



ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G06F 17/40 (2024.08); G06N 3/08 (2024.08); G06Q 50/06 (2024.08)

(21)(22) Заявка: 2024111616, 26.04.2024

(24) Дата начала отсчета срока действия патента:  
26.04.2024

Дата регистрации:  
07.10.2024

Приоритет(ы):

(22) Дата подачи заявки: 26.04.2024

(45) Опубликовано: 07.10.2024 Бюл. № 28

Адрес для переписки:

109147, Москва, ул. Марксистская, 4,  
Публичное акционерное общество "Мобильные  
Телесистемы"

(72) Автор(ы):

Ведяшкин Денис Вячеславович (RU),  
Алехин Алексей Николаевич (RU),  
Барбетов Алексей Игоревич (RU),  
Кузнецов Роман Алексеевич (RU),  
Шигаев Роман Евгеньевич (RU),  
Наумкин Роман Юрьевич (RU),  
Царапкин Игорь Владимирович (RU),  
Соломянюк Алексей Николаевич (RU),  
Дерябин Александр Юрьевич (RU)

(73) Патентообладатель(и):

Публичное акционерное общество  
"Мобильные Телесистемы" (RU)

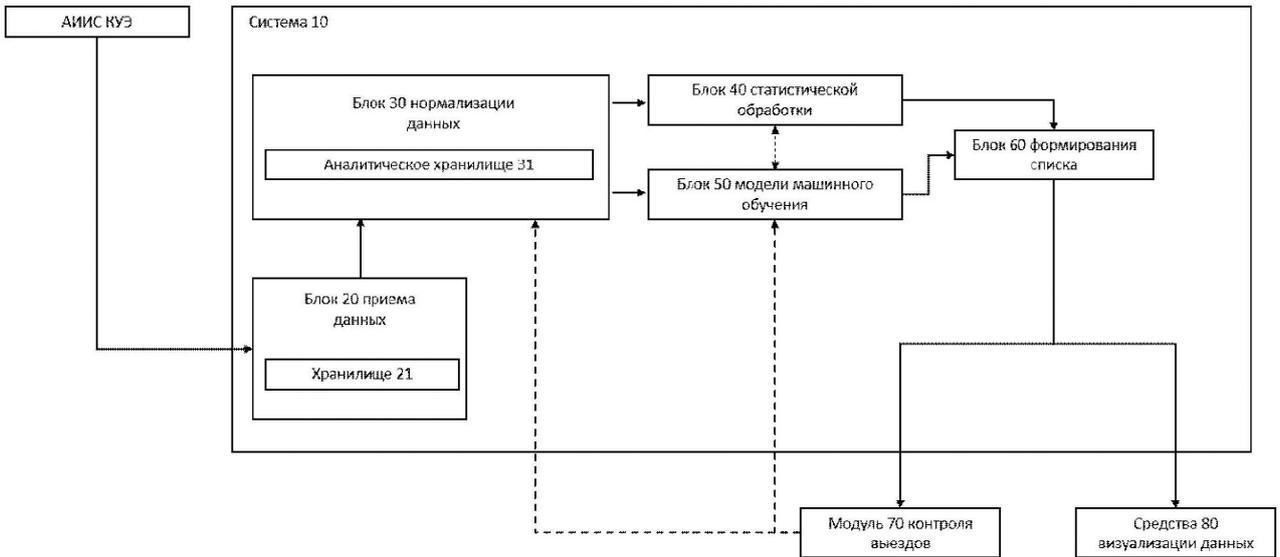
(56) Список документов, цитированных в отчете  
о поиске: US 2018/0299495 A1, 18.10.2018. US  
2019/0122132 A1, 25.04.2019. US 2020/0182653  
A1, 11.06.2020. US 2016/0161539 A1, 09.06.2016.  
RU 2314542 C2, 10.01.2008.

(54) СИСТЕМА АНАЛИЗА ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ, СООТВЕТСТВУЮЩИЙ СПОСОБ  
И МАШИНОЧИТАЕМЫЙ НОСИТЕЛЬ ДАННЫХ

(57) Реферат:

Изобретение относится к области вычислительной техники для учета и анализа потребляемой электроэнергии для выявления умышленных потерь в электрических сетях. Технический результат заключается в повышении точности обнаружения фактов хищения электроэнергии. Технический результат достигается за счет вычисления блоком статистической обработки метрик для каждого идентификатора абонента на основании временных рядов значений в заранее заданном интервале времени: метрика P1 среднего дневного потребления, метрика P2 среднего ночного потребления, метрика P3 соотношения дневного и ночного потребления, метрика P4 размаха значений дневного потребления, метрика P5 сезонной кратности; анализа блоком статистической обработки упомянутых

вычисленных метрик с присвоением идентификатору абонента статуса, соответствующего подозрительному потреблению, при обнаружении того, что значение упомянутых метрик заранее заданное количество раз попало в диапазон значений подозрительного потребления в течение заранее заданного интервала времени; анализа данных, принятых от блока статистической обработки, и/или нормализованных данных, принятых от блока нормализации данных, блоком модели машинного обучения путем решения задачи бинарной классификации абонентов в две группы: нормального потребления и подозрительного потребления, с присвоением идентификаторам абонентов из группы подозрительного потребления статуса, соответствующего подозрительному потреблению. 3 н. и 12 з.п. ф-



Фиг.1

RU 2828166 C1

RU 2828166 C1



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.  
*G06F 17/40* (2006.01)  
*G06N 3/08* (2006.01)  
*G06Q 50/06* (2012.01)

(12) **ABSTRACT OF INVENTION**

(52) CPC  
*G06F 17/40* (2024.08); *G06N 3/08* (2024.08); *G06Q 50/06* (2024.08)

(21)(22) Application: **2024111616, 26.04.2024**  
(24) Effective date for property rights:  
**26.04.2024**  
Registration date:  
**07.10.2024**  
Priority:  
(22) Date of filing: **26.04.2024**  
(45) Date of publication: **07.10.2024** Bull. № 28  
Mail address:  
**109147, Moskva, ul. Marksistskaya, 4, Publichnoe  
aktsionernoe obshchestvo "Mobilnye Telesistemy"**

(72) Inventor(s):  
**Vediashkin Denis Viacheslavovich (RU),  
Alekhin Aleksei Nikolaevich (RU),  
Barbetov Aleksei Igorevich (RU),  
Kuznetsov Roman Alekseevich (RU),  
Shigaev Roman Evgenevich (RU),  
Naumkin Roman Iurevich (RU),  
Tsarapkin Igor Vladimirovich (RU),  
Solomianiuk Aleksei Nikolaevich (RU),  
Deriabin Aleksandr Iurevich (RU)**  
(73) Proprietor(s):  
**Mobile TeleSystems Public Joint Stock  
Company (RU)**

(54) **ELECTRIC POWER CONSUMPTION ANALYSIS SYSTEM, CORRESPONDING METHOD AND COMPUTER-READABLE DATA MEDIUM**

(57) Abstract:

FIELD: computer engineering.

SUBSTANCE: invention relates to computer engineering for accounting and analysis of consumed electric power for detection of deliberate losses in electric networks. Technical result is achieved by calculating a statistical processing unit of metrics for each subscriber identifier based on time series of values in a predetermined time interval: metric P1 of average daily consumption, metric P2 of average night consumption, metric P3 of ratio of day and night consumption, metric P4 of range of values of day consumption, metric P5 of seasonal ratio; analysis by a statistical processing unit of said calculated metrics with assignment of a subscriber identifier a status corresponding to a suspicious consumption, when it is

