

**ПАРОТУРБИННЫЕ, ГАЗОТУРБИННЫЕ,  
ПАРОГАЗОВЫЕ УСТАНОВКИ  
И ИХ ВСПОМОГАТЕЛЬНОЕ ОБОРУДОВАНИЕ**

**МОДЕЛЬ РАННЕГО ОБНАРУЖЕНИЯ АВАРИЙНЫХ СИТУАЦИЙ  
НА ОБОРУДОВАНИИ ЭЛЕКТРОСТАНЦИЙ НА ОСНОВЕ МЕТОДА  
НАИМЕНЬШИХ ПОТЕНЦИАЛОВ**

© 2021 г. А. А. Коршикова<sup>а</sup>, \*, О. М. Идзон<sup>а</sup>

<sup>а</sup>ООО “Инконтрол”, ул. Ленинская Слобода, д. 23, стр. 2, Москва, 115280 Россия

\*e-mail: korshikova@inctrl.ru

Поступила в редакцию 02.12.2020 г.

После доработки 12.01.2021 г.

Принята к публикации 20.01.2021 г.

Рассматривается метод обнаружения и предсказания аномальности функционирования оборудования энергоблоков на примере газотурбинной установки. Формулируется задача обнаружения аномальности функционирования как математическая задача моделирования критерия аномальности, принимающего значения от 0 до 1. Сделано предположение о возможности и эффективности применения методов предиктивной аналитики для прогнозирования будущего состояния технологического оборудования на основе существующего парка измерений без какого бы то ни было его увеличения. Предполагается, что, даже когда каждое отдельное измерение не выходит из диапазона, принятого за диапазон нормального функционирования, по их совокупной динамике можно судить о развивающемся дефекте, т.е. о переходе диагностируемого технологического оборудования (ДТО) в аномальную зону функционирования. Для решения поставленной задачи предложен подход, основанный на вычислении значения “показателя аномальности”, который можно интерпретировать как условный потенциал, создаваемый точками в многомерном пространстве признаков, характеризующих состояние оборудования в данный момент времени. Обучая модель на признаках, задающих области состояний (состояние нормальной эксплуатации и состояния для разного рода фиксированных дефектов), можно затем применить обученную модель для определения типа состояния: чем ближе значение показателя аномальности к значениям, присущим той или иной области функционирования, тем больше вероятность того, что состояние ДТО соответствует этой области. Показано, что в связи с определенными объективными обстоятельствами отсутствует практическая возможность обучения модели на данных, которые получены в периоды аномальной эксплуатации с конкретными видами дефектов ДТО. Это сводит задачу к адаптации метода к ситуации, когда в качестве обучающей имеется только область нормальной эксплуатации. Предложенная модель была обучена и протестирована в период нормального функционирования оборудования. Результат тестирования позволяет говорить о непротиворечивости предложенного метода (отсутствии ложноположительных реакций).

*Ключевые слова:* технологическое оборудование, обнаружение аномалий, предиктивная аналитика, метод наименьших потенциалов, области нормального и аномального функционирования, пространство состояний, показатели функционирования

**DOI:** 10.1134/S0040363621090034

Во время работы технологического оборудования неизбежно происходят события, которые могут оказать негативное влияние на его производительность или даже привести к выходу из строя. Модель, предсказывающая будущую аварийную ситуацию, позволила бы своевременно предотвратить ее развитие и повысить тем самым эффективность использования технологического оборудования. Построением и исследованием таких моделей занимается предиктивная аналитика [1, 2].

Главная идея предиктивной аналитики в энергетике, по мнению авторов, состоит в том, что

возникновение аварии можно с некоторой вероятностью предсказать, основываясь на непрерывном анализе совокупности данных, которые характеризуют функционирование диагностируемого оборудования и измеряются штатными средствами контроля.

Для определения состояния технологического оборудования применяют различные методы классификации [3], базирующиеся на многоклассовой и одноклассовой классификации.

Главная идея методов, основанных на многоклассовой классификации, состоит в построении

классификатора нормальных и аномальных данных. В качестве примеров таких методов можно привести:

нейросетевые классификаторы [4];

статистические классификаторы, в частности байесовский подход, анализ распределений (применительно к оборудованию энергоблока рассматривается в статьях [5, 6]);

машинное обучение (деревья решений, SVM и пр.) [7].

Методы, в которых обучение базируется на прецедентах, предоставляют возможность диагностировать конкретные типы аварий, но их главный недостаток – это сложность, а часто и невозможность получения данных о возникших на энергоблоке дефектах и произошедших авариях. Поэтому для первичного выявления возможных отклонений в состоянии технологического оборудования перспективно применение методов, обучение которых не требует данных о прецедентах.

