Determining root causes of drilling problems by combining cases and general knowledge

Samad Valipour Shokouhi[1], Agnar Aamodt[2] and Pål Skalle[1], Frode Sørmo[3]

[1] Department of Petroleum Technology (IPT)
[2] Department of Computer and Information Science (IDT)
Norwegian University of Science and Technology (NTNU)
NO-7491, Trondheim, Norway
[3] Verdande Technology AS
Stiklestadveien 1- Trondheim, Norway
valipour@ntnu.no, agnar.aamodt@idt.ntnu.no, pal.skalle@ntnu.no,
frode@verdandetechnology.com

Abstract. Oil well drilling is a complex process which frequently leads to operational problems. In order to deal with some of these problems, knowledge intensive case based reasoning (KiCBR) has clearly shown potential. An important problem in drilling is hole cleaning, in which a high number of observed parameters and other features are involved. This paper presents how to determine the root causes of poor hole cleaning episodes by means of KiCBR. The effect of general domain knowledge was demonstrated in a comparative study, in which improved results in terms of similarity assessment and explanation capability were achieved.

Keywords: Case-based, knowledge intensive, oil well drilling

1 Introduction

Drilling of oil wells is an expensive offshore operation, costing typically 200 000 US$ per day. Any loss of time caused by unwanted events is costly. During drilling all material drilled out need to be removed, i.e. transported to the surface, a process which is referred to as hole cleaning. Often some of the material remains in the well, and hole cleaning is still among the most important problems to deal with during drilling. It is also one of the most studied phenomena within the petroleum industry. Insufficient hole cleaning can in extreme cases lead to loss of the well or a part of it, i.e. stop of the drilling process and blocking of the hole. Due to the number of parameters influencing hole cleaning and the complex mechanisms involved, the phenomenon has not yet been fully understood [1].

Case-based reasoning (CBR) is an approach to problem solving and decision making where a new problem is solved by finding one or more similar previously solved problems, called cases, and re-using them in the new problem situation. Application-oriented research in the area of case based reasoning has moved mature research results into practical applications. Skalle et al [2] employed case based reasoning to improve efficiency of oil well drilling. Their focus was on lost circulation, which means that some of the drilling fluid that always fills the gap between the drill string and the well wall gets lost into fractures in the geological formation. They built fifty cases on the basis of information from one North Sea operator. A general domain model was used to match non-identical features that were related in the model. Mendes et al [3] presented an application of CBR in offshore well design. The result of that work was a formalization of the methodology for planning of an oil well in a case-based reasoning context. They used fuzzy set theory for the matching of index features. Popa et al [4] presented an application of CBR for planning and execution of well interventions, in order to improve the decision-making process. Abdollahi et al [5] explained the applicability of CBR for diagnosis of well integrity problems in order to reduce the risk of uncontrolled release of formation fluids into the well.

In the above systems general knowledge has been used in the case retrieval process, for feature matching. None of the systems, or other CBR applications in this domain, have taken advantage of general knowledge in order to help identify a problem solution. In the study presented here a model-based method has been implemented as a complementary tool in order to determine the root cause of a hole cleaning problem. In addition, parts of the model are also used to enhance matching quality. An experiment has been undertaken to study the effect of the causal model combined with cases, in comparison with cases only.

The rest of the paper is structured as follows: In chapter 2 we explain the hole cleaning problem in some more detail, related to the functionality of our system. Chapter 3 explains the case structure and similarity methods. In chapter 4 results from the study of the effect of the causal model is reported. The types of input to the reasoning system, and their relationships with causes of hole cleaning problems are described in chapter 5. The last chapter summarizes and concludes the paper.

2 The hole cleaning problem

A drilling process consists of many steps, of which the actual drilling into the geological formation and the continuous cleaning of the borehole are core subprocesses. Fig. 1 illustrates the process at an abstract level. The hole cleaning issues arise when the drilling direction moves from vertical to deviated and horizontal hole angles. Horizontal drilling is getting more and more common, due to the increasing distance from the rig to productive wells. ("All the easy wells are already drilled", as the phrase goes). Accumulation of solids at a given depth is a common source of pack off, which is a serious situation indicated by the building up of material inside the hole wall, with reduced hole diameter as a result.

Many studies have been carried out by other researchers related to the cleaning of deviated and horizontal holes [6], [7], [8], [9], [10], [11]. However, the results of the studies have so far not provided clear operational recommendations. One reason may be that such studies are focused on the role and effect of individual parameters. A CBR approach, on the other hand, allows us to view a larger set of parameters as a unit, without assuming particular restrictions on the parameters, such as parameter independence.

