

УДК 004.852

doi: 10.17586/2226-1494-2019-19-6-1094-1105

## ИДЕНТИФИКАЦИЯ СТАДИИ ДЕГРАДАЦИИ ОБОРУДОВАНИЯ В СИСТЕМАХ СЕРВИСНОГО ОБСЛУЖИВАНИЯ ПРЕВЕНТИВНОГО ТИПА

А.В. Тимофеев<sup>a</sup>, В.М. Денисов<sup>b,c</sup>

<sup>a</sup> ТОО «Эквалайзрум», Астана, 010000, Казахстан

<sup>b</sup> ООО «Флагман-Гео», Санкт-Петербург, 197376, Российская Федерация

<sup>c</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

Адрес для переписки: [timofeev.andrey@gmail.com](mailto:timofeev.andrey@gmail.com)

### Информация о статье

Поступила в редакцию 29.07.19, принята к печати 09.09.19

Язык статьи — русский

**Ссылка для цитирования:** Тимофеев А.В., Денисов В.М. Идентификация стадии деградации оборудования в системах сервисного обслуживания превентивного типа // Научно-технический вестник информационных технологий, механики и оптики. 2019. Т. 19. № 6. С. 1094–1105. doi: 10.17586/2226-1494-2019-19-6-1094-1105

### Аннотация

**Предмет исследования.** Предложен новый способ организации систем превентивного сервисного обслуживания (включая системы обслуживания по состоянию и системы предсказательного обслуживания), основанных на использовании современных методов машинного обучения и функционирующих с привлечением оригинального, непараметрического метода идентификации текущей фазы деградации обслуживаемого оборудования. **Метод.** Предложенный подход заключается в сведении задачи идентификации текущей фазы деградации оборудования к интервальному оцениванию величины параметра «индекс здоровья» оборудования, представляющего собой ступенчатую функцию. Аргументами этой функции является некоторый набор измеримых параметров, объективно характеризующих состояние оборудования. Текущая фаза деградации оборудования определяется с использованием классификационного подхода, в рамках которого на основании анализа наблюдаемых данных, принимается решение о том, какому классу (фазе состояния) эти данные соответствуют. В качестве данных, используемых для идентификации стадии деградации оборудования, рассматриваются измерения от группы сенсоров в общем случае различной физической природы, которые размещены как на поверхности, так и внутри контролируемого оборудования. Математически предложенный подход сводится к взвешенной комбинации двух классификаторов. Один из классификаторов этой комбинации основан на решении группы задач бинарной классификации. Второй классификатор базируется на оценивании параметра «оставшийся срок полезного использования» методом непараметрической регрессии. **Основные результаты.** В отличие от традиционных предложенный подход использует минимум априорной информации о принципах функционирования и устройстве обслуживаемого оборудования и основан на использовании ступенчатой функции «индекс здоровья» оборудования. Новизна подхода заключена в одновременном использовании ступенчатой функции «индекс здоровья» и взвешенной комбинации двух классификаторов различной структуры. Разработанный метод показал положительные результаты при его тестировании на базе данных C-MAPPS Dataset, которая содержит данные об отказах турбовентиляторных двигателей, смоделированных с использованием термодинамической имитационной модели. Предотказный статус оборудования идентифицирован с вероятностью 99 %. **Практическая значимость.** Полученные результаты и алгоритмы могут быть использованы в системах превентивного обслуживания с целью высоконадежной идентификации текущей стадии деградации оборудования.

### Ключевые слова

предсказательное обслуживание, обслуживание по состоянию, машинное обучение, ML PdM, XGBoost, SVM-регрессия

doi: 10.17586/2226-1494-2019-19-6-1094-1105

## IDENTIFICATION OF EQUIPMENT DEGRADATION PHASE IN PREVENTATIVE MAINTENANCE SYSTEMS

A.V. Timofeev<sup>a</sup>, V.M. Denisov<sup>b,c</sup>

<sup>a</sup> EqualiZoom, LLP, Astana, 010000, Kazakhstan

<sup>b</sup> Flagman Geo, OOO, Saint Petersburg, 197376, Russian Federation

<sup>c</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

Corresponding author: [timofeev.andrey@gmail.com](mailto:timofeev.andrey@gmail.com)

### Article info

Received 29.07.19, accepted 09.09.19

Article in Russian

**For citation:** Timofeev A.V., Denisov V.M. Identification of equipment degradation phase in preventative maintenance systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2019, vol. 19, no. 6, pp. 1094–1105 (in Russian). doi: 10.17586/2226-1494-2019-19-6-1094-1105

#### Abstract

**Subject of Research.** The paper proposes a novel organization technique for preventive maintenance systems (including condition-based and predictive maintenance systems) based on the use of modern machine learning methods. The systems are operating using an original, non-parametric identification method for the current degradation phase of serviced equipment. **Method.** The proposed method comprises reducing the task of the current phase identification of the equipment degradation phase to interval estimation of the value of the so-called “health index” parameter of the equipment. This parameter is represented as a step function with the arguments in terms of a set of the measurable equipment objective parameters. The current equipment degradation phase is determined by classification approach. At this, based on the analysis of the observed data, it is decided upon what class (state phase) these data correspond to. Measurements from a group of sensors, in general, of various physical nature, which are located both on the surface and inside the equipment being monitored are used as data for identification of the equipment degradation stage. Mathematically, the proposed approach is reduced to a weighted combination of two classifiers. One of the classifiers of this combination is based on solving a group of binary classification problems. The second classifier is based on “Remaining Useful Life” parameter estimation by the method of nonparametric regression. **Main Results.** As distinguished from traditional approaches, the proposed approach uses a minimum of a priori information about the principles of operation and the internal structure of the equipment being serviced. The approach is based on the usage of the “health index” equipment parameter presented in the form of a step function. The novelty of the approach lies in the simultaneous use of the “health index” step function and the weighted combination of two classifiers with various structure. The proposed method showed good results when being tested on the C-MAPPS Dataset database, which contains data on failures of turbofan engines modeled using a thermodynamic simulation model. The pre-failure status of the equipment is identified with the probability of 99%. **Practical Relevance.** The obtained results and algorithms can be used in preventive maintenance systems aimed at reliable identification of the equipment degradation current stage.

#### Keywords

predictive maintenance, condition-based maintenance, machine learning, ML PdM, XGBoost, SVM-regression

### Введение

В последние несколько лет в рамках внедрения концепции «Индустриализация 4.0» в области планирования технического обслуживания сложной техники значительный интерес вызывают подходы, основанные на использовании современных методов машинного обучения с целью организации сервисного плана по каждому, отдельно взятому образцу оборудования конкретного типа. В основном речь идет об организации сервисного обслуживания превентивного типа в части обслуживания по состоянию [1] и прогностического обслуживания на основе машинного обучения [2]. В англоязычной литературе организация превентивного сервисного обслуживания по состоянию получила название Condition-Based Maintenance (CBM), а прогностическое обслуживание на основе машинного обучения соответственно — Machine Learning based Predictive Maintenance (ML PdM). Под прогнозированием здесь понимается вероятностная оценка длительности временного интервала, по истечении которого произойдет немедленный отказ оборудования. Этот интервал получил вполне устоявшееся название: *RUL* (Remaining Useful Life) [3]. На предприятиях Российской Федерации вместо термина CBM зачастую используется семантически эквивалентное ему понятие «стратегия ремонтов по текущему техническому состоянию». Текущее состояние оборудования определяется для каждой конкретной единицы обслуживаемого оборудования при помощи группы инструментальных технологий, обязательно включающих в себя удаленный мониторинг состояния оборудования с использованием группы сенсоров различной физической природы. В настоящее время CBM-стратегия сервисного обслуживания получила широкое распространение на железнодорожном транспорте для обслуживания силовых двигательных установок, а также всего спектра путевого (напольного) оборудования [4].

В рамках техник CBM и ML PdM организация сервисного плана для конкретной единицы оборудования основана на идентификации стадии ее текущего износа, которая играет роль начальной точки отсчета при использовании различных математических моделей, описывающих процесс деградации оборудования. Вообще, в рамках различных идеологий превентивного обслуживания исторически изучаются различные модели деградации. В частности [1], рассматриваются подходы, основанные:

- на физической модели деградации (Physical Model-Based Methodology) [5];
- на использовании формализованных знаний (Knowledge-Based Methodology) [6];
- на анализе данных (Data-Driven Methodology – DDM) [1, 7–16].

Тенденции последних лет доказывают, что именно DDM обладает наибольшим потенциалом и практической эффективностью, которая продолжает наращиваться на базе использования как статистических методов анализа многомерных данных, так и методов машинного обучения. Характерно, что в ноябре 2018 г. Пентагон (США) принял программу по масштабной разработке и внедрению в практику обслуживания и ремонта вооружений армии и флота США именно ML PdM-подход, являющийся одним из вариантов DDM [17]. При реализации DDM на основе статистического анализа используется множество методов, в том числе метод главных компонент, линейный и квадратичный дискриминантный анализ, канонический вариационный анализ, различные модели пространства состояний (байесовские сети, скрытые марковские

и полумарковские сети), а также всевозможные модели регрессионного типа. В рамках использования DDM на базе машинного обучения успешно применяются различные варианты нейронных сетей, модели на основе метода опорных векторов (SVM), ансамблевые методы (ADABOOST, Gradient Boosting и пр.), а также некоторые иные методики.

Для описания степени износа оборудования используется [7–16] скалярная функция, которая получила название «индикатор здоровья» (Health Indicator, *HI*). Аргументом этой функции является некоторый набор объективных параметров оборудования, а ее значение косвенно характеризует степень износа оборудования или узла при данном значении аргумента. Показатель *HI* предназначен для интегральной характеристики степени износа оборудования или его отдельных узлов. В настоящее время общепринятого подхода к конструированию *HI* не сформировано, ввиду объективно существующего различия как в принципах функционирования оборудования различных типов, так и в методах получения объективной информации относительно его текущего статуса. В зависимости от используемой модели представления *HI*, разработано и опробовано значительное число подходов оценивания его статуса. В том числе регрессионные и авторегрессионные модели, различные варианты кластерного анализа, модели на основе нейронных сетей, скрытые марковские модели и ансамбли классификаторов.

