

УДК 622.248.5:004.85

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ «БЕЗ УЧИТЕЛЯ» ДЛЯ ПРЕДУПРЕЖДЕНИЯ ПРИХВАТОВ БУРИЛЬНОЙ И ОБСАДНОЙ КОЛОНН

Щербаков Роман Эдуардович^{1,2},
res718@bk.ru

Ковалев Артем Владимирович¹,
Kovalevav@tpu.ru

¹ Национальный исследовательский Томский политехнический университет,
Россия, 634050, г. Томск, пр. Ленина, 30.

² АО «Газпромнефть НТЦ»,
Россия, 190000, г. Санкт-Петербург, наб. реки Мойки, 75–79, литер Д.

Актуальность. С каждым годом истощение легкодоступных запасов углеводородов определяет необходимость разработки месторождений, характеризующихся сложными горно-геологическими условиями. Строительство скважин в данных условиях часто сопровождается различными осложнениями и авариями. Повышенные риски их возникновения обусловлены усложнением конструкции скважины, траектории ствола, а также недостаточно достоверными данными о горно-геологических условиях бурения. Прихват бурильной или обсадной колонны является одной из самых сложных аварий, которая оказывает существенное влияние на эффективность бурения и стоимость строительства скважины. Зачастую при возникновении прихвата требуется проведение дорогостоящих мероприятий по его ликвидации. В настоящее время признаки прихвата определяются по параметрам бурения инженерным составом, однако ввиду особенностей человеческого восприятия и сменного режима работы это происходит несвоевременно. Недостатки используемого в настоящий момент подхода предопределили использование алгоритмов машинного обучения. Интеллектуальная система может автоматически анализировать тенденции изменения параметров бурения, обнаруживать аномалии в реальном времени, заблаговременно прогнозировать вероятность возникновения аварии и предупреждать инженера по бурению на ранней стадии, что позволит реализовать превентивные мероприятия до того, как колонна будет прихвачена.

Цель: создание алгоритма, позволяющего прогнозировать вероятность возникновения прихвата бурильной или обсадной колонны в процессе бурения скважины на основании анализа данных геолого-технологических исследований.

Методы: анализ современных достижений в области определения аномалий во временных рядах при помощи алгоритмов машинного обучения; создание алгоритма прогнозирования прихвата на языке программирования Python с использованием библиотек с открытым исходным кодом.

Результаты. Авторами разработан алгоритм прогнозирования прихвата, проведена оценка его эффективности на тестовой выборке, а также масштабируемость на прочие аварии и осложнения, возникающие в процессе бурения. Определены перспективные направления дальнейших исследований.

Ключевые слова:

Технологические операции строительства скважин, прогнозирование прихватов колонны, машинное обучение, многомерные временные ряды, распознавание аномалий, иерархическая временная память.

Введение

Прихват бурильной или обсадной колонны является одной из самых сложных аварий, которая оказывает существенное влияние на эффективность бурения и стоимость строительства скважины. Зачастую при возникновении прихвата требуется проведение дорогостоящих мероприятий по ликвидации последствий аварии, которые могут включать в себя ловильные работы, бурение бокового ствола или, в крайнем случае, ликвидацию аварийной и бурение новой скважины. Поэтому различные отраслевые оценки, утверждающие, что стоимость проведения мероприятий по ликвидации последствий прихвата бурильного инструмента может превышать несколько сотен миллионов долларов в год, не являются надуманными. Инциденты, связанные с прихватом бурильного инструмента, составляют значительную часть непроизводительного времени, при этом оценки варьируются от 25 до 35 % [1].

Понимание механизмов возникновения прихвата и его признаков помогает на ранней стадии реализовывать мероприятия по снижению последствий надви-

гающей аварии. Традиционно признаки прихвата определяются по параметрам бурения, регистрируемым станцией геолого-технологических исследований (ГТИ). Характерными признаками дифференциального прихвата являются увеличение сопротивления продольному перемещению и повышение крутящего момента при срагивании колонны после пребывания без движения. В том случае, если помимо вышеуказанных признаков отмечается повышение давления на стояке, можно говорить о вероятности возникновения механического прихвата [2].

Зачастую признаки предстоящего прихвата не обнаруживаются заблаговременно. В основном это происходит в результате особенностей человеческого восприятия информации. Кроме того, признаки предстоящей аварии не распознаются заблаговременно в результате сменного графика работы персонала (членов буровой бригады, инженерного состава на буровой площадке и в оперативном офисе) – по статистике большинство аварий происходит в течение нескольких часов после «пересменки» [1].

Внедрение машинного обучения в анализ изменения параметров бурения становится все более популярным по вышеуказанным причинам. Различные алгоритмы машинного обучения предлагают автоматизированные решения для преодоления этих ограничений, что позволяет адаптировать человеческие знания и практический опыт для решения данной проблемы.

Автоматический анализ тенденций изменения параметров бурения, определение едва заметных, но критических отклонений и их взаимосвязей, которые зачастую не являются линейными, позволяют заблаговременно прогнозировать вероятность возникновения аварии и предупреждать инженера по бурению на ранней стадии. Данные системы позволяют избежать серьезной аварии и уменьшить последствия по освобождению прихваченной колонны.

Обзор литературы

Одной из первых зарубежных публикаций, в которой использовались статистические методы для прогнозирования вероятности возникновения прихвата колонны, была работа Хемпкинса и др. [2]. Советскими учеными также разрабатывались статистические методы, в частности метод последовательной диагностической процедуры [3]. Современные подходы к раннему определению признаков прихватов описаны в работе [5].

Ряд алгоритмов машинного обучения способны предсказывать неизвестные взаимосвязи между входными и выходными параметрами. Прогнозирование вероятности возникновения аварии требует создания более точной модели ввиду высокой нелинейности и нестабильности взаимосвязей между входными и выходными значениями.

К настоящему моменту описано несколько подходов для прогнозирования вероятности прихвата колонны при помощи различных алгоритмов машинного обучения:

- статистический анализ многомерных временных рядов [6];
- логистическая регрессия [7];
- нейронная сеть [8–10];
- метод опорных векторов [10, 12].

В упомянутых ранее статистических методах использовались данные с большого количества скважин. Эти статистические методы обычно использовали суточные отчеты о бурении, отчеты об окончании строительства скважины и другую информацию, которая отражает одно измерение в день для параметров, используемых для прогнозирования прихвата. Входные данные включали информацию о траектории ствола скважины (глубина по стволу и по вертикали, зенитный угол и интенсивность искривления), данные о свойствах бурового раствора (удельный вес, реологические параметры раствора, водоотдача, водоотдача при высоком давлении и высокой температуре, pH и концентрация хлоридов), параметры бурения (нагрузка на долото, крутящий момент на поверхности, нагрузка на крюке, давление на насосе, расход буро-

вого раствора, частота вращения ротора или системы верхнего привода) [3].

Основным ограничением применения этого подхода к прогнозированию прихвата в реальном времени является то, что условия, приводящие к прихвату, часто возникают в течение нескольких десятков минут или часов. Поэтому требуется более высокая частота обновления данных, по сравнению с суточными отчетами о бурении и отчетами об окончании строительства скважины. Кроме того, в практике строительства скважин замер параметров бурового раствора производится с недостаточной периодичностью, что осложняет раннее обнаружение признаков прихвата. Разработанные системы для непрерывного автоматического измерения основных параметров бурового раствора в настоящий момент не получили широкого распространения [13].

В случае использования подхода к обнаружению прихвата при помощи алгоритмов обучения «с учителем» (логистическая регрессия, нейронная сеть, градиентный бустинг) возникает проблема разметки обучающей выборки, поскольку требуется предварительный сбор информации для обучения алгоритма, проверка качества входных данных, их предварительная обработка, ручная разметка (выделение интервалов с признаками предстоящей аварии). В случае изменения распределения параметров бурения при использовании модели в условиях другого месторождения требуется повторная подготовка обучающей выборки, обучение модели и оценка ее работы в условиях реального бурения.

