

Повышение распознаваемости аварийных режимов релейной защитой методами деревьев решений

КУЛИКОВ А.Л.¹, ЛОСКУТОВ А.А.¹, БЕЗДУШНЫЙ Д.И.²

¹НГТУ им. Р.Е. Алексеева, Нижний Новгород, Россия

²ПАО Сбербанк России, Нижний Новгород, Россия

В условиях внедрения в электрические сети современных активных элементов и устройств (гибкие электропередачи, управляемые шунтирующие реакторы, накопители энергии, источники распределенной генерации и др.) традиционные методы распознавания аварийных режимов для корректного функционирования устройств релейной защиты и автоматики (РЗА) иногда становятся неэффективными. С развитием и массовым внедрением стандарта МЭК 61850, а также программных комплексов имитационного моделирования появляется доступ к большому объему информации о режимах электрической сети, что позволяет разрабатывать принципиально новые алгоритмы для РЗА, в том числе на основе методов машинного обучения. В статье исследуется новый подход к распознаванию повреждения на линиях электропередачи с ответвлениями путем одновременного анализа нескольких информационных признаков и применения групповых алгоритмов машинного обучения: дерево решений, случайный лес и градиентный бустинг. Для получения обучающих выборок применяются имитационное моделирование и метод Монте-Карло. Результаты тестирования исследуемых методов показали требуемую гибкость, способность использования большого количества информационных параметров (многопараметричность), а также лучшие результаты распознавания повреждений по сравнению с дистанционным органом релейной защиты. Реализация алгоритма решающего дерева для микропроцессорных устройств РЗА не потребует модернизации аппаратной части, а связана с доработкой специального программного обеспечения.

К л ю ч е в ы е с л о в а: распознавание коротких замыканий, машинное обучение, решающее дерево, случайный лес, градиентный бустинг, релейная защита и автоматика, имитационное моделирование, алгоритм

В соответствии с новой технической политикой в электросетевом комплексе [1] необходимо внедрять и совершенствовать мероприятия по повышению эффективности функционирования энергосистем. Это касается работы как существующих энергосистем, так и перспективных, в которых могут устанавливаться современные активные элементы (гибкие электропередачи, управляемые шунтирующие реакторы, накопители энергии и др.). Отдельное внимание уделяется математическому (имитационному) моделированию нормальных и аварийных режимов работы при проектировании, строительстве и реконструкции электрических сетей, а также при верификации инженерных расчетов и выборе параметров срабатывания устройств релейной защиты и автоматики (РЗА) [2–5]. Кроме того, большой интерес проявляется к технологиям анализа больших данных (*Big Data*) и искусственного интеллекта для повышения технического совершенства устройств учета электрической энергии [6], автоматики и защиты электрических сетей [7].

Надежная защита элементов энергосистемы от аварийных режимов (короткие замыкания (КЗ), асинхрон-

ных ход, небаланс активной мощности и др.) является важнейшим условием ее бесперебойного функционирования. Актуальным направлением для исследования являются разработка и анализ методов обнаружения и распознавания аварийных режимов. Это становится возможным с массовым внедрением в электрические сети интеллектуальных электронных устройств (ИЭУ) релейной защиты и автоматики с поддержкой стандарта МЭК 61850, которые имеют прямой доступ к информационным параметрам (токи, напряжения, сопротивления и пр.) режима электрической сети. Сложность будущих интеллектуальных сетей потребует применения современных методов цифровой обработки сигналов токов и напряжений, систем сбора информации о текущих режимах, идентификации информационных параметров, а также прогнозных моделей поведения энергосистем в различных режимах. Задача распознавания повреждений в перспективных электрических системах, в том числе с источниками распределенной генерации, будет еще более сложной, поскольку имеет место быстротечность переходных процессов, а также непостоянство параметров электрических сигналов на различных генерирующих установках и нагрузках.

С развитием цифровизации в электроэнергетике распознавание аварийных режимов, в частности связанных с КЗ, может быть реализовано с применением алгоритмов машинного обучения [8–13]. Такой подход может оказаться наиболее востребованным в ситуациях, когда для обеспечения селективной работы устройства РЗА алгоритм защиты должен быть в состоянии различить КЗ на различных участках электрической сети, которые характеризуются близкими значениями параметров режимов.

Цель работы – исследование подхода к распознаванию повреждения на ЛЭП с ответвлениями путем одновременного анализа нескольких информационных признаков и применения методов машинного обучения: дерево решений, случайный лес и градиентный бустинг. В основе получения обучающих выборок лежат имитационное моделирование и метод Монте-Карло [14].

Имитационное моделирование участка электрической сети. Несмотря на внедрение цифровых технологий и мероприятий по совершенствованию алгоритмов защиты, в электрических сетях сохраняются режимы, в которых типовое применение комплектов РЗА не удовлетворяет требованиям чувствительности и селективности. В качестве упрощенного примера рассмотрим фрагмент электрической сети на рис. 1.

Пусть защита линии ω_1 выполнена с использованием комплекта дифференциальной защиты линии (ДЗЛ). Для того чтобы обеспечить блокировку ДЗЛ при КЗ за трансформатором T_1 , необходим дополнительный разрешающий пусковой орган, в качестве которого, как правило, применяют дистанционное реле [15]. Его уставочная характеристика выбирается таким образом, чтобы реле срабатывало при КЗ на линии и не срабатывало при КЗ за трансформатором. Чаще всего выбор характеристики, отвечающей указанным требованиям, возможен, однако короткие замыкания на линии и за трансформатором становятся трудно раз-

личимыми, когда трансформатор обладает достаточно большой номинальной мощностью (малым сквозным сопротивлением), а ответвление расположено рядом с подстанцией, на которой установлена защита.

Для иллюстрации, в соответствии со схемой рис. 1 были реализованы множественные имитационные эксперименты КЗ на линиях ω_1 и ω_2 в программном комплексе *MATLAB*. Имитировались двухфазные и трехфазные КЗ на различной удаленности от места установки защиты. При моделировании проводилось разделение режимов на две категории: α -режимы и β -режимы [16]. Под α -режимами понимаются КЗ и другие аварийные режимы, при которых РЗА должна срабатывать, а β -режимами считаются альтернативные режимы, от которых защита должна быть отстроена.

На рис. 2 приведены замеры комплексного сопротивления дистанционного реле при КЗ в различных точках ЛЭП ω_1 и за трансформатором T_1 (рис. 1), причем отмечается значительная степень перекрытия параметров режимов (рис. 2,а). На рис. 2,б показаны гистограммы действующих значений токов через защиту при тех же условиях и видах КЗ.

Очевидно, что выбор характеристики срабатывания дистанционного реле, способной обеспечить надежное срабатывание защиты при КЗ на линии, и отсутствие срабатывания при КЗ за трансформатором не представляется возможным. Любое решение будет связано с компромиссом между чувствительностью и селективностью. Рассмотрим две граничные характеристики срабатывания дистанционного реле (рис. 3).

Область срабатывания на рис. 3,а является наиболее чувствительной, так как в нее попадают 100 % КЗ на линии ω_1 . Однако внутри такой характеристики также оказывается 26,15 % КЗ на линии ω_2 , расположенной за трансформатором. Характеристика на рис. 3,б является селективной, не допускает ложного срабатывания дистанционного реле, но обеспечивает лишь 35,25 % КЗ на линии ω_1 .