Широкое распространение получил подход, когда значения параметра  $\{HI(t)\}$ , определенные в различные моменты времени  $t$ , используются в качестве множества наблюдений, на основе которых формируются статистические точечные оценки параметра *RUL*, значения которых в свою очередь используются в качестве основы для построения сервисного плана. При этом полагается, что параметр *RUL* должен быть определен с максимально возможной точностью. Однако далеко не во всех системах сервисного обслуживания использование высокоточной (несмещенной, с минимумом дисперсии) точечной статистической оценки *RUL* в качестве основания для построения сервисного плана имеет практический смысл. Напротив, в значительном числе практических систем сервисного обслуживания параметр *RUL* играет лишь вспомогательную роль, уступая основную роль собственно параметру *HI*, представленному в виде ступенчатой функции от определенного набора объективных параметров оборудования. Действительно, с одной стороны, при построении сервисного плана на основе параметра *RUL* в любом случае придется разрабатывать и использовать какой-либо способ разбиения *RUL* на интервалы, соответствующие степени деградации *HI*, так как только степень деградации *HI* является основанием для определения группы адекватных этой степени сервисных актов. С другой стороны, задача построения сервисного плана в любом случае сведется к оцениванию параметра *HI*. А с другой стороны, — использование для сервисного планирования оценок параметра *HI* обосновано сравнительно невысокой скоростью развития процессов деградации реального оборудования. В этом случае для формирования сервисного плана достаточно определить текущий статус параметра *HI*, который с высокой вероятностью останется неизменным на протяжении интервала времени, которое потребуется для проведения сервисного обслуживания. В настоящей работе предложен новый непараметрический DDM-метод идентификации текущего значения параметра *HI*, имеющего вид ступенчатой функции, который не нуждается ни в априорной информации о принципах функционирования обслуживаемого оборудования, ни в информации о принципах функционирования сенсорной сети, которая используется для контроля физических параметров оборудования. Таким образом, предложенный подход, с одной стороны, является типичным представителем DDM для идентификации *HI*, а с другой, — представляет собой простое, удобное и понятное сервисным инженерам средство для формирования сервисных планов.

### Необходимые понятия и определения

В данном разделе будут введены понятия и определения, необходимые для формализации постановки задачи и описания предложенного метода ее решения. Необходимость данного раздела обусловлена еще и тем, что в русскоязычном сегменте научной периодики еще не вполне устоялись базовые термины, понятия и определения, необходимые для анализа систем превентивного и прогнозного сервисного обслуживания.

**Определение *HI* ступенчатого типа в системах сервисного обслуживания.** Рассмотрим ситуацию, когда необходимо построить план сервисного обслуживания для некоторого парка однотипного оборудования типа  $\beta$ , основываясь на динамическом оценивании параметра *RUL* для каждой единицы этого парка. Условимся называть такой парк оборудования  $\beta$ -парком оборудования или  $\beta$ -ПО. Сервисное обслуживание основано на допущении того, что число возможных состояний обслуживаемого оборудования — конечно. Состояние оборудование оценивается с точностью до некоего статусного параметра, определяющего его техническое состояние. В специальной литературе устоялось название этого параметра: *HI*. Традиционно рассматриваются различные типы представления *HI*, однако в большинстве практических систем сервисного обслуживания достаточно рассматривать *HI* в виде некой ступенчатой функции. Число значений этой функции — сравнительно невелико, редко превосходит величину 6, а в подавляющем числе случаев равно трем. В настоящей работе *HI* будет рассматриваться исключительно в форме ступенчатой функции, способ определения которой будет описан далее.

Пусть каждая единица обслуживаемого оборудования  $d_j \in \mathbf{D}_\beta, j = 1, \dots, m$  оснащена сенсорной сетью, состоящей из  $w$  сенсоров  $s_i^{(j)}, i = 1, \dots, w$  в общем случае различной физической природы, размещенных как на

поверхности оборудования, так и в его объеме *однотипно* (в точках с одинаковыми относительными координатами) для каждого  $d_j \in \mathbf{D}_\beta$ . Здесь  $\mathbf{D}_\beta$  — конечное множество, состоящее из  $m$  единиц оборудования типа  $\beta$ .

Выход каждого сенсора  $s_i^{(j)}$ , в момент времени  $t$ , представляет собой некоторую величину  $r_i^{(j,\beta)}(t) \in \mathbf{R}^1$ . Здесь  $\mathbf{R}^1$  — множество действительных чисел. Итак, в момент  $t$ ,  $HI$  единицы оборудования  $d_j$  типа  $\beta$  в общем случае зависит от некоторого набора объективных, наблюдаемых с использованием сенсорной сети, множеств скалярных параметров  $R_i^{(j,\beta)} = \{r_i^{(j,\beta)}(t) | i = 1, \dots, w\}$ . Здесь  $R_i^{(j,\beta)}$  — обозначение множества скалярных параметров, которое зависит от момента времени  $t$ , индекса сенсора  $j$  и типа оборудования  $\beta$ . Значения параметров  $r_i^{(j,\beta)}(t)$  в свою очередь обусловлены влиянием множества ненаблюдаемых факторов, которые носят как случайный, так и системный характер. Например, они зависят от функционального и климатического режимов эксплуатации оборудования, а также от качества материалов, из которых изготовлен конкретный образец оборудования. В процессе эксплуатации единицы оборудования  $d_j$  величины компонентов набора наблюдаемых параметров  $r_i^{(j,\beta)}$  — девируют во времени, что обусловлено как постепенной деградацией единицы оборудования  $d_j$ , так и влиянием случайных факторов. Можно представить этот процесс в виде простой модели:

$$r_i^{(j,\beta)}(t_{n+1}) = r_i^{(j,\beta)}(t_n) + \Delta_{\beta,i}(\varphi_{t_{n+1}}) + \xi_i^{(j,\beta)}(t_{n+1}), n = 1, 2, \dots \quad (1)$$

Здесь  $\Delta_{\beta,i}(\varphi)$  — априорно неизвестная, системная, неслучайная часть приращения, обусловленная постепенным износом оборудования и зависящая от ненаблюдаемой функции износа,  $\varphi$ ,  $\xi_i^{(j,\beta)}(t)$  — случайная часть приращения, которая зависит от группы параметров, определяющих режим эксплуатации оборудования, часть из которых носит ярко выраженный случайный характер. Для простоты условимся считать, что  $\forall i, j, t: \mathbf{E} \xi_i^{(j,\beta)}(t) = 0$ . Здесь символ  $\mathbf{E}$  обозначает операцию математического ожидания. Таким образом, для единицы оборудования  $d_j$  можно записать  $HI$  в следующем виде:  $HI(R_i^{(j,\beta)}) \equiv HI_{t,j,\beta}$ . Из (1) легко видеть, что:

$$\forall i, j, t, n: \mathbf{E} r_i^{(j,\beta)}(t_n) = \sum_{k=1}^n \Delta_{\beta,i}(\varphi_{t_k}) = \rho_{n,i}^{(\beta)}$$

Обозначим:

$$\overline{R}_i^{(\beta)} = \{\rho_{n,i}^{(\beta)} | i = 1, \dots, w\} \quad (2)$$

$\overline{R}_i^{(\beta)}$  представляет собой набор данных от сенсоров с индексами  $i = 1, \dots, w$ , наблюдаемый в момент времени  $t$  и **усредненный по всем возможным единицам оборудования типа  $\beta$** , которые тестируются на устойчивость к неисправности типа  $\lambda$ . Производитель оборудования в процессе его разработки и тестирования создает корпус из  $b$  единиц оборудования типа  $\beta$ , который тестируется на устойчивость к неисправности типа  $\lambda$ . В этом случае величины  $\rho_{n,i}^{(\beta)}$  могут быть заменены их статистическими оценками по этому корпусу, а выражение (2) может быть переписано в следующем, приближенном, виде:

$$\overline{R}_i^{(\beta)} = \{\sum_{j=1}^b r_i^{(j,\beta)}(t) / b | i = 1, \dots, w\} \quad (3)$$

$\overline{R}_i^{(\beta)}$  характеризует  $HI$  в момент времени  $t$ . В общем случае в (3), вместо операции усреднения, могут быть использованы иные статистики, проецирующие множества  $R_i^{(j,\beta)}, j = 1, \dots, b$  на  $\mathbf{R}^1$ .

В результате статистического исследования производителем выясняется средняя наработка до отказа по неисправности данного типа, которую обозначим символом  $t_\beta(\lambda)$ . В этом случае интервал  $[0, t_\beta(\lambda)]$  называется средним временем жизни оборудования типа  $\beta$  по отношению к неисправности типа  $\lambda$ . Далее экспертными методами интервал  $[0, t_\beta(\lambda)]$  разбивается на  $g$  непересекающихся субинтервалов следующим образом:

$$I_p^{(\lambda,\beta)} = [t_\beta(\lambda) - \sum_{k=1}^p a_k, t_\beta(\lambda) - \sum_{k=1}^{p-1} a_k], a_1 = 0, \sum_{k=1}^g a_k = t_\beta(\lambda), \cup_k I_p^{(\lambda,\beta)} = [0, t_\beta(\lambda)], \cap_k I_p^{(\lambda,\beta)} = \emptyset$$

