributed state estimation in order to detect irregularities in consumption.

To improve the efficiency of fraud detection and resource utilization, in (Di Martino et al., 2013) was implemented a tool that automatically detects suspicious behavior analyzing customers historical consumption curve. This approach has the drawback of requiring a base previously tagged by the experts, in order to use it in the training stage.

In this paper we set out to analyze the behavior of the proposed framework to fraud classification and compare it by using labels based on the inspection results instead of labels defined by experts. This new approach does not require that the company personnel conduct a manual study of the customers' consumption curve, since it use labels resulting from inspections in the past. We investigate performance improvement originated by training with individual algorithms and their combinations with labels of fraud and no fraud (based on inspections) and the importance of choosing the appropriate performance measure to solve the problem.

The paper is organized as follows. Section 2 describes general aspects of the class imbalance problems, Section 3 describes the strategies to be compare, Section 4 presents the obtained results and, finally Section 5 concludes the work.

2 THE CLASS IMBALANCE PROBLEM AND THE CHOICE OF PERFORMANCE MEASURE

When working on fraud detection problems, we can not assume that the number of people who commit fraud are near the same than those who do not, usually they are a minority class. This situation is known as class imbalance problem, and it is particularly important in real world applications where it is costly to misclassify examples from the minority class. In this cases, standard classifiers tend to be overwhelmed by the majority class and ignore the minority class, hence obtaining suboptimal classification performance. In order to confront this type of problem, different strategies can be used on different levels: (i) changing class distribution by resampling; (ii) manipulating classifiers; (iii) and on the ensemble of them, as proposed in (Di Martino et al., 2013).

Table 1: Confusion matrix.

	Labeled as	
	Positive	Negative
Positive	TP (True Positive)	FN (False Negative)
Negative	FP (False Positive)	TN (True Negative)

Another problem which arises when working with imbalanced classes is that the most widely used metrics for measuring the performance of learning systems, such as *Accuracy* and *ErrorRate*, are not appropriate because they do not take into account misclassification costs, since they are strongly biased to favor the majority class (Garcia et al., 2012). Then others measures have to be considered:

- *Recall* is the percentage of correctly classified positive instances, in this case, the fraud samples.
$$Recall = \frac{TP}{TP+FN}$$
- *Precision* is defined as the proportion of labeled as positive instances that are actually positive.
$$Precision = \frac{TP}{TP+FP}$$
Where TP, FN and FP are defined in Table 1.
- The combination of this two measurements, the $F_{measure}$, represents the geometric mean between them, weighted by the parameter β,
$$F_{measure} = \frac{(1+\beta^2)Recall \times Precision}{\beta^2 Recall + Precision} \quad (1)$$

Depending on the value of β we can prioritize *Recall* or *Precision*. For example, if we have few resources to perform inspections, it can be useful to prioritize *Precision*, so the set of samples labeled as positive has high density of true positive.

When working with inspection labels the imbalance problem is worst, in terms of unbalance, than dealing with experts labels. In the experts labels method, the ratio of suspect to no suspect is near 10%, while in the one based on inspection labels, the ratio is near 0.4%.

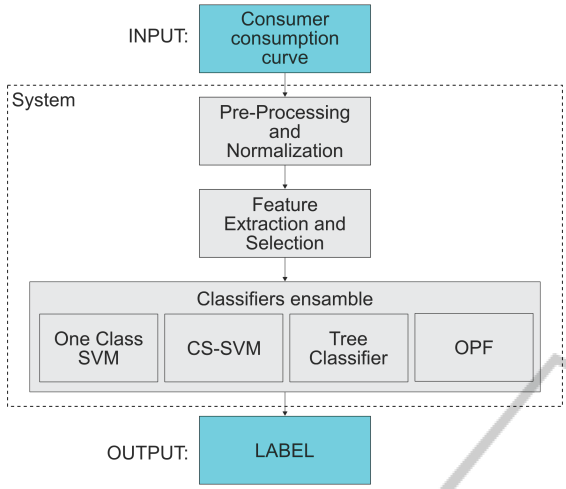

Figure 2: Block Diagram of the automatic fraud detection system.

3 FRAMEWORK

The system presented consists basically on three modules: Pre-Processing and Normalization, Feature Extraction and Selection, and Classification. Figure 2 shows the system configuration. The system input corresponds to the last three years of the monthly consumption curve of each costumer.

The first module, Pre-Processing and Normalization, modifies the input data so that they all have normalized mean and implements some filters to avoid peaks from billing errors. A feature set was proposed taking into account UTE's technician expertize in fraud detection by manual inspection and recent papers on non technical loss detection (Alcetegaray and Kosut, 2008), (Muniz et al., 2009), (Nagi and Mohamad, 2010). Di Martino et al. use a list of the features extracted from the monthly consumption records (Di Martino et al., 2013). In this work, we use the framework illustrated in Figure 2 and a subset of the same set of features used in (Di Martino et al., 2013) but doing a selection of them taking into account the label type (based on inspection or expertise's criterion).

