

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ ДЕФИЦИТОВ МОЩНОСТИ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ

Бояркин Денис Александрович

Аспирант, Иркутский национальный исследовательский
технический университет, 664074, г. Иркутск, ул. Лермонтова 83,
e-mail: denisboyarkin@isem.irk.ru

Крупенёв Дмитрий Сергеевич

К.т.н., с.н.с., Институт систем энергетики им. Л.А. Мелентьева СО РАН,
664130 г. Иркутск, ул. Лермонтова 130, e-mail: krupenev@isem.irk.ru

Якубовский Дмитрий Викторович

Аспирант, Институт систем энергетики им. Л.А. Мелентьева СО РАН,
664130 г. Иркутск, ул. Лермонтова 130, e-mail: yakubovskii.dmit@mail.ru

Аннотация. В статье рассматривается вопрос повышения вычислительной эффективности методики оценки балансовой надёжности электроэнергетических систем на основе метода Монте-Карло. При использовании данного метода скорость и точность выполнения расчёта зависит от числа анализируемых случайных состояний моделируемой системы. Под анализом понимается решение задачи потокораспределения для каждого случайно сгенерированного состояния. Это достаточно трудоёмкий процесс, поэтому сокращение числа анализируемых состояний при сохранении точности оценки позволит повысить скорость расчёта в целом. Для этого предлагается использовать методы машинного обучения, задача которых – определять дефицит мощности энергосистемы без обращения к оптимизационным методам. При расчётах были использованы метод опорных векторов и случайного леса, оценена жизнеспособность предлагаемого подхода при оценке балансовой надёжности тестовой электроэнергетической системы.

Ключевые слова: электроэнергетические системы, оценка надёжности, метод Монте-Карло, машинное обучение.

Цитирование: Бояркин Д.А., Крупенев Д.С., Якубовский Д.В. Использование методов машинного обучения для определения дефицитов мощности электроэнергетических систем // Информационные и математические технологии в науке и управлении. 2018. № 4 (12). С. 61–69. DOI: 10.25729/2413-0133-2018-4-06

Введение. Бесперебойная и качественная поставка электроэнергии – это одни из основных характеристик работы электроэнергетической системы (ЭЭС). Для их обеспечения необходимо заблаговременно предпринимать комплекс соответствующих мер, таких как ввод резервных мощностей, усиление связей и др. Такие меры должны быть своевременно запланированными, быть технически и экономически эффективными. Для этого предпринимаемые решения должны основываться на корректной оценке надёжности энергосистемы для прогнозируемых режимов потребления. Таким образом, заблаговременная оценка балансовой надёжности (БН) ЭЭС является необходимым условием, как для функционирования, так и развития современных энергосистем. Особенно

актуальной она становится с развитием таких направлений в энергетике, как распределённая генерация, интеллектуализация, возобновляемые источники энергии, которые усложняют ЭЭС и, соответственно, вносят свои коррективы в процесс оценки надёжности ЭЭС.

Методика оценки балансовой надёжности (БН) ЭЭС, основанная на методе Монте-Карло, состоит из трех этапов [4]:

- 1) формирование расчетных состояний ЭЭС,
- 2) расчет режимов сформированных состояний ЭЭС для определения дефицитов мощности,
- 3) вычисление показателей надёжности по результатам расчётов, полученных на первых двух этапах.

Задача оценки балансовой надёжности ЭЭС, как правило, решается для годового интервала, при этом учитываются плановые ремонты энергетического оборудования, случайные отклонения нагрузки и другие факторы функционирования ЭЭС.

Сложность при решении данной задачи представляет большая размерность. Так, для достижения приемлемой точности оценки надёжности ЭЭС необходимо сформировать порядка 10^3 состояний ЭЭС на каждый час расчетного периода, для года это 8760 часов. Далее необходимо рассчитать дефициты мощности для каждого состояния. Как результат на выполнение этих операций тратится необоснованно много времени, что в свою очередь препятствует решению последующих задач, например, оптимизации балансовой надёжности ЭЭС.

В настоящее время методы машинного обучения все шире находят применение в различных областях, так, например, в [2, 11] данные методы используются для интеллектуальной оценки режимной надёжности электроэнергетических систем. В нашей же статье предлагается применить эти методы для ускорения расчета дефицитов мощности при оценке БН ЭЭС. Для этого предлагается использовать методы машинного обучения с классификатором для определения дефицитности сгенерированных состояний, тем самым сократив число состояний, дефицит которых необходимо определять на втором этапе.

В статье рассмотрены два метода классификации – метод опорных векторов [1] и метод случайного леса [8], оценивается их эффективность и точность исходя из постановки задачи, а также применимость данной методики в целом. Выбор данных методов обусловлен высокой скоростью их обучения, так как в рамках постановки задачи общая скорость вычислений является критичным параметром.

1. Постановка задачи оптимизации. Как уже было отмечено, методика оценки БН ЭЭС состоит из трех вычислительных этапов расчёта. На первом этапе формируются состояния ЭЭС при моделировании случайных событий (отказы генерирующих установок, линий электропередачи (ЛЭП), нерегулярные колебания нагрузки потребителей), которые могут произойти в некотором режиме. Моделирование одного случайного события K_i , $i=1, \dots, h$, вероятность которого равна p_i , $i=1, \dots, h$, происходит при разыгрывании одного случайного числа r_i , $i=1, \dots, h$, из равномерно распределенного множества на интервале $[0, 1]$. Если при этом r_i находится в интервале $[0, p_i]$ считается, что событие наступило, если нет, то не наступило, т.е.:

$$K_i = \begin{cases} 1, & \text{если } r_i \in [0, p_i] \\ 0, & \text{если } r_i \in (p_i, 1] \end{cases}, i = 1, \dots, h. \quad (1)$$

Распределенное множество представляет собой последовательность случайных чисел, сгенерированных с помощью генераторов псевдослучайных (ГПСЧ) [3] или квазислучайных чисел. В наших расчётах использовались Вихрь Мерсенна [10] и ЛП_т-последовательности [7].

На этапе расчета режимов сформированных случайных состояний ЭЭС для каждого состояния рассчитывается оптимальный по минимуму дефицита мощности установившийся режим. При оценке балансовой надёжности ЭЭС математическую формулировку данной задачи можно представить в следующем виде [9]:

при оценке дефицита мощности u -го состояния ЭЭС, $u = 1, \dots, N$ требуется найти:

$$\sum_{i=1}^f y_i \rightarrow \max, \quad (2)$$

учитывая балансовые ограничения

$$g_i - y_i + \sum_{j=1}^f (1 - a_{ji} z_{ji}) z_{ji} - \sum_{j=1}^f z_{ij} = 0 \\ i = 1, \dots, f, i \neq j \quad (3)$$

и линейные ограничения-неравенства на переменные

$$y_i \leq y_i^u, \\ g_i \leq g_i^u, \quad i = 1, \dots, f, j = 1, \dots, f, i \neq j, \\ z_{ij} \leq z_{ij}^u, \\ y_i \geq 0, g_i \geq 0, z_{ij} \geq 0, \quad i = 1, \dots, f; j = 1, \dots, f; i \neq j \quad (4)$$

где g_i – потребляемая мощность в узле i , МВт; g_i^u – располагаемая генерирующая мощность в узле i , МВт; y_i – величина покрытия нагрузки в узле i , МВт; y_i^u – нагрузка в узле i , МВт; z_{ij} – поток мощности из узла i в узел j , МВт; z_{ij}^u – пропускная способность ЛЭП между узлами i и j , МВт; a_{ij} – заданные коэффициенты удельных потерь мощности при её передаче из узла i в узел j , $i \neq j$; $i = 1, \dots, f$; $j = 1, \dots, f$; $u = 1, \dots, N$.