Our application is targeted at reducing the risk of unwanted downtime (i.e. stopped drilling). The drill plan acts as guidance to expected drilling behavior. The real-time data from the drilling process is the main source of a situation description, which is matched with a past case in order to identify possible hole cleaning problems ahead of the drill bit. When a sufficiently similar past case is found, the root cause for that problem is presented to the user. In KiCBR this is supported by the causal model, linking input features to root causes.

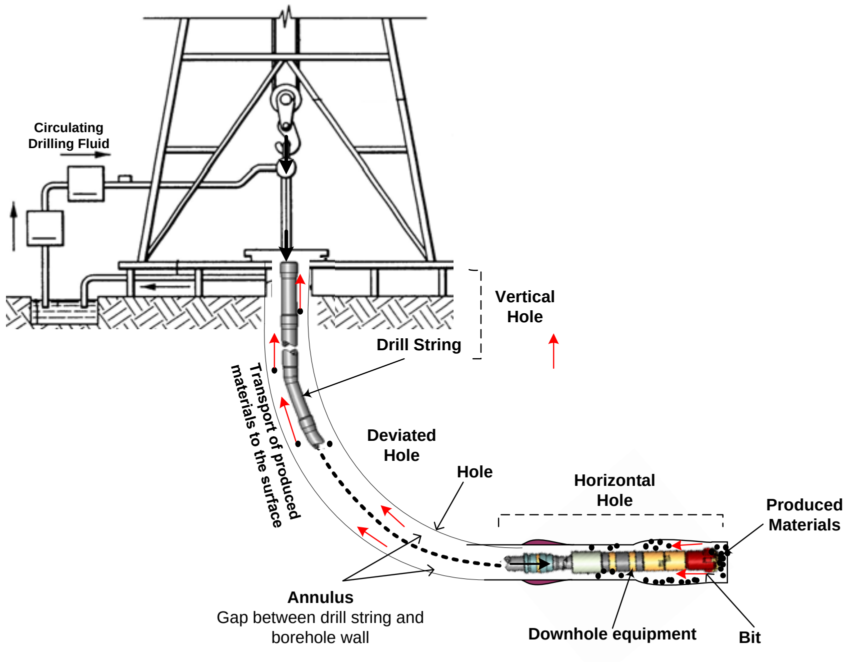

Fig. 1. Schematic drawing of an oil well being drilled.

3 Knowledge assessment

3.1 Case matching

The system is an architecture for knowledge intensive case based problem solving. It is designed for finding the root cause of a hole cleaning problem based on either the case base or the general knowledge module alone – and in combination. To build the system, three knowledge models are needed:

- *A taxonomy:* *extracting important terms from the domain.*
- *A causal model*: building a model that describes causes and effects.
- *A set of cases*: concrete past problem solving experiences.

A case's features consist of administrative data, wellbore formation characteristics, plan data, static and variable drilling data, the drilling activity performed before case occurrence, response action and conclusion. The case structure is illustrated in Fig. 2.

The CBR cycle consists of four steps; retrieve, reuse, revise and retain. The retrieval task starts with a (partial) problem description, and ends when a best matching previous case has been found [12]. A similarity assessment process has been defined that can be run with or without the use of the causal model. The similarity method is an adaptation and extension of the Creek method [13]. Our method consists of two different similarity properties, one being direct or linear indexing, the other being concept abstraction. The latter is used when the model based module is utilized.

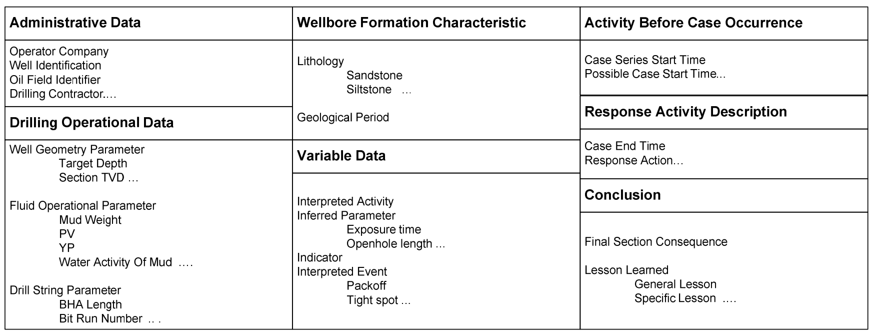

Fig. 2. Case structure

Basic similarity is computed by the following equation.

