ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

Automated operational states detection for drilling systems control in critical conditions

Galina V. Veres and Zoheir A. Sabeur *

University of Southampton IT Innovation Center - Faculty of Physical and Applied Sciences
Gamma House, Enterprise Road, Southampton - UK

Abstract. Critical events in industrial drilling should be overcome by engineers while they maintain safety and achieve their targeted operational drilling plans. Geophysical drilling requires maximum awareness of critical situations such as “Kicks”, “Fluid loss” and “Stuck pipe”. These may compromise safety and potentially halt operations with the need of staff rapid evacuations from rigs. In this paper, a robust method for the detection of operational states is proposed. Specifically, Echo State Networks (ESNs) were benchmarked and tested rigorously despite the challenging unbalanced datasets used for training. Nevertheless, these challenges were overcome and led to acceptable ESNs performances.

1 Introduction

Real-time operations decision-making is performed by drilling managers for executing complex drilling procedures and actions while meeting their daily drilling depth targets. Despite their collective expert knowledge, drillers need further support from more advanced information systems, specializing in automated detection of drilling states. The automation of state detection and reporting shall improve future drilling operations, particularly for handling critical events. The extraction of drilling information with the detection models shall assist drillers to reduce their time for the analysis of large volumes of well-bores drilling sensor observation data and decision-making. Also, the integration of operational state detection models and storage of their results in drilling systems should improve audit-trails of drilling activities. These are conducted by multiple shifts of drilling teams during the 24hr operational cycles. The on-line access and learning from past decisions made during critical events by various teams shall potentially save operational costs and increase safety in drilling rigs.

2 Operational drilling in critical conditions

Critical drilling conditions do not usually arise abruptly in time. Drilling systems normally undergo transitional conditions prior to reaching criticality. Therefore, the ability to detect and control operational procedures and actions shall be very important to achieve in order to minimize critical situations during drilling operations.

* This work was partly supported by the European Commission FP7 IST Programme under contract TRIDEC IP FP7-258723.

Nevertheless, due to the lack of real-time knowledge and analysis on geological formation, it is very challenging for drilling teams to avoid critical situations. This is a difficult position of responsibility to take while assuring safety and low cost of operations. The time constraints for achieving sound decisions is also very tight, while the diagnosis made at each step of drilling procedures could encounter relative high uncertainties, and lead to costly mitigation measures for avoiding crises [1].

Ten drilling sensor parameters were considered to represent drilling processes in this study. These are used for monitoring the drilling system' functions and therefore can be used as indicators of critical conditions. However, it is challenging for drilling experts to diagnose, assess and make sense of the large volume of generated data from these 10 channels during very limited time to decision-making. The ten drilling Operation States (OS) which have been considered in this study include: Drilling rotary (DrlRot), Drilling sliding (DrlSld), Clean downwards (CleanDN), Clean upwards (CleanUP), Wash upwards (WashUP), Wash downwards (WashDN), Move in hole (MoveDN), Move out of hole (MoveUP), Circulation on (CirclHL) , Make connection (MakeCN). The OS can be reasonably detected under one drilling run, using machine learning techniques and the support of additional principal states such as String at Hook, Rock Penetrate, Fluid Circulate and String Rotate and Translate [2]. However, the challenge is to maintain such detection success rates in multiple drilling runs within a drilling phase and beyond.

In this paper, benchmarking is performed for the deployment and training of data-driven models to carry out OS detection under multiple runs per drilling phase. The models are trained using only 30% of the first drilling run data at each given drilling phase. Testing is then performed on the remaining unseen data at various subsequent runs of each phase. Drilling OS are then detected directly in order to avoid additional uncertainties caused by estimated principal states such as in hierarchical method used in [2]. Powerful Echo State Networks (ESNs) were deployed with significantly reduced training times. Techniques for using multi-class unbalanced data were also required due to the unbalanced datasets both for training and testing found during statistical analysis of the labeled data.

