Learning CRF Models from Drill Rig Sensors for Autonomous Mining

Sildomar T. Monteiro, Fabio Ramos and Peter Hatherly
Australian Centre for Field Robotics
School of Aerospace, Mechanical and Mechatronic Engineering
The University of Sydney
Sydney, NSW 2006, Australia
s.monteiro@acfr.usyd.edu.au

Abstract

This paper investigates an approach that combines ensemble methods with graphical models to analyse multiple sensor measurements in the context of mine automation. Drill sensor measurements used for drilling automation have the potential to provide an estimate of the subsurface geological properties of the rocks being drilled. A Boosting algorithm is used as a local classifier mapping drill measurements to corresponding geological categories. A Conditional Random Field then uses this local information in conjunction with neighbouring measurements to jointly reason about their categories. Model parameters are learned from training data by maximizing the pseudo-likelihood. The probability distribution of classified borehole sections is calculated using belief propagation. We present experimental results of applying the method to classify rock types from sensor data collected from a semi-autonomous drill rig at an iron ore mine in Australia.

1 Introduction

Modern semi-autonomous drill rigs are equipped with various sensors which provide measurements while drilling (MWD) used to control and monitor the drilling process. A major challenge for autonomous mining is to build accurate representations of the in-ground geology to determine the quantity and quality of the minerals of interest. The idea is to relate the drilling measurements to geotechnical properties of the rocks being drilled. Characterizing subsurface geology from drill measurements can be of substantial value for the mining industry. The accurate assessment of lithology and rock strength can be used to maximise the recovery of the desired rock types and improve blasting design by accurately determining the optimal explosive load and distribution. Previous works focused on determining empirical indices for rock strength based on drilling parameters, e.g. [1]. Machine learning techniques have also been applied to this problem, e.g. [2, 3]. However, previous methods do not model spatial dependencies of nearby geology.

In this paper, we investigate the problem of jointly estimating the geology of neighbouring regions by modelling measurements logged while drilling from multiple sensors. A classification method that takes into account spatial relationships is investigated in the framework of conditional random fields (CRFs). CRFs can handle arbitrary dependencies between observations which give them substantial flexibility in modelling complex geological dependencies in the data. The CRF framework is applied in conjunction with boosting algorithms. Boosting is used in this work to non-linearly map the drill measurements to the estimated geology. The set of labels classified by Boosting are used in the CRF framework to learn model parameters discriminatively. The resulting CRF model specifies the spatial relationship between MWD data providing an improved geological classification of the rock layers.

2 Conditional random fields for rock classification

CRFs are undirected graphical models that directly model the conditional probability of the hidden states given the observations rather than the joint probability thus avoiding the difficult task of specifying a generative model [4]. The CRF model can be applied to jointly reason about drilling measurements and neighbouring sections in the axial direction of a borehole by using a chain-like structure as illustrated in Fig. 1. The edges in the graph represent potential functions mapping sensor measurements to non-negative numbers. Drill measurements are considered observed variables and are represented by shadowed nodes. Borehole section categories, which are not observed, correspond to latent (hidden) variables and are represented by clear nodes. The relationships between nearby borehole sections are represented by edges connecting them.

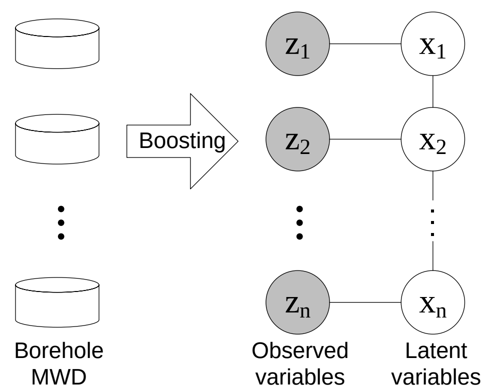

Figure 1: Graphical model of the CRF framework representing the spatial association between neighbouring borehole sections. The observations z_i correspond to drill measurements and the latent variables x_n indicate the corresponding classes.

Inference in CRFs can estimate either the marginal distribution of each hidden variable or the most likely configuration of all hidden variables (*i.e.*, MAP estimation). The computation of the partition function can be exponential in the number of variables, thus exact solutions might be unfeasible. However, exact inference can be computed in polynomial time for the particular graphical model used in this paper (linear chain). Both inference tasks can be solved using *belief propagation* (BP), which works by sending local messages through the graph structure of the model [5].

