Explaining Radio Access Network User Dissatisfaction with Multiple Regression Models

Adrien Schaffner
LivingObjects
Toulouse, France
adrien.schaffner@livingobjects.com

Louise Travé-Massuyès
LAAS-CNRS & ANITI, University of Toulouse
Toulouse, France
louise.trave@cnrs.fr

Simon Pachy
LivingObjects
Toulouse, France
simon.pachy@livingobjects.com

Bertrand Le Marec
LivingObjects
Toulouse, France
bertrand.lemarec@livingobjects.com

***Abstract*—The evaluation of user satisfaction is an essential performance indicator for network operators. It can be impacted by several causes, including causes linked to the network. In addition to constantly surveying and monitoring the network, network operators count the complaints received at customer services to know the evolution of the dissatisfaction rate. The difficulty is to link the subjective comment of a customer with an objective behavior of the network. Experience shows that an indicator taken from complaints, gives a good trend on the level of network quality perceived by customers, but it is difficult to transpose into concrete actions because it is often unrelated to the key performance indicators on which engineers base their action plans. The objective of this work is to design a model that links the complaint rate, expressed by the *Customer Satisfaction Rate* indicator, with a set of key performance indicators so that performance engineers better understand customer expectations and act primarily on the indicators that give the most dissatisfaction. The model hence makes it possible to link quality of experience and quality of service.**

***Index Terms*—Regression models, Data analysis, Knowkedge extraction, Radio access networks, Quality-of-Service/Quality-of-Experience relationship, Quality via Quality-of-Experience and customer reports.**

I. INTRODUCTION

In the space of a few years, the telecom market has undergone numerous technological and regulatory transformations which have engendered a price war from which operators are now trying to get out. They try to better differentiate themselves by moving towards a better customer experience and better support. The evaluation criteria most often adopted to establish a comparison of mobile networks are field measurement campaigns or user satisfaction surveys. User satisfaction surveys are expressed by the number of complaints received, the presence or absence of unfair terms in contracts, the commercial network and telephone assistance, connection time as well as call drop rate and their management noted by a supervisory authority, such as ARCEP (Regulatory Authority for Electronic Communications and Posts) in France or FCC (Federal Communications Commission) in United States.

The Customer Satisfaction Rate (CSR) is a good performance indicator that helps operators to effectively manage and control their business and decision making. The CSR provides the number of complaints relative to the number of customers for a given area. However, predicting customer behavior, their level of satisfaction (or dissatisfaction) has always been a challenge for operators. It is therefore important to link the CSR to a set of Key Performance Indicators (KPI) that can easily be interpreted by performance engineers to act on the relevant causes of dissatisfaction. This paper presents how to learn this link from data in the form of a regression model while selecting a set of explanatory KPIs from an oversized, but yet relevant, set. The regression model captures the relationship between Quality of Experience (QoE) and Quality of Service (QoS).

The contents of the paper are organized as follows. Section II analyzes related work and positions the method of this paper with respect to the state of the art. Section III formulates the problem as a regression problem and provides the identified issues. Section IV presents two regression methods, Ordinary Least Squares (OLS) and east Absolute Shrinkage and Selection Operator (LASSO), that are later used in the method. Section V describes the application referring to explaining the customer complaint indicator CSR that has been driving the design of the method. It also presents the data that has been used and the KPIs that have been considered as candidate explanatory variables. Section VI explains the bricks of the fusion method and the fusion method itself. The results of applying the fusion method to the CSR problem are interpreted in Section VII. Finally, Section VIII concludes the paper.

II. RELATED WORK

Much research investigated about customer complaint behavior since long [1] [2]. The idea of using complaint data to solve problems in design, marketing, installation, distribution and after sale use and maintenance, is quite natural. Understanding of customer complaint and market behavior has also been investigated so as to provide a framework for interpreting the data and extrapolating it to an entire customer base [3]. Especially in the mobile telecom industry, studies on customer complaint behaviour are numerous and continue

 ISBN: 978-1-61208-944-7

today, significantly accentuated by the emphasis on machine learning techniques.