3 Automated OS detection

3.1 Multi-class imbalance problem

Multi-class imbalance problems are difficult to solve and present an important challenge for building reliable decision-support applications in engineering [3]. The most useful approach to address this issue involves the decomposition of the problem into a multiple *two-classes-classification problem* to handle each imbalance binary sub-problem accordingly. For the case of drilling OS detection, two decomposition schemes were adopted [4]: *one-against-one* (OAO) and *one-against-all* (OAA). With the OAA scheme, each class is trained against all classes. While with the OAO scheme, each class is trained against every single other class, i.e. a system of multiple classifiers has to be implemented. Considering the complex dynamics of drilling time series, the ESN was adopted as a classifier. A single ESN with multiple outputs that represent a number of classes was used for the OAA scheme, while multiple ESNs with a single binary output were deployed using the OAO scheme.

Two approaches were selected for dealing with the imbalance problems in this study: *a Prior Duplicate Sampling* (PDS) and *Inverse Random Under Sampling* (IRUS). PDS is based on duplicating data examples of minority class by a fixed number of times in a training set [5]. After this duplication, the number of examples in minority and majority classes becomes equal. IRUS is a recently proposed method for solving class imbalance problems. It showed significant performance gains on 22 UCI public datasets when compared with many other existing class-imbalance learning methods [6]. The main idea is to severely under sample the majority class (negative class), thus creating a large number of distinct negative training sets. In this study, both PDS and IRUS approaches are used as data pre-processing tools. PDS can be realized using a single ESN as a classifier under both OAA and OAO schemes. IRUS approach results in forming the new training sets. These sets can be considered as Multiple Time Series and used as inputs to a single ESN for training. In this case, OAO scheme is applied to make the decision about a sample class.

Extensive simulation experiments based on the different combinations of the approaches mentioned above were conducted. Thus, the most promising strategies for drilling OS detection are OAA, OAO, OAO with a prior duplicate sampling (OAO_PDS), and OAO with inverse random under sampling (OAO_IRUS) and ESN as classifier.

3.2 Echo state networks

ESNs are reservoir computing networks which are conceptually simple, computationally inexpensive, and able to learn complex dynamical behaviour [7]. The hidden layer (reservoir) consists of randomly connected neurons, some of which are connected in cycles. It allows the resulting states to be "echoes" of the past inputs in ESN. The hidden layer neurons are also randomly connected to the input signal which drives the network. Only the output weights are learnt; while all other weights including feedbacks are randomly selected and remain static. Detailed description of ESNs can be found in [7]. The decision function is used depending on a type of approach for modeling pattern classes. For OAA, the class label selected corresponds to the largest output value after ESN activation. The majority voting is used for OAO.

4 Simulation results

4.1 Experimental setup

In order to evaluate the effectiveness of the proposed approaches, OS detection was performed on sensor measurement data for monitoring drilling operations at a single well on a drilling rig. The measurements received from a mud-logging system on the drilling rig included: Block Position, Bit Depth, Hole Depth, Weight on Bit, Mud Flow, Pump Pressure, Rate of Penetration, Rotary Torque, Hook Load and Rotary Speed. Six additional features were also considered: 1) Hole Depth - Bit Depth, 2) Hole Depth + Block Position, 3) Bit Depth + Block Position, 4) Rotary Torque * Rotary Speed, 5) Pump Pressure * Mud Flow, 6) Rate of Penetration * Weight on Bit. The drilling operation was achieved in three phases: 12.25", 8.75" and 6.00", and models were trained and tested for each phase respectively. The following

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

assumptions were made about the data:

- Drilling Runs are known for both training and testing. Altogether 11 drilling runs took place in the experiment.
- Training was restricted to 30% of the data represented in the first Drilling Run at a given phase
- The OS labels were provided for training sets at each phase

We applied a plain ESN with 300 hidden units and 16 inputs for each algorithm. Single Time Series learning mode were used for OAA, OAO and OAO_PDS algorithms, while Multiple Time Series learning mode were used for OAO_IRUS. The spectral radius was set to 0.45 without input and teaching scaling with no feedback. One ESN was trained with 10 binary outputs for OAA and 45 ESNs with 1 binary output for other algorithms. For each drilling phase, a new single or set of multiple ESNs was trained. In OAO scheme ties are broken as follows. If a label in ties occurred at the previous step, it is used as an output. Otherwise, the ties are randomly broken. Table 1 below shows the unbalanced OS labels for Drilling Runs. The most representative class is DrlRot, followed by DrlSl then MakeCN. The rarest classes include CleanUp, WashDN and WashUp.

