

A Quality Control Method for Fraud Detection on Utility Customers without an Active Contract

Bernat Coma-Puig
Universitat Politècnica de Catalunya
Barcelona, Spain
bcoma@cs.upc.edu

Josep Carmona
Universitat Politècnica de Catalunya
Barcelona, Spain
jcarmona@cs.upc.edu

ABSTRACT

Fraud detection in energy consumption has proven to be a difficult problem for current techniques. In general, the approaches used in this area are restricted to compute a fraud score for each client based on its behaviour. The problem gets much more complicated on customers with no contract, since the company does not have enough information from them to compute an accurate profile. On this paper, we introduce a semi-autonomous method that combines different machine learning algorithms and human knowledge to alleviate the lack of information to build a framework that detects fraud nimbly.

CCS CONCEPTS

•General and reference → Reliability; •Computing methodologies → Classification and regression trees;

KEYWORDS

Fraud Detection, Utility Company, Quality Control Method

ACM Reference format:

Bernat Coma-Puig and Josep Carmona. 2018. A Quality Control Method for Fraud Detection on Utility Customers without an Active Contract. In *Proceedings of SAC 2018: Symposium on Applied Computing*, Pau, France, April 9–13, 2018 (SAC 2018), 4 pages. DOI: 10.1145/3167132.3167384

1 INTRODUCTION

Gas Natural-Fenosa (an international utility company from Spain), altogether with the Universitat Politècnica de Catalunya, have been working on a system that embraces data mining techniques to automatically detect fraudulent customers. The system has achieved good results in spite of its complexity (see [2]).

The customers with no contract require a special attention in our system¹. In these cases, the electric installation is disabled, and therefore there are no consumption records (the basis of our framework) but the customer can still commit fraud because much of the infrastructure (e.g. electric wiring) is not removed. To face this

¹Even though we may still refer them as customer without contract, they actually are old customers that cancelled their contract as well as disabled supply points that have never had a customer with contract

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SAC 2018, Pau, France

© 2018 Copyright held by the owner/author(s). 978-1-4503-5191-1/18/04...\$15.00
DOI: 10.1145/3167132.3167384

problem, we included new features to compensate the lack of information (e.g. socio-economical information of the area considered), which only help slightly to solve this problem.

In this paper we propose a method that combines the classification model (i.e. an *XGBoost* model, [1]), an explanatory method (i.e. the *Lime* algorithm [4]), as well as the human knowledge to improve the classification task. Our experiments over a real data based on field verifications (a term called *campaigns*) done in the past prove that the addition of *Lime* and the human factor help to avoid justifiable misclassification by the poor quality of the available data.

2 RELATED WORK

Detection of fraud or irregularities (the *non-technical losses* or NTL) has received considerable attention in the last decades; see for instance [3, 5] for the particular case of electricity. In this section we focus on recent approaches. Differences with our work were already reported in [2].

The *Lime* model explanation method (see Figure 1) introduces an approach to faithfully explain the predictions of any model with a local interpretable model. This method analyses the characteristics of a prediction in comparison to similar elements and shows to the user an explanation of that prediction. In our case, a weight to each feature is assigned, indicating its contribution to the final score (e.g. if $feature_n$ has a weight of 0.15, that would indicate that the value of that feature contributes +0.15 in the final score).

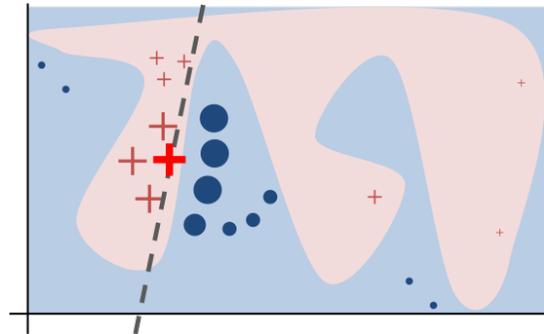


Figure 1: Image from [4]. *Lime* generates a local linear model that is not good globally (i.e. coloured zones), but is useful to explain the classification of an element (e.g. big red cross).

3 THE SYSTEM

The initial approach of our system was presented at [2]. In this section we provide an up-to-date explanation of the system, focusing on the clients with no active contract.

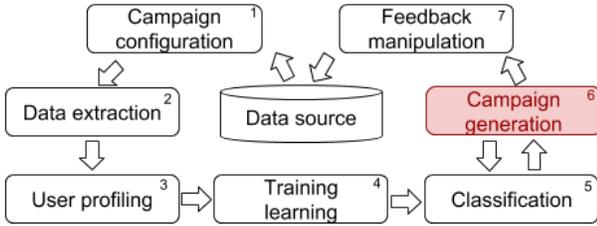


Figure 2: System architecture. Campaigns are generated iteratively. The campaign generation, highlighted in red, is done manually, but can be (partially) automatized with the double-check proposed in this paper.