Задача расчёта режима является трудоёмкой задачей, поэтому сокращение числа рассчитываемых состояний позволит повысить вычислительную эффективность всей методики в целом. Предлагается разбить набор случайных состояний, полученных при выполнении первого этапа, на меньшую и на большую выборки. Определение размера каждой из выборок определяется пользователем. Расчёт показателей для состояний первой выборки не будет отличаться от изначального решения за исключением того, что сами состояния, как и результат расчёта сформируют обучающую выборку, на основе которой будет построен классификатор. Затем, для каждого состояния ЭЭС из большей выборки классификатор предсказывает его дефицитность и, в случае если оно бездефицитно, расчёт дефицита не производится, дефицит мощности системы принимается за 0. Если алгоритм классифицирует состояние как дефицитное, то точное значение дефицита мощности определяется при расчёте на втором этапе. Тем самым снижается число обращений к методам оптимизации на втором этапе.

2. Постановка задачи классификации. Рассматриваемая задача классификации дефицитности состояния по конфигурации работоспособности оборудования ЭЭС является типичной задачей обучения по прецедентам, где требуется построить алгоритм, аппроксимирующий целевую зависимость $X_A \rightarrow L$ между множеством объектов X_A и множеством всех ответов L по объектам обучающей выборки S . Для задачи классификации

дефицитности X_A представляет собой множество всех состояний ЭЭС, а $L \in [-1, 1]$, где -1 означает бездефицитное состояние, а 1 – дефицитное состояние.

Каждый объект данных $x_a \in X_A$, $a = 1, \dots, A$, представляет собой вектор в d -мерном пространстве и характеризует состояние ЭЭС, где d – количество признаков объекта, описывающих это состояние, A – число всех возможных состояний ЭЭС.

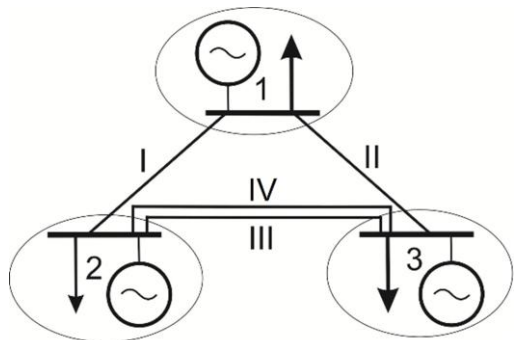


Рис. 1. Схема исследуемой ЭЭС

3. Экспериментальные исследования. В [5] были определены наилучшие способы генерации случайных состояний ЭЭС. Абсолютным лидером стал способ на основе ЛП_т-последовательностей, наилучшие результаты среди генераторов псевдослучайных чисел (ГСПЧ) показал Вихрь Мерсенна. Используя данные методы для формирования случайных состояний, мы вводим в модель оценки балансовой надёжности использование алгоритмов машинного обучения, такие как метод опорных векторов и случайный лес.

Были проведены экспериментальные исследования имитационного моделирования надёжности ЭЭС при использовании вышеприведенных ГПСЧ и методов машинного обучения. Для исследований была взята трёхузловая схема ЭЭС (рис. 1), она обладает наибольшей наглядностью и хорошо отражает работу реальных систем, поэтому наилучшим образом подходит в качестве тестовой схемы. Также её размерность позволяет провести полный перебор для получения эталонного значения дефицита мощности за приемлемое время, в то время как для сложных схем это невозможно. В табл. 1 представлены основные характеристики узлов ЭЭС, в табл. 2 основные характеристики связей.