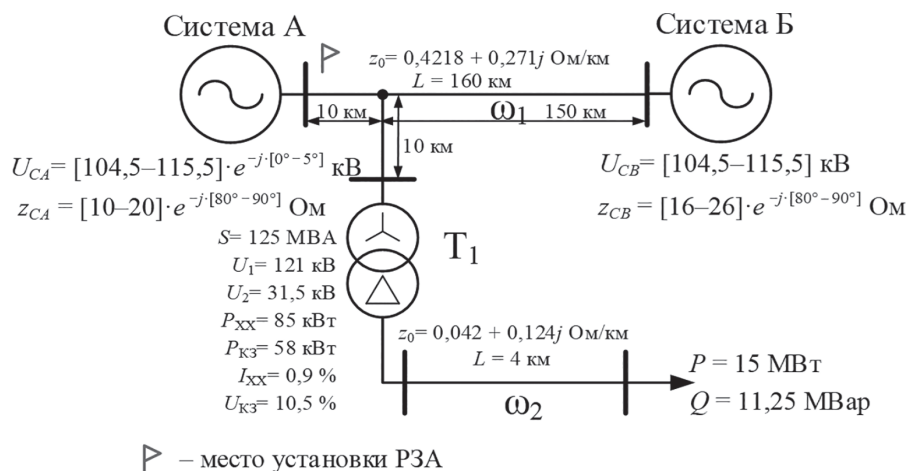


Рис. 1. Расчетная схема фрагмента электрической сети

Fig. 1. Scheme of the electrical network fragment

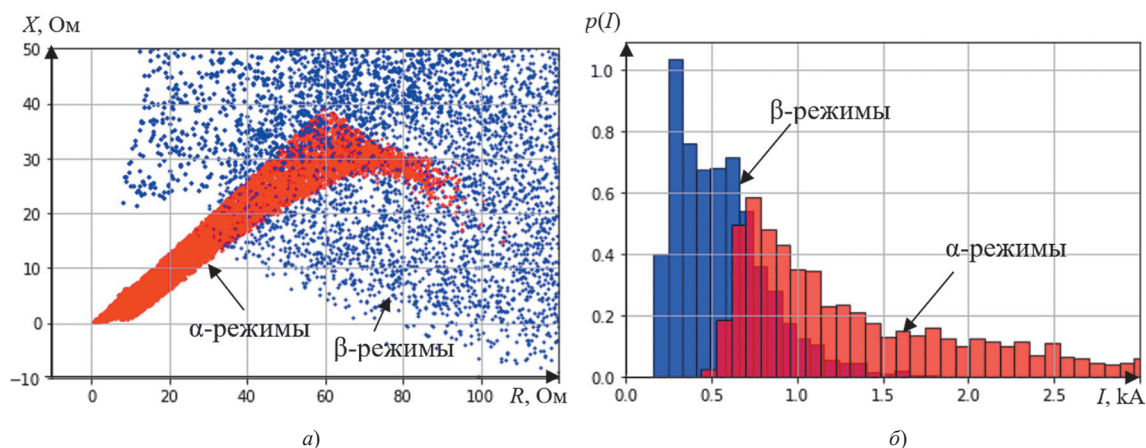


Рис. 2. Результаты имитационного моделирования: *a* – замеры комплексного сопротивления; *б* – гистограммы действующего значения токов
 Fig. 2. Results of simulation: *a* – complex resistance measurements; *б* – histograms of the effective value of currents

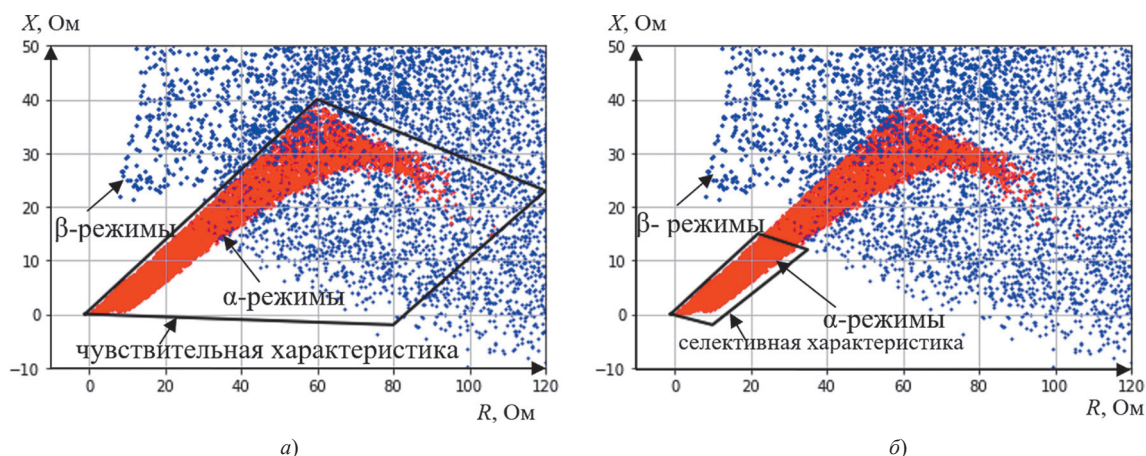


Рис. 3. Граничные характеристики срабатывания дистанционного реле: *a* – чувствительная характеристика; *б* – селективная характеристика
 Fig. 3. Boundary characteristics of remote relay operation: *a* – sensitive characteristic; *б* – selective characteristic

Статистические показатели эффективности алгоритма релейной защиты. Для оценки надежности функционирования РЗА в условиях эксплуатации используется усредненный статистический показатель её работы h , называемый процентом правильных действий [17]:

$$h = n_{п.с} \cdot 100 / (n_{п.с} + n_{о.с} + n_{и.с} + n_{л.с}) \quad (1)$$

или

$$h = [n_{п.с} - (n_{и.с} + n_{л.с})] \cdot 100 / (n_{п.с} + n_{о.с}), \quad (2)$$

где $n_{п.с}$, $n_{и.с}$, $n_{л.с}$ – число правильных, излишних и ложных срабатываний соответственно; $n_{о.с}$ – число отказов срабатывания.

Такой показатель является статистическим, отражает функционирование устройств РЗА в условиях эксплуатации и не может быть применен для оценки распознающей способности алгоритмов защиты по данным имитационного моделирования (например рис. 2).

Однако использование статистического подхода, а также методов, широко распространенных в задачах

машинного обучения [18, 19] и радиотехнических приложениях [20], позволяет ввести требуемые критерии оценки эффективности РЗА. При этом принимается, что значения переменных $n_{п.с}$, $n_{и.с}$, $n_{о.с}$ получены по данным имитационного моделирования.

Введем вероятностные показатели распознавания алгоритма РЗА при условии наличия КЗ (α -режим). Используя переменные выражений (1) и (2), к таким показателям можно отнести условные вероятности:

правильного распознавания КЗ алгоритмом РЗА

$$D = n_{п.с} / (n_{п.с} + n_{о.с}); \quad (3)$$

нераспознавания КЗ алгоритмом РЗА

$$D^* = n_{о.с} / (n_{п.с} + n_{о.с}). \quad (4)$$

Поскольку одному и тому же условию (наличие КЗ, α -режим) соответствуют взаимоисключающие решения, то

$$D + D^* = 1. \quad (5)$$

Качественными показателями распознавания алгоритмом РЗА нормального режима (отсутствие КЗ, β -режим) являются условные вероятности:

ложного решения алгоритмом на срабатывание РЗА

$$F = n_{л.с} / (n_{л.с} + n_{н.с}); \quad (6)$$

правильного решения алгоритмом на несрабатывание РЗА

$$F^* = n_{н.с} / (n_{л.с} + n_{н.с}), \quad (7)$$

где $n_{н.с}$ – число правильных несрабатываний в общем числе имитационных экспериментов, причем

$$F + F^* = 1. \quad (8)$$

Следует отметить, что в расчетах по выражениям (3)–(8) не используется показатель $n_{н.с}$, поскольку он не относится непосредственно к процессу распознавания исследуемой схемно-режимной ситуации, а соотносится с иными режимами функционирования электрической сети.

При такой постановке задачи в теории машинного обучения и радиотехнических приложениях для оценки распознающей способности обучаемых алгоритмов [18–21] используют «рабочую характеристику приемника» (ROC-кривая, от англ. *receiver operating characteristic*) [18]. Рабочую характеристику приемника применительно к алгоритму РЗА можно построить, отложив на плоскости по оси ординат значения D , а по оси абсцисс – переменную F для различных порогов (уставок) срабатывания защиты (рис. 4). Поскольку величина $n_{н.с}$ в серии имитационных экспериментов характеризует чувствительность защиты, рабочая характеристика приемника наглядно показывает баланс между чувствительностью устройства РЗА и его несрабатыванием в допустимых (нормальных) режимах при различных значениях уставки.

Используя терминологию машинного обучения, можно сказать, что зависимость (рис. 4) представляет собой характеристику классификатора, в роли которого выступает дистанционное реле [18, 22, 23], и дает наглядное представление о качестве работы алгоритма бинарной (релейной) классификации.

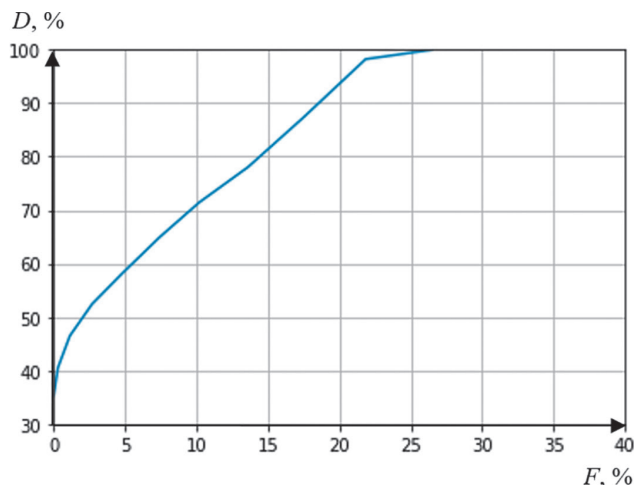


Рис. 4. Соотношение доли правильных и ложных срабатываний при различных характеристиках дистанционного органа

Fig. 4. The ratio of the proportion of correct and false positives for various characteristics of the remote control

Применение алгоритма решающего дерева для распознавания режимов электрической сети. Используем в задаче распознавания аварийных режимов имитационное моделирование и методы машинного обучения, причем в качестве алгоритма распознавания КЗ выберем «решающее дерево». Вычислительная процедура представляет собой определенный набор (совокупность) простейших решающих правил, объединенных в иерархическую структуру (дерево). Пример такой структуры показан на рис. 5.

Каждое решающее правило представляет собой сравнение некоторого информационного признака с заданным порогом срабатывания (уставкой). Листовые узлы дерева ассоциированы с одним из возможных классов. Задача распознавания режима электрической сети по его параметрам с помощью обученного решающего дерева требует движения от корня дерева к одному из его листовых узлов по траектории, зависящей от выполнения или невыполнения условий в узлах.

