Dynamic Multi-criteria Classifier Selection for Illegal Tapping Detection in Oil Pipelines

Victor Henrique Alves Ribeiro*†, Pedro Henrique Domingues‡, Paulo Rodrigo Cavalin†,
Gilberto Reynoso-Meza*, Helon Vicente Hultmann Ayala‡, and Luís Fernando Alzuguir Azevedo‡

*Industrial and Systems Engineering Graduate Program, *Pontifícia Universidade Católica do Paraná*, Curitiba, Brazil
Email: victor.henrique@pucpr.edu.br, g.reynosomeza@pucpr.br
†IBM Research, Rio de Janeiro, Brazil
Email: victor.henrique@ibm.com, pcavalin@br.ibm.com
‡Department of Mechanical Engineering, *Pontifícia Universidade Católica do Rio de Janeiro*, Rio de Janeiro, Brazil
Email: phd.engmec@gmail.com, helon@puc-rio.br, lfaa@puc-rio.br

***Abstract*—Illegal tapping of fuel pipelines has recently become one of the most relevant safety problems faced by the industry. Hundreds of illegal interventions have been reported around the world, causing a significant number of deaths, relevant impacts on the environment, and capital loss. Therefore, it is important to develop systems that are able to detect such scenarios at an early stage, enabling a fast counteract. To this end, machine learning algorithms can train models on available data for detecting future issues. Most recently, ensemble learning and dynamic classifier selection (DCS) techniques have been achieving promising results in supervised learning tasks. Such models are usually trained based on a single criterion. However, it is desirable to take into account both the number of false positives (FP) and false negatives (FN) for the illegal tapping detection task, since they are conflicting and both lead to financial losses and/or accidents. Therefore, this work proposes a novel DCS technique based on multiple criteria, namely overall local class-specific accuracy (OLCA), which employs multi-criteria decision making for dynamically selecting the best classifier for a new sample given the local true positive and negative ratios. A numerical experiment is conducted for assessing the generalization performance of the proposed method in an oil pipeline, with the goal of detecting illegal taping using pressure transient signals. Results show that OLCA is able to reduce the number of both FP and FN when dynamically selecting the classifiers of a baseline Random Forest ensemble.**

***Index Terms*—Ensemble learning, dynamic classifier selection, multi-criteria decision making, time series classification, oil pipelines, illegal tapping.**

I. Introduction

In 2018, the entity committed to ensure the safe use of petroleum substances in Europe (Concawe) has published the Report no. 6/18 [1]. Such report alerted the seriousness of the illegal tapping problem, showing that the sharp increase in the number of cold pipeline spills is due to the increased number of fuel theft events. Pipeline leakage causes environmental and economic consequences [2]. For instance, the Niger Delta region in Nigeria reaches 12 billion American dollars annually with losses related to crude oil theft [3]. Nevertheless, the development of illegal tapping detection systems has recently become a relevant field in both academy and industry [4].

Machine learning algorithms can be employed for building such a system. Therefore, a classifier can be trained to detect leak and illegal tapping of fuel pipelines given a data set with available examples. Recent studies have built system using techniques such as k-nearest neighbors (k-NN) [5], support vector machines (SVMs) [5]–[8], and artificial neural networks (ANNs) [6], [9]–[11].

In addition to machine learning, ensemble learning can be used for generating and combining multiple classifiers, as they are known for improving the results of single models [12]. Moreover, dynamic classifier selection (DCS) has attained attention in literature for building dynamic ensembles that use different base models given the characteristics of the sample to be classified [13], [14]. However, to the best of the authors knowledge, ensemble learning has not yet been explored for developing an illegal tapping detection system.

In addition to this, usual learning algorithms rely on the optimization of a single criterion during the learning process. However, multiple conflicting criteria must be met when solving complex engineering problems. The development of an illegal tapping detection system must take into account not only a global accuracy, but also the number of false positives (FPs) and false negatives (FNs), two conflicting types of errors that lead to great financial losses and environmental issues. Therefore, this paper proposes the development of a novel dynamic multi-criteria classifier selection technique.

To test the illegal tapping detection system based on the new proposed technique, a numerical experiment is prepared using simulated data based on a real-world oil pipeline. Prior to the machine learning approach, signal preprocessing and feature engineering is performed to enable a better generalization of the models. Next, an ensemble of base classifiers is built on training data, and the dynamic multi-criteria selection technique is employed on a separate testing data. Results show that the proposed method is able to successfully reduce both the numbers of FPs and FNs when compared to an "off-

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES), the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), and the *Fundação Araucária* (FAPPR) - Brazil - Finance Codes: 159063/2017-0-PROSUC, 310079/2019-5-PQ2, 437105/2018-0-Univ, 405580/2018-5-Univ, and PRONEX-042/2018.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

the-shelf" ensemble learning technique, namely random forest (RF) [15].

The remainder of this paper is organized as follows: Section II introduces the problem of illegal tapping; Section III proposes a novel DCS technique based on multi-criteria decision making (MCDM); Section IV details the numerical experiment; and Section V discusses the results; Finally, the paper is concluded with some final remarks and future research.

II. Illegal Tapping

Criminals have been performing fluid and gas theft in pipelines around the world, making illegal tapping an worldwide issue [4]. In addition to financial losses, they are also responsible for environmental consequences due to leakage [16]. Briefly, there are two approaches for illegal tapping in pipelines: "offensive theft" and "sophisticated theft" [17].

