

Паттерный анализ данных телеметрии скважины

М. И. Гудилов

Санкт-Петербургский государственный
электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

GudilovMI@yandex.ru

Н. А. Жукова

Санкт-Петербургский Федеральный
исследовательский центр
Российской академии наук

nazhukova@mail.ru

Аннотация. Представлен подход к обнаружению аномалий в работе ЭЦН при добыче нефти, основанный на кластеризации временных рядов и библиотеке обобщённых аномалий. Метод позволяет точно определять тип отклонения и оперативно реагировать на нештатные ситуации. Интеграция в SCADA-системы обеспечивает раннее предупреждение и способствует повышению надёжности и эффективности добычи.

Ключевые слова: кластеризация; телеметрия; DTW; временные ряды

I. ВВЕДЕНИЕ

На сегодняшний день механизированные способы добычи нефти являются основой функционирования подавляющего большинства нефтедобывающих скважин на территории Российской Федерации. Доля таких скважин превышает 90 %, что подчёркивает масштаб и важность задач, связанных с управлением режимами их работы. Среди различных технологий механизированной добычи особое место занимает эксплуатация с применением установок электроцентробежных насосов (УЭЦН), отличающихся широким диапазоном применимости. Однако, несмотря на технологические преимущества, эксплуатация УЭЦН сопряжена с определёнными ограничениями, связанными как с геологическими, так и с техническими факторами.

В частности, при наличии нестабильных или крайне малых дебитов, а также при осложнённых гидродинамических условиях (например, значительное содержание механических примесей, наличие свободного газа, высокие температуры или вязкость продукции) непрерывная работа оборудования становится неэффективной или даже невозможной. В подобных случаях на практике применяется режим периодического кратковременного включения (ПКВ), при котором насос функционирует циклами: период накопления сменяется периодом работы. Такой режим позволяет обеспечить стабильность процесса добычи на фоне нестабильных притоков.

Тем не менее, режим ПКВ характеризуется высокой чувствительностью к изменяющимся условиям эксплуатации. Повышенное поступление свободного газа или механических примесей в насос может вызывать нестабильность в работе оборудования, проявляющуюся в виде срывов подачи, турбинного вращения, клинов и других аномалий. Неоптимальный выбор продолжительности фаз накопления и работы способствует ускоренному износу оборудования, росту энергозатрат и снижению эффективности использования фонда скважин.

Традиционно анализ эффективности режима ПКВ выполняется специалистами вручную, что требует значительного времени, высокой квалификации и доступа к полным телеметрическим данным. При этом человеческий фактор, субъективность оценки и высокий объём обрабатываемой информации создают предпосылки для пропуска критически важных аномалий. В этих условиях актуальным становится внедрение автоматизированных методов анализа, способных выявлять отклонения от нормального режима в потоке данных, поступающем в режиме реального времени, и формировать обоснованные рекомендации по корректировке параметров работы скважин.

Одним из таких методов является применение алгоритма динамического выравнивания временных рядов (Dynamic Time Warping, DTW), позволяющего выявлять структурные аномалии в телеметрических сигналах, даже если они имеют различную длительность или временные сдвиги [1]. В сочетании с кластеризацией и построением библиотеки типичных аномалий DTW может стать основой интеллектуальной системы мониторинга, способной в автоматическом режиме анализировать поведение скважин, выделять участки с нетипичной динамикой и предлагать корректирующие действия. Настоящая работа посвящена разработке и практической реализации такого подхода, ориентированного на массовое применение в условиях действующих нефтепромыслов.

II. ФОРМИРОВАНИЕ БИБЛИОТЕКИ АНОМАЛИЙ

В рамках предложенного подхода автоматизированного анализа работы скважин в режиме периодического кратковременного включения (ПКВ) основное внимание уделяется двум взаимосвязанным задачам: формированию библиотеки типовых аномалий и последующему сопоставлению текущих данных телеметрии с этими эталонными паттернами. Реализация метода включает несколько последовательных этапов, начиная с обработки первичных сигналов и заканчивая алгоритмической идентификацией аномальных циклов работы оборудования путем сравнения с циклами работы скважины из библиотеки.