Идея методов, основанных на одноклассовой классификации, состоит в построении границы области нормальных данных. Все данные, лежащие вне этой области, считаются аномальными. Примерами таких методов могут быть:

нейросетевые классификаторы (автоэнкодеры и пр.) [8];

машинное обучение (one-class SVM, one-class Fisher Discriminant, MSET [9] и пр.).

Однако приведенные методы довольно сложны в реализации при высокой неоднозначности результата для энергетического оборудования.

Предиктивная диагностика, которая ориентирована на одноклассовую классификацию, может приводить к технической некорректным результатам анализа конкретной ситуации, так как удаление данных по этой ситуации от границ известной (ранее определенной) области нормальной эксплуатации не обязательно будет входом в реальную область аномальных данных и, соответственно, сигналом об аварийной ситуации. Возможно, это будет входом в ранее не зафиксированную нормальную область. Поэтому в такой системе возможны ложные сообщения о дефектной ситуации и для уменьшения их вероятности крайне желательно “дообучение” системы на верифицированных данных, которые имело диагностируемое оборудование в выбранный определенный период его работы с известными дефектами (далее обучение на периоде).

Описанный в настоящей статье метод предназначен для диагностики технологического оборудования энергоблоков, оснащенных средствами индивидуальных измерений, которые, с одной стороны, могут непосредственно диагностировать возникновение дефектов/неисправностей, но, с другой стороны, оказываются не в состоянии об-

наружить дефекты на ранних стадиях их развития, а также тогда, когда эти дефекты вообще не проявляются при индивидуальных измерениях вплоть до возникновения аварии. Метод дает шанс на раннее обнаружение дефектов с помощью анализа как самих измеряемых параметров в их совокупности, так и статистических производных этих параметров. А в частном случае метод не требует обязательного наличия верифицированных периодов ДТО с известными неисправностями/дефектами, т.е. обучение может проводиться только на нормальных периодах работы ДТО. Проверка адекватности модели при таком подходе заключается в том, что в результате ее испытания на другом (необучающем) периоде нормальной эксплуатации критерий аномальности не должен выходить за заданные уставки, т.е. не выдавать ложноположительных прогнозов аномальности.

## ПОСТАНОВКА ЗАДАЧИ

Исходными данными служат архивы программно-технических комплексов автоматизированных систем управления энергоблоков, хранящие массивы данных за длительный период работы ДТО:  $\mathbf{P}(t) = [P_1(t), \dots, P_m(t)]^T$ , где  $\mathbf{P}(t)$  – вектор показателей в момент времени  $t$ ;  $P_1(t), \dots, P_m(t)$  – показатели функционирования ДТО в момент времени  $t$ ;  $m$  – количество этих показателей. Под показателями функционирования ДТО далее понимаются не только индивидуальные измерения, но и их статистические производные (средние значения за период, дисперсии, корреляционные функции и др.). Векторы  $\mathbf{P}(t_r)$ , взятые в моменты времени  $t_r$  ( $r = 1, 2, \dots$ ), образуют точки в  $m$ -мерном пространстве. Кроме того, дано конечное множество прецедентов (аварий, неисправностей, дефектов), каждый из которых можно классифицировать, т.е. отнести к определенному виду дефекта ДТО. Каждому прецеденту приписана метка времени его обнаружения  $t_{i,j}$ , здесь  $t_{i,j}$  – моменты обнаружения  $i$ -го прецедента класса (типа дефекта)  $j$  ( $j = 1, \dots, l$ ;  $l$  – количество классов;  $i = 1, \dots, k$ ;  $k$  – количество прецедентов в классе  $j$ ).

Можно предположить, что область пространства с точками нормального функционирования ДТО отличается от областей пространства, которые образованы точками функционирования этого же объекта в периоды времени, совпадающие с моментами возникновения прецедента  $t_{i,j}$ , так же как различаются между собой и области пространства прецедентов разных классов. Более того, для технологического оборудования энергоблоков вполне обоснованно считать эти множества линейно разделимыми. Период функционирования оборудования длительностью от  $(t_{i,j} - \Delta t_{b,j})$  до  $(t_{i,j} + \Delta t_{a,j})$  можно назвать аномальным периодом эксплуата-

ции в классе  $j$ . Здесь  $\Delta t_{b,j}$  – период времени до прецедента, когда наличие развивающегося дефекта начинает оказывать влияние на совокупность признаков функционирования ДТО;  $\Delta t_{a,j}$  – период времени от возникновения прецедента до устранения его последствий.

