

(19)



**Евразийское
патентное
ведомство**

(11) **045839**(13) **B1**(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента
2023.12.29

(21) Номер заявки
202293438

(22) Дата подачи заявки
2022.12.23

(51) Int. Cl. **G06V 30/00** (2022.01)
G06F 40/295 (2020.01)
G06Q 30/00 (2023.01)

(54) **СПОСОБ И УСТРОЙСТВО ПОДБОРА ПРОДУКТА НА ОСНОВАНИИ ИНДЕКСА РЕЛЕВАНТНОСТИ**

(31) **2022123554**

(32) **2022.09.02**

(33) **RU**

(43) **2023.12.26**

(56) **US-A1-20190080207**
US-A1-20180137551
US-A1-20160225053
RU-C1-2706473

(71)(73) Заявитель и патентовладелец:
**ПУБЛИЧНОЕ АКЦИОНЕРНОЕ
ОБЩЕСТВО "СБЕРБАНК
РОССИИ" (ПАО СБЕРБАНК) (RU)**

(72) Изобретатель:
**Ермолаев Алексей Александрович,
Дмитриев Сергей Сергеевич, Сусов
Михаил Валерьевич (RU)**

(74) Представитель:
Герасин Б.В. (RU)

(57) Представленное изобретение относится, в общем, к области вычислительной техники, а в частности, к способу и устройству подбора продукта на основании N-стадийной категоризации параметров продукта и определения его индекса релевантности. Техническим результатом является повышение точности при подборе похожего продукта. Указанный технический результат достигается благодаря осуществлению способа автоматизированного подбора похожего продукта, выполняемого по меньшей мере одним вычислительным устройством, содержащего этапы, на которых: получают данные о по меньшей мере одном продукте; извлекают из полученных данных набор сущностей продукта, содержащий по меньшей мере одну сущность, экземпляр сущности, характеризующий название продукта, и атрибуты сущности, характеризующие параметры продукта; определяют типы параметров продукта посредством их классификации, причем типы параметров включают: фиксированные параметры, отклонение которых недопустимо; параметры с допустимыми отклонениями; параметр, приоритет которого выше при наименьшем его значении, или параметр, приоритет которого выше при наибольшем его значении; извлекают из набора сущностей продукта данные о экземпляре сущности и на его основе осуществляют поиск похожих продуктов; для каждого найденного похожего продукта извлекают набор сущностей, содержащий атрибуты сущностей, характеризующие параметры продукта; назначают атрибутам сущностей похожих продуктов коэффициенты релевантности посредством сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров; на основе полученных коэффициентов релевантности определяют индексы релевантности похожих продуктов; формируют список похожих продуктов и размечают похожие продукты в списке в соответствии с индексами релевантности похожих продуктов; определяют на основе индекса релевантности похожих продуктов наиболее похожий продукт.

B1**045839****045839****B1**

Область техники

Представленное изобретение относится, в общем, к области вычислительной техники, а в частности, к способу и устройству подбора продукта на основании N-стадийной категоризации параметров продукта и определения его индекса релевантности.

Уровень техники

Из уровня техники известны различные решения, предназначенные для поиска похожего товара или услуги на основе параметров товара или услуги и предпочтений пользователя.

Например, известен способ автоматизированного поиска наименьших цен на товары и услуги, раскрытый в патенте RU 2721333 С1, опубл. 19.05.2020. Известный способ содержит этапы: заполнения либо обновления баз данных, каталогизированных в СУБД, содержащих информацию о ранее запомненных повторяющихся товарах и услугах, индексированных уникальным идентификационным кодом конкретного товара и/или уникальным идентификатором товара; самостоятельного отправления запросов используемым продавцами внешним серверам; установки времени отклика на запрос каждого из продавцов; повторного обращения к устройствам продавцов с установленной продолжительностью времени отклика на запрос не более 10 минут; формирования списка продавцов, которым запросы от пользователей не переадресовываются; восстановления в СУБД информации о ценах для каждого из выбранных товаров или услуг; вычисления СУБД стоимости выбранного товара на момент запроса; ранжирования продавцов; предоставления СУБД ответа на запрос пользователя, содержащего информацию о ряде потенциальных продавцов, ранжированных по возрастанию цен и сумме цен, и демонстрации результатов в табличной форме. Также известны система и способ рекомендаций товаров пользователю, раскрытый в патенте US 6266649 В1, опубл. 24.07.2001. В известном решении осуществляют: (а) создание неспецифичной для пользователя структуры данных, которая сопоставляет отдельные элементы базы данных с соответствующими наборами подобных элементов, в которых сходство между элементами основано, по меньшей мере, на коллективных интересах сообщества пользователей; (b) определение элементов, которые, как известно, представляют интерес для пользователя; (с) для каждого из множества элементов, идентифицированных на этапе (b), доступ к структуре данных для идентификации соответствующего набора подобных элементов; (d) объединение наборов подобных элементов, определенных на этапе (с), для создания комбинированного набора дополнительно подобных элементов; а также (е) рекомендацию пользователю по меньшей мере некоторых из похожих элементов комбинированного набора, созданного на этапе (d); при этом этап (а) выполняется в автономном режиме, а этапы (b)-(е) выполняются по существу в режиме реального времени в ответ на онлайн-действие пользователя.