$$sim(C_{IN}, C_{RE}) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} sim(f_i, f_j) \times relevancefactor_{f_j}}{\sum_{j=1}^{m} relevancefactor_{f_j}} \tag{1}$$

C_{IN} and C_{RE} are the input and retrieved cases, n is the number of findings in C_{IN}, m is the number of findings in C_{RE}, f_i is the i^{th} finding in C_{IN}, f_j the j^{th} finding in C_{RE}, and $sim(f_1,f_2)$ is simply given as:

For symbolic concepts:

$$sim(f_1, f_2) = \begin{cases} 1 & if\ f_1 = f_2 \\ 0 & if\ f_1 \neq f_2 \end{cases} \tag{2}$$

For linear concepts:

$$sim(f_1, f_2) = 1 - \left|\frac{f_1 - f_2}{Max - Min}\right| \tag{3}$$

The relevance factor is a number that represents the weight of a feature for a stored case. The linear approach explicitly computes the values of similarity according to the minimum and maximum values of each concept. For example, minimum and maximum for true vertical depth has been set to zero and 8000 meter respectively. TVD for case 1 and 6 are 2869 and 2242 meter respectively, which provide 92 % similarity. Some of the indexing attributes will have both a symbolic and a linear description. An example of the categorization of numerical values is shown in Table 1. If a numerical value is available, linear similarity will be used and the symbolic terms will only be used in the model-based part.

Table 1. True Vertical Depth abstracted to symbolic entities.

True Vertical Depth (TVD)	*Very shallow Well*	<1000 meter
	Shallow Well	1000-2000 meter
	Medium Deep Well	2000-3000 meter
	Deep Well	3000-4000 meter
	Very Deep Well	>4000 meter

3.2 Root causes assessment

The main objective is to determine the root cause starting out from three types of features: Direct *observations* – i.e. measurements, *inferred parameters* – i.e. values derived from observations, and interpreted *events* – i.e. particular concepts describing important states which require particular awareness or action. The features and causes are related through intermediate state concepts, see Fig. 3.

The model used is a semantic net-based model of entities linked by relations. Each relation is labeled. The root causes and the case features are all represented as entities in this model, and the model-based reasoner works by finding paths from the entities representing case findings to the entities representing root causes. Fig. 8 shows an example of two such paths.

The goal of the model-based reasoner is to determine which root causes or intermediate states are entailed or likely provided the features. Only some paths provide support for such a conclusion. In order to determine legal paths, plausible inheritance was used. This method is a generalization of normal subclass inheritance that allows inheritance of relationships over other relation types than ‘subclass of’ relations. Plausible inheritance is governed by a set of rules declaring which relation-types can be inherited over which relation-types. In this paper, causal relationships are transitive, and any relationship can be inherited over ‘subclass of ‘ relationships. For more information, see [13].

Assume there is a legal path from a finding observations related to the root cause (see Fig. 3). Its strength is the product of the strength of each relation leading from finding to the target entity [14].

$$Path\ strength = \prod_{i=1}^{n} relation\ strength_i \qquad (4)$$

where n is the number of serial relations. Sometime there is more than one explanatory path from different finding to each target entity (the root cause entity). The total explanation strength for each target entity is determined with Eq. (5). This calculated explanation strength will be a good indicator of being the possible root cause.

$$Explanation\ strength = 1 - \prod_{i=1}^{m} (1 - path\ strength_i \times weight) \qquad (5)$$

where m is the number of paths. The strength of the indicating entities was decided based on a survey among five experts to reduce subjectiveness of these values. Weight of each indicating group observation parameter, inferred parameters and events are fourth, half and one respectively.

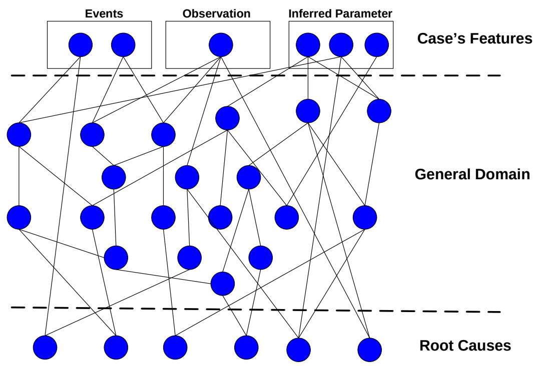

Fig. 3. Schematic model of the causal knowledge.