Given the increased competitiveness in this field, many studies have focused on a problem related to customer complaints, which is customer churn. Due to the direct effect on the revenues of the companies, especially in the telecom field, companies are seeking to develop means to predict potential customer to churn. Over the years, many machine learning algorithms have been used to produce churn prediction models and building feature's engineering and selection methods [4] [5] [6]. In the churn problem, not only complaint data but Henley segmentation, call details, line information, bill and payment information, account information, demographic profiles, service orders, etc. are potentially important. In this huge set of features, [7] identifies a subset of relevant features and applies several prediction techniques including Logistic Regressions, Multilayer Perceptron Neural Networks, Support Vector Machines and the Evolutionary Data Mining Algorithm in customer churn as predictors, based on the subset of features. [8] uses classification like the Random Forest algorithm, as well as, clustering techniques to identify the churn customers and provide the factors behind the churning of customers by categorizing the churn customers in groups.

In this paper, the focus is put on using solely complaint data to solve problems in maintenance. To do so, this work aims at linking the complaint rate with a set of technical KPIs that point at the cause of the complaints and suggest reconfiguration or repair actions on the network. This problem is much less explored in the literature than that of the churn. Literature can be exemplified by [9] that achieves correlation analysis and prediction between mobile phone users complaints and telecom equipment failures in three steps involving hierarchical clustering, pattern mining, and decision trees. On the other hand, [10] uses four machine learning algorithms, Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Decision Tree (DT) experimented on a database of 10,000 Korean mobile market subscribers and the variables of gender, age, device manufacturer, service quality, and complaint status. It found that ANN's prediction performance outperformed other algorithms. This last work takes into account much more data than those fixed by the objective of this paper. In addition, the first focuses on equipment failure while we want to handle the KPIs which are the data used on a daily basis by network monitoring operators. Last but not least, the algorithms used are certainly good for prediction, but they are limited in their ability to explain predictions. The relation between the prediction and the inputs of the model remains implicit. On the contrary, the objective of this work is to clearly explain this link so that it provides useful information. This is why, the approach has been based on simple regression models.

III. Problem formulation

Multiple linear regression [11] is a classic family of learning algorithms that postulates that a variable is expressed as the weighted sum of other variables. Multiple linear regression defines the conditions and the model according to which a quantitative variable y is explained by several other quantitative variables $x_j, j = 1, \dots, p$. y is considered *dependent* or *endogenous* and the variables $x_j, j = 1, \dots, p$ are said to be *explanatory* or *predictor* variables. Multiple linear regression assumes that the variation of each explanatory variable has an influence, with not necessarily equal proportions, on the behavior of the dependent variable. The function that relates the dependent variable to the explanatory variables is linear.

Summarizing, multiple linear regression is a learning method that postulates that a variable y (here y=CSR) is expressed as the weighted sum of other variables (here the KPIs). Formally, for a number p of KPIs $x_j, j = 1, \dots, p$, the goal is to learn weights $\beta_0, \beta_1, ..., \beta_p$ such as:

$$y = \beta_0 + \beta_1 x_1 + ...\beta_p x_p \tag{1}$$

For this, we have a dataset gathering n observed samples, $n > p + 1$, each of dimension $(p + 1)$ and identified by the index i:

$$(x_1^i, x_2^i, \dots, x_p^i, y^i),\ i = 1, \dots, n, \tag{2}$$

that we use to estimate the parameters β_k $(k = 0, \dots, p)$ assumed to be constant. Each sample is assumed to satisfy the relation (1) with an error ϵ_i:

$$y^i = \beta_0 + \beta_1 x_1^i + ...\beta_p x_p^i + \epsilon_i,\ i = 1, \dots, n. \tag{3}$$

Under some statistical assumptions of the error terms ϵ_i, in particular independence and identical distribution, the vector of the parameters $\beta = (\beta_1, \dots, \beta_p)^T$ and the nuisance parameter σ^2 defining the variance of the error $\epsilon = (\epsilon_1, \dots, \epsilon_n)^T$, i.e., $var(\epsilon) = \sigma^2 I$, can be estimated by classical methods like least squares minimization [12] or, assuming that the error terms follow a centered normal distribution, likelihood maximization [13].

The model obtained after estimation of the parameters can be evaluated by the coefficient of determination R^2.

$$R^2 = \frac{SSR}{SST} = \frac{\sum_i^n (\hat{y}^i - \bar{y})^2}{\sum_i^n (y^i - \bar{y})^2} \tag{4}$$

where $\hat{y}^i$ is the prediction for the i-th sample, $\bar{y}$ is the mean, SSR is the sum of squares due to regression, i.e., the variability from the mean $\bar{y}$ that the regression manages to explain, and SST is the sum of squares total, i.e., the variability of the observed variables around the mean. R^2 represents the proportion of variance for the dependent variable that is explained by independent variables in the regression model. The closer the value of R^2 is to 1, the better the regression. In practice, the threshold value for R^2 for considering a good regression is highly dependent on the problem.