Совокупность всех имеющихся описаний прецедентов называется обучающей выборкой. Требуется по ней выявить общие зависимости, позволяющие определять аномальное функционирование ДТО с указанием класса аномальности. При этом должен быть сформирован обобщенный критерий аномальности, по изменению которого можно с некоторой достоверностью определять возможность возникновения прецедента.

### ПОСТРОЕНИЕ МОДЕЛИ ДЛЯ ОБЩЕГО СЛУЧАЯ

Пусть даны несколько множеств из  $n$ -мерного евклидова пространства признаков состояния ДТО. Каждая точка этого пространства образована значениями признаков, полученными через определенные отрезки времени. Первое множество точек составлено в период исправного состояния оборудования. Другие множества образованы в аномальные периоды работы оборудования, причем каждое из них связано с определенным видом аномальности, характеризующим конкретную неисправность. Далее множества нормальной и аномального функционирования называются базовыми множествами. На рис. 1 изображены в двумерном пространстве ( $P_1$  и  $P_2$  – 1-я и 2-я координаты вектора) расстояния от точки процесса в текущий момент времени  $t$  до точек подмножеств нормального (индекс “норм”) и аномального (индексы “a1” и “a2”) функционирования ДТО;  $N_{\text{норм}}, M_{a1}, M_{a2}$  – количество точек в подмножествах нормального и аномального функционирования ДТО, а значения  $N_{\text{норм}}, M_{a1}, M_{a2}$  соответствуют длительности нормального и аномальных периодов.

С каждой точкой из базовых множеств можно связать некоторую функцию, аналогичную по форме электрическому потенциалу, т.е. максимальную в этой точке и убывающую по всем направлениям от нее (точка таким образом окажется как бы источником потенциала). Такой функцией может быть, например,

$$\varphi(R) = \frac{W}{1 + R^2},$$

где  $W$  – “вес” точки;  $R$  – расстояние между точкой-источником и точкой, в которой вычисляется потенциал, которое рассчитывается по формуле

$$R(t) = \sqrt{\sum_{i=1}^m [P(t)_i - P_i]^2};$$

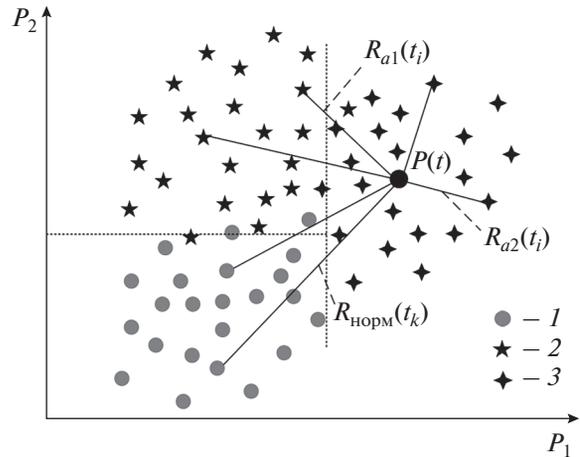


Рис. 1. Пространство параметров функционирования оборудования.

Обучающее множество – точки подмножества функционирования ДТО: 1 – нормального; 2 – аномального с типом дефекта 1; 3 – аномального с типом дефекта 2. Пунктирные линии иллюстрируют линейную делимость нормальной и аномальных областей

здесь  $P(t)_i$  –  $i$ -я координата (признак ДТО)  $t$ -й точки  $m$ -мерного множества (точки в момент времени  $t$ ), относительно которой определяется потенциал  $\varphi(t)$  точки с координатами  $P_i$ .

Значение функции  $\varphi(t)$  в каждой точке пространства можно считать мерой близости этой точки к точке-источнику.

Пусть источниками служат точки пространства нормального функционирования. Тогда средний потенциал, создаваемый в данной точке пространства всеми точками, т.е. суммарный потенциал, деленный на число точек (потенциал нормального функционирования), будет характеризовать близость данной точки ко всему базовому множеству в целом:

$$\varphi_{\text{норм}}^{\text{cp}} = \frac{1}{N_{\text{норм}}} \sum_{t=1}^{N_{\text{норм}}} \varphi[R(t)].$$

Аналогично можно определить в этой же точке пространства потенциалы  $j$ -го базового множества, где  $j = 1, \dots, l$  – тип неисправности [ $l$  – количество типов неисправностей (классов дефектов) в обучающей выборке]:

$$\varphi_j^{\text{cp}} = \frac{1}{M_{a j}} \sum_{t=1}^{M_j} \varphi[R(t)].$$

Здесь  $M_{a j}$  – количество точек, образующих базовое множество аномального функционирования с дефектом класса  $j$ .