Здесь  $p$  — количество субинтервалов,  $a_k$  — ширина  $k$ -того субинтервала, измеряемая в числе циклов измерения. Будем обозначать эту величину следующим образом:  $|I_k^{(\lambda,\beta)}|$ , т. е.  $a_k = |I_k^{(\lambda,\beta)}|$ . Число субинтервалов  $g$  и принцип разбиения зависит от множества факторов, но основная цель разбиения такова: выделить субинтервалы, которые значимо для сервисного обслуживания идентифицируют степень деградации оборудования. Каждый из этих субинтервалов, очевидно, соответствует фиксированному значению (ступени) параметра  $HI$ , имеющего вид ступенчатой функции со значениями из множества  $H = \{h_1, \dots, h_g\} \subseteq \mathbf{R}^1$ . Каждому значению  $h_p \in H$  ступенчатой функции  $HI$ , согласно экспертному решению производителя, предписывается соответствующая группа сервисных операций, предназначенных продлить срок эксплуатации оборудования, находящегося в техническом состоянии, которое приблизительно характеризуется соответствующим значением  $HI$ . Как правило, субинтервал  $I_g^{(\lambda,\beta)}$  — всегда соответствует нормальному режиму работы оборудования ( $HI = h_g$ ), а субинтервал  $I_1^{(\lambda,\beta)}$  — всегда соответствует предотказному режиму ( $HI = h_1$ ). Исходя из логики построения сервисных систем, субинтервал  $I_1^{(\lambda,\beta)}$  соответствует аварийной ситуации тогда, когда до полного отказа данной единицы оборудования остались считанные циклы его функционирования: необ-



ходимо немедленное вмешательство в форме акции сервисного обслуживания. Каждому интервалу  $I_p^{(\lambda, \beta)}$  соответствует множество объективных характеристик

$$a_k = |I_k^{(\lambda, \beta)} \cap R^{(\beta)}(p) = \{\overline{R}_t^{(\beta)} \mid t \in I_p^{(\lambda, \beta)}\}. \quad (4)$$

Согласно (1), на любом непустом интервале времени параметры  $r_i^{(j, \beta)}(t)$  в среднем — девируют, поэтому на интервале  $I_p^{(\lambda, \beta)}$  девируют и элементы множества  $R^{(\beta)}(p)$ . Фактически, множество  $R^{(\beta)}(p)$  является набором объективных характеристик  $HI$  на интервале  $I_p^{(\lambda, \beta)}$ , усредненных по всему множеству возможных единиц оборудования типа  $\beta$ . Для параметра  $HI$ , в форме функции ступенчатого вида, который определен для некой обобщенной (усредненной) единицы оборудования типа  $\beta$ , имеем:

$$\forall_{p=1}^g \forall t \in I_p^{(\lambda, \beta)} : (HI_t(R^{(\beta)}(p)) = h_p), h_p \in H.$$

Функция  $HI(\cdot)$  — неизвестна. В общем случае считается, что параметр  $HI$ , в процессе деградации оборудования, не должен достигать значения  $h_1$ , которое соответствует интервалу  $I_1^{(\lambda, \beta)}$ , так как в этом случае высок риск выхода из строя по неисправности типа  $\lambda$ .

Пусть  $T_o = t_1^{(o)}, t_2^{(o)}, t_3^{(o)}, \dots, t_{\beta}(\lambda)$  — конечная последовательность строго возрастающих моментов времени получения данных от сенсорной сети, размещенной на **обобщенной единице оборудования** типа  $\beta$ . Данные, поступившие от сенсорной сети обобщенной единицы оборудования и соответствующие множеству моментов измерения из  $T_o$ , представляют собой кортеж  $\mathbf{r}^{(o, \beta)}(T_o) = (\overline{R}_{t_1^{(o)}}^{(\beta)}, \overline{R}_{t_2^{(o)}}^{(\beta)}, \dots, \overline{R}_{t_{\beta}(\lambda)}^{(\beta)})$ , который будем называть *траекторией* измерений группы сенсоров для **обобщенной единицы оборудования** или *o-траекторией*. *Трендом деградации* оборудования назовем следующую последовательность:

$$\mathbf{HI}^{(o, \beta)}(T_o) = (HI_{t_1^{(o)}}(\overline{R}_{t_1^{(o)}}^{(\beta)}), HI_{t_2^{(o)}}(\overline{R}_{t_2^{(o)}}^{(\beta)}), \dots, HI_{t_{\beta}(\lambda)}(\overline{R}_{t_{\beta}(\lambda)}^{(\beta)})).$$

Тренд деградации  $\mathbf{HI}^{(o, \beta)}(T_o)$  характеризует динамику деградации усредненной единицы оборудования типа  $\beta$  для  $HI$  в виде ступенчатой функции. Каждая ступень  $HI$  четко соответствует определенному субинтервалу  $I_p^{(\lambda, \beta)}$ . Динамика деградации конкретной единицы оборудования  $d_j$  в силу множества причин, которые формально отражены в (1), имеет случайный характер и поэтому может существенно отличаться от тренда деградации, который, по определению, — детерминирован.

**Особенности построения сервисного плана при использовании  $HI$  ступенчатого типа.** В (4) множество  $R^{(\beta)}(p)$  для субинтервала  $I_p^{(\lambda, \beta)}$  определено для обобщенной единицы оборудования путем усреднения характеристик  $R_i^{(j, \beta)}$  по всему множеству возможных единиц  $d_j$  оборудования типа  $\beta$ . Понятие  $R^{(\beta)}(p)$  в большей части — теоретический концепт, который необходим для определения базовых представлений о способе формирования ступенчатого показателя  $HI$  и тренде деградации оборудования типа  $\beta$ . В реальной практике сервисного обслуживания приходится работать с множествами  $R_i^{(j, \beta)}(p) = \{R_i^{(j, \beta)} \mid t \in I_p^{(\lambda, \beta)}\}$ , которые определены для конкретных единиц оборудования  $d_j$ .

Процесс построения сервисного плана для каждой единицы оборудования из  $\beta$ -ПО, обычно состоит из двух фаз. На первой, основываясь на уже прошедшем наблюдении корпусе объективных данных, поступивших от сенсорной сети, для каждой единицы оборудования производится оценка  $HI$ . На следующей фазе множество единиц оборудования  $\beta$ -ПО разбивается на подмножества, каждое из которых содержит единицы оборудования с одним и тем же значением параметра  $h_p \in H$ . Условимся называть эти подмножества  $h_p$ -блоками. Далее, каждой единице оборудования из соответствующего  $h_p$ -блока назначается одна и та же группа сервисных операций, которые включаются в общий план сервисного обслуживания, предназначенные для того, чтобы увеличить параметр  $RUL$  для каждой единицы  $h_p$ -блока.

Важный нюанс состоит в том, что для реализации сервисного обслуживания необходимо знать не только текущую оценку параметра  $HI$ , но и интервал времени, в течение которого значение этого параметра будет актуально. Действительно, теоретически можно предположить ситуацию, когда для некоторых  $d_j$  и  $p$ ,  $HI_{t_k}(R_i^{(j, \beta)}(p)) \neq HI_{t_{k+1}}(R_i^{(j, \beta)}(p))$  для смежных моментов времени контроля  $t_k$  и  $t_{k+1}$ , т. е. параметр  $HI$  оказался на границе двух субинтервалов  $I_p^{(\lambda, \beta)}$ ,  $I_{p+1}^{(\lambda, \beta)}$ . Назовем эту ситуацию проблемой перехода. Для того чтобы идентифицировать проблему перехода, в системах со ступенчатыми  $HI$  в теории может быть использована оценка параметра  $RUL$ , каким-либо образом определенная для каждой единицы оборудования из  $\beta$ -ПО. При этом считается [3], что параметр  $RUL$  должен быть оценен с максимально высокой точностью, так как высокая точность оценивания  $RUL$  позволит надежно определить: какому именно субинтервалу  $I_p^{(\lambda, \beta)}$  в рамках развития тренда деградации единицы  $d_j$ , в действительности, соответствует параметр  $HI$ . Однако для довольно обширного класса реальных систем сервисного обслуживания высокая точность оценки величины  $RUL$  является избыточным требованием. В этих системах, в особенности на субинтервале  $I_2^{(\lambda, \beta)}$ , который непосредственно предшествует предотказному субинтервалу  $I_1^{(\lambda, \beta)}$ , процессы деградации развиваются достаточно плавно. Для систем данного класса интервал времени  $\Delta t$ , в течение которого должна быть реализована группа сервисных операций для конкретной единицы оборудования согласно сервисному плану, существенно меньше величины  $(t_{k+1} - t_k)$  интервала, по истечении которого будет сформирована очередная

оценка  $HI_{t_{k+1}}(R_{t_{k+1}}^{(j,\beta)}(p))$ , т. е.  $\Delta t \ll (t_{k+1} - t_k)$ . Таким образом, для систем такого класса группа операций сервисного обслуживания будет завершена прежде, чем параметр  $HI$  для единицы  $d_j$  успеет деградировать до следующей ступени, которая в общем случае будет требовать выполнения другой группы сервисных операций. Таким образом, в рассматриваемых условиях проблема перехода неактуальна. Следовательно, использование высокоточной оценки  $RUL$  не имеет практического смысла. При этом оценка  $RUL$  может оказаться вполне полезной для оценивания текущего статуса  $HI$ , а также для ранжирования сервисных операций внутри  $h_p$ -блока. Важно отметить, что правильно сконфигурированная система сервисного обслуживания должна обладать настолько развитой базой сервисных ресурсов, чтобы обеспечивать сервисное обслуживание на субинтервале  $I_2^{(\lambda,\beta)}$ , не допуская проведения сервисных работ в течение субинтервала  $I_1^{(\lambda,\beta)}$ .

В реальных системах сервисного обслуживания отсутствует практическая необходимость рассматривать более чем три значения параметра  $HI$ :  $H = \{h_A, h_W, h_N\} \subseteq \mathbf{R}^1$ . Здесь переобозначено:

—  $h_A \equiv h_1$ : «Alarm» — необходимо незамедлительное сервисное обслуживание, субинтервал  $I_1^{(\lambda,\beta)} \equiv I_A^{(\lambda,\beta)}$ ;

—  $h_W \equiv h_2$ : «Warning» — приближение к предотказному состоянию, данной единице оборудования настоятельно рекомендовано сервисное обслуживание, субинтервал  $I_2^{(\lambda,\beta)} \equiv I_W^{(\lambda,\beta)}$  (это основной интервал для проведения сервисного обслуживания: планирование необходимо провести таким образом, чтобы все сервисные операции были выполнены в течение интервала  $I_W^{(\lambda,\beta)}$ , в крайнем случае часть операций можно завершить в течение  $I_A^{(\lambda,\beta)}$ , но этого следует избегать).