The goal of the obtained regression model is to extract knowledge, i.e., to determine the KPIs that influence the CSR and to quantify their influence from the coefficients of the regression.

In practice, the problems to be faced are the following :

- Business experience tells us that each of the explanatory KPIs can only worsen the condition of the telecom

 ISBN: 978-1-61208-944-7

network and therefore logically increases the CSR (e.g., an increase in the call drop rate, in the expert's mind, naturally increases the CSR). It is hence important to take care of the signs of the coefficients obtained by the regression.
- The number of candidate KPIs for explanation is high and can lead to irrelevant models.

IV. Two classical linear regression methods

This section presents the principle of two classical multiple regression methods then leveraged in the proposed fusion method presented in Section VI.

A. Ordinary Least Squares

When trained with data, Ordinary Least Squares (OLS) method selects parameter values β_j, $j = 1, \ldots, p$ of the linear expression (1) by the principle of *least squares*. It minimizes the sum of the squares of the differences between the observed dependent variable value in the data $y^i, i = 1, \ldots, n$, and the value predicted by the linear function of the independent variables $\hat{y}^i, i = 1, \ldots, n$. The optimization criterion, or loss function, is thus given by:

$$\begin{aligned} \mathcal{L} &= \min_{\beta_0, \beta_1, \ldots, \beta_p} \sum_{i=1}^{n} (y^i - \hat{y}^i)^2 \\ &= \min_{\beta_0, \beta_1, \ldots, \beta_n} \sum_{i=1}^{n} (y^i - \beta_0 - \sum_{j=1}^{p} x_{i,j})^2 \end{aligned} \tag{5}$$

In geometrical terms, this can be seen as the sum of the squared distances, parallel to the axis of the dependent variable, between each data point in the set and the corresponding point on the regression surface. The smaller the differences, the better the model fits the data.

The OLS estimator is consistent, i.e., has convergence to the real parameters values as the training data is increased, when the regressors are exogenous. It is optimal in the class of linear unbiased estimators when the errors are homoscedastic, i.e., they have the same variance, and are serially uncorrelated. Under these conditions, the method of OLS provides minimum-variance mean-unbiased estimation when the errors have finite variances. Under the additional assumption that the errors are normally distributed, OLS is the maximum likelihood estimator.

In this work, the function `ols` of the Python module `statsmodels` has been used to implement OLS.

B. Least Absolute Shrinkage and Selection Operator

Least Absolute Shrinkage and Selection Operator (LASSO) is a regression method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the resulting model. In other words, the LASSO method handles the complexity of the model with L1 regularization [14], so that the variables not having a contribution to the model are automatically removed from the regression. This means that it adds the "absolute value of magnitude" of coefficients as penalty term to the loss function $\mathcal{L}$:

$$\begin{aligned} \mathcal{L} &= \min_{\beta_0, \beta_1, \ldots, \beta_p} \sum_{i=1}^{n} (y^i - \hat{y}^i)^2 + \lambda \sum_{j=1}^{p} \mid \beta_j \mid \\ &= \min_{\beta_0, \beta_1, \ldots, \beta_n} \sum_{i=1}^{n} (y^i - \beta_0 - \sum_{j=1}^{p} x_{i,j})^2 + \lambda \sum_{j=1}^{p} \mid \beta_j \mid \end{aligned} \tag{6}$$

LASSO shrinks the less important feature's coefficients to zero thus, removing some explanatory variables altogether. This method works well for feature selection, particularly in case of a huge number of explanatory variables.

If λ is set to zero, then LASSO gets back OLS whereas a very large value increases zero coefficients hence it under-fits.

In this work, the fonction `lassocv` of the Python module `statsmodels` has been used to implement LASSO.

V. Data and pre-processing

The goal is to predict the CSR and the influencing factors on a global scale, and not on each specific site, so that the operator retrieves aggregated information useful for decision making. The project was hence conducted using data at the level of French *departments* (France has 93 departments which define as many territorial communities) by setting as many regression problems as French departments.

As for the features used, the advice of telecom experts was followed and led to a mixture of signals for both 2G, 3G, and 4G for six classes: traffic (like `downlink data traffic`), availability (like `signaling failure rate`), drop rates, accessibility, performance (like `data_failure rate`), and mobility (like `handover_drop_rate`). In total, 50 KPIs were in the list of explanatory variables, to divide between Data and Voice. Data and Voice are indeed considered to be truly independent from a customer perspective. However, the technical KPIs used by experts to explain voice and data performance have an important common basement. Among the 35 KPIs of the voice list and the 30 KPIs of the data list, 15 KPIs were common to the two lists.