On the one hand, the offensive approach focuses on a quick withdrawal of the product. This is easily identified if the pressure transmission cables are not previously damaged. It also has a great spill potential if badly executed. On the other hand, the sophisticated approach uses low flow rates for the theft activity. This makes the detection and location both time consuming and difficult. Despite the lower environmental risk of the sophisticated theft, the volume of stolen product generally exceeds that of the other approach if the tapping is not detected at an early stage [17].

To counteract such a criminal activity, leak detection techniques can be used for illegal tapping detection, such as pattern recognition, or machine learning [16]. To this end, the most common approach is based on the detection of negative pressure waves [18]. The resulting models are designed to be sensitive to variations in the pressure signals, caused by leak or illegal tapping, that roam through the pipeline. A review of recent literature on illegal tapping and leak detection systems is detailed next.

A. Illegal Tapping and Leak Detection Systems

Rostek, Morytko and Jankowska [9] developed a diagnostic and prediction system based on ANN, which is divided into: (1) early fault detection by virtual sensors; and (2) leak isolation through fault state classification. The system has been applied to six blocks of a professional plant and proved to be able to detect 11 out of 12 failures at least two days in advance, being efficient in distinguishing between three classes of leaks. The authors also reported a number of works related to early detection and prediction of leaks in fluidized-bed boilers using ANN.

Xiao et al [6] designed a new method for small leak detection based on variational mode decomposition (VMD) and ambiguity correlation classification (ACC). Briefly, the method applies the VMD to the acquired sensor's signal and uses the probability density function to make an adaptive de-noising algorithm, which processes noise components and reconstruct de-noised ones. The ambiguity function image is used for the reconstruction component analysis, and the ACC is built based on the correlation coefficient. The method proved to be efficient in detecting small leaks from $1mm$ and $2mm$ wide holes, achieving a better performance than SVMs and ANNs.

Zadkarami, Shahbazian and Salahshoor [10] implemented a fault detection and isolation system capable of recognizing the leakage and suggesting its location, as well as severity. The system is based on a simulation software that provides the pipeline inlet pressure and outlet flow rates used for extracting three types of features for training an ANN: statistical, wavelet transform and a merge of these two. The resulting system was applied to a $20km$ long real pipeline in southern Iran, yielding a correct classification rate for severity and location identification ability of 92%, with a small false alarm rate.

Rahmati et al [11] trained an ANN based on gas flow pattern for a gas leak detection system. The pipeline is divided into sections and modeled based on the inlet and outlet pressures of each segment, which allows the generation of gas flow data used to train and evaluate the performance of the ANN. The method was validated using real gas flow data measured through wireless sensor network and industrial internet of things.

Kayaalp et al [5] presented a water pipeline real-time monitoring system based on wireless sensor network and a multi-label learning method. The study consisted of acquiring pressure data from wireless pressure sensor nodes and using three multi-label learning methods (random k-label sets, binary relevance k-NN and binary relevance with SVM) to detect and locate the water leakage. The results showed that: (1) multi-label classification methods can be successful for detecting and locating pipelines leaks; and (2) the random k-label sets classification method yielded the best results in almost all measures, with an accuracy of 98%.

Li et al [19] developed a novel leakage location algorithm based on the attenuation of negative pressure wave. The approach is attractive for: (1) deducing the negative pressure wave propagation equation using momentum and continuity equations, avoiding the problem of velocity disturbance by the pipeline liquid flow rate; and (2) relying on pressure change rather than time difference, which is difficult to determine. The method performed better than a traditional negative pressure wave method in most cases, presenting errors between 1.161% and 0.355%.

Xie, Xu and Dubljevic [7] proposed a leak detection and localization system based on a real pipeline modeled by nonlinear coupled first-order hyperbolic partial differential equations. At first, a discrete-time Luenberger observer is designed by solving the operator Riccati equation and allowing the reconstruction of the pressure and mass flow velocity evolution with limited measurements, which permitted the generation of various upstream and downstream velocity profiles for normal and leakage conditions. Then, a SVM model is trained and tested from statistical features extracted from the velocity profile data, presenting an overall accuracy of 99%.

He et al [20] built: (1) the framework for a big data, cloud computing, and internet of things technology based monitoring and accidental leak handling system; (2) a new leak location method based on negative pressure waves; and (3) a strategy for emergency shutdown after the leaking identification, which calculates the volume of spilled product and selects the strategy that reduces it. The experimental results in multi-product pipelines proved that the framework can accurately estimate the leak starting time, location, coefficient, and volume over the period of time required for a negative pressure wave to reach the full pipe length.

Finally, Liu et al [8] implemented a leak detection system based on Markov feature extraction from the pressure data, least square SVM, and a two-stage decision scheme. The former switches between a short or long term detection model for a rapid and precise identification of the pipeline status. The proposed system obtained an average accuracy of 99.17% and 92% in conventional and small leakage, respectively, and a maximum false alarm rate of 10%.

III. Dynamic Multi-criteria Classifier Selection

To build an illegal tapping detection system, this Section proposes a novel DCS technique based on multiple criteria. First, a brief introduction to DCS is given. Next, the importance of MCDM and its applicability to machine learning are discussed. Finally, the new proposed technique is detailed.