Недостатком известных решений является их низкая точность в выборе похожих товаров.

Раскрытие изобретения

Технической проблемой или задачей, поставленной в данном изобретении, является создание нового эффективного, простого и надежного метода подбора похожих продуктов за счет определения их индекса релевантности. Техническим результатом является повышение точности при подборе похожего продукта.

Указанный технический результат достигается благодаря осуществлению способа автоматизированного подбора похожего продукта, выполняемого по меньшей мере одним вычислительным устройством, содержащего этапы, на которых:

получают данные о по меньшей мере одном продукте;

извлекают из полученных данных набор сущностей продукта, содержащий по меньшей мере одну сущность, экземпляр сущности, характеризующий название продукта, и атрибуты сущности, характеризующие параметры продукта;

определяют типы параметров продукта посредством их классификации, причем типы параметров включают: фиксированные параметры, отклонение которых недопустимо; параметры с допустимыми отклонениями; параметр, приоритет которого выше при наименьшем его значении, или параметр, приоритет которого выше при наибольшем его значении;

извлекают из набора сущностей продукта данные о экземпляре сущности и на его основе осуществляют поиск похожих продуктов;

для каждого найденного похожего продукта извлекают набор сущностей, содержащий атрибуты сущностей, характеризующие параметры продукта;

назначают атрибутам сущностей похожих продуктов коэффициенты релевантности посредством сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров;

на основе полученных коэффициентов релевантности определяют индексы релевантности похожих продуктов;

формируют список похожих продуктов и размечают похожие продукты в списке в соответствии с индексами релевантности похожих продуктов;

определяют на основе индекса релевантности похожих продуктов наиболее похожий продукт. В одном из частных примеров осуществления способа данные о продукте могут представлять собой информацию о товаре или услуги, причем данные о продукте извлекаются из сметы.

В другом частном примере осуществления способа этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

определяют, что атрибут сущности относится к типу параметра, приоритет которого выше при большем его значении;

определяют, что значения атрибутов сущностей не совпадают;

определяют значение отклонения (в меньшую или большую сторону) атрибута сущности похожего продукта от атрибута продукта, причем если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в меньшую сторону, то такому атрибуту сущности похожего продукта назначают минимальное значение коэффициента релевантности для данного типа параметра, а если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в большую сторону, то такому атрибуту сущности похожего продукта назначают максимальное значение коэффициента релевантности для данного типа параметра.

В другом частном примере осуществления способа дополнительно выполняют этап автоматизированной замены в документе данных о продукте данными о похожем продукте.

В другом предпочтительном варианте осуществления заявленного решения представлено устройство подбора продукта, содержащее по меньшей мере одно вычислительное устройство и по меньшей мере одно устройство памяти, содержащее машиночитаемые инструкции, которые при их исполнении по меньшей мере одним вычислительным устройством выполняют вышеуказанный способ.

Краткое описание чертежей

Признаки и преимущества настоящего технического решения станут очевидными из приводимого ниже подробного описания изобретения и прилагаемых чертежей, на которых:

на фиг. 1 представлена общая схема системы обработки данных.

на фиг. 2 представлен пример общего вида вычислительного устройства.

Осуществление изобретения

Ниже будут описаны понятия и термины, необходимые для понимания данного изобретения.

В данном изобретении под системой подразумевается, в том числе компьютерная система, ЭВМ (электронно-вычислительная машина), ЧПУ (числовое программное управление), ПЛК (программируемый логический контроллер), компьютеризированные системы управления и любые другие устройства, способные выполнять заданную, четко определенную последовательность операций (действий, инструкций).

Под устройством обработки команд подразумевается электронный блок, вычислительное устройство, либо интегральная схема (микروпроцессор), исполняющая машинные инструкции (программы).