Поскольку множества нормального и аномального функционирования при расчете счита-

ются линейно разделимыми, естественно будет отнести точку к тому базовому множеству, чей потенциал в этой точке наибольший.

Однако, как уже было отмечено, получить данные по аномальным классам крайне затруднительно, а чаще всего вообще невозможно. Поэтому в дальнейшем будет рассмотрена модификация изложенного общего подхода, а именно построение модели только по данным нормального периода эксплуатации.

### ПОСТРОЕНИЕ МОДЕЛИ ПО ДАННЫМ ТОЛЬКО НОРМАЛЬНОГО ПЕРИОДА ЭКСПЛУАТАЦИИ

В этом частном случае имеется лишь одно базовое множество – множество нормального функционирования  $N_{\text{норм}}$  с количеством точек  $N_{\text{норм}}$ . Тогда алгоритм обучения модифицируется следующим образом. Для каждой  $k$ -й точки базового множества, где  $k \in (1, N_{\text{норм}})$ , нужно вычислить ее средний потенциал относительно остальных  $N_{\text{норм}} - 1$  точек:

$$\varphi_k^{\text{cp}} = \frac{1}{N_{\text{норм}} - 1} \sum_{t=1}^{N_{\text{норм}}} \varphi[R(t)_k],$$

где  $t \in (1, N_{\text{норм}})$ ;  $t \neq k$ ;  $R(t)_k$  – расстояние от  $t$ -й до  $k$ -й точки.

Далее рассчитывается минимальный потенциал базового множества  $N_{\text{норм}}$

$$\varphi^{\text{min}} = \min[\varphi_k^{\text{cp}}, k \in (1, N_{\text{норм}})].$$

Тогда в режиме run-time потенциал в каждый новый момент времени будет определяться по формуле:

$$\varphi^{\text{cp}}(t) = \frac{1}{N_{\text{норм}}} \sum_{i=1}^{N_{\text{норм}}} \varphi[R(t)_i],$$

где  $t$  – текущий момент времени;  $R(t)_k$  – расстояние от точки, образованной значениями признаков в текущий момент времени  $t$ , до  $k$ -й точки в обучающем периоде.

На этом построение модели можно считать законченным. Теперь, если окажется, что потенциал  $\varphi^{\text{cp}}(t)$  для точек, образованных текущими значениями параметров ДТО относительно точек обучающего множества, меньше, чем значение  $\varphi^{\text{min}}$ , можно считать, что ДТО вошло в аномальную зону функционирования. Отслеживая тенденцию дальнейшего уменьшения  $\varphi^{\text{cp}}(t)$ , можно прогнозировать возможность возникновения аварии. Вычисленный потенциал изменяется в пределах  $(1, 0)$  и, по сути, является критерием аномальности.

При реальной эксплуатации модели можно задать граничную уставку потенциала и такой период времени  $T_a$ , что при  $\varphi(t) < (\varphi^{\text{min}} - \Delta a)$  в течение периода времени, превышающего  $T_a$ , будет срабатывать сигнализация ( $\Delta a$  – максимальное допустимое снижение условного потенциала). Далее специалист-технолог должен оценить правильность поведения модели и при совпадении предсказания и фактического состояния ДТО принять необходимые меры по устранению неисправности. Если же предсказание модели окажется ложным, то модель должна быть дообучена путем расширения пространства нормальной эксплуатации вследствие включения в него новых точек.

### АЛГОРИТМ РАСЧЕТА КРИТЕРИЯ АНОМАЛЬНОСТИ ФУНКЦИОНИРОВАНИЯ ДТО

В связи с тем что предлагаемый метод предполагает вычислительную нагрузку, экспоненциально растущую с ростом количества точек базового множества, представляется необходимым ввести ограничения на их количество, не потеряв при этом точки, определяющие конфигурацию этого множества.

#### *Формирование базового множества точек в пространстве признаков*

Пусть:  $N_{\text{норм}}$  – общее количество точек (моментов времени), полученных из архивных данных в период нормального функционирования ДТО (множество  $N_{\text{норм}}$ );

$N_F$  – количество точек в указанном периоде, не имеющих ни одного недостоверного признака (см. далее раздел *Учет недостоверности*);

$N_L$  – предельное количество точек, при котором все вычисления, требуемые методом, заканчиваются за приемлемое время;

$m$  – количество признаков (координат пространства базового множества),  $m \ll N_L$ .