The available data for each department covered a full year. While both daily and weekly values were considered, it was eventually decided to stick with daily ones, to retain a bigger dataset in the training and avoid losing information by averaging over 7 days.

In a context where the number of explanatory variables is high, it is quite often the case that several variables provide the same information or that some variables remain almost constant, or also that some variables have been poorly sensored. In order to remedy this problem, classic data pre-treatment solutions were applied in a first step resulting in:

- Removing strongly correlated variables, more precisely those with correlation coefficient higher or equal to 0.8;
- Removing variables of low variance through the dataset, more precisely those whose relative standard deviation was lower or equal to 10% of the highest;

 ISBN: 978-1-61208-944-7

- Removing variables with more than 10% missing values. Interpolation was used to fill the gaps for the remaining variables.

In addition, all variables were scaled so that they could be ranked according to the magnitude of their corresponding weights in the regressions.

VI. A FUSION REGRESSION METHOD WITH SELECTION OF EXPLANATORY VARIABLES

Despite the pre-processing carried out and the elimination of a subset of the KPIs proposed by experts in the field, the number of KPIs remains high, which suggests that still several of them have no direct impact on the CSR. In order to tackle this problem, the idea is to apply the following three approaches and then obtain consolidated results by fusioning the results of each of them:

- Multicollinearity analysis with OLS (MCOL),
- Iterative reduction via p-value with OLS (ITER),
- Structure learning with LASSO (LASSO).

Each of the methods has its own way to tackle the problem of selecting the most relevant explanatory variables, as explained in Sections VI-A, VI-B, and VI-C. To obtain the benefits of the three methods and smooth out the inconsistencies, the three methods are then fusioned as explained in Section VI and illustrated in Fig. 1. This strategy follows the analysis of [15] whose results suggest the need to examine models using multiple variable selection methods, because when they do not agree, they each may expose different aspects of the complicated theoretical relationships among predictors.

Methods MCOL and ITER rely on the classical Ordinary Least Squares method (OLS) presented in Section IV-A whereas LASSO, Least Absolute Shrinkage and Selection Operator, uses the method of this name in its original version of linear regression as presented in Section IV-B.

A. Multicollinearity analysis with OLS

The MCOL method builds on OLS adding an additional preprocessing step that selects a subset of features based on multicollinearity analysis.

In a regression, multicollinearity is a problem that arises when some explanatory variables in the model measure the same phenomenon. Strong multicollinearity is problematic because it can increase the variance of the regression coefficients and make them unstable and difficult to interpret. Strongly correlated predictor coefficients will vary considerably from sample to sample. They may even present the wrong sign.

Multicollinearity does not affect the goodness of the fit or the quality of the forecast. However, the individual coefficients associated with each explanatory variable cannot be interpreted reliably whereas this interpretation is exactly what we are looking for.

Multicollinearity and correlation should not be confused. If collinear variables are de facto strongly correlated with each other, two correlated variables are not necessarily collinear. There is collinearity when two or more variables measure the ”same thing”.

Classically, in case of quantitative explanatory variables, multicollinearity can be assessed by the *variance inflation factor* (VIF). The VIF for an explanatory variable is equal to the ratio of the overall model variance to the variance of a model that includes only that single explanatory variable. This ratio is calculated for each explanatory variable. The VIF estimates how much the variance of a coefficient is ”increased” due to a linear relationship with other predictors. Thus, a VIF of 1.7 tells us that the variance of this particular coefficient is 70% greater than the variance that should be observed if this factor was absolutely not correlated with the other predictors. The ideal case is obviously when all VIFs are equal to 1, indicating that there is no multicollinearity.

In the case study, multicollinearity analysis was performed considering the 35 and 30 KPIs indicated by the experts in the Voice and Data lists respectively. The VIF threshold was chosen to be 5, beyond which the corresponding KPI was eliminated. Fig. 2 shows the results obtained on a specific cell.

B. Iterative reduction via p-value with OLS

After training a regression model, a *p-value* for each KPI can be obtained: it tests the null hypothesis that the coefficient is equal to zero, in other words, whatever its value, the KPI brings no information whatsoever to the model. A low p-value (typically 0.05 or less) indicates that one can reject the null hypothesis: a predictor that has a low p-value is probably a meaningful addition to the model as it changes the model prediction. Conversely, a larger p-value implies that changes in the predictor do not bring changes in the response.