It is well known that finding a small set of relevant features can improve the final classification performance; this is why we implemented a feature selection stage. We used two types of evaluation methods: filter and wrapper. Filters methods looks for subsets of features with low correlation between them and high correlation with the labels, while wrapper methods evaluate the performance of a given classifier for the given subset of features. In the wrapper methods, we used as performance measure the $F_{measure}$, also, the evaluations were performed using 10 fold cross validation over the training set.

As searching method, we used *Bestfirt*, for which we found in this application a good balance between performance and computational costs.

Different feature subsets were selected from the original set proposed in (Di Martino et al., 2013) for both approaches. For example, for the experts' labels approach, the features include:

- Consumption ratio for the twelve months and the average consumption.
- Difference between fourth Wavelet coefficient from the last and previous years.
- Euclidean distance of each customer to the *mean customer*, where the *mean customer* is calculated by taking the mean for each month between all the customers.
- Module of the first Fourier coefficient of the total consumption.

While for inspection label approach, the features include:

- Difference between the first two Fourier coefficients from the last and previous years.
- Variance of the consumption curve.

Some features are selected in both approaches, such as:

- Consumption ratio for the last three and six months and the average consumption.
- Difference between fifth Wavelet coefficient from the last and previous years.
- Slope of the straight line that fits the consumption curve.

The performance analysis considers, SVM algorithm, one-class classifier (O-SVM) and cost-sensitive learning (C-SVM), Optimum Path Forest (OPF) (Ramos et al., 2010), a decision tree proposed by Roos Quinlan, C4.5 and Iterative Combination proposed in (Di Martino et al., 2012). The latter method performs an optimal combination of the before mentioned classifiers. The choice of combination's weights is done exhaustively in order to maximize the $F_{measure}$.

4 EXPERIMENTS AND RESULTS

In this work we used a data set of 456 industrial profiles obtained from the UTE's database. Each profile is represented by the customers monthly consumption in the last 36 months and has two labels, one dictated manually by technicians previous the inspection and

Table 2: Fraud detection with experts label training.

Description	*Recall* (%)	*Precision* (%)	$F_{measure}$ (%)[$\beta = 1$]
OPF	39	27	32
Tree (C4.5)	38	23	29
O-SVM	51	22	30
CS-SVM	35	20	26
Iterative Combination	77	22	35

Table 3: Fraud detection with inspection label training.

Description	*Recall* (%)	*Precision* (%)	$F_{measure}$ (%)[$\beta = 1$]
OPF	36	34	35
Tree (C4.5)	33	37	35
O-SVM	71	31	44
CS-SVM	74	33	46
Iterative Combination	77	33	46

another based on the inspection results. Training was done considering both labels separately and performance evaluation was done given the inspection labels, using a 10-fold cross validation scheme.

Tables 2 and 3 shows the results obtained when experts and inspection labels are used to train the different classifiers respectively. The Iterative Combination technique with expert label training obtains the best result for fraud detection clearly overpassing the other methods, however the number false positive (FP) is relatively high, since

$$\frac{FP}{TP} = \frac{1}{Precision} - 1 \approx 4.$$

On the other hand, if we use the inspection labels the Iterative Combination also obtains the best results for fraud detection, but reducing in a half the number of FP ($\frac{FP}{TP} \approx 2$).

If we compare both approaches, we see that learning from the inspection labels could get better results (in the $F_{measure}$ sense) than learning from the labels set by experts. The former has the additional advantage of not requiring that the experts made the manual labeled of the training base.

The results for the method performed manually by experts, i.e. validating the expert labels with inspection labels, are $Recall = 38\%$, $Precision = 51\%$ and $F_{measure} = 44\%$.

Comparing the F_{masure} obtained manually by the experts (44%) and automatically by the Iterative Combination (46%) both are similar. However, the former consider other features as the history's fraud detection, contracted power, number of estimated readings, etc. and not only the monthly consumption, as the automatic one.

5 CONCLUSIONS AND FUTURE WORK

In this work we compare the performance of a strategy based on learning from expert labeling: suspect/no-suspect, with one using inspection labels: fraud/no-fraud. In the $F_{measure}$ sense with all the tested classifiers the classification with inspection label obtains better results than using experts labels. Among them the Iterative Combination obtains the best result and also better than the manual method.

In future work we propose to include new categorical attributes as the history's fraud detection, contracted power, number of estimated readings, etc. We also want to explore a semi-supervised approach that allows to learn from data with and without previous inspection labels.