Fig. 3 shows three different clusters, namely *events* e.g. 'Pack Off'; *inferred parameters* e.g. 'Open Hole Exposure Time' (OHET); and *observation* e.g. 'True Vertical Depth' and 'Mud Weight' (density of the drilling fluid). The importance of each cluster on hole cleaning evaluation are 0.25, 0.5 and 1 for *observation*, *inferred parameters* and *events* respectively.

Observation factors include well plan data (drilling fluid and drill string parameters, well geometry), formation characteristic and case occurrence description.

In this section 'Pack Off' (of the event cluster) and 'Open Hole Exposure Time' (of the inferred parameter cluster) are exemplified.

Fig. 4 shows a 'Pack Off' event interpreted from real time data. Observed data collected from sensors, like flow rate and stand pipe pressure, cannot explain the situation alone. They are more useful for case classification and for finding the root cause when combined. In the Explanations (right part of the figure), 'Flow rate' is the pump rate of drilling fluid for transportation of produced material from the bottom of the hole to the surface. 'Stand Pipe Pressure' is the pressure measured at the surface which may increase due to any obstacle inside the hole. Increasing of the 'Stand Pipe Pressure' will indicate a 'Pack Off' situation while the variables such as 'Flow Rate' are constant.

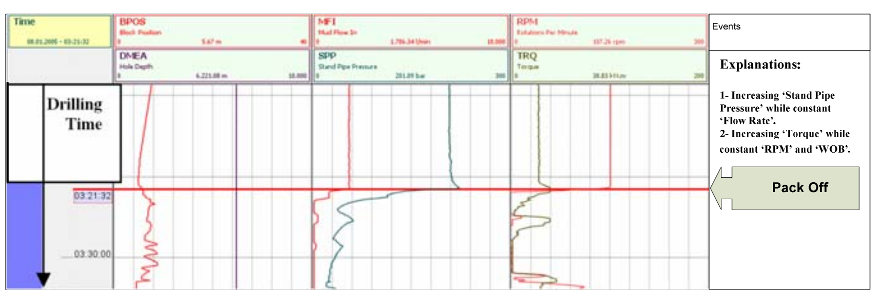

Fig. 4. 'Pack Off' recognition from observed data

'Open Hole Exposure Time' (OHET) is one of the inferred parameters in this study. OHET is the time period when the formation is in contact with drilling fluid, which again may cause a problematic situation during the drilling operation. Higher exposure time can contribute to higher problems. This time is being updated for desired points as the position of the drill bit changes. Desired points, i.e. the points where the cases were tagged depicted by drilling time and drilling depth are shown in Fig. 5.

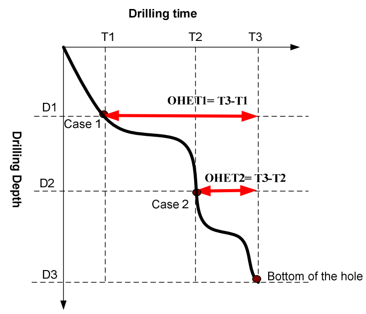

Fig. 5. Computation of the 'Open Hole Exposure Time' (OHET) for case 1 and case 2 when bit has reached D3.

4 Case matching results

The case base contains cases related to poor hole cleaning problems experienced in North Sea wells. To simplify the discussion about quality and applicability of the KiCBR in solving hole cleaning problems, seven cases are presented in this section. As mentioned, a symbolic and linear similarity framework was utilized. The case matching results for the case based module alone (CBR), the model-based approach alone (Model-Based) and for the integrated model- and case-based reasoning (KiCBR) will be presented. To evaluate the methods, a standard cross-validation technique is used, taking one case at a time out the case base and matching against the 6 remaining cases. Fig. 6 shows the case matching results for case 7 and case 5 as unsolved cases. For case 5, the retrieved case with highest similarity was case 2 with 18% similarity using the CBR method. When the KiCBR method was applied instead, case 3 was retrieved with 39% similarity.

In order to differentiate between the retrieved cases, they were grouped into three levels according to severity (how much drilling downtime they caused). The three levels of downtime are; insignificant, significant and highly significant repair time.

For instance, evaluation of downtime for case 7 revealed that this case required highly significant repair time while cleaning the hole. However, the CBR method retrieved case 2, which had insignificant repair time. On the other hand, the KiCBR method retrieved case 3 which is more similar than case 7.

In another example, the case matching process was run for case 5 as an unsolved case. When using the CBR matching method, case 2 was retrieved, while case 6 was retrieved using KiCBR. Case 2 and case 6 are grouped in the same class in terms of the downtime during the drilling, but detail study showed that case 6 had significant downtime later in the operation around the same area, and this is similar to the situation in case 5. This means that case 6 is more similar than case 2.

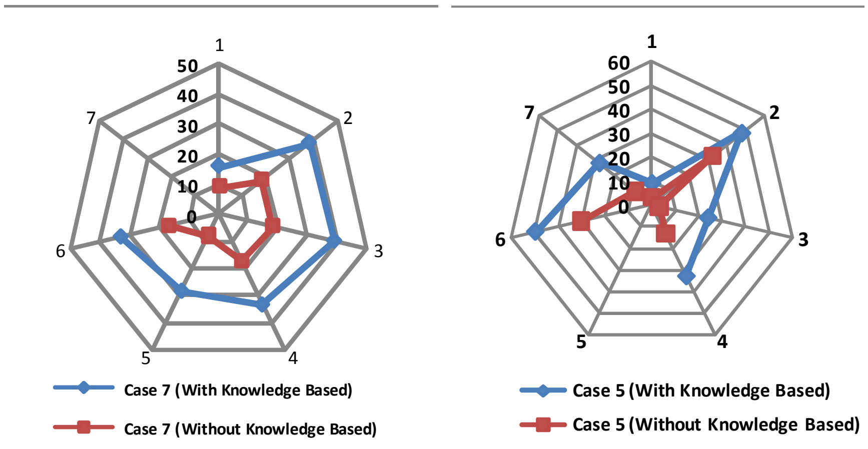

Fig. 6. Case matching results (in %) for case based module alone and combined case based and model based module for case 2 (left) and case 5 (right), matched against the remaining 6 cases. Lines between points are only used for better illustration.

The results show not only improvement for similarity assessment but also good prediction in problem solving. The effect of including general knowledge was monitored by changing not only the similarity but also the retrieved cases.

Similarity assessments are summarized in Fig. 7. The similarity growth was fluctuating from about 20 % to 100 % or even higher. As shown in Fig. 7, the most similar case by means of case-based, model-based and KiCBR are:

Unsolved case	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Retrieved by Case-based	Case 3	**Case 6**	Case 4	**Case 3**	Case 2	**Case 2**	Case 2
Retrieved by Model-based	**Case 7**	Case 7	**Case 7**	Case 5	Case 4	Case 4	**Case 3**
Retrieved by KiCBR	Case 3	**Case 6**	**Case 7**	Case 6	**Case 6**	**Case 2**	**Case 3**

Bold items in the above table represent the best case for each unsolved case according to downtime and detail studies. KiCBR was able to retrieve the optimal case in 5 out of 7 cases, while model-based and case-based retrieved only 3 optimal cases.

In summary, two important phenomena can be observed from the above tests. First, the general knowledge can generally increase the similarity for all cases in different rates. Second, general knowledge may also change which case obtains the highest similarity.

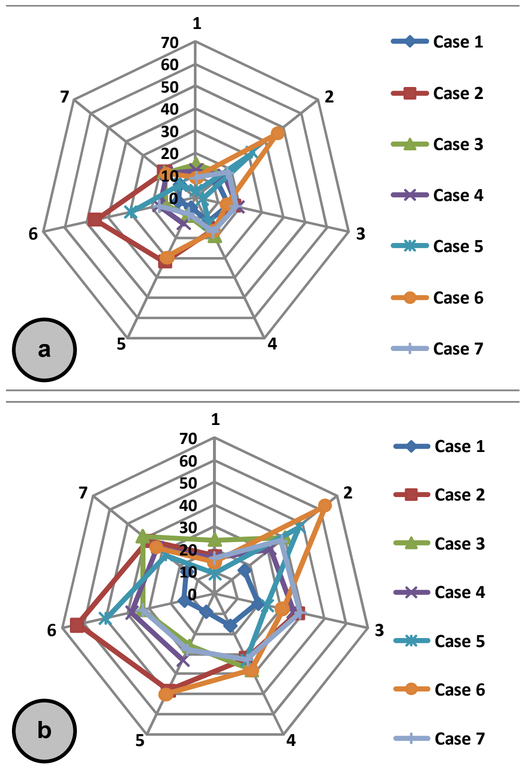

Fig. 7. Case matching using CBR without knowledge model *(a)* and with knowledge model *(b).*

5 Determining root causes of drilling problems

Many parameters are involved in the drilling process, and deviation of one factor may lead to hole cleaning issues and other problematic situations. Like in medicine, different diagnosis leads to different remedies, and as in medicine, finding the root cause of the problem from observable symptoms is a major challenge in drilling engineering.

The general domain knowledge serves as explanatory support for the case retrieval and reuse processes, through a model-based reasoning (MBR) method. In this study, the failure type/main root causes were divided into seven groups e.g. 'Hole Cleaning', 'Hole Collapse', 'Swelling', 'Erosion of Weakened Wellbore', 'Thick Filter Cake', 'Lost Circulation', and 'Dissolving'.

Fig. 8 illustrate some of the parameters involved in hole cleaning. In this figure two of the plausible inheritance paths were highlighted with solid and dotted lines.

First plausible inheritance path:

> 'High Mud Weight' implies 'High Mud Solids Content' causes occasionally 'Low ROP' causes 'Wear Of Shale' subclass of 'Erosion Of Weakened Wellbore'.

The starting entity 'High Mud Weight' is an observed parameter. The strength of this explanatory path is 0.11 calculated by Eq. (4).

Second plausible inheritance path:

> 'Open Hole Exposure Time' has subclass 'Long Exposure Time' causes sometimes 'Erosion Of Weakened Wellbore'.

This link starts out from 'Open Hole Exposure Time' which is an inferred parameter. The path strength for this explanatory path is 0.5.

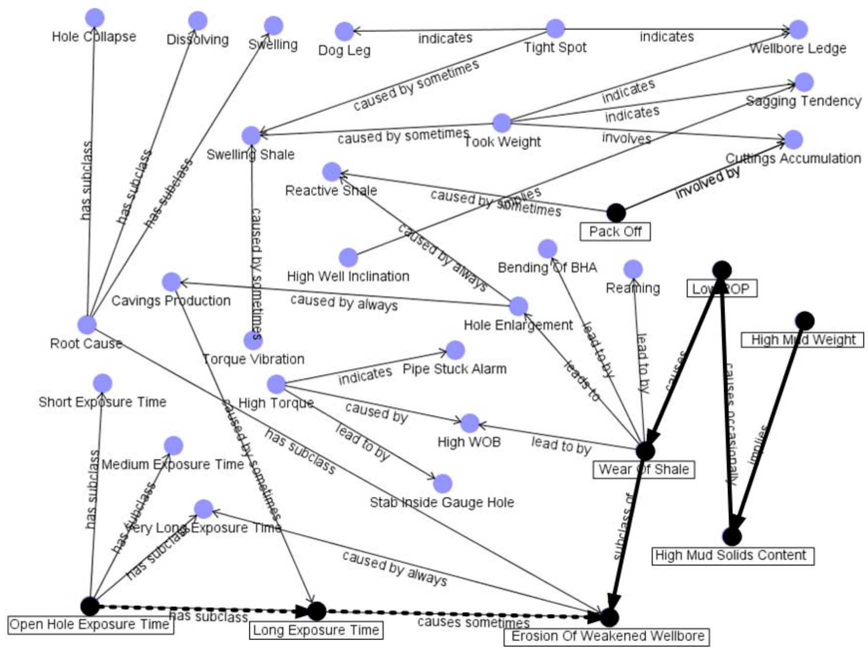

Fig. 8. Part of the hole cleaning concepts with their relations.

For each root cause, all plausible inheritance paths from each inferred or observed parameter in the cases is combined using Eq. (5), which determines the explanation strength. This calculation yields a number between 0 and 1 for each root cause, with a higher value indicating higher support for that root cause. Fig. 9 presents the value for each root cause for each of the seven cases.

Textual sources written during or after the drilling operation (Daily Drilling Report (DDR) and End of well report (EWR)) as well as real-time sensor logs showed us that six of the seven cases were highly representative of hole cleaning problems. As shown in Fig. 9 derived path strength of all seven cases points at poor hole cleaning except for case 1. In case 1, no events and inferred parameters took place. Therefore, explanation strength is based on just observed parameters, which results in a fairly low value of the path strength.

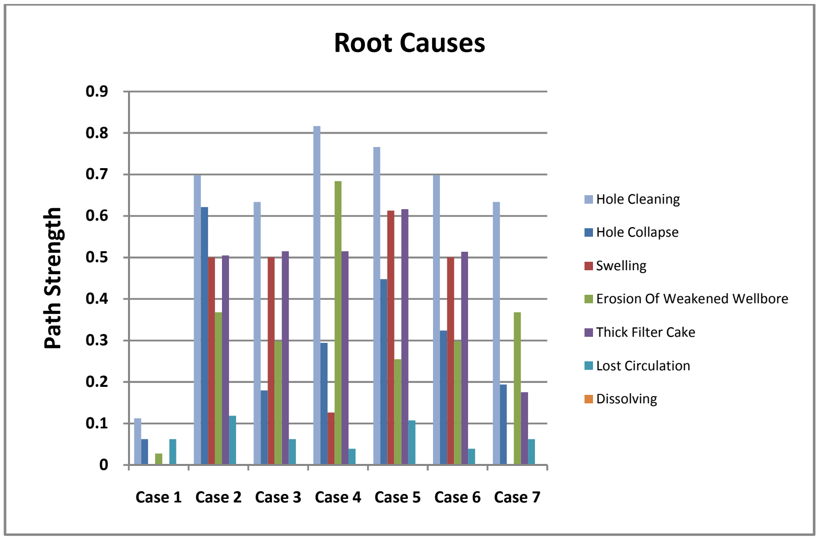

Fig. 9. Path strength of 7 cases based on general model to determine level of the hole cleaning problem

Once the root cause is found, it can be treated by applying a repair action. Each problem needs to be treated differently. A preliminary assessment of well data was performed to determine the specific root cause. In figure 10, the results for two cases (case 2 and 4) are shown. The plausible inheritance model provides strongest support for the 'Hole Collapse' and 'Erosion Of Weakened Wellbore' to be the specific root causes of poor hole cleaning for case number 2 and 4 respectively. Dissolving is zero for all the cases since there was not any salty rock in the studied holes. The presence of claystone (i.e. a type of rock) and about 26 days of 'Open Hole Exposure Time' caused the claystone to react with drilling fluid and the formation around the hole wall was eroded.

One of the main purposes of introducing knowledge based system is to advise the user of how to modify the controllable drilling parameters with respect to the associated root cause. Whenever the cause of a problem is revealed, the proper remedy can be applied. 'Hole Collapse' is one of the major causes of poor hole cleaning, mostly resolved by adjusting the density of the drilling fluid (mud).

Indications so far show that the KiCBR method may be better at retrieving the correct case, but even where this method is used, the explanation facilities of the model-based approach is valuable, as it allows the user to see what factors contribute to the problem by providing explanations. The model-based approach also calculates the support for different root causes independently, allowing it to conclude that multiple problems can be present. This is important as multiple problems requires multiple or complex remedies. For instance, for case 2 in Fig. 10 'Swelling' has high support, although 'Hole collapse' has even higher support. Chances are, both of these problems are present.

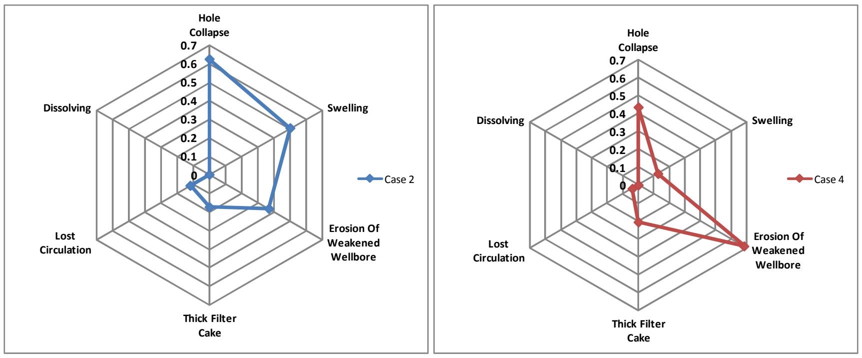

Fig. 10. Finding of root causes by means of knowledge model for case 2 (left) and case 4 (right).

6 Conclusion

The application of a relatively new methodology to reduce downtime during the oil well drilling has been considered. A combination of symbolic and linear similarity was utilized. Case similarity was changed by combining case based and model based reasoning.

KiCBR obtained a higher similarity and accuracy than case based reasoning alone. Similarity between an unsolved case and cases in the case base increased in average by typically 50 % after introducing the knowledge module in the reasoning process.

The most probable root cause could be determined on basis of the knowledge model. The root cause determined with the model-based approach had a good correlation with the expert analysis from real-time sensor data.

Further work

The results point out that combing knowledge intensive with case based reasoning improved the case matching routine. Furthermore, knowledge model serves as explanatory support for finding root causes. But in this study few cases were available and the results have to be tested out with many cases. Our aim is to implement this platform on more cases and perform a broader and more detailed assessment of the methodology.

Acknowledgments

The authors acknowledge the data support provided by StatoilHydro during the research work. The authors would also like to express appreciation to the staffs of Verdande Technology As for their help and cooperation in this research.

References

1. Datta, B.K., Ratnayake, C., Saasen, A., Omland, T.H., Hole Cleaning and Pressure-Loss Prediction From a Bulk Transport Perspective. Paper SPE 96315 presented at Offshore Europe, 6-9 September 2005, Aberdeen, United Kingdom, 2005.
2. Skalle, P., Sveen, J., Aamodt, A., Improved Efficiency of Oil Well Drilling through Case-Based Reasoning. Proceedings of PRICAI 2000,the Sixth Pacific Rim International Conference on Artificial Intelligence. Melbourne,, August- September 2000.
3. Mendes, J.R.P., Morooka, C. K., Guilherme, I. R., Case-based reasoning in offshore well design. Journal of Petroleum Science and Engineering, 2003. 40: p. 47- 60.
4. Popa, A., Popa, C., Malamma, M., Hicks, J., Case-based reasoningapproach for well failluer diagnostics and planning. SPE paper 114 229, presented at the 2008 SPE Western Regiounal and Pacific Section AAPG Joint Mtg, Bakersfield, March 31 - April 2, 2008.
5. Abdollahi, J., Carlsen, I.M., Randhol, P., Tenold, E., Haga, H.B., Jakobsen, T., A Case-Based Approach to Understand the Complexity of Causal Connections Related to Well Integrity Problems. IADC/SPE 111129, Presented at the 2008 IADC/SPE Drilling Conference held in Orlando, Florida, U.S.A., 4–6 March 2008.
6. Yu, M., Takach, N. E., Nakamura, D. R., Shariff, M. M., An experimental study of hole cleaning under simulated downhole conditions. Paper SPE 109840 presented at SPE Annual Technical Conference and Exhibition, 11-14 November, Anaheim, California, 2007.
7. Lapierre, S., Courville, G., Song, J., Achieving Technical Limits: Expanded Application of Real-Time Pressure-While-Drilling Data Helps Optimize ROP and Hole Cleaning in Large-Diameter, Directional Intervals. Paper SPE 99142-MS presented at IADC/SPE Drilling Conference, 21-23 February, Miami, 2006.
8. Adari, R.B., Miska, S., Kuru, E., Bern, P., Selecting Drilling Fluid Properties and Flow Rates For Effective Hole Cleaning in High-Angle and Horizontal Wells. Paper SPE 63050-MS presented at SPE Annual Technical Conference and Exhibition, 1-4 October 2000, Dallas 2000.
9. Sanchez, R.A., Azar, J. J., Bassal, A. A., Martins, A.L., Effect of Drillpipe Rotation on Hole Cleaning During Directional-Well Drilling. SPE Journal, 1999. 4(2): p. 101-108.
10. Charlez, P.A., Rock Mechanics:petroleum Applications. 1997: Editions Technip. 661.
11. Peden, J.M., Ford, J.T., Oyeneyin, M.B., Comprehensive Experimental Investigation of Drilled Cuttings Transport in Inclined Wells Including the Effects of Rotation and Eccentricity. Paper SPE 20925-MS presented at European Petroleum Conference , 21-24 October 1990, The Hague, 1990.
12. Aamodt, A., Plaza, E., Case-Based Reasoning: Fundamental Issues, Methodological Variations, and System Approaches. Artificial Intelligence Communications, 1994. 7 (1): p. pp. 39-59.
13. Sørmo, F., Aamodt, A., Knowledge elaboration for improved CBR. In Sixteenth International Joint Conference on Artificial Intelligence, Workshop ML-5: Automating the Construction of Case-Based Reasoners. Stockholm, Sarabjot Singh Anand, Agnar Aamodt, David W. Aha (eds.), 1999: p. pp. 39-43.
14. Aamodt, A., Knowledge-Intensive Case-Based Reasoning in CREEK. Lecture Notes in Computer Science, 16 November 2004. 3155/2004: p. pp. 1-15.