

УДК 622.276.054.23

ПРИМЕНЕНИЕ ЭВРИСТИЧЕСКИХ АЛГОРИТМОВ В АНАЛИЗЕ ДАННЫХ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ДИАГНОСТИРОВАНИЯ ЭЛЕКТРОЦЕНТРОБЕЖНЫХ НАСОСНЫХ УСТАНОВОК

Валиахметов Рустам Илдарович¹,
valiakhmetovri@gmail.com

Ямалиев Виль Узбекович¹,
rusoil@mail.ru

Шубин Станислав Сергеевич¹,
shubinss@outlook.com

Алфёров Алексей Викторович²,
aleksey.alferov@gmail.com

¹ Уфимский государственный нефтяной технический университета,
Россия, 450062, г. Уфа, ул. Космонавтов, 1.

² Институт механики им. Р.Р. Мавлютова УНЦ РАН,
Россия, 450054, г. Уфа, пр. Октября, 71.

Обеспечение эффективного контроля и предотвращение отказов электроцентробежных насосных установок ввиду их широкого распространения является актуальной и востребованной задачей. Применение систем автоматизированного контроля электроцентробежных насосных установок позволяет повысить качество и скорость принимаемых решений о их техническом состоянии. Все методы диагностирования установок электроцентробежных насосов направлены на контроль состояния составных узлов и сводятся к анализу временных рядов, являющихся временными развёртками параметров эксплуатации. Традиционно применяемые линейные методы исследования временных рядов в последние десятилетия были существенно расширены нелинейными методами, среди которых значительное развитие получили эвристические алгоритмы.

Цель работы: повышение эффективности определения технического состояния установок электроцентробежных насосов в процессе эксплуатации.

Методы исследования. Предложенный в работе подход основывается на решении задачи диагностирования путём декомпозиции на следующие подзадачи: автоматическая сегментация, формализация и интерпретация полученных данных. Сегментация рассмотрена как задача кластеризации, цель которой – установление автокорреляционных связей между значениями временного ряда с формированием темпоральных кластеров и адаптивной аппроксимации в рамках установленных участков при сохранении локальных особенностей сигналов. Для каждого выделяемого класса отклонений работы электроцентробежных насосных установок сформированы решающие правила на основании экспертных знаний. Основными отличиями от классического подхода к задаче диагностирования являются: отсутствие необходимости участия эксперта при решении задачи кластеризации; обеспечение адаптивной аппроксимации в рамках выделенных временных участков; возможность реализации интерпретируемых подходов распознавания неисправностей.

Результаты. Предложен подход, основанный на решении задачи диагностирования путём её декомпозиции на следующие подзадачи: автоматическая сегментация, формализация и интерпретация полученных данных. Определён необходимый перечень контролируемых параметров эксплуатации электроцентробежных насосных установок, позволяющий реализовать процесс технического диагностирования.

Ключевые слова:

Насос, центробежный, диагностика, анализ, состояние, статистика, классификация.

Введение

В настоящее время решение задачи обеспечения интеллектуального контроля процесса эксплуатации скважин представляет большой интерес для специалистов производственных служб. Различные подходы к диагностике неисправностей нефтедобывающего оборудования, от аналитических методов до искусственных нейронных сетей, рассмотрены во множестве литературных источников [1–8].

Основным преимуществом установок электроцентробежных насосов (УЭЦН) по отношению к другим способам механизированной добычи является более высокая эффективность и широкий диапазон подач, но в связи с ограничениями и усло-

виями, накладываемыми конструкцией скважины к компоновке узлов УЭЦН [9, 10], им присущ ряд неисправностей, зачастую сопровождающих эксплуатацию, в качестве примера рассмотрим следующие:

- «клин» рабочих органов, что вызывает превышение номинальных значений токовых характеристик в погружном электродвигателе (ПЭД), т. е. «перегруз» ПЭД, следствием чего является повреждение шлицевой муфты (скручивание и слом вала);
- высвобождение растворенного в нефти газа в свободную газовую фазу в проточной части секции насосного агрегата (ЭЦН) и входного модуля (газосепаратора). Следствием является сни-

жение подачи насосного агрегата либо критический режим нарушения целостности потока флюида с образованием газовой пробки («срыв подачи»).

В связи с отсутствием эффективных подходов для диагностирования УЭЦН возникла необходимость разработки обобщённого алгоритма по определению её технического состояния. Предложенный в работе подход основывается на решении задачи диагностирования путём её декомпозиции на следующие подзадачи: автоматическая сегментация, формализация и интерпретация полученных данных, ниже приведена структурная схема предложенного подхода (рис. 1).

Описание предложенного решения

Основными отличиями от классического подхода к задаче диагностирования являются: отсутствие необходимости участия эксперта при решении задачи кластеризации; обеспечение адаптивной аппроксимации в рамках выделенных временных участков; возможность реализации интерпретируемых подходов распознавания неисправностей. В рамках предложенного подхода, с точки зрения теории машинного обучения, можно выделить два основных функциональных блока:

- формирование признаков (адаптивная сегментация и аппроксимация, укладка в алфавит и т. п.);
- решающий алгоритм (интерпретация на основании словаря и правил).

Далее более подробно рассмотрим основные элементы предложенного подхода.