Whenever a new KPI is added or deleted in the training phase, the model obtained will get different regression coefficients, but also different p-values thus it makes sense to add or remove high p-value KPIs one at a time, in line with a backward elimination strategy in stepwise regression. The algorithm is then as follows:

- train a model with all KPIs,
- check which KPIs have a high p-value,
- remove the one with the highest p-value.

Although stepwise regression methods are recognized as undesirable for explanatory purposes [16] they may, however, provide efficient means to examine multiple models for further investigation.

C. Structure learning with LASSO

The lasso method is well known in the literature and has already proved itself in numerous regressions. Here is a quick reminder of the presentation of Section IV-B : in the standard regression like OLS, coefficients are obtained through minimization of the residual squared sum. The LASSO method is similar but adds a penalization term to reduce the number of KPIs kept during the regression. The penalization takes the form of an L1 norm of the coefficients which reduces the available domain of values, allowing some coefficients to be precisely zero, thus letting one remove the matching KPIs.

 ISBN: 978-1-61208-944-7

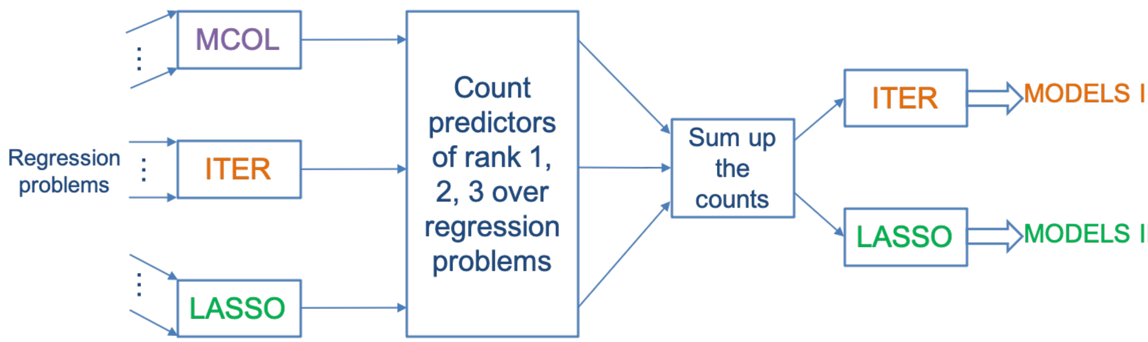

Fig. 1. Steps of the fusion regression method

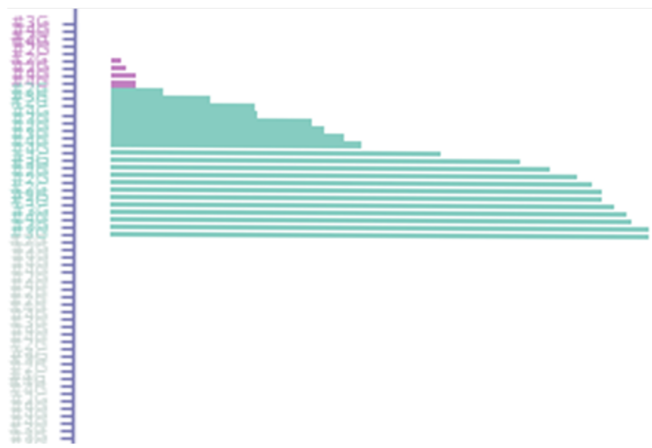

Fig. 2. KPI selection and relevancy on a specific cell: grey KPIs are those discarded by pre-processing and multicollinearity analysis, green KPIs are those of minor impact on the CSR, magenta KPIs are those that are preponderant according to the obtained regression model.

D. Fusioning the methods

In the regression model given by (1), explanatory variables $x_j, j = 1, \dots, p$, can be ranked according to the magnitude of their corresponding weight $\beta_1, \dots, \beta_p$. The idea developed in this work uses this ranking and includes four steps for the fusion regression method:

- *Step 1* – For every regression problem, learn three regression models with the three selected methods involving explanatory variable selection, namely MCOL, ITER, and LASSO;
- *Step 2* – For each method, count the number of times a given explanatory variable (KPI) has rank 1, 2, or 3 over the whole set of regression problems;
- *Step 3* – Sum up the counts over the three methods and select the explanatory variables whose count exceeds a threshold $\mathcal{T}$;
- *Step 4* – For every regression problem, learn two regression models with ITER and LASSO considering only the explanatory variables selected at the previous step and deduce the most impacting variables and the final model.