—  $h_N \equiv h_3$ : «Norma» — нормальное состояние оборудования, сервисное обслуживание не требуется, субинтервал  $I_3^{(\lambda,\beta)} \equiv I_N^{(\lambda,\beta)}$ .

### Постановка задачи

Пусть  $T_j = t_1^{(j)}, t_2^{(j)}, t_3^{(j)}, \dots, t_{|T_j|}^{(j)}$  — конечная последовательность строго возрастающих моментов времени получения данных от сенсорной сети  $\{s_i^{(j)} | i = 1, \dots, w\}$ , размещенной на экземпляре оборудования  $d_j \in \mathbf{D}_\beta$ ,  $|T_j|$  — число моментов времени (элементов) в последовательности  $T_j$ . Элемент последовательности  $T_j$  с индексом  $i$  будем обозначать  $\varepsilon_i(T_j)$ . Таким образом, данные, поступившие от датчика  $s_i^{(j)}$  для множества моментов измерения из  $T_j$ , представляют собой кортеж  $\mathbf{r}_i^{(j,\beta)}(T_j) = (r_i^{(j,\beta)}(t_1^{(j)}), r_i^{(j,\beta)}(t_2^{(j)}), r_i^{(j,\beta)}(t_3^{(j)}), \dots, r_i^{(j,\beta)}(t_{|T_j|}^{(j)}))$ , который будем называть *траекторией*  $i$ -того сенсора для  $j$ -того экземпляра оборудования  $d_j \in \mathbf{D}_\beta$  или  $(i, j)$ -траекторией. Длиной траектории  $\mathbf{r}(T)$  называется число компонентов кортежа символом  $|\mathbf{r}(T)|$ .

Рассмотрим конечное множество  $D_\beta^{(L)} \subseteq \mathbf{D}_\beta$ , состоящее из  $q$  единиц оборудования типа  $\beta$ .  $\mathbf{D}_\beta / D_\beta^{(L)} \neq \emptyset$ . Условимся называть  $D_\beta^{(L)}$  множеством единиц оборудования типа  $\beta$ , которое используется для обучения системы. Исходно задано обучающее множество:

$$R_L(\lambda | \mathbf{T}(D_\beta^{(L)})) = \{\mathbf{r}_i^{(j,\beta)}(T_j) | j = 1, \dots, q, \forall_j : d_j \in D_\beta^{(L)}, i = 1, \dots, w\}.$$

состоящее из прошедших наблюдение траекторий, каждая из которых *заканчивается выходом из строя* соответствующего экземпляра оборудования  $d_j \in D_\beta^{(L)}$  по одному и тому же типу неисправности с кодом  $\lambda$  или  $\lambda$ -завершенными траекториями. Другими словами, в моменты времени  $t_{|T_j|}^{(j)}$  оборудование  $d_j \in D_\beta^{(L)}$  выходит из строя по одной и той же причине, которая имеет код  $\lambda$ . Множество  $\mathbf{T}(D_\beta^{(L)}) = \{T_j | j = 1, \dots, m, d_j \in D_\beta^{(L)}\}$  состоит из последовательностей  $T_j$ , соответствующих каждой единице оборудования  $d_j \in D_\beta^{(L)}$ . Данные измерений от отдельных сенсоров могут поступать несинхронно, в различном темпе (интервалы съема данных могут существенно различаться по величине). Обозначим символом  $t_{fail}^{(\beta)}(d_j)$  момент выхода единицы оборудования  $d_j \in D_\beta^{(L)}$  из строя. По определению множества  $R_L(\lambda | \mathbf{T}(D_\beta^{(L)}))$ ,  $\forall T_j \in \mathbf{T}(D_\beta^{(L)}) : t_{fail}^{(\beta)}(d_j) = t_{|T_j|}^{(j)}$ .

В соответствии с ранее сделанными предположениями, для интервала  $[0, t_{fail}^{(\beta)}(d_j)]$  определены субинтервалы  $I_A^{(\lambda,\beta)}$ ,  $I_W^{(\lambda,\beta)}$ , и  $I_N^{(\lambda,\beta)}$ ,  $I_A^{(\lambda,\beta)} \cup I_W^{(\lambda,\beta)} \cup I_N^{(\lambda,\beta)} = [0, t_{fail}^{(\beta)}(d_j)]$ ,  $I_A^{(\lambda,\beta)} \cap I_W^{(\lambda,\beta)} \cap I_N^{(\lambda,\beta)} = \emptyset$ .

Обозначим  $\mathbf{I} = \{I_N^{(\lambda,\beta)}, I_W^{(\lambda,\beta)}, I_A^{(\lambda,\beta)}\}$  — множество субинтервалов, а  $Index(\mathbf{I})$  — множество индексов множества  $\mathbf{I}$ , в качестве которых используются соответствующие значения параметра  $HI$ . То есть  $Index(\mathbf{I}) = \{h_A, h_W, h_N\} \subseteq \mathbf{R}^1$ . В процессе функционирования оборудования типа  $\beta$  наблюдается множество:

$$R_T(\lambda | \mathbf{T}(D_\beta^{(T)})) = \{\mathbf{r}_i^{(j,\beta)}(T_j) | j = 1, \dots, p, \forall_j : d_j \in D_\beta^{(T)}, i = 1, \dots, w\}.$$

Здесь  $D_\beta^{(T)} \subseteq \mathbf{D}_\beta$ ,  $D_\beta^{(T)} \cap ID_\beta^{(T)} = \emptyset$ ,  $|D_\beta^{(T)}| = p$ . Назовем подпоследовательностью  $T_j^*$  последовательности  $T_j$  произвольную непустую последовательность смежных элементов из последовательности  $T_j$  такую, что  $|T_j^*| \leq |T_j|$ ,  $t_{|T_j^*|}^{(j)} \leq t_{|T_j|}^{(j)}$ . Если  $T_j^*$  есть подпоследовательность  $T_j$ , то будем обозначать это так:  $T_j^* \leftarrow T_j$ . Максимально близким для момента времени  $t$  снизу назовем такой  $\varepsilon_i(T_j)$ ,  $\varepsilon_i(T_j) \leq t$ , что  $\forall_i ! \exists i : (\varepsilon_i(T_j) > \varepsilon_i(T_j^*))$ . Для некоторого момента  $t$ , подпоследовательность  $T_j^* \leftarrow T_j$  такая, что  $t_{|T_j^*|}^{(j)} = \varepsilon_i(T_j^*)$  будет обозначаться  $T_j^*(t)$ . Обозначим:  $\langle \mathbf{a} \rangle_k$  —  $k$ -той компонента кортежа  $\mathbf{a}$ . Если  $\mathbf{b}$  является субкортежем кортежа  $\mathbf{a}$ , то будем писать:  $\mathbf{b} \subseteq \mathbf{a}$ . Если  $T_j^* \leftarrow T_j$ , то  $\mathbf{r}_i^{(j,\beta)}(T_j^*(t)) \subseteq \mathbf{r}_i^{(j,\beta)}(T_j)$  называется субтраекторией траектории  $\mathbf{r}_i^{(j,\beta)}(T_j)$ .

Для момента времени  $t$ ,  $d_j \in D_{\beta}^{(T)}$ ,  $1 \leq j \leq p$  рассмотрим такое множество наблюдений  $R_j(\lambda|t, \mathbf{T}(D_{\beta}^{(T)}))$ , что

$$R_j(\lambda|t, \mathbf{T}(D_{\beta}^{(T)})) = \{r_i^{(j,\beta)}(T_j^*(t)) | T_j^*(t) \leftarrow T_j, i = 1, \dots, w\}.$$

Пусть  $t_s$  — момент принятия классификационного решения, тогда множество наблюдений  $R_j(\lambda|t, \mathbf{T}(D_{\beta}^{(T)}))$  определяет набор данных, который будет наблюден для  $d_j \in D_{\beta}^{(T)}$  до момента времени.

Необходимо на основании данных  $R_L^{(\lambda,\beta)} \equiv R_L(\lambda|\mathbf{T}(D_{\beta}^{(L)}))$ , для оборудования типа  $\beta$  и неисправности типа  $\lambda$ , синтезировать классификатор  $C_{\lambda,\beta}(\mathbf{r}_j|R_L^{(\lambda,\beta)})$ ,  $\mathbf{r}_j = R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$ , состояния произвольной единицы оборудования  $d_j \in D_{\beta}^{(T)}$ , который, момент времени принятия классификационного решения  $t_s$  для конкретной единицы оборудования  $d_j \in D_{\beta}^{(T)}$  отображает множество  $R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$  на множество  $Index(\mathbf{I})$ :

$$R_j(\lambda | t_s, \mathbf{T}(D_{\beta}^{(T)})) \xrightarrow{C_{\lambda,\beta}(\mathbf{r}_j|R_L^{(\lambda,\beta)})} Index(\mathbf{I}).$$

То есть  $\forall_{j=1}^p : C_{\lambda,\beta}(\mathbf{r}_j | R_L^{(\lambda,\beta)}) = k, k \in Index(\mathbf{I})$ .

### Алгоритм классификации состояния оборудования

Решение рассматривается в виде ансамбля из двух классификаторов. Один из которых представляет собой обычный классификатор, отображающий  $R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$  на  $Index(\mathbf{I})$  и построен на базе традиционной методологии (формирование пространства признаков, формирование обучающих выборок по каждому классу, обучение). Второй основан на использовании точечного оценивания величины  $RUL$  по непараметрической регрессионной схеме, с дальнейшей идентификацией субинтервала из множества  $\mathbf{I}$ , которому принадлежит точечная оценка  $RUL$ . Интегральное классификационное решение выносится с использованием взвешенного оптимальным образом правила, которое будет подробно описано ниже. Для удобства записи определим для  $r_i^{(j,\beta)}(t)$  функцию  $Q$  такую, что  $Q(r_i^{(j,\beta)}(t)) = t$ .