Устройство обработки команд считывает и выполняет машинные инструкции (программы) с одного или более устройств хранения данных. В роли устройства хранения данных могут выступать, но не ограничиваясь, жесткие диски (HDD), флеш-память, ПЗУ (постоянное запоминающее устройство), твердотельные накопители (SSD), оптические приводы.

Программа - последовательность инструкций, предназначенных для исполнения устройством управления вычислительной машины или устройством обработки команд.

База данных (БД) - совокупность данных, организованных в соответствии с концептуальной структурой, описывающей характеристики этих данных и взаимоотношения между ними, причем такое собрание данных, которое поддерживает одну или более областей применения (ISO/IEC 2382:2015, 2121423 "database").

Содержимое документа - данные, представляющие информацию документа и предназначенные для восприятия человеком (ГОСТ Р 52292-2004).

Сущность - любой различимый объект (объект, который мы можем отличить от другого), информацию о котором необходимо хранить в базе данных.

Экземпляр сущности - это конкретный представитель данной сущности.

Например, представителем сущности "Сотрудник" может быть "Сотрудник Иванов".

Экземпляры сущностей должны быть различимы, т.е. сущности должны иметь некоторые свойства, уникальные для каждого экземпляра этой сущности.

Атрибут сущности - поименованная характеристика сущности, определяющая его свойства и принимающая значения из некоторого множества значений. Каждый атрибут обеспечивается именем, уникальным в пределах сущности (<имя сущности>.<Имя атрибута>). Например: Самолет. Размах крыла; Кошка. Вес.

Сигнал - материальное воплощение сообщения для использования при передаче, переработке и хранении информации.

Логический элемент - элемент, осуществляющий определенные логические зависимости между входными и выходными сигналами. Логические элементы обычно используются для построения логических схем вычислительных машин, дискретных схем автоматического контроля и управления. Для всех видов логических элементов, независимо от их физической природы, характерны дискретные значения входных и выходных сигналов.

В соответствии со схемой, приведенной на фиг. 1, система обработки данных содержит: устройство 1 пользователя и устройство 10 подбора продукта.

Устройство 1 пользователя может представлять собой портативный или стационарный компьютер, телефон, смартфон, планшет или прочее вычислительное устройство, оснащенное проводными и/или беспроводными средствами связи, предназначенными для обмена данными с по меньшей мере одним устройством 10 подбора продукта.

Устройством 10 подбора продукта может быть реализовано на базе по меньшей мере одного вычислительного устройства и оснащено: модулем 11 извлечения сущностей; базой данных (БД) 12; модулем 13 классификации; модулем 14 поиска; модулем 15 сопоставления данных; и модулем 16 разметки данных.

Модуль 11 извлечения сущностей может быть реализован на базе инструмента для оптического распознавания символов (англ. optical character recognition, OCR) и включать набор моделей для анализа семантики естественных языков word2vec, заранее обученные математические модели - глубокие нейронные сети различной архитектуры (CNN (convolutional neural network), BiLSTM (Bidirectional Long Short-Term Memory), Transformer) и вычислительные средства для обработки естественного языка (Natural Language Processing, NLP). Модуль 13 классификации, модуль 14 поиска, модуль 15 сопоставления данных и модуль 16 разметки данных могут быть реализованы с использованием логических элементов на транзисторах, например, "И" "ИЛИ" "НЕ" "И-НЕ" "ИЛИ-НЕ" и пр., размещенных на печатной плате широко известным методом для обеспечения передачи между упомянутыми модулями сигналов, содержащих данные, а также для обеспечения обработки полученных данных. Соответственно, на первом этапе работы устройства 10 подбора продукта в модуль 11 извлечения сущностей поступают данные о по меньшей мере одном продукте. Данные о по меньшей мере одном продукте могут быть получены от по меньшей мере одного устройства 1 пользователя с соответствующей командой на подбор похожих продуктов. Данные о продукте могут представлять собой информацию о товаре или услуги и включать: название продукта и параметры продукта, например, для товара - размеры, вес, объем, цвет, масса, срок (эксплуатации, годности), идентификатор (ID) страны производителя и др., идентифицирующие товар признаки. Для услуги, например, по переводу документа с иранского языка на японский язык, параметры продукта могут включать: срок исполнения услуги; стоимость перевода; языки перевода; ID компании, рейтинг компании и пр.

Например, данные о по меньшей мере одном продукте могут быть указаны в документе, например, в сметном расчете и представлены в виