Plan Recommendation for Well Engineering

Richard Thomson[1, 2], Susan Craw[1], Stewart Massie[1], Hatem Ahriz[1] and Ian Mills[2]

[1] IDEAS Research Institute,
Robert Gordon University
{s.craw, s.massie, h.ahriz}@.rgu.ac.uk

[2] Exceed Ltd,
12 Bon Accord Square, Aberdeen
{thomson.r, mills.i}@xcd.com

Abstract.
Good project planning provides the basis for successful offshore well drilling projects. In this domain planning occurs in two phases: an onshore phase develops a project plan; and an offshore phase implements the plan and tracks actual progress compared to the initial plan that was created onshore. The reuse of project plans is already prevalent within the oil and gas industry. A Performance Tracker tool has been built to support users in the reuse of project plans. It uses a case-based reasoning approach for plan recommendation that retrieves and reuses past well drilling project plans. Cases are composed of problem parts that store project initiation data, and solution parts that record the tasks and subtasks of actual plans. An initial evaluation shows that the nearest neighbour similarity based retrieval identifies relevant projects with similar project initiation data, and that the retrieved tasks and subtasks are relevant for the new project. The Performance Tracker can be viewed as a recommender system where the recommendations are plans based on past well drilling projects. Thus the data that is routinely captured as part of the performance tracking during offshore implementation is utilised as experiences that are retrieved and reused in similar well engineering contexts to propose project plans during onshore planning for new wells.

Keywords: recommendation, oil well drilling, nearest neighbour, project planning

1. Introduction

The oil and gas industry is one of the world's largest industries and is estimated to be worth $10400 trillion, based on current discovered oil reserves and the average price of oil. In the oil and gas industry wells are holes produced by a boring process for the purpose of finding and producing hydrocarbons. Wells have various categorisations and are constructed to either, obtain geological data prior to drilling (exploration well, test well and appraisal well), to research possible oil fields (wildcat well) or to extract the raw materials from the ground (oil well, gas well, production well, aquifer producers and gas injectors).

This paper focuses on the reuse of project plans for subsea drilling and as such, only offshore well drilling projects will be discussed. Offshore wells are constructed using rigs with various types of equipment used for drilling, casing the hole and extraction. The process of drilling a well can be split into 5 segments: planning where the tasks and subtasks required to construct the well are identified, boring the hole to reach the reservoir, preparing the hole for the extraction of the hydrocarbons by casing the hole with cement, extracting and refining the hydrocarbons, and lastly plugging the well when the reservoir is empty or the reservoir has stopped producing enough hydrocarbons to be seen as a viable use of resources.

Due to the contractual nature of recruitment within the oil industry, knowledge retention can be challenging and hiring individuals with the experience necessary is expensive, meaning the price of retaining corporate memory is high. The reuse of knowledge using a centralised system can help minimise this cost.

Section 2 will explain the process of planning and monitoring a subsea well construction project. Section 3 describes how the Performance Tracker will use the data produced during the planning process to identify similarities between projects. Section 4 describes the architecture of the Performance Tracker system. Section 5 evaluates the effectiveness and performance of the approach. Section 6 discusses any related work. Lastly, Section 7 reflects on the conclusions that were made and any planned future development.

2. Problem Domain

Wells share some characteristics with other wells and as a result follow the same drilling process. It is then possible to reuse project plans of related projects in order to cut the time and money allocated to the planning process. The planning process is split into two stages:

1. The project plan is developed by the onshore team after the discovery of a well. (Onshore Planning Stage)
2. The plan is implemented by the offshore team where data is monitored and recorded. (Offshore Planning Stage)

Onshore Planning Stage: At this point a new well has been identified and approval for drilling has been received. A new project plan is developed in a spreadsheet after which, the plan is scrutinised during a Drill Well on Paper exercise where the type of rig is chosen, the project is costed and potential causes of non-productive time are identified. The final project plan is then created, comprising the list of tasks and their associated subtasks, containing time targets. This plan is then ready to be used during the second phase of the process.

Offshore Planning Stage: The project plan created by the onshore team is now put into operation offshore. The project plan now changes its primary function from a planning tool to a monitoring tool. The offshore project manager will input the operational data which can then be compared to that of the planned data, created during the onshore planning phase. Where there is discrepancy between the planned task times and the actual task times, it is classified as either invisible lost time which states inefficiency within the well drilling operation, or non-productive time which is time spent rectifying unforeseen problems during the operation e.g. tool failures. The project plan is then refined taking into account any lost time encountered. Once the project has been completed, the project plan is used to evaluate the project before being retained by the company.