Другой подход заключается в использовании аналитических моделей, прогнозирующих скручивающие и осевые нагрузки, воздействующие на бурильную колонну (T&D) в соответствии с плановой траекторией скважины. Сравнение прогнозов аналитической модели с фактическими данными позволяет выявить ранние признаки прихвата [14]. Таким образом, на этапе подготовки к строительству скважины разрабатывается модель скручивающих и осевых нагрузок, действующих на бурильную колонну, прогнозирующая вес на крюке и момент на устье при бурении для каждой точки по стволу скважины на основе информации о компоновке низа бурильной колонны, реологии бурового раствора и плановой траекторий ствола скважины. Использование данных T&D необходимо для автоматического вычисления несоответствия параметров бурения и прогнозирования возможного прихвата, но утверждается, что точность модели значительно повышается при добавлении дополнительных данных, таких как давление на стояке и расход бурового раствора.

Методика, предложенная далее в этой работе, не отрицает важности имитационных моделей. Однако предлагается иной подход, при котором не требуется разработка предварительной модели, поскольку система использует самообучающуюся и самонастраивающуюся модель и автоматически выявляет аномалии в поступающих данных в режиме реального времени, что позволяет предотвратить предстоящую аварию.

Описание исходных данных

Основным фактором успешного бурения скважины без возникновения аварий, связанных с прихватом бурильного инструмента, является возможность непрерывно отслеживать многочисленные тенденции изменения параметров бурения для раннего обнаружения аномалий в их поведении. В анализе тенденций важно не столько абсолютное значение какого-либо из параметров, сколько направление изменения одного из этих параметров по отношению к другим.

В процессе строительства скважины ключевыми параметрами являются:

- нагрузка на крюке (HKLD);
- момент на устье (TRQ);
- давление на стояке (SPP).

Понимание ожидаемой тенденции изменения этих параметров имеет большое значение для раннего обнаружения и, следовательно, для снижения последствий прихвата бурильного инструмента [1].

Для корректного анализа тенденций изменения параметров в процессе бурения необходимо регистрировать вес на крюке без вращения бурильной колонны, потому как в тот момент, когда бурильная колонна неподвижна, трение действует против осевого перемещения и его влияние на измеренное значение веса на крюке максимально [2].

Основной проблемой при бурении наклонно-направленных, горизонтальных скважин, а также скважин с большим отходом от вертикали является образование слоя шлама в затрубном пространстве. Ввиду недостаточной очистки затрубного пространства происходит накопление шламовой подушки, что приводит к постоянному повышению трения и закупориванию затрубного пространства до тех пор, пока колонна не потеряет подвижность и циркуляция не будет потеряна. Данная ситуация может быть выявлена на ранней стадии путем постоянного автоматизированного мониторинга и анализа тенденции изменения измеренных ключевых параметров (HKLD, TRQ и SPP) в режиме реального времени для каждой пробуренной свечи. Это позволит проводить превентивные мероприятия на ранней стадии, чтобы избежать потенциальных дорогостоящих потерь времени и ресурсов на ликвидацию последствий механического прихвата.

Дифференциальный прихват характеризуется повышением трения между бурильным инструментом или обсадной колонной и стенкой открытого ствола скважины в результате перепада давления. Обычно из-за чрезмерного перепада давления прилипание колонн происходит в пористых и проницаемых пластах, таких как песчаник или известняк, в которых в процессе бурения образуется толстая фильтрационная корка на стенке скважины. Ранними признаками дифференциального прихвата являются посадки и затяжки после наращивания, регистрируемые датчиком веса на крюке, повышение регистрируемого момента на устье при бурении ротором, а также «скачки» момента после запуска вращения вслед за продолжительным неподвижным состоянием колонны.

Рекомендации по предотвращению прихвата бурильного инструмента в основном касаются технологических аспектов процесса строительства скважины – поддержание нейтральной точки выше УБТ, добавление расширителя в КНБК и повышенный контроль параметров бурового раствора за счет обеспечения надлежащих реологических свойств раствора, снижения фильтрации бурового раствора, уменьшения толщины фильтрационной корки [2].

После того как определены механизмы возникновения различных типов прихватов и их признаки, следует определить основные требования, предъявляемые к регистрируемым параметрам бурения.

Одна из основных проблем регистрации данных параметров бурения заключается в том, что тип, частота и качество доступных данных не согласуются от скважины к скважине, что связано с применением различных станций геолого-технологических исследований (ГТИ).

Успешная идентификация ранних признаков предстоящей аварии будет зависеть от сбора конкретных потоков данных в режиме реального времени и использования этих данных для формирования прогнозов условий бурения. В зависимости от компании, предоставляющей сервис ГТИ, и количества датчиков, установленных на буровой установке, количество потоков данных, поступающих на сервер посредством языка разметки передачи информации WITSML, может варьироваться от 10 до 100. В целом наличие большого количества потоков данных позволяет лучше интерпретировать скважинные условия, однако большинство буровых установок снабжены только базовым набором датчиков.

Чтобы гарантировать, что система оповещения будет работать на различных типах скважин, предлагается разработка алгоритма для мониторинга скважины и предоставления предупреждений, даже если доступны только «критические» потоки данных. Эти критические потоки должны соответствовать следующим критериям [1]:

- полнота (степень, в которой каждая точка данных датчика содержит ожидаемые показания);
- чувствительность (степень, в которой значения показаний датчика соответствуют определенным пороговым значениям качества);
- однородность (степень, в которой потоковый объем точек данных датчиков согласован с течением времени);
- структура (степень, в которой точки потоковых данных соответствуют согласованной структуре).

Согласно предоставленным компанией ООО «Газпромнефть НТЦ» исходным данным, в соответствии с обозначенными ранее критериями, приведен список критических параметров, используемых в данном исследовании:

- Входные данные: параметры, непосредственно контролируемые бурильщиком на поверхности при проводке скважины:
 - расход бурового раствора на входе (л/с);
 - частота вращения бурильной колонны (об/мин);
 - высота талевого блока (м).

- Выходные данные – измерения, которые показывают реакцию скважины на входные параметры бурения:

- давление на стояке (атм);
- момент при вращении на поверхности (Н*м);
- нагрузка на крюке (т);
- глубина долота, глубина скважины (м).

Различные параметры в скважинах, как правило, регистрируются с частотой 1 точка/5 секунд (0,2 Гц). Анализ показал, что для надежного прогнозирования

вероятности возникновения аварии требуется не менее 1 точки/10 секунд (0,1 Гц) для каждого параметра [14].

На рис. 1 приведен пример исходных данных ГТИ, представленных временными рядами параметров бурения (глубина долота и скважины, вес на крюке, давление на стояке), со скважины, на которой произошел дифференциальный прихват. На рисунке отмечены три зоны, в которых бурильная колонна была прихвачена.

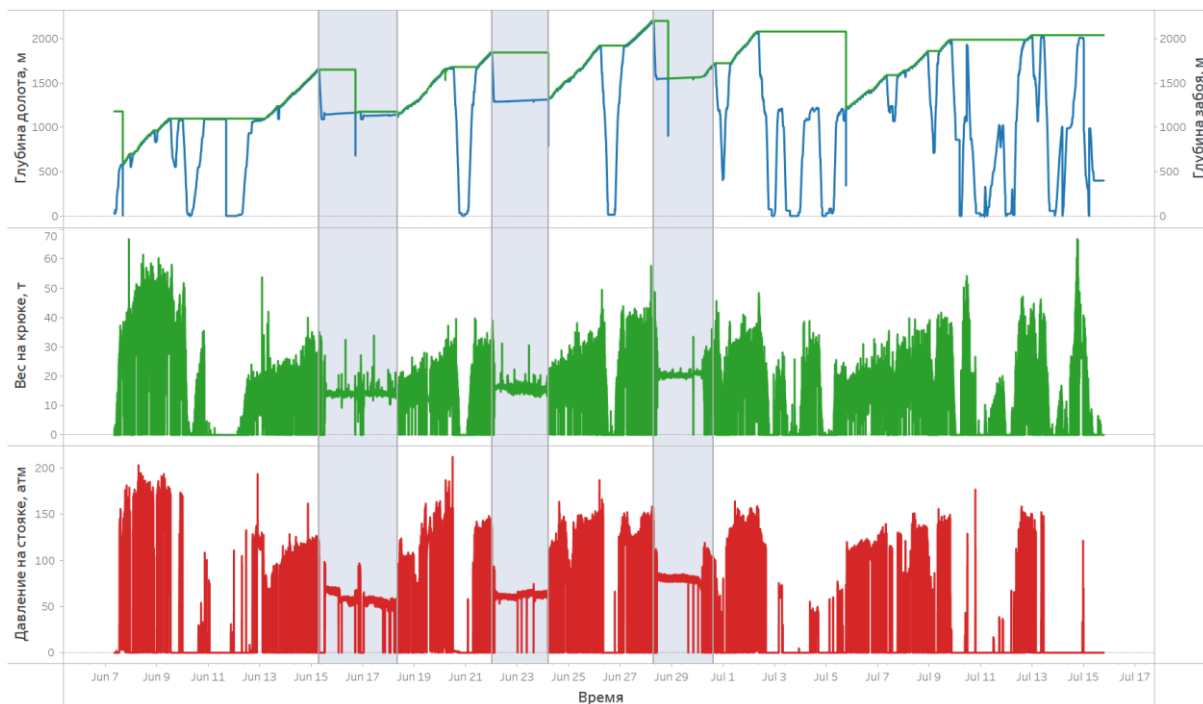


Рис. 1. Исходные данные параметров бурения обучающей выборки

Fig. 1. Drilling parameters time series of training dataset

Стоит отметить, что в соответствии с анализом различных инцидентов, связанных с прихватом бурильного инструмента, можно утверждать, что во всех случаях не было обнаружено единого опережающего признака предстоящей аварии, поэтому требуется разработать алгоритм, комплексно анализирующий различные параметры бурения.

Методика проведения исследований

Как отмечалось ранее, применение алгоритмов машинного обучения позволяет исключить человеческий фактор, тем самым сократить непроизводительное время и повысить эффективность бурения скважин.

На рис. 2 показана блок-схема предлагаемой системы прогнозирования прихвата бурильной или обсадной колонны [1].

Как только данные ГТИ в режиме реального времени получены с буровой площадки, выполняется контроль качества данных и их обработка (удаление пропусков, выбросов и пр.). На следующем шаге выполняется классификация данных посредством определения технологической операции, выполняемой на буровой установке (бурение ротором или слайдированием, проработка, наращивание и пр.) (табл. 1) [14].

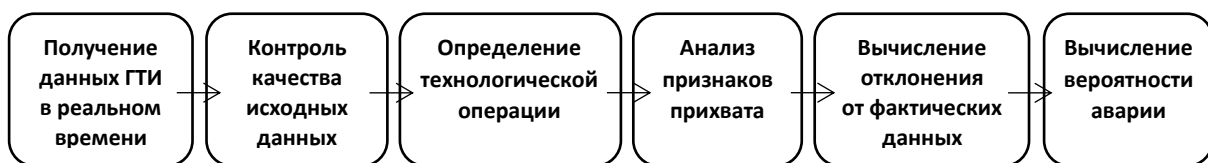


Рис. 2. Блок-схема системы прогнозирования вероятности прихвата

Fig. 2. Stuck pipe model flowchart

Таблица 1. Условия классификации технологических операций, выполняемых на буровой установке
Table 1. Classification conditions of rig states

Технологическая операция Rig state	Наличие вращения СВП Rotation	Наличие циркуляции бурового раствора Flowrate	Направление движения крюкоблока Bit movement	Увеличение глубины скважины Hole depth increase	Наличие нагрузки на крюке Hookload
Роторное бурение Rotary drilling	Да/Yes	Да/Yes	Вниз/Down	Да/Yes	Да/Yes
Бурение слайдированием Sliding drilling	Нет/No	Да/Yes	Вниз/Down	Да/Yes	Да/Yes
Спуск с циркуляцией Wash downwards	Нет/No	Да/Yes	Вниз/Down	Нет/No	Да/Yes
Подъем с циркуляцией Wash upwards	Нет/No	Да/Yes	Вверх/Up	Нет/No	Да/Yes
Прямая проработка Ream downwards	Да/Yes	Да/Yes	Вниз/Down	Нет/No	Да/Yes
Обратная проработка Ream upwards	Да/Yes	Да/Yes	Вверх/Up	Нет/No	Да/Yes
Спуск без циркуляции Run in hole	Нет/No	Нет/No	Вниз/Down	Нет/No	Да/Yes
Подъем без циркуляции Pull out of hole	Нет/No	Нет/No	Вверх/Up	Нет/No	Да/Yes
Нарращивание Make connection	Нет/No	Нет/No	—	Нет/No	Нет/No
Промывка над забоем Circulate hole	Нет/No	Да/Yes	Без движения Static	Нет/No	Да/Yes
Вращение над забоем Rotating hole	Да/Yes	Да/Yes	Без движения Static	Нет/No	Да/Yes

Затем данные подаются на вход ряду алгоритмов, предназначенных для поиска различных признаков, сигнализирующих о предстоящей аварии, связанной с прихватом бурового инструмента. В случае обнаружения одного или нескольких признаков прихвата колонны модель рассчитает вероятность возникновения аварии. Если она превысит определенное пороговое значение, модель сообщит инженеру о высокой вероятности возникновения аварии.

Как уже упоминалось ранее, нет единого опережающего признака, определяющего вероятность возникнове-

ния аварии. Следовательно, требуется, чтобы метод раннего обнаружения был достаточно гибким, чтобы генерировать действительные предупреждения, не зависящие от отклонения одного конкретного параметра.

Методы машинного обучения могут быть достаточно гибкими, чтобы обеспечить минимальное количество ложных срабатываний. Для задач выявления аномалий во временных рядах используется огромное множество различных алгоритмов, классификация которых представлена на рис. 3 [16].

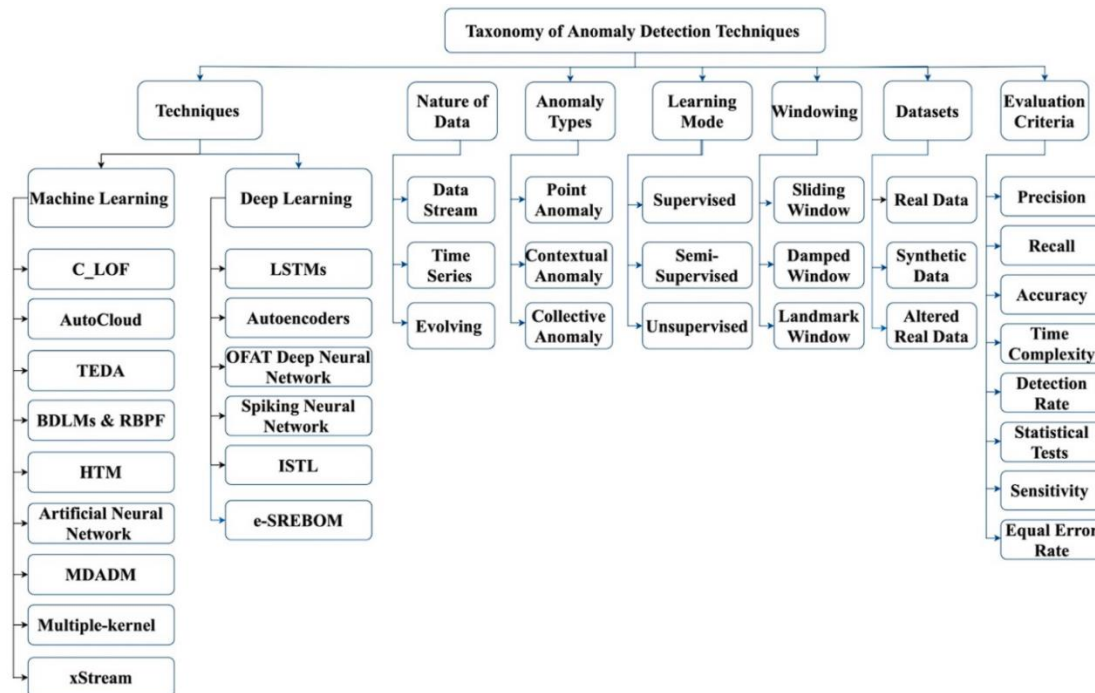


Рис. 3. Классификация алгоритмов машинного обучения, методов и критериев оценки качества выявления аномалий во временных рядах

Fig. 3. Taxonomy of anomaly detection techniques using machine learning

Аномалии – это наблюдения, которые значительно отличаются от других наблюдений, что, возможно, обусловлено другим механизмом работы системы. Аномалии бывают точечными и контекстуальными.

Точечная аномалия возникает, когда какая-либо точка в потоке данных значительно отличается от ожидаемой закономерности распределения этих данных. Такие точки называют выбросами [16].

Обнаружение аномалий такого типа включает в себя наблюдение за каждой точкой в общем потоке данных. На рис. 4 показана точечная аномалия [16].

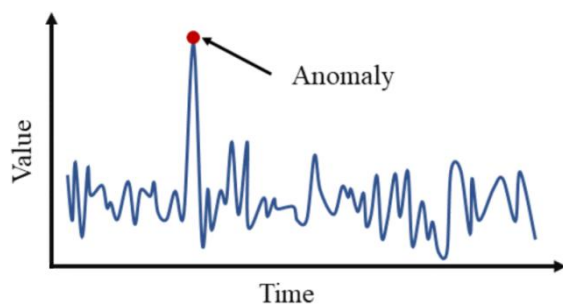


Рис. 3. Пример выявления точечной аномалии

Fig. 3. Point anomaly detection

Другой тип аномалии наблюдается для любого шаблона данных, который аномален в одном из сценариев распределения данных. Распространенным примером являются большие дорожные пробки, которые могут быть контекстуально аномальной активностью движения после полуночи из-за аварии, плохой видимости или других причин, связанных с погодными условиями. Пример контекстуальной аномалии показан на рис. 5 Рис. 4[16].

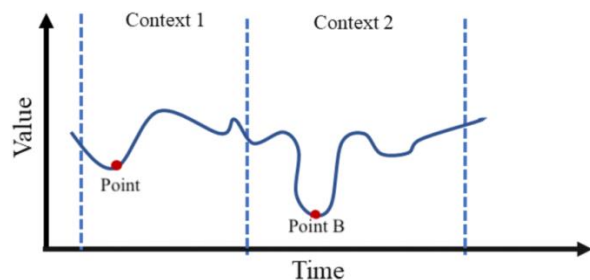


Рис. 4. Пример выявления контекстуальной аномалии

Fig. 4. Contextual anomaly detection

В соответствии с классификацией, приведенной на Рис. 3, алгоритм может использовать схему обучения «с учителем», «с частичным привлечением учителя» (полуавтоматическое обучение), или обучаться «без учителя» [17].

Подход к обнаружению аномалий при работе алгоритма «с учителем» предполагает обнаружение аномалий путем создания набора правил, выявленных в процессе обучения алгоритма на размеченных наборах данных, которые помогают прогнозировать будущие данные.

Подход к обнаружению аномалий при работе алгоритма «с частичным привлечением учителя» предполагает обнаружение аномалий путем создания пер-

воначального набора правил, выявленных на небольшой размеченной выборке данных, в дальнейшем алгоритм корректирует первоначальные правила путем обработки поступающих неразмеченных данных.

Подход к обнаружению аномалий при работе алгоритма «без учителя» предполагает обнаружение аномалий на основе неразмеченного потока данных, таким образом, алгоритм автоматически определяет, какие данные считать нормальными и аномальными. Данный подход не требует отдельных этапов обучения и тестирования.

Как уже упоминалось ранее, при обучении «с учителем» существует проблема сбора и разметки обучающей выборки. Для решения данной проблемы предлагается использовать систему автоматического обнаружения аварий и осложнений в процессе бурения на основе метода обучения «без учителя». Данный подход позволит сократить временные и материальные затраты на подготовку и обучение модели. Кроме того, стоит отметить, что предлагается реализовать систему, которая обучается непосредственно в процессе работы, таким образом, алгоритму не требуется изначально иметь весь набор данных, потому как с каждой новой точкой он будет корректировать свою работу в режиме реального времени.

Обнаружение признаков аварии позволяет заблаговременно проводить превентивные мероприятия, поэтому требуется, чтобы эта информация была достаточно преждевременной для ее использования. Существует компромисс между ранним обнаружением и ложными срабатываниями, поскольку алгоритм, который часто делает неточные обнаружения, скорее всего, будет проигнорирован.

Резюмируя, можно выделить следующие требования к предлагаемому решению:

- 1) Прогнозы должны быть сделаны в режиме реального времени – алгоритм должен идентифицировать состояние x_t как нормальное или аномальное до получения последующего x_{t+1} .
- 2) Алгоритм должен непрерывно обучаться без необходимости изначально иметь и хранить весь объем данных.
- 3) Алгоритм должен обучаться «без учителя» – без предварительной разметки данных или ручной настройки параметров.
- 4) Алгоритм должен динамически адаптироваться к изменениям состояния системы и смене распределения данных, поскольку базовые статистики потока данных параметров бурения зачастую нестационарны.
- 5) Алгоритм должен распознавать аномалии в параметрах бурения как можно раньше.
- 6) Алгоритм должен иметь минимальное количество как ложноположительных, так и ложноотрицательных срабатываний.

Анализ приведенных выше критериев позволяет сделать вывод о том, что алгоритмы обнаружения аномалий в неизменных временных рядах принципиально отличаются от тех, что требуются для работы в режиме реального времени, поэтому большинство алгоритмов, приведенных ранее, не применимы в контексте данной проблемы.

Для решения поставленной задачи и выполнения поставленных требований предлагается использовать следующий алгоритм – Hierarchical Temporal Memory (HTM).

Hierarchical Temporal Memory (Иерархическая Временная Память) – это частная модель мозга, которая моделирует некоторые структурные и алгоритмиче-

ские свойства неокортекса. Данный метод машинного обучения основан на большом количестве шаблонов последовательностей, хранящихся в нем, формируемых с помощью потока данных временных рядов. На рис. 6 представлена схема работы алгоритма в случае использования одномерного временного ряда [18].

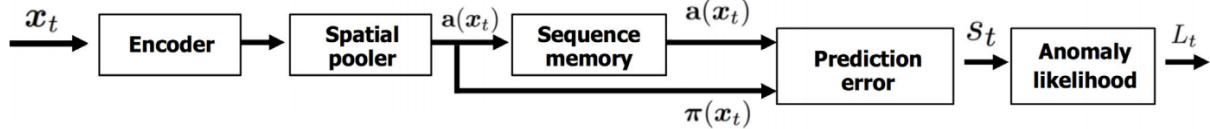


Рис. 5. Принципиальная схема работы алгоритма HTM для обработки одномерного временного ряда
 Fig. 5. Primary functional steps in HTM algorithm for univariate-sensing time-series

Текущий входной сигнал x_t подается на декодер (Encoder), который переводит входные значения в двоичный код, затем выполняется процесс разреженного пространственного объединения (Spatial pooler). Результирующий вектор $a(x_t)$ представляет собой разреженный двоичный вектор, отображающий текущий входной сигнал в алгоритм иерархической временной памяти (Sequence memory) – ключевой компонент данной системы. Данный компонент моделирует временные закономерности (x_t) и выводит прогноз в виде другого разреженного вектора $\pi(x_t)$. Таким образом, $\pi(x_t)$ является предсказанием для $a(x_{t+1})$ [18]. Однако $a(x_t)$ и $\pi(x_t)$ напрямую не отражают аномалии во временных рядах. Для создания надежной системы обнаружения аномалий в алгоритм вводится два дополнительных шага. Во-первых, необработанная оценка аномалии (Prediction error) вычисляется по двум разреженным векторам – предсказанному и действительному значению. Так как необработанные оценки аномалии могут быть зашумлены, с целью снижения количества ложных срабатываний вычисляется вероятность аномалии (Anomaly likelihood) в соответствии с гауссовым распределением, при превышении вероятности аномалии порогового значения состояние системы отмечается аномальным.