The steps of the fusion regression method are illustrated in Fig. 1. The output of the method takes the form of two sets of models called MODELS I and MODELS II, from which knowledge about most influencing explanatory variables can be extracted as explained in Section VII.

The fusion method is exemplified with the CSR prediction problems set at the level of French departments.

Step 1-2 are illustrated in Fig. 3 that gives the results for the Voice performance problem. For each explanatory KPI, the blue, orange, and grey bars provide the number of times the KPI is ranked 1, 2 or 3 by the MCOL, ITER, and LASSO method, respectively. Let us note a good convergence of the count referring to ITER and LASSO.

Step 3 is illustrated in Fig. 4. It aggregates the counts for each method and sums them up. It hence represents the sum of the counts of the number of times an explanatory KPI is ranked 1, 2, or 3 by one of the methods MCOL, ITER, and LASSO indifferently. A threshold is chosen, here at 45, and the explanatory KPIs that count above this threshold are selected. There are 7 KPIs that count above the threshold, framed in red.

Step 4 considers the 7 ”survivor” KPIs as the most relevant on the prediction of the CSR. This is why step 4 reconsiders every regression problem by restricting explanatory variables to these 7 KPIs. Only the ITER and LASSO methods are considered because of the good convergence of their results. The obtained results are provided in Fig. 5 that represents the KPIs ranked 1, 2, and 3 over ITER and LASSO and over all the French departments.

VII. Making Sense of the Predictions

Let us recall that the objective of this work is to design a model that makes it possible to link the CSR indicator with a set of objective performance indicators so that performance engineers better understand customer expectations and act primarily on the indicators that give the most dissatisfaction. The results of the prediction problems can be analyzed in two ways: at the level of each French department, and aggregated for the whole France.

 ISBN: 978-1-61208-944-7

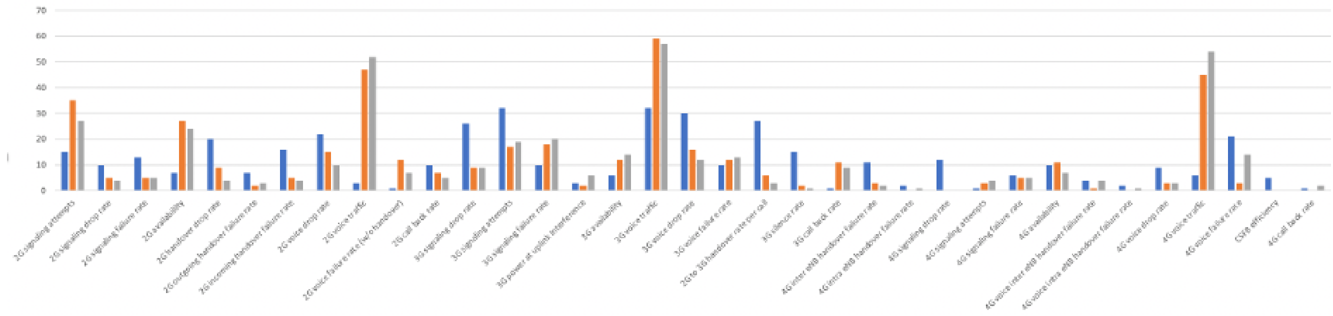

Fig. 3. Steps 1-2 of the fusion method for the Voice performance problem: count of the number of times an explanatory KPI is ranked 1, 2, or 3 by MCOL (blue), ITER (orange), LASSO (grey).