A. Dynamic Classifier Selection

Much progress has been achieved in the field of pattern recognition with the combination of multiple classifiers (or ensemble of classifiers). In addition to this, recent literature started performing dynamic selection of the classifiers to improve the predictive performance in many different tasks [13]. The dynamic selection is performed by firstly defining a region of competence based on the new samples. To define such a region of competence, researchers have usually relied on techniques such as clustering and k-NN. Finally, given a selection criteria, one or more base models are selected to predict the new sample given their local prediction scores [14].

Moreover, literature started using more than one selection criteria. For instance, selection has already been performed given the N most accurate and J most diverse classifiers [21], [22]. Also, classifiers are selected given an aggregate function computed with data complexity and accuracy in dynamic selection on complexity (DSOC) [23]. However, few techniques perform an analysis from the multi-criteria perspective, which can benefit the performance of predictive models by taking into account the trade-off between conflicting objectives [24]. META-DES uses additional characteristics from the data set to select the base learners [25]. Most recently, the hesitant fuzzy MCDM has been applied for dynamic ensemble selection [26]. Despite this, the trade-off relation between conflicting class-specific scores has not yet been studied. Therefore, this work aims at further exploring MCDM for DCS.

B. Multi-criteria Decision Making

When dealing with complex engineering problems, it is usually important to take into account more than one criterion. For instance, the occurrence of both FPs and FNs must be minimized when developing an illegal tapping detection system. On the one hand, FP relates to the detection of an illegal tapping where nothing has occurred in reality, which leads to financial losses due to unnecessary work. On the other hand, FN, indicates the number of times where an illegal tapping is not detected, which leads to financial loss and possible environmental consequences due to undetected theft and leakage. Therefore, when developing such a detection system, considering a single criterion could lead to "wrongly neglecting certain aspects of realism" [27].

With such, the given task is considered a multi-criteria problem. In such problems, the criteria are usually conflicting, meaning that it is not possible to improve the results of one of them without affecting the performance of others. Therefore, there will exist situations where it is not possible to select the best classifier, since some of them will present better results for minimizing the number of FNs while presenting a higher number of FPs, and vice-versa.

MCDM can be used for solving the selection problem, which is responsible for aiding the process of decision making in problems where multiple criteria are taken into account [27]. Ranking techniques have been used with success in the literature [28], where algorithms compute an overall score given the multiple differences between each solution. Nevertheless, this work proposes a new DCS technique that employs MCDM, namely overall local class-specific accuracy (OLCA).

C. Overall Local Class-specific Accuracy

OLCA is based on the existing overall local accuracy (OLA) algorithm [29], where the accuracy on a region of competence is used for selecting a single classifier. However, this work makes use class-specific accuracy instead of the global one. Additionally, a MCDM technique is applied for ranking and selecting a preferred classifier. According to the existing taxonomy [14], the proposed algorithm is detailed in terms of region of competence definition, selection criteria, and selection approach.

An overview of OLCA is drawn in Figure 1. The first stage is the definition of a region of competence given a new sample, as described in Subsection III-C1. Next, given the region of competence and a pool of base models, which can be generated through algorithms such as RF [15], Bagging [30] or Boosting [31], the selection criteria are computed according to Subsection III-C2. Finally, the selection approach detailed in Subsection III-C3 results in a final output.

1) Region of Competence Definition: This work employs the k-NN for defining the region of competence. When a new sample is analyzed, k-NN selects 20 neighbors from the dynamic selection data set [14]. Next, the selection criteria are computed on the region of competence for selecting the preferred classifier.

2) Selection Criteria: In contrast to current DCS literature, where a single predictive performance criterion is analyzed for the dynamic selection, this work makes use of two criteria, true positives ratio (TPR) and true negatives ratio (TNR). In a binary classification problem, such metrics are related to