Тогда:

если  $N_F \leq N_L$ , то базовое множество тождественно множеству архивных данных нормально функционирующего ДТО;

иначе базовое множество нужно формировать следующим образом, включая в него:

все точки множества  $N_{\text{норм}}$ , имеющие  $\min_{i=1, \dots, N_{\text{норм}}} P(t)_i$  – минимальное значение координаты (признака)  $P_i$  среди всех  $N_{\text{норм}}$  точек множества  $N_{\text{норм}}$ ,  $i = 1, \dots, m$ ;

все точки множества  $N_{\text{норм}}$ , имеющие  $\max_{i=1, \dots, N_{\text{норм}}} P(t)_i$  – максимальное значение координаты (признака)  $P_i$  среди всех  $N_{\text{норм}}$  точек множества  $N_{\text{норм}}$ ,  $i = 1, \dots, m$ ;

из оставшихся  $N_{\text{норм}} - 2m$  точек следует через равные интервалы  $\Delta t = \frac{N_{\text{норм}} - 2m}{N_L - 2m}$  до- брать точки до количества  $N_L$ .

Итоговое количество точек базового множества можно обозначить  $N_B$ . Тогда из всего изложенного следует:

$$N_B = N_F \quad \text{при} \quad N_F \leq N_L;$$

$$N_B = N_L \quad \text{при} \quad N_F > N_L.$$

#### Определение весов $W(t)$

Вес является атрибутом точки, т.е. атрибутом каждого из моментов времени базового множества. Представляется разумным давать большие веса тем моментам времени, в которые с большей уверенностью можно предположить нормальный режим работы, чем в другие моменты записи исторических данных. Таковыми могут быть, например, периоды работы ДТО сразу после ремонта.

#### Формирование признаков

Признаками, характеризующими функционирование ДТО, служат, прежде всего, его измеряемые параметры.

Для расширения возможности улавливания эффектов аномального функционирования ДТО возможно дополнить перечень признаков производными от измеряемых параметров, например их дисперсиями  $D(t)$ . Предполагается, что такое расширение может оказаться полезным, когда развивающийся дефект первоначально оказывает влияние не столько на сами параметры или их совокупность, сколько на их изменчивость:

$$D(t) = \frac{1}{T_w} \sum_{t-T_w}^t [P(t)_{\text{ср}} - P(t)]^2,$$

где  $P(t)_{\text{ср}} = \frac{1}{T_w} \sum_{t-T_w}^t P(t)$ . Здесь  $T_w$  – ширина окна определения статистических характеристик временного ряда;  $t$  – дискретные моменты времени, на которых фиксируются значения признаков ДТО. Ширину окна можно выбрать равной наибольшему периоду в спектре колебаний параметров ДТО на суточном отрезке стационарного режима в нормальном периоде его эксплуатации, например  $T_w = 8$  ч.

У некоторых параметров поддерживаются заданные значения благодаря воздействию автоматизированных систем регулирования. При этом возможно, что в довольно широком диапазоне развития некоторого теоретически оказывающего влияние на такой параметр дефекта сам параметр существенно изменяться не будет. Однако будет меняться положение поддерживающего этот параметр регулирующего органа. Это означает, что в перечень признаков ДТО желательно включать значения указателей положения регуляторов.

#### Фильтрация помех усреднением на интервале

Временные ряды признаков функционирования ДТО известны в виде дискретного ряда значений через равные интервалы времени.

В статистике выделяют три компонента временного ряда: тренд, сезонные колебания и случайную составляющую. Для решения поставленной задачи представляет интерес именно тренд, а случайная составляющая является вредным шумом, который и должен быть отфильтрован усреднением на интервале, сезонные же колебания для большинства признаков практически отсутствуют.

Усреднение признаков на временном интервале:

$$\overline{P(t)}_i = P(t)_i \quad \text{при} \quad t = 1, \dots, T_m;$$

$$\overline{P(t)}_i = \frac{1}{T_m} \int_{t-T_m}^t P(t)_i dt \quad \text{при} \quad t > T_m.$$

Здесь  $T_m$  – интервал усреднения;  $i = 1, \dots, m$  – количество признаков.

Так как в архивах программно-технического комплекса автоматизированной системы управления технологическим процессом данные хранятся с дискретностью по времени, то интеграл на самом деле является суммой значений каждого признака.

Значение  $T_m$  определяется периодом помехи в спектре частот временных рядов технологических параметров. Исходя из имеющихся архивных данных, представляется разумным принять  $T_m = 10$  мин.

