

На правах рукописи



Жуков Алексей Витальевич

**МЕТОДЫ КЛАССИФИКАЦИИ И ВОССТАНОВЛЕНИЯ  
РЕГРЕССИИ НА ОСНОВЕ КОМПОЗИЦИЙ РЕШАЮЩИХ  
ДЕРЕВЬЕВ С ПРИЛОЖЕНИЕМ К ЗАДАЧЕ ОЦЕНКИ  
РЕЖИМНОЙ НАДЕЖНОСТИ ЭЭС**

Специальность 05.13.18 — Математическое моделирование, численные  
методы и комплексы программ

**АВТОРЕФЕРАТ**

диссертации на соискание учёной степени  
кандидата технических наук

Иркутск — 2022



## Общая характеристика работы

**Актуальность темы.** Математические методы классификации и восстановления регрессии на основе машинного обучения всё чаще применяются в задачах современной электроэнергетики. Это связано как с усложнением современных электроэнергетических систем (ЭЭС), так и с целым рядом свойств методов машинного обучения. Современные ЭЭС становятся всё более сложными для аналитического анализа, но в то же время, внедрение большого количества устройств мониторинга позволяет получать колоссальный объём информации о состоянии ЭЭС. Такое положение дел открывает возможности для анализа ЭЭС как систем типа «чёрный ящик» с помощью методов машинного обучения, применяя уже имеющиеся методы науки о данных. Однако такие методы не могут полностью заменить классические подходы анализа ЭЭС, а лишь позволяют компенсировать некоторые их недостатки. К таким недостаткам зачастую относятся низкая скорость работы, необходимость в полной и точной информации о параметрах ЭЭС, в то время как в реальности доступна лишь их часть. Чтобы преодолеть упомянутые недостатки, задачу можно свести к построению модели классификации или восстановления регрессии, которая устанавливает соответствие режимных параметров ЭЭС и параметров среды с целевой переменной на основе выборки данных с известным значением целевой переменной. Другими словами, методы машинного обучения строят аппроксимацию целевой зависимости, позволяя получать приближенное решение, но за меньшее время и используя меньший набор входных переменных. Одной из таких задач является оценка режимной надёжности. Под режимной надёжностью (РН) ЭЭС понимается свойство системы сохранять заданные режимы функционирования при изменении условий, отказах элементов и внезапных возмущениях. Оценка РН заключается в определении допустимых значений режимных параметров, а также моделировании отказов с учётом выбранных критериев РН. Зачастую задача оценки режимной надёжности сопряжена с необходимостью большого объёма вычислений, поэтому традиционно при моделировании используются методы сокращения количества рассматриваемых случаев, а также рассмотрение упрощённых линейных моделей ЭЭС. Альтернативой являются методы на основе машинного обучения, которые способны построить модель достаточно точно аппроксимирующую традиционные показатели режимной надёжности. Таким образом, необходимость в полном моделировании ЭЭС возникает лишь на этапе построения модели, но отсутствует на этапе применения, что позволяет использовать такие модели в темпе процесса. В то же время, существует разрыв между современными математическими методами классификации и восстановления регрессии общего назначения и методами, применяемыми в конкретных задачах современной энергетики в целом и оценки режимной надёжности в частности. Так, на практике зачастую ограничиваются традиционными статистическими методами анализа данных. К таким подходам можно отнести авторегрессию интегрированного скользящего среднего и его модификации, построенные с помощью методологии Бокса – Дженкинса.

Такие методы имеют чёткое математико-статистическое обоснование и подробно описанную методику подбора параметров. Однако, за последние несколько десятилетий были созданы методы на основе машинного обучения, которые во многих задачах существенно превосходят традиционные методы анализа временных рядов по точности и времени работы. К таким методам относятся, например, искусственные нейронные сети (ИНС), благодаря их способности аппроксимировать сложные нелинейные зависимости. Наряду с методами на основе ИНС, в задачах восстановления регрессии эффективно применяется метод опорных векторов (МОВ), а также композиционные методы на основе таких математических моделей, как решающее дерево (РД). Композиционными называют методы, объединяющие в себе несколько алгоритмов машинного обучения для получения наиболее точного результата, чем каждый из алгоритмов отдельно. Одним из таких методов является случайный лес (СЛ), который всё чаще применяется для оценки режимной надёжности, ввиду малого количества параметров модели, а также эмпирически подтверждённой высокой точности и стойкостью к переобучению. Исследованию такого подхода посвящены работы Негневицкого М., Kamwa I., Vittal V. и др. Эффективность методов на основе СЛ подтвердили испытания, проведённые при поддержке системного оператора ЭЭС Канады и Дании<sup>1</sup>. Однако значимым недостатком большинства существующих методов, является предположение, что рассматриваемый процесс подчиняется одному закону на всей области определения, то есть является стационарным. В то время как реальные системы ввиду своей сложности зачастую имеют нестационарную природу. С целью преодоления этих сложностей создано множество как специализированных, так и общих методов. К первым можно отнести методы анализа нестационарных нелинейных динамических систем на основе интегральных уравнений в работах Апарцина А.С., Сидорова Д.Н., Smola A., Schölkopf B. и др. Ко вторым относятся разрабатываемые в настоящее время гибридные и потоковые методы на основе машинного обучения, которые способны адаптировать модель системы к изменениям в темпе процесса, что позволяет учесть его нестационарный характер. Исследованию подобных методов посвящены труды Моттля В.В., Туркова П.А., Saffari A., Tsybmal A., Žliobaitė I. и др. Такие подходы показали высокую эффективность во многих областях науки, однако их применение в задачах современной электроэнергетики представлено крайне скудно. Отсюда вытекает необходимость в адаптации данного подхода к задачам современной электроэнергетики.

Таким образом, диссертация посвящена совершенствованию методов классификации и восстановления регрессии в задачах современной энергетики. В качестве такой задачи подробно рассматривается оценка режимной надёжности современной ЭЭС со сложной многоконтурной структурой, в которой

---

<sup>1</sup>Kaci A. et al. Synchrophasor data baselining and mining for online monitoring of dynamic security limits //IEEE transactions on power systems. – 2014. – Vol. 29. – No. 6. – pp. 2681-2695;  
Liu C. et al. Dynamic Security Assessment of Western Danish Power System Based on Ensemble Decision Trees, IET Conference Proceedings – 2014. – p. 12.78-12.78

преобладают линии электропередачи средней протяженности, что соответствует электроэнергетическим системам крупных городов и мегаполисов. Другой чертой рассматриваемой ЭЭС является стохастический характер её поведения, что происходит вследствие внедрения возобновляемых источников энергии, наличия активных потребителей и рыночных факторов. Высокая сложность анализа таких ЭЭС приводит к необходимости разработки новых методов оценки режимной надёжности, соответствующих современным требованиям по их точности, ресурсоёмкости, достоверности и быстрдействию.