Таблица 1. Основные характеристики узлов ЭЭС

№ узла	Единичная мощность генерирующего агрегата (МВт)	Количество генерирующих агрегатов (шт.)	Аварийность генерирующего агрегата (д.е.)	Нагрузка (МВт)
1	50	5	0,05	450
	100	2	0,05	
2	50	5	0,05	400
	100	1	0,05	
3	50	5	0,05	490
	100	1	0,05	
	200	1	0,05	

Таблица 2. Основные характеристики связей ЭЭС

№ связи	Вектор связи	Пропускная способность, МВт.	Количество ЛЭП в связи, шт.	Аварийность на 100 км, о.е.	Длина, км
1	1-2	150	1	0,001	400
2	1-3	150	1	0,001	400
3	2-3	100	2	0,001	400

Для оценки эффективности применения представленных методов машинного обучения была проведена серия опытов оценки надёжности представленной ЭЭС с определением математического ожидания (м.о.) дефицита мощности. На первом этапе м.о. дефицита мощности было рассчитано для каждого способа генерации случайных состояний без использования методов машинного обучения. Значение м.о. дефицита мощности для данной ЭЭС, найденное методом полного перебора было принято за эталонное. Так как ЭЭС состоит из 24 разыгрываемых элементов, было проанализировано $2^{24} = 16777216$ состояний. Эталонное значение м.о. дефицита мощности установилось на $\sim 61,78$ МВт.

В основу вычислений заложен метод статистических испытаний, которым естественным образом присуще приближенное вычисление искомых величин. При этом точность расчётов зависит, в первую очередь, от числа испытаний. Для характеристики ошибки на практике используют вероятную ошибку [6]:

$$\delta_{\text{вер}} = 0,675 \frac{\overline{D\xi}}{N} = 0,675 \frac{\text{с.к.о.}\xi}{N}, \quad (6)$$

где N – число испытаний, ξ – искомая случайная величина, D – дисперсия случайной величины.

Фактическая абсолютная ошибка зависит от использованных в расчёте случайных чисел и может оказаться как больше, так и меньше $\delta_{\text{вер}}$. Так что $\delta_{\text{вер}}$ показывает не границу ошибки, а её порядок [6].

Воспользовавшись данной формулой, получим график значений вероятных ошибок для модели (рис. 2).

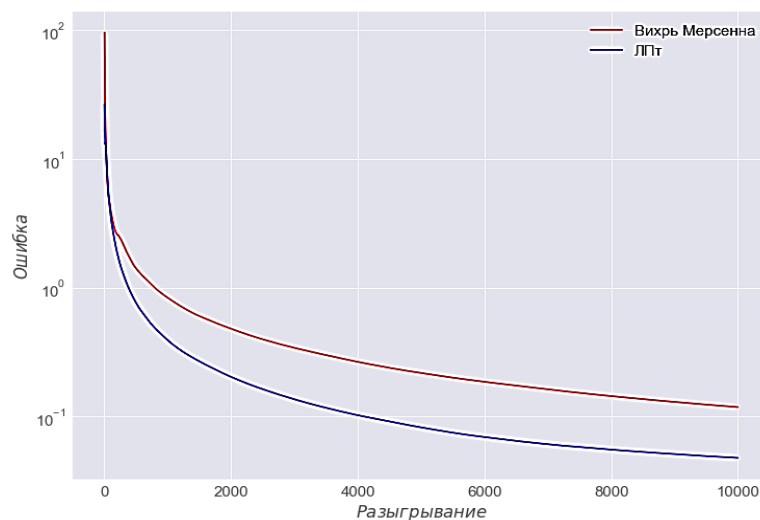


Рис. 2. Значения вероятной ошибки вычислений. Для отображения использован логарифмический масштаб по оси абсцисс

Как видно из графика, показанного на рис. 2, значения вероятной ошибки резко уменьшаются с увеличением числа испытаний и уже после 500-й итерации она асимптотически приближается к нулю. Также по графику видно незначительное преимущество модели, использующей в основе ЛП_τ-последовательности, её значения вероятной ошибки быстрее приближаются к 0, чем аналогичная модель на основе вихря Мерсенна. Так как вероятная ошибка показывает порядок ошибки, из графика можно сделать вывод, что модель на основе ЛП_τ-последовательности быстрее сойдётся к искомому значению м.о. дефицита мощности, чем на основе Вихря Мерсенна.