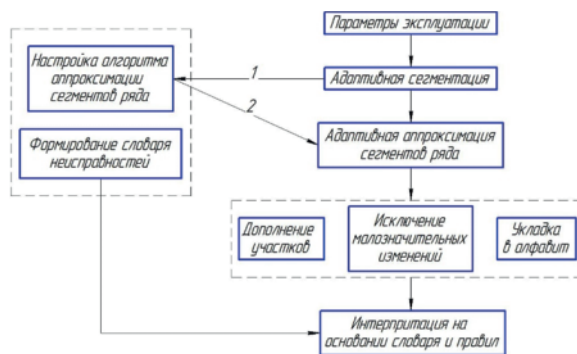


Рис. 1. Структурная схема алгоритма идентификации технического состояния УЭЦН

Fig. 1. Block diagram of the algorithm identifying the technical condition of Electrical Submersible Pump (ESP)

Параметры эксплуатации. В основу выбора параметров диагностирования положен принцип унификации набора параметров эксплуатации УЭЦН, что связано с различными комплектациями установок. Унифицированный набор данных включает в себя восемь технологических параметров: ток по трём фазам $I_{A,B,C}$; напряжение по трём фазам $U_{A,B,C}$; сопротивление изоляции R ; частота питающего тока F . Запись и сбор данных значений тока и частоты осуществляется на станции управления УЭЦН. К чи-

слу дополнительных параметров относятся такие, как замерной дебит жидкости – Q_{liq} ; давление на приёме – P_{int} ; температура жидкости на приёме – T_{int} ; температура масла электродвигателя – T_{motor} . Снятие изменений показаний температуры погружного электродвигателя, температуры и давления на приёме осуществляется при помощи погружного блока телеметрической системы, снятие замеров дебита осуществляется автоматической групповой замерной установкой.

Адаптивная сегментация. Представление сигнала посредством сегментации является хорошим инструментом для ранжирования изменений в контролируемой системе, однако в то же время разделение временного ряда на внутренние однородные сегменты является сложной задачей. Изменения во временном ряду обычно не локализованы на каком-либо конкретном моменте времени, т. е. невозможно определить строгие границы сегментов. Таким образом, задачу сегментации можно определить как задачу ограниченной кластеризации, т. е. значения должны быть сгруппированы на основе сходства, но с ограничением, что все точки в кластере должны происходить в последовательные моменты времени.

В рамках разработанного подхода был применён модифицированный алгоритм кластеризации *Gath–Geva*, предложенный и описанный в работах [11–17]. Особенностью данного алгоритма является то, что в нем реализована одновременная идентификация (локализация) на основании вероятностного метода главных компонент для установления однородных сегментов и элементы нечёткой логики для представления этих сегментов во временной области. Суть алгоритма сводится к идее, что данные могут быть смоделированы как линейная комбинация гауссовых функций, путём минимизации суммы взвешенных квадратов расстояний между исходными точками ($z_k = [t_k, x_k^T]^T$) и прототипом кластера (η_i). Алгоритм начинается с создания приближенного представления временного ряда, итерационно поглощая малозначительные сегменты до достижения критерия остановки.

Задан временной ряд T , принята конечная погрешность $\varepsilon > 0$ и инициализированные значения кластеров $W_i, v_i^x, \sigma_{i,x}^2, \mu_{i,k}$.

Шаг 1: расчёт $\eta_i = \{v_i^t, A_i, v_i^t, \sigma_{i,x}^2, \alpha_i\}$ – первоначальные параметры i -го кластера

- первоначальной вероятности кластера

$$\alpha_i = \frac{1}{N} \sum_{k=1}^N \mu_{i,k},$$

где $\mu_{i,k}$ – степень принадлежности наблюдения $z_k = [t_k, x_k^T]^T$ к i -у кластеру.

- центров кластеров

$$v_i^x = \frac{\sum_{k=1}^N (\mu_{i,k})^m (x_k - W_i(Y_{i,k}))}{\sum_{k=1}^N (\mu_{i,k})^m}.$$

- матрицы весов W_i

$$\tilde{W}_i = F_i W_i (\sigma_{i,x}^2 I + M_i^{-1} W_i^T F_i W_i),$$

$$F = \frac{\sum_{k=1}^N (\mu_{i,k})^m (x_k - v_i^x)(x_k - v_i^x)^T}{\sum_{k=1}^N (\mu_{i,k})^m} - \text{ковариационная матрица.}$$

- отклонение функции принадлежности $\sigma_{i,x}^2$

$$\sigma_{i,x}^2 = \frac{1}{q} \text{tr}(F_i - F_i W_i M_i^{-1} \tilde{W}_i^T).$$

- норма расстояния ($n \times n$ матрица) – гауссова функция

$$A_i = \sigma_{i,x}^2 I + \tilde{W}_i \tilde{W}_i^T.$$

- параметры модели: центр кластера и стандартное отклонение

$$v_i^t = \frac{\sum_{k=1}^N (\mu_{i,k})^m t_k}{\sum_{k=1}^N (\mu_{i,k})^m}, \quad \sigma_{i,x}^2 = \frac{\sum_{k=1}^N (\mu_{i,k})^m (t_k - v_i^t)^2}{\sum_{k=1}^N (\mu_{i,k})^m}.$$

Шаг 2: Вычисление $D^2(z_k, \eta_i)$ – мера обратно пропорциональная вероятности того, что (точка данных) принадлежит к i -му кластеру $p((z_k | \eta_i))$:

$$p(z_k | \eta_i) = \frac{1}{D_2(z_k, \eta_i)} = \alpha_i \frac{1}{\sqrt{2\pi\sigma_{i,t}^2}} \exp\left(-\frac{(t_k - v_i^t)^2}{2\sigma_{i,t}^2}\right) \times \frac{1}{(2\pi)^{\frac{r}{2}} \sqrt{\det(A_i)}} \exp\left(-\frac{1}{2}(x_k - v_i^x) A_i^{-1} (x_k - v_i^x)^T\right).$$

Шаг 3: Обновление матрицы степеней принадлежности данных z_k к i -му кластеру

$$\mu_{i,k}^{(l)} = \frac{1}{\sum_{j=1}^c \left(\frac{D(z_k, \eta_i)}{D(z_k, \eta_j)} \right)^{\frac{2}{m-1}}}$$

пока $\|U^{(l)} - U^{(l-1)}\| < \varepsilon$.