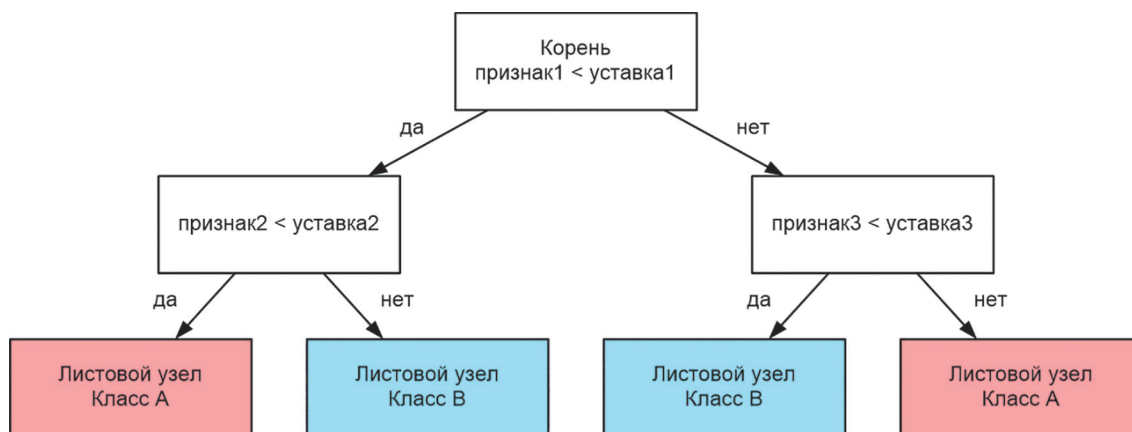


Рис. 5. Пример структуры решающего дерева

Fig. 5. An example of the decision tree structure

Обучение алгоритма решающего дерева предполагает выполнение следующей рекурсивной процедуры:

1. Нахождение информационного признака и его порога срабатывания, которые обеспечат разделение обучающей выборки таким образом, что суммарная неоднородность классов в дочерних узлах максимально уменьшится по сравнению с неоднородностью в исходном узле. В качестве меры неоднородности чаще всего выбираются энтропия или индекс Джини [24, 25].

Энтропия. Расчет энтропии до и после разбиения выполняется по выражениям (9) и (10), разность энтропий равна количеству информации, получаемой от данного узла дерева [21]:

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i); \quad (9)$$

$$H(X|Y) = -\sum_{i=1}^m p(y_i) \sum_{j=1}^n p(x_j|y_i) \log_2 p(x_j|y_i), \quad (10)$$

где $H(X)$ – исходная энтропия узла; $H(X|Y)$ – суммарная энтропия после разбиения по условию Y ; $p(x_i)$ – доля объектов обучающей выборки, имеющих класс i ; $p(y_0)$ – доля объектов обучающей выборки, не удовлетворяющих условию в узле; $p(y_1)$ – доля объектов обучающей выборки, удовлетворяющих условию в узле; $p(x_i|y_0)$ – доля объектов, имеющих класс i , среди объектов, не удовлетворяющих условию в узле; $p(x_i|y_1)$ – доля объектов, имеющих класс i , среди объектов, удовлетворяющих условию в узле.

Индекс Джини:

$$Gini(X) = 1 - \sum_{i=1}^n p(x_i)^2; \quad (11)$$

$$Gini(X|Y) = \sum_{i=0}^1 p(y_i) \left(1 - \sum_{j=1}^n p(x_j|y_i)^2 \right), \quad (12)$$

где $Gini(X)$ – исходная неоднородность узла; $Gini(X|Y)$ – суммарная неоднородность после разбиения по условию Y .

2. После разделения обучающей выборки на подмножества по выбранному критерию каждое такое подмножество вновь подвергается разделению в соответствии с п.1. Процедура выполняется рекурсивно, пока в узлах не окажутся представители лишь одного класса либо не будет достигнута заданная максимальная глубина дерева. Максимальная глубина дерева решений – это максимально возможная длина от корня до листа.

К особенностям алгоритма дерева решений следует отнести интерпретируемость, возможность использовать большое количество информационных признаков, отсутствие необходимости масштабировать данные перед обучением алгоритма и невозможность изменять порог принятия решения обученной модели с целью ее очувствления/загрубления.

Формирование признакового пространства. В результате выполнения каждой итерации имитационной

модели формируются комплексные значения тока и напряжения трех фаз в месте установки разрабатываемой защиты (рис. 1). На основе полученных данных рассчитываются параметры режима, которые используются в качестве информационных признаков в процессе обучения алгоритма решающего дерева. Перечень параметров режима, а также расчетные их выражения сведены в табл. 1.

Пример обучения алгоритма решающего дерева. Проведем обучение и проанализируем результаты функционирования алгоритма решающего дерева. Зададим требования к структуре алгоритма: глубина решающего дерева равна 3, минимальное число выборочных значений информационного признака, попадающих в листовую узел, равно 1. В качестве меры неоднородности воспользуемся индексом Джини.

Разделим множество смоделированных аварийных и нормальных режимов на обучающую и тестовую выборки в пропорции 67 % к 33 %. Обучающая выборка будет использоваться для обучения алгоритма решающего дерева, а тестовая – для оценки качества его функционирования.

В результате множественных имитационных экспериментов, направленных на обучение алгоритма защиты, получаем решающее дерево, обладающее структурой (рис. 6).

В корневом узле (рис. 6) решающее дерево реализует разделение нормальных и аварийных режимов по реактивному сопротивлению X_{Ca} , а вероятностные распределения этого информационного признака приведены на рис. 7.

Хотя вероятностные распределения информационного признака для α - и β -режимов накладываются друг на друга (рис. 7), уставочное значение (порог) позволяет отделить область, характерную для КЗ на линии ω_1 , от области, характерной для КЗ на ω_2 . В результате такого разделения в узел 1 решающего дерева попадают только режимы, удовлетворяющие условию узла 0. Как можно видеть из анализа рис. 6 и 7, в узел 1 попадают режимы в пропорции 82 % (α -режим) и 18 % (β -режим).

В результате обучения алгоритма решающего дерева признаком, по которому обеспечивается разделение режимов в узле 1, также оказалось реактивное сопротивление, но на этот раз X_{Bc} . На рис. 8 представлены вероятностные распределения информационного признака X_{Bc} в узле 1 и уставочное (пороговое) значение признака.

Режимы, расположенные слева от порогового значения на рис. 8, попадают в узел 3, в то время как режимы, находящиеся справа, – в узел 4. Из рис. 8 видно, что в узел 4 попадают только β -режимы, следовательно, дальнейший их анализ не имеет смысла.

На вход узла 3 выборочные значения, характерные для α - и β -режимов, поступают в соотношении 96,6 % к 3,4 %. Здесь по результатам обучения алгоритма ре-

Информационные признаки для распознавания режима
Information features used for mode recognition

Признак	Формула
Активное сопротивление. Фазы AB	$R_{AB} = \text{Re}[(U_A - U_B)/(I_A - I_B)]$
Активное сопротивление. Фазы BC	$R_{BC} = \text{Re}[(U_B - U_C)/(I_B - I_C)]$
Активное сопротивление. Фазы CA	$R_{CA} = \text{Re}[(U_C - U_A)/(I_C - I_A)]$
Реактивное сопротивление. Фазы AB	$X_{AB} = \text{Im}[(U_A - U_B)/(I_A - I_B)]$
Реактивное сопротивление. Фазы BC	$X_{BC} = \text{Im}[(U_B - U_C)/(I_B - I_C)]$
Реактивное сопротивление. Фазы CA	$X_{CA} = \text{Im}[(U_C - U_A)/(I_C - I_A)]$
Активная мощность фазы A	$P_A = \text{Re}(U_A I_A^*)$
Активная мощность фазы B	$P_B = \text{Re}(U_B I_B^*)$
Активная мощность фазы C	$P_C = \text{Re}(U_C I_C^*)$
Реактивная мощность фазы A	$Q_A = \text{Im}(U_A I_A^*)$
Реактивная мощность фазы B	$Q_B = \text{Im}(U_B I_B^*)$
Реактивная мощность фазы C	$Q_C = \text{Im}(U_C I_C^*)$
Действующее значение тока фазы A	$I_{дA} = (1/\sqrt{2}) I_A $
Действующее значение тока фазы B	$I_{дB} = (1/\sqrt{2}) I_B $
Действующее значение тока фазы C	$I_{дC} = (1/\sqrt{2}) I_C $
Действующее значение напряжения фазы A	$U_{дA} = (1/\sqrt{2}) U_A $
Действующее значение напряжения фазы B	$U_{дB} = (1/\sqrt{2}) U_B $
Действующее значение напряжения фазы C	$U_{дC} = (1/\sqrt{2}) U_C $
Действующее значение тока прямой последовательности	$I_{д1} = (1/3) I_A + I_B e^{j120^\circ} + I_C e^{j240^\circ} $
Действующее значение тока обратной последовательности	$I_{д2} = (1/3) I_A + I_B e^{j240^\circ} + I_C e^{j120^\circ} $
Действующее значение напряжения прямой последовательности	$U_{д1} = (1/3) U_A + U_B e^{j120^\circ} + U_C e^{j240^\circ} $
Действующее значение напряжения обратной последовательности	$U_{д2} = (1/3) U_A + U_B e^{j240^\circ} + U_C e^{j120^\circ} $

шающего дерева наиболее информативным признаком для разделения режимов вновь оказалось реактивное сопротивление X_{CA} , вероятностные распределения которого для α - и β -режимов показаны на рис. 9.