Решению различных аспектов рассматриваемой задачи посвящены работы Воропая Н.И., Гамма А.З., Глазуновой А.М., Готман Н.Э., Колосок И.Н., Курбацкого В.Г., Манова Н.А., Манусова В.З., Старцевой Т.Б., Чукуреева Ю.Я., Шумиловой Г.П. и др.

**Целью** данной работы является совершенствование методов классификации и восстановления регрессии в задачах современной электроэнергетики.

**Объект исследования** диссертационной работы – современная ЭЭС.

**Предметом исследования** диссертационной работы являются интеллектуальные методы и модели классификации и прогнозирования на основе композиций решающих деревьев.

Для достижения поставленной цели решены следующие **задачи**:

1. Анализ современных композиционных моделей и методов классификации и восстановления регрессии для анализа сложных процессов в ЭЭС.
2. Разработка модели и эффективного численного метода потоковой классификации по ретроспективным данным в темпе процесса.
3. Разработка метода оценки режимной надёжности с применением предложенного метода.
4. Разработка комплекса программ для проведения экспериментального исследования предложенных моделей и методов.

**Научная новизна:**

1. Предложена оригинальная непараметрическая модель на основе Случайного леса для классификации потоковых данных.
2. Для предложенной модели разработан эффективный численный метод потоковой классификации на основе модифицированного метода Случайный лес, использующий адаптивное взвешивание решающих деревьев, а также удаление неэффективных элементов композиции.
3. Разработан новый метод оценки режимной надёжности в темпе процесса с использованием предложенного потокового метода, устойчивый к изменениям в поведении ЭЭС.

**Практическая значимость** результатов исследования заключается в следующем:

1. Разработанная модель и метод позволяют решать задачи классификации, возникающие при анализе нестационарных процессов в ЭЭС.

2. Разработанный метод позволяет оценивать надёжность режима ЭЭС с целью недопущения аварийных режимов, а также оценивать необходимые инъекции реактивной мощности для отведения системы от границ неустойчивости и может быть задействован для поддержки принятия решений диспетчером ЭЭС.
3. Разработанные модель, метод и комплекс программ могут применяться для решения более широкого круга задач, связанных с анализом нестационарных данных в иных задачах современной электроэнергетики.

**Методология и методы исследования.** В работе используются методы математического моделирования, обработки сигналов, анализа временных рядов, машинного обучения и распознавания образов. Программная реализация выполнена с использованием объектно-ориентированного подхода средствами языков C++, Python и R, а также пакета прикладных программ Matlab.

**Основные положения, выносимые на защиту:**

1. Поточковая непараметрическая модель на основе Случайного леса, а также численный метод для её обучения, позволяющие эффективно решать задачу классификации в условиях нестационарных исходных данных. Соответствует п. 1 паспорта спец. 05.13.18<sup>2</sup>.
2. Численный метод оценки режимной надёжности в темпе процесса на основе предложенной модели, отличающийся способностью адаптации к нестационарности, позволяющий повысить эффективность оценки по сравнению с другими моделями. Соответствует п. 3 паспорта спец. 05.13.18.<sup>3</sup>
3. Численный метод оценки режимной надёжности на основе полностью рандомизированных решающих деревьев, способный давать численную оценку состояния как ЭЭС в целом, так и каждого узла ЭЭС, а также вычислять рекомендуемые инъекции реактивной мощности, необходимые для повышения устойчивости системы. Соответствует п. 3 паспорта спец. 05.13.18.<sup>3</sup>
4. Комплекс программ для проведения экспериментального исследования разработанных численных методов в задаче оценки режимной надёжности. Соответствует п. 4 паспорта спец. 05.13.18.<sup>4</sup>

**Достоверность** полученных результатов обеспечивается корректным использованием методов математического моделирования, обработки сигналов, анализа временных рядов, машинного обучения и подтверждается результатами экспериментальных исследований, которые находятся в соответствии с результатами, полученными другими авторами.

---

<sup>2</sup>П.1 «Разработка новых математических методов моделирования объектов и явлений»

<sup>3</sup>П.3 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий»

<sup>4</sup>П.4 «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента»

**Апробация работы.** Основные результаты работы докладывались на: конференции «XVI Байкальская международная школа-семинар. Методы оптимизации и их приложения» (о. Ольхон, 2014), VII Международном симпозиуме «Обобщенные постановки и решения задач управления (GSSCP-2014)» (Геленджик – Дивноморское, 2014), конференции «Математические методы распознавания образов» (Светлогорск, Россия, 2015), международной конференции «IEEE PowerTech» (Эйдинховен, Нидерланды, 2015), международной конференции «International Youth Conference on Energy» (Пиза, Италия, 2015), 5-й международной конференции «Анализ изображений, социальных сетей и текстов» (Екатеринбург, 2016), «2nd International Workshop on Mathematical Modeling of Smart Energy and Power Systems» (Чанша, Китай, 2019); Результаты диссертации получены при поддержке грантов РФФ № 14-19-00054, 19-49-04108 (рук. Н.И.Воропай) и РФФИ № 20-48-383004, 18-31-00206 (рук. А.В.Жуков), РФФИ-ГФЕН № 19-58-53011 (рук. Д.Н.Сидоров). По мере получения результаты работы обсуждались на семинарах ИГУ, ИСЭМ СО РАН, ИСЗФ СО РАН, а также на конференциях научной молодежи ИСЭМ СО РАН.

**Личный вклад.** В работах, опубликованных в соавторстве, соискателю принадлежит разработка математических моделей и методов, реализация и тестирование методов в программно-вычислительных комплексах, анализ результатов. На защиту выносятся материалы, полученные лично соискателем. Конфликта интересов с соавторами нет.

**Публикации.** Основные результаты по теме диссертации изложены в 21 печатном издании, 2 из которых изданы в журналах, рекомендованных ВАК [1; 2], в 2 монографиях (главы 3,5,7 [17] и глава 5 [18]), 4 статьи в рецензируемых журналах [3—6], 13 – в материалах конференций [7—10; 19; 11; 20; 21; 12—16]. Получено свидетельство о государственной регистрации программы для ЭВМ № 2016661345. Зарегистрир. 6 октября 2016 года [22]. Материалы настоящей диссертации внедрены в учебный процесс на кафедре электроснабжения и электротехники ИРНТУ.

## Содержание работы

Во **введении** показана актуальность исследований, проводимых в рамках данной диссертационной работы, сформулированы цель и задачи, необходимые для её достижения, представлена общая характеристика работы. Сформулированы положения, выносимые на защиту.

**Первая глава** содержит обзор, постановку задачи оценки режимной надёжности ЭЭС, а также описание современных методов и моделей классификации и восстановления регрессии для анализа сложных процессов, протекающих в электроэнергетических системах (ЭЭС). Особое внимание уделяется методам на основе машинного обучения, а также их применению к задаче оценки режимной надёжности, в частности, устойчивости по напряжению.