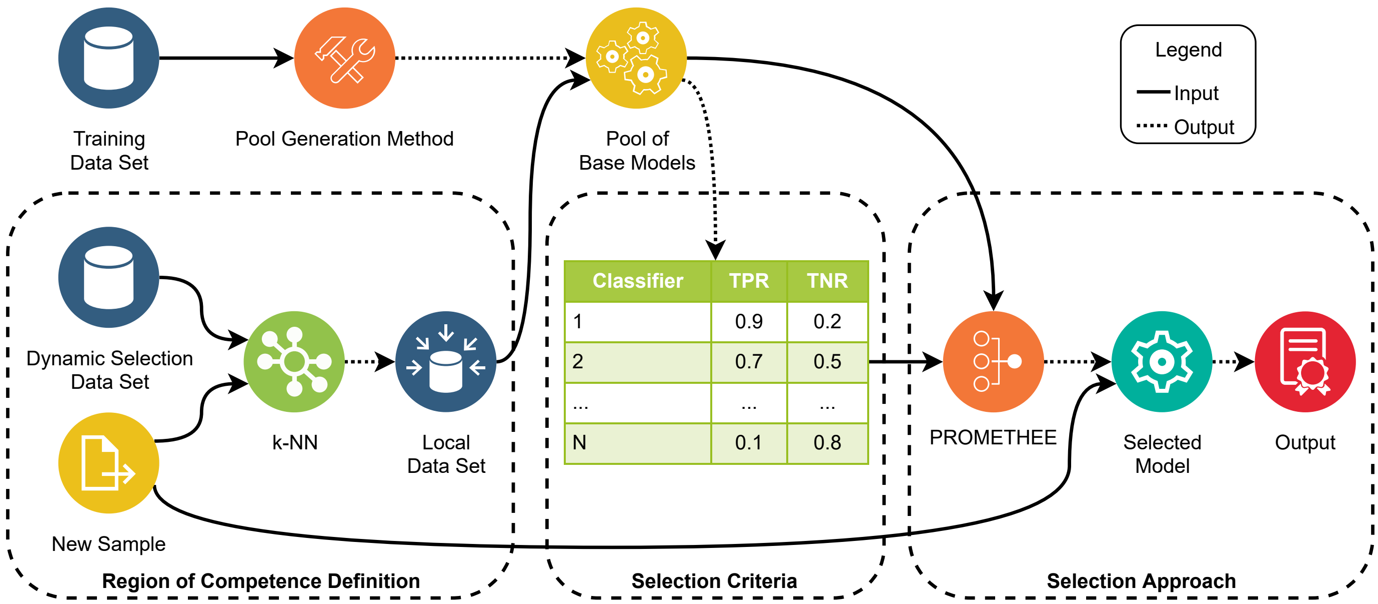

Fig. 1. An overview of the proposed overall local class-specific accuracy (OLCA) dynamic classifier selection (DCS) technique.

the class-specific accuracy. On the one hand, TPR indicates the ratio of positive samples that are correctly classified as such. On the other hand, TNR indicates the ratio of negative samples that are correctly classified as such. TPR and TNR are computed according to Equations 1 and 2, respectively. In the equations, true positive (TP) is the number of positive samples correctly predicted as such, true negative (TN) is the number of negative samples correctly predicted as such, FN is the number of incorrectly predicted positive samples, and FP is the number of incorrectly predicted negative samples.

$$TPR = \frac{TP}{TP + FN} \tag{1}$$

$$TNR = \frac{TN}{TN + FP} \tag{2}$$

3) Selection Approach: When dealing with both TPR and TNR, there will be situations where a single classifier cannot be considered the best solution. A classifier that presents a better TPR than others can be outperformed by them in terms of TNR. Therefore, a MCDM technique can be employed for selecting a preferred classifier. The preference ranking organization method for enriched evaluation (PROMETHEE) [32], [33] has been used with success in many real-world problems [28]. Such a technique computes a final score for each solution based on the outranking of such solutions in terms of multiple criteria. That is, the more a solution outranks other ones, the better its final score. To this end, pairwise comparisons are performed between all the solutions according to each criteria, where a preference function is employed to quantify the outranking scores. Therefore, it is necessary to define significant and insignificant difference values for each criterion. This work employs such a technique and configures it according to Table I, where both criteria have the same weight, differences lower than 1% are deemed insignificant, and differences higher that 10% are considered significant for both TPR and TNR.

TABLE I
MATRIX WITH WEIGHTS (W), INSIGNIFICANT (I), AND SIGNIFICANT (S) DIFFERENCES FOR THE TRUE POSITIVES RATIO (TPR) AND TRUE NEGATIVES RATIO (TNR) CRITERIA.

Criterion	W	I	S
TPR	1.00	0.01	0.10
TNR	1.00	0.01	0.10

IV. EXPERIMENT

This section presents the experiment for solving an illegal tapping detection problem using the OLCA algorithm. First, the case study for the numerical experiment is depicted. Next, the problem's data set is described, followed by the feature engineering procedure. Finally, the used models and evaluation metrics are detailed.

A. Case Study

To generate an illegal tapping scenario for evaluating the proposed technique, simulations were made with the Synergi Pipeline Simulator software[1] (version 10.6). The simulation model addresses the "sophisticated theft" problem, which is more difficult to detect. Such a model is based on an oil pipeline operating under actual field conditions, detailed in Figure 2.

The model considers two centrifugal boosters (A/B) with nominal flow rates of $2,960m^3/h$, five centrifugal pumps

[1] https://www.dnvgl.com/services/pipeline-simulator-and-surge-analysis-software-analyse-pipeline-design-and-performance-synergi-pipeline-simulator-5376

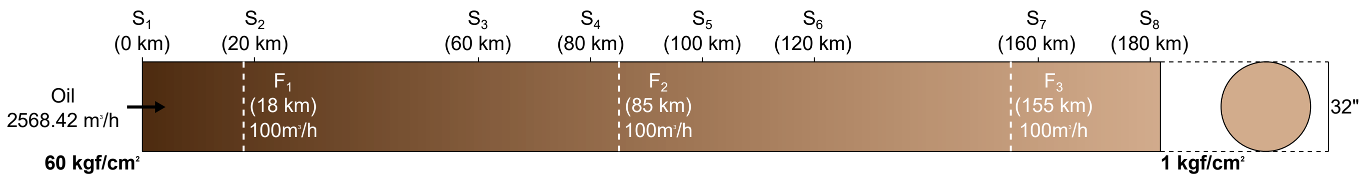

Fig. 2. Simulation model scheme of the oil pipeline.

(A/B/C/D/E) with nominal flow rates of $3,000m^3/h$, and minimum flow rates of $454m^3/h$, all arranged in series. The oil has a density of $0.8986kg/m^3$, viscosity of $66.36cP$, vapor pressure of $0.50kgf/cm^2abs$, and Bulk modulus of $15,751kgf/cm^2$. The total pipeline length is $181.83km$, and temperature is set to $20°C$.