Уставочное (пороговое) значение для узла 3 разделило режимы таким образом, что слева от него оказалось практически 100 % α -режимов, а справа – режимы оказались в пропорции 88 % к 12 %. Из-за ограничения по максимальной глубине дальнейшее построение решающего дерева не проводится.

Аналогичным образом реализуется анализ функционирования алгоритма решающего дерева для режимов, расположенных справа от порогового значения корневого узла 0. Характерные вероятностные распределения информационных признаков узлов, а также уставочные значения приведены на рис. 10.

Анализ эффективности функционирования алгоритма решающего дерева. Для оценки эффективности функционирования обученного алгоритма ре-

шающего дерева привлекалась тестовая выборка, реализации которой не участвовали в процессе обучения алгоритма. Отнесение тестируемого режима к одному из двух классов при наличии обученного решающего дерева выполнялось от корня дерева с продвижением вдоль ветвей и сравнением требуемых параметров классифицируемого режима (информационных признаков) с пороговыми значениями узлов. Оказавшись в листовом узле, принималось решение в пользу того класса, вероятность которого является наибольшей по результатам обучения.

При проведении имитационных экспериментов с тестовой выборкой была получена матрица ошибок (табл. 2), характеризующая процесс распознавания режимов алгоритмом РЗА, включающим решающее дерево (рис. 6). В табл. 2 КЗ на линии ω_2 (рис. 1) обозначены как класс 0, тогда как КЗ на линии на ω_1 – как класс 1.

Из табл. 2 видно, что алгоритм решающего дерева в устройстве РЗА (рис. 1) позволяет обнаружить 98,7 %

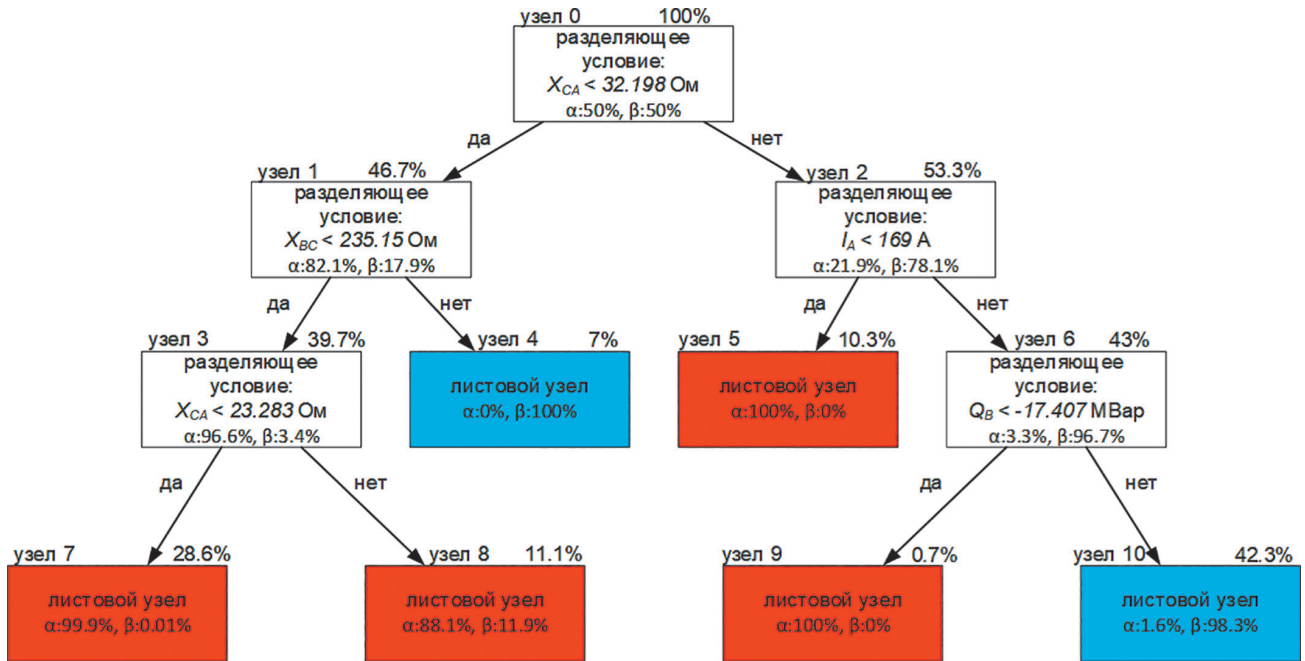


Рис. 6. Пример структуры решающего дерева

Fig. 6. An example of the decision tree structure

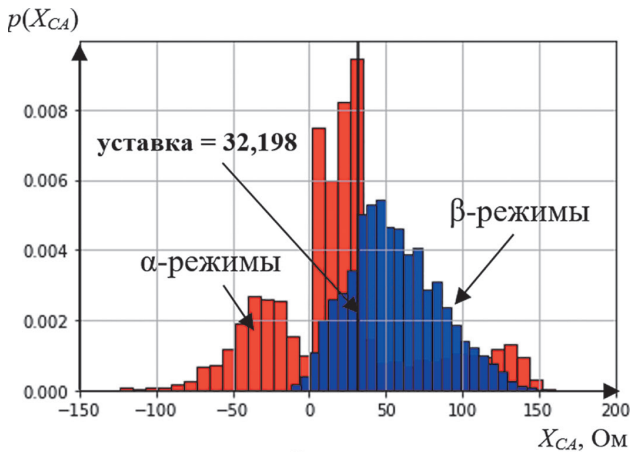


Рис. 7. Вероятностные распределения значений признака X_{CA} в корневом узле решающего дерева

Fig. 7. Probability distributions of feature X_{CA} at the root node of the decision tree

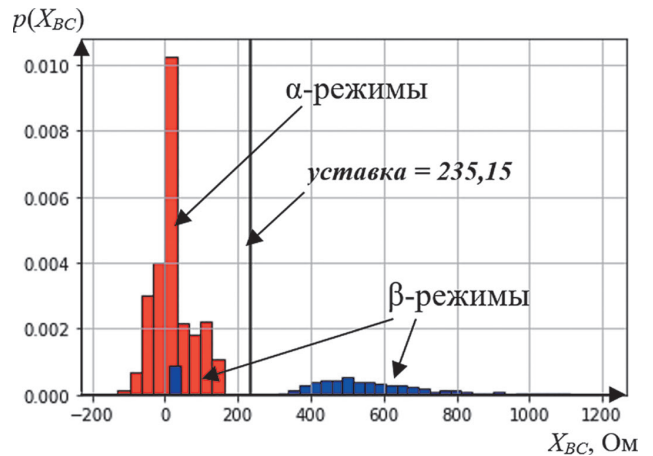


Рис. 8. Вероятностные распределения значений признака X_{BC} в узле 1 решающего дерева

Fig. 8. Probability distributions of feature X_{BC} in node 1 of the decision tree

КЗ на ЛЭП ω_1 , при этом вероятность ложного срабатывания устройства РЗА при КЗ за трансформатором составляет около 2%.

Соотнесем получившийся результат (табл. 2) с результатом работы дистанционного органа, разместив D и F решающего дерева и ранее полученную рабочую характеристику (рис. 4) дистанционного органа на одном графике (точка 1 на рис. 11).

Анализ рис. 11 наглядно показывает преимущество алгоритма решающего дерева по сравнению с дистанционным органом при распознавании устройством РЗА (рис. 1) аварийных и нормальных режимов. Сле-

дует отметить, что недостатком алгоритма решающего дерева является отсутствие возможности оптимизировать уставочные значения (рис. 6), изменяя тем самым соотношение чувствительность/селективность.

Структурная оптимизация параметров алгоритма решающего дерева. Выше был рассмотрен алгоритм решающего дерева с фиксированными, заранее заданными параметрами: глубиной дерева (max_depth) и минимальным значением количества объектов в листовом узле ($min_samples_leaf$). Однако в практике построения систем распознавания на основе машинного обучения подбор параметров, как правило, выпол-

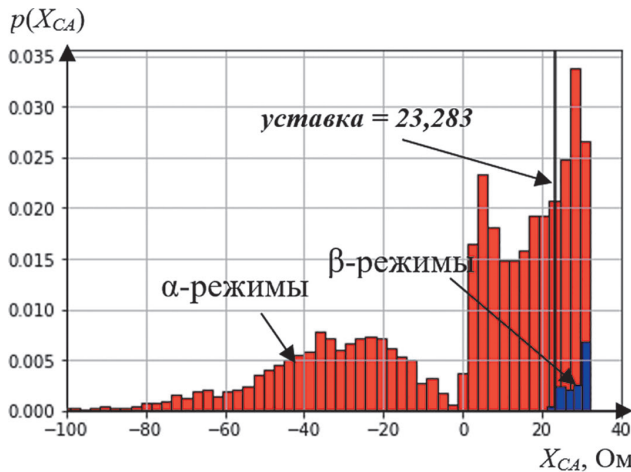


Рис. 9. Вероятностные распределения значений признака X_{CA} в узле 3 решающего дерева

Fig. 9. Probability distributions of feature X_{CA} in node 3 of the decision tree

няется итерационного, чтобы обеспечить наилучшее качество работы алгоритма классификации. Наиболее простым способом подобрать оптимальные параметры алгоритма классификации является метод поиска по сетке (Grid Search) [24, 25]. Метод состоит в последовательном переборе всех комбинаций параметров из заданных диапазонов. Для каждой такой комбинации оценивается качество работы алгоритма классификации по заранее выбранной целевой метрике, после чего выбирается наилучшая комбинация. При этом целесо-

образно применять технику кросс-валидации, которая заключается в многократном воспроизведении следующей последовательности шагов:

- случайным образом часть данных, используемых для обучения, разбивается на обучающую и валидационную выборку;
- реализуется обучение алгоритма классификации на обучающей выборке;
- оценивается целевая метрика с использованием валидационной выборки.