Результаты расчёта дефицита мощности представлены на рис. 3 (графики МТ – Вихрь Мерсенна и Sobol - ЛП_τ-последовательности).

Как видно из графика, показанного на рис. 3, наилучшим способом генерации случайных состояний является способ на основе ЛП_τ-последовательностей (на графике обозначено как Sobol). Значение м.о. дефицита мощности, рассчитанное с применением данного способа стабилизировалось на четырёхтысячной итерации, тогда как при использовании вихря Мерсенна (МТ) значение не стабилизируется и после 10000 итераций, что подтверждает выводы, полученные в [5], а также результаты расчёта вероятной ошибки.

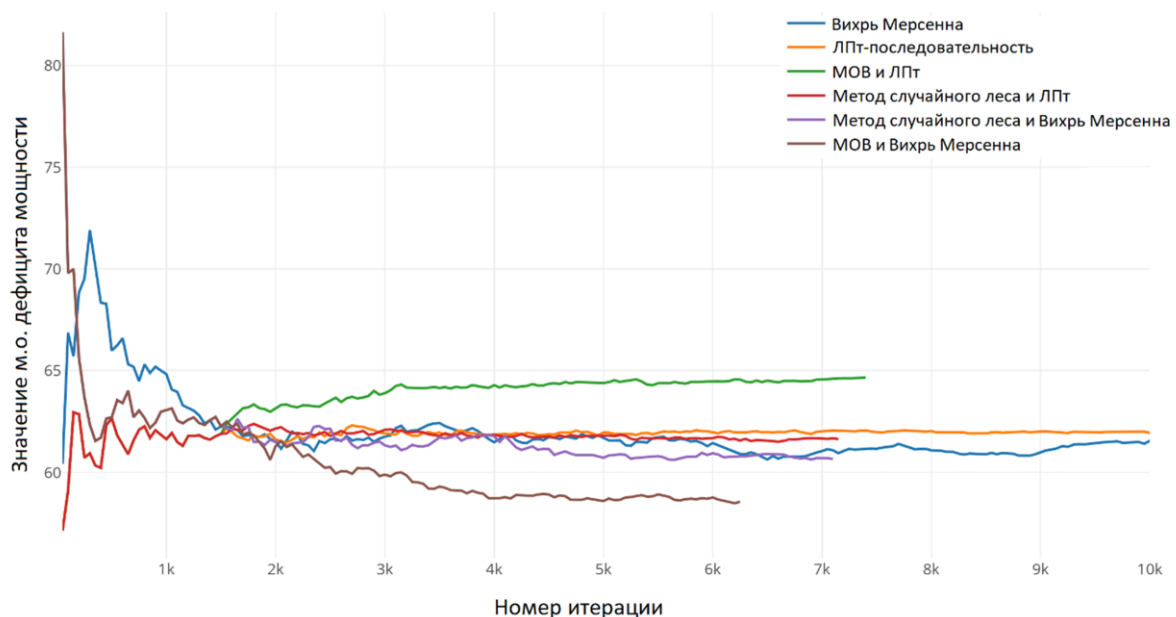


Рис. 3. Стабилизация значения м.о. дефицита мощности

На втором этапе были применены метод опорных векторов с линейным ядром и метод случайного леса. Подбор параметров моделей осуществлялся при отслеживании значения AUC ROC метрики при проведении пятикратной кросс-валидации. Для метода опорных векторов анализировалось использование различных значений штрафного коэффициента. Наилучшим оказалось значение, равное ~0.07742. Для метода случайного леса подбирались глубина деревьев и количество рассматриваемых случайных признаков при проведении расщепления. Наилучшим значением для обоих параметров оказалось 13. По результатам тестирования качества классификации полученных моделей на отложенной выборке были получены значения метрик, представленные в таблице 3.

Таблица 3. Значения метрик качества модели на отложенной выборке

Метод	Средняя доля правильных ответов	Значение F1-метрики
Метод опорных векторов	0.7439	0.8531
Случайный лес	0.9482	0.9650

Как видно из таблицы 3 метод случайного леса демонстрирует более высокое качество классификации (в терминах средней доли правильных ответов и F1-метрики) чем метод опорных векторов.

Использование методов машинного обучения, как видно из рисунка, позволило сократить число случайных состояний, необходимых для достижения результата, в среднем до 6800. При этом наилучшие результаты показывает метод случайного леса, что подтверждает проведенные оценки качества моделей. Использование данного метода в совокупности с ЛП_τ-последовательностями позволяет не вычислять режим с помощью затратных методов оптимизации на втором этапе для около 3000 состояний без серьезного влияния на точность оценки.

Заключение. Вычислительная эффективность методов, используемых в программных средствах оценки надёжности ЭЭС, является одним из критериев пригодности их для применения в практических задачах. При оценке надёжности ЭЭС методом Монте-Карло основным фактором, влияющим на вычислительную эффективность методики, является число анализируемых случайных состояний ЭЭС, необходимых для достижения определенной точности значения дефицита мощности. Сокращение числа таких состояний системы позволит повысить вычислительную эффективность расчётов в целом.

На первом этапе были проанализированы два способа генерации случайных состояний: Вихрь Мерсенна и ЛП_τ-последовательности без использования методов машинного обучения. Были проведены расчёты с использованием данных генераторов, а результаты визуализированы на графике (рис. 2). В итоге наиболее эффективным генератором случайных состояний для оценки надёжности ЭЭС оказался способ на основе ЛП_τ-последовательности.

На втором этапе были применены метод опорных векторов и метод случайного леса, проанализировано влияние их применения на эффективность расчёта. Результат этого анализа показал, что оба метода позволяют добиться сокращения числа анализируемых состояний, необходимых для определения дефицита мощности системы, при этом наилучшую эффективность показал метод случайного леса в совокупности с ЛП_τ-последовательностями.

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00234.

СПИСОК ЛИТЕРАТУРЫ

1. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. М: Наука. 1974. 416 с.
2. Воропай Н.И., Курбацкий В.Г., Томин Н.В., Панасецкий Д.А., Сидоров Д.Н., Жуков А.В., Ефимов Д.Н., Осак А.Б., Спиряев В.А., Домышев А.В. Комплекс интеллектуальных средств для предотвращения крупных аварий в электроэнергетических системах. Новосибирск: Наука. 2016. 332 с.

3. Кнут Д. Искусство программирования / Том 2. Получисленные алгоритмы. М: Вильямс. 2001.
 4. Ковалёв Г.Ф., Лебедева Л.М. Надежность систем электроэнергетики. Новосибирск: Наука. 2015. 224 с.
 5. Крупенёв Д.С., Бояркин Д.А., Якубовский Д.В. Формирование случайных состояний электроэнергетических систем при оценке их надежности методом статистических испытаний // Надежность и безопасность энергетики. 2017. Т. 10. № 1. С. 33–41.
 6. Соболев И.М. Метод Монте-Карло. М: Наука. 1968.
 7. Соболев И.М. Многомерные квадратурные формы и функции Хаара. М: Наука. 1969.
 8. Breiman L. Random forests // Machine Learning. 2001. Vol. 45. No. 1. Pp. 5–32.
 9. Krupenev D., Pershabinsky S. Algorithm for the adequacy discrete optimization by using dual estimates when planning the development of electric power systems // Proceeding of The 17th international scientific conference EPE. 2016.
 10. Matsumoto M., Nishimura T. Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator // ACM Transactions on Modeling and Computer Simulations. 1998. Vol. 8. Iss. 1. Pp. 3–30.
 11. Panasetky D., Tomin N., Voropai N. Development of software for modelling decentral-ized intelligent systems for security monitoring and control in power systems // IEEE Eindhoven PowerTech. 2015. Pp. 1–6.
-

UDK 004.942

**MACHINE LEARNING METHODS APPLYING FOR DEFICIENCY OF ELECTRICITY
SYSTEM POWER DETERMINATION**

Denis A. Boyarkin

Graduate student, National Research Irkutsk State Technical University (NR ISTU)
83, Lermontov Str., 664074, Irkutsk, Russia, e-mail: denisboyarkin@isem.irk.ru

Dmitriy S. Krupenev

Candidate of engineering sciences, senior researcher,
Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences
130, Lermontov Str., 664033, Irkutsk, Russia, e-mail: krupenev@isem.irk.ru

Dmitrii V. Yakubovskii

Graduate student, Melentiev Energy Systems Institute
Siberian Branch of the Russian Academy of Sciences
130, Lermontov Str., 664033, Irkutsk, Russia, e-mail: yakubovskii.dmit@mail.ru

Abstract. The article considers the question of increasing the computational efficiency of the methodology for electric power system adequacy assessment based on the Monte Carlo method. When using this method, the speed and accuracy of the calculation depends on the count of analyzed random states of the modeled system. Analysis means the solution of the flow distribution problem for each randomly generated state. This is a fairly time-consuming process, so reducing the number of analyzed states while maintaining the accuracy of the assessment will increase the speed of the calculation as

a whole. For this, it is proposed to use methods of machine learning, whose task is to determine the power system's deficit without optimization methods applying. In calculations, the support vector machine and random forest methods were used, the viability of the proposed approach was evaluated by solving adequacy assessment task of the test energy power system.

Keywords: electric power systems, adequacy assessment, Monte Carlo method, machine learning.

References

1. Vapnik V.N., Chervonenkis A.Y. Teoria raspoznavaniya obrazov [Theory of Pattern Recognition]. Moscow. Nauka. 1974. 416 p. (in Russian)
2. Voropai N.I., Kurbatskiy V.G., Tomin N.V., Panasetsky D.A., Sidorov D.N., Zhukov A.V., Efimov D.N., Osak A.B., Spiryaev V.A., Domyshev A.V. Kompleks intellektual'nyh sredstv dlya predotvrasheniya krupnyh avariyy v elektroenergeticheskikh sistemah [Intellectual Tools Complex for Prevention of Major Accidents in Electric Power Systems]. Novosibirsk. Nauka. 2016. 332 p. (in Russian)
3. Knuth D. The Art of Computer Programming. Vol. 2. Seminumerical Algorithms. Addison-Wesley. 1981. 688 p.
4. Kovalev G.F., Lebedeva L.M. Nadeshnost' sistem elektroenergetiki [Electric Power Systems Reliability]. Novosibirsk. Nauka. 2015. 224 p. (in Russian)
5. Krupenev D.S., Boyarkin D.A., Yakubovskiy D.V. [Generation of random states of electric power systems at assessment of their reliability by the Monte Carlo method]. Safety and Reliability of Power Industry [Nadeshnost' i bezopasnost' energetiki]. 2017. no. 10(1). Pp. 33–41. (in Russian)
6. Sobol I.M. Metod Monte-Karlo [Monte-Carlo method]. Moscow. Nauka. 1968. 64 p. (in Russian)
7. Sobol I.M. Mnogomernye kvadrurnye formy i funktsii Haara [Multidimensional Quadrature Formulas and Haar Functions]. Moscow. Nauka. 1969. 288 p. (in Russian)
8. Breiman L. Random forests // Machine Learning. 2001. Vol. 45. no. 1. Pp. 5–32.
9. Krupenev D.S., Pershabinsky S.M. Algorithm for the adequacy discrete optimization by using dual estimates when planning the development of electric power systems // Proceeding of The 17th international scientific conference EPE. 2016.
10. Matsumoto M., Nishimura T. Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator // ACM Transactions on Modeling and Computer Simulations. 1998. Vol. 8. Iss. 1. Pp. 3–30.
11. Panasetsky D.A., Tomin N.V., Voropai N.I. Development of software for modelling decentralized intelligent systems for security monitoring and control in power systems // IEEE Eindhoven PowerTech. 2015. Pp. 1–6.