На рис. 2 представлен пример работы данного алгоритма – разделение кластеров проведено по

критерию равновероятной принадлежности точки к двум смежным кластерам.

Адаптивная аппроксимация сегментов ряда. Проведение анализа данных временных рядов характеризуется как сложная задача, особенно в контексте автоматизации процесса анализа. Основной проблемой является представление данных, оно же и ключевым моментом для эффективных решений. Наличие шума в сигнале усложняет задачу распознавания поведения УЭЦН; решением может служить анализ трендов контролируемых параметров. В данной работе применён метод представления информации о динамике процесса на основании анализа трендов.

Использование представления в виде трендов является естественным способом отображения для систем с потоковым поступлением данных. Среди основных преимуществ подобного представления можно выделить следующие:

- качественное описание – неопределённость, неполнота и неоднородность измерений делают качественное описание сигналов хорошим инструментом для решения задач контроля и управления;
- соотнесение изменений во времени;
- компактное представление знаний – учитывая большой объем информации, поступающей со станций управления, для анализа необходимо использовать только самую важную информацию.

В совокупности вышеописанные преимущества позволяют интерпретировать состояние системы с помощью экспертных правил, которые сформированы на основании опыта инженерно-технических работников. Одним из самых широко используемых способов представления временных рядов является кусочно-линейная аппроксимация. Большинство алгоритмов сегментации временных рядов могут быть сгруппированы в три категории:

- *Sliding Window Algorithm (SW)*, или алгоритм скользящего окна: сегмент увеличивается до тех пор, пока не превышает некоторую ошибку. Процесс повторяется со следующими точками временного ряда, не включёнными во вновь аппроксимированный сегмент.

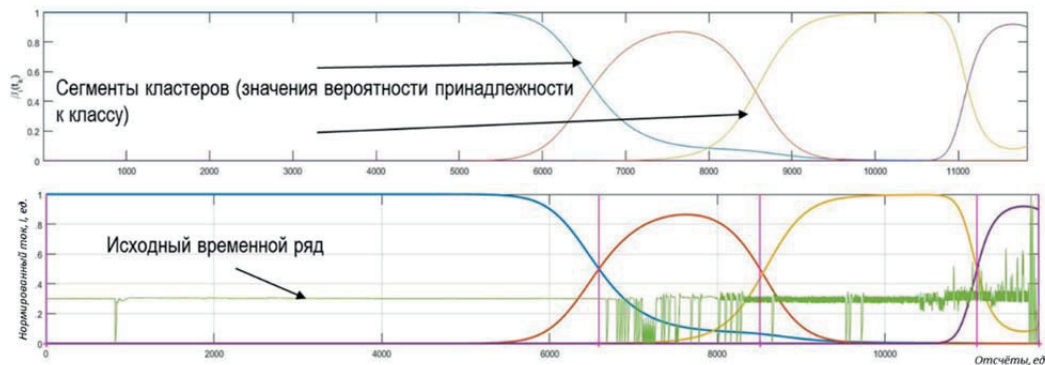


Рис. 2. Пример работы алгоритма кластеризации: зелёным – контролируемый параметр, вертикальные линии – границы кластеров, остальные кривые – вероятность принадлежности точки к конкретному кластеру

Fig. 2. Example of clustering algorithm: controlled parameter are in green, vertical lines are the clusters borders, other curves are the probability of belonging to a particular cluster point

- *The Top-Down Algorithm (TD)*, или алгоритм спуска сверху вниз: временные ряды рекурсивно «нарезаются» до тех пор, пока не удовлетворён некоторый критерий остановки.
- *The Bottom-Up Algorithm (BU)*, или алгоритм снизу–вверх: начиная от лучшего возможного приближения сегменты объединяются до тех пор, пока не будет достигнут некоторый критерий остановки.

Вышеописанные алгоритмы имеют ряд характерных недостатков, в полной мере описанных в ряде публикаций. В реализованном подходе был применён алгоритм *SWAB (Sliding Window And Bottom-up)*, фактически совмещающий в себе два алгоритма: *SW* и *BU*, совместное применение которых позволило снять их ограничения, алгоритм описан в работе [18–21].

При применении рассмотренного выше алгоритма к аппроксимации временных реализаций параметров эксплуатации УЭЦН было соблюдено условие сохранения информативности сигнала, т. е. сохранение локальных особенностей сигнала, вызванных влиянием осложняющих факторов. Для обеспечения такой возможности было проведено статистическое исследование различных по характеру поведения временных выборок, на основании которых был получен статистический критерий, позволяющий проводить настройку алгоритма *SWAB*. Наиболее информативным и обеспечивающим сепарабельность по «характеру» поведения сигнала является среднеквадратичное отклонение (СКО), прочие статистические моменты не проявили стабильность в уровнях показаний для характерных сигналов. В результате полученное значение СКО позволило задать уровень ошибки алгоритма *SWAB* для всего набора данных, использованных в исследовании.

На рис. 3 приведён пример аппроксимации различных по «поведению» сигналов с использованием алгоритма *SWAB* (нормированные значения потребляемого тока). Как видно из рис. 3, а, в сигнале присутствует «дребезг» контролируемого параметра, в свою очередь уровень ошибки принят та-

ким, что при аппроксимации не допускается избыточное разбиение на дополнительные участки, но при этом все основные особенности сигнала сохранены. На рис. 3, б присутствуют значительные отклонения в сигнале, т. е. хорошо локализованные «всплески», имеющие информационную ценность, для этого случая настройка алгоритма была принята такой, чтобы сохранить эти локальные особенности. Приведённые примеры хорошо иллюстрируют сильные стороны применённого алгоритма адаптивной аппроксимации.