The CRF parameters can be learned by maximizing the conditional likelihood of labelled training data. Unfortunately, global optimisation using a numerical gradient algorithm runs an inference procedure at each iteration, which can be intractably inefficient in our case. We therefore resort to maximizing the *pseudo-likelihood* of the training data [6]. We optimise the pseudo-likelihood by minimizing the negative of its log using unconstrained L-BFGS [7]. This is equivalent to maximizing the pseudo-likelihood and can be computed extremely efficiently without running an inference algorithm.

A pairwise feature is used to associate measurements from neighbouring sections. The measurements are modelled as a local feature. In such complex multimodal problems, instead of learning the CRF model directly from the raw observations (MWD data) it is advantageous to extract local features from the data using a classification algorithm [8].

2.1 Boosting local features

We applied boosting to provide the local feature; non-linearly mapping drill measurements to rock categories for each local borehole section. In Boosting, many weak learners are trained on various distributions of the input data and then combined to produce a single classifier committee [9]. We used a single node decision tree, also known as a decision stump, as weak learner [10]. For AdaBoost, a popular Boosting binary classifier, the multiclass problem was solved by using a one-against-all approach. In addition, we also implemented LogitBoost, which can handle multiple classes directly [11]. Three ways of combining boosting and CRFs were investigated:

- The first uses the multiclass output of LogitBoost as continuous features in a CRF;
- In the second, AdaBoost's weak learners for all classes are used as features. The weak learners' weight vector α is reset and the CRF learns corresponding weights instead;
- The third uses AdaBoost's binary outputs for each class as features.

3 Experimental results

The effectiveness of the proposed method is evaluated using MWD data collected from a blasthole drill rig at an open pit mine located in Western Australia. In this study, a total of 12 drill sensor measurements were recorded for analysis: bit air pressure, pull-down pressure, rotation pressure, pull-down rate, head speed, feed down pressure, feed up pressure, reverse rotation, forward rotation, rotation relief pressure, feed relief pressure, and hold back pressure. The dataset consists of 28 boreholes 12 m deep drilled in a straight line approximately 3 m apart. Since each sensor has different sampling rate, the measurements need to be re-sampled according to time stamp and then grouped into appropriate sections, 10 cm intervals in this study.

After drilling, the boreholes were tested using geophysical sensors: calliper, natural gamma, magnetic susceptibility and density (gamma gamma) logging tools. The detailed geology was determined by site geologists using a combination of geophysical, chip and core logs. Determining the detailed geology in terms of lithology, mineralogy and rock strength is a complex task and requires interpretation of the available data. This non-objective process creates uncertainty in the labels. The geology of the target area can be categorised in four classes: two waste rocks, banded iron formation (BIF) and shale, and two iron ore zones, Zone A and Zone B. Each borehole section was labelled accordingly. The labelled data set was used as reference to learn the CRF models.

The performance of the CRF models was evaluated by calculating accuracy, precision and recall using k-fold cross-validation. Instead of taking a random sample, in each round of training all sections from one of the boreholes were left out for validation while a model was learned using the remaining data. This cross-validation approach was devised to simulate the real-world scenario of trying to predict the geology of newly drilled boreholes, while possessing a model learned from previous holes.

A quantitative analysis of the algorithms' performance is presented in Table 1.

Table 1: Performance results of the proposed methods for 28 boreholes classified into 4 categories using cross-validation

		BIF	Zone A	Zone B	Shale	Overall[a]
LogitBoost	Accuracy	0.9322	0.8746	0.8900	0.8250	0.7837
	Precision	0.8362	0.7328	0.7304	0.8207	0.7800
	Recall	0.8981	0.8464	0.7174	0.7117	0.7934
CRF+LogitBoost	Accuracy	0.9457	0.8912	0.8915	0.8203	0.7935
	Precision	0.8772	0.7567	0.7309	0.8048	0.7924
	Recall	0.9043	0.8867	0.7191	0.7075	0.8044
CRF+AdaBoost[b]	Accuracy	0.9630	0.8641	0.9118	0.8265	0.7992
	Precision	0.9436	0.7265	0.8058	0.7714	0.8118
	Recall	0.9043	0.7747	0.7424	0.7844	0.8015
CRF+AdaBoost[c]	Accuracy	0.9521	0.9118	0.9073	0.8501	0.8246
	Precision	0.8923	0.8101	0.7454	0.8383	0.8215
	Recall	0.9074	0.8776	0.7961	0.7582	0.8348

[a] Overall precision and recall is the mean of all classes. Otherwise, overall accuracy is the ratio of the number of correctly classified sections over the total number of samples.
[b] CRF with AdaBoost decision stumps.
[c] CRF with AdaBoost classes.