3. The Performance Tracker

During the onshore process described in section 2, it is common practice for the onshore team to employ the time consuming process of manually retrieving old project plans from within the company, and adapting these to suit the new project.

Currently it is up to the onshore team members to recall the old project plans based on their own past experiences which can be problematic as individuals may forget about a more suitable project or be unable to obtain the project plans due to it being stored on a local machine. Users may also have a poor or unclear perception of the project leading to an inadequate choice of project. Furthermore, an inexperienced project team that may lack past knowledge to effectively reuse project plans.

The Performance Tracker system addresses these issues by supporting the user during the onshore process by providing a CBR recommendation function for the retrieval and reuse of past project plans from a central project plan case base. This approach will use CBR to reduce the time and work required when selecting a suitable project plan and using a case base to provide a greater set of potential relevant project plans.

3.1 Case Representation

A project case *c* will be made up of a problem part *p* and a solution part *s*. *p* will contain the feature values of the project initiation data that will be used to identify the similarities between cases. The project initiation data stores the core project information required to begin the planning stage consisting of a project description, the geographical location of the well and the planned drilling depth.

p is made up of five base features: a textual *description* of the project; the *rig* that was chosen during the onshore process; the *well* that is going to be drilled; the estimated drilling *depth* in feet; and the estimated *duration* of the project in days. The *rig* and *well* are represented by a set of sub-features which when combined provides their overall problem representation, as shown in figure 1. Oil rigs are chosen due to their availability and appropriateness for a project type therefore, similar rigs will be used for similar projects. Wells of similar depths and locations tend to encounter the same geological features, requiring a similar sequence of tasks. This practice promotes the use of performance comparison between similar wells within an oil field where tasks times and issues encountered are compared.

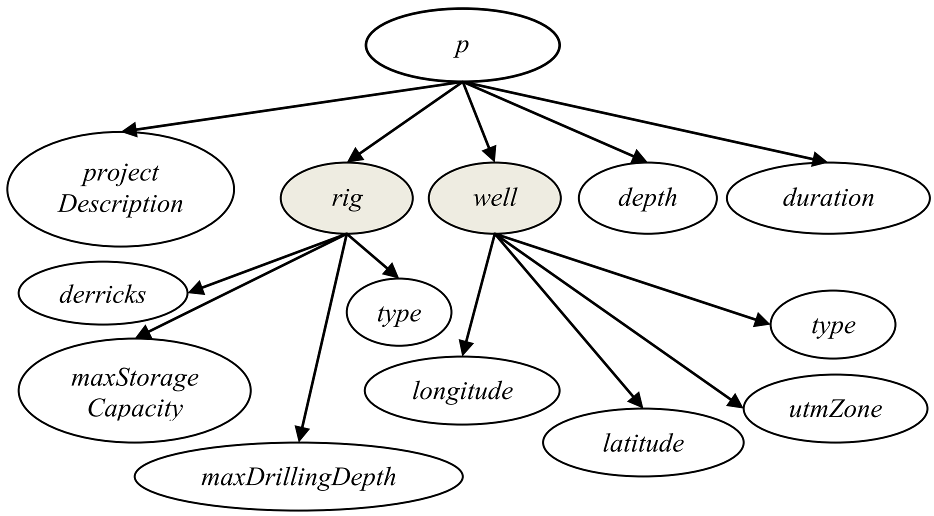

Figure 1. Project Problem Structure

The case solution *s* contains tasks and subtasks that were required to complete the project (figure 2). Each task is decomposed into a number of smaller subtasks that are used to plan and monitor a task in greater detail. The subtasks store the time data used to identify the lost time classifications which are required for project refinement. Subtasks will also include information of the resources that were used during the operation of a subtask as well as the risks that may potentially occur. A small sample of a case is shown in figure 3 where the left hand box contains the project problem features and the right hand contains a set of tasks making up the project solution.

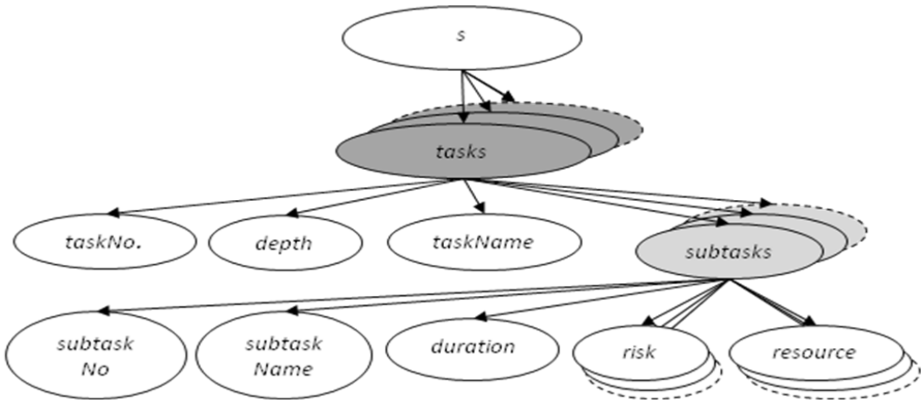

Figure 2 Project Solution Structure

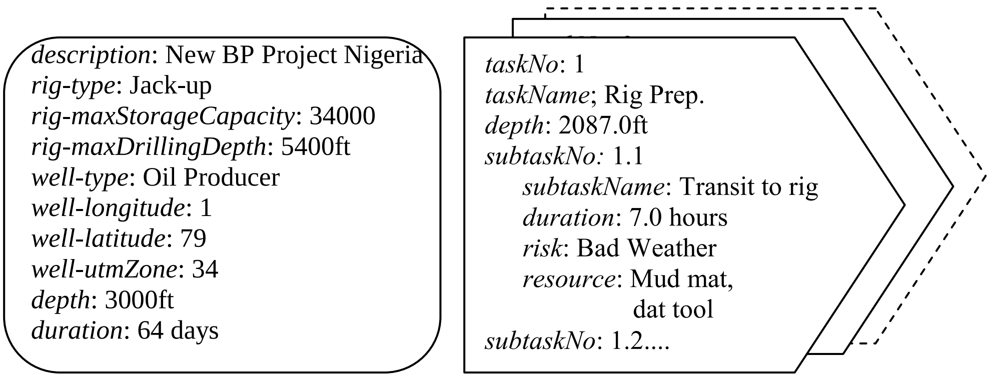

Figure 3 Example of Project Initiation *p* and solution *s*

3.2 Similarity

Cases are selected by using the Nearest Neighbour algorithm [1] in which the case similarity score is determined by using a weighted feature average and the closest *k* cases are then recommended to the user.

3.2.1 Local Similarity

The problem part of the case representation, shown in figure 1 consists of numerical (*depth*, *duration*, *longitude*, *latitude*), symbolic (*rigType*, *wellType*) and textual features (*projectDescription*). The process for calculating the similarity for each feature type will now be outlined.

Numerical Similarity: Similarity between numerical features will be derived from the Normalised Manhattan distance:

$$sim(qv, cv) = 1 - \frac{|cv - qv|}{r}$$

Where *qv* is the numerical feature value of the query, *cv* is the numerical feature value of the case and *r* is the predefined feature range; e.g. the latitude feature will have values ranging from -90° to 90°, hence range *r* will be 180.

Symbolic Similarity: An oil rig may be suitable for various project types therefore a Boolean similarity value of 1 for a match and 0 for no match, will not be suitable. There should be a varying degree of similarity so that partial suitability is taken into account. The same can be said for the type of well that will be constructed. For this reason individual rig and well similarity matrices have been created. **T**he *rigType* matrix in figure 4 was developed in conjunction with a domain expert and shows the similarities between different rig types. This matrix will serve as a look-up table to provide the similarity between *rigType* feature values. A similar process is applied to identifying the similarity between *wellType* values.

	Jackup	**Semi Sub**	**Platform**	**Drillship**
Jackup	1	0.8	0.7	0.6
Semi Sub.	0.8	1	0.9	0.8
Platform	0.7	0.9	1	0.5
Drillship	0.6	0.8	0.5	1

Figure 4 *rigType* Similarity Matrix

Text Similarity: The *projectDescription* gives an overview of who the project is for, and an explanation of what the project's aims are. This is also the space where the user can add any comments about the project or any special conditions, such as "test deep water project" or "HSE recovery". The completed *projectDescription* is treated as a bag of words, as the presence of a word is of great importance, whereas word position is unimportant. The Jaccard Coefficient assesses the overlap of words within the two word sets – the *projectDescriptionQuery* (*qv*) and the case feature *projectDescriptionCase* (*cv*) - and not the position in which they appear within the text. *projectDescription* similarity is defined as:

$$sim(qv, cv) = \frac{|qv \cap cv|}{|qv \cup cv|}$$

Due to the industry specific nature of terms used within documents and the more personalised approach to shorthand used (for example CC, circ and ccution are all used in place of circulation), the use of a generic lexicon such as WordNet would prove to be unsuitable for this type of synonymic analysis. For this reason a domain specific lexicon has been created to analyse similarity and meaning behind these industry specific terms.