3.1 dataset

The company has more than 370000 customers with no contract, where almost a hundred thousand of them are supplies that have never had a contract (e.g. new building) and around 270000 are customers that cancelled their contract. The information available from them are:

- Consumption data: For those customers that once had a contract contains information of the energy consumed by the customer. It includes meter readings (depending on the meter it can be more or less trustful), billing extractions as well as other consumption information.
- Static profile data: Within this concept we include information eminently static related to the contract of the customer with the company (e.g. the tariff), information of the customer (e.g. its address) and characteristics of the equipment (e.g. whether the meter is inside or outside the house).
- Historical operations: The company saves the results of all types of campaigns and other type of operations. This information is both used to get labelled data to train models and to generate new features (e.g. if the customer has committed fraud during the period with no contract).
- External information: In addition to the data from the company, we have included socio-economic information that complements the information available from the company (e.g. the unemployment of the city, ratio of tourism, among others).

3.2 The classification problem

Our method, uses the data explained above to create a supervised model that scores the customers (see Figure 2):

- (1) We extract all the data to create the profiling variables.
- (2) A vector of values that defines the customer profile in the present, as well as the feedback profiles (i.e. the profiles of the customer when it was visited) are done.
- (3) We create a supervised model with an optimized configuration and assign a calibrated fraud score for every customer in the target area for a new campaign.
- (4) We create a campaign of a desired size N by selecting the N customers with highest fraud score that comply the requirements established by the company (e.g. a minimum number of customers for each province).
- (5) When the campaign results return from the field, the feedback is added to the system automatically

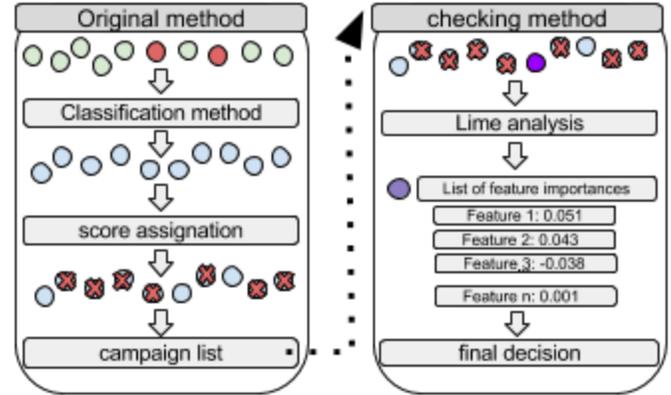


Figure 3: Global vision of the system: After the score assignation, we use Lime to explain locally the score assigned by the global model to double-check the inclusion of the customer in the campaign and understand the biases and defects of our method.

The results of a campaign can be summarized as: i) fraud, ii) incident (i.e. cases where the company has lost energy, but are not considered fraud), iii) unknown (i.e. cases where the installation was not checked (e.g. the customer did not allow the checking)) and iv) correct (i.e. the meter was checked and there was no problem).

Both the fraud and the incident are categorized as positive-labelled cases (in both cases the company loses energy), and the correct cases as negative-labelled. Those cases with no result, in spite of its higher probability of fraudulence, remain unlabelled.

4 THE PROBLEM

The detection of fraud in customers with no contract has an important problem: the lack of information. The customers with no contract are neither billed nor heavily controlled, since the supply is disabled, and therefore there should not be any type of consumption. In spite of that, the customer that cancelled the contract, or other people that can access the wiring, can commit fraud by tampering the installation.

The fundamentals of our framework is to profile the behaviour of the customers to detect abnormal behaviours, focusing on its consumption data. The profile of a no-contract customer, with no consumption information, has little predictive value. Despite the inclusion of new features (e.g. we included a wide range of information related to historical campaigns, as well as several features related to the last period of contract for those customers that cancelled it), the picture of the behaviour may still be poor. This poor information may cause misclassification.

The method explained in this paper palliates this problem by combining the classification method, Lime and the knowledge of a specialist. This process, as seen in Figure 3, is:

- (1) First of all, the classification method would assign a score to the customers, using the framework shown in Figure 2.
- (2) After that, the Lime algorithm would be used to understand locally the score assigned to those customers that are suspicious by the classification method (i.e. customers with a high score). The coefficients of the features (as seen in Section 2) would indicate their importance in the final score.