CircHL	CleanDN	CleanUP	DrlRot	DrlSld	MakeCN	MoveDN	MoveUP	WashDN	WashUP
7%	6.2%	1.4%	38.3%	23.4%	13.2%	2.8%	3.4%	2%	1.9%

Table 1: Distribution of OS labels for Drilling Runs.

4.2 Algorithms comparison

Four algorithms were compared in this section. These are: OAA, OAO, OAO_PDS and OAO_IRUS. Micro-averaged and macro-averaged F-measures for overall performance at a drilling phase together with Correct Classification Rate (CCR) for each OS were adopted. The F-measure metric is equal to the harmonic mean of recall (ρ) and precision (π), and its values are in the interval [0,1]. The overall π and ρ are obtained by summing over all individual classes:

$$\pi = \frac{TP}{TP+FP} = \frac{\sum_{i=1}^{M} TP_i}{\sum_{i=1}^{M}\left(TP_i + FP_i\right)},\ \rho = \frac{TP}{TP+FN} = \frac{\sum_{i=1}^{M} TP_i}{\sum_{i=1}^{M}\left(TP_i + FN_i\right)} \qquad (1)$$

Where M is the number of classes, while TP, FP and FN are the number of their true positives, false positives and false negatives respectively. The micro-averaged and macro-averaged F-measures are computed as follows:

$$F_1 = \frac{2\pi\rho}{\pi+\rho} \quad \text{and} \quad F_2 = \frac{\sum_{i=1}^{M} F_i}{M},\ \text{where}\ F_i = \frac{2\pi_i\rho_i}{\pi_i+\rho_i}. \qquad (2)$$

Micro-averaged F-measure F_1 gives equal weights to each label and is therefore considered as an average over all labels/classes pairs. It tends to be dominated by the classifier's performance on common classes. Macro-averaged F-measure F_2 gives equal weights to each class regardless of its frequency.

Table 2, below, shows the overall performance of four algorithms for the 8.75" Drilling Phase. The best performing algorithm both for common and rare classes overall is OAO_IRUS, while the poor performing algorithm is OAA. OAO

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

overperforms OAO_PDS on rare classes, while OAO_PDS shows better performance on common classes.

Algorithm	OAA	OAO	OAO_PDS	OAO_IRUS
F_1	51%	71.9%	74%	78.8%
F_2	28.7%	37.5%	32.4%	40.8%

Table 2: Overall performance for 8.75" Drilling Phase.

The overall performance does not show how good the algorithms are for different classes. However, the most important operational states are DrlRot, DrlSld and MakeCN. The aim is that these should be classified with the highest possible accuracy. The misclassification of the remaining states is in fact less critical. Figure 1 shows CCR for each OS. All algorithms except OAA show good performance (>90%) for DrlSld, except OAO_PDS for MakeCN. The best performance is achieved with OAO_PDS and OAO_IRUS for DrlRot. All algorithms showed low CCR (<40%) for CleanDN/CleanUP and WashDN/WashUP. CleanUP, WashUP and MoveUP are often miss-detected as CleanDN, WashDN and MoveDN respectively and vice versa. CircHL is another OS with a low detection level, though significant improvement (100% or more) in CCR can be achieved using the OAO or OAO_IRUS algorithms. It is difficult to select the best algorithm for drilling applications, since the respective performances of these algorithms varies from OS to OS; and from Drilling Phase to another. Based on CCR for each OS, OAO and OAO_IRUS show more robust performance for the majority of considered OSs.