Pressure transients for the illegal tapping points are simulated at the beginning (F_1), middle (F_2) and end (F_3) of the pipeline, with valves positioned at 18, 85 and $155km$, respectively. The opening rate yields $100m^3/h$ of theft flow in all points. Eight different sensors (S_i, $i = 1, ..., 8$), are positioned at 0, 20, 60, 80, 100, 120, 160 and $180km$ to record the pressure data.

B. Data Set

The data set is comprised of an annotated multivariate time series captured at a frequency of $10Hz$. As described previously, the data has been acquired by simulating a real scenario composed of eight pressure sensors along the oil pipeline. In total, seven different scenarios have been simulated: normal operation, a choke in the end of the pipeline, the actuation of a safety valve, the addition of a new pump, and the illegal tapping in the beginning, middle and end of the pipeline.

To evaluate the proposed model, the data set has been split into training and testing sets. The test set contains the data from the illegal tapping in the middle of the pipeline, while the training set contains all the remaining scenarios. This is performed to evaluate the model's generalization capacity for illegal tapping at unknown positions. Additionally, the training set is further split into 70% for training the base classifiers and 30% for creating the dynamic selection set. No validation set is used for hyper-parameter tuning. Figure 3 plots the sensor data for the test scenario, where oil theft occurs in the middle of the pipeline.

C. Feature Engineering

Despite having information available from eight pressure sensors disposed along the pipeline, it is not recommended to use the raw data for the data-driven task. Without a pre-processing stage, it would be necessary to collect data from all possible theft locations for training an accurate model. Therefore, feature engineering is employed for transforming the available data into knowledge that can be generalized. To this end, principal component analysis (PCA) [34] is employed for reducing the dimension from the eight sensors to only two variables.

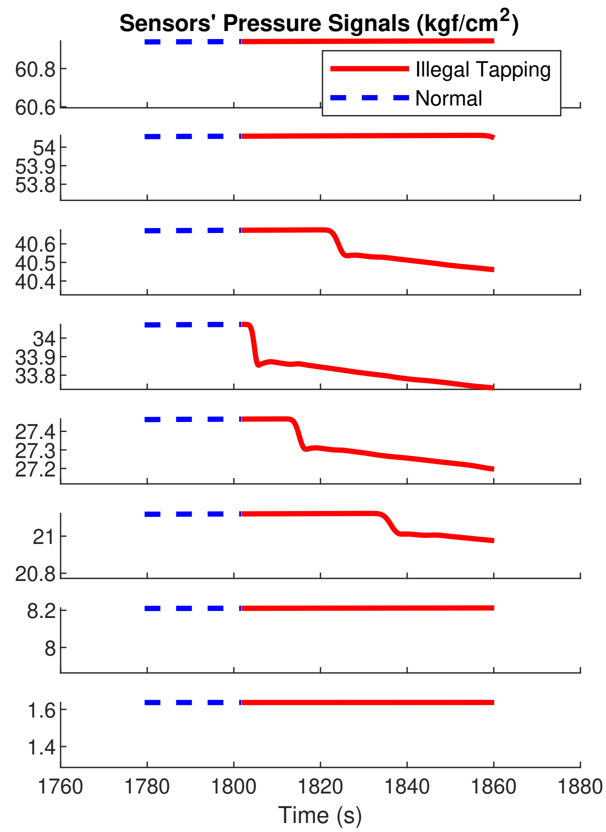

Fig. 3. Pressure signals from the eight available sensors for an unseen illegal tapping scenario.

First, since the pressure values from each sensor have different ranges, the raw data is transformed so that each sensor contributes similarly when using PCA. This is done according to Equation 3, where $x_{i,j-W}$ are the adjusted values for the pressure signals $p_{i,j}$ at position i and time step j, using window length of $W + 1 = 100$ samples (or 10 seconds) for J total observations. Next, the dimensionality reduction technique is also employed in a window of 10 seconds. With such, the pressure variation can be identified despite in which sensor it occurs first.

$$x_{i,j-W} = \frac{p_{i,j}}{\sum_{k=j}^{j-W} p_{i,k}}, i \in [1, \ldots, 8], W + 1 \leq j \leq J \quad (3)$$

Additionally, to add more information to the predictors, statistical features are extracted from the signal [35]. To this end, the 10 seconds-long moving average, standard deviation, maximum, and minimum values are collected for the first two principal components. As a result, a total of 10 features are available for the data-driven learning task. The resulting features are summarized in Table II.

TABLE II
SUMMARY OF THE RESULTING FEATURES.

Feature	Description
A	1^{st} Principal Component
B	2^{nd} Principal Component
C	Moving Average of Feature *A*
D	Moving Average of Feature *B*
E	Moving Standard Deviation of Feature *A*
F	Moving Standard Deviation of Feature *B*
G	Moving Minimum of Feature *A*
H	Moving Minimum of Feature *B*
I	Moving Maximum of Feature *A*
J	Moving Maximum of Feature *B*

D. Models

This work uses RF [15] for generating the base models of the ensemble. Such technique uses random samples of the training data set to generate multiple diverse decision trees. Additionally, random features are selected at each node of the trees to improve diversity. Such technique is heavily used in literature due to its high performance in learning tasks [36]. In total, 100 decision trees are generated with maximum number of splits equal to $n-1$ for n observations and selection of $\sqrt{f}$ random features for a total of f features.