Рассмотрим субкортеж  $r_i^{(j,\beta)}I_p^{(\lambda,\beta)} = r_i^{(j,\beta)}(t_1^{(j,\beta)}, r_i^{(j,\beta)}t_2^{(j,p)}, \dots)$ ;  $r_i^{(j,\beta)}I_p^{(\lambda,\beta)} \subseteq \mathbf{r}_i^{(j,\beta)}(T_j)$ ,  $\mathbf{r}_i^{(j,\beta)}(T_j) \in R_L(\lambda|\mathbf{T}(D_{\beta}^{(L)}))$ . Субкортеж  $r_i^{(j,\beta)}I_p^{(\lambda,\beta)}$  будем называть субтраекторией  $\lambda$ -завершенной траектории  $\mathbf{r}_i^{(j,\beta)}(T_j)$ . Каждую  $\lambda$ -завершенную траекторию из  $R_L(\lambda|\mathbf{T}(D_{\beta}^{(L)}))$  разбиваем на субтраектории  $r_i^{(j,\beta)}I_p^{(\lambda,\beta)}$  так, что:  $\forall i, j, I_p^{(\lambda,\beta)} \in \mathbf{I}: r_i^{(j,\beta)}(I_p^{(\lambda,\beta)}): (\forall k: Q(< r_i^{(j,\beta)}I_p^{(\lambda,\beta)}) > k) \in I_p^{(\lambda,\beta)}$ . Таким образом, все компоненты субтраектории  $r_i^{(j,\beta)}I_p^{(\lambda,\beta)}$  соответствуют субинтервалу  $I_p^{(\lambda,\beta)}$ . Будем называть эти субтраектории  $h_p$  — траекториями, где  $h_p$  — соответствующий множеству  $I_p^{(\lambda,\beta)}$  индекс из множества  $Index(\mathbf{I})$ .

Согласно условиям постановки задачи, размеченная база данных имеет следующий вид:

$$U_L = \{(r_i^{(j,\beta)}(I_A^{(\lambda,\beta)}), h_A), (r_i^{(j,\beta)}(I_W^{(\lambda,\beta)}), h_W), (r_i^{(j,\beta)}(I_N^{(\lambda,\beta)}), h_N) | i = 1, \dots, w, j = 1, \dots, q\}. \quad (5)$$

Рассмотрим следующую последовательность задач бинарной классификации:

— Задача 1: для единицы оборудования  $d_j$ , по наблюдениям  $R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$  определить:  $HI = h_A$  или  $HI \neq h_A$ ? При этом классы определены следующим образом:

- /1/-класс, название:  $HI_A$ :  $[HI = h_A]: \{(r_i^{(j,\beta)}(I_A^{(\lambda,\beta)}), 1) | i, j\} = U_L^A$ ;
- /-1/-класс, название:  $HI_{-A}$ :  $[HI \neq h_A]: \{(r_i^{(j,\beta)}(I_W^{(\lambda,\beta)}), -1) \cup (r_i^{(j,\beta)}(I_N^{(\lambda,\beta)}), -1) | i, j\} = U_L^{-A}$ .

Задача 2: для единицы оборудования  $d_j$ , по наблюдениям  $R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$  определить:  $HI = h_W$  или  $HI \neq h_W$ ? При этом классы определены следующим образом:

- /1/-класс, название:  $HI_W$ :  $[HI = h_W]: \{(r_i^{(j,\beta)}(I_W^{(\lambda,\beta)}), 1) | i, j\} = U_L^W$ ;
- /-1/-класс, название:  $HI_{-W}$ :  $[HI \neq h_W]: \{(r_i^{(j,\beta)}(I_A^{(\lambda,\beta)}), -1) \cup (r_i^{(j,\beta)}(I_N^{(\lambda,\beta)}), -1) | i, j\} = U_L^{-W}$ .

Задача 3: для единицы оборудования  $d_j$ , по наблюдениям  $R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$  определить:  $HI = h_N$  или  $HI \neq h_N$ ? При этом классы определены следующим образом:

- /1/-класс, название:  $HI_N$ :  $[HI = h_N]: \{(r_i^{(j,\beta)}(I_N^{(\lambda,\beta)}), 1) | i, j\} = U_L^N$ ;
- /-1/-класс, название:  $HI_{-N}$ :  $[HI \neq h_N]: \{(r_i^{(j,\beta)}(I_A^{(\lambda,\beta)}), -1) \cup (r_i^{(j,\beta)}(I_W^{(\lambda,\beta)}), -1) | i, j\} = U_L^{-N}$ .

Обозначим:  $\mathbf{I}_1 = \{I_A^{(\lambda,\beta)}, I_N^{(\lambda,\beta)} \cup I_W^{(\lambda,\beta)}\}$ ,  $\mathbf{I}_2 = \{I_W^{(\lambda,\beta)}, I_N^{(\lambda,\beta)} \cup I_A^{(\lambda,\beta)}\}$ ,  $\mathbf{I}_3 = \{I_N^{(\lambda,\beta)}, I_A^{(\lambda,\beta)} \cup I_W^{(\lambda,\beta)}\}$ .

Здесь и далее будет использоваться сокращение:  $\mathbf{r}_j = R_j(\lambda|t_s, \mathbf{T}(D_{\beta}^{(T)}))$ . Рассмотрим три бинарных классификатора  $B_l(\mathbf{r}_j|R_L^{(\lambda,\beta)})$ ,  $l = 1, 2, 3$ , такие, что:

$$\forall l, j : \mathbf{r}_j \xrightarrow{B_l(\mathbf{r}_j|R_L^{(\lambda,\beta)})} Index(\mathbf{I}_l).$$

Для повышения эффективности классификации, вместо величин  $\mathbf{r}_j$ , используется набор значений некоторых скалярных функций, вычисленных на  $\mathbf{r}_j$ . Эти функции принято называть признаками (feature). Исследования показали высокую эффективность признаков, в основе которых лежат различные варианты функции скользящего среднего, рассчитанного по длине «окна»  $N$ , а также функции, которые характеризуют дисперсионные характеристики этих величин. Будем обозначать этот набор следующим образом:  $F[\mathbf{r}_j]$ ,  $F[\cdot] \subseteq R^m$ . Соответственно, бинарные классификаторы будем записывать в виде  $B_l(F[\mathbf{r}_j]|R_L^{(\lambda,\beta)})$  и  $\forall l, j : F[\mathbf{r}_j] \xrightarrow{B_l(F[\mathbf{r}_j]|R_L^{(\lambda,\beta)})} Index(\mathbf{I}_l)$ . Для каждого классификатора из множества  $\mathbf{B} = \{B_l(F[\mathbf{r}_j]|R_L^{(\lambda,\beta)})\}$ ,

$l = 1, 2, 3$  определена функция надежности классификации  $\mu$ , которая характеризует достоверности принятия классификационного решения. Например, если классификатор  $C(x)$  принимает решение о том, что образец прошедших наблюдение данных  $x$  соответствует классу 1 с надежностью 0,2, а классу  $-1$  с надежностью 0,8, это записывается следующим образом:  $\mu(C(x)=1) = 0,2$ ,  $\mu(C(x)=-1) = 0,8$ . Часто эту функцию, порой, называют вероятностью принадлежности классифицируемого объекта к определенному классу, хотя функция  $\mu$  не имеет прямого отношения к вероятностной мере в ее «частотной» интерпретации. Тем не менее, формально функция  $\mu$  обладает основными свойствами функции вероятности и широко используется в ML для оценивания достоверности классификационного решения. Принципы определения функции  $\mu$  отличны для различных методов классификации, но описание этих принципов выходит за рамки настоящей статьи. Для бинарных классификаторов из множества  $\mathbf{B}$  можно записать:

$$\forall \mathbf{r}_j \forall l : (\sum_{k \in Index(\mathbf{I}_l)} \mu(B_l(F[\mathbf{r}_j] | R_L^{(\lambda, \beta)})) = k) = 1, \mu(B_l(F[\mathbf{r}_j] | R_L^{(\lambda, \beta)})) \in [0, 1].$$

Обозначим:

$$l^*(F[\mathbf{r}_j]) = Arg Sup_l (\mu(B_l(F[\mathbf{r}_j] | R_L^{(\lambda, \beta)})) = k),$$

$$\forall \mathbf{r}_j, k(l) \in Index(\cup_p \mathbf{I}_p) : k(l) = B_l(F[\mathbf{r}_j] | R_L^{(\lambda, \beta)}).$$

Классификационное решение для группы классификаторов из множества  $\mathbf{B}$  рассматривается в виде простой функции  $k(l^*(F[\mathbf{r}_j]))$ . В данном случае, каждый бинарный классификатор множества  $\mathbf{B}$  обучается на соответствующих множествах  $F[U_L^A]UF[U_L^{-A}]$ ,  $F[U_L^W]UF[U_L^{-W}]$  и  $F[U_L^N]UF[U_L^{-N}]$ . Здесь символом  $F[U]$  обозначено множество, в котором, вместо величины  $r_i^{(j, \beta)} I_p^{(\lambda, \beta)} \subseteq r_i^{(j, \beta)}(T_j)$  использован вектор признаков  $F[r_i^{(j, \beta)}]$ .