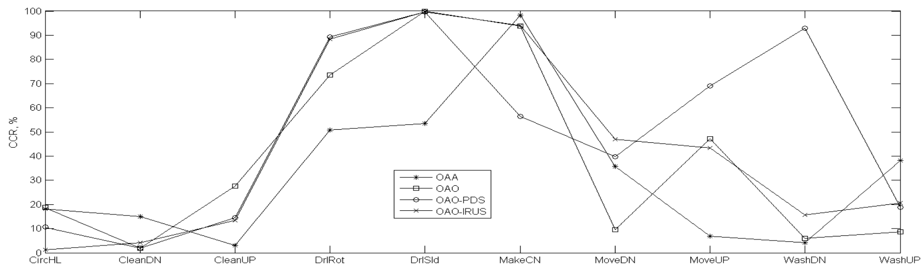

Fig. 1: CCR per OS for Drilling Phase 8.75".

The performance of the algorithms was also studied through runs at a given phase. Figure 2 shows how F1 and F2 measures changes for each algorithm as drilling progresses after training. The OAA algorithm tends to degrade in performance both for common and rare classes. Similar observation is displayed with the OAO for common classes. After significant drop in F2 from Run3 to Run4, it stays approximately stable for both Run5 and Run6. All algorithms show unsatisfactory performance for Run6 and common classes. The drop in both F1 and F2 in the algorithms took place for Run4. The best preforming algorithm for progressive drilling is OAO_IRUS.

ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0.
Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

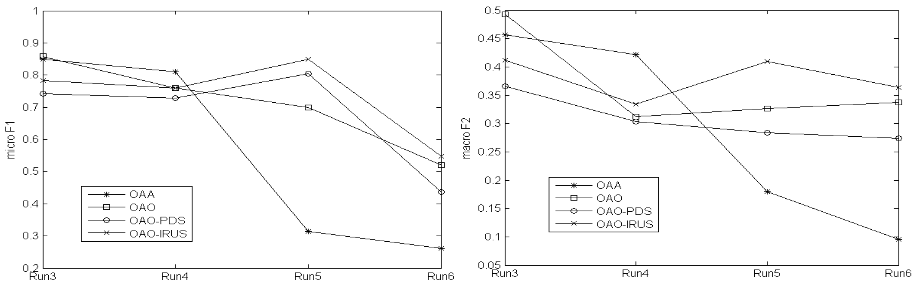

Fig. 2: F1 and F2 per Drilling Run for Drilling Phase 8.75".

5 Conclusions and future development

A set of ESNs were implemented for the automated detection of industrial drilling operational states. Their overall performance is satisfactory for the detection of critical states and low for the less relevant ones. This may be due to the self-imposed benchmarking for training below 30% of the labeled dataset. Also the inclusion of high level features representing the direction of operations may improve the overall results. The ESNs were also tested on unseen multiple runs to explore the limits of their performances under one drilling phase. This approach was strategic to put in place in order to launch future deployment of ESNs as Multiple fused ESNs under an adaptive OS framework with dynamic error estimations.

References

[1] H. Luo, J. Yin, and X. Jia. "The application of associated rules algorithm in the prediction of drilling troubles," in *Proc. 2011 3d Int. Workshop on Intelligent Systems and Application (ISA),* pp.1-3.

[2] B. Esmael, R. Fruhwirth, A. Arnaout, and G. Thonhauser. " Operations Recognition at Drill-Rigs," *EGU General Assembly 2012,* Vienna, Austria, April, 2012.

[3] S. Wang, and X. Yao. "Multi-class imbalance problems: analysis and potential solution," *IEEE Trans. Systems, Man and Cybernetics, part B,* vol. 42(4), pp.1119-1130, 2012.

[4] G. Ou and Y.L. Murphey. "Multi-class pattern classification using neural networks," *Pattern Recognition,* vol. 40, pp.4-18, 2007

[5] Y.L. Murphey, and H. Guo. "Neural learning from unbalanced data," *Applied Intelligence*, vol. 21, pp. 117-128, 2004

[6] M. A. Tahir, J. Kittler, and F. Yan. "Inverse random under sampling for class imbalance problem and its application to multi-label classification," *Pattern Recognition*, vol. 45, pp. 3738-3750, 2012

[7] H. Jaeger. “The "echo state" approach to analyzing and training recurrent neural networks ,” German National Research Center for Information Technology, GMD Report 148, 2001