#### Нормализация выборок данных

Необходимость нормализации выборок данных обусловлена различной природой используемых признаков, которые могут изменяться в широком диапазоне и, будучи разнообразными по физическому смыслу, отличаться один от другого на несколько порядков. Например, температура выражается трехзначным числом, тогда как перепад давлений – однозначным.

Работа моделей с такими показателями окажется некорректной: дисбаланс между значения-

ми признаков может привести к совершенно неадекватным результатам.

На практике при вычислении расстояний между точками или векторами чаще всего используется  $Z$ -масштабирование координат этих векторов:

$$\frac{P(t)_i}{\sqrt{D_i}} = \frac{P(t)_i - P(t)_{i_{cp}}}{\sqrt{D_i}},$$

где

$$P_{i_{cp}} = \frac{1}{N_B} \sum_{i=1}^{N_B} P(t)_i;$$

$$D_i = \frac{1}{N_B} \sum_{i=1}^{N_B} [P_{i_{cp}} - P(t)_i]^2;$$

$N_B$  – итоговое количество точек базового множества;  $t = 1, \dots, N_B$ ;  $i = 1, \dots, m$ .

После нормализации все числовые значения входных признаков приводятся к одинаковой области их изменения – некоторому узкому диапазону. Это позволяет свести их вместе в одной модели и обеспечить корректную работу вычислительных алгоритмов.

#### Учет достоверности измерений

Для построения адекватной модели необходимо иметь возможность тем или иным образом отфильтровывать недостоверные измерения. Так как все современные программно-технические комплексы, служащие источниками данных для работы любой предиктивной модели, несут информацию не только о значениях измеряемых параметров, но и об их достоверности (так называемый бит

достоверности), то, используя эту информацию, можно различными способами обрабатывать недостоверные значения и тем самым не допускать ошибок обучения, если эти измерения относятся к обучающему периоду (базовому множеству), или ложных прогнозов, когда эти измерения и основанные на них показатели работы ДТО берутся при работе модели в режиме run-time.

При формировании базового множества все точки, в числе координат которых имеется хотя бы один недостоверный признак, отбрасываются.

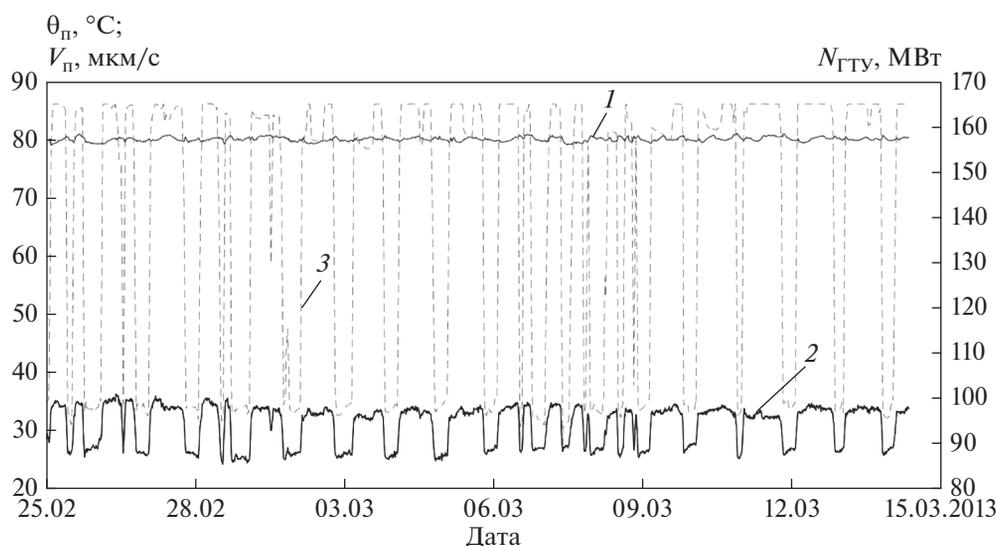
При работе модели в режиме run-time наиболее рационально заменять значение недостоверного параметра на его последнее достоверное значение.

## РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Для повышения чувствительности модели и конкретизации технологического узла, имеющего развивающийся дефект, весь набор измерений был разделен на группы в соответствии с технологией.