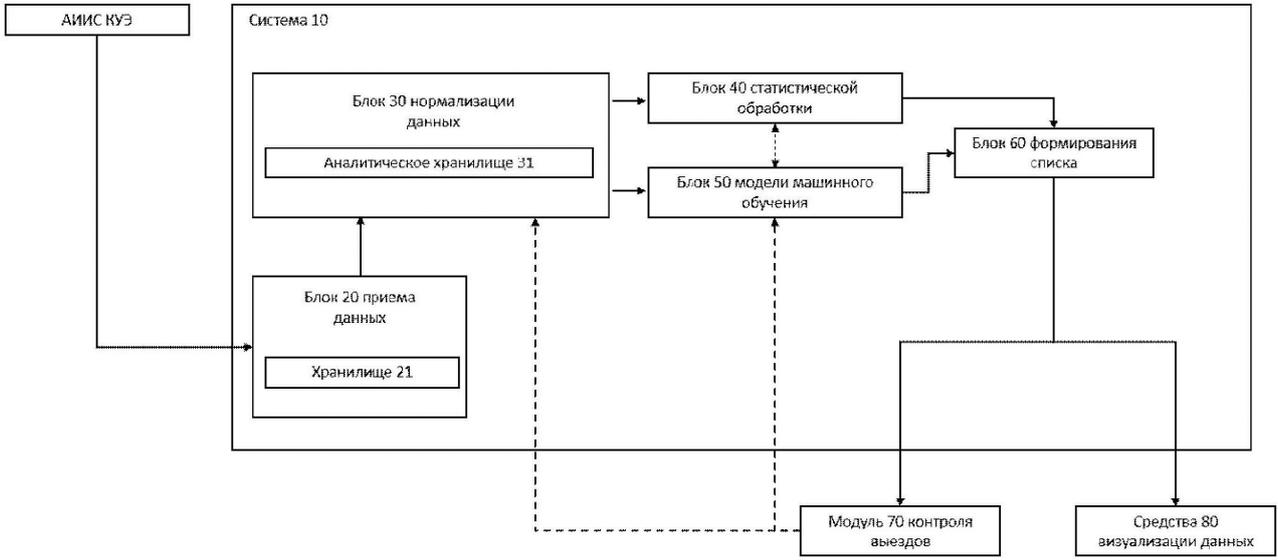
detected that the value of said metrics a predetermined number of times fell into the range of values of suspicious consumption during a predetermined time interval; analysis of data received from the statistical processing unit, and/or normalized data received from the data normalization unit, by the machine learning model unit by solving the problem of binary classification of subscribers into two groups: normal consumption and suspicious consumption, with assignment of identifiers of subscribers from the group of suspicious consumption of status corresponding to suspicious consumption.

EFFECT: higher accuracy of electric power theft detection.

15 cl, 2 dwg

RU 2 828 166 C1

RU 2 828 166 C1



Фиг.1

RU 2828166 C1

RU 2828166 C1

## ОБЛАСТЬ ТЕХНИКИ

[1] В целом изобретение относится к системам и способам учета и анализа потребляемой электроэнергии, и в частности изобретение относится к системам и способам для выявления потерь в электрических сетях.

## УРОВЕНЬ ТЕХНИКИ

[2] Потери в электрических сетях могут возникать как в процессе потребления, так и в процессе распределения электроэнергии. Они могут быть вызваны неумышленными факторами, такими как утечки, потери в линиях передачи и распределения, неэффективность оборудования и тому подобное, а также умышленными факторами, связанными с действиями потребителей, например, потребитель может использовать электроэнергию без оплаты в обход прибора учета, занижать заявленный объем потребления путем внесения изменений в прибор учета, использовать бытовой тариф для коммерческой деятельности и т.п. Устранение неумышленных факторов является для поставщиков электроэнергии привычной задачей, решение которой лежит в своевременной модернизации и обслуживании оборудования.

[3] Устранение умышленных факторов потерь является задачей, которая до сих пор не имеет очевидного и прямолинейного решения. Большое разнообразие видов несанкционированных действий при потреблении электроэнергии и отсутствие возможности физически контролировать деятельность каждого абонента обуславливает низкую эффективность существующих подходов к выявлению умышленных потерь при потреблении электроэнергии. В настоящее время применяется несколько разных подходов: анализ специалистом в данной области большого объема разнообразных данных о потребителях; системы, основанные на заранее заданных правилах, предоставляемые с умными приборами учета; системы, основанные на случайности, основанные на уведомлении компании поставщика электроэнергии другими потребителями об очевидных случаях несанкционированных действий (например, уведомление о том, что сосед очевидным образом занимается хищением электроэнергии).

[4] Известные подходы ненадежны, негибки, плохо масштабируемы и часто выдают ложно положительные заключения, что снижает доверие и лояльность добросовестных потребителей и приводит к потере времени и ресурсов на безрезультатную проверку.

[5] Таким образом, технической задачей, на решение которой направлено настоящее изобретение, является создание системы лишенной упомянутых недостатков.

## СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[6] Для решения упомянутой технической задачи предлагается программно-аппаратная платформа, выполненная с возможностью облачных вычислений комплекса технологий обработки и анализа больших данных с использованием методов машинного обучения моделей распознавания умышленных потерь в электрических сетях.

[7] В соответствии с первым аспектом изобретения предложена система для анализа потребления электроэнергии, включающая: блок приема данных, связанный с возможностью передачи данных с блоком нормализации, и выполненный с возможностью приема первичных данных от системы учета потребления электроэнергии, при этом первичные данные включают по меньшей мере идентификатор абонента, информацию о приборе учета абонента и информацию о потреблении электроэнергии для каждого абонента электрической сети. Упомянутая система также включает блок нормализации данных с аналитическим хранилищем, связанный с возможностью передачи данных с блоком статистической обработки и блоком модели машинного обучения, и выполненный с возможностью нормализации первичных данных с формированием по меньшей мере двух временных рядов значений потребления

электроэнергии для каждого идентификатора абонента, при этом по меньшей мере один из упомянутых временных рядов характеризует дневное потребление электроэнергии и по меньшей мере один из упомянутых временных рядов характеризует ночное потребление электроэнергии. Система также включает блок статистической обработки, связанный с возможностью передачи данных с блоком модели машинного обучения и блоком формирования списка, и выполненный с возможностью: вычисления для каждого идентификатора абонента в заранее заданном интервале времени на основании упомянутых временных рядов значений по меньшей мере следующих метрик: метрика P1 среднего дневного потребления, метрика P2 среднего ночного потребления, метрика P3 соотношения дневного и ночного потребления, метрика P4 размаха значений дневного потребления, метрика P5 сезонной кратности; анализа упомянутых вычисленных метрик с присвоением идентификатору абоненту статуса, соответствующего подозрительному потреблению, при обнаружении того, что значение по меньшей мере одной из упомянутых метрик заранее заданное количество раз попало в диапазон значений подозрительного потребления в течение заранее заданного интервала времени; направления идентификаторов абонентов, которым присвоен статус, соответствующий подозрительному потреблению, в блок формирования списка. Еще одним элементов упомянутой системы является блок модели машинного обучения, связанный с возможностью передачи данных с блоком формирования списка, и выполненный с возможностью: анализа данных, принятых от блока статистической обработки, и/или нормализованных данных, принятых от блока нормализации данных, путем решения задачи бинарной классификации абонентов в две группы: нормального потребления и подозрительного потребления, присвоения идентификаторам абонентов из группы подозрительного потребления статуса, соответствующего подозрительному потреблению направления идентификаторов абонентов, которым присвоен статус, соответствующий подозрительному потреблению, в блок формирования списка, который в свою очередь выполнен с возможностью формирования и сортировки списка идентификаторов абонентов, рекомендованных к проверке, на основании данных, принятых от блока статистической обработки и блока модели машинного обучения.

[8] В одном из вариантов выполнения блок статистической обработки может быть дополнительно выполнен с возможностью направления идентификаторов абонентов, которым присвоен статус, соответствующий подозрительному потреблению, в блок модели машинного обучения для проверки.

[9] В еще одном варианте выполнения блок статистической обработки дополнительно выполнен с возможностью присвоения идентификатору абонента статуса, соответствующего подозрительному потреблению, при обнаружении по меньшей мере одного из следующего: 1) попадания метрики P3 абонента заранее заданное количество раз в диапазон значений подозрительного потребления в течение заранее заданного интервала времени; 2) отсутствия изменения по существу метрики P3 на протяжении заранее заданного интервала времени, равенства по существу или стремления к единице метрики P5 на протяжении заранее заданного интервала времени и превышения математического ожидания эталонной выборки метриками P1 и P2 в два и более раз; 3) попадания метрики P2 в диапазон значений подозрительного потребления со значением ниже математического ожидания эталонной выборки, попадания метрики P3 в диапазон значений подозрительного потребления со значением выше математического ожидания эталонной выборки, попадания метрики P4 в диапазон значений подозрительного потребления со значением выше математического ожидания эталонной выборки, и попадания метрики P1 в диапазон значений подозрительного

потребления только в заранее заданные дни недели; 4) попадания любой из упомянутых метрик заранее заданное количество раз в диапазон значений подозрительного потребления в течение заранее заданного интервала времени после обнаружения информации о факте замены прибора учета абонента на основании упомянутой информации о приборе учета абонента; 5) изменения по меньшей мере двух из упомянутых метрик по меньшей мере на 50% в течение заранее заданного интервала времени.