ACKNOWLEDGEMENTS

This work was supported by the program Sector Productivo CSIC UTE. Authors would like to thank UTE, especially Juan Pablo Kosut and Fernando Santomauro, for providing datasets and share fraud detection expertise.

REFERENCES

Alcetegaray, D. and Kosut, J. (2008). One class svm para la detección de fraudes en el uso de energía eléctrica. *Trabajo Final Curso de Reconocimiento de Patrones, Dictado por el IIE- Facultad de Ingeniería- UdelaR*.

Biscarri, F., Monedero, I., Leon, C., Guerrero, J. I., Biscarri, J., and Millan, R. (2008). *A data mining method based on the variability of the customer consumption - A special application on electric utility companies*, volume AIDSS, pages 370–374. Inst. for Syst. and Technol. of Inf. Control and Commun.

Depuru, S. S. S. R., Wang, L., and Devabhaktuni, V. (2011). Support vector machine based data classification for detection of electricity theft.

Di Martino, J., Decia, F., Molinelli, J., and Fernández, A. (2012). Improving electric fraud detection using class imbalance strategies. *1st International Conference In Pattern Recognition Aplications and Methods*, pages 135–141, volumen 2.

Di Martino, M., Decia, F., Molinelli, J., and Fernndez, A. (2013). A novel framework for nontechnical losses detection in electricity companies. In Latorre Carmona, P., Snchez, J. S., and Fred, A. L., editors, *Pattern Recognition - Applications and Methods*, volume

204 of *Advances in Intelligent Systems and Computing*, pages 109–120. Springer Berlin Heidelberg.

dos Angelos, E., Saavedra, O., Corts, O., and De Souza, A. (2011). Detection and identification of abnormalities in customer consumptions in power distribution systems.

Filho, J. R., Gontijo, E. M., Delaiba, A. C., Mazina, E., Cabral, J. E., and Pinto, J. O. P. (2004). Fraud identification in electricity company customers using decision tree.

Galvn, J., Elices, E., Noz, A. M., Czernichow, T., and Sanz-Bobi, M. (1998). System for detection of abnormalities and fraud in customer consumption.

Garcia, V., Sanchez, J., and Mollineda, R. (2012). On the suitability if numerical performance evaluation measures for class imbalance problems. *1st International Conference In Pattern Recognition Aplications and Methods*, pages 310–313, volumen 2.

Jiang, R., Tagaris, H., and Laschusz, A. (2002). Wavelets based feature extraction and multiple classifiers for electricity fraud detection.

Leon, C., Biscarri, F. X. E. L., Monedero, I. X. F. I., Guerrero, J. I., Biscarri, J. X. F. S., and Millan, R. X. E. O. (2011). Variability and trend-based generalized rule induction model to ntl detection in power companies.

Lo, Y.-L., Huang, S.-C., and Lu, C.-N. (2012). Non-technical loss detection using smart distribution network measurement data. In *Innovative Smart Grid Technologies - Asia (ISGT Asia), 2012 IEEE*, pages 1–5.

Markoc, Z., Hlupic, N., and Basch, D. (2011). Detection of suspicious patterns of energy consumption using neural network trained by generated samples.

Monedero, I., Biscarri, F., Len, C., Guerrero, J., Biscarri, J., and Milln, R. (2010). Using regression analysis to identify patterns of non-technical losses on power utilities. In Setchi, R., Jordanov, I., Howlett, R., and Jain, L., editors, *Knowledge-Based and Intelligent Information and Engineering Systems*, volume 6276 of *Lecture Notes in Computer Science*, pages 410–419. Springer Berlin Heidelberg.

Muniz, C., Vellasco, M., Tanscheit, R., and Figueiredo, K. (2009). Ifsa-eusflat 2009 a neuro-fuzzy system for fraud detection in electricity distribution.

Nagi, J. and Mohamad, M. (2010). Nontechnical loss detection for metered customers in power utility using support vector machines. *IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 25, NO. 2*.

Ramos, C., de Sousa, A. N., Papa, J., and Falcao, A. (2010). A new approach for nontechnical losses detection based on optimum-path forest. *IEEE TRANSACTIONS ON POWER SYSTEMS*.

Romero, J. (2012). Improving the efficiency of power distribution system through technical and non-technical losses reduction.

Sforna, M. (2000). Data mining in power company customer database.

Yap, K. S., Hussien, Z., and Mohamad, A. (2007). Abnormalities and fraud electric meter detection using hybrid support vector machine and genetic algorithm.

Yap, K. S., Tiong, S. K., Nagi, J., Koh, J. S. P., and Nagi, F. (2012). Comparison of supervised learning techniques for non-technical loss detection in power utility.