С целью повышения обобщающей способности обучение каждого классификатора из  $\mathbf{B}$  должно производиться в рамках концепции скользящего подхода (Cross-Validation) [18] в варианте контроля по отдельным объектам (leave-one-out CV или LOO CV). В качестве алгоритмов бинарной классификации классификатора могут рассматриваться различные методы, в том числе SVM, XGBoost [19], Gradient-Boosting [20], а также нейронные сети различных топологий. Классификационное решение  $k(l^*(F[\mathbf{r}_j]))$ , теоретически, позволяет определить текущее значение  $HI$  для единицы оборудования типа  $\beta$  при анализе на неисправность типа  $\lambda$ . С целью повышения надежности классификационного решения общий алгоритм его принятия был усложнен. В дополнение к методу  $k(l^*(F[\mathbf{r}_j]))$ , предлагается рассматривать метод непосредственного оценивания  $RUL$  на базе непараметрического регрессионного анализа, при этом регрессия строится по данным  $R_L(\lambda | \mathbf{T}(D_\beta^{(L)}))$ . Использование непараметрического регрессионного анализа обусловлено тем, что для различных типов оборудования, число которых может быть весьма значительным, разработка параметрических моделей не представляется разумной стратегией. Поэтому непараметрический регрессионный анализ, который исходит из минимальных предположений о типе восстанавливаемой регрессионной зависимости для величины  $RUL$ , представляется наиболее целесообразным с практической точки зрения. Как указывалось ранее, в рамках предложенного подхода, восстанавливается регрессионная зависимость величины  $RUL$  от  $\mathbf{r}_j$  с дальнейшей идентификацией конкретного субинтервала из множества  $\mathbf{I}$ , которому принадлежит точечная оценка  $RUL$ . Регрессионная функция обучается на данных  $R_L(\lambda | \mathbf{T}(D_\beta^{(L)}))$ . Таким образом, восстанавливается функция  $RUL(\mathbf{r}_j)$ , а дальнейшее решение принимается согласно следующей импликации:

$$(\exists I_p^{(\lambda, \beta)} \in \mathbf{I} : RUL(\mathbf{r}_j) \in I_p^{(\lambda, \beta)}) \Rightarrow (HI = h_p, h_p \in Index(\mathbf{I})).$$

Для удобства будем записывать это решение в следующем виде:

$$\forall j : \mathbf{r}_j \xrightarrow{C_{RUL}(\mathbf{r}_j)} Index(\mathbf{I}).$$

Для повышения эффективности восстановления функции  $RUL(\mathbf{r}_j)$  используются специальные методы сглаживания данных  $\{\mathbf{r}_j\}$ . Для сглаживания могут быть использованы различные варианты фильтра Калмана, методов скользящего среднего и kernel smoothers [21]. Будем обозначать функцию сглаживания данных  $\mathbf{r}_j$  символом  $S[\mathbf{r}_j]$ . В этом случае можно записать:

$$\forall j : S[\mathbf{r}_j] \xrightarrow{C_{RUL}(S[\mathbf{r}_j])} Index(\mathbf{I}).$$

Исследовалось несколько методов непараметрического регрессионного анализа для восстановления функции  $C_{RUL}(\cdot)$ . Наилучшие результаты показали методы SVR (Support Vector Regression) [22] и GBR (Gradient Boosting Regression) [23], которые обеспечили вполне сравнимое качество восстановления регрессионной зависимости  $C_{RUL}(\cdot)$  по данным  $R_L(\lambda | \mathbf{T}(D_\beta^{(L)}))$ , состоящим из  $\lambda$ -завершенных траекторий.

Пусть классификаторы  $C_{RUL}(S[\mathbf{r}_j])$  и  $k(l^*(F[\mathbf{r}_j]))$  на наборе данных  $F[\mathbf{r}_j]$  выдают следующие результаты:  $C_{RUL}(S[\mathbf{r}_j]) = i$  и  $k(l^*(F[\mathbf{r}_j])) = j$ . В общем случае  $i \neq j$ . Символами  $e_{RUL}\{i\}$  и  $e_{k(l^*)}\{j\}$  соответственно обозначим средние тотальные ошибки классификации на классах  $i$  и  $j$  для классификаторов  $C_{RUL}(\cdot)$  и  $k(l^*(\cdot))$ .



Оценки  $e_{RUL}\{i\}$  и  $e_{k(I^*)}\{j\}$  получены при тестировании классификаторов  $C_{RUL}(S[\mathbf{r}_j])$  и  $k(I^*)$  на множестве  $U_L$  из (5). Результирующее решение рассмотрим в следующем виде:

$$p(\mathbf{r}_j) = \begin{cases} [C_{RUL}(S[\mathbf{r}_j]) = k(I^*(F[\mathbf{r}_j])) = I] \Rightarrow I \\ [C_{RUL}(S[\mathbf{r}_j]) \neq k(I^*(F[\mathbf{r}_j]))] \Rightarrow \begin{cases} [e_{RUL}\{C_{RUL}(S[\mathbf{r}_j])\} > e_{k(I^*)}\{k(I^*(F[\mathbf{r}_j]))\}] \Rightarrow C_{RUL}(S[\mathbf{r}_j]), \\ [e_{RUL}\{C_{RUL}(S[\mathbf{r}_j])\} \leq e_{k(I^*)}\{k(I^*(F[\mathbf{r}_j]))\}] \Rightarrow k(I^*(F[\mathbf{r}_j])) \end{cases} \end{cases} \quad (6)$$

$p(\mathbf{r}_j) \in \text{Index}(\mathbf{I})$ . Нетрудно видеть, что структура решения (6) вполне соответствует структуре интегрального решения, которое исследовалось в работе [24]. Поэтому можно утверждать, что решение (6), согласно теореме 1 [25], обеспечивает минимум средней классификационной ошибки для экспоненциальной функции потерь.

### Результаты тестирования метода на C-MAPPS Dataset

Предложенный подход был подвергнут тестированию на широко известной базе данных C-MAPPS Dataset (NASA) [25], которая была представлена в [26] и содержит информацию об отказах турбовентиляторных двигателей, смоделированных с использованием термодинамической имитационной модели. Используемая имитационная модель имеет корректное физико-статистическое основание и признана научным сообществом адекватной. Набор данных C-MAPPS Dataset традиционно используется в качестве тестовой базы для сравнения различных систем ML PdM. Например, результаты такого тестирования отражены в работах [27–31], а также во множестве других. Модель, которая легла в основу генерации набора C-MAPPS Dataset, имеет 34 параметра. Различные комбинации значений этих параметров позволили ее авторам моделировать последствия отказов и типовых деградаций основных вращающихся компонентов турбовентиляторных двигателей в виде реакции моделируемой системы (для различных технических режимов и условий эксплуатации). Предполагается, что данные, составляющие набор C-MAPPS Dataset, представляют собой последовательность измерений состояния группы однотипных турбовентиляторных двигателей. Данные измерения получены от сенсорной сети  $\{s_i^{(j)}\}_{i=1, \dots, w}$ , которая однотипно размещена на поверхностях группы этих двигателей. Набор C-MAPPS Dataset состоит из четырех блоков (FD001, FD002, FD003, FD004), которые отличаются друг от друга конфигурацией деградационных режимов и типами моделируемых неисправностей. В исследованиях использовался блок FD001, который имеет следующие характеристики: моделируется одна причина деградации  $\lambda$ , а именно: выход из строя компрессора высокого давления; мощность сенсорной сети  $w = 26$ ; мощность множества двигателей  $D_\beta$  типа  $\beta$  равняется 100 как для обучающего набора (train\_FD001.txt), так и для набора, предназначенного для тестирования (test\_FD001.txt). В обучающем наборе длительность каждой  $\lambda$ -завершенной траектории всегда превосходит величину 100, но различна для каждой единицы оборудования (иногда в разы). Длительность тестирующих траекторий также довольно различна и также превосходит величину 100. Интервалы, определяющие длительности ступеней параметра  $HI$ , были определены следующим образом:  $I_A^{(\lambda, \beta)} = [0, 25]$ ,  $I_W^{(\lambda, \beta)} = [25, 125]$  и  $I_N^{(\lambda, \beta)} = [125, \infty]$  (аналогичная структура разбиения была предложена в [32]).

Целью тестирования является следующая задача: для каждой траектории  $\mathbf{r}_i^{(j, \beta)}(T_j)$  из тестового набора test\_FD001.txt необходимо оценить значение параметра  $HI$  единицы оборудования  $j$ , соответствующее моменту времени  $t_{|T_j|}^{(j)}$  завершения  $j$ -той траектории, и сравнить полученное значение параметра  $HI$  с его истинным значением, которое содержится в проверочном наборе RUL\_FD001.txt. Для моделирования использовалась среда Python 3.7, в том числе базовые пакеты *scipy* и *pandas*, а также пакет поддержки методов машинного обучения *sklearn*. Исходя из выбранного разбиения  $I_A^{(\lambda, \beta)}$ ,  $I_W^{(\lambda, \beta)}$ ,  $I_N^{(\lambda, \beta)}$ , на основании методики, изложенной в предыдущем разделе статьи, из набора RUL\_FD001 формируется структурированный обучающий набор  $U_L$ . Исходные данные для восстановления регрессии  $C_{RUL}(\cdot)$ , практически без модификации, представлены в том же наборе данных: RUL\_FD001. Для сглаживания данных использовался метод kernel smoothers. Результаты сглаживания данных хорошо иллюстрируются рисунком. Здесь по оси абсцисс отложен параметр  $t$  — номер цикла измерения (измеряется в единицах, сокращение — ед.), по оси ординат — значение признака  $\mathbf{r}_j$  (безразмерная величина — бр.).