[10] В еще одном варианте выполнения блок статистической обработки дополнительно выполнен с возможностью вычисления диапазона значений эталонного потребления и диапазона значений подозрительного потребления для каждой из метрик на основании исторических данных, сохраненных в аналитическом хранилище.

[11] В еще одном варианте выполнения в качестве обучающей выборки для модели машинного обучения используют исторические данные, в которых указаны результаты выездных проверок абонентов.

[12] В еще одном варианте выполнения блок приема данных выполнен с возможностью приема первичных данных от автоматизированной информационно-измерительной системы контроля и учёта (АИИС КУЭ), выполненной с возможностью сбора первичных данных от приборов учета электроэнергии.

[13] В еще одном варианте выполнения блок нормализации данных выполнен с возможностью поиска, удаления ошибок телеметрии и восстановления удаленных и/или утерянных данных посредством аппроксимации.

[14] В соответствии со вторым аспектом настоящего изобретения предложен способ анализа потребления электроэнергии, включающий: получение первичных данных блоком приема данных от системы учета потребления электроэнергии, при этом первичные данные включают по меньшей мере идентификатор абонента, информацию о приборе учета абонента и информацию о потреблении электроэнергии для каждого абонента электрической сети; нормализацию первичных данных блоком нормализации данных с формированием по меньшей мере двух временных рядов значений потребления электроэнергии для каждого идентификатора абонента, при этом по меньшей мере один из упомянутых временных рядов характеризует дневное потребление электроэнергии и по меньшей мере один из упомянутых временных рядов характеризует ночное потребление электроэнергии; вычисление блоком статистической обработки по меньшей мере следующих метрик для каждого идентификатора абонента на основании упомянутых временных рядов значений в заранее заданном интервале времени: метрика P1 среднего дневного потребления, метрика P2 среднего ночного потребления, метрика P3 соотношения дневного и ночного потребления, метрика P4 размаха значений дневного потребления, метрика P5 сезонной кратности; анализ блоком статистической обработки упомянутых вычисленных метрик с присвоением идентификатору абонента статуса, соответствующего подозрительному потреблению, при обнаружении того, что значение по меньшей мере одной из упомянутых метрик заранее заданное количество раз попало в диапазон значений подозрительного потребления в течение заранее заданного интервала времени; анализ данных, принятых от блока статистической обработки, и/или нормализованных данных, принятых от блока нормализации данных, блоком модели машинного обучения путем решения задачи бинарной классификации абонентов в две группы: нормального потребления и подозрительного потребления, с присвоением идентификаторам абонентов из группы подозрительного потребления статуса, соответствующего подозрительному потреблению; формирование и сортировка списка абонентов, рекомендованных к

проверке, блоком формирования списка на основании данных, полученных от блока статистической обработки и блока машинного обучения.

5 [15] В соответствии с третьим аспектом настоящего изобретения предложен машиночитаемый носитель данных содержащий программные инструкции, при исполнении которых компьютером обеспечивается выполнение упомянутого способа анализа потребления электроэнергии.

10 [16] Предложенное изобретение обеспечивает возможность выявления различных типов умышленных потерь электроэнергии, например, потребление электроэнергии в обход прибора учета, вмешательство в измерительные цепи тока или напряжения, другие аномалии в приборах учета и подключение юридических лиц по тарифам для населения. Предложенная система также обеспечивает повышение точности выявления несанкционированных действий при потреблении электроэнергии и снижение количества ложно положительных выявлений подозрительных потребителей. Кроме того, предлагаемая система является легко масштабируемой.

### 15 **КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ**

[17] Эти и другие признаки, аспекты и преимущества предложенного изобретения будут лучше понятны при прочтении следующего подробного описания, приведенного со ссылками на сопровождающие чертежи.

20 [18] На фиг.1 показана принципиальная схема работы системы для анализа потребления электроэнергии.

[19] На фиг.2 схематически изображена блок-схема способа анализа потребления электроэнергии.

### **ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ**

25 [20] Настоящее описание и раскрытые в нем варианты выполнения предложенного изобретения приведены исключительно в качестве примера, а не для ограничения изобретения.

30 [21] В представленном ниже подробном описании раскрыты признаки, приведенные для лучшего понимания различных вариантов выполнения предложенного изобретения. Некоторые признаки могут быть опциональными, при этом для специалиста в данной области техники будет понятно, что настоящее изобретение может быть осуществлено без конкретных указанных признаков или с использованием известных из уровня техники соответствующих эквивалентов. Следует также отметить, что подробное описание некоторых широко известных способов, процедур, компонентов и тому подобного было опущено для сохранения простоты и ясности изложения сущности заявленного изобретения.

35 [22] В целях описания предложенного изобретения понятие «умышленные потери» в энергосети следует понимать, как потери, связанные с ошибками или недобросовестностью в процессе измерения и учета потребления электроэнергии. Эти потери могут быть вызваны различными факторами, такими как неправильное считывание счетчиков, незаконное подключение к сетям, нечестность потребителей или производителей, а также ошибки в расчетах и отчетности. Умышленные потери могут быть также связаны с несанкционированными действиями потребителей, например, потребитель может потреблять электроэнергию в обход прибора учета, занижать заявленный объем потребления путем внесения изменений в прибор учета, использовать бытовой тариф для коммерческой деятельности и т.п. Умышленные потери также называют в уровне техники системными или несанкционированными для отделения от неизбежных потерь, называемых в уровне техники техническими и возникающих из-за физических процессов, таких как преобразование энергии и ее

передача по сетям, связанных с невозможностью достижения стопроцентной эффективности работы оборудования и систем.

5 [23] Как показано на фиг.1, предложенная система 10 для выявления умышленных потерь при распределении электроэнергии включает блок 20 приема данных с хранилищем 21 первичных данных, блок 30 нормализации данных с соответствующим аналитическим хранилищем 31, блок 40 статистической обработки, блок 50 модели машинного обучения, блок 60 формирования списка абонентов, подлежащих проверке. В некоторых вариантах выполнения система 10 может дополнительно включать средства связи с терминалами пользователя и/или администратора, а также с иными внешними  
10 программно-аппаратными комплексами.

[24] Система 10 может быть выполнена в виде совокупности технических средств и программного обеспечения, используемых для формирования, получения, хранения, обработки и передачи информации, а также технических средств, выполненных с возможностью взаимодействия с внешними системами мониторинга и управления  
15 обеспечением электроэнергией на основании упомянутой информации.

[25] Например, как показано на фиг.1, заявленная система 10 для выявления умышленных потерь может быть выполнена с возможностью взаимодействия с автоматизированной информационно-измерительной системой контроля и учёта (АИИС КУЭ) энергетической компании для получения первичных данных учета  
20 потребления электроэнергии, а также для отключения абонента или ограничения нагрузки на приборе учета до момента выяснения причин аномалий в получаемых данных. АИИС КУЭ может включать различные устройства и компоненты, такие как счетчики электроэнергии, концентраторы данных, маршрутизаторы, а также программное обеспечение для обработки и анализа информации сырых данных.

25 [26] Известные в уровне техники АИИС КУЭ позволяют автоматизировать сбор и учет данных о потреблении электроэнергии.

[27] Для взаимодействия с АИИС КУЭ, в частности, для сбора и извлечения первичных данных учета потребления электроэнергии система 10 включает блок 20 приема данных и хранилище 21 первичных данных.

30 [28] Первичные данные, принимаемые блоком 20 приема данных от АИИС КУЭ могут иметь, например, следующую структуру:

35

40

45

Поле	Тип данных	Описание
address_too	(nvarchar(255), null)	Адрес хозяйственной организации (Полный адрес индекс, улица, дом, кв и тд)
type_pu	(nvarchar(255), null)	Тип Прибора Учета (ПУ) (например, марка прибора)
phase_num	(nvarchar(255), null)	Количество фаз
kt	(nvarchar(255), null)	Коэффициент трансформации
factory_count_num	(nvarchar(255), null)	Заводской номер ПУ (Серийный номер)
accounting_affil	(nvarchar(255), null)	Принадлежность учета ПУ
accuracy_class	(nvarchar(255), null)	Класс точности ПУ
mpi_years	(nvarchar(255), null)	Межповерочный интервал ПУ
last_verific_date	(nvarchar(255), null)	Дата прошлой проверки ПУ
actual_verific_day	(nvarchar(255), null)	Дата текущей проверки ПУ
days_until_verific	(nvarchar(255), null)	Количество дней до следующей поверки ПУ
physic_voltage_level	(nvarchar(255), null)	Физический уровень напряжения
max_power_too	(int, null)	Максимальная мощность
date_of_acceptance_apfc	(datetime2(7), null)	Дата принятия точки учета к расчету

	fl_ul	(nvarchar(255), null)	Указание на тип потребителя (физическое/юридическое лицо)
5	whol_retail	(nvarchar(255), null)	Указание на тип тарифа (Оптовый/розничный рынок)
	e_e_volume_fixated_ pu	(float, null)	Объем электроэнергии, зафиксированный ПУ
	[dt]	(datetime, null)	Дата показаний
10	[N_Rate]	(int, null)	0 - Суммарное потребление электроэнергии за сутки 1 - Дневное потребление электроэнергии 2 - Ночное потребление электроэнергии
15	[Val]	(float, null)	Показания ПУ, кВт*ч
	[State]	(int, null)	Состояние ПУ
	[id_point]	(int, null)	Идентификатор ПУ в сети
20	[PointName]	(varchar(300), null)	Наименование ПУ
	[Direction]	(varchar(3), not null)	Направление (in ; out)
	[coef]	(float, not null)	Коэффициент трансформации
	[SN]	(varchar(50), null)	Серийный номер ПУ
25	[NameType]		Тип ПУ
	[PhoneNum]	(int, null)	Телефонный номер ПУ
	[NumUSD]	(int, null)	Номер адресации ПУ

30 [29] После приема данных блоком 20 приема данных, данные в автоматическом режиме могут быть направлены блоком 20 приема данных в блок 30 нормализации данных и аналитическое хранилище 31 для нормализации.

[30] Блок 20 приема данных также может быть выполнен с возможностью получения дополнительных данных об абоненте, помещении, в котором расположен прибор учета (например, его тип, площадь, год постройки, класс энергоэффективности и тому подобное) из других источников.

[31] Техническое несовершенство АИИС КУЭ и линий связи между АИИС КУЭ и блоком 20 приема данных может приводить к нерегулярному поступлению данных, а также к потере и/или искажению по меньшей мере части данных при передаче. Подобные 40 искажения данных могут существенно влиять на результаты работы системы 10 для выявления умышленных потерь, поэтому для устранения возможных искажений данные из блока 20 передают в блок 30 нормализации данных, выполненный с возможностью восстановления значений потребления электроэнергии.

[32] Одним из примеров подобных искажений могут быть ошибки телеметрии, 45 которые могут характеризоваться, например, пропусками значений показаний от прибора учета, либо значениями, которые искажают монотонно возрастающую последовательность измерений прибора учета. При этом в монотонно возрастающей последовательности измерений возможно множество вариантов ошибок телеметрии:

одиночные провалы, одиночные скачки, резкие скачки, смена прибора учета, периодические провалы и скачки и т.д.

[33] Блок 30 нормализации данных может быть выполнен с возможностью проверки данных в строгой хронологии и удаления строк из исходных данных, которые попадают под критерии ошибок.

[34] Этап поиска и удаления ошибок может быть повторен неоднократно.

[35] Удаленные в процессе поиска ошибок или пропущенные по другим причинам данные могут быть восстановлены блоком 30 нормализации данных. Значения могут быть восстановлены, например, кумулятивно от последнего непустого значения перед пропущенным диапазоном до первого непустого значения после. Данные могут быть восстановлены, например, путем аппроксимации по формуле:

$$\text{Val}_{\text{new}} = \text{Val}_{\text{min}} + (\text{Val}_{\text{max}} - \text{Val}_{\text{min}}) / \text{Cnt}_{\text{row}} * \text{Num}_{\text{row}}$$

где  $\text{Val}_{\text{new}}$  - новое расчетное значение;

$\text{Val}_{\text{min}}$  - последнее непустое значение перед пропущенным диапазоном;

$\text{Val}_{\text{max}}$  - первое непустое значения после пропущенного диапазона;

$\text{Cnt}_{\text{row}}$  - количество пропущенных строк диапазона;

$\text{Num}_{\text{row}}$  - порядковый номер строки диапазона.

[36] После аппроксимации получают монотонно возрастающие значения измерений по каждому абоненту за весь период нормализации. Могут быть сформированы среднесуточные показания как разница между последующим показанием и текущим. Самое последнее значение среднесуточного показания может быть указано отдельно как разница текущего и предыдущего показания.

[37] Нормализованные данные могут сохраняться блоком 30 нормализации данных в аналитическом хранилище 31 в виде, например, таблицы, в которой каждому идентификатору абонента соответствует своя последовательность значений показаний потребления электроэнергии.

[38] Аналитическое хранилище 31 может быть выполнено в виде базы данных любого типа, например, в виде программно-аппаратного комплекса, включающего по меньшей мере одно запоминающее устройство в виде, например, машиночитаемого носителя данных, и по меньшей мере один блок управления базой данных, выполняющий функции системы управления базой данных. Очевидно, что аналитическое хранилище 31 может также представлять собой сеть хранения данных (Storage Area Network, SAN), сетевое хранилище данных (Network Attached Storage, NAS), распределенную базу данных или любую их комбинацию.

[39] Нормализованные данные из блока 30 далее передают в блок 40 статистической обработки для вычисления метрик потребления для каждого абонента и опционально в блок 50 модели машинного обучения в качестве исходных данных.

[40] Далее, в блоке 40 статистической обработки вычисленные метрики потребления сравнивают со статистическими для присвоения оценок потребления абонентам.

[41] Блок 50 модели машинного обучения в одном варианте выполнения изобретения выполнен с возможностью сравнения нормализованных данных с данными абонентов, которые ранее были уличены в несанкционированных действиях при потреблении электроэнергии, например, в ходе выезда специалиста энергетической компании.

[42] Как упомянуто выше, блок 40 статистической обработки выполнен с возможностью вычисления метрик для каждого из абонентов. Примерами упомянутых метрик могут быть, не ограничиваясь перечисленным: среднее дневное потребление P1, среднее ночное потребление P2, соотношение P3 дневного и ночного потребления,

размах  $P_4$  значений дневного потребления, сезонная кратность  $P_5$ .

[43] В настоящем описании термин «метрика» следует понимать, как численное значение, характеризующее потребление электроэнергии абонентом в заранее заданном интервале времени. Среднее дневное потребление  $P_1$ , например, может вычисляться как среднее значение кВт/ч за 16 часов потребления электроэнергии абонентом, например, с 8 часов утра (08:00) до полуночи (00:00). Среднее ночное потребление  $P_2$ , например, может быть вычислено как среднее значение кВт/ч за 8 часов потребления электроэнергии абонентом, например, с полуночи (00:00) до восьми часов утра (08:00). Метрика  $P_3$  соотношения дневного и ночного потребления может представлять собой отношение уровня дневного потребления абонента к уровню ночного потребления и вычисляться по формуле  $P_3=P_1/P_2$ . Метрика  $P_4$  размаха значений дневного потребления может представлять собой разницу среднего из четырех максимальных и минимальных значений. Метрика  $P_5$  сезонной кратности может представлять собой отношение среднедневного потребления наиболее холодного месяца года, например, января, и наиболее теплого месяца года, например, июня. Альтернативно метрика  $P_5$  сезонной кратности может представлять собой отношение среднедневного потребления в месяце с наиболее продолжительной долготой дня (световым днем) и в месяце с наименее продолжительной долготой дня (световым днем).

[44] Как упомянуто выше, блок 40 статистической обработки выполнен с возможностью сравнения, после вычисления метрик для каждого из абонентов, метрик абонента со статистическими значениями этих метрик с присвоением оценок метрикам потребления для этого абонента.

[45] Для этого блок 40 статистической обработки может быть выполнен с возможностью определения двух диапазонов значений для каждой из упомянутых метрик: диапазон значений для эталонного потребления и диапазон значений для подозрительного потребления. Упомянутые диапазоны могут быть определены, например, на основании накопленных исторических данных, включающих значения метрик всех абонентов в прошедший период.