Традиционно задача оценки режимной надёжности включает в себя, прежде всего, математическую модель ЭЭС, которая задаётся набором нелинейных дифференциальных и алгебраических уравнений как

$$\dot{x} = f(x,y), \quad 0 = g(x,y), \quad (1)$$

где  $x \in \mathbb{R}^n$  – переменные состояния ЭЭС (активная и реактивная мощность  $P$  и  $Q$ ),  $y \in \mathbb{R}^m$  – алгебраические переменные (такие как напряжение  $U$ , фазовый угол  $\delta$  и другие), дифференциальная часть  $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ , алгебраическая часть  $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ .

Различные модели представляют  $f$  и  $g$  по-разному в зависимости от конкретной задачи и объекта исследования. В этой работе в рамках задачи оценки режимной надёжности в качестве объекта рассматривается современная ЭЭС со сложной многоконтурной структурой, в которой преобладают линии электропередачи средней протяженности. Другой чертой исследуемого объекта является стохастический характер поведения такой ЭЭС. Что происходит вследствие внедрения возобновляемых источников энергии, наличия активных потребителей и рыночных факторов. Другим словами, поведение может меняться со временем, а значит анализируемые процессы могут быть нестационарными.

Анализ опасных ситуаций в ЭЭС со сложной многоконтурной структурой показал, что неустойчивость по напряжению является одной из основных причин крупных аварий [6]. Как отмечено в [17], большинство системных аварий развивалось по сценарию среднесрочной неустойчивости по напряжению, когда аварийные процессы протекали относительно медленно (от одной до пяти минут). Поэтому для рассмотрения задачи оценки режимной надёжности в этом случае можно ограничиться математическим описанием ЭЭС без уравнений динамики с использованием лишь моделирования потокораспределения как основного метода исследования, то есть для описания системы используется лишь алгебраическая часть  $g$  (1). В общем случае устойчивость по напряжению может быть описана якобианом установившегося режима ЭЭС. При этом линеаризованные уравнения установившегося режима в точке решения могут быть представлены в следующем виде:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} \frac{\partial P}{\partial \delta} & \frac{\partial P}{\partial U} \\ \frac{\partial Q}{\partial \delta} & \frac{\partial Q}{\partial U} \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta U \end{bmatrix}, \quad (2)$$

где  $\Delta P, \Delta Q, \Delta \delta, \Delta U$ , – соответственно, векторы инкрементных изменений инъекций активной и реактивной мощностей, а также углов и напряжений в узлах ЭЭС. Существует ряд численных показателей (индексов) близости ЭЭС к коллапсу по напряжению, которые рассчитываются как на основе якобиана, так и напрямую на основе параметров режима. Однако зачастую оказывается, что расчёт таких индексов не удаётся произвести в темпе процесса в силу вычислительной сложности. Чтобы решить эту проблему предложено на основе выборки, полученной с помощью традиционных методов моделирования ЭЭС, строить аппроксимацию

индекса методами машинного обучения. Таким образом, строится алгоритмическая модель. Часто её структура не задана априори, а определяется данными, то есть она является непараметрической.

В общем виде задача для стационарного случая может быть сформулирована следующим образом: имеем обучающую выборку  $L = \{X, Y\}$  из  $n$  примеров,  $X$  – множество параметров ЭЭС;  $x_i$  – измеряемый параметр системы;  $Y$  – множество состояний системы. В случае классификации  $Y \in 1, \dots, c$ , где  $c$  – количество классов. Таким образом,  $Y$  – конечное множество номеров меток классов (например, надёжный и ненадёжный). Если же решается задача восстановления регрессии, то  $Y \in \mathbb{R}$ , то есть задаётся непрерывной величиной, например, как вероятность аварии. При этом существует неизвестная целевая зависимость – отображение  $y^* : X \rightarrow Y$ , значения которой известны только на объектах конечной обучающей выборки  $X^m = \{(x_1, y_1), \dots, (x_m, y_m)\}$ . Требуется построить алгоритмическую модель  $a : X \rightarrow Y$ , способную классифицировать (или найти значение отклика) новый набор параметров системы  $x \in X$ . Под алгоритмической моделью имеется в виду модель, представленная в форме алгоритма. В качестве критерия качества алгоритмической модели часто используется ошибка на проверочной выборке  $X^l = \{(\check{x}_1, \check{y}_1), \dots, (\check{x}_l, \check{y}_l)\}$ , называемая ошибкой обобщения.

Для прогнозирования нестационарных процессов часто используются потоковые методы, либо гибридные методы, сочетающие в себе специализированные способы предобработки (например, преобразование Гильберта-Хуанга), а также методы искусственного интеллекта (в частности методы машинного обучения). Далее будем называть последние интеллектуальными методами. Также обозначим как интеллектуальные модели построенные с их помощью. Кроме традиционных интеллектуальных методов, таких, как авторегрессия проинтегрированного скользящего среднего (ARIMA), используются методы на основе искусственных нейронных сетей (ИНС), решающих деревьев (РД), метода опорных векторов (МОВ) и других.

Основой для многих эффективных композиционных моделей является РД, которое представляет из себя логический алгоритм классификации (и восстановления регрессии), построенный на основе рекурсивного разделения пространства признаков. Логическим он называется вследствие того, что содержит в каждом узле предикат, при этом выбор предиката происходит при построении дерева на каждом уровне при помощи максимизации некоторого критерия информативности. К методам построения РД относятся такие методы как ID3, C4.5, Cart и др.<sup>5</sup>.

Пример такого дерева для количественного признака  $A \in \mathbb{R}$  и номинального признака  $B \in \{1, 2, \dots, n\}$ ,  $n \in \mathbb{N}$  показан на рисунке 1, где  $V$  и  $S$  – соответствующие возможные значения  $A$  и  $B$ , по которым разделяется пространство признаков.

---

<sup>5</sup>Sharma H., Kumar S. A survey on decision tree algorithms of classification in data mining //International Journal of Science and Research (IJSR). – 2016. – Vol. 5. – No. 4. – Pp. 2094-2097.

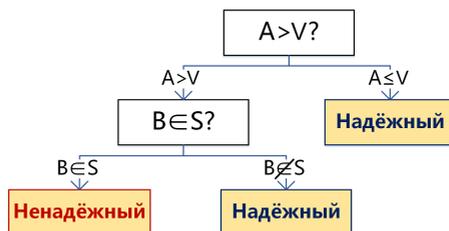


Рис. 1 — Пример изображения модели решающего дерева для дискретной оценки надёжности.

Одним из наиболее эффективных интеллектуальных методов, используемых в задачах электроэнергетики, является СЛ, предложенный Л. Брэйманом, который позволяет аппроксимировать классические показатели надёжности при меньших вычислительных затратах и большей устойчивости к неверным входным данным <sup>6</sup>.