To improve the predictive performance of the illegal tapping detection problem, this work employs OLCA for dynamically selecting the classifiers. With such, two models are evaluated: the original RF and the proposed RF with OLCA. Both models are trained and evaluated on the same data sets, where an improvement is expected for the proposed approach.

E. Evaluation Metrics

Finally, to evaluate the experimental procedure, four metrics are analyzed: TNR, TPR, Accuracy, and F1 score. As already mentioned, the first two are related to class-specific accuracy. Global accuracy (ACC) is employed for measuring the ratio of correctly assigned classes, computed according to Equation 4. The F1 score measures the harmonic mean between precision, or positive predictive value (PPV), and recall (TPR). Such metric is used for measuring the accuracy on imbalanced problems, and is computed according to Equation 5.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$F1 = \frac{2 \cdot PPV \cdot TPR}{PPV+TPR} \quad (5)$$

where

$$PPV = \frac{TP}{TP+FP} \quad (6)$$

V. Results and Discussion

The results from the previous experiment are detailed in this Section, which have been acquired using 100 base learners for the RF and the PROMETHEE parameters from Table I. First, Figure 4 plots the results of the time series, followed by the confusion matrix in Figure 5. Finally, Table III displays the numerical results.

Figure 4 plots the results on the test data set, where the 10 computed features are visualized and colored according to the model's predictions. The features are detailed in Table II. In the blue dashed line, the features remain linear, and are correctly identified as the negative class (no illegal tapping). Next, the black stars show the illegal tapping samples that were miss-classified. In the red line, features start to modify their behavior and are correctly detected as the positive class (illegal tapping). The main difficulty of the task is to detect the first illegal tapping samples, where there has not yet been much variations in the signals.

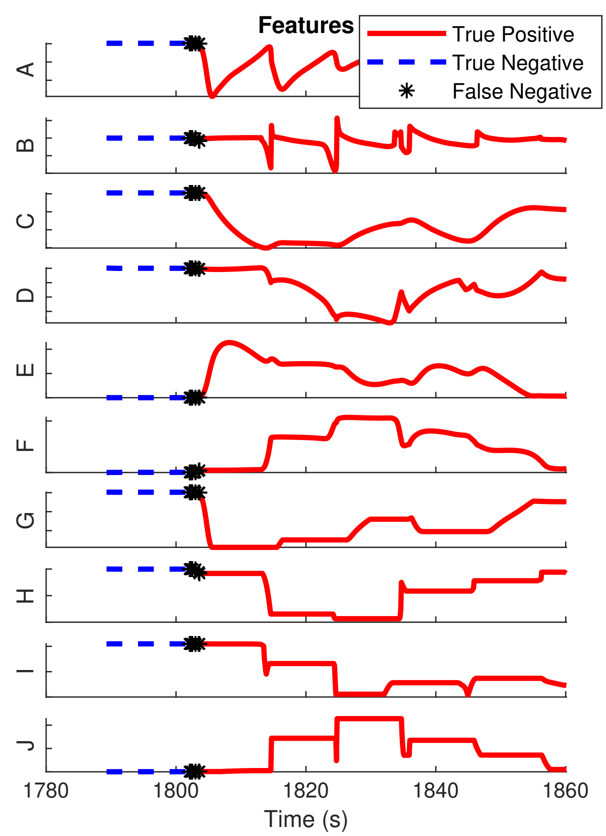

Fig. 4. Feature signals and detection results of the proposed method for the unseen illegal tapping scenario.

Figure 5 details the confusion matrix for the RF and the RF with OLCA models. On the one hand, when only the RF is used, 862 samples are correctly identified as the positive class and 12465 as the negative class. Only 1 FP and 14 FN occurs. On the other hand, when OLCA is employed, the number of FP falls to zero, while the number of FN falls to 11. Both models present great results for the task. However, the proposed dynamic selection technique is able to further improve the results.

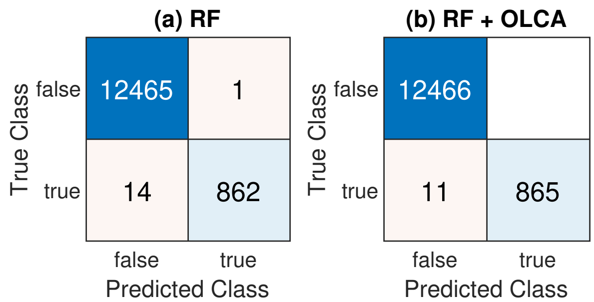

Fig. 5. Confusion matrix built on the test set for the (a) original random forest (RF) and (b) RF with overall local class-specific accuracy (OLCA) models.

Finally, Table III brings the numerical results for the task. The RF + OLCA method achieves the best results in terms of all evaluation metrics. The error is reduced by approximately 27% on both Accuracy and F1 Score. Moreover, the number of false positives is reduced by approximately 21% while the number of false negatives is completely reduced. This result confirms the significance of the proposed method for the given task, where it is able to improve the results even when excellent results are already achieved. Therefore, OLCA shows potential in real-world engineering tasks, and should be tested on new problems.

TABLE III
EXPERIMENT RESULTS FOR THE RANDOM FOREST (RF) AND RF WITH OVERALL LOCAL CLASS-SPECIFIC ACCURACY (OLCA) MODELS.