- (3) All this information can be interpreted easily by a human specialist to detect customers with a score unjustifiably high, and drop them out from the definitive campaign list. This process can be done easily and fast, because Lime would indicate which are the features that define the customer and justify its score. This method allows us to improve the campaign generated (e.g. dropping non-suspicious customers with high score) and also it is useful to improve future campaigns, as one can detect biases and erroneous patterns that can be corrected.

5 THE TEST

The dataset used in our tests includes data from 2016 and 2017 is:

- Training data: 5278 fraudulent customers and 71676 non-fraudulent customers from 2016.
- Validation data: 1816 fraudulent customers and 19800 non-fraudulent customers from the first 3 months of 2017.
- Test data: 1200 fraudulent customers and 17213 non-fraudulent customers from April, May and June of 2017.

The method explained in Section 4 is executed as follows:

- (1) A first study using the training and the validation data will be used to tune the XGBoost algorithm.
- (2) Then, a model will be trained using both the training and the validation data. This model will be used to assign a score to the test customers. The customers with higher score will be analysed by our method.
- (3) Statistical information about our experiment will be explained.

The score analysis requires a human specialist. Due to the lack of space, we simplify this human process by different rules that could be seen as an approximation of the human decision. That is:

- (1) Pre-discard: The customer would not be included into the campaign if Lime considers suspicious (i.e. the feature has a positive weight) that the customer was visited in the last 3 months. This rule is introduced to the system to avoid over-control (i.e. to visit constantly a customer) and to optimize the campaign (i.e. a longer fraud period maximizes benefits).
- (2) Analysis: If the customer was not discarded, a Lime model will be created, and according to the information provided by the local model, a decision will be made:
 - (a) Include: If the top 3 positive features (i.e. the features that Lime indicated as fraudulent) are trustful (e.g. the customer has historically committed fraud).
 - (b) Discard: If the top 3 non-fraudulent features can be trusted as indicators of no fraud (e.g. the customer has never committed fraud).
 - (c) Doubtful case: If neither of the previous rules are met (or Lime provides both fraudulent and non-fraudulent trustful top features), a deeper analysis is required. The options available that will be used depending on the situation are:
 - Revision: A manual check of the exact value of a non-top Lime feature that can complement the information proportioned by Lime to do the final decision (e.g. in a customer that committed fraud a long time ago, we can check when was the last visit with no fraud).
 - Substitution: We can modify the value of a feature that Lime considers important but one would disagree (e.g.

the province where the customer lives). A new score and Lime analysis would be done to do the final decision.

- Remove: Similar to the substitution process, we can choose to remove the feature whose value is unusually high in the customer we are analysing. After the new Lime model and score, one can decide its inclusion in the final campaign.

5.1 The results

With the aim of explaining our method in this paper, we will simulate a campaign, selecting those customers from the test dataset with a higher score. We have kept the proportion between the amount of customers usually selected in a campaign between the 370000 customers and the amount of customers used in this test. The customers selected in this campaign have a score range from 0.60 to 0.45. For each case, we are going to proceed according to the method explained in Section 5.

The final results of the application of our method in terms of the simulated campaign can be seen in Table 3. Our method included in the final selection 25% of the fraudulent customers, and only 10% of non-fraudulent customers, improving the accuracy of the simulated campaign by 13%, as seen in the bold row in Table 3. If we focus on the customers dropped out, the % of non-fraudulent customers not included in the campaign is almost three times the % of fraudulent customers included in the campaign (i.e. we discard more non-fraudulent than fraudulent customers). Moreover, we can see that, in spite of a high percentage of fraudulent customers were discarded by the "3 months rule", one can understand that most of them are recidivist, and would be visited in the future. It is also worth explaining that the benefits in a real world scenario campaign would be increased, since the proportion of the fraud/no-fraud is more imbalanced than in our test.

As already explained in Section 4, this method also helps us to understand several aspects of our method that can be improved. Some of the aspects that Lime shed light are:

- Most of the top influential features (i.e. the features with a higher absolute value) were those that are considered fraud proof, and there were only a few variables indicative of non-fraud. This fact makes us think that we may improve our system if we include better non-fraudulent features.
- Visits in the building: In the no-contract scenario considered in this paper, the visits to the customer are key factors to detect fraud in our model. Lime constantly used both the visits and the density of fraud in the buildings as explainers of a high score. One can understand that this information can be combined to improve the system; we can create features that refers to the visits in a building, avoiding the visit of the same building in consecutive months, especially in those buildings that the meters are in the public area (e.g. in the building corridor).