Пространство признаков, которое исходно имело размерность 26, с использованием методов снижения размерности (метод «Minimum Redundancy Maximum Relevancy»: mRmR [33]) было сжато до 12. Для восстановления регрессии  $C_{RUL}(\cdot)$  использовался метод SVR, а в качестве группы бинарных классификаторов  $\mathbf{V}$  использовались XGBoost, обученные на соответствующих поднаборах из  $U_L$ , в рамках концепта Cross-Validation в варианте контроля по отдельным объектам (LOO CV). В соответствии с условиями постановки задачи система выдает решение в следующем виде:  $HI = h$ ,  $h \in \text{Index}(\mathbf{I}) = \{h_A, h_W, h_N\}$ . В результате тестирования были получены следующие результаты:

$$\mathbf{P}(HI = h_W | RUL \in I_W^{(\lambda, \beta)}) = 0,99, \mathbf{P}(HI = h_A | RUL \in I_A^{(\lambda, \beta)}) = 0,99, \mathbf{P}(HI = h_N | RUL \in I_N^{(\lambda, \beta)}) = 0,98.$$

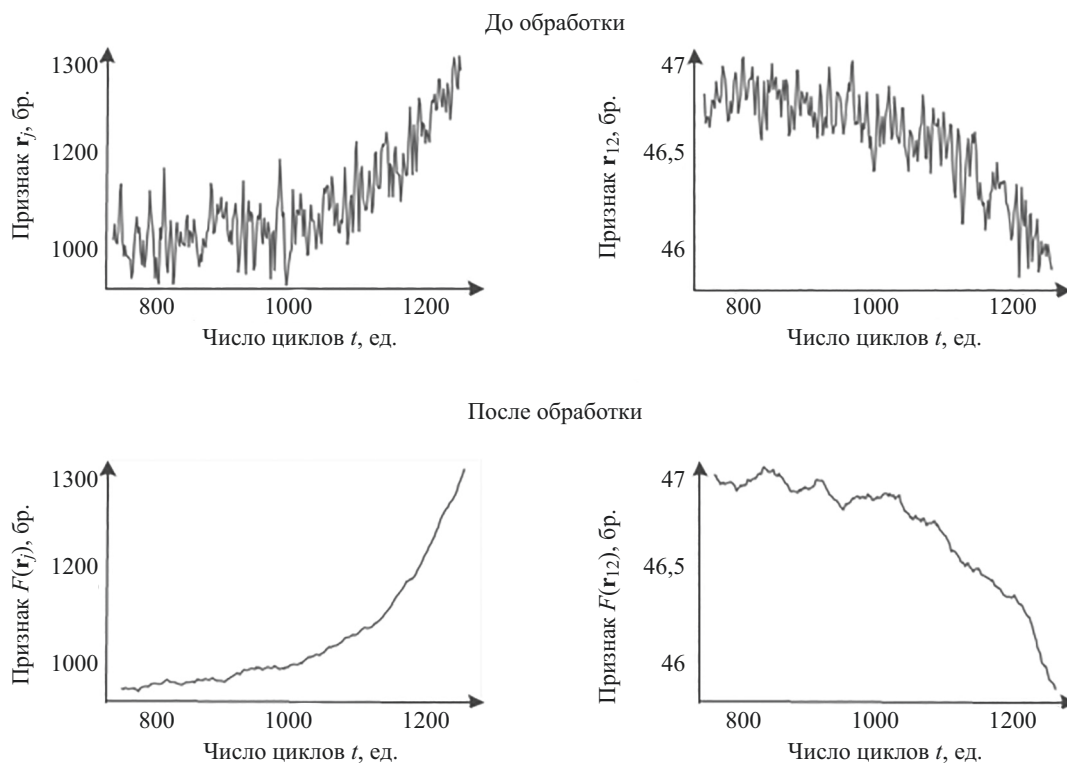


Рисунок. Сглаженные и исходные данные от двух элементов (7 и 12) сенсорной сети

Полученный результат вполне адекватен для практического использования предложенного метода, так как система определяет текущее значение  $HI$  с высокой достоверностью, при этом особенно важно то, что для практически важного случая  $RUL \in I_W^{(\lambda, \beta)}$  вероятность корректной оценки значения  $HI$  равняется 0,99. Средняя по множеству  $Index(\mathbf{I})$  вероятность правильной классификации равна 0,98. Для сравнения рассмотрим работу [33], в которой также используется ступенчатая  $HI$ , но иная схема классификации (не ансамблевая, как в настоящей работе, а обычная, построенная на классификаторах типа дерева решений, метод ближайшего соседа, SVM и naïve Bayes). Подход из работы [33] на том же тестовом блоке данных из C-MAPPS Dataset обеспечил величину средней по множеству  $Index(\mathbf{I})$  вероятности правильной классификации 0,78, а для случая  $RUL \in I_W^{(\lambda, \beta)}$  вероятность корректной оценки значения  $HI$  равняется 0,7.

### Заключение

Результаты моделирования показывают, что предложенный метод идентификации стадии деградации оборудования в системах сервисного обслуживания превентивного типа, основанный на использовании ступенчатой функции «индекс здоровья» и взвешенной комбинации двух классификаторов различной структуры, показал обнадеживающие результаты при его тестировании на базе данных C-MAPPS Dataset. Надежность идентификации стадии деградации оборудования, оцененная в результате этого тестирования, превосходит известные методы идентификации. Предложенная методика проста в реализации, хорошо масштабируется и обеспечивает хорошо интерпретируемые результаты.

### Литература

1. Peng Y., Dong M., Zuo M.-J. Current status of machine prognostics in condition-based maintenance: a review // *International Journal of Advanced Manufacturing Technology*. 2010. V. 50. N 1-4. P. 297–313. doi: 10.1007/s00170-009-2482-0
2. Grall A., Dieulle L., Berenguer C., Roussignol M. Continuous-time predictive-maintenance scheduling for a deteriorating system // *IEEE Transactions on Reliability*. 2002. V. 51. N 2. P. 141–150. doi: 10.1109/TR.2002.1011518
3. Ramasso E. Investigating computational geometry for failure prognostics // *International Journal of Prognostics and Health Management*. 2014. V. 5. N 1. P. 005.
4. Fumeo E., Oneto L., Anguita D. Condition based maintenance in railway transportation systems based on big data streaming analysis // *Procedia Computer Science*. 2015. V. 53. P. 437–446. doi: 10.1016/j.procs.2015.07.321

### References

1. Peng Y., Dong M., Zuo M.-J. Current status of machine prognostics in condition-based maintenance: a review. *International Journal of Advanced Manufacturing Technology*, 2010, vol. 50, no. 1-4, pp. 297–313. doi: 10.1007/s00170-009-2482-0
2. Grall A., Dieulle L., Berenguer C., Roussignol M. Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE Transactions on Reliability*, 2002, vol. 51, no. 2, pp. 141–150. doi: 10.1109/TR.2002.1011518
3. Ramasso E. Investigating computational geometry for failure prognostics. *International Journal of Prognostics and Health Management*, 2014, vol. 5, no. 1, pp. 005.
4. Fumeo E., Oneto L., Anguita D. Condition based maintenance in railway transportation systems based on big data streaming analysis. *Procedia Computer Science*, 2015, vol. 53, pp. 437–446. doi: 10.1016/j.procs.2015.07.321