Для построения модели СЛ на основе обучающей выборки  $L$  строятся подвыборки с повторением, на основе которых создаются  $K$  рандомизированных решающих деревьев, которые являются базовыми алгоритмами композиции. Рандомизация заключается в том, что на каждом шаге построения дерева случайно равномерно выбирается подмножество атрибутов размера  $M_{try}$ , по которым происходит разбиение. Итоговое значение выбирается путём голосования по большинству. Так для конкретного объекта  $x \in X$  решение в случае классификации выглядит как  $H(x) = \operatorname{argmax}_{y \in Y} \sum_{k=1}^K [\hat{h}_k(x) = y]$ , где  $\hat{h}_k(x)$  решающее правило  $k$ -го решающего дерева композиции, а сама функция  $H$  является так называемой корректирующей операцией, так как определяет способ получения итоговой метки класса из набора меток базовых алгоритмов. Из-за описанных выше свойств и высокой эффективности СЛ в исследуемом круге задач, он выбран как основа для создания потоковой математической модели классификации.

**Интеллектуальные методы оценки режимной надёжности ЭЭС.** Рассмотрены различные интеллектуальные методы и модели исследования режимной надёжности ЭЭС. По способу построения модели, методы исследования режимной надёжности можно разделить на методы полного моделирования и приближенные методы. Методы полного моделирования, ввиду комбинаторной природы проблемы, имеют большую вычислительную сложность, и соответственно, требуют много времени и поэтому часто не могут быть эффективно использованы в режиме реального времени. Хотя в некоторых случаях эта задача может быть решена за счёт технологических средств, таких, как параллельные вычисления. Далее мы будем рассматривать только приближенные методы, которые также можно условно разделить на **традиционные** и **интеллектуальные**.

<sup>6</sup>Kamwa I., Samantaray S. R., Joos G. Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices //IEEE Transactions on Smart Grid. – 2010. – Vol. 1. – No. 2. – Pp. 144-158.

Традиционно, чтобы уменьшить вычислительные затраты при решении этой задачи, системы диспетчерского управления, как правило, используют один или несколько индикаторов оценки надёжности, таких, как матрицы чувствительности, коэффициенты распределения, показатели, получаемые при быстром расчёте потокораспределения и т.п. Такие приближённые подходы позволяют снизить количество критических ситуаций в ЭЭС, которые необходимо исследовать в результате моделирования. Также возможно формирование показателя надёжности режима ЭЭС в виде дискретного набора классов (например, надёжный/ненадёжный). Ещё одним преимуществом интеллектуального подхода является возможность некоторых моделей адаптироваться к изменениям в системе. Ведь такие изменения могут привести к потере качества традиционных моделей или необходимости их полной перестройки, в то время как потоковые методы машинного обучения позволяют обновлять модель в процессе работы. При этом высокая скорость работы даёт возможность использовать их как для статической, так и для динамической оценки режимной надёжности.

Во **второй главе** предложены непараметрическая модель классификации, работающая в темпе процесса, и численный метод её обучения, а также методы решения задачи оценки режимной надёжности.

**Предлагаемая методика решения задач классификации и прогнозирования с помощью моделей Случайного леса.** Аппарат композиций решающих деревьев является мощным средством для решения задач классификации, регрессии и отбора наиболее информативных признаков, которые часто возникают при анализе поведения сложных систем, таких, как современные ЭЭС. Таким образом, пользуясь всеми возможностями композиций деревьев решений и опираясь на индустриальный стандарт CRISP-DM<sup>7</sup>, можно представить общий способ решения целого круга актуальных задач электроэнергетики, который состоит из следующих шагов: предобработка данных; извлечение признаков; отбор признаков с использованием оценки информативности входных переменных на основе композиции РД; подбор оптимальных параметров модели поиском по сетке; обучение итоговых моделей.

Первой стадией является подготовка данных: устранение выбросов и пропусков. Хотя Случайный лес содержит возможность устранения этих нежелательных явлений, это не рассматривается в рамках данной работы. Однако, допускается, что данные могут быть повреждены в результате неполадок в каналах связи или действий злоумышленников.

Второй стадией является подготовка набора признаков, которые потенциально могут быть использованы для построения решения. Этот этап наиболее сильно связан с предметной областью. Что касается задачи оценки РН, то так как работа ЭЭС зависит от погодных условий, состава сетевого и генерирующего оборудования, объемов резерва активной и реактивной мощности, текущего режима функционирования системы (значений узловых напряжений, потоков

---

<sup>7</sup>Chapman P. et al. CRISP-DM 1.0 Step-by-step data mining guide. – 2000.

мощности по межсистемным связям, загрузки сетевого оборудования и т.п.), все эти показатели могут быть использованы в качестве признаков.

Следующим этапом является отбор признаков для получения «оптимального» (в смысле максимальной точности) набора. К тому же, уменьшение количества признаков дополнительно уменьшает влияние проклятья размерности. Традиционно для её сокращения используется метод главных компонент или многомерное шкалирование. Однако, из-за сложной нелинейной природы данных, эти методы не всегда применимы. Поэтому было предложено использовать Случайный лес, так как он позволяет работать с данными большой размерности. Его эффективность подтверждена экспериментами в таких областях как анализ генома<sup>8</sup>, где размерность часто превосходит количество примеров обучающей выборки.

Что касается непосредственно отбора признаков, то традиционные методы часто требуют большого количества вычислений. Поэтому при работе в режиме оффлайн предлагается использовать метод рекурсивного удаления признаков (РУП), который опирается на оценку значимости признаков Случайного леса. Таким образом, удаётся за «разумное» время отобрать набор признаков и моделей, обеспечивающих решение задачи с наименьшей ошибкой.

Для получения итоговых моделей классификации предложено использовать долю правильных ответов, а прогнозирования - среднеквадратичную ошибку (СКО). Для получения несмещённой оценки качества моделей применяется блочная перекрёстная проверка.

Однако не все задачи возможно решать, применяя данную схему. Проблемы вызывают задачи классификации и восстановления регрессии в условиях так называемой смены концепта, когда наблюдаемые процессы имеют нестационарную природу. Если для задач исследования достаточно медленных процессов возможно применить периодическое обновление модели с нуля, то для быстрых процессов с требованиями реального времени это становится крайне проблематичным. Для этих случаев нестационарных данных предлагается использовать описанный далее потоковый метод Proximity Driven Streaming Random Forest (PDSRF). Под потоковостью в данном случае имеется в виду его возможность адаптироваться к изменениям в поступающих данных в темпе процесса.

### **Разработка потокового метода классификации.**

Ввиду того, что поведение реальных ЭЭС может иметь нестационарный характер, необходимо построить потоковую модель классификации и регрессии, которая могла бы поддерживать высокую точность работы в условиях так называемой смены концепта. Для создания такого метода в качестве модели была взята композиция рандомизированных решающих деревьев - СЛ, которая была дополнена адаптивными весами (см. рис. 2).