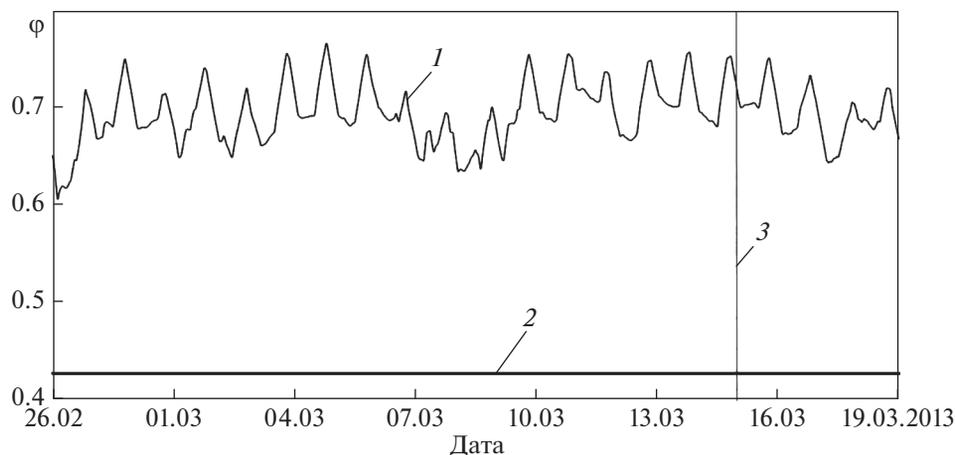
Экспериментальные исследования проводили на исторических данных функционирования газотурбинной установки (ГТУ), записанных в течение трех лет. В каждый момент времени с интервалом  $\Delta t = 10$  мин регистрировалось  $m = 45$  показателей функционирования ГТУ (активная мощность, вибрации подшипников, температуры баббитов подшипников и др.). На рис. 2 показан фрагмент исходных данных.

На рис. 3 приведен график зависимости рассчитанного критерия аномальности  $\phi$  (кривая 1)



**Рис. 2.** Показатели работы газовой турбины в интервале времени от 25.02.2013 до 15.03.2013, охарактеризованном специалистами-технологами как период нормального функционирования.

1 – температура подшипника  $\theta_p$ ; 2 – относительная вибрация подшипника  $V_p$ ; 3 – активная мощность ГТУ  $N_{ГТУ}$



**Рис. 3.** Критерий аномальности  $\phi$  (потенциал) функционирования газовой турбины в интервале времени от 26.02.2013 до 19.03.2013, включающем в себя нормальный (до 15.03.2013) и тестовый (после 15.03.2013) периоды функционирования. 1 – рассчитанный критерий аномальности (потенциал); 2 – определенный ранее минимальный потенциал; 3 – 15 марта 2013 г.

функционирования газовой турбины. Горизонтальной прямой показано минимальное значение потенциала, полученное при обучении модели. На основании экспертного мнения, исходя из оценки каждого отдельного измерения в качестве тестового периода был выбран еще один период нормальной эксплуатации оборудования. Из графика видно, что в тестовый период эксплуатации показатель аномальности диагностирует нормальное состояние исследуемого объекта. Таким образом, в тестовом периоде разработанный метод показывает непротиворечивые результаты определения нормального/аномального функционирования ДТО.

## ВЫВОДЫ

1. Предложенный метод, обладая простотой и наглядностью, продемонстрировал адекватность и непротиворечивость результатов определения нормального/аномального функционирования ДТО на имеющихся архивных данных.

2. При наличии соответствующих исходных (архивных) данных о фиксированных неисправностях в ДТО модель может быть обучена на их поиск. Однако и при отсутствии информации, достаточной для обучения модели для поиска конкретных неисправностей (что есть практически всегда и на всех энергоблоках), модель может успешно работать, обучаясь только на данных, которые получены в период нормальной эксплуатации ДТО.

3. К недостаткам метода следует отнести экспоненциально возрастающую вычислительную сложность по мере увеличения массива обучающих данных. Однако эта проблема присутствует практически во всех имеющихся моделях предик-

тивной аналитики, а в данной статье предложен метод ее решения.

4. Представленный метод, как и все другие методы предиктивной аналитики, применяемые для диагностики технологического оборудования, универсален по отношению к виду этого оборудования. Единственное условие – оборудование должно иметь достаточное для диагностики количество измеряемых параметров, характеризующих его работу.

5. Предложенный метод требует дополнительного апробирования на других архивных выборках, содержащих (что желательно) верифицированный экспертом период аномальной эксплуатации в режимах как офлайн, так и онлайн на текущих данных, оптимизации обучающего набора данных, совершенствования в части прогнозирования сроков выхода на предаварийное и аварийное техническое состояние, определения вероятностных оценок факта возникновения и развития дефекта во времени.