Model	Accuracy	F1 Score	TNR	TPR
RF	0.9989	0.9914	0.9840	0.9999
RF + OLCA	0.9992	0.9937	0.9874	1.0000
Error reduction	−27%	−27%	−21%	−100%

Such results can also be compared to current literature in illegal tapping and leak detection systems. The accuracy of the proposed method achieved highly accurate models, as the ones presented in Section II-A [5], [7], [8], [10], [19], [20]. Moreover, the number of false alarms has been completely reduced.

One important point to focus attention on is the application of PCA. Such tool enables the detection of the illegal tapping regardless of where in the pipeline it occurs. Without such technique for processing the signals, it would be necessary to acquire data from all possible tapping locations for a learning algorithm to generate a highly accurate model. Hence, dimensionality reduction techniques play an important role in the development of such a detection system.

Finally, as expected from using a dynamic classifier selection method, the local information aids in the selection of better classifiers for a given sample. Moreover, both the number of FPs and FNs have been reduced from using a multi-criteria approach. Therefore, it can be concluded that the proposed method is a promising tool for designing leak and illegal tapping detection systems. Nevertheless, such method can also be employed for solving other classification problems, and further study is necessary.

VI. Conclusions

This paper proposes a novel DCS algorithm based on MCDM for solving an illegal tapping detection problem in oil pipelines. Illegal tapping is a serious problem, which leads to great financial and environmental consequences. However, machine learning models can aid the development of a robust detection system. Ensemble learning can be used as well for achieving even better results. In addition to this, to further improve the results of a RF ensemble, local TPR and TNR scores are taken into account for dynamically selecting the best classifiers using the MCDM technique PROMETHEE. Experiments are performed using simulated data based on an oil pipeline operating under actual field conditions, where an unseen illegal tapping scenario must be correctly detected. Results prove the success of the proposed solution.

It is important to notice, however, that it is not only a strong machine learning model that aids the development of the detection system. Knowledge in engineering and signal processing plays an important role in the definition of features to be used for the specific problem. In the case of the illegal tapping detection task, the use of PCA enables not only the reduction in the dimensionality of the problem, but also reduces the number of necessary samples and scenarios that need to be simulated for inferring such intelligent models. Therefore, when implementing the system in the field condition, there will be a cost reduction in the data collection stage.

Finally, the proposed dynamic multi-criteria classifier selection technique, OLCA, has proven to further improve the results of a strong "off-the-shelf" RF ensemble model. With the novel technique, the numbers of FPs and FNs are both reduced. By reducing the number of FPs, companies and government benefit from less financial burden caused by unnecessary work. Additionally, by reducing the number of FNs, fuel losses due to theft and environmental consequences due to leakage are also minimized. Such results shows the importance of tackling complex engineering problems in terms of not only one, but multiple, often conflicting, criteria.

Future work on the method shall focus on the exploration of novel multi-criteria approaches for DCS and dynamic ensemble selection (DES). To this end, different methods can be used for the region of competence definition, selection criteria, and selection approach. Region of competence definition can be performed with other techniques, such as clustering. Different selection criteria can be used, such as complexity and diversity measures, as well as different performance metrics for regression and multi-class classification problems. Finally, different MCDM techniques can be explored for classifier and ensemble selection. Nevertheless, in order to fully understand the advantages of using a multi-criteria approach, a special attention shall be given to statistical comparison of such techniques with existing single-criterion techniques for dynamic selection in future studies.

Regarding illegal tapping detection, future work shall be focused on the application of the proposed method on field. In this context, an analysis of the sensors sensitivity is recommended to guarantee the negative pressure waves are correctly acquired in the actual field condition, as it was possible in the numerical experiment. Additionally, future work shall focus on the application of data-driven techniques for illegal tapping location detection.

References

[1] M. Cech, P. Davis, F. Gambardella, A. Haskamp, P. H. González, M. Spence, and J.-F. Larivé, *Performance of european cross-country oil pipelines: statistical summary of reported spillages in 2016 and since 1971*. CONCAWE, 2018.

[2] Q. Chen, G. Shen, J. C. Jiang, X. Diao, Z. Wang, L. Ni, and Z. Dou, "Effect of rubber washers on leak location for assembled pressurized liquid pipeline based on negative pressure wave method," *Process Safety and Environmental Protection*, vol. 119, pp. 181–190, 2018.

[3] T. Ngada and K. Bowers, "Spatial and temporal analysis of crude oil theft in the niger delta," *Security Journal*, vol. 31, no. 2, pp. 501–523, 2018.

[4] R. Landstorfer, "Illegal hot taping – a case study on how to detect and locate product theft," in *Pipeline Technology Conference 2019*. EITEP Institute, 2019.

[5] F. Kayaalp, A. Zengin, R. Kara, and S. Zavrak, "Leakage detection and localization on water transportation pipelines: a multi-label classification approach," *Neural Computing and Applications*, vol. 28, no. 10, pp. 2905–2914, 2017.

[6] Q. Xiao, J. Li, Z. Bai, J. Sun, N. Zhou, and Z. Zeng, "A small leak detection method based on vmd adaptive de-noising and ambiguity correlation classification intended for natural gas pipelines," *Sensors*, vol. 16, no. 12, p. 2116, 2016.