The number of weak learners of the boosting algorithms was determined experimentally. Nevertheless, the boosting algorithm is quite resilient to overfitting and we observed that using more weak learners does not degrade performance severely. In the case of LogitBoost, the inputs are multiclass and the CRF has to simply associate neighbouring class labels. In the case of AdaBoost, the CRF also has to associate between the several binary classes provided by AdaBoost, or learn the decision stumps' thresholds. Clearly, the mapping to solve the multiclass problem performed by the CRF is different than the multiclass model provided by LogitBoost. The CRF may be more flexible and thus able to learn a more accurate multi-classification model based on the data.

4 Conclusions

All algorithms tested were able to learn a model from the available training area and generalise the results to the whole data set. The CRF approach presents a "smoothing" effect which correlates better with the expected geology of the area. While all CRF methods were superior to LogitBoost alone, the "CRF with AdaBoost classes" outperformed the other methods by more than 2% overall. Relearning the decision stumps' weights within the CRF did not improve performance.

This work addresses the problem of modelling measurements logged while drilling from multiple sensors while incorporating spatial information into the rock classification. A CRF framework was investigated to estimate geology using drill sensor measurements. The CRFs take the neighbourhood dependencies into account and improve the classification accuracy of a base boosting classifier. Modelling spatial relationships is useful to exploit the fact that local lithology can be highly homogeneous locally.

In this study, reference data for parameter learning was obtained by geophysical logging analysis provided by experienced geologists. Given the complex analysis required to provide ground-truth labels, a possible future direction is to investigate how to make use of the large volume of unlabeled or dubiously labelled data to improve the models.

Acknowledgments

This work was supported by the Rio Tinto Centre for Mine Automation and the ARC Centres of Excellence programme funded by the Australian Research Council (ARC) and the New South Wales State Government.

References

[1] Teale, R. (1965) The concept of specific energy in rock drilling. *International Journal of Rock Mechanics and Mining Sciences*, **2**, 57–73.

[2] King, R. L., Hicks, M. A., and Signer, S. P. (1993) Using unsupervised learning for feature detection in a coal mine roof. *Engineering Applications of Artificial Intelligence*, **6**, 565–573.

[3] Utt, W. K. (1999) Neural network technology for strata strength characterization. *Proc. International Joint Conference on Neural Networks*, vol. 6, pp. 3806–3809.

[4] Lafferty, J., McCallum, A., and Pereira, F. (2001) Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Proc. of the 18th International Conference on Machine Learning*, pp. 282–289.

[5] Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc.

[6] Kumar, S. and Hebert, M. (2003) Discriminative random fields: A discriminative framework for contextual interaction in classification. *Proc. of the 9th IEEE International Conference on Computer Vision*, vol. 2, pp. 1150–1157.

[7] Liu, D. and Nocedal, J. (1989) On the limited memory BFGS method for large scale optimization. *Math. Programming*, **45**, 503–528.

[8] Friedman, S., Fox, D., and Pasula, H. (2007) Voronoi random fields: Extracting the topological structure of indoor environments via place labeling. *Proc. of the International Joint Conference on Artificial Intelligence*, pp. 2109–2114.

[9] Freund, Y. and Schapire, R. E. (1996) Experiments with a new boosting algorithm. *Proc. of the 13th International Conference on Machine Learning*, pp. 148–156.

[10] Viola, P. and Jones, M. (2004) Robust real-time object detection. *International Journal of Computer Vision*, **57**, 137–154.

[11] Friedman, J., Hastie, T., and Tibshirani, R. (2000) Additive logistic regression: A statistical view of boosting. *Annals of Statistics*, **28**, 337–407.