Композиционный подход позволяет адаптироваться к изменениям в данных несколькими различными способами. В предлагаемом подходе используется

---

<sup>8</sup>Chen X., Ishwaran H. Random forests for genomic data analysis //Genomics. – 2012. – Vol. 99. – No. 6. – p. 323-329.

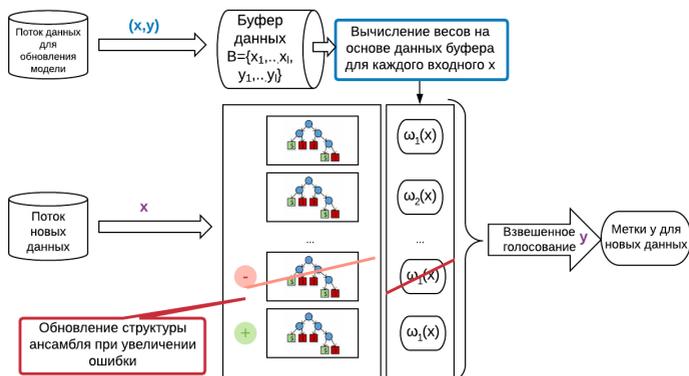


Рис. 2 — Схема работы предлагаемого подхода классификации и восстановления регрессии.

комбинация манипуляции с данными (как и в оригинальном СЛ), адаптации агрегирующей операции и изменения структуры композиции.

В отличие от оригинального СЛ в качестве агрегирующей операции используется взвешенное голосование. Каждому базовому алгоритму ставится в соответствие весовая функция  $\omega_t(x) \in [0, 1]$ , которая тем меньше, чем больше предполагаемая ошибка алгоритма в точке  $x$ . Эта функция аппроксимируется на основе ретроспективных данных буфера  $B$ . При этом мы предполагаем, что характер изменения  $\omega_t(x)$  по времени достаточно медленный, то есть используем базовое предположение о том, что данные сохраняют стационарность наибольшую часть времени.

Так, для конкретного объекта  $x \in X$ , решение в случае классификации вычисляется в виде следующей задачи оптимизации

$$H(x) = \operatorname{argmax}_{y \in Y} \sum_{k=1}^K \omega_k(x) [\hat{h}_k(x) = y], \quad (3)$$

где  $\hat{h}_k(x)$  – решающее правило  $k$ -го решающего дерева ансамбля.

Для вычисления весов алгоритмов ансамбля необходимо хранить определенный набор измерений. Для этих целей мы используем скользящее окно постоянной длины по времени (как предложено в периодически обновляемом Случайном лесе [19]). Длина окна является гиперпараметром метода и оценивается с помощью перекрестной проверки. Так как мы имеем дело с потоковыми методами, то на размер окна также должны быть наложены ограничения по используемой памяти и времени работы, которые определяются условиями конкретной задачи.

В целях сокращения затрат по памяти и времени выполнения, в качестве базовых алгоритмов используются полностью рандомизированные деревья. При

этом каждое из деревьев в отдельности используется в режиме оффлайн, без каких-либо модификаций.

Таким образом в диссертации предложен авторский метод PDSRF, эффективно комбинирующий принципы методов СЛ и Accuracy Weighted Ensemble<sup>9</sup>. В качестве базовых алгоритмов композиции используется полностью рандомизированные деревья (ПРД)<sup>10</sup>, что позволяет ускорить процесс обучения за счет использования только отдельных деревьев в режиме «офф-лайн» и обновления композиции простой заменой деревьев с низкой точностью. При этом замена происходит только в том случае, если средняя точность всего ансамбля меньше заданного порога. Также ограничивается количество замен с целью лимитировать вычислительные затраты и не допустить переобучения ансамбля.

Оценка ошибки конкретного классификатора на новых примерах выполняется из следующих эвристических предположений:

- классификаторы имеют близкие ошибки на «похожих» примерах,
- функция ошибки для конкретного классификатора меняется достаточно редко.

Это позволяет приблизить функцию ошибки с помощью взвешенного метода  $k$  ближайших соседей. В качестве обучающей выборки извлекаются примеры из буфера, а в качестве значений отклика – ошибки классификаторов на примерах буфера. Таким образом, для того, чтобы найти веса для классификации конкретного примера, первоначально происходит поиск похожих примеров в буфере, затем усреднение и полученное значение определяется как  $\omega_i = \frac{1}{E_i^2(x) + \Delta}$ , где  $E_i(x)$  – предполагаемая ошибка  $i$ -го классификатора ансамбля на примере  $x \in X$ , а  $\Delta$  – малый положительный параметр, который задает максимальный возможный вес.

Приближение строится с помощью взвешенного метода  $k$  ближайших соседей. При этом взят буфер скользящего окна  $B^q = (x_1, y_1), \dots, (x_q, y_q)$  с размером  $q$ . Задана функция расстояния  $\rho(x, \hat{x})$ . Для примера  $\hat{x} \in X$  все примеры буфера упорядочиваются в порядке возрастания расстояния  $\rho(\hat{x}, x_{1;x}) \leq \rho(\hat{x}, x_{2;x}) \leq \dots \leq \rho(\hat{x}, x_{q;x})$ , тогда искомая функция может быть оценена как

$$E_i(\hat{x}) = \frac{\sum_{j=1}^k \tilde{E}_i(x_{j;\hat{x}}) \rho(\hat{x}, x_{j;\hat{x}})}{\sum_{j=1}^k \rho(\hat{x}, x_{j;\hat{x}})}, \quad (4)$$

где  $k$  – количество ближайших соседей,  $i$  – номер алгоритма в композиции.

Функция расстояния для метода  $k$  ближайших соседей может быть определена разными способами. Однако, с целью сокращения вычислительной сложности, целесообразно использовать так называемую внутреннюю функцию

<sup>9</sup>Wang H. et al. Mining concept-drifting data streams using ensemble classifiers //Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. – AcM, 2003. – С. 226-235.

<sup>10</sup>Geurts P., Ernst D., Wehenkel L. Extremely randomized trees //Machine learning. – 2006. – Vol. 63. – No. 1. – pp. 3-42.

близости. Эта функция использует структуру деревьев для получения значения близости следующим образом: если два примера попадают в одну и ту же ячейку пространства признаков соответствующую терминальному узлу дерева, то значение увеличивается на единицу. По завершении процесса итоговая функция определяется по количеству деревьев, чем больше число, тем ближе находятся примеры друг к другу.

Таким образом, автором предложены метод и модель для работы с нестационарными данными, которые могут быть использованы для решения задачи классификации в широком множестве задач современной электроэнергетики. В дальнейшем возможно расширение данного метода для решения задач восстановления регрессии.

**Предлагаемый метод оценки режимной надёжности ЭЭС.** Для оценки режимной надёжности построены интеллектуальные модели на основе машинного обучения, которые приближают классические показатели режимной надёжности системы. В качестве одного из таких показателей был выбран L-индекс <sup>11</sup>, отражающий близость системы к коллапсу по напряжению, который рассчитывается для каждого узла в отдельности, а общая оценка для всей системы вычисляется как максимум по всем индексам. Значение L-индекса лежит в интервале [0,1] и может быть рассмотрено как вероятность коллапса. Но в данном случае предлагается использовать дискретный набор состояний, который построен на его основе (Таблица 1).

Таблица 1 — Набор состояний полученный на основе L-индекса.

№	Индекс надёжности	Метка класса
1	$L \leq 0.3$	Нормальный режим
2	$0.3 < L < 0.5$	Предаварийный режим
3	$0.5 < L < 0.7$	Аварийный режим (корректируемый)
4	$L \geq 0.7$	Аварийный режим (не корректируемый)

Таким образом, происходит обучение модели классификации для получения дискретной оценки режимной надёжности, которая может быть легко интерпретирована человеком-диспетчером.

Кроме того, как показал анализ исследований других авторов <sup>12</sup>, существует общий аналитический алгоритм оптимизации по реактивной мощности на основе L-индекса, который состоит в поиске значений инъекций реактивной мощности, позволяющих минимизировать сумму L-индексов по нагрузочным узлам ЭЭС. Применение данного подхода позволяет повысить устойчивость по напряжению и «отвести» систему в утяжелённых режимах от границ неустойчивости.

<sup>11</sup>Kessel P., Glavitsch H. Estimating the voltage stability of a power system //IEEE Transactions on power delivery. – 1986. – Т. 1. – №. 3. – С. 346-354.

<sup>12</sup>Gong X. et al. Research on the method of calculating node injected reactive power based on L indicator //Journal of Power and Energy Engineering. – 2014. – Т. 2. – №. 04. – С. 361.

Тем не менее, несмотря на относительную простоту расчёта, этот метод требует большого объема вычислений, что отражается на скорости работы расчётной модели. Для решения этой проблемы было предложено использовать композиционный метод ПРД для аппроксимации значений реактивной мощности. При этом задача оценки инъекций реактивной мощности решается для каждого узла отдельно. В работе рассматриваются аварийные процессы, протекающие относительно медленно (от одной до пяти минут), поэтому математически они могут быть описаны без уравнений динамики с использованием лишь моделирования потокораспределения как основного метода исследования. Таким образом, ЭЭС моделируется в виде последовательности установившихся режимов. Для обучения предложено генерировать модельные данные для общеизвестных тестовых схем, таких, как IEEE118 и воспроизводить два сценария: постепенное утяжеление режима ЭЭС с отключением случайно выбранных узлов ЭЭС; постепенное утяжеление режима ЭЭС с подключением/отключением мощного потребителя. Для каждого сценария разработаны вычислительные эксперименты, результаты которых изложены в третьей главе. Таким образом, с помощью предлагаемых моделей планируется снизить влияние нестационарности и увеличить скорости работы системы.

**Третья глава** посвящена тестированию предложенных модели, алгоритма и разработке комплекса программ. Результаты предложенных подходов сравниваются с современными методами.

### **Разработка комплекса программ для проведения экспериментального исследования**

Для выполнения поставленных задач автором разработан комплекс программ (см. рис. 3), который позволяет сравнить предложенные методы с существующими. Такой комплекс программ включает

- реализацию предложенного метода потоковой классификации в виде программы на языке C++ (подробнее Forest Factory [22]);
- программу для тестирования предложенного алгоритма потоковой классификации и сравнения с аналогами в классе потоковых алгоритмов;
- программу оценки режимной надёжности ЭЭС на языке R;
- программу моделирования установившегося режима для проведения расчётов режима ЭЭС на основе библиотеки PSAT на языке Matlab.

Программа классификации по потоковым данным имеет расширяемую объектно-ориентированную структуру. Так как к программе могут быть предъявлены дополнительные требования по времени выполнения, поэтому в качестве языка разработки выбран C++ (стандарт C++11). Для осуществления векторных и матричных расчётов используется библиотека Eigen версии 3. Остальные части программного комплекса разработаны на интерпретируемых языках. Для проведения моделирования установившегося режима ЭЭС использована библиотека

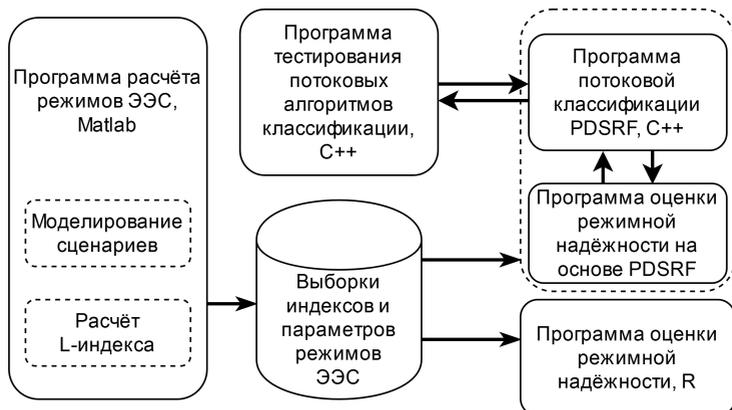


Рис. 3 — Структура разработанного программного комплекса.

PSAT (Power System Analysis Toolbox)<sup>13</sup> на языке Matlab и расчёт с использованием метода Ньютона-Рафсона. Расчёты производились с помощью пакета прикладных программ Matlab. Программа для тестирования метода оценки режимной надёжности написана на языке R.

**Тестирование предложенного метода классификации.** Для оценки эффективности предложенного метода проведено сравнение его точности на известном тестовом наборе данных CoverType. В результате эксперимента получены результаты, приведенные в Таблице 2. Установлено, что предложенный алгоритм превосходит другие использованные в эксперименте новейшие алгоритмы по точности. Также нужно заметить, что предложенное взвешивание позволяет значительно повысить точность работы. И хотя такая мера сопряжена с дополнительными вычислениями, их количество удаётся сократить благодаря использованию внутренней функции близости СЛ.

**Тестирование предложенной потоковой модели для оценки устойчивости по напряжению.** Для тестирования предложенных модели и метода PDSRF для оценки устойчивости по напряжению была использована тестовая схема IEEE 118. При этом использовалось пропорциональное утяжеление с коэффициентом утяжеления  $\alpha=0,005$  по всем узлам. На каждом шаге расчёта установившегося режима случайным образом моделировались 30 аварийных событий – отключения трансформаторов и линий электропередач. Таким образом, созданы две выборки для обучения и тестирования предлагаемого алгоритма. Тестирование показало, что методы машинного обучения высоко эффективны для классификации

<sup>13</sup>Milano F. An open source power system analysis toolbox //IEEE Transactions on Power systems. – 2005. – Т. 20. – №. 3. – С. 1199-1206.

<sup>14</sup>Brzezinski D., Stefanowski J. Reacting to different types of concept drift: The accuracy updated ensemble algorithm //IEEE Transactions on Neural Networks and Learning Systems. – 2014. – Vol. 25. – no. 1. – Pp. 81-94.