Итоговый результат определяется как усредненное значение целевой метрики по всем итерациям.

В ходе имитационного моделирования изменялись параметры алгоритма классификации (табл. 3), которые использовались для наилучшего разделения режимов электрической сети (рис. 1).

В качестве целевой метрики использовалась величина M , характеризующая соотношение числа правильных распознаваний и ошибок классификации режимов [25]. С учетом использования переменных выражений (1)–(7) целевая метрика принимает вид:

$$M = 2 \left[\frac{n_{п.п.}}{n_{п.п.} + n_{л.п.}} \right] \left[\frac{n_{п.п.}}{n_{п.п.} + n_{о.п.}} \right] \left[\frac{n_{п.п.}}{n_{п.п.} + n_{л.п.}} + \frac{n_{л.п.}}{n_{п.п.} + n_{о.п.}} \right]. \quad (13)$$

В результате применения метода поиска по сетке были найдены параметры алгоритма решающего дерева, обеспечивающие наилучшее значение целевой метрики: максимальная глубина дерева 9; минимальное количество объектов в листовом узле 3. Такому ал-

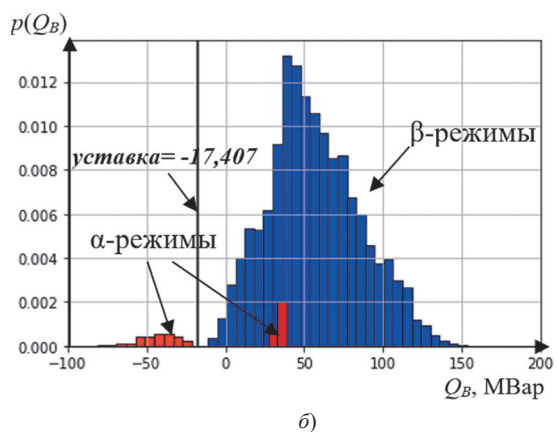
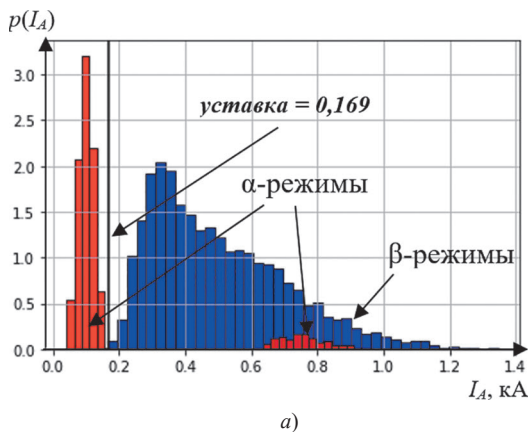


Рис. 10. Вероятностные распределения значений признаков в узлах дерева: а – узел 2; б – узел 6

Fig. 10. Probability distributions of feature values in tree nodes: a – node 2; б – node 6

Таблица 2

Матрица ошибок решающего дерева (рис. 6)
Decision Tree Error Matrix (Fig. 6)

Характеристики классов		Предсказанный класс	
		0	1
Истинный класс	0	97,995	2,005
	1	1,292	98,708

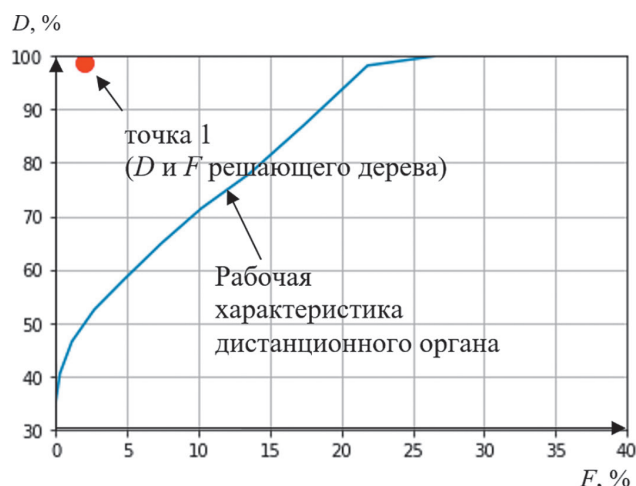


Рис. 11. Соотношение характеристик эффективности функционирования алгоритма решющего дерева и дистанционного органа релейной защиты

Fig. 11. Correlation between the performance characteristics of the decision tree algorithm and remote relay protection unit

Таблица 3

Диапазоны значений параметров решющего дерева, используемых в методе поиска по сетке
Ranges of decision tree parameters used in the grid search method

Параметр	Значение
Глубина решющего дерева	3, 5, 7, 9, 11, 13, 15
Количество объектов в листовом узле	1, 3, 5, 10

горитму классификации режимов электрической сети соответствует матрица ошибок (табл. 4)

Сопоставительный анализ табл. 2 и 4 показывает, что оптимизация параметров алгоритма решющего дерева применительно к устройству РЗА (рис. 1) позволила не только увеличить долю обнаруженных КЗ на линии, но и снизить вероятность ложного срабатывания защиты при КЗ за трансформатором.

Структурная схема устройства РЗА, реализующего алгоритм решющего дерева. На рис. 12 приведена структурная схема устройства РЗА, реализующего принятие решения по алгоритму решющего дерева.

Перед началом функционирования устройства РЗА алгоритм решющего дерева подлежит обучению на выборке информационных параметров (например

табл. 1), полученной по результатам предварительного имитационного моделирования. В ходе моделирования имитируются все возможные режимы, характерные для выбранного фрагмента электрической сети. Для каждого из режимов, имеющего свой номер-идентификатор, результаты обучения в виде уставочных значений признаков, а также вариантов (параметров) подключения информационных признаков к узловым схемам сравнения сохраняются в блоке памяти.

В процессе функционирования устройства РЗА на его вход непрерывно поступают мгновенные значения (комплексы) токов и напряжений. В блоке оценки параметров режима вычисляются параметры (информационные признаки), участвующие в реализации алгоритма решющего дерева (например, в табл.1). В соответствии с текущим режимом фрагмента электрической сети на вход блока памяти поступает управляющий сигнал (например, в виде номера режима), обеспечивающий выдачу с блока памяти на вход блока оценки параметров режима информации о параметрах коммутации для передачи информационных признаков в требуемые схемы сравнения (узлы решющего дерева). Дополнительно по управляющему сигналу на входе блока памяти реализуется выдача уставочных значений на схемы сравнения для реализации принятия решения алгоритмом решющего дерева (устройством РЗА). Соединение схем сравнения выполняется в соответствии с выбранной структурой алгоритма решющего дерева (например рис. 6).

На выходах схем сравнения, соответствующих листовым узлам алгоритма решющего дерева, формируются сигналы, по которым принимается решение относительно распознаваемого режима (α - или β -режим) и которые поступают в блок логики. В блоке логики формируется окончательное решение относительно распознаваемого режима (например по мажоритарному принципу), которое выдается на выход устройства РЗА.

Следует отметить, что реализация алгоритма решющего дерева, например для микропроцессорных устройств РЗА, не потребует модернизации аппаратной части, а связана с доработкой специального программного обеспечения.