3.2.2 Global Similarity

The Global Similarity *GSim* of a case is assessed by calculating the weighted average of the local feature similarities. The feature weights are set to bias the results towards "more important" features.

$$GSim(q,c) = \frac{\sum_{i=0}^{n} w_i \, sim_i(q,c)}{\sum_{i=0}^{n} w_i}$$

Where w_i is the weight of the ith feature and $sim_i(q,c)$ is the local similarity of the ith feature of the query q and of the case c.
The individual feature weights were set after consultation with Oil Field Managers. The feature weightings for *p* are shown in figure 5. The features of the *well* and the *rig* are given higher importance. The *depth* will help determine the suitability of a *rig* for the project as a *rig* will have a maximum depth that it can drill. Due to the presence of vague or short *projectDescription*, w^i for this feature has been set to the lowest value.

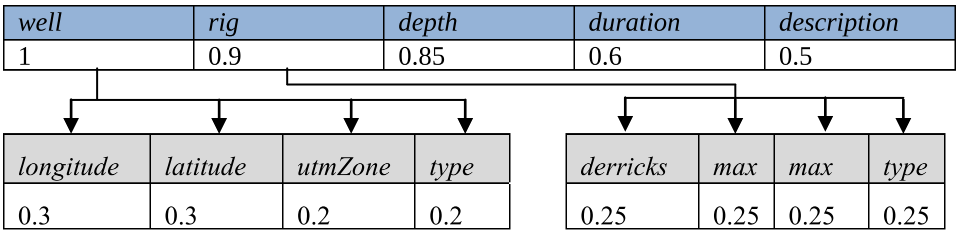

Figure 5 Feature Weightings

4 Architecture

The Performance Tracker shown in figure 6 consists of 2 main components, the onshore project plan recommendation tool and the offshore monitoring tool, both of which reside within a central server. The onshore project plan recommendation tool uses the JColibri2 framework in order to apply the modelling rules to the cases. JColibri2 proved to be a suitable framework because it contains modules for both pre-processing text and calculating the similarity of the three feature types used by the Performance Tracker.

New project initiation data is sent from the web portal to the Case Modeller where modelling rules are applied prior to the connection with the case base, creating a query from the initiation data that corresponds to the problem part of a new case. The Similarity Assessor applies the similarity metrics to identify the *k* most similar cases in the case base. The Solution Extraction Module can now take the solution part of the *k* cases and display these as recommendations in the Web Portal. Once the user has adapted the project plan of the selected case, it is stored as a new project plan within the database.

The Performance Tracker system is delivered as a web based tool and is accessible both onshore and offshore. This ensures that the project plans in the database can be refined during the offshore monitoring process where task times are constantly being adapted to correspond with the live project data. All projects are stored and updated in the database for persistent storage. The case base will be stored in server memory providing quick access for the onshore plan recommendation tool when a query is made. In order to ensure that both new projects created by the onshore process and refined projects generated by the offshore monitoring are consistent in both the case base and the database a weekly synchronisation process was developed, adding any newly created projects as new cases whilst updating current cases, as a result of the offshore monitoring. The synchronisation process reduces the number of queries made to the database during a retrieval resulting in a reduction in retrieval times. The project case base stored in memory will be used as the data source for case similarity matching and case retrieval.