Таблица 2 — Сравнение результатов тестирования различных алгоритмов на наборе данных CoverType. Результаты для алгоритмов HOT, AUE2, AUE1, Lev, DWM, Oza, AWE, Win получены с использованием настроек описанных в работе 14.

Метод	Средняя точность, %
PDSRF	87,42
HOT	86,48
AUE2	85,20
PDSRF (без взвешивания)	82,79
AUE1	81,24
Lev	81,04
DWM	80,84
Oza	80,40
AWE	79,34
Win	77,19

режимов ЭЭС в условиях стационарных данных и показывают точность близкую к 100%. Однако реальные системы намного более сложны, поэтому уточним используемую для расчёта режимов ЭЭС модель.

**Тестирование предложенной потоковой модели для оценки режимной надёжности в условиях смены концепта.** Чтобы добавить в сценарий изменение поведения ЭЭС (так называемое изменение концепции) на фоне пропорционального утяжеления моделировались подключения мощных потребителей (см. рис. 4). Из Таблицы 3 видно, что предложенный подход позволяет достичь высокой точности даже в условиях нестационарности. Отметим, что потоковый метод PDSRF превосходит по точности такие методы, как градиентный бустинг над РД (ГБ) и ИНС экстремального обучения (ИНС ЭО).

Таблица 3 — Результаты тестирования различных алгоритмов на схеме IEEE 118 с присутствием в данных смещения концепта.

Метод	Точность, %	Каппа, %
PDSRF	<b>97,24</b>	<b>95,30</b>
МОБ	81,54	64,92
СЛ	96,01	93,24
ГБ	93,64	89,41
ИНС ЭО	80,80	65,08

**Тестирование композиционных моделей для оценки инъекций реактивной мощности.**

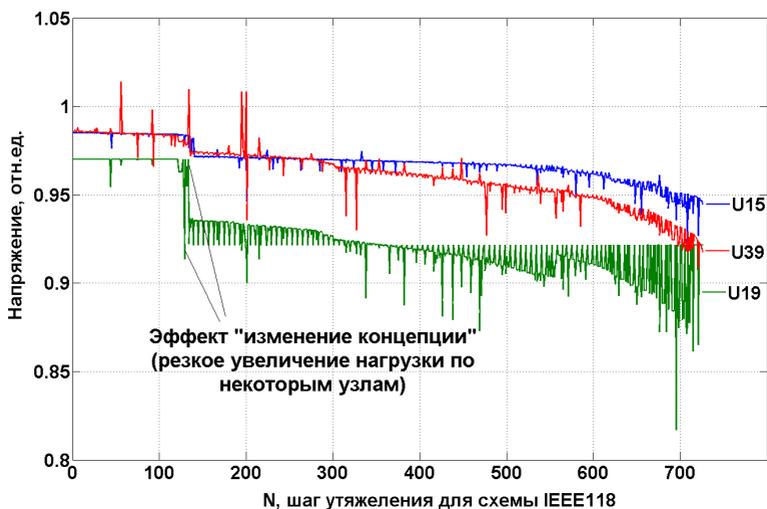


Рис. 4 — Изменение напряжений узлов 15, 19, 39, для которых смоделировано резкое возрастание мощности нагрузки.

В процессе моделирования тестовой схемы IEEE 118 получена база данных для обучения и тестирования моделей оценки инъекций реактивной мощности, которая включала следующие входные данные: параметры режима (напряжение узлов, потоки активной и реактивной мощности) и локальные значения  $L$ -индекса  $L_j$ , а также выходные – значения дополнительных инъекций реактивной мощности  $\Delta Q_i$  для 91 нагрузочного узла схемы IEEE 118. Инъекции реактивной мощности  $\Delta Q_i$  можно получить, минимизируя сумму по всем локальным индексам  $L_j$ , однако решение такой задачи напрямую является вычислительно сложным и занимает существенное время. Поэтому применяется аппроксимация  $\Delta Q_i$  методами машинного обучения, а именно ПРД.

Результаты тестирования моделей представлены в Таблице 4. В качестве сравнительных критериев использовались: СКО - среднеквадратичная ошибка усреднённая по всем узлам, САО - средняя абсолютная ошибка усреднённая по всем узлам, а также  $t_{train}$  и  $t_{test}$  - время обучения и тестирования, соответственно. Как видно из Таблицы 4, метод ПРД даёт более низкую ошибку при определении инъекций  $\Delta Q_i$ . При этом классическая реализация алгоритма случайного леса показывает несколько худшие результаты: СКО = 0,1502; САО = 0,1271. Также указаны время обучения и тестирования. Установлено, что для обучения полной модели на основе ПРД требуется порядка 4 с., а отклик при тестировании составляет 0,00149 секунд.

Таким образом, низкая ошибка при определении дополнительных инъекций  $\Delta Q_i$  и высокая скорость отклика полученной модели позволяют эффективно

Таблица 4 — Эффективность алгоритмов машинного обучения при определении дополнительных объёмов инъекций  $\Delta Q_i$  для нагрузочных узлов тестовой схемы IEEE118 на основе оценки локальных  $L$ -индексов.

Метод	СКО	САО	$t_{train}, c$	$t_{test}, c$
ПРД	0,1299	0,1116	4,812	0,00149
МОВ	0,1498	0,1254	3,441	0,00167
СЛ	0,1502	0,1271	0,811	0,00153
ГБ	0,1463	0,1234	4,671	0,00282
ИНС ЭО	0,1517	0,1282	0,021	0,00153

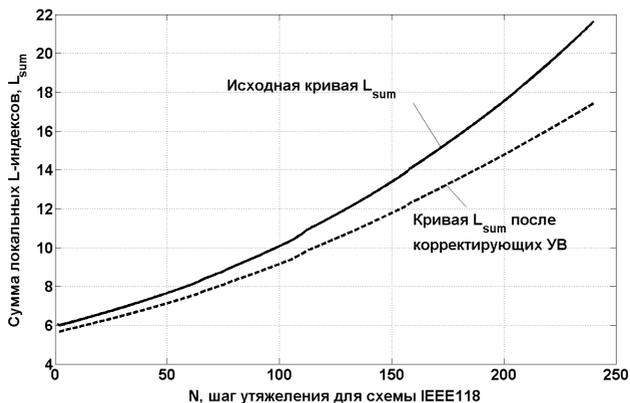


Рис. 5 — Кривые суммы локальных  $L$ -индексов до и после корректирующих УВ для схем IEEE118

использовать её в реальном времени даже для ЭЭС большой размерности. Полученные значения дополнительных инъекций  $\Delta Q_i$  использованы для компенсации дефицита реактивной мощности, в результате чего удалось добиться сокращения суммы  $L$ -индексов по нагрузочным узлам, прежде всего для утяжелённых и опасных режимов схемы IEEE118 (см. рис.5).