[46] Исторические значения метрики  $P_i$ , хранимые в аналитическом хранилище 31 могут быть поданы на вход блока 40 статистической обработки для анализа. В ходе анализа блок 40 статистической обработки осуществляет оценку математического ожидания  $m_0$  и дисперсии  $\sigma_0$  значений метрики  $P_i$  для выборки, включающей абсолютно всех абонентов энергосети, которую в рамках настоящего описания можно назвать «общей».

[47] После этого определяют заранее заданный доверительный интервал метрики  $P_i$  для выбранного распределения, например, для нормального распределения. Далее блок 40 статистической обработки может осуществлять фильтрацию и выделение абонентов, значения метрики  $P_i$  которых находятся вне заранее заданного доверительного интервала, в отдельную выборку, которую можно назвать «подозрительной». Для подозрительной выборки также осуществляется оценка математического ожидания  $m_2$  и дисперсии  $\sigma_2$ .

[48] Выборку, полученную путем исключения подозрительной выборки из общей выборки, в рамках настоящего описания можно назвать «эталонной». Для эталонной выборки также осуществляют расчет математического ожидания  $m_1$  и дисперсии  $\sigma_1$ .

[49] Далее может быть выбран интересующий интервал времени  $[t_1, t_2]$  для выполнения анализа метрики  $P_i$  и рассчитан критерий Стьюдента (t-критерий) для сравнения

математического ожидания  $m_1$  эталонной выборки и математического ожидания  $m_2$  подозрительной выборки.

[50] Для параметров  $m$  и  $\sigma$  может быть построена функция плотности распределения случайной величины, например, по нормальному закону (функция Гаусса):

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\left(\frac{(x-m)^2}{2\sigma^2}\right)}$$

[51] Наибольшее значение  $f(x)$  принимает при  $x=m$ . При удалении от  $m$  график функции будет неограниченно приближаться к оси  $Ox$ . Площадь под графиком функции  $f(x)$  будет равна вероятности попадания значений в конкретный интервал. Авторы настоящего изобретения в ходе испытаний пришли к тому, что 95%-ый доверительный интервал обеспечивает достаточную точность анализа.

[52] Поскольку площадь под кривой вычисляется через интеграл, на практике для определения вероятности может использоваться интегральная функция распределения:

$$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^x e^{-\left(\frac{(x-m)^2}{2\sigma^2}\right)} dx$$

[53] Для каждой выборки вычисляются свои значения  $m$  и  $\sigma$ . Далее средние значения (математические ожидания) сравниваются по критерию Стьюдента для определения статистической значимости их различий. Вычисляется эмпирическое значение критерия Стьюдента

$$t_{\text{эмп.}} = \frac{\bar{x} - \bar{y}}{(n_1 - 1)s_x^2 + (n_2 - 1)s_y^2} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

где  $\bar{x}$ ,  $\bar{y}$  - выборочные средние;  $n_1$  и  $n_2$  - объемы выборок.

[54] Число степеней свободы  $k = n_1 + n_2 - 2$ , уровень значимости  $\alpha = 0,05$ . По таблице распределения Стьюдента для заданных значений  $\alpha$  и  $k$  может быть найден  $t_{\text{крит.}}$

[55] Если  $t_{\text{эмп.}} > t_{\text{крит.}}$ , то рассматриваемые выборки принадлежат разным генеральным совокупностям, и мы имеем значимые статистические различия средних значений. Поэтому данные, выходящие за пределы 95%-го интервала в процессе наблюдений, требуют пристального внимания и углубленного анализа.

[56] Полученные в результате анализа исторических данных диапазоны эталонного потребления (95%-ый доверительный интервал) и подозрительного потребления (5%-ый доверительный интервал) для каждой из упомянутых метрик могут быть использованы блоком 40 статистической обработки отдельно или совместно с другими заранее заданными правилами для дифференцирования анализируемых абонентов по группам в зависимости от типа потребления. В частности, блок 40 статистической обработки может помещать абонентов в группу, соответствующую подозрительному потреблению, путем присвоения соответствующим абонентам статуса или флага, соответствующего подозрительному потреблению, в аналитическом хранилище 31.

[57] Блок 40 статистической обработки может быть выполнен с возможностью присвоения абоненту статуса, соответствующего подозрительному потреблению, при экстремальном отклонении от вычисленного или заранее заданного «нормального» соотношения  $P3$  дневного и ночного показаний потребления. Как описано выше, блок 40 статистической обработки может определять «нормальное» соотношение  $P3$  путем

анализа исторических данных и определения диапазона значений для эталонного потребления, и диапазона значений для подозрительного потребления. При заранее заданном количестве попаданий текущих значений метрики P3 абонента в диапазон значений подозрительного потребления в течение заранее заданного периода времени, например, в течение трех месяцев, блок 40 статистической обработки может присвоить этому абоненту статус, соответствующий подозрительному потреблению.

[58] Блок 40 статистической обработки также может быть выполнен с возможностью присвоения абоненту статуса, соответствующего подозрительному потреблению, при кратном завышении и постоянном потреблении в дневное и ночное время, в частности статус может быть присвоен абоненту, потребление которого соответствует следующим условиям: 1) постоянный уровень P3 соотношения дневного и ночного показаний потребления; 2) кратно завышенный уровень потребления в дневное и ночное время; 3) отсутствие сезонных колебаний. Блок 40 статистической обработки может быть выполнен с возможностью выделения абонентов, текущие значения метрики соотношения P3 дневного и ночного потребления которых по существу не изменяются на протяжении заранее заданного интервала времени, например, в течение 6 месяцев. Блок 40 статистической обработки может быть также выполнен с возможностью выделения абонентов, значения метрики сезонная кратность P5 которых по существу равны или стремятся к 1 на протяжении заранее заданного интервала времени, например, в течение 6 месяцев. Блок 40 статистической обработки также может быть выполнен с возможностью выделения абонентов, значения метрик среднего дневного потребления P1 и среднего ночного потребления P2 которых в два и более раз (более предпочтительно в 4 и более раз) превышают математическое ожидание эталонной выборки, определенное блоком 40 статистической обработки путем анализа исторических данных, в течение, например, 6 месяцев. Абонентам, потребление электроэнергии которых соответствует упомянутым условиям, может быть присвоен статус, соответствующий подозрительному потреблению.

[59] Блок 40 статистической обработки также может быть выполнен с возможностью выявления абонентов, паттерны потребления которых характерны для ведения коммерческой деятельности. В частности, статус «подозрительного потребления» может быть присвоен абоненту, потребление которого соответствует следующим условиям: 1) заниженный уровень метрик показаний ночного потребления P2; 2) высокий уровень соотношения P3 дневного и ночного показаний потребления; 3) повышенный уровень оценки «размаха» P4; 4) пики потребления приходятся на «рабочие» дни. Как описано выше, блок 40 статистической обработки может определять «нормальное» значение метрики среднее ночное потребление P2 путем анализа исторических данных и определения диапазона значений для эталонного потребления, и диапазона значений для подозрительного потребления. Кроме того, блок 40 статистической обработки может сравнивать текущие показания абонента с математическим ожиданием эталонной выборки и выделять абонента, если его текущие значения метрики P2 попадают в диапазон значений подозрительного потребления и ниже математического ожидания эталонной выборки. Аналогично, блок 40 статистической обработки также может определять диапазон значений для эталонного потребления и диапазон значений для подозрительного потребления для метрики соотношение P3 дневного и ночного потребления. При этом блок 40 статистической обработки может выделять абонента, если его текущие значения метрики P3 попадают в диапазон значений подозрительного потребления и выше математического ожидания эталонной выборки. Аналогично блок 40 статистической обработки может также определять диапазон значений для эталонного

потребления и диапазон значений для подозрительного потребления для метрики размах Р4 значений дневного потребления. При этом блок 40 статистической обработки может выделять абонента, если его текущие значения метрики Р4 попадают в диапазон значений

5 Блок 40 статистической обработки может также выделять абонента, в случае попадания его текущих значений среднего дневного потребления Р1 в диапазон значений подозрительного потребления только в заранее заданные дни недели. Абонентам, потребление электроэнергии которых соответствует упомянутым условиям, может быть присвоен статус, соответствующий подозрительному потреблению.