5. Li Y., Billington S., Zhang C., Kurfess T., Danyluk S., Liang S. Adaptive prognostics for rolling element bearing condition // *Mechanical Systems and Signal Processing*. 1999. V. 13. N 1. P. 103–113. doi: 10.1006/mssp.1998.0183
6. Frelicot C. A fuzzy-based prognostic adaptive system // *Journal European des Systemes Automatises*. 1996. V. 30. N 2-3. P. 281–299.
7. Kothamasu R., Huang S.H., VerDuin W.H. System health monitoring and prognostics — a review of current paradigms and practices // *International Journal of Advanced Manufacturing Technology*. 2006. V. 28. N 9. P. 1012–1024. doi: 10.1007/s00170-004-2131-6
8. Goebel K., Saha B., Saxena A., Celaya J.R., Christophersen J.P. Prognostics in battery health management // *IEEE Instrumentation and Measurement Magazine*. 2008. V. 11. N 4. P. 33–40. doi: 10.1109/MIM.2008.4579269
9. Dong M., He D., Banerjee P., Keller J. Equipment health diagnosis and prognosis using hidden semi-Markov models // *International Journal of Advanced Manufacturing Technology*. 2006. V. 30. N 7-8. P. 738–749. doi: 10.1007/s00170-005-0111-0
10. Wang W.Q., Golnaraghi M.F., Ismail F. Prognosis of machine health condition using neuro-fuzzy systems // *Mechanical Systems and Signal Processing*. 2004. V. 18. N 4. P. 813–831. doi: 10.1016/S0888-3270(03)00079-7
11. Shetty P., Mylaraswamy D., Ekambaram T. A hybrid prognostic model formulation and health estimation of auxiliary power units // *Journal of Engineering for Gas Turbines and Power*. 2008. V. 130. N 2. P. 021601. doi: 10.1115/1.2795761
12. Lee J.-M., Yoo C., Lee I.-B. On-line batch process monitoring using a consecutively updated multiway principal component analysis model // *Computers and Chemical Engineering*. 2003. V. 27. N 12. P. 1903–1912. doi: 10.1016/S0098-1354(03)00151-0
13. Garga A.K., McClintic K.T., Campbell R.L., Yang C.-C., Lebold M.S., Hay T.A., Byington C.S. Hybrid reasoning for prognostic learning in CBM systems // *IEEE Aerospace Conference Proceedings*. 2001. V. 6. P. 2957–2969. doi: 10.1109/AERO.2001.931316
14. Ramasso E. Investigating computational geometry for failure prognostics in presence of imprecise health indicator: Results and comparisons on C-MAPSS datasets // *Proc. 2<sup>nd</sup> European Conference of the PHM Society 2014 (PHME14)*. 2014. V. 5.
15. Medjaher K., Zerhouni N., Baklouti J. Data-driven prognostics based on health indicator construction: Application to pronostia's data // *Proc. 12<sup>th</sup> European Control Conference (ECC 2013)*. 2013. P. 1451–1456.
16. Susto G.A., Schirru A., Pampuri S., McLoone S., Beghi A. Machine learning for predictive maintenance: A multiple classifier approach // *IEEE Transactions on Industrial Informatics*. 2015. V. 11. N 3. P. 812–820. doi:10.1109/TII.2014.2349359
17. Summary of the 2018 Department of defense artificial intelligence strategy. harnessing ai to advance our security and prosperity [Электронный ресурс]. URL: <https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF> (дата обращения: 18.10.2019)
18. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection // *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*. 1995. V. 2(12). P. 1137–1143.
19. Chen T., Guestrin C. XGBoost: A scalable tree boosting system // *Proc. 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13–17, 2016*. P. 785–794. doi:10.1145/2939672.2939785
20. Mason L., Baxter J., Bartlett P.L., Frean M. Boosting algorithms as gradient descent // *Proc. 13<sup>th</sup> Annual Neural Information Processing Systems Conference*. 1999. P. 512–518.
21. Li Q., Racine J.S. *Nonparametric econometrics: Theory and practice*. Princeton University Press, 2007. 768 p.
22. Zemel R.S., Pitassi T. A gradient-based boosting algorithm for regression problems // *Advances in Neural Information Processing Systems*. 2001. P. 696–702.
23. Smola A.J., Schölkopf B. A tutorial on support vector regression // *Statistics and Computing*. 2004. V. 14. N 3. P. 199–222. doi: 10.1023/B:STCO.0000035301.49549.88
24. Тимофеев А.В. Параметрическая оптимизация структуры консолидирующего решения в биометрических системах
5. Li Y., Billington S., Zhang C., Kurfess T., Danyluk S., Liang S. Adaptive prognostics for rolling element bearing condition. *Mechanical Systems and Signal Processing*, 1999, vol. 13, no. 1, pp. 103–113. doi: 10.1006/mssp.1998.0183
6. Frelicot C. A fuzzy-based prognostic adaptive system. *Journal European des Systemes Automatises*, 1996, vol. 30, no. 2–3, pp. 281–299.
7. Kothamasu R., Huang S.H., VerDuin W.H. System health monitoring and prognostics — a review of current paradigms and practices. *International Journal of Advanced Manufacturing Technology*, 2006, vol. 28, no. 9, pp. 1012–1024. doi: 10.1007/s00170-004-2131-6
8. Goebel K., Saha B., Saxena A., Celaya J.R., Christophersen J.P. Prognostics in battery health management. *IEEE Instrumentation and Measurement Magazine*, 2008. vol. 11, no. 4, pp. 33–40. doi: 10.1109/MIM.2008.4579269
9. Dong M., He D., Banerjee P., Keller J. Equipment health diagnosis and prognosis using hidden semi-Markov models. *International Journal of Advanced Manufacturing Technology*, 2006, vol. 30, no. 7-8, pp. 738–749. doi: 10.1007/s00170-005-0111-0
10. Wang W.Q., Golnaraghi M.F., Ismail F. Prognosis of machine health condition using neuro-fuzzy systems. *Mechanical Systems and Signal Processing*, 2004, vol. 18, no. 4, pp. 813–831. doi: 10.1016/S0888-3270(03)00079-7
11. Shetty P., Mylaraswamy D., Ekambaram T. A hybrid prognostic model formulation and health estimation of auxiliary power units. *Journal of Engineering for Gas Turbines and Power*, 2008, vol. 130, no. 2, pp. 021601. doi: 10.1115/1.2795761
12. Lee J.-M., Yoo C., Lee I.-B. On-line batch process monitoring using a consecutively updated multiway principal component analysis model. *Computers and Chemical Engineering*, 2003, vol. 27, no. 12, pp. 1903–1912. doi: 10.1016/S0098-1354(03)00151-0
13. Garga A.K., McClintic K.T., Campbell R.L., Yang C.-C., Lebold M.S., Hay T.A., Byington C.S. Hybrid reasoning for prognostic learning in CBM systems. *IEEE Aerospace Conference Proceedings*, 2001, vol. 6, pp. 2957–2969. doi: 10.1109/AERO.2001.931316
14. Ramasso E. Investigating computational geometry for failure prognostics in presence of imprecise health indicator: Results and comparisons on C-MAPSS datasets. *Proc. 2<sup>nd</sup> European Conference of the PHM Society 2014 (PHME14)*. 2014, vol. 5.
15. Medjaher K., Zerhouni N., Baklouti J. Data-driven prognostics based on health indicator construction: Application to pronostia's data. *Proc. 12<sup>th</sup> European Control Conference (ECC 2013)*, 2013, pp. 1451–1456.
16. Susto G.A., Schirru A., Pampuri S., McLoone S., Beghi A. Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 2015, vol. 11, no. 3, pp. 812–820. doi:10.1109/TII.2014.2349359
17. *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity*. Available at: <https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF> (accessed: 18.10.2019)
18. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*, 1995, vol. 2(12), pp. 1137–1143.
19. Chen T., Guestrin C. XGBoost: A scalable tree boosting system. *Proc. 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13–17, 2016*, pp. 785–794. doi:10.1145/2939672.2939785
20. Mason L., Baxter J., Bartlett P.L., Frean M. Boosting algorithms as gradient descent. *Proc. 13<sup>th</sup> Annual Neural Information Processing Systems Conference*, 1999, P. 512–518.
21. Li Q., Racine J.S. *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, 2007, 768 p.
22. Zemel R.S., Pitassi T. A gradient-based boosting algorithm for regression problems. *Advances in Neural Information Processing Systems*, 2001, pp. 696–702.
23. Smola A.J., Schölkopf B. A tutorial on support vector regression. *Statistics and Computing*, 2004, vol. 14, no. 3, pp. 199–222. doi: 10.1023/B:STCO.0000035301.49549.88

- идентификации личности // Научно-технический вестник информационных технологий, механики и оптики. 2014. Т. 14. № 3(91). С. 96–102.
25. Saxena A., Goebel K. Turbofan engine degradation simulation data set [Электронный ресурс]. URL: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository> (дата обращения: 14.06.2018)
  26. Saxena A., Goebel K., Simon D., Eklund N. Damage propagation modeling for aircraft engine run-to-failure simulation // Proc. International Conference on Prognostics and Health Management (PHM 2008). 2008. P. 22–34. doi: 10.1109/PHM.2008.4711414
  27. Al-Salah T.H., Zein-Sabatto S., Bodruzzaman M. Decision fusion software system for turbine engine fault diagnostics // Proc. IEEE SoutheastCon, (SOUTHEASTCON 2012). 2012. P. 1–6. doi: 10.1109/SECon.2012.6197000
  28. Coble J., Hines J.W. Applying the general path model to estimation of remaining useful life // International Journal of Prognostics and Health Management. 2011. V. 2. P. 1–13.
  29. Hu C., Youn B., Wang P., Taek Yoon J. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life // Reliability Engineering and System Safety. 2012. V. 103. P. 120–135. doi: 10.1016/j.res.2012.03.008
  30. Jianzhong S., Hongfu Z., Haibin Y., Pecht M. Study of ensemble learning-based fusion prognostics // Proc. Prognostics and System Health Management Conference, PHM'10. 2010. P. 5414582. doi: 10.1109/PHM.2010.5414582
  31. Liu K., Gebrael N.Z., Shi J. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis // IEEE Transactions on Automation Science and Engineering. 2013. V. 10. N 3. P. 652–664. doi: 10.1109/TASE.2013.2250282
  32. Pitz K., Anderl R. Implementing clustering and classification approaches for big data with MATLAB // Advances in Intelligent Systems and Computing. 2019. V. 880. P. 458–480. doi: 10.1007/978-3-030-02686-8\_35
  33. Peng H.C., Long F., Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2005. V. 27. N 8. P. 1226–1238. doi: 10.1109/TPAMI.2005.159
  24. Timofeev A.V. Parametric optimization of the multimodal decision-level fusion scheme in automatic biometric person's identification systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2014, vol. 14, no. 3(91), pp. 96–102. (in Russian)
  25. Saxena A., Goebel K. *Turbofan Engine Degradation Simulation Data Set*. Available at: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository> (accessed: 14.06.2018)
  26. Saxena A., Goebel K., Simon D., Eklund N. Damage propagation modeling for aircraft engine run-to-failure simulation. *Proc. International Conference on Prognostics and Health Management (PHM 2008)*, 2008, pp. 22–34. doi: 10.1109/PHM.2008.4711414
  27. Al-Salah T.H., Zein-Sabatto S., Bodruzzaman M. Decision fusion software system for turbine engine fault diagnostics. *Proc. IEEE SoutheastCon, (SOUTHEASTCON 2012)*, 2012, pp. 1–6. doi: 10.1109/SECon.2012.6197000
  28. Coble J., Hines J.W. Applying the general path model to estimation of remaining useful life. *International Journal of Prognostics and Health Management*, 2011, vol. 2, pp. 1–13.
  29. Hu C., Youn B., Wang P., Taek Yoon J. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering and System Safety*, 2012, vol. 103, pp. 120–135. doi: 10.1016/j.res.2012.03.008
  30. Jianzhong S., Hongfu Z., Haibin Y., Pecht M. Study of ensemble learning-based fusion prognostics. *Proc. Prognostics and System Health Management Conference, PHM'10*, 2010, pp. 5414582. doi: 10.1109/PHM.2010.5414582
  31. Liu K., Gebrael N.Z., Shi J. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. *IEEE Transactions on Automation Science and Engineering*, 2013, vol. 10, no. 3, pp. 652–664. doi: 10.1109/TASE.2013.2250282
  32. Pitz K., Anderl R. Implementing clustering and classification approaches for big data with MATLAB. *Advances in Intelligent Systems and Computing*, 2019, vol. 880, pp. 458–480. doi: 10.1007/978-3-030-02686-8\_35
  33. Peng H.C., Long F., Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, vol. 27, no. 8, pp. 1226–1238. doi: 10.1109/TPAMI.2005.159

#### Авторы

**Тимофеев Андрей Владимирович** — доктор технических наук, научный директор, ТОО «Эквалайзум», Астана, 010000, Казахстан, Scopus ID: 56689367600, ORCID ID: 0000-0001-7212-5230, [timofeev.andrey@gmail.com](mailto:timofeev.andrey@gmail.com)  
**Денисов Виктор Маевич** — доктор технических наук, доцент, генеральный директор, ООО «Флагман Гео», Санкт-Петербург, 197376, Российская Федерация; профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, Scopus ID: 56689367600, ORCID ID: 0000-0002-1907-5198, [070255@gmail.com](mailto:070255@gmail.com)

#### Authors

**Andrey V. Timofeev** — D.Sc., Director for Science, EqualZoom, LLP, Astana, 010000, Kazakhstan, Scopus ID: 56689367600, ORCID ID: 0000-0001-7212-5230, [timofeev.andrey@gmail.com](mailto:timofeev.andrey@gmail.com)  
**Victor M. Denisov** — D.Sc., Associate Professor, CEO, Flagman Geo, ООО, Saint Petersburg, 197376, Russian Federation; Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, Scopus ID: 56689367600, ORCID ID: 0000-0002-1907-5198, [070255@gmail.com](mailto:070255@gmail.com)