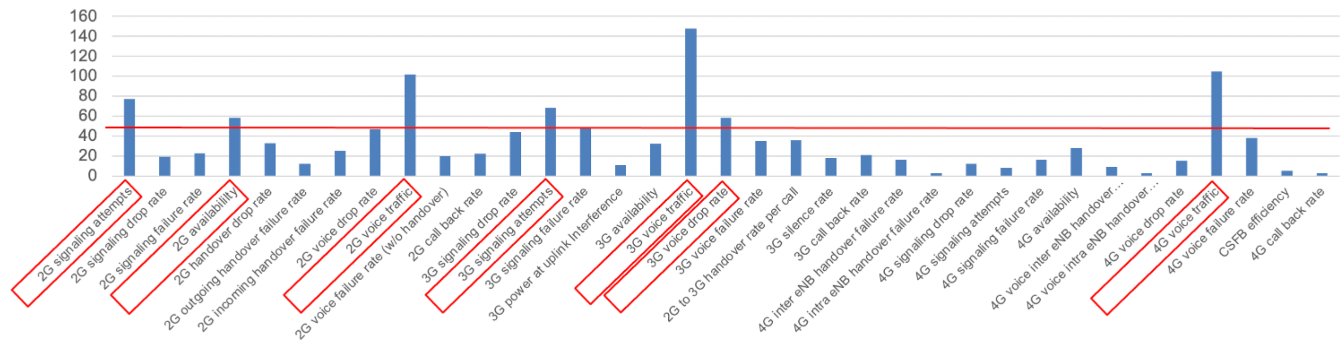

Fig. 4. Step 3 of the fusion method for the Voice performance problem: sum of the counts of the number of times an explanatory KPI is ranked 1, 2, or 3 by MCOL, ITER, LASSO. KPIs framed in red count above the threshold.

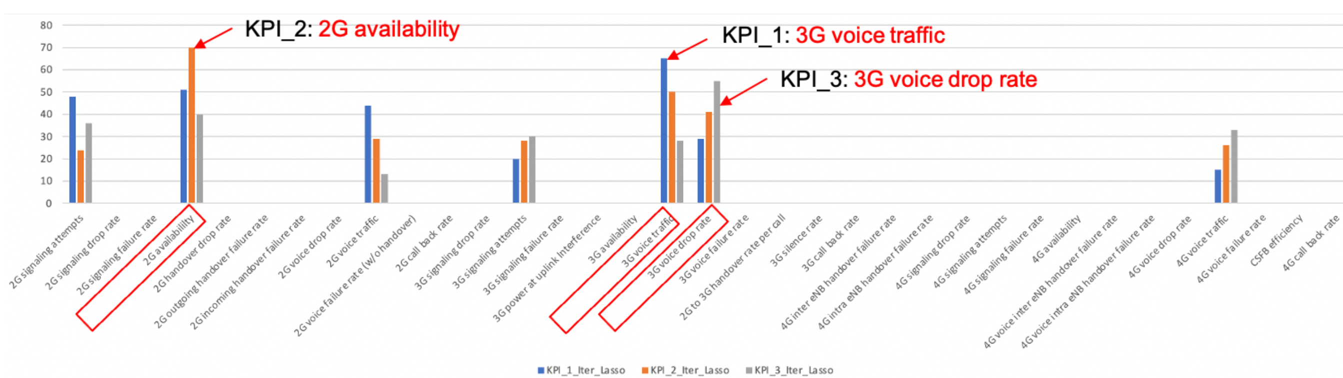

Fig. 5. KPIs ranked 1, 2, and 3 over ITER and LASSO and over all the French departments.

A. Interpretation at the level of each French department

This is done by associating a profile to each department. For this purpose, the results of the ITER method applied to the 7 survivor KPIs have been used and the department profile has been obtained by clustering the coefficients of the obtained models. This leads to the map in Fig. 6 where the departments that have similar profile are depicted with the same color. A similar profile indicates that the KPIs that must be mainly incriminated are the same, and so are the reasons explaining client complaints.

B. Aggregated interpretation

This interpretation is provided by step 4 of the fusion method. The three top KPIs over ITER and LASSO and over all the French departments appear in red in Fig. 5 and are: `3G voice traffic`, `2G availability`, `3G voice drop rate`.

This indicates that complaints are highly related to network behavior. Among the various metrics used to measure network behavior, it appears that `2G availability` that represents network maintenance processes and `3G voice drop rate` that represents network call drops are the most

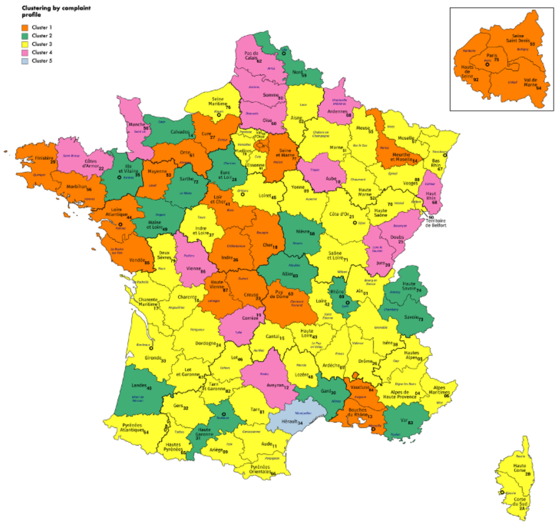

Fig. 6. Map with departments colored by profiles given by the weight of top KPIs influencing the CSR.

significant KPIs to express customer dissatisfaction, which is intuitively understandable. Traffic represented by the `3G voice traffic` KPI is also a relevant metric to assess the impact of network operation on customer satisfaction. It can be related to network unavailability, loss of coverage, and network engineering issues. Let us also notice that other metrics like accessibility failure rate or mobility issues are less significant than call drops or traffic issues.