Иерархическая временная память состоит из слоев нейронов, организованных в набор столбцов (рис. 7). Сеть принимает поток входных данных, закодированных в виде разреженных векторов, моделирует последовательности высокого порядка (последовательности с долгосрочными зависимостями), используя композицию из двух отдельных разреженных представлений. Текущий ввод, x_t и контекст предыдущей последовательности (x_{t-3} , x_{t-2} , x_{t-1}) одновременно кодируются с использованием динамически обновляемого разреженного распределенного представления. Сеть использует эти представления для прогнозирования следующего временного шага в виде разреженного вектора. При получении входных данных для следующего временного шага сеть использует разницу между прогнозируемым значением и фактическим вводом для обновления своих синаптических соединений. Обучение происходит на каждом временном шаге, но поскольку представления очень разрежены, обновляется лишь крошечный процент синапсов [18].

Дополнительные детали алгоритма HTM и свойства его представления выходят за рамки данной работы, но подробно описаны в следующих работах [18, 18].

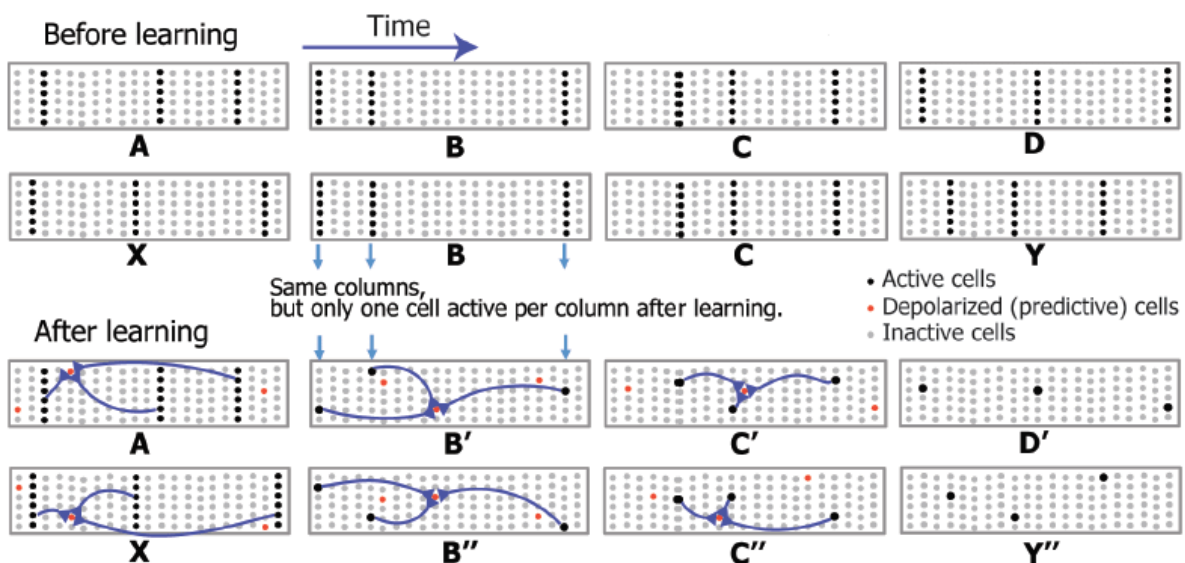


Рис. 6. Блок иерархической временной памяти
 Fig. 6. Hierarchical temporal memory block

Так как входные данные содержат большое количество потоков данных с различных датчиков буровой установки, такие среды не могут быть эффективно обработаны с помощью одной сложной модели, так как сложность обучения и вывода изменяется квадратично с увеличением входной размерности. В таком случае следует разделить большую систему на набор меньших моделей, каждая из которых моделирует подмножество среды. Меньшие модели облегчают обучение и повышают производительность, но могут потерять потенциально полезные корреляции между другими моделями [19].

Входной набор данных представляет многомерный временной ряд, поэтому используется следующая принципиальная схема работы алгоритма (рис. 8).

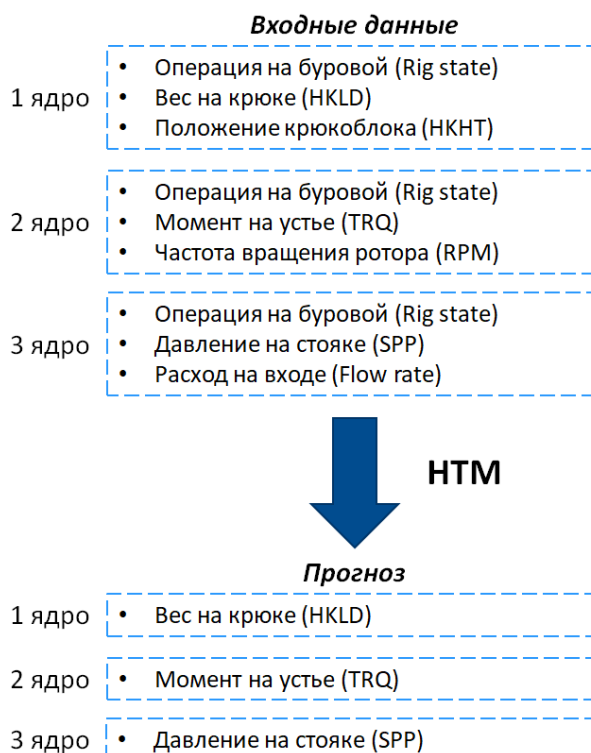
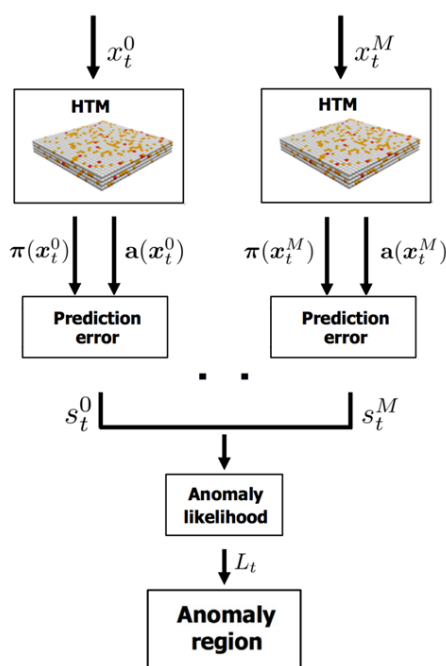


Рис. 7. Принципиальная схема алгоритма обработки многомерного временного ряда параметров бурения скважины
 Fig. 7. Functional diagram illustrating a complex system for multivariate-sensing time-series

Таблица 2. Гиперпараметры модели иерархической временной памяти

Table 2. Hierarchical temporal memory hyperparameters

Наименование параметра Parameter name	Значение Value
Количество ячеек/Number of columns N	2048
Количество слоев в ячейке Number of cells per column M	16
Порог активации сегмента Dendritic segment activation threshold Θ	19
Исходное состояние синапса Initial synaptic permanence	0,23075451521239015
Увеличение связи синапсов Synaptic permanence increment p^+	0,04329160631961319
Уменьшение связи синапсов Synaptic permanence decrement p^-	0,007285708940843109
Уменьшение связи синапсов для предсказанной неактивной ячейки Synaptic permanence decrement for predicted inactive segments p^-	0,00102650387546339

Реализуется три ядра алгоритма, на каждый из которых подаются на вход выделенные части доступного набора данных. Для прогнозирования вероятности прихвата бурильного инструмента отслеживаются такие ключевые параметры, как нагрузка на крюке, момент на устье, давление на стояке, для различных технологических операций. Для того чтобы суммировать результаты работы отдельных ядер и сократить объем вычислений, рассчитывается логарифм вероятности возникновения аварии по каждому отдельному параметру, затем результаты предсказаний ядер суммируются. При превышении суммарного значения порогового значения состояние системы отмечается аномальным.