10 [60] Блок 40 статистической обработки также может быть выполнен с возможностью определения факта замены прибора учета абонента, например, на основании того, что значения поля Val были обнулены или, например, изменилось значение в одном или более из полей last\_verific\_date, actual\_verific\_day, days\_until\_verific, State, id\_point, PointName, SN, NumUSD. При этом абоненту может быть присвоен статус,

15 соответствующий подозрительному потреблению, если было зафиксировано заранее заданное количество попаданий значений по меньшей мере одной из упомянутых метрик Р<sub>i</sub> в диапазон значений для подозрительного потребления в течение заранее заданного интервала времени после определения факта замены прибора учета этого абонента, например, в течение трех месяцев после замены.

20 [61] Блок 40 статистической обработки также может быть выполнен с возможностью присвоения абоненту статуса, соответствующего подозрительному потреблению, в случае фиксации пропорционального изменения значений по меньшей мере двух упомянутых метрик Р<sub>i</sub> на 50 и более процентов.

25 [62] Под выделением абонента в рамках настоящего описания следует понимать, например, помещение идентификатора абонента в соответствующий список в базе данных. В еще одном варианте выполнения выделение абонента может быть выполнено путем установки соответствующего флага в специализированной ячейке базы данных, связанной с идентификатором абонента. В рамках настоящего изобретения следует отличать операцию выделения абонента от операции присвоения абоненту статуса,

30 соответствующего подозрительному потреблению. Операция выделения абонента осуществляется для указания на то, что одна из метрик потребления абонента является аномальной. При этом для присвоения «подозрительного» статуса в большинстве случаев необходимо, чтобы потребление абонента характеризовалось совокупностью аномальных метрик для уменьшения ложноположительных заключений.

35 [63] Идентификаторы абонентов, которым присвоен статус, соответствующий подозрительному потреблению, могут направляться блоком 40 статистической обработки в блок 60 формирования списка абонентов, подлежащих проверке, и опционально в блок 50 модели машинного обучения. Кроме того, совместно с упомянутыми идентификаторами блок 40 статистической обработки может направлять

40 численные представления совокупных степеней отклонений метрик потребления абонентов.

[64] Как указано выше, система 10 также включает блок 50 модели машинного обучения.

45 [65] Модель машинного обучения может быть основана на известных в уровне техники моделях, которые могут быть обучены под задачу бинарной классификации и выдавать однозначные решения по каждому абоненту, например – взломан прибор учета абонента или нет.

[66] Альтернативно, модель машинного обучения может быть выполнена с

возможностью определения вероятности умышленных потерь по каждому абоненту, например – вероятность взлома прибора учета абонента равняется 0.8 (80%), что подчеркивает высокую степень риска.

5 [67] Для обучения модели могут быть использованы исторические данные, накапливаемые в аналитическом хранилище 31 системы 10, в котором могут фиксироваться данные о результатах выездных проверок абонентов. Эти данные играют ключевую роль в процессе обучения, позволяя модели адаптироваться к изменяющимся условиям и повышать свою точность. Модель машинного обучения может быть  
10 выполнена с возможностью предсказания вероятности мошеннических действий абонента на основании исторических данных электропотребления. Это позволяет оперативно выявлять потенциальные угрозы и принимать меры по их предотвращению, что является критически важным для обеспечения безопасности и надежности системы.

[68] Блок 50 модели машинного обучения выполнен с возможностью приема данных от блока 30 нормализации данных в качестве входных данных. После приема, данные  
15 могут быть подвергнуты предварительной очистке и подготовке: увеличение исходной выборки (аугментации), путем небольшого изменения исходных данных, создание и отбор наиболее полезных признаков при помощи статистических методов, полученных из исходных данных и, в конечном итоге, обучение на их основе модели машинного обучения. Это позволяет улучшить качество входных данных и повысить эффективность  
20 работы модели. На выходе модели машинного обучения блок 50 модели машинного обучения может получать скор модели от 0.0 до 1.0, отражающий вероятность умышленных потерь, например, вероятность взлома прибора учета. Чем выше значение, тем больше вероятность того, что потери связаны с умышленными действиями абонента.

[69] Для постоянного повышения точности модели блок 50 модели машинного  
25 обучения может быть выполнен с возможностью переобучения модели при получении новых данных от АИИС КУЭ и/или от модуля 70 контроля выездов, выполненного с возможностью сбора данных по результатам выездных проверок абонентов, что позволяет постоянно повышать точность модели, адаптируя ее к новым условиям и данным, что является ключевым для поддержания высокого качества работы.

30 [70] Модуль 70 контроля выездов может быть связан с возможностью передачи данных с аналитическим хранилищем 31 блока 30 нормализации данных и выполнен с возможностью внесения изменений в записи, хранящиеся в аналитическом хранилище 31. В частности, модуль 70 контроля выездов может обеспечивать возможность указания нового соответствующего статуса для идентификаторов пользователей, для которых  
35 был подтвержден или опровергнут статус подозрительного потребления в результате выездной проверки по месту установки прибора учета абонента.

[71] Блок 50 модели машинного обучения выполнен с возможностью направления идентификаторов абонентов, скор которых выше заранее заданного порога, в блок 60  
формирования списка абонентов, подлежащих проверке.

40 [72] В одном варианте выполнения изобретения блок 50 модели машинного обучения дополнительно выполнен с возможностью приема данных от блока 40 статистической обработки в качестве входных данных. Таким образом, в данном варианте выполнения блок 50 модели машинного обучения осуществляет проверку и фильтрацию результатов работы блока 40 статистической обработки и направляет идентификаторы абонентов,  
45 скор которых выше заранее заданного порогового значения, в блок 60 формирования списка абонентов, подлежащих проверке.

[73] Блок 60 формирования списка абонентов, подлежащих проверке, выполнен с возможностью приема выходных данных из блока 40 статистической обработки и

блока 50 модели машинного обучения и формирования списка абонентов, подлежащих проверке, на основе принятых данных. Упомянутый список абонентов может включать все идентификаторы абонентов, принятые от блоков 40 и 50. Альтернативно, блок 60 формирования списка абонентов может быть выполнен с возможностью сравнения 5 списков идентификаторов абонентов, принятых от блока 40 и блока 50, и включения в список абонентов, подлежащих проверке, только тех идентификаторов, которые указаны и списке, принятом от блока 40, и в списке, принятом от блока 50.

[74] Блок 60 формирования списка абонентов, подлежащих проверке, может быть выполнен с возможностью сортировки идентификаторов абонентов в соответствии с 10 заранее заданными критериями. Например, идентификаторы абонентов могут быть отсортированы по убыванию численного представления совокупной степени отклонения метрик потребления.

[75] Блок 60 формирования списка абонентов может быть также выполнен с возможностью направления сформированного списка во внешние средства 80 15 визуализации данных для представления пользователю, например, в узел визуализации данных для представления с помощью инструментов визуализации в личном кабинете пользователя, например, выполненном в виде страницы в сети Интернет.

[76] В соответствии с еще одним аспектом изобретения предложен способ анализа 20 потребления электроэнергии, осуществляемый описанной выше системой 10. Как показано на фиг.2, статистическая обработка данных, включающая операции, выполняемые блоком 40 статистической обработки данных, может быть выполнена параллельно обработке посредством блока 50 модели машинного обучения. При этом список абонентов, рекомендованных к проверке, формируется блоком 60 формирования списка на основании идентификаторов абонентов, полученных и от блока 40 25 статистической обработки и от блока 50 модели машинного обучения.

[77] В одном варианте выполнения изобретения, модель машинного обучения может быть применена к множеству идентификаторов абонентов, получаемых от блока 40 30 статистической обработки. Таким образом, блок 50 модели машинного обучения может быть использован для анализа гипотез о подозрительном потреблении, сформированных при статистической обработке блоком 40. При этом при формировании списка абонентов, рекомендованных к проверке, могут быть учтены результаты упомянутого анализа, например, путем исключения идентификаторов абонентов, статус 35 подозрительного потребления которых, не был подтвержден моделью. В другом варианте осуществления упомянутые результаты анализа могут быть использованы при сортировке списка идентификаторов абонентов.