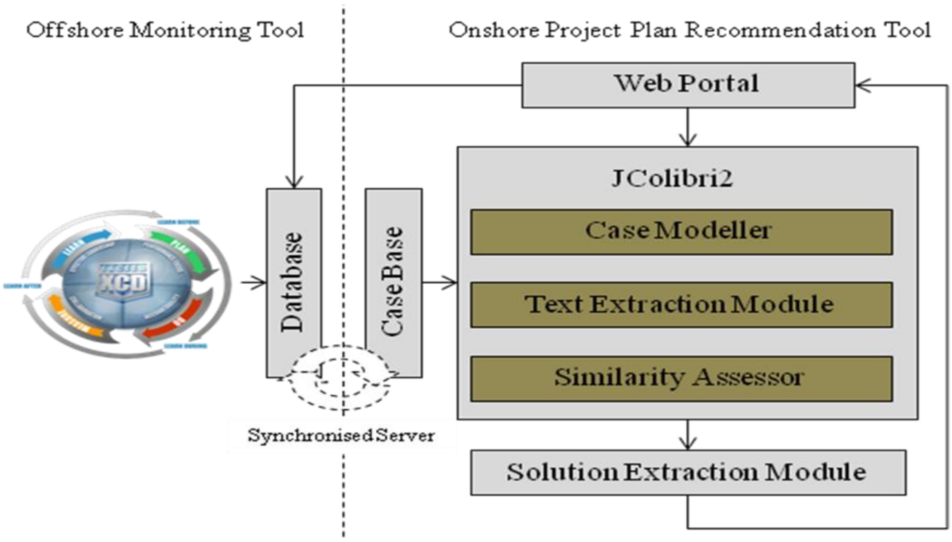

Figure 6 Performance Tracker Architecture

5 Evaluation

This section describes the results of three recommendation tasks that were made on the Performance Tracker system. Testing will be split into two parts, the first part will run three queries on the Synchronised version where the case base is built once and then stored in memory for future queries. After each run a comparison between the Project Initiation feature values and project plans of the most similar case and query will be made. Following the evaluation of similarity, there will be an analysis of the differences in retrieval time between the synchronised tracker version and the baseline system using the standard JColibri cycle where the case base is built for each query.

The case base was created using existing project data stored in Excel spreadsheets. The case base comprises of 200 cases for 200 wells, drilled by 87 rigs in 9 locations. The drilling depths range from 958 feet to 23,060 feet. Three cases are extracted from the cases base, providing the project initiation data for three queries in turn and leaving a case-base of 199 cases.

5.1 Retrieved Tasks

Query 1: The first query was based on an oil well construction project off the Nigerian coast. Figure 7 shows that the 5 most similar cases retrieved were all based on the same well type and had initiation data similar to the query. The right hand column shows the overlap between the tasks of each retrieved case and those of the extracted case. All of the project plans returned were very similar to those in the extracted case but would require adaptations to the depth and task times in order to be suitable.

Project Rank	Well Type	Initiation Data Similarity (%)	Project Plan Overlap (%)
1	Oil Well	83%	97%
2	Oil Well	81%	84%
3	Oil Well	81%	80%
4	Oil Well	78%	81%
5	Oil Well	74%	80%

Figure 7 Result from Query 1

The top ranked case contained an almost identical project plan to the extracted case with only one difference. The base plan for both cases was the same besides, a newly created task that addressed a lost time issue, which can be seen in figure 8. The added task accounted for the 3% reduction in the overlap percentage. This highlights the potential for the possible adaptation of project plans based on lost time.