Первоначально производится сегментация сигнала загрузки погружного электродвигателя (ПЭД) на отдельные циклы – так называемые «зубцы», соответствующие рабочему циклу. При этом используются данные о частоте тока, представляющие собой сигнал прямоугольной формы (однополярный меандр) (рис. 1). Такая структура позволяет чётко выделить границы между фазами накопления и активной работы. Значения, лежащие между верхним и нижним

уровнями амплитуды, исключаются из анализа, а каждый зубец определяется как временной интервал между моментами начала и окончания активной работы. Подобная процедура позволяет устранить искажения, вызванные точками на участках пуска/останова насоса.

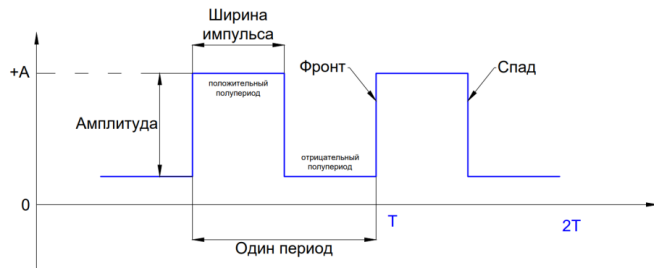


Рис. 1. Сигнал прямоугольной формы (меандр)

С учётом того, что сигналы телеметрии могут записываться с различной временной дискретностью, все выделенные зубцы дополнительно интерполируются и нормализуются, чтобы обеспечить единообразие масштаба и временной сетки (рис. 2).

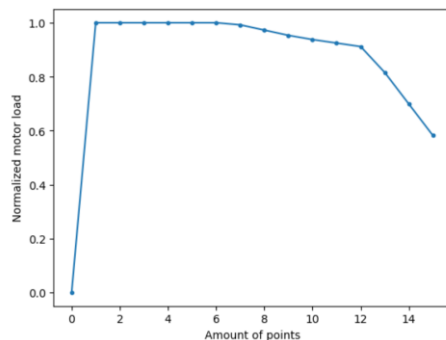


Рис. 2. Пример «зубца» после нормализации

Для решения задачи идентификации кластеров с аномальными рабочими циклами из исходного набора данных сигналов загрузки двигателя была выбрана KMeans кластеризация для временных рядов с использованием метрики DTW [2].

Окончательное количество кластеров определяется с использованием метрики оценки силуэта. Модель получает список возможных значений количества кластеров в качестве входных данных, выполняет кластеризацию для каждого значения и вычисляет оценку силуэта. Затем оптимальное количество кластеров выбирается на основе наивысшей полученной оценки силуэта. Кластеры, демонстрирующие аномальное поведение сигнала, выбираются экспертом, а средние значения кластера в каждой точке добавляются в библиотеку аномальных циклов.

III. ВЫЯВЛЕНИЕ АНОМАЛИЙ

Для анализа каждое новое включение насоса сравнивается с шаблонами из библиотеки с помощью метрики DTW, что позволяет точно определять тип отклонения и оперативно реагировать на нештатные ситуации.

Первый этап состоит в построение матрицы расстояния. На втором этапе строится матрица

трансформации D , каждый элемент которой вычисляется по следующей формуле:

$$D_{ij} = d_{ij} + \min \begin{cases} D_{i-1, j-1} \\ D_{i-1, j} \\ D_{i, j-1} \end{cases}$$

После заполнения матрицы трансформаций находится оптимальный путь трансформации, который минимизирует расстояние между временными рядами. Конечный элемент принимается, как общая стоимость выравнивания двух временных рядов.

Таким способом рассчитывается нормализованная стоимость выравнивания для всех аномалий в библиотеке. На основе предопределенного порога определяется сходство (да/нет) с соответствующей аномалией.

Данный метод анализа телеметрии позволяет обнаружить следующие типы отклонений в работе скважины:

- отклонения, вызванные свободным газом, влияющим на работу ЭЦН;
- заклинивание вала насоса, вызванное воздействием механических примесей на работу ЭЦН;
- высокий пусковой ток, необходимый для запуска двигателя.

Таким образом, предложенная модель сочетает в себе устойчивую сегментацию сигналов, мощные средства сравнения временных рядов и элементы машинного обучения в виде кластеризации, что обеспечивает её высокую применимость в задачах цифрового мониторинга нефтяных скважин.