[78] Один или более этапов работы системы 10, осуществляемых блоками упомянутой системы, могут быть описаны в виде программных инструкций и храниться в 40 машиночитаемом носителе. Машиночитаемый носитель данных может включать энергонезависимые носители и энергозависимые носители. Энергозависимые носители могут включать, среди прочего, полупроводниковые запоминающие устройства и динамические запоминающие устройства. Энергонезависимые носители могут включать, среди прочего, оптические диски и магнитные диски. При этом при исполнении хранимых на упомянутом машиночитаемом носителе программных инструкций процессором, 45 вычислительным устройством или распределенным вычислительным устройством, обеспечивается выполнение упомянутых этапов работы описанной выше системы анализа потребления электроэнергии.

[79] На представленной на фиг.1 блок-схеме изображенные функциональные блоки не обязательно указывают на разделение между схемами аппаратных средств. Таким

образом, один или более функциональных блоков могут быть реализованы, например, в едином комплексе программно-аппаратных средств (например, процессоре для обработки сигналов общего назначения, микроконтроллере, оперативном запоминающем устройстве, жестком диске и т.п.). Функциональность одного или более 5 блоков может быть реализована, например, с помощью машиночитаемого носителя, на котором записан программный код, при выполнении которого процессором компьютер выполняет соответствующие функции. Упомянутые в настоящем описании программы могут быть автономными программами, могут быть включены как подпрограммы в операционную систему, могут быть функциями в инсталлированном 10 пакете программ и тому подобным.

[80] Используемый в настоящем описании элемент, изложенный в единственном числе, не должен пониматься как исключающий множественное число упомянутых элементов, если только такое исключение не указано явно. Кроме того, ссылки на "один вариант выполнения" предложенного изобретения не должны быть интерпретированы 15 как исключающие существование дополнительных вариантов выполнения, которые также включают указанные отличительные признаки. Более того, если обратное не указано явным образом, варианты выполнения изобретения "включающие" или "содержащие" элемент или множество элементов, имеющих конкретное свойство, могут дополнительно включать такие элементы, которые не имеют этого свойства.

20

#### (57) Формула изобретения

1. Система (10) для анализа потребления электроэнергии, включающая:

блок (20) приема данных, связанный с возможностью передачи данных с блоком (30) нормализации и выполненный с возможностью приема первичных данных от 25 системы учета потребления электроэнергии, при этом первичные данные включают по меньшей мере идентификатор абонента, информацию о приборе учета абонента и информацию о потреблении электроэнергии для каждого абонента электрической сети;

блок (30) нормализации данных с аналитическим хранилищем (31), связанный с 30 возможностью передачи данных с блоком (40) статистической обработки и блоком (50) модели машинного обучения и выполненный с возможностью нормализации первичных данных с формированием по меньшей мере двух временных рядов значений потребления электроэнергии для каждого идентификатора абонента, при этом по меньшей мере один из упомянутых временных рядов характеризует дневное потребление 35 электроэнергии и по меньшей мере один из упомянутых временных рядов характеризует ночное потребление электроэнергии;

блок (40) статистической обработки, связанный с возможностью передачи данных с блоком (50) модели машинного обучения и блоком (60) формирования списка и выполненный с возможностью:

40 вычисления для каждого идентификатора абонента в заранее заданном интервале времени на основании упомянутых временных рядов значений по меньшей мере следующих метрик: метрика P1 среднего дневного потребления, метрика P2 среднего ночного потребления, метрика P3 соотношения дневного и ночного потребления, метрика P4 размаха значений дневного потребления, метрика P5 сезонной кратности;

45 анализа упомянутых вычисленных метрик с присвоением идентификатору абонента статуса, соответствующего подозрительному потреблению, при обнаружении того, что значение по меньшей мере одной из упомянутых метрик заранее заданное количество раз попало в диапазон значений подозрительного потребления в течение заранее заданного интервала времени;

направления идентификаторов абонентов, которым присвоен статус, соответствующий подозрительному потреблению, в блок (60) формирования списка; блок (50) модели машинного обучения, связанный с возможностью передачи данных с блоком (60) формирования списка и выполненный с возможностью:

5 анализа данных, принятых от блока (40) статистической обработки, и/или нормализованных данных, принятых от блока (30) нормализации данных, путем решения задачи бинарной классификации абонентов в две группы: нормального потребления и подозрительного потребления,

10 присвоения идентификаторам абонентов из группы подозрительного потребления статуса, соответствующего подозрительному потреблению;

направления идентификаторов абонентов, которым присвоен статус, соответствующий подозрительному потреблению, в блок (60) формирования списка;

15 блок (60) формирования списка, выполненный с возможностью формирования и сортировки списка идентификаторов абонентов, рекомендованных к проверке, на основании данных, принятых от блока (40) статистической обработки и блока (50) модели машинного обучения.

2. Система по п.1, в которой блок (40) статистической обработки дополнительно выполнен с возможностью направления идентификаторов абонентов, которым присвоен статус, соответствующий подозрительному потреблению, в блок (50) модели машинного обучения для проверки.

3. Система по п.1, в которой блок (40) статистической обработки дополнительно выполнен с возможностью присвоения идентификатору абонента статуса, соответствующего подозрительному потреблению, при обнаружении по меньшей мере одного из следующего:

25 1) попадания метрики P3 абонента заранее заданное количество раз в диапазон значений подозрительного потребления в течение заранее заданного интервала времени;

30 2) отсутствия изменения по существу метрики P3 на протяжении заранее заданного интервала времени, равенства по существу или стремления к единице метрики P5 на протяжении заранее заданного интервала времени и превышения математического ожидания эталонной выборки метриками P1 и P2 в два и более раз;

35 3) попадания метрики P2 в диапазон значений подозрительного потребления со значением ниже математического ожидания эталонной выборки, попадания метрики P3 в диапазон значений подозрительного потребления со значением выше математического ожидания эталонной выборки, попадания метрики P4 в диапазон значений подозрительного потребления со значением выше математического ожидания эталонной выборки и попадания метрики P1 в диапазон значений подозрительного потребления только в заранее заданные дни недели;

40 4) попадания любой из упомянутых метрик заранее заданное количество раз в диапазон значений подозрительного потребления в течение заранее заданного интервала времени после обнаружения информации о факте замены прибора учета абонента на основании упомянутой информации о приборе учета абонента;

5) изменения по меньшей мере двух из упомянутых метрик по меньшей мере на 50% в течение заранее заданного интервала времени.

45 4. Система по любому из предшествующих пунктов, в которой блок (40) статистической обработки дополнительно выполнен с возможностью вычисления диапазона значений эталонного потребления и диапазона значений подозрительного потребления для каждой из метрик на основании исторических данных, сохраненных в аналитическом хранилище (31).

5. Система по п.1, в которой в качестве обучающей выборки для модели машинного обучения используют исторические данные, в которых указаны результаты выездных проверок абонентов.

6. Система по п.1, в которой блок (20) приема данных выполнен с возможностью приема первичных данных от автоматизированной информационно-измерительной системы контроля и учёта (АИИС КУЭ), выполненной с возможностью сбора первичных данных от приборов учета электроэнергии.

7. Система по п.1, в которой блок (30) нормализации данных выполнен с возможностью поиска, удаления ошибок телеметрии и восстановления удаленных и/или утерянных данных посредством аппроксимации.

8. Способ анализа потребления электроэнергии, включающий: получение первичных данных блоком (20) приема данных от системы учета потребления электроэнергии, при этом первичные данные включают по меньшей мере идентификатор абонента, информацию о приборе учета абонента и информацию о потреблении электроэнергии для каждого абонента электрической сети;

нормализацию первичных данных блоком (30) нормализации данных с формированием по меньшей мере двух временных рядов значений потребления электроэнергии для каждого идентификатора абонента, при этом по меньшей мере один из упомянутых временных рядов характеризует дневное потребление электроэнергии и по меньшей мере один из упомянутых временных рядов характеризует ночное потребление электроэнергии;

вычисление блоком (40) статистической обработки по меньшей мере следующих метрик для каждого идентификатора абонента на основании упомянутых временных рядов значений в заранее заданном интервале времени: метрика P1 среднего дневного потребления, метрика P2 среднего ночного потребления, метрика P3 соотношения дневного и ночного потребления, метрика P4 размаха значений дневного потребления, метрика P5 сезонной кратности;

анализ блоком (40) статистической обработки упомянутых вычисленных метрик с присвоением идентификатору абонента статуса, соответствующего подозрительному потреблению, при обнаружении того, что значение по меньшей мере одной из упомянутых метрик заранее заданное количество раз попало в диапазон значений подозрительного потребления в течение заранее заданного интервала времени;

анализ данных, принятых от блока (40) статистической обработки, и/или нормализованных данных, принятых от блока (30) нормализации данных, блоком (50) модели машинного обучения путем решения задачи бинарной классификации абонентов в две группы: нормального потребления и подозрительного потребления, с присвоением идентификаторам абонентов из группы подозрительного потребления статуса, соответствующего подозрительному потреблению;

формирование и сортировка списка абонентов, рекомендованных к проверке, блоком (60) формирования списка на основании данных, полученных от блока (40) статистической обработки и блока (50) машинного обучения.