Интерпретация на основании словаря и правил. В качестве конечного представления исходных данных сформирован алфавит (таблица), пример реализации приведён на рис. 4. Дополнительно к алфавиту при формировании слов используются числовые значения, характеризующие изменение амплитуды контролируемого параметра за рассматриваемый период.

Таблица. Описание примитивов для аппроксимирующих отрезков

Table. Description of primitives for approximating segments

Символ Symbol	Описание Description	Диапазон первой производной Range of the first derivative
n	горизонтальная прямая horizontal straight line	$[-0,05...0,05]$
l	восходящая прямая ascending straight line	$[0,05...0,15]$
d	нисходящая прямая descending straight line	$[-0,15...-0,05]$
U	быстрая восходящая прямая fast ascending straight line	$>0,15$
D	быстрая нисходящая прямая fast descending straight line	$<-0,15$
z	быстрая нисходящая прямая в 0 fast descending straight line to 0	$<-0,15$

На основании аппроксимации временной последовательности с учётом приведённых правил в

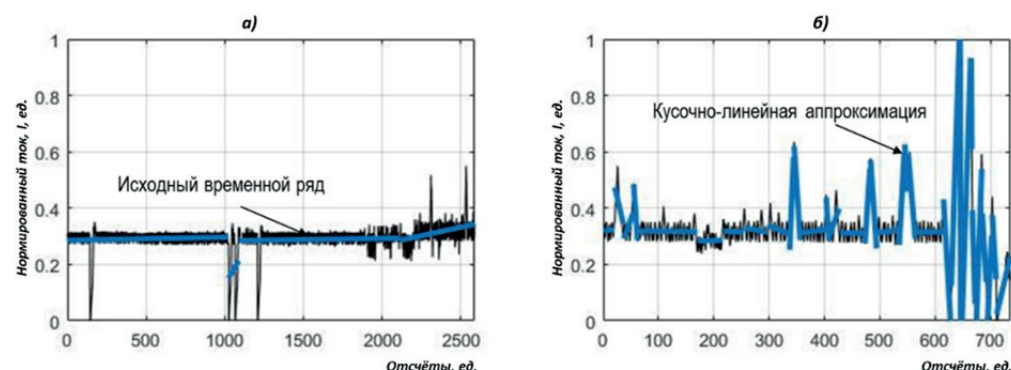


Рис. 3. Примеры различного поведения временных реализаций потребляемого тока: а) сигнал с «дребезгом» контролируемого параметра; б) сигнал с локализованными «всплесками»

Fig. 3. Examples of different behavior of temporal realizations of current consumption: а) signal with «bounce» of the controlled parameter; б) signal with localized «spikes»

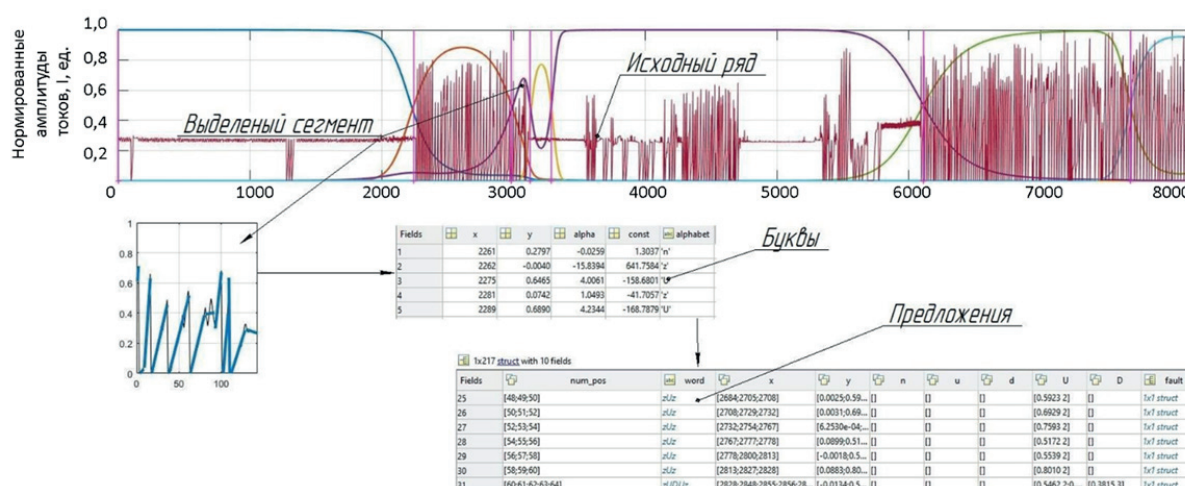


Рис. 4. Пример интерпретации трендов на основании предложенного подхода

Fig. 4 Example of interpretation of trends on the basis of the proposed approach

таблице формируется символьная (буквенно-числовая) интерпретация временной последовательности [22–24]. В качестве разделителей в полученных последовательностях используется буква «z», обозначающая отклонение установки и являющаяся признаком окончания предложений. Далее в рамках полученных предложений проводится объединение повторяющихся букв и удаление малозначительных изменений с целью сокращения размерности полученных выражений и как следствие сокращение словаря характерных шаблонов поведения сигналов (рис. 4).

На основании формирования запроса из словаря проводится поиск характерных комбинаций букв в полученных предложениях. Словарь является набором характерных типов шаблонов, свидетельствующих о потенциальном наличии неисправности, например, шаблон «zUz» характерен для запусков с последующей быстрой остановкой. Для установленных совпадений проводится запись максимального и минимального значений амплитуды, что в дальнейшем используется для интерпретации и поиска неисправностей, согласно разработанным правилам.

Обсуждение результатов

В качестве примера интерпретации рассмотрим следующее. На рис. 5 приведены результаты интерпретации показаний токов, полученных при работе УЭЦН в условиях выноса мехпримесей, отказ данной установки произошёл по причине слома вала в одной из секции насоса [21–22]. Как видно из графика, работа установки характеризуется наличием хорошо локализованных «всплесков» значений потребляемого тока.

Для формирования конечного заключения о техническом состоянии установки были использованы «правила» поиска неисправности. В рамках данного примера рассмотрим «правило», указы-

вающее на наличие механических примесей, ниже приведено его эвристическое описание:

- наличие механических примесей в рабочей среде сказывается на поведении значений потребляемого тока погружного электродвигателя, а именно в его временных реализациях зачастую встречаются «выбросы», что связано с «подклиниваем» насоса (рис. 5);
- в районе 11800 отсчёта (красная линия на графике) пусковое значение тока снизились более чем на 30 % от максимального значения, что характерно для случаев, сопряжённых с разрушением элементов трансмиссии (вала, шлицов).

Результатом применения вышеприведённого «правила» стала возможность локализации искомых характерных особенностей сигнала во временной реализации значений потребляемого тока УЭЦН. Для диагностирования иных неисправностей УЭЦН необходимо формирование «правил», отличных от приведённых, что является сложной и нетривиальной задачей. В продолжении работы предполагается модернизация процесса интерпретации путём замещения «строгих» правил на их «нечёткие» аналоги (нечёткая логика) либо анализ на основании математического аппарата искусственной нейронной сети.

Выводы

1. Обозначены основные способы реализации процесса контроля технического состояния установок электроцентробежных насосов, установлена необходимость декомпозиции задачи диагностирования на ряд подзадач, а именно автоматическая сегментация, формализация и интерпретация полученных данных об эксплуатации насосной установки.
2. Рассмотрена задача ранжирования изменений в контролируемой системе как задача ограни-

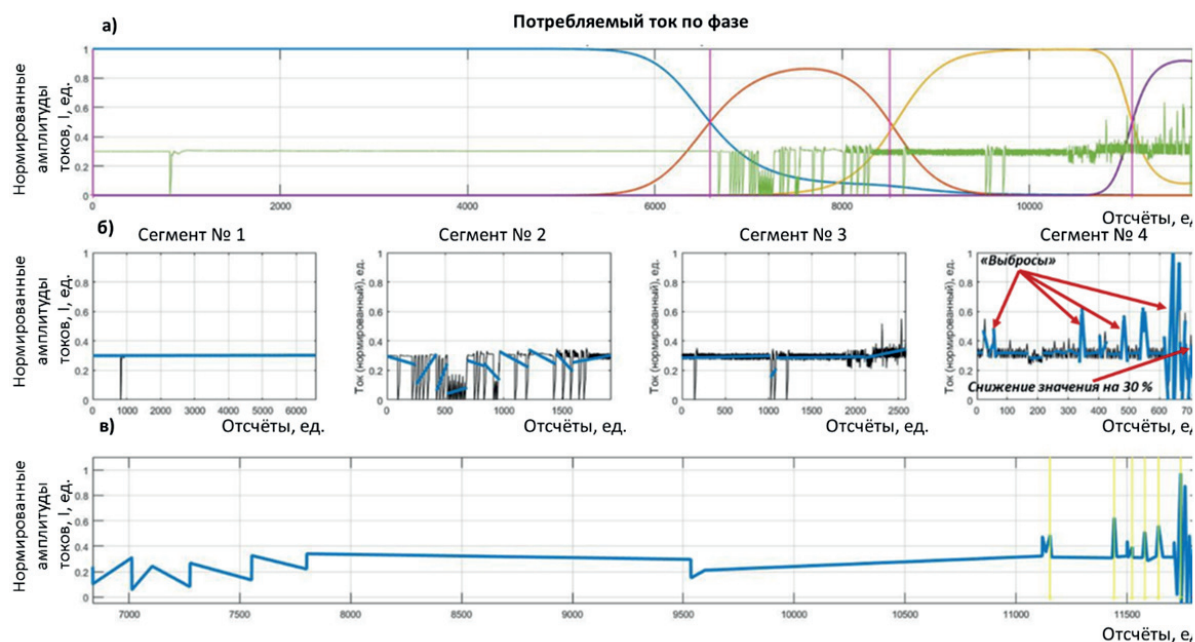


Рис. 5. Результаты анализа показаний потребляемого тока при влиянии механических примесей: желтые вертикальные линии – «подклинки», красная вертикальная линия – «слом вала» (частота проведения замеров составляет 0,3 Гц): а) исходные значения нормированного тока по фазе; б) аппроксимация значений тока в рамках выделенных сегментов; в) восстановленный сигнал после аппроксимации с локализацией искомых особенностей сигнала

Fig. 5. Results of analysis of current consumption readings under the influence of mechanical impurities: yellow vertical lines are the «wedging»; red vertical line is the «scrapped shaft» (the sampling frequency is 0,3 Hz): а) initial values of the normalized current phase; б) approximation of the current value within the selected segments; в) reconstructed signal after approximating the desired features with localization signal

- ченной кластеризации, позволившая применить подход на основании модифицированного метода Gath–Geva, в которой реализована одновременная идентификация установления гомогенных сегментов и элементов нечёткой логики для их представления в online-режиме.
3. Применённый и настроенный для работы с временными реализациями параметров эксплуатации УЭЦН метод аппроксимации (SWAB) позволил сохранить локальные особенности сигнала, вызванные влиянием осложняющих факторов.
 4. Определён необходимый перечень контролируемых параметров эксплуатации УЭЦН, позволяющий реализовать процесс технического диагностирования.
 5. Предложена символьная (буквенно-числовая) интерпретация временной реализации сигнала, позволившая формировать наборы паттернов

СПИСОК ЛИТЕРАТУРЫ

1. Ямалиев В.У., Салахов Т.Р., Шубин С.С. Применение элементов детерминированного хаоса к решению задач технического диагностирования УЭЦН. // Электронный научный журнал Нефтегазовое дело. – 2014. – № 4. – С. 174–191. URL: <http://ogbus.ru/article/primeneniye-elementov-teorii-determinirovannogo-chaosa-k-resheniyu-zadach-texnicheskogo-diagnosticsirovaniya-uchsnapplying-the-elements-of-theory-of-deterministic-chaos-for-the-problem-of-technical-diagnosi/> (дата обращения: 12.12.2017).

их поведения конкретным состояниям контролируемого объекта.

6. Предложенный в работе подход позволил реализовать следующие качественные улучшения по отношению к классическим способам обеспечения контроля за работой УЭЦН:
 - отсутствие необходимости участия эксперта при решении задачи кластеризации;
 - обеспечение адаптивной аппроксимации в рамках выделенных временных участков;
 - возможность реализации интерпретируемых подходов распознавания неисправностей.
7. Установлено, что имеется возможность улучшения процесса идентификации состояний УЭЦН путём замещения «строгих» правил на их «нечёткие» аналоги (нечёткая логика) либо путём использования математического аппарата искусственной нейронной сети.

2. Ямалиев В.У., Салахов Т.Р., Шубин С.С. Оценка технического состояния установок электроцентробежных насосов в процессе эксплуатации методом нейросетевой классификации // Нефтегазовое дело. – 2013. – № 114. – С. 102–109.
3. Устройство для оценки технического состояния установок электроцентробежных насосов в процессе эксплуатации: пат. 2525094 Рос. Федерация: МПК E21B 47/008 № 2013115609/03; заявл. 05.04.2013; опубл. 10.08.2014, Бюл. № 22. – 8 с.
4. Формирование расчетной модели рабочего процесса в ступени электроцентробежного насоса при стендовых испытаниях /

- В.У. Ямалиев, К.В. Литвиненко, Р.И. Валиахметов, С.С. Шубин, А.С. Лякин, И.И. Морозов // Вестник ЦКР Роснедра. – 2015. – № 2. – С. 28–35.
5. Боловин Е.В., Глазырин А.С. Метод идентификации параметров погружных асинхронных электродвигателей установок электроприводных центробежных насосов для добычи нефти // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2017. – Т. 328. – № 1. – С. 123–131.
6. Tkachuk R.Y., Glazyrin A.S., Polishchuk V.I. Induction Motor Drives Parameters Identification Using Genetic Algorithms // 7 International Forum on Strategic Technology (IFOST-2012). – Tomsk, September 18–21, 2012. – Tomsk: TPU Publ. house, 2012. – V. 2. – P. 586–589.
7. Оценка остаточного ресурса изоляции погружного электродвигателя установок электрических центробежных насосов добычи нефти при воздействиях импульсных перенапряжений / В.В. Сушков, В.В. Тимошкин, И.С. Сухачев, С.В. Сидоров // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2017. – Т. 328. – № 10. – С. 74–80.
8. Оптимальное управление многомашинным комплексом системы поддержания пластового давления на нефтяных месторождениях / В.В. Сушков, М.К. Велиев, В.В. Тимошкин, Т.Д. Гладких // Нефтяное хозяйство. – 2017. – № 2. – С. 82–84.
9. Богданов А.А. Погружные центробежные электронасосы. – М.: Гомптехиздат, 1957. – 375 с.
10. Богданов Е.А. Основы технической диагностики нефтегазового оборудования. – М.: Высшая школа, 2006. – 277 с.
11. Modified Gath–Geva clustering for fuzzy segmentation of multivariate time-series / J. Abonyi, B. Feil, N. Sandor, A. Peter // Fuzzy Sets and Systems. – 2005. – № 149. – P. 39–56.
12. Geva A.B. Hierarchical-fuzzy clustering of temporal-patterns and its application for time-series prediction // Pattern Recognition Letters. – 1999. – V. 20 (14). – P. 1519–1532.
13. Kun Yu, Tian Ran Lin, Ji Wen Tan. A bearing fault diagnosis technique based on singular values of EEMD spatial condition matrix and Gath–Geva clustering // Applied Acoustics. – 2017. – V. 121. – P. 33–45.
14. Jianchuan Yin, Nini Wang. An online sequential extreme learning machine for tidal prediction based on improved Gath–Geva fuzzy segmentation // Neurocomputing. – 2016. – V. 174. – P. A. – P. 85–98.
15. Abonyi J. Fuzzy Model Identification for Control. – Cambridge: Birkhauser, 2003. – 239 p.
16. Dobos L., Bankó Z., Abonyi J. Segmentation-based optimal experiment design // 10 International Symposium of Hungarian Researchers on Computational Intelligence and Informatics, CINTI 2009. – Hungary, 2009. – P. 279–288.
17. Integrated Process and Control System Model for Product Quality Control-Application to a Polypropylene Plant / B. Balasko, S. Nemeth, G. Nagy, J. Abonyi // Chemical Product and Process Modeling. – 2008. – V. 3. – P. 1934–2659.
18. Keogh E., Pazzani M.J. An enhanced representation of time series which allows fast and accurate classification clustering and relevance feedback // 4 Int. Conf. on KDD. – USA, 1998. – P. 239–243.
19. An online algorithm for segmenting time series / E. Keogh, S. Chu, D. Hart, M. Pazzani // IEEE Internat. Conf. on Data Mining. – 2001. URL: <https://pdfs.semanticscholar.org/14e8/6f39831e30b4037ab99b5de5e5d86608ea16.pdf> (дата обращения: 12.12.2017).
20. Hu B., Chen Y., Keogh E. Classification of streaming time series under more realistic assumptions // Data Mining and Knowledge Discovery. – 2016. – V. 30 (2). – P. 403–437.
21. Dau H.A., Begum N., Keogh E. Semi-Supervision Dramatically Improves Time Series Clustering under Dynamic Time Warping // Proc. of the 25 ACM International Conference on Information and Knowledge Management. – USA, 2016. – P. 999–1008.
22. Beyond one billion time series: indexing and mining very large time series collections with iSAX2+ / A. Camerra, J. Shieh, T. Palpanas, T. Rakthanmanon, E. Keogh // Knowledge and information systems. – 2014. – V. 39 (1). – P. 123–151.
23. iSAX 2.0: Indexing and mining one billion time series / A. Camerra, T. Palpanas, J. Shieh, E. Keogh // Data Mining (ICDM), 2010 IEEE 10 International Conference. – Australia, 2010. – P. 58–67.
24. Experiencing SAX: a novel symbolic representation of time series / J. Lin, E. Keogh, L. Wei, S. Lonardi // Data Mining and knowledge discovery. – 2007. – V. 15 (2). – P. 107–144.