С целью отслеживания динамики изменения состояния системы в скользящем окне вычисляется количество точек за последние 30 минут, в которых поведение системы было аномальным. При превышении частоты появления аномальных интервалов порогового значения будет сгенерировано предупреждение для инженера по бурению.

На обучающей выборке были определены следующие гиперпараметры модели (табл. 2).

Результаты исследования

Для выполнения данной исследовательской работы компанией ООО «Газпромнефть НТЦ» был предоставлен исходный набор данных временных рядов параметров бурения скважины, на которой произошла серия дифференциальных прихватов. Предоставленные данные были переданы реализованному алго-

ритму на вход, при выполнении программного кода данные передавались строчка за строчкой, таким образом, изначально алгоритм не имел возможности использовать весь набор данных, тем самым обеспечивалось обучение алгоритма в режиме реального времени.

Для оценки качества работы алгоритма использовались такие метрики, как среднеквадратичная ошибка прогнозирования определенного параметра, время получения предупреждения до того, как авария произошла.

Результаты работы алгоритма на обучающей выборке представлены на рис. 9. На верхнем треке отображены значения веса на крюке в течение рейса, на нижнем отмечены зоны, для которых вероятность возникновения аварии превысила предопределенный порог, и в данной точке отмечаются признаки прихвата.

Параметры используемой модели, а также архитектура алгоритма определялись на данном наборе данных путем запуска серии экспериментов с различными конфигурациями.

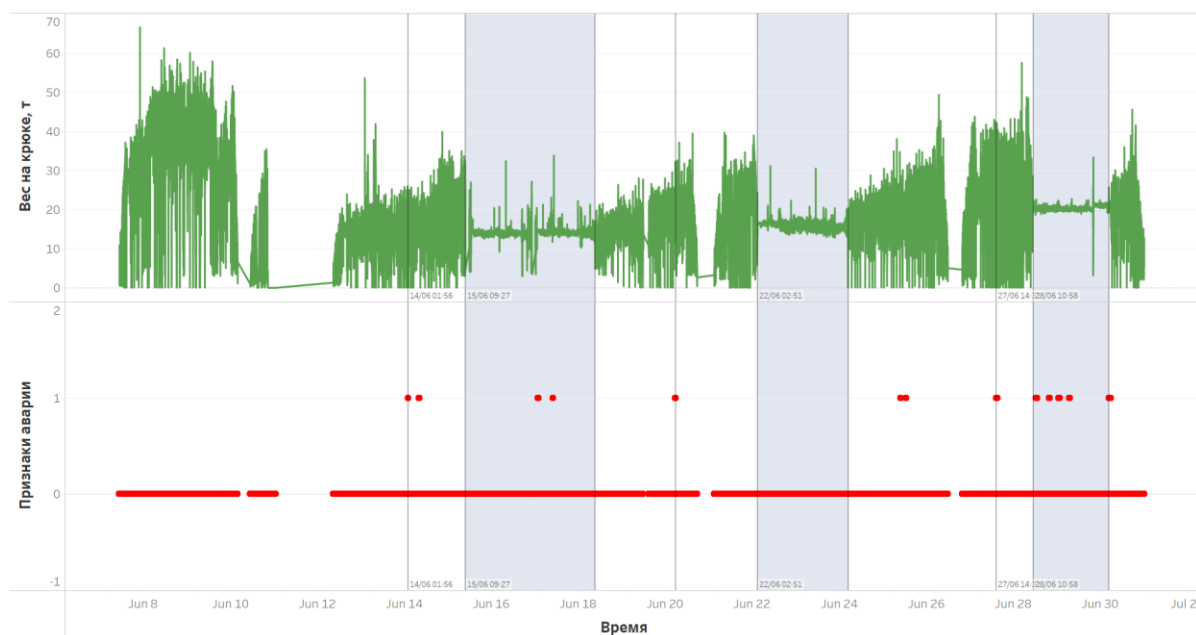


Рис. 8. Результаты работы алгоритма иерархической временной памяти на обучающей выборке

Fig. 8. Hierarchical temporal memory results on training dataset

На первом интервале произошедшего прихвата предупреждение было сгенерировано за 31 час до инцидента. На втором интервале предупреждение было сгенерировано до подъема КНБК, последующих предупреждений непосредственно до момента прихвата не было, поэтому будем считать, что модель корректно не распознала признаки предстоящей аварии. На третьем интервале произошедшего прихвата предупреждение было сгенерировано за 20 часов до инцидента. Медианная оценка среднеквадратичной ошибки прогнозирования равна 2,42 т, что является достаточным показателем для выявления признаков предстоящей аварии. Также немаловажной характеристикой работы модели является количество ложных срабатываний, при работе алгоритма с данными предоставленной выборки было обнаружено четыре ложных срабатывания. При бурении под кондуктор признаков прихвата не наблюдалось, ложные срабатывания также отсутствовали.

С целью проверки качества работы алгоритма на тестовой выборке были зафиксированы параметры модели. Тестовая выборка включала в себя две скважины, на которых в определенный момент времени произошел дифференциальный прихват. Результаты работы алгоритма на тестовой выборке представлены на рис. 10, 11.

Как видно из представленного графика, предупреждение о вероятности возникновения прихвата было сгенерировано за 30 часов до возникновения инцидента. Медианная оценка среднеквадратичной ошибки прогнозирования составляет 1,02 т, что является достаточным показателем для выявления признаков предстоящей аварии. При работе алгоритма с данными тестовой выборки скважины № 1 было обнаружено четыре ложных срабатывания.

Анализ результатов работы модели на тестовой скважине № 2 показал, что предупреждение о вероятности возникновения прихвата было сгенерировано за 13 часов до возникновения инцидента. Медианная оценка среднеквадратичной ошибки прогнозирования составляет 0,89 т, что является достаточным показателем для выявления признаков предстоящей аварии. При работе алгоритма было сгенерировано семь ложных срабатываний.

Заключение

В данной работе были приведены требования, предъявляемые к алгоритмам, обучающимся «без учителя», для обнаружения аномалий в потоковых данных в режиме реального времени. На основании недостатков разрабатываемых ранее подходов к про-

гнозированию прихвата колонны был предложен новый метод автоматического обнаружения инцидентов. Алгоритм, основанный на иерархической временной памяти, способен обнаруживать точечные и контек-

стуальные аномалии в зашумленных данных. Он отвечает требованиям непрерывного обнаружения аномалий в режиме реального времени без предварительного доступа ко всему объему входных данных.