9. Способ по п.8, в котором множество идентификаторов абонентов, которым блоком (40) статистической обработки присвоен статус, соответствующий подозрительному потреблению, направляют в блок (50) машинного обучения для проверки перед направлением в блок (60) формирования списка.

10. Способ по п.8, в котором при анализе блоком (40) статистической обработки упомянутых вычисленных метрик и упомянутых первичных данных идентификатору абонента также присваивают статус, соответствующий подозрительному потреблению,

при обнаружении по меньшей мере одного из следующего:

1) попадания метрики РЗ абонента заранее заданное количество раз в диапазон значений подозрительного потребления в течение заранее заданного интервала времени;

2) отсутствия изменения по существу метрики РЗ на протяжении заранее заданного интервала времени, равенства по существу или стремления к единице метрики Р5 на протяжении заранее заданного интервала времени и превышения математического ожидания эталонной выборки метриками Р1 и Р2 в два и более раз;

3) попадания метрики Р2 в диапазон значений подозрительного потребления со значением ниже математического ожидания эталонной выборки, попадания метрики РЗ в диапазон значений подозрительного потребления со значением выше математического ожидания эталонной выборки, попадания метрики Р4 в диапазон значений подозрительного потребления со значением выше математического ожидания эталонной выборки и попадания метрики Р1 в диапазон значений подозрительного потребления только в заранее заданные дни недели;

4) попадания любой из упомянутых метрик заранее заданное количество раз в диапазон значений подозрительного потребления в течение заранее заданного интервала времени после обнаружения информации о факте замены прибора учета абонента на основании упомянутой информации о приборе учета абонента;

5) изменения по меньшей мере двух из упомянутых метрик по меньшей мере на 50% в течение заранее заданного интервала времени;

11. Способ по любому из предшествующих пунктов, в котором диапазон значений эталонного потребления и диапазон значений подозрительного потребления для каждой из метрик вычисляется блоком (40) статистической обработки на основании исторических данных, сохраненных в аналитическом хранилище (31).

12. Способ по п.8, в котором в качестве обучающей выборки для модели машинного обучения используют исторические данные, в которых указаны результаты выездных проверок абонентов.

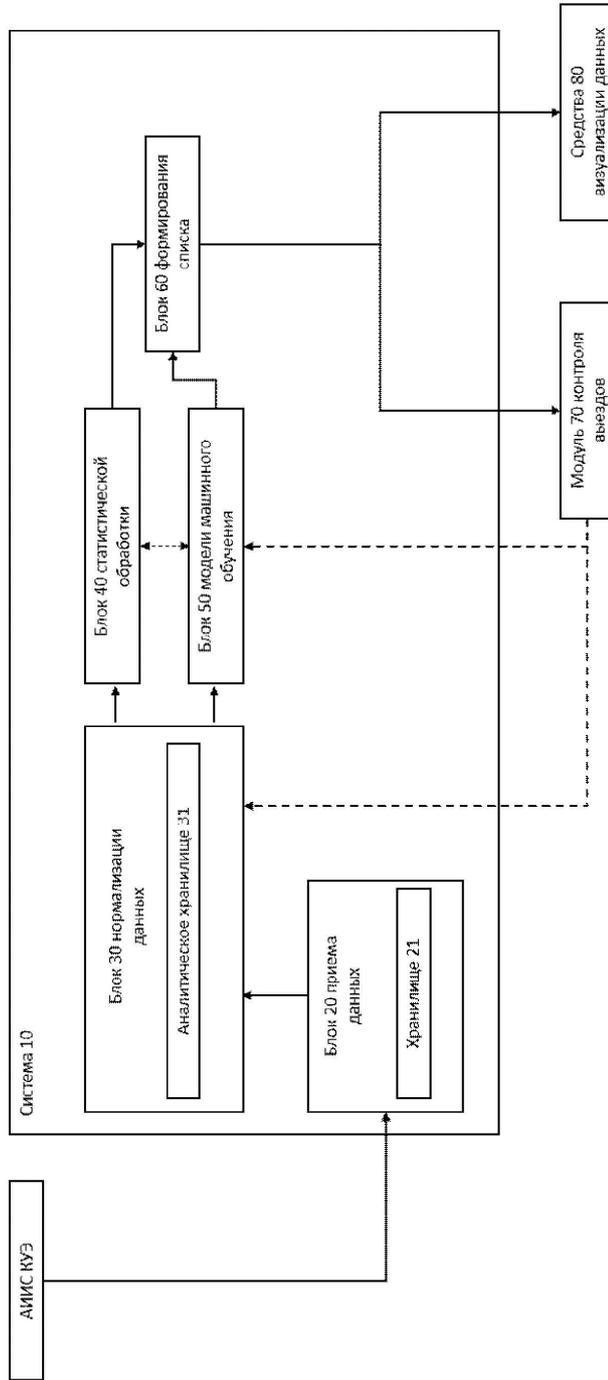
13. Способ по п.8, в котором сбор первичных данных от приборов учета электроэнергии осуществляют посредством автоматизированной информационно-измерительной системы контроля и учёта (АИИС КУЭ).

14. Способ по п.8, в котором при нормализации первичных данных осуществляют поиск и удаление ошибок телеметрии, при этом удаленные и/или утерянные данные восстанавливают посредством аппроксимации.

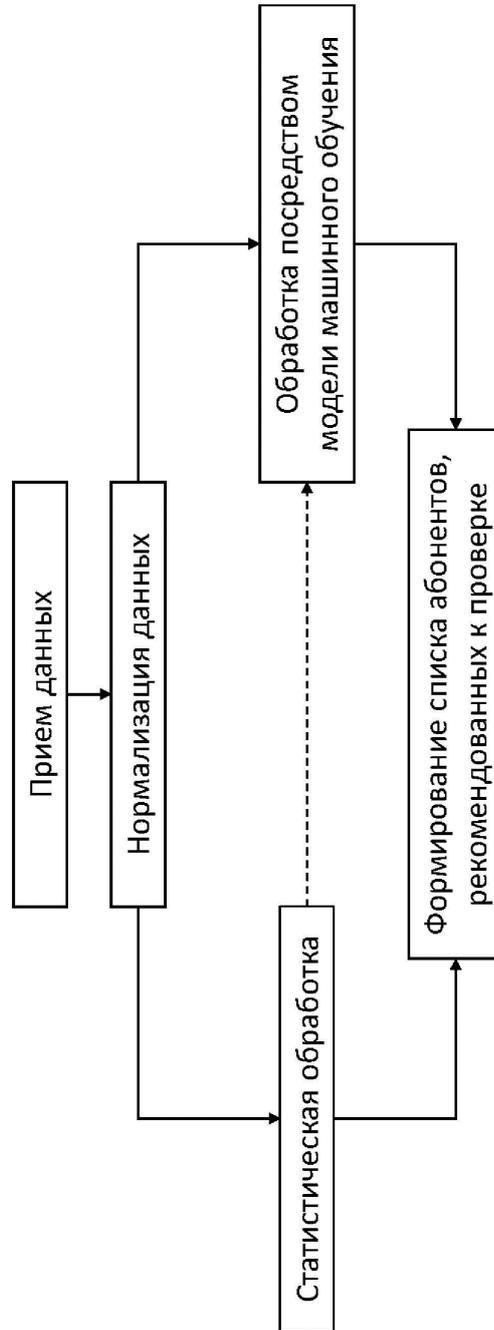
15. Машиночитаемый носитель данных, содержащий программные инструкции, при исполнении которых компьютером обеспечивается выполнение способа по любому из пп. 8-14.

40

45



ФИГ.1



Фиг.2