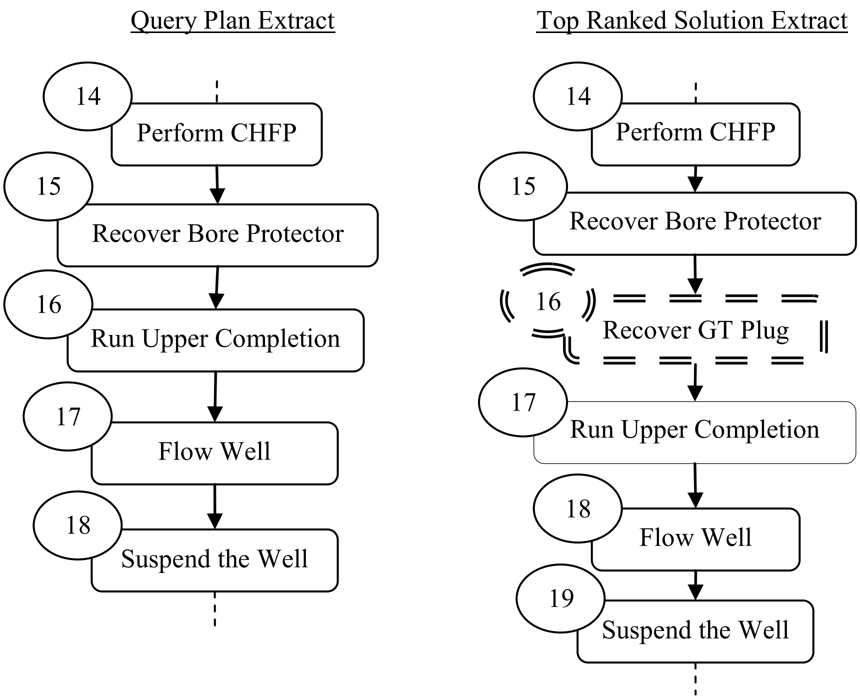

Figure 8 Query/Solution Overlap

Query 2: The second query contained initiation data for an oil well project with no *projectDescription* and no *well – type*. The similarity of the cases in the case base was low mainly because the project description and *wellType* were missing from the query. The similarity of the retrieved plans fluctuated greatly. It is evident that the type of *well* being constructed impacts the similarity between the retrieved project plans. Although the 2nd most similar retrieved case had a lower Project Initiation similarity value, the project problem was very similar in respect of *depth*, *well* and the *duration*. The low ranking came due to the dissimilarity between the *rig* and the *projectDescription*. When a *projectDescription* was added to the problem, the second ranked case was promoted as the most similar case.

Project Rank	Well Type	Initiation Data Similarity (%)	Project Plan Overlap (%)
1	Oil Well	71%	79%
2	Oil Well	63%	81%
3	Wildcat	63%	44%
4	Gas Well	62%	77%
5	Test Well	53%	39%

Figure 9 Result for Query 2

Query 3: The third query used initiation data for the JADA oil field test well, containing a sizable *projectDescription* including the phrases “test well”, “JADA field” and “oil and gas producing well” which provides a larger vocabulary for evaluating similarity within the text matching module.

Figure 10 again illustrates that the type of well being constructed has a large impact on the project plan overlap. The top ranked case used the same rig to drill an oil well that was closely located to the query which resulted in a high similarity of project initiation data. However, tasks only overlapped during the rig set-up and abandonment stages. The cases ranked as the 2nd and 5th most similar had the highest level of project overlap with the *duration* and *wellType* being the main reasons for this. As the 2nd highest ranking case contained the project plan for the test well of an oil field adjacent to the JADA field it is not unsuprising that this case higher similarity value than the top ranked case. There are two reasons for the rankings being as they are, the first is the inclusion of “JADA field” within the *projectDescription*, and the second was the differences in the depth of the hole being drilled. It may be prudent to allocate a higher weighting to the *wellType* feature for future versions of the tool.

Project Rank	Well Type	Initiation Data Similarity (%)	Project Plan Overlap (%)
1	Oil Well	87%	34%
2	Test Well	85%	94%
3	Oil Well	81%	34%
4	Oil Well	80%	37%
5	Test Well	74%	86%

Figure 10 Result for Query 3

5.2 Retrieval Speed

It was previously mentioned that the synchronised version of the tracker eliminates the case base building step from the base system in order to improve the run time of a query. These tests were used to identify whether there was any significant difference between the retrieval times of each system as a result of pre-building the case base. For this, twenty identical queries were each run on the baseline system and the synchronised system, using a low bandwidth to replicate the offshore network connection. The baseline system had an average retrieval time of 1 minute 34 seconds with times ranging between 1 minute 27 seconds and 1 minute 45 seconds. The synchronised system had an average retrieval time of 5.7 seconds. The fastest retrieval time was 2.15 seconds and the slowest retrieval time was 1 minute 34 seconds. The slowest retrieval time was the first retrieval as the case base had to initially be built from the database. Subsequent retrievals did not require the case base to be rebuilt and due to the process of synchronisation being a background process, the retrieval times were kept low. The difference between retrieval times is likely to grow as the case base increases in size.

6 Related Work

Case-based planning is not a new concept with applications ranging from holiday planners to planning for logistics as used in CaPER [2, 3]. Conventional CBR planners construct a project by adding project elements step-by-step whilst being in consultation with the user. The Performance Tracker does not operate in the same way as many of these Case-based planners, as entire sets of project plans are recommended based on the similarity of the problem. The Performance Tracker has more in common with CBR recommender systems and applies the single shot, proposal type of recommendation as described by Smyth [4]. This approach uses a specific user problem and based on the user criteria a set of cases are returned.