Поступила 21.12.2017 г.

Информация об авторах

Валиахметов Р.И., соискатель кафедры машин и оборудования нефтегазовых промыслов Уфимского государственного нефтяного технического университета.

Ямалиев В.У., доктор технических наук, профессор кафедры машин и оборудования нефтегазовых промыслов Уфимского государственного нефтяного технического университета.

Шубин С.С., кандидат технических наук, соискатель кафедры машин и оборудования нефтегазовых промыслов Уфимского государственного нефтяного технического университета.

Алфёров А.В., соискатель Института механики им. Р.Р. Мавлютова УНЦ РАН,

УДК 622.276.054.23

APPLICATION OF HEURISTIC ALGORITHMS IN ANALYZING DATA TO SOLVE THE PROBLEM OF DETECTION OF ELECTRIC CENTRIFUGAL PUMPING UNITS

Rustam I. Valiakhmetov¹,
valiakhmetovri@gmail.com

Vil U. Yamaliev¹,
rusoil@mail.ru

Stanislav S. Shubin¹,
shubinss@outlook.com

Aleksey V. Alferov²,
aleksey.alferov@gmail.com

¹ Ufa State Petroleum Technological University,
1, Kosmonavtov Street, Ufa, 450062, Russia.

² Mavlyutov Institute of Mechanics,
71, Oktyabrya avenue, Ufa, 450054, Russia.

Ensuring effective control and preventing failures of electrical submersible pumps, because of their wide distribution is the relevant and demanded task. The use of automated control systems of electric centrifugal pumping units allows improving the quality and speed of decisions made about their technical condition. All methods of diagnosing the installation of electrical submersible pumps are aimed at monitoring the state of the composite nodes, and are reduced to the analysis of time series, which are the time scans of the operation parameters. Traditionally applied linear methods of time series research in the last decades have been substantially expanded by nonlinear methods, among which heuristic algorithms were developed significantly.

The main aim is to increase the efficiency of determining the technical state of installations of electric centrifugal pumps during operation.

Methods. The approach proposed in this paper is based on solving the diagnostic problem by decomposition into the following subtasks: automatic segmentation, formalization and interpretation of the data obtained. Segmentation is considered as a clustering problem, the purpose of which is the establishment of autocorrelation links between the values of the time series with the formation of temporal clusters and adaptive approximation within the established areas while preserving the local features of the signals. For each allocated class of deviations in operation of electric centrifugal pumping units, the decisive rules are formed based on expert knowledge. The main differences from the classical approach to the problem of diagnosis are: the lack of the need for expert participation in solving the clustering problem; providing adaptive approximation within the allocated time intervals; the possibility of implementing interpretable approaches to fault recognition.

Results. The authors have proposed the approach based on solving the diagnostic problem by its decomposition into the following subtasks: automatic segmentation, formalization and interpretation of the data obtained. The necessary list of controlled parameters of operation of the electric centrifugal pumping units is determined. The list allows technical diagnosis.

Key words:

Pump, centrifugal, diagnostics, analysis, status, statistics, classification.

REFERENCES

1. Yamaliev V.U., Salakhov T.R., Shubin S.S. Primenenie elementov teorii determinirovannogo khaosa k resheniyu zadach tekhnicheskogo diagnostirovaniya UECN [Application of the theory of determined chaos for solving the task of ESP technical diagnosis]. *Elektronnyy nauchnyy zhurnal Neftegazovoe delo*, 2014, no. 4, pp. 174–191. Available at: <http://ogbus.ru/article/primenenie-elementov-teorii-determinirovannogo-khaosa-k-resheniyu-zadach-tekhnicheskogo-diagnostirovaniya-uecn> applying-the-elements-of-theory-of-deterministic-chaos-for-the-problem-of-technical-diagnosi/ (accessed 12 December 2017).
2. Yamaliev V.U., Salakhov T.R., Shubin S.S. Otsenka tekhnicheskogo sostoyaniya ustanovok elektrosentrobezhnykh nasosov v protsesse ekspluatatsii metodom neyrosetevoy klassifikatsii [Estimation of technical state of ESP units in operation by neural network classification]. *Neftegazovoe delo*, 2013, no. 114, pp. 102–109.
3. Yamaliev V.U., Salakhov T.R., Shubin S.S. *Ustroystvo dllya otsenki tekhnicheskogo sostoyaniya ustanovok jelektrotsentrobezhnykh nasosov v protsesse jekspluatatsii* [Device for estimating technical state of ESP units in operation]. Patent RF, no. 2013115609/03, 2013.
4. Yamaliev V.U., Litvinenko K.V., Valiakhmetov R.I., Shubin S.S., Lyaskin A.S., Morozov I.I. Formirovanie raschetnoy modeli rabocheho protsessa v stupeni elektrosentrobezhnogo nasosa pri stendovykh ispytaniyakh [Formation of design model of operational process in electrical submersible pump step at bench tests]. *Vestnik CKR Rosnedra*, 2015, no. 2, pp. 28–35.
5. Bolovin E.V., Glazyrin A.S. Method for identifying parameters of submersible induction motors of electrical submersible pump units for oil production. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2017, vol. 328, no. 1, pp. 123–131.
6. Tkachuk R.Y., Glazyrin A.S., Polishchuk V.I. Induction Motor Drives Parameters Identification Using Genetic Algorithms. 7th International Forum on Strategic Technology (IFOST-2012). Tomsk, September 18–21, 2012. Tomsk, TPU Publ. house, 2012. Vol. 2, pp. 586–589.