To improve client experience, the network operators should therefore prioritize to base their action plans on:

- reducing unavailability periods by, for instance, optimizing the maintenance process,
- improving the call drop rate by modifying network parameter settings, optimizing site engineering, or building new sites.

VIII. Conclusion and perspectives

This paper proposes a method to obtain a regression model with explanatory power. In many applications, the number of variables that could be thought to be explanatory for a given dependent variable is huge. However, many of them are correlated or collinear and others do not really impact the predicted variable. The method presented in this paper leverages the benefits of three methods to select relevant explanatory variables and deduce a robust regression model.

The method has been tested on telecom data to obtain a model that indicates the link of the complaint rate with a set of objective performance indicators so that performance engineers better understand customer expectations and act primarily on the indicators that give the most dissatisfaction. The final results can be used to cluster French departments according to their profile as a function of the top influencing KPIs. They can also be used on a global scale to exhibit the top KPIs at country level.

Future work will consider mapping the top KPIs returned by the model to actual actions to be performed on the network so that customer satisfaction is increased, i.e. CSR is decreased. This mapping could benefit from ideas coming from the combination of the theories of prospect theory and satisfaction games found in the literature, such as [17].

Acknowledgment

We would like to thank the telecom expert Didier Rana for his remarks and relevant suggestions throughout this work.

References

[1] J. Jacoby and J. J. Jaccard, “The sources, meaning, and validity of consumer complaint behavior: A psychological analysis.” *Journal of retailing*, vol. 57, no. 3, pp. 4–24, 1981.

[2] J. Singh and R. E. Wilkes, “When consumers complain: A path analysis of the key antecedents of consumer complaint response estimates,” *Journal of the Academy of Marketing science*, vol. 24, no. 4, pp. 350–365, 1996.

[3] J. Goodman and S. Newman, “Understand customer behavior and complaints,” *Quality Progress*, vol. 36, no. 1, pp. 51–55, 2003.

[4] W.-H. Au, K. C. Chan, and X. Yao, “A novel evolutionary data mining algorithm with applications to churn prediction,” *IEEE transactions on evolutionary computation*, vol. 7, no. 6, pp. 532–545, 2003.

[5] A. K. Ahmad, A. Jafar, and K. Aljoumaa, “Customer churn prediction in telecom using machine learning in big data platform,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–24, 2019.

[6] A. Amin, F. Al-Obeidat, B. Shah, A. Adnan, J. Loo, and S. Anwar, “Customer churn prediction in telecommunication industry using data certainty,” *Journal of Business Research*, vol. 94, pp. 290–301, 2019.

[7] B. Huang, M. T. Kechadi, and B. Buckley, “Customer churn prediction in telecommunications,” *Expert Systems with Applications*, vol. 39, no. 1, pp. 1414–1425, 2012.

[8] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, “A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector,” *IEEE Access*, vol. 7, pp. 60 134–60 149, 2019.

[9] Q. Yang, G. Ji, and W. Zhou, “The correlation analysis and prediction between mobile phone users complaints and telecom equipment failures under big data environments,” in *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE, 2017, pp. 201–206.

[10] C. Choi, “Predicting customer complaints in mobile telecom industry using machine learning algorithms,” Ph.D. dissertation, Purdue University, 2018.

[11] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.

[12] G. S. Watson, “Linear least squares regression,” *The Annals of Mathematical Statistics*, pp. 1679–1699, 1967.

[13] I. J. Myung, “Tutorial on maximum likelihood estimation,” *Journal of mathematical Psychology*, vol. 47, no. 1, pp. 90–100, 2003.

[14] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[15] P. Ruengvirayudh and G. P. Brooks, “Comparing stepwise regression models to the best-subsets models, or, the art of stepwise,” *General linear model journal*, vol. 42, no. 1, pp. 1–14, 2016.

[16] G. Smith, “Step away from stepwise,” *Journal of Big Data*, vol. 5, no. 1, pp. 1–12, 2018.

[17] S. Papavassiliou, E. E. Tsiropoulou, P. Promponas, and P. Vamvakas, “A paradigm shift toward satisfaction, realism and efficiency in wireless networks resource sharing,” *IEEE Network*, vol. 35, no. 1, pp. 348–355, 2020.