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Рассмотрены современные композиционные модели и методы классификации и восстановления регрессии применяемые для анализа сложных процессов в ЭЭС.
2. Разработана непараметрическая модель на основе Случайного леса для классификации в нестационарных данных ЭЭС.
3. Разработан и программно реализован численный метод потоковой классификации на основе предложенной модели. Разработанный метод апробирован на популярном наборе данных. Проведено сравнение с

другими общеизвестными потоковыми методами. Результаты показали высокую точность предложенного метода.

4. Предложен метод оценки режимной надёжности в темпе процесса на основе разработанной модели. Метод апробирован на модельных данных схемы IEEE118. Проведено сравнение эффективности существующих интеллектуальных методов с предложенным. Во всех рассмотренных случаях предложенный метод показывает наибольшую точность.
5. Разработан комплекс программ для проведения экспериментального исследования разработанных математических моделей и численных методов в задаче оценки режимной надёжности. Получено авторское свидетельство о регистрации программной реализации предложенного вычислительного алгоритма.

Работа выполнена в рамках госзадания Министерства науки и высшего образования РФ, проект No. FZZS-2020-0039.

### **Список публикаций по теме диссертации**

#### **Журналы из списка ВАК**

1. *Жуков, А. В.* Модификация алгоритма случайного леса для классификации нестационарных потоковых данных [Текст] / А. В. Жуков, Д. Н. Сидоров // Вестник Южно-Уральского государственного университета. — 2016. — Т. 9, № 4. — С. 86—95.
2. Power System Parameters Forecasting Using Hilbert-Huang Transform and Machine Learning [Текст] / А. V. Zhukov [и др.] // «Известия Иркутского государственного университета». Серия «Математика». — 2014. — Т. 9. — С. 75—90.

#### **Публикации индексируемые Scopus и Web of Science**

3. Ensemble methods of classification for power systems security assessment [Текст] / А. Zhukov [et al.] // Applied computing and informatics. — 2019. — Vol. 15, no. 1. — P. 45—53.
4. Random Forest Based Model for Preventing Large-Scale Emergencies in Power Systems [Текст] / А. Zhukov [et al.] // International Journal of Artificial Intelligence™. — 2015. — Vol. 13, no. 1. — P. 211—228.
5. Toward zero-emission hybrid AC/DC power systems with renewable energy sources and storages: A case study from lake Baikal region [Текст] / А. Zhukov [et al.] // Energies. — 2020. — Vol. 13, no. 5. — P. 1226.
6. A suite of intelligent tools for early detection and prevention of blackouts in power interconnections [Текст] / N. I. Voropai [и др.] // Automation and Remote Control. — 2018. — Т. 79, № 10. — С. 1741—1755.
7. Optimization of isolated power systems with renewables and storage batteries based on nonlinear Volterra models for the specially protected natural area of lake Baikal [Текст] / А. Zhukov [et al.] // Journal of Physics: Conference Series. Vol. 1847. — IOP Publishing. 2021. — P. 12—37.

8. Development of software for modelling decentralized intelligent systems for security monitoring and control in power systems [Текст] / A. Zhukov [et al.] // PowerTech, 2015 IEEE Eindhoven. — IEEE. 2015. — P. 1—6.
9. A hybrid artificial neural network for voltage security evaluation in a power system [Текст] / A. Zhukov [et al.] // Energy (IYCE), 2015 5th International Youth Conference on. — IEEE. 2015. — P. 1—8.
10. A random forest-based approach for voltage security monitoring in a power system [Текст] / A. Zhukov [et al.] // PowerTech, 2015 IEEE Eindhoven. — 2015. — P. 1—6.
11. A hybrid wind speed forecasting strategy based on Hilbert-Huang transform and machine learning algorithms [Текст] / A. Zhukov [et al.] // Power System Technology (POWERCON), 2014 International Conference on. — IEEE. 2014. — P. 2980—2986.
12. Zhukov, A. Random forest based approach for concept drift handling [Текст] / A. Zhukov, D. Sidorov, A. Foley // International Conference on Analysis of Images, Social Networks and Texts. — Springer. 2016. — P. 69—77.
13. Development of computational intelligence-based algorithms of preventing voltage collapse in power systems with a complex multi-loop structure [Текст] / A. Zhukov [et al.] // 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC). — 10/2016. — P. 1—5.
14. On-line power systems security assessment using data stream random forest algorithm modification [Текст] / A. Zhukov [et al.] // Innovative Computing, Optimization and Its Applications. — Springer, 2018. — P. 183—200.
15. Development of automatic intelligent system for on-line voltage security control of power systems [Текст] / A. Zhukov [et al.] // 2017 IEEE Manchester PowerTech. — IEEE. 2017. — P. 1—6.
16. Machine learning techniques for power system security assessment [Текст] / A. Zhukov [et al.] // IFAC-PapersOnLine. — 2016. — Vol. 49, no. 27. — P. 445—450.

## Монографии

17. Комплекс интеллектуальных средств для предотвращения крупных аварий в электроэнергетических системах [Текст] / А. Жуков [и др.]. — Новосибирск : Наука, 2016. — С. 332.
18. Жуков А.В., Томин Н. В., Курбацкий В. Г., Сидоров Д. Н. Современные методы и модели прогнозирования скорости ветра и выработки мощности ветроэнергетических установок [Текст] // Региональные аспекты ветроэнергетики / Под ред. Стенникова, В. А. and Курбацкого, В. Г. – Новосибирск : СО РАН, 2020. – С. 113–138.

## **Прочие публикации**

19. Random Forest Based Model for emergency state monitoring in power systems [Текст] / A. Zhukov [et al.] // Mathematical Method for Pattern Recognition: Book of abstract of the 17th All-Russian Conference with Interneational Participation. — Svetlogorsk : TORUS PRESS, 2015. — P. 274.
20. Zhukov, A. Power Systems Parameters Forecasting Using the Hilbert-Huang Transform and Machine Learning [Текст] / A. Zhukov // Тезисы докладов XVI Байкальской международной школы-семинара "Методы оптимизации и их приложения". — ИСЭМ СО РАН. Иркутск, 2014. — С. 170.
21. Random forest-based models for preventing large emergencies in power systems [Текст] / A. Zhukov [и др.] // Сборник трудов VII Международного симпозиума "Обобщенные постановки и решения задач управления"(GSSCP-2014). — ИПУ им. В.А.Трапезникова РАН. г. Геленджик, Россия, 2014. — С. 99—103.

## **Свидетельства о регистрации программ для ЭВМ**

22. Жуков, А. В. Программа для классификации потоковых данных «Forest Factory»: свидетельство № 2016661345. / А. В. Жуков. — 2016 – заявл. 11.08.2016, зарегистр. 06.10.2016. — Правообладатель ФГБОУ ВО "Иркутский государственный университет". Реестр программ для ЭВМ.