In spite of the relatively slow uptake of CBR and recommender systems within the oil industry there are still a number of systems which have now been developed to support the well drilling life cycle. CBR solves problems based on past solutions which would make the identification of reasons for lost time in running projects based on past projects a logical choice of area of application. Skalle *et al* [5] identified the usefulness of CBR when reducing lost time by analysing one problem area, stuck drill strings. Although the research focused on one area, and primarily on research, it is very important since that $250 million per annum was wasted on this form of downtime alone, highlighting the potential for using past experiences in order to solve a stuck drill string problem quickly.

Drill Edge uses CBR to identify possible reasons for a lost time problem during the offshore monitoring stage of a project, and then advises users on how the project could be refined to solve the problem. Drill Edge builds on the studies carried out for Creek and TrollCreek knowledge intensive CBR frameworks [6]. TrollCreek was developed to identify lost time based on data from the "Drilling Club" [7, 8]. This research proved that a CBR methodology can be used effectively for reusing past experiences within a drilling environment, particularly within the context of lost time reduction.

The main difference between the Performance Tracker and Drill Edge is nature of the systems. Drill Edge is a proactive system that uses real time well performance data to identify problem cases with similar feature values, with a view to preventing the reoccurrence of lost time [9]. Experts are only consulted during the reuse phase of the R^4 cycle (retrieve, revise, reuse, retain) [10]. The Performance Tracker is a reactive system that performs a retrieval based on data inputted by the expert, making the system less obtrusive to the existing drilling process. This late position of expert involvement in the Drill Edge recommendation process may lead to the unwanted flagging of events or cause unneeded lost time by investigating a flagged lost time event.

7 Conclusion

We have presented an approach to apply similarity matching to offshore well drilling project data to effectively recommend past project plans. We have shown that by choosing appropriate project initiation features and feature weightings it is possible to retrieve a set of suitable project plans that can be manually adapted to for a new well construction project. Furthermore, the process of project plan refinement during the offshore monitoring stage has demonstrated that the project plan is constantly being reused and becoming a more robust solution, throughout the life of a project.

We have also shown that the retrieval time required for a recommendation can be reduced by storing the case base within memory which is updated through a regular synchronisation process.

Drawbacks came when a user was required to manually adapt a small number of task depths of retrieved project plans. To limit the amount of adaptation required, the addition of a sea bed depth feature and a true vertical depth feature is being developed to provide a broader analysis of depth similarity. It was also shown that the type of well being drilled had a large impact in the suitability of a project plan. The next version of the Performance Tracker will include a CBR function to offer solutions to users during project monitoring based on past Lessons Learned.

References

1 T. Cover, P. Hart: Nearest Neighbor Pattern Classification, Institute of Electrical and Electronics Engineers Transactions on Information Theory, Volume 13, Page 21-27 (1967)

2 S. Stewart, C. Vogt.: A Case-Based Approach to Understanding Vacation Planning. Leisure Sciences, Volume 21, Issue 2, Page 79-95 (1999)

3 High Performance Case-Based Planning, http://www.cs.umd.edu/projects/plus/Caper/

4 B. Smyth.: Case-Based Recommendation. The Adaptive Web, LNCS 4321, Page 342 – 376, Springer (2007)

5 P.Skalle, A. Aamost, J. Sveen.: Case-Based Reasoning, a method for gaining experience and giving advice on how to avoid and how to free stuck drill strings. IADC Middle East Drilling Conference, Dubai, (1998).

6 A. Aamodt.: *Knowledge*-Intensive Case-Based Reasoning in CREEK. 6th International Conference on Case-Based Reasoning, Workshop Proceedings. DePaul University, page 62-71 (2005)

7 A. M. Islam, P. Skalle.: Review of wellbore instability cases in drilling through case base resoning (CBR) method. Proceedings of the International Conference on Mechanical Engineering, (2007).

8 P. Skalle, A. Aamodt: Knowledge-Based decision support in oil well drilling. Intelligent Information Processing II IFIP International Federation for Information Processing, Volume 163, Page 443-455, (2005).

9 Verdande: Harvesting knowledge to improve drilling, performance in real time, A perspective on lowering NPT in difficult and critical wells (http://www.verdandetechnology.com/images/stories/verdande/publications/vt_whitepaper.pdf) (2005).

10 A. Aamodt and E.Plaza, Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, AI Communications. IOS Press, Volume 7:1, Page 39-59, (1994)