[7] J. Xie, X. Xu, and S. Dubljevic, "Long range pipeline leak detection and localization using discrete observer and support vector machine," *AIChE Journal*, 2019.

[8] J. Liu, D. Zang, C. Liu, Y. Ma, and M. Fu, "A leak detection method for oil pipeline based on markov feature and two-stage decision scheme," *Measurement*, vol. 138, pp. 433–445, 2019.

[9] K. Rostek, Ł. Morytko, and A. Jankowska, "Early detection and prediction of leaks in fluidized-bed boilers using artificial neural networks," *Energy*, vol. 89, pp. 914–923, 2015.

[10] M. Zadkarami, M. Shahbazian, and K. Salahshoor, "Pipeline leakage detection and isolation: an integrated approach of statistical and wavelet feature extraction with multi-layer perceptron neural network (mlpnn)," *Journal of Loss Prevention in the Process Industries*, vol. 43, pp. 479–487, 2016.

[11] M. Rahmati, H. Yazdizadeh, and A. Yazdizadeh, "Leakage detection in a gas pipeline using artificial neural networks based on wireless sensor network and internet of things," in *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)*. IEEE, 2017, pp. 659–664.

[12] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018.

[13] A. S. Britto Jr, R. Sabourin, and L. E. Oliveira, "Dynamic selection of classifiers—a comprehensive review," *Pattern Recognition*, vol. 47, no. 11, pp. 3665–3680, 2014.

[14] R. M. Cruz, R. Sabourin, and G. D. Cavalcanti, "Dynamic classifier selection: Recent advances and perspectives," *Information Fusion*, vol. 41, pp. 195–216, 2018.

[15] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[16] D. Vogt and M. Ihring, "Product theft detection and localization on pipelines using theft pattern recognition," in *Pipeline Technology Conference 2019*. EITEP Institute, 2019.

[17] M. Dudek and ttc members, "Liquid leak detection focused on theft protection," in *PSIG Annual Meeting*. Pipeline Simulation Interest Group, 2005.

[18] R. Ramadevi, J. Jaiganesh, and N. Krishnamoorthy, "Leak detection methods—a technical review," in *International Conference on Communications and Cyber Physical Engineering 2018*. Springer, 2018, pp. 125–139.

[19] J. Li, Q. Zheng, Z. Qian, and X. Yang, "A novel location algorithm for pipeline leakage based on the attenuation of negative pressure wave," *Process Safety and Environmental Protection*, vol. 123, pp. 309–316, 2019.

[20] G. He, Y. Li, Y. Huang, L. Sun, and K. Liao, "A framework of smart pipeline system and its application on multiproduct pipeline leakage handling," *Energy*, vol. 188, p. 116031, 2019.

[21] R. G. Soares, A. Santana, A. M. Canuto, and M. C. P. de Souto, "Using accuracy and diversity to select classifiers to build ensembles," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. IEEE, 2006, pp. 1310–1316.

[22] M. C. de Souto, R. G. Soares, A. Santana, and A. M. Canuto, "Empirical comparison of dynamic classifier selection methods based on diversity and accuracy for building ensembles," in *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*. IEEE, 2008, pp. 1480–1487.

[23] A. L. Brun, A. S. Britto, L. S. Oliveira, F. Enembreck, and R. Sabourin, "Contribution of data complexity features on dynamic classifier selection," in *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 4396–4403.

[24] V. H. A. Ribeiro and G. Reynoso-Meza, "A holistic multi-objective optimization design procedure for ensemble member generation and selection," *Applied Soft Computing*, vol. 83, p. 105664, 2019.

[25] R. M. Cruz, R. Sabourin, G. D. Cavalcanti, and T. I. Ren, "Meta-des: A dynamic ensemble selection framework using meta-learning," *Pattern recognition*, vol. 48, no. 5, pp. 1925–1935, 2015.

[26] J. Elmi and M. Eftekhari, "Dynamic ensemble selection based on hesitant fuzzy multiple criteria decision making," *Soft Computing*, pp. 1–13, 2020.

[27] B. Roy, "Paradigms and challenges," in *Multiple criteria decision analysis*. Springer, 2016, pp. 19–39.

[28] M. Behzadian, R. B. Kazemzadeh, A. Albadvi, and M. Aghdasi, "Promethee: A comprehensive literature review on methodologies and applications," *European journal of Operational research*, vol. 200, no. 1, pp. 198–215, 2010.

[29] K. Woods, W. P. Kegelmeyer, and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, no. 4, pp. 405–410, 1997.

[30] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[31] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[32] J.-P. Brans, P. Vincke, and B. Mareschal, "How to select and how to rank projects: The promethee method," *European journal of operational research*, vol. 24, no. 2, pp. 228–238, 1986.

[33] J.-P. Brans and Y. De Smet, "Promethee methods," in *Multiple criteria decision analysis*. Springer, 2016, pp. 187–219.

[34] I. Jolliffe, *Principal component analysis*. Springer, 2011.

[35] Z. Lan, O. Sourina, L. Wang, and Y. Liu, "Real-time eeg-based emotion monitoring using stable features," *The Visual Computer*, vol. 32, no. 3, pp. 347–358, 2016.

[36] K. Fawagreh, M. M. Gaber, and E. Elyan, "Random forests: from early developments to recent advancements," *Systems Science & Control Engineering: An Open Access Journal*, vol. 2, no. 1, pp. 602–609, 2014.