IV. РЕЗУЛЬТАТ

В задаче поиска аномальных закономерностей алгоритмом DTW результаты получены после кластеризации 3767 рабочих периодов параметров загрузки двигателя для 63 скважин, работающих в режиме ПКВ. В рамках тестирования в алгоритм было заложено разбиение периодов от 4 до 8 и окончательный паттерн выбирался по наибольшему силуэтному баллу.

На основании табл. 1 получено, что оптимальными являются четыре кластера с различным поведением временного ряда с максимальным силуэтным баллом = 0,118.

ТАБЛИЦА 1. МЕТРИКА SILHOUETTE SCORE ДЛЯ РАЗЛИЧНОГО КОЛ-ВА КЛАСТЕРОВ

Number of clusters	Silhouette Score
4	0.107
5	0.105
6	0.118
7	0.037
8	0.049

Полученные кластеры показаны на рис. 3. Эти кластеры характеризуют то, как ведет себя нагрузка двигателя в большинстве случаев в нестабильных условиях эксплуатации.

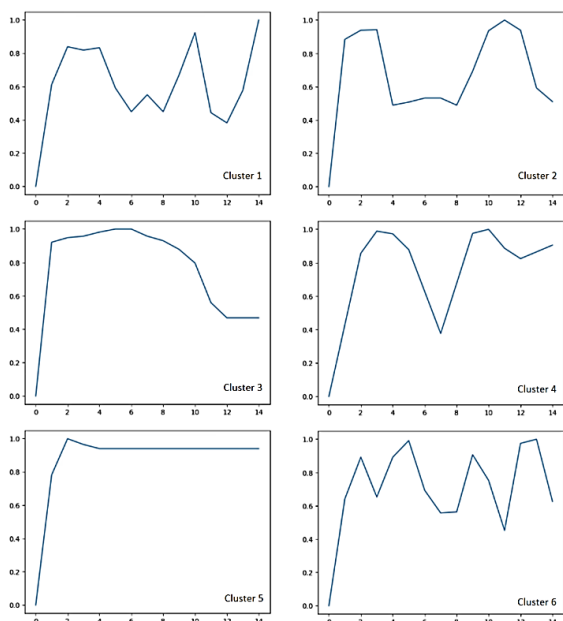


Рис. 3. Кластеры аномальной работы скважин по данным телеметрии загрузки двигателя

На рис. 4 показан пример применения алгоритма DTW. Маркером отмечены циклы со сбоями подачи из-за большого количества свободного газа на входе в насос.



Рис. 4. Кластеры аномальной работы скважин по данным телеметрии загрузки двигателя

V. ЗАКЛЮЧЕНИЕ

Разработанный подход к обнаружению аномалий в работе скважин, эксплуатируемых в режиме периодического кратковременного включения, демонстрирует высокую эффективность при анализе телеметрических данных. Использование алгоритма динамического выравнивания временных рядов в сочетании с кластеризацией и формированием библиотеки типичных аномалий позволяет точно идентифицировать нарушения в работе оборудования, связанные с геологическими и технологическими осложнениями.

Предложенная методика обеспечивает автоматическое сравнение каждого нового цикла работы насоса с эталонными шаблонами, что значительно сокращает временные и трудовые затраты на анализ данных, минимизирует влияние человеческого фактора и повышает оперативность реагирования на нештатные ситуации. Практическое применение алгоритма на данных 63 скважин подтвердило его работоспособность и способность выявлять ключевые типы отклонений, такие как сбой подачи из-за свободного газа, заклинивание вала и повышенные пусковые токи.

Интеграция разработанного решения в существующие SCADA-системы открывает широкие перспективы для масштабирования и промышленного внедрения, способствуя переходу к интеллектуальному управлению добычей и повышению общей надёжности нефтепромыслового оборудования.

СПИСОК ЛИТЕРАТУРЫ

- [1] Stan Salvador and Philip Chan Fast DTW: Toward Accurate Dynamic Time Warping in Linear Time and Space
- [2] https://tslearn.readthedocs.io/en/stable/gen_modules/clustering/tslearn.clustering.TimeSeriesKMeans.htm