7. Sushkov, V.V., Timoshkin, V.V., Sukhachev, I.S., Sidorov, S.V. Evaluation of submersible electric motor insulation residual resource in oil production electric centrifugal pumps under the influence of impulse overvoltages. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2017, vol. 328, no. 10, pp. 74–80. In Rus.
8. Sushkov V.V., Veliev M.K., Timoshkin V.V., Gladkih T.D. Optimum control for multiple computer complex of oil field formation pressure maintenance system. *Oil Industry*, 2017, no. 2, pp. 82–84. In Rus.
9. Bogdanov A.A. *Pogruzhnye ctsentrobezhnye jelektronasosy* [Electrical submersible pumps]. Moscow, Gotoptehizdat Publ., 1957. 375 p.
10. Bogdanov E.A. *Osnovy tekhnicheskoy diagnostiki neftegazovogo oborudovaniya* [Bases of technical diagnostics of oil-and-gas equipment]. Moscow, Vysshaya shkola Publ., 2006. 277 p.
11. Abonyi J., Feil B., Sandor N., Peter A. Modified Gath–Geva clustering for fuzzy segmentation of multivariate time-series. *Fuzzy Sets and Systems*, 2005, no. 149, pp. 39–56.
12. Geva A.B. Hierarchical-fuzzy clustering of temporal-patterns and its application for time-series prediction. *Pattern Recognition Letters*, 1999, vol. 20 (14), pp. 1519–1532.
13. Kun Yu, Tian Ran Lin, Ji Wen Tan. A bearing fault diagnosis technique based on singular values of EEMD spatial condition matrix and Gath–Geva clustering. *Applied Acoustics*, 2017, vol. 121, pp. 33–45.
14. Jianchuan Yin, Nini Wang. An online sequential extreme learning machine for tidal prediction based on improved Gath–Geva fuzzy segmentation. *Neurocomputing*, 2016, vol. 174, P. A, pp. 85–98.
15. Abonyi J. *Fuzzy Model Identification for Control*. Cambridge, Birkhauser, 2003. 239 p.
16. Dobos L., Bankó Z., Abonyi J. Segmentation-based optimal experiment design. *10th International Symposium of Hungarian Researchers on Computational Intelligence and Informatics, CINTI 2009*. Hungary, 2009. pp. 279–288.
17. Balasko B., Nemeth S., Nagy G., Abonyi J. Integrated Process and Control System Model for Product Quality Control-Application to a Polypropylene Plant. *Chemical Product and Process Modeling*, 2008, vol. 3, pp. 19342–659
18. Keogh E., Pazzani M.J. An enhanced representation of time series which allows fast and accurate classification clustering and relevance feedback. *4th Int. Conf. on KDD*. USA, 1998. pp. 239–243.
19. Keogh E., Chu S., Hart D., Pazzani M. An online algorithm for segmenting time series. *IEEE Internat. Conf. on Data Mining*, 2001. Available at: <https://pdfs.semanticscholar.org/14e8/6f39831e30b4037ab99b5de5e5d86608ea16.pdf> (accessed 12 December 2017).
20. Hu B., Chen Y., Keogh E. Classification of streaming time series under more realistic assumptions. *Data Mining and Knowledge Discovery*, 2016, vol. 30 (2), pp. 403–437.
21. Dau H.A., Begum N., Keogh E. Semi-Supervision Dramatically Improves Time Series Clustering under Dynamic Time Warping. *Proc. of the 25th ACM International on Conference on Information and Knowledge Management*. USA, 2016. pp. 999–1008.
22. Camerra A., Shieh J., Palpanas T., Rakthanmanon T., Keogh E. Beyond one billion time series: indexing and mining very large time series collections with iSAX2+. *Knowledge and information systems*, 2014, vol. 39 (1), pp. 123–151.
23. Camerra A., Palpanas T., Shieh J., Keogh E. iSAX 2.0: Indexing and mining one billion time series. *Data Mining (ICDM)*, 2010 IEEE 10th International Conference. Australia, 2010. pp. 58–67.
24. Lin J., Keogh E., Wei L., Lonardi S. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and knowledge discovery*, 2007, vol. 15 (2), pp. 107–144.

Received: 21 December 2017.

Information about the authors

Rustam I. Valiakhmetov, graduate student, Ufa State Petroleum Technological University.

Vil U. Yamaliev, Dr. Sc., professor, Ufa State Petroleum Technological University.

Stanislav S. Shubin, Cand. Sc., postdoc, Ufa State Petroleum Technological University.

Aleksey V. Alferov, graduate student, Mavlyutov Institute of Mechanics.