## СПИСОК ЛИТЕРАТУРЫ

1. **Siegel E.** Predictive analytics: The power to predict who will click, buy, lie, or die. John Wiley & Sons Inc., 2016. <https://doi.org/10.1002/9781119172536>
2. **Waller M.A., Fawcett S.E.** Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management // J. Business Logistics. 2013. V. 34. № 2. P. 77–84. <https://doi.org/10.1111/jbl.12010>
3. **Khan S.S., Madden M.G.** One-class classification: taxonomy of study and review of techniques // The Knowledge Engineering Review Cambridge University. 2014. V. 29. Is. 3. P. 345–374. <https://doi.org/10.1017/S026988891300043X>

4. **Trofimov A.G., Kuznetsova K.E., Korshikova A.A.** Abnormal operation detection in heat power plant using ensemble of binary classifiers // *Advances in Neural Computation, Machine Learning, and Cognitive Research. II. Neuroinformatics 2018. Studies in Computational Intelligence* / Ed. by B. Kryzhanovsky, W. Dunin-Barkowski, V. Redko, Y. Tiumentsev. Springer, 2019. V. 799. P. 227–233.  
[https://doi.org/10.1007/978-3-030-01328-8\\_27](https://doi.org/10.1007/978-3-030-01328-8_27)
5. **Коршикова А.А., Трофимов А.Г.** Модель раннего обнаружения аварийных ситуаций на оборудовании электростанций на основе методов машинного обучения // *Теплоэнергетика*. 2019. № 3. С. 49–56.  
<https://doi.org/10.1134/S0040363619030044>
6. **Korshikova A.A., Trofimov A.G.** Predictive model for calculating abnormal functioning power equipment // *Cyber-Physical Systems: Industry 4.0 Challenges. Studies in Systems, Decision and Control* / Ed. by A. Kravets, A. Bolshakov, M. Shcherbakov. Springer, 2020. V. 260. P. 249–259.  
[https://doi.org/10.1007/978-3-030-32648-7\\_20](https://doi.org/10.1007/978-3-030-32648-7_20)
7. **Random** forest as a predictive analytics alternative to regression in institutional research / H. Lingjun, R.A. Levine, J. Fan, J. Beemer, J. Stronach // *Practical Assessment, Research & Evaluation*. 2018. V. 23. № 1. [Электрон. ресурс.] Режим доступа: <https://doi.org/10.7275/1wpr-m024>
8. **Dietterich T.G.** Ensemble learning. *The Handbook of Brain Theory and Neural Networks* / Ed. by M.A. Arbib. 2nd ed. Cambridge: MIT Press, 2002. V. 3. P. 110–125.
9. **Липатов М.** Новая реальность диктует новые правила // *Прана. Новости*. [Электрон. ресурс.] Режим доступа: <https://prana-system.com/novosti/novosti/novaya-realnost-diktuets-novye-pravila/> (Дата обращения 23.11.2020.)

## Model of Emergency Conditions' Early Detection in Power Plant Equipment Based on the Least Potentials Method

A. A. Koshikova<sup>a,\*</sup> and O. M. Idzon<sup>a</sup>

<sup>a</sup> *OOO Incontrol, Moscow, 115280 Russia*

<sup>\*</sup> *e-mail: korshikova@inctrl.ru*

**Abstract**—A method is considered for detecting and predicting the abnormality in operation of power unit equipment by an example of a gas-turbine unit (GTU). A problem of detecting abnormality in operation is formulated as the mathematical problem of modeling an abnormality criterion taking the values from 0 to 1. It has been assumed that the predictive analytics methods can be effective for predicting the future state of process equipment based on the existing scope of measurements without any increase. It is assumed that, even when each individual measurement is within the range taken as the range of normal functioning, their cumulative dynamics enables us to judge a developing defect, i.e., about the transition of the diagnosed process equipment (DPE) to the zone of abnormal operation. To solve this problem, an approach is proposed based on calculating the value of the “abnormality indicator,” which can be interpreted as a conditional potential created by points in a multidimensional space of indicators that characterize the state of equipment at the given time. By learning the model against the indicators that set the regions of states (the state of normal operation and the state for various kinds of fixed defects), one can then apply the trained model to determine the type of state: the closer the value of the abnormality indicator to the values inherent in a particular region of functioning, the greater the probability that the state of DPE corresponds to this region. It is shown that, due to certain objective circumstances, there is no practical possibility of training the model against the data obtained during abnormal operation with specific types of defects in DPE. This reduces the problem to adapting the method to the case when we have only the region of normal operation for learning the model. The proposed model was trained and tested during normal operation of the equipment. The test results indicate that the proposed method is consistent (i.e., it does not yield false positive response).

**Keywords:** process equipment, detection of abnormalities, predictive analytics, least potentials method, regions of normal and abnormal operation, state space, functioning indicators