Распознавание режимов электрической сети с применением групповых алгоритмов машинного обучения. Эффективным способом повышения вероятности правильной классификации режимов электри-

Таблица 4

Матрица ошибок решющего дерева в результате оптимизации параметров
Decision Tree Error Matrix as a Result of Parameter Optimization

Характеристики классов		Предсказанный класс	
		0	1
Истинный класс	0	98,619	1,381
	1	0,758	99,242

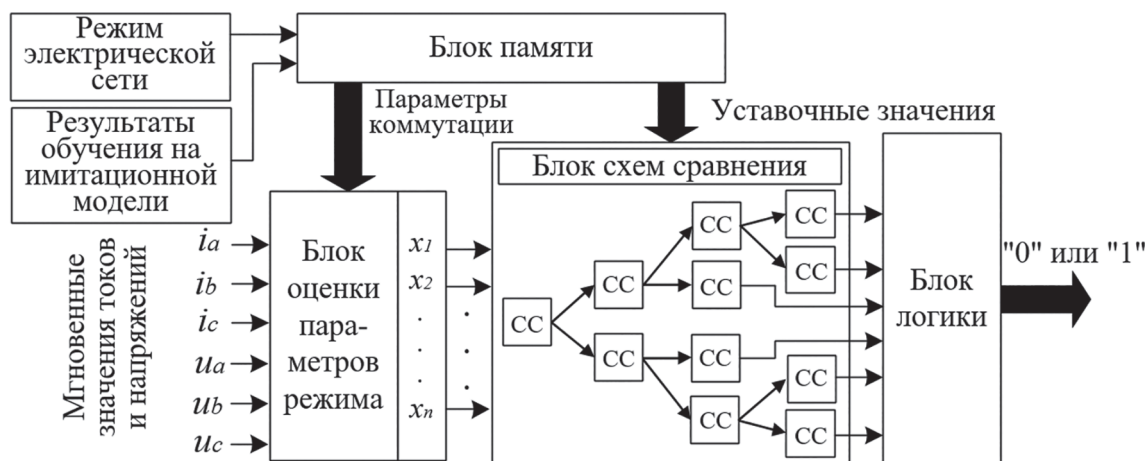


Рис. 12. Структурная схема устройства РЗА, реализующего алгоритм решающего дерева

Fig. 12. Structural diagram of the relay protection device that implements the decision tree algorithm

ческой сети является совместное использование нескольких алгоритмов машинного обучения. Особенно явно такой эффект проявляется, когда в групповой алгоритм входят алгоритмы, существенно отличающиеся принципами функционирования. При использовании множества различных вычислительных процедур возможны ситуации, когда их ошибки расчетов взаимно компенсируют друг друга.

Для формирования групповых алгоритмов классификации режимов электрической сети ниже анализируется применение двух алгоритмов машинного обучения: случайный лес и градиентный бустинг [25]. В качестве базовых алгоритмов, входящих в состав групповых, в обоих случаях используются рассмотренные выше решающие деревья.

Случайный лес [25] представляет собой групповой алгоритм машинного обучения, состоящий из некоторого количества решающих деревьев, каждое из которых обучено независимо друг от друга на различных подмножествах исходных данных. При классификации режимов электрической сети вероятность каждого класса определяется как доля входящих в групповой алгоритм решающих деревьев, указавших на заданный класс.

Основными параметрами, алгоритма случайного леса являются:

количество решающих деревьев в групповом алгоритме ($n_estimators$);

объем обучающей выборки, который будет использоваться для обучения каждого дерева ($max_samples$). При этом формирование подвыборки для каждого решающего дерева реализуется случайным образом по схеме с повторениями;

количество информационных признаков режимов электрической сети, которое будет использоваться для обучения каждого дерева ($max_features$);

параметры, характеризующие решающие деревья, входящие в групповой алгоритм.

Определим параметры группового алгоритма случайного леса применительно к решаемой задаче классификации режимов электрической сети (рис. 1). Примем максимальное число информационных признаков ($max_features$) в групповом алгоритме равным общему количеству информационных признаков (табл. 1), а максимальное число выборочных отсчетов ($max_samples$) – равное числу отсчетов в обучающей выборке. При такой схеме за счет выборки с повторениями для обучения каждого дерева будет использовано в среднем 63 % ранее не использованных значений из обучающей выборки. Количество отдельных алгоритмов ($n_estimators$) будет определяться в результате подбора параметров группового алгоритма.

Как и в случае с обучением алгоритма решающего дерева, воспользуемся методом поиска по сетке для определения параметров группового алгоритма случайного леса, обеспечивающих максимальные вероятности правильного распознавания режимов электрической сети. Список изменяемых параметров приведен в табл. 5.

Таблица 5

Значения параметров группового алгоритма случайного леса, используемых при реализации метода поиска по сетке

Parameter values of the group random forest algorithm used in the implementation of the grid search method

Параметр	Перечень значений
Максимальная глубина решающего дерева	3, 5, 7, 9, 11, 13, 15
Минимальное количество объектов в листовом узле	1, 3, 5, 10, 15
Количество деревьев в групповом алгоритме	50, 100, 150

Матрица ошибок решающего дерева в результате оптимизации параметров
Decision Tree Error Matrix as a Result of Parameter Optimization

Характеристики классов		Предсказанный класс	
		0	1
Истинный класс	0	99,153	0,847
	1	0,535	99,465

В отличие от решающих деревьев групповой алгоритм случайного леса позволяет построить рабочую характеристику (рис. 13), а площадь под кривой рабочей характеристики может быть использована в качестве метрики при подборе параметров группового алгоритма классификации.

В результате реализации метода поиска по сетке были найдены параметры группового алгоритма случайного леса, обеспечивающие наилучшее распознавание режимов электрической сети (рис. 1): максимальная глубина решающего дерева 9; минимальное количество объектов в листовом узле 1; количество деревьев в групповом алгоритме 100.

Матрица ошибок для обученного группового алгоритма случайного леса при пороговом (уставочном) значении 0,5 соответствует табл. 6.

На рис. 13 показана рабочая характеристика группового алгоритма случайного леса, а также значения величин D и F , характерные для алгоритма решающего дерева.

Из рис. 13 и анализа табл. 6 видно, что применение группового алгоритма случайного леса снижает вероятность ошибок классификации режимов электрической сети по сравнению с одиночным алгоритмом решающего дерева, а также дает возможность регулировать уставку (порог) срабатывания для очувствления либо загробления релейной защиты. Поскольку полученный групповой алгоритм случайного леса включает 100 одиночных алгоритмов решающего дерева, то становится очевидным, что повышение распознающей способности устройства релейной защиты достигается существенным его усложнением.

Градиентный бустинг является альтернативным способом формирования групповых алгоритмов классификации режимов электрической сети [25]. Идея бустинга заключается в том, что каждый последующий отдельный распознающий алгоритм, входящий в группу, должен обучиться так, чтобы компенсировать ошибки всех предыдущих. Тогда при реализации процедуры классификации электрического режима с применением группового алгоритма необходимо сложить решения всех входящих в состав группы алгоритмов.

При применении градиентного бустинга базовыми алгоритмами являются алгоритмы регрессии, причем чаще всего используются модификации решающих деревьев для задачи регрессии (*Regressor decision tree*) [25]. Однако алгоритмы регрессии могут иметь различные значения выходных параметров, а сумма их конеч-

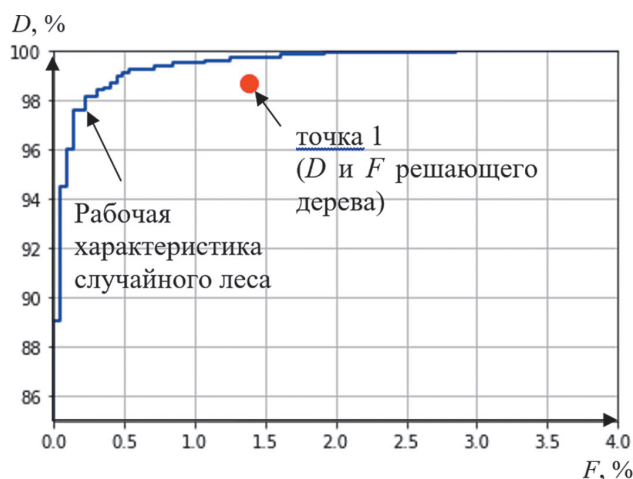


Рис. 13. Показатели распознающей способности алгоритмов решающего дерева и случайного леса

Fig. 13. Indicators of recognition ability of decision tree and random forest algorithms

ных решений может располагаться в диапазоне от $-\infty$ до $+\infty$. Для того чтобы получить вероятности классов, лежащие в диапазоне $[0, 1]$, применяют ограничивающую функцию, как правило, сигмоидальную (рис. 14), соответствующую выражению

$$\text{sigm}(x) = \frac{1}{1 + e^{-x}}. \quad (14)$$

При обучении группового алгоритма классификации на основе градиентного бустинга задают функцию

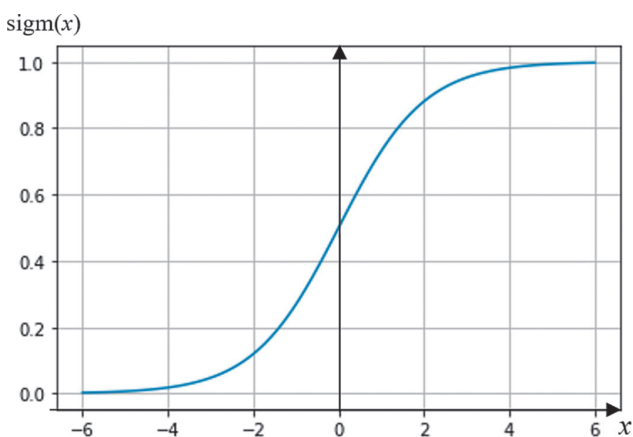


Рис. 14. График функции сигмоиды

Fig. 14. Graph of the sigmoid function

Таблица 7

потерь, зависящую от соотношения решений, принятых групповым алгоритмом, и истинных значений классов. В ходе процедуры обучения группового алгоритма градиентного бустинга обеспечивают минимизацию функции потерь, в качестве которой выбирают зависимость $\log \text{loss}(x)$, соответствующую равенству

$$\log \text{loss}(x) = -\sum_{i=0}^{N-1} y_i \log(p_i) + (1 - y_i)(1 - p_i), \quad (15)$$

где p_i – предсказанная моделью вероятность того, что групповой алгоритм отнесет i -й режим электрической сети к классу «1»; y_i – истинный класс i -го объекта (0 или 1); N – объем выборки.

Обучение группового алгоритма градиентного бустинга также выполняется с использованием итерационной процедуры и метода поиска по сетке [25]. На каждой итерации рассчитывается значение функции потерь (15) и ее градиента. Дополнительно вводится параметр коэффициента скорости обучения (*learning rate*). При использовании регулируемого коэффициента скорости обучения достигается положительный эффект, направленный на недопущение ни одного пропуса локального минимума функции потерь.

Помимо коэффициента скорости обучения, в число основных параметров группового алгоритма градиентного бустинга входят: количество деревьев в групповом алгоритме (*n_estimators*), а также параметры, описывающие каждое решающее дерево в отдельности.

Значения параметров группового алгоритма градиентного бустинга, которые использовались при обучении с применением метода поиска по сетке, приведены в табл. 7.

В результате оптимизационного подбора параметров градиентного бустинга с целевой метрикой, в качестве которой выступала площадь под кривой рабочей характеристики, были получены значения: максимальная глубина решающего дерева 7, минимальное количество объектов в листовом узле 15, количество деревьев в групповом алгоритме 100, коэффициент скорости обучения 0,1.

Матрица ошибок, сформированная в результате тестирования группового алгоритма градиентного бустинга, соответствует табл. 8.

На рис. 15 приведены рабочие характеристики групповых алгоритмов градиентного бустинга и случайного леса.

Диапазоны значений параметров градиентного бустинга

Ranges of gradient boosting parameters

Параметр	Перечень значений
Максимальная глубина решающего дерева	3, 5, 7, 9, 11, 13, 15
Минимальное количество объектов в листовом узле	1, 3, 5, 10, 15
Количество деревьев в групповом алгоритме	50, 100, 150
Коэффициент скорости обучения	0,01; 0,05; 0,1

Анализ характеристик (рис. 15) групповых алгоритмов градиентного бустинга и случайного леса позволяет сделать вывод об их сопоставимой эффективности. Градиентный бустинг имеет несколько большую чувствительность при низких значениях F , однако при допустимой вероятности ложного срабатывания более 0,2 % более предпочтительным является случайный лес. Ввиду близости кривых рабочих характеристик (рис. 15) можно предположить, что совокупное использование различных групповых алгоритмов машинного обучения позволяет получить предельные значения параметров, характеризующих потенциальные воз-

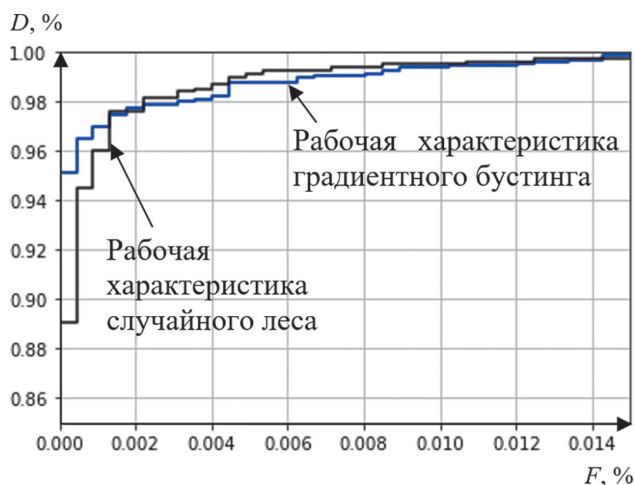


Рис. 15. Показатели распознающей способности групповых алгоритмов градиентного бустинга и случайного леса

Fig. 15. Indicators of recognition ability of group algorithms for gradient boosting and random forest

Таблица 8

Матрица ошибок группового алгоритма градиентного бустинга
Error matrix of the group gradient boosting algorithm

Характеристики классов		Предсказанный класс	
		0	1
Истинный класс	0	98,797	1,203
	1	0,49	99,51

возможности устройства РЗА по распознаванию режимов электрической сети.

Важно отметить, что несмотря на высокую распознающую способность групповых алгоритмов градиентного бустинга и случайного леса, их применение в устройствах РЗА весьма сомнительно. Это обусловлено тем, что одновременная реализация за время срабатывания РЗА (десятки миллисекунд) сотни отдельных алгоритмов решающего дерева потребует крайне высокой производительности и стоимости соответствующих устройств. Однако для решения задач распознавания режимов электрической сети, не требующих высокого быстродействия, применение групповых алгоритмов машинного обучения является несомненно перспективным.

Выводы. Множественное имитационное моделирование на основе метода Монте-Карло позволяет не только получать распределения параметров токов и напряжений в различных режимах функционирования электрической сети, но и создает основу для синтеза статистических алгоритмов РЗА, основанных на методах машинного обучения.

Применение метода решающих деревьев для организации РЗА на ЛЭП с ответвлениями позволило сформировать алгоритм многопараметрической защиты, отличающийся повышенной чувствительностью при КЗ на ответвлении за трансформатором.

Совокупное использование групповых алгоритмов градиентного бустинга и случайного леса, основанных на машинном обучении, позволяет получить предельные значения параметров, характеризующих потенциальные возможности устройства РЗА по распознаванию режимов электрической сети. Однако их аппаратная реализация весьма сомнительна ввиду очень высоких требований к производительности вычислительных компонентов устройств РЗА.

Исследование выполнено в рамках государственного задания в сфере научной деятельности (тема №FSWE-2022-0005).

СПИСОК ЛИТЕРАТУРЫ

1. Положение ПАО «Россети» «О единой технической политике в электросетевом комплексе» Приложение №4 к решению Совета директоров ПАО «МРСК Центра». Протокол от 25.05.2021 № 19/21.
2. Шарыгин М.В., Куликов А.Л. Статистические методы распознавания режимов в релейной защите и автоматике сетей электроснабжения. – Электрические станции, 2018, № 2 (1039), с. 32–39.
3. Лоскутов А.А., Митрович М., Осокин В.Ю. Повышение распознаваемости режимов функционирования системы электроснабжения на основе методов машинного обучения. – Релейная защита и автоматизация, 2020, № 4 (41), с. 26–34
4. Лоскутов А.А., Пелевин П.С., Митрович М. Разработка логической части интеллектуальной многопараметрической релейной защиты. – Электричество, 2020, № 5. с. 12–18.
5. Куликов А.Л., Лоскутов А.А., Совина А.Н. Использование машинного обучения и искусственных нейронных сетей для распознавания витковых замыканий в силовых трансформаторах. – Электричество, 2022, № 10, с. 34–44.

6. Рибейро П.Ф. и др. Обработка сигналов в интеллектуальных сетях энергосистем. М.: Техносфера, 2020, 496 с.
7. Rebizant W., Szafran J., Wiszniewski A. Digital Signal Processing in Power System Protection and Control. London: Springer, 2011, 316 p.
8. Hasan A.N., Pouabe P.S., Twala B. The Use of Machine Learning Techniques to Classify Power Transmission Line Fault Types and Locations. – International Conference on Optimization of Electrical and Electronic Equipment, 2017, pp. 221–226, DOI:10.1109/OPTIM.2017.7974974.
9. Witten I.H., Frank E. Data Mining: Practical Machine Learning Tools and Techniques. Amsterdam: Elsevier, 2005, 525 p.
10. Michie D., Spiegelhalter D., Taylor C. Machine Learning, Neural and Statistical Classification. Ellis Horwood, 1994, 290 p.
11. Bishop C.M. Pattern Recognition and Machine Learning. Berlin: Springer, 2006, 738 p.
12. Breiman L. et al. Classification and Regression Trees. New York: Routledge, 1984, 368 p.
13. Yongli Z., Limin H., Jinling L. Bayesian Networks-Based Approach for Power Systems Fault Diagnosis. – IEEE Transactions on Power Delivery, 2006, No. 21, pp. 636–669, DOI: 10.1109/TPWRD.2005.858774.
14. Ермаков С.М. Метод Монте-Карло в вычислительной математике. СПб., 2009, 192 с.
15. Шаф дифференциальной защиты линии типа ШЭ2607 091. Руководство по эксплуатации. ЭКРА.656453.126 РЭ. Чебоксары: ООО НПП «ЭКРА», 2008.
16. Лямец Ю.Я. и др. Эффекты многомерности в релейной защите. – Электричество, 2011, № 9, с. 48–54.
17. Андреев В.А. Релейная защита и автоматика систем электроснабжения. М.: Высшая школа, 2006, 639 с.
18. Fawcett T. An Introduction to ROC Analysis. – Pattern Recognition Letters, 2006, No. 8, pp. 861–874, DOI:10.1016/j.patrec.2005.10.010.
19. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Berlin: Springer, 2001, 745 p.
20. Ширман Я.Д., Манжос В.Н. Теория и техника обработки радиолокационной информации на фоне помех. М.: Радио и связь, 1981, 416 с.
21. Kulikov A., Loskutov A., Bezdushniy D. Relay Protection and Automation Algorithms of Electrical Networks Based on Simulation and Machine Learning Methods. – Energies, 2022, vol. 15, No. 18, p. 6525, DOI:10.3390/en15186525.
22. Xia Y.Q., Li K.K., David A.K. Adaptive Relay Setting for Stand-Alone Digital Distance Protection. – IEEE Transactions on Power Delivery, 1994, vol. 9, No.1. pp. 480–491, DOI: 10.1109/61.277720.
23. Qi W., Swift G., McLaren P. Distance Protection Using an Artificial Neural Network. – Sixth International Conference on Developments in Power System Protection (Conf. Publ. No. 434). Nottingham, UK, 1997. pp. 286–290, DOI: 10.1049/cp:19970083.
24. Hutter F., Kotthoff L., Vanschoren J. Automated Machine Learning. Cham, Switzerland: Springer, 2019, 219 p.
25. Mueller A.S., Guido S. Introduction to Machine Learning with Python. Sebastopol, CA: O'Reilly, 2016, 338 p.