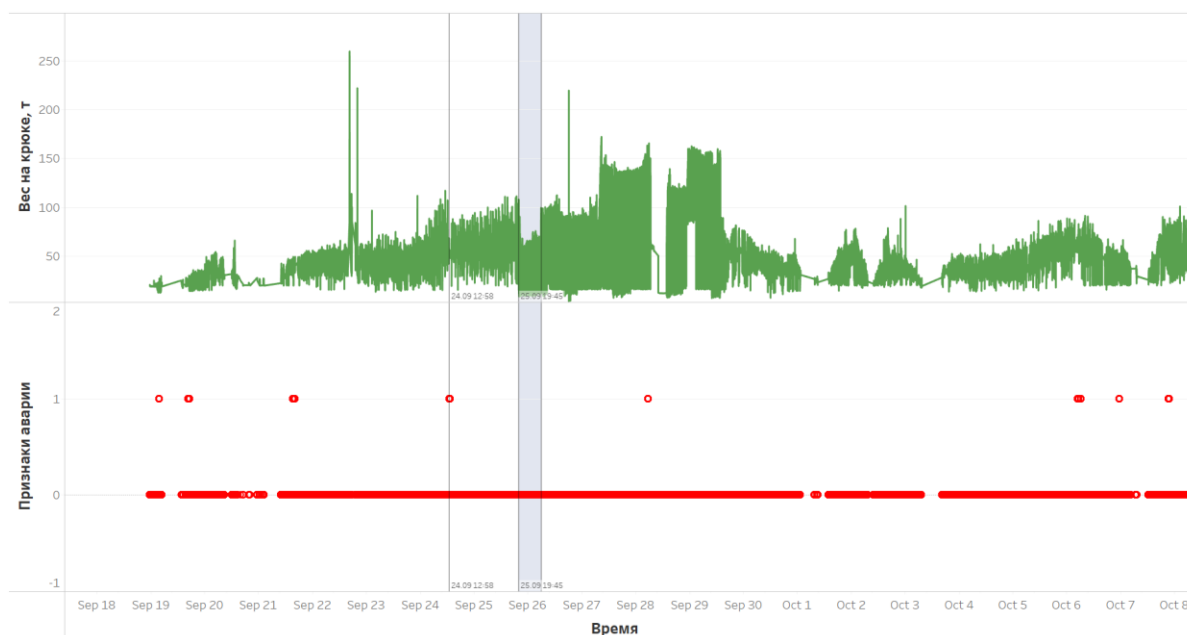


Рис. 9. Результаты работы алгоритма иерархической временной памяти на тестовой выборке, скважина № 1
Fig. 9. Hierarchical temporal memory results on test dataset, well no. 1



Рис. 10. Результаты работы алгоритма иерархической временной памяти на тестовой выборке, скважина № 2
Fig. 10. Hierarchical temporal memory results on test dataset, well no. 2

Результаты работы алгоритма доказывают его применимость в задаче прогнозирования вероятности возникновения аварий в процессе бурения.

В соответствии с полученными результатами определены направления дальнейших исследований:

1) совершенствование методики автоматического прогнозирования прихвата;

- 2) разработка методов автоматической оценки качества входных данных в реальном времени;
- 3) проведение опытно-промышленных испытаний;
- 4) разработка пользовательского интерфейса;
- 5) масштабирование предложенного метода на другие типы осложнений и аварий с целью повышения безопасности и эффективности работ по строительству скважин.

СПИСОК ЛИТЕРАТУРЫ

1. Stuck pipe early warning system utilizing moving window machine learning approach / A.S. Omogbolahan, A.M. Beshir, Z.A. Majed, F.I. Ajikobi // Abu Dhabi International Petroleum Exhibition & Conference. – Abu Dhabi, UAE, 2019. URL: <https://doi.org/10.2118/197674-MS> (дата обращения 02.03.2022).
2. Митчелл Дж. Безаварийное бурение. – Ижевск: Институт компьютерных исследований, 2017. – 364 с.
3. Multivariate statistical analysis of stuck drillpipe situations / W.B. Hemphkins, R.H. Kingsborough, W.E. Lohec, C.J. Nini // SPE Drilling Engineering. – 1987. – V. 03. – P. 237–244. URL: <https://doi.org/10.2118/14181-PA> (дата обращения 02.03.2022).
4. Мирзаджанзаде А.Х., Аветисов А.Г., Булатов А.И. Методические указания по применению статистических методов в бурении нефтяных и газовых скважин. – Краснодар: ВНИИКР-нефть, 1983. – 316 с.
5. Липатов Е.Ю. Исследование и разработка технологии и технических средств для предупреждения и ликвидации прихвата бурильной колонны (на примере месторождений Среднего Приобья): моногр. – Тюмень: ТюмГНГУ, 2019. – 128 с.
6. Weakley R.R. Use of stuck pipe statistics to reduce the occurrence of stuck pipe // SPE Annual Technical Conference and Exhibition. – New Orleans, USA, 1990. URL: <https://doi.org/10.2118/20410-MS> (дата обращения 02.03.2022).
7. Wisnie A.P., Zhiwei Zhu. Quantifying stuck pipe risk in Gulf of Mexico oil and gas drilling // SPE Annual Technical Conference and Exhibition. – New Orleans, USA, 1994. URL: <https://doi.org/10.2118/28298-MS> (дата обращения 02.03.2022).
8. Кодиров Ш.Ш., Шестаков А.Л. Разработка искусственной нейронной сети для прогнозирования прихватов колонн бурильных труб // Вестник ЮУрГУ. Компьютерные технологии, управление, радиоэлектроника. – 2019. – № 3. – С. 20–32. DOI: 10.14529/ctcr190302
9. Borozdin S., Dmitrievsky A., Eremin N. Drilling problems forecast system based on neural network // SPE Annual Caspian Technical Conference. – Virtual, 2020. URL: <https://doi.org/10.2118/202546-MS> (дата обращения 02.03.2022).
10. Абу-Абед Ф.Н. Снижение риска при строительстве газовых скважин на базе нейросетевой модели // Газовая промышленность. – 2014. – № S712 (712). – С. 100–102.
11. Intelligent prediction of differential pipe sticking by support vector machine compared with conventional artificial neural networks: an example of Iranian offshore oil fields / J. Reza, K. Reza, A.S. Mahdi, A. Emamzadeh // SPE Drilling & Completion. – 2012. – V. 27. – P. 586–595. URL: <https://doi.org/10.2118/163062-PA> (дата обращения 02.03.2022).
12. Chamkalani Ali, Pordel Shahri Mojtaba, Saeed Poordad. Support vector machine model: a new methodology for stuck pipe prediction // SPE Unconventional Gas Conference and Exhibition. – Muscat, Oman, 2013. URL: <https://doi.org/10.2118/164003-MS> (дата обращения 02.03.2022).
13. Vajargah Ali Karimi, Sullivan Gregory, Eric van Oort. Automated fluid rheology and ECD management // SPE Deepwater Drilling and Completions Conference. – Galveston, USA, 2016. URL: <https://doi.org/10.2118/180331-MS> (дата обращения 02.03.2022).
14. Stuck-pipe prediction by use of automated real-time modeling and data analysis / K. Salminen, C. Cheatham, M. Smith, Kh. Valiullin // SPE Drilling & Completion. – 2017. – V. 32. – P. 184–193. URL: <https://doi.org/10.2118/178888-PA> (дата обращения 02.03.2022).
15. Arnaout Mohammad Arghad. Distributed multi-sensor fusion system for drilling rig state detection: Doctoral Thesis. – Leoben, 2014. – 132 p.
16. A review of machine learning and deep learning techniques for anomaly detection in IoT data / R. Al-amri, R.K. Murugesan, M. Man, A.F. Abdulateef, M.A. Al-Sharafi, A.A. Alkhatani // Applied Sciences. – 2021. – V. 12. URL: <https://doi.org/10.3390/app11125320> (дата обращения 02.03.2022).
17. Машинное обучение / Б. Хенрик, Д. Ричардс, М. Феверолф, И. Рузмайкина. – СПб: Изд-во «Питер», 2021. – 336 с.
18. Unsupervised real-time anomaly detection for streaming data / S. Ahmad, A. Lavin, S. Purdy, Z. Agha // Neurocomputing. – 2017. – V. 262. – P. 134–147. URL: <https://doi.org/10.1016/j.neucom.2017.04.070> (дата обращения 02.03.2022).
19. Ahmad S., Hawkins J. Properties of sparse distributed representations and their application to Hierarchical Temporal Memory // arXiv. 2015. URL: <https://arxiv.org/abs/1503.07469> (дата обращения 02.03.2022).
20. Multivariate-time-series-driven real-time anomaly detection based on Bayesian network sensors / N. Ding, H. Gao, H. Bu, H. Ma, H. Si // Sensors. – 2018. – V. 10. URL: <https://doi.org/10.3390/s18103367> (дата обращения 02.03.2022).

Поступила: 02.03.2022 г.

Информация об авторах

Щербаков Р.Э., аспирант Инженерной школы природных ресурсов Национального исследовательского Томского политехнического университета; главный специалист Центра управления строительством скважин АО «Газпромнефть НТЦ».