6 CONCLUSIONS

In this paper, we presented an automatic semi-autonomous method that allow us to make our autonomous method to detect fraud more robust and solid. To achieve that purpose, it uses the Lime algorithm, a method that helps to interpret a machine learning method using local models. This information, in the hands of the appropriate specialist, may help to palliate biases that could misclassify

Score	Top-3 Fraudulent	Top-3 Non Fraudulent	Analysis	Inclusion	Fraud
0.502	(1) newerMeter>-1, W:0.0427 (2) #Fraud2>0, W:0.023 (3) buildFraud>=0, W:0.0222	(1) lastReading<=0, W:-0.0191 (2) #ImpVisit2<=0, W:-0.010 (3) #MeterComp<=0, W:-0.009	According to the top fraudulent features, this customer has a meter (1), has committed fraud during its non-contract period (2) and there is fraud in its building (3) (all three features a good proof of fraud). In spite of that, the first no fraudulent feature is a good non-fraudulent feature. We checked manually how many frauds has committed this customer (an example of the revision case explained in 4). It has committed 8 frauds during its no-contract period, so we decided to include it to the final list.	Yes	Yes

Table 1: An example of how we simulated the human behaviour. The information is extracted from the Lime algorithm. The structure (value of the feature, weight) indicates how Lime weights the importance of the feature, where a positive weight indicates a fraudulent factor and a negative weight indicates a no-fraud factor. This table is complemented with Table 2.

Feature	definition
buildFraud	It indicates the proportion of customers that has committed fraud in the building.
newerMeter	How old the customer’s newer meter is. The age equal to -1 means that all the customer’s meters were removed. Having a meter might be a fraudulent feature, since it would be easier to manipulate the installation.
LastReading	How many months has passed from its last reading. If the customer still has a meter installed, a technician or the smart meter can send the reading to the company. This might be understood as a non-fraudulent element.
#Fraud2	How many frauds of type 2 (an internal category of the company) has done this customers during its no contract period. This is absolutely understood as a fraudulent proof.
#ImpVisits2	How many impossible visits (i.e unknown results in campaigns, explained in Section 3) of type 2 (an internal category of the company) has the customer had. This can be considered as a fraudulent behaviour.
#MetersComp	The number of meters from the company the customer has. If the customer owns a meter it can be more difficult to control by the company, so the fraud would be easier to commit and more difficult to control.

Table 2: Summarize of the features from Table 1.

Original campaign	% customers included	% Fraud cases	% No fraud	% Precision
campaign without our method	100	72	28	72
With our method	% customers included	% Fraud included	% No fraud included	% Precision
Postponed/discarded	58	60	52	63
Included in the campaign	21	25	10	85
Discarded in the campaign	22	15	38	50

Table 3: Results on the application of our checking method in our simulated campaign. The % of fraudulent customers from the original selection included in the campaign is higher than the non-fraudulent customers, and therefore we increase the success of the campaign (from a precision of 72% to a 85%)

suspicious customers, as well as be used to understand better the defects of the system for its improvement.

In addition to the possible improvement of the system, this quality control method is fast and agile, and it does not contradicts with the original purpose of the project; to achieve an autonomous system that detects fraud automatically.

In the future, our purpose will be to improve this method, making it more dispensable the hand of the specialist in the quality control method.

ACKNOWLEDGEMENTS

This work has been supported by funds from the Spanish Ministry for Economy and Competitiveness (MINECO) and the European Union (FEDER funds) via grants GRAMM (TIN2017-86727-C2-1-R) and Gas Natural Fenosa Distribución.

REFERENCES

- [1] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. *CoRR* abs/1603.02754 (2016). <http://arxiv.org/abs/1603.02754>
- [2] Bernat Coma-Puig, Josep Carmona, Ricard Gavaldà, Santiago Alcoverro, and Victor Martin. 2016. Fraud Detection in Energy Consumption: A Supervised Approach. In *2016 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2016, Montreal, QC, Canada, October 17-19, 2016*. 120–129. DOI: <http://dx.doi.org/10.1109/DSAA.2016.19>
- [3] Rong Jiang, H. Tagaris, A. Lachsz, and M. Jeffrey. 2002. Wavelet based feature extraction and multiple classifiers for electricity fraud detection. In *Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. IEEE/PES*, Vol. 3. 2251–2256 vol.3. DOI: <http://dx.doi.org/10.1109/TDC.2002.1177814>
- [4] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*. 1135–1144.
- [5] M Sforna. 2000. Data mining in a power company customer database. *Electric Power Systems Research* 55, 3 (2000), 201 – 209. DOI: [http://dx.doi.org/10.1016/S0378-7796\(00\)00086-9](http://dx.doi.org/10.1016/S0378-7796(00)00086-9)