Поступила в редакцию [02.05.2023]
Принята к публикации [25.05.2023]

Авторы:



Куликов Александр Леонидович – доктор техн. наук, профессор, профессор кафедры «Электроэнергетика, электроснабжение и силовая электроника», Нижегородский государственный технический университет им. Р.Е. Алексеева, Нижний Новгород, Россия.



Лоскутов Антон Алексеевич – кандидат техн. наук, доцент, доцент кафедры «Электроэнергетика, электроснабжение и силовая электроника»; старший научный сотрудник лаборатории «Автономные гибридные электроэнергетические комплексы», Нижегородский государственный технический университет им. Р.Е. Алексеева, Нижний Новгород, Россия.



Бездушный Дмитрий Игоревич – кандидат техн. наук, ведущий инженер по разработке, ПАО Сбербанк России, Нижний Новгород, Россия.

Elektrichestvo, 2023, No. 7, pp. 20–34

DOI:10.24160/0013-5380-2023-7-20-34

Improving the Recognition of Emergency Modes by Relay Protection Using Decision Tree Methods

KULIKOV Aleksandr L. (Nizhny Novgorod State Technical University n.a. Alekseev, Nizhny Novgorod, Russia) – Professor of the Electric Power Engineering, Power Supply and Power Electronics Dept., Dr. Sci. (Eng.), Professor.

LOSKUTOV Anton A. (Nizhny Novgorod State Technical University n.a. Alekseev, Nizhny Novgorod, Russia) – Docent of the Electric Power Engineering, Power Supply and Power Electronics Dept., Cand. Sci. (Eng.), Docent.

BEZDUSHNYY Dmitriy I. (PJSC Sberbank of Russia, Nizhny Novgorod, Russia) – Data Scientist, Cand. Sci. (Eng.).

In the context of introducing modern active elements and devices into electrical networks (flexible power transmission lines, controlled shunt reactors, energy storage devices, distributed generation sources, etc.), conventional methods for recognizing emergency modes to secure correct operation of relay protection and automatic control devices (RPA) sometimes become ineffective. With the development and mass-scale application of the IEC 61850 standard, as well as simulation software systems, access to a large amount of information about the electrical network operation modes becomes available, which opens the possibility to develop fundamentally new algorithms for RPA, including those based on machine learning methods. The article explores a new approach to fault recognition in power lines with branches by simultaneously analyzing several information features and applying group machine learning algorithms: decision tree, random forest, and gradient boosting. To obtain training samples, simulation and the Monte Carlo method are used. The results of testing the studied methods have shown the required flexibility, the ability to use a large number of information parameters, as well as better results of fault recognition in comparison with the distance protection relay pickup device. The implementation of the decision tree algorithm for RPA microprocessor devices will not require hardware upgrades, but only involves refinement of special software.

Key words: short circuit recognition, machine learning, decision tree, random forest, gradient boosting, relay protection and automatic control devices, simulation, algorithm

The research was carried out within the framework of the state assignment in the field of scientific activity (topic no. FSWE-2022-0005).

REFERENCES

1. **Polozhenie** PAO «Rosseti» «O edinoy tekhnicheskoy politike v elektrosетevom komplekse» Prilozhenie №4 k resheniyu Soveta direktorov PAO «MRSK TSentra». Protokol ot 25.05.2021 № 19/21 (The Regulation of PJSC ROSSETI "On the Unified Technical Policy in the Electric Grid Complex" Appendix No. 4 to the decision of the Board of Directors of PJSC IDGC of Centre. Protocol No. 19/21 dated 05/25/2021.).

2. **Sharygin M.V., Kulikov A.L.** *Elektricheskie stantsii* – in Russ. (Power Plants), 2018, No. 2 (1039), pp. 32–39.

3. **Loskutov A.A., Mitrovich M., Osokin V.Yu.** *Releynaya zashchita i avtomatizatsiya* – in Russ. (Relay Protection and Automation), 2020, No. 4 (41), pp. 26–34.

4. **Loskutov A.A., Pelevin P.S., Mitrovich M.** *Elektrichestvo* – in Russ. (Electricity), 2020, No. 5, pp. 12–18.

5. **Kulikov A.L., Loskutov A.A., Sovina A.N.** *Elektrichestvo* – in Russ. (Electricity), 2022, No. 10, pp. 34–44.

6. **Ribeyro P.F. et al.** *Obrabotka signalov v intellektual'nyh setyah energosistem* (Signal Processing in Intelligent Power Grid Networks). M.: Tekhnosfera, 2020, 496 p.
7. **Rebizant W., Szafran J., Wiszniewski A.** *Digital Signal Processing in Power System Protection and Control*. London: Springer, 2011, 316 p.
8. **Hasan A.N., Pouabe P.S., Twala B.** The Use of Machine Learning Techniques to Classify Power Transmission Line Fault Types and Locations. – International Conference on Optimization of Electrical and Electronic Equipment, 2017, pp. 221–226, DOI:10.1109/OPTIM.2017.7974974.
9. **Witten I.H., Frank E.** *Data Mining: Practical Machine Learning Tools and Techniques*. Amsterdam: Elsevier, 2005, 525 p.
10. **Michie D., Spiegelhalter D., Taylor C.** *Machine Learning, Neural and Statistical Classification*. Ellis Horwood, 1994, 290 p.
11. **Bishop C.M.** *Pattern Recognition and Machine Learning*. Berlin: Springer, 2006, 738 p.
12. **Breiman L. et al.** *Classification and Regression Trees*. New York: Routledge, 1984, 368 p.
13. **Yongli Z., Limin H., Jinling L.** Bayesian Networks-Based Approach for Power Systems Fault Diagnosis. – IEEE Transactions on Power Delivery, 2006, No. 21, pp. 636–639, DOI: 10.1109/TPWRD.2005.858774.
14. **Ermakov S.M.** *Metod Monte-Karlo v vychislitel'noy matematike* (Monte Carlo Method in Computational Mathematics). SPb., 2009, 192 p.
15. **Shkaf** *differentzial'noy zashchity linii tipa SHE2607 091. Rukovodstvo po ekspluatatsii*. EKRA.656453.126 RE (Differential Protection Cabinet of the Line Type ShE2607 091. Operating Manual. EKRA.656453.126 RE). Cheboksary: OOO NPP «EKRA», 2008.
16. **Lyamets Yu.Ya. et al.** *Elektrichestvo – in Russ. (Electricity)*, 2011, No. 9, pp. 48–54.
17. **Andreev V.A.** *Releynaya zashchita i avtomatika sistem elektrosnabzheniya* (Relay Protection and Automation of Power Supply Systems). M.: Vysshaya shkola, 2006, 639 p.
18. **Fawcett T.** An Introduction to ROC Analysis. – Pattern Recognition Letters, 2006, No. 8, pp. 861–874, DOI:10.1016/j.patrec.2005.10.010.
19. **Hastie T., Tibshirani R., Friedman J.** *The Elements of Statistical Learning*. Berlin: Springer, 2001, 745 p.
20. **Shirman Ya.D., Manzhos V.N.** *Teoriya i tekhnika obrabotki radiolokatsionnoy informatsii na fone pomekh* (Theory and Technique of Radar Information Processing Against the Background of Interference). M.: Radio i svyaz', 1981, 416 p.
21. **Kulikov A., Loskutov A., Bezdushniy D.** Relay Protection and Automation Algorithms of Electrical Networks Based on Simulation and Machine Learning Methods. – Energies, 2022, vol. 15, No. 18, p. 6525, DOI:10.3390/en15186525.
22. **Xia Y.Q., Li K.K., David A.K.** Adaptive Relay Setting for Stand-Alone Digital Distance Protection. – IEEE Transactions on Power Delivery, 1994, vol. 9, No.1. pp. 480–491, DOI: 10.1109/61.277720.
23. **Qi W., Swift G., McLaren P.** Distance Protection Using an Artificial Neural Network. – Sixth International Conference on Developments in Power System Protection (Conf. Publ. No. 434). Nottingham, UK, 1997. pp. 286–290, DOI: 10.1049/cp:19970083.
24. **Hutter F., Kotthoff L., Vanschoren J.** *Automated Machine Learning*. Cham, Switzerland: Springer, 2019, 219 p.
25. **Mueller A.S., Guido S.** *Introduction to Machine Learning with Python*. Sebastopol, CA: O'Reilly, 2016, 338 p.

Received [02.05.2023]
Accepted [25.05.2023]