Ковалев А.В., кандидат технических наук, доцент отделения нефтегазового дела Инженерной школы природных ресурсов Национального исследовательского Томского политехнического университета.

UDC 622.248.5:004.85

USING UNSUPERVISED MACHINE LEARNING ALGORITHM TO PREVENT THE STICKING OF DRILLING AND CASING STRINGS

Roman E. Shcherbakov^{1,2},
res718@bk.ru

Artem V. Kovalev¹,
Kovalevav@tpu.ru

¹ National Research Tomsk Polytechnic University,
30, Lenin avenue, Tomsk, 634050, Russia.

² «Gazpromneft NTC»,
75–79, liter D, Moika River embankment, St. Petersburg, 190000, Russia.

The relevance. Drilling failures and accidents will continue to attract attention in drilling for oil and gas as more complex wells are being drilled across depleted zones to reach deeper reservoir targets. Stuck pipe incident continue to be a major contributor to non-productive time in drilling operations for oil and gas. When a stuck pipe incident occurs, costly corrective actions may include fishing operations, side-tracking the hole, or completely having to drill a new well. Stuck pipe warning signs are often undetected early enough for the deployment of effective mitigation strategies due to human mistakes and crew changes during drilling operations. The unsupervised machine learning algorithm is programmed to automatically detect abnormalities in real-time drilling parameter trends and predict potential stuck pipes, communicate observations in the form of alerts to engineers in advance to allow proactive corrective actions. Early detection of a stuck pipe and mitigating the incident in real time not only help to prevent its occurrence, but also help in making informed decisions to the appropriate freeing mechanism to adopt if it occurs.

The main aim: create the stuck pipe detection model which predicts failure probability during the well drilling using mud logging service data.

Objects: multivariate-sensing time-series data of mud logging service.

Methods: analysis of current anomaly detection techniques achievements in the field of using machine learning; developing the stuck pipe detection model with open-source Python frameworks.

Results. The authors have developed stuck pipe detection model with HTM algorithm, evaluated performance with test dataset. Promising areas of further research were identified.

Key words:

Drilling operation, stuck pipe, machine learning, multivariate-sensing time-series, anomaly detection, hierarchical temporal memory.

REFERENCES

- Omogbolahan A.S., Beshir A.M., Majed Z.A., Ajikobi F.I. Stuck pipe early warning system utilizing moving window machine learning approach. *Abu Dhabi International Petroleum Exhibition & Conference*. Abu Dhabi, UAE, 2019. Available at: <https://doi.org/10.2118/197674-MS> (accessed 2 March 2022).
- Mitchell J. *Trouble-free drilling: stuck pipe prevention*. Austin, Drillbert Engineering, Incorporated, 2002. 300 p.
- Hempkins W.B., Kingsborough R.H., Lohec W.E., Nini C.J. Multivariate statistical analysis of stuck drillpipe situations. *SPE Drilling Engineering*, 1987, vol. 3, pp. 237–244. Available at: <https://doi.org/10.2118/14181-PA> (accessed 2 March 2022).
- Mirzadzhanzade A.H., Avetisov A.G., Bulatov A.I. *Metodicheskie ukazaniya po primeneniyu statisticheskikh metodov v burenii neflyanykh i gazovykh skvazhin* [Methodological guidelines on the application of statistical methods in oil and gas drilling]. Krasnodar, VNIKRneft Publ., 1983. 62 p.
- Lipatov E.Yu. *Issledovanie i razrabotka tekhnologii i tekhnicheskikh sredstv dlya preduprezhdeniya i likvidatsii prikhvata burilnoy kolonny (na primere mestorozhdeniy Srednego Priobya)*. Monografiya [Research and development of technology and technical means for prevention and liquidation of drill string sticking (on the example of Middle Priobye fields. Monography)]. Tyumen, TyumGNGU Publ., 2015. 128 p.
- Weakley R.R. Use of stuck pipe statistics to reduce the occurrence of stuck pipe. *SPE Annual Technical Conference and Exhibition*. New Orleans, Louisiana, 1990. Available at: <https://doi.org/10.2118/20410-MS>
- Wisnie A.P., Zhiwei Zhu. Quantifying stuck pipe risk in Gulf of Mexico oil and gas drilling. *SPE Annual Technical Conference and Exhibition*. New Orleans, USA, 1994. Available at: <https://doi.org/10.2118/28298-MS> (accessed 2 March 2022).
- Qodirov Sh.Sh., Shestakov A.L. Development of artificial neural network for predicting drill pipe sticking. *Bulletin of the South Ural State University. Computer Technologies, Automatic Control, Radio Electronics*, 2019, vol. 19, no. 3, pp. 20–32. In Rus. DOI: 10.14529/ctcr190302
- Borozdin S., Dmitrievsky A., Eremin N. Drilling problems forecast system based on neural network. *SPE Annual Caspian Technical Conference*. Virtual, 2020. Available at: <https://doi.org/10.2118/202546-MS> (accessed 2 March 2022).
- Abu-Abed F.N. Snizhenie riska pri stroitelstve gazovykh skvazhin na baze neyrosetevoy modeli [Risk minimization during construction of gas wells based on the neural network model]. *Gas Industry*, 2014, no. S712 (712), pp. 100–102.
- Reza J., Reza K., Mahdi A.S., Emamzadeh A. Intelligent prediction of differential pipe sticking by support vector machine compared with conventional artificial neural networks: an example of Iranian offshore oil fields. *SPE Drilling & Completion*, 2012, vol. 27, pp. 586–595. Available at: <https://doi.org/10.2118/163062-PA> (accessed 2 March 2022).
- Chamkalani Ali, Pordel Shahri Mojtaba, Saeed Poordad. Support vector machine model: a new methodology for stuck pipe prediction. *SPE Unconventional Gas Conference and Exhibition*. Muscat, Oman, 2013. Available at: <https://doi.org/10.2118/164003-MS> (accessed 2 March 2022).
- Vajargah A.K., Sullivan G., Van Oort E. Automated fluid rheology and ECD management. *SPE Deepwater Drilling and Completions Conference*. Galveston, USA, 2016. Available at: <https://doi.org/10.2118/180331-MS> (accessed 2 March 2022).
- Salminen K., Cheatham C., Smith M., Valiullin K. Stuck-pipe prediction by use of automated real-time modeling and data analysis. *SPE Drilling & Completion*, 2017, vol. 32, pp. 184–193. Available at: <https://doi.org/10.2118/178888-PA> (accessed 2 March 2022).

15. Arnaout Mohammad Arghad. *Distributed multi-sensor fusion system for drilling rig state detection*. Doctoral Thesis. Leoben, 2014. 132 p.
16. Al-amri R., Murugesan R.K., Man M., Abdulateef A.F., Al-Sharafi M.A., Alkahtani A.A. A review of machine learning and deep learning techniques for anomaly detection in IoT data. *Applied Sciences*, 2021, vol. 12. Available at: <https://doi.org/10.3390/app11125320> (accessed 2 March 2022).
17. Brink H., Richards J.W., Fetherolf M. *Real-world machine learning*. New York, Manning, 2017. 264 p.
18. Ahmad S., Lavin A., Purdy S., Agha Z. Unsupervised real-time anomaly detection for streaming data. *Neurocomputing*, 2017, vol. 262, pp. 134–147.
19. Ahmad S., Hawkins J. *Properties of sparse distributed representations and their application to Hierarchical Temporal Memory*. 2015. Available at: <https://arxiv.org/abs/1503.07469> (accessed 2 March 2022).
20. Ding N., Gao H., Bu H., Ma H., Si H. Multivariate-time-series-driven real-time anomaly detection based on Bayesian network. *Sensors*, 2018, vol. 10. <https://doi.org/10.3390/s18103367> (accessed 2 March 2022).

Received: 2 March 2022.

Information about the authors

Roman E. Shcherbakov, post-graduate student, National Research Tomsk Polytechnic University; principal engineer, «Gazpromneft NTC».

Artem V. Kovalev, Cand Sc., associate professor, National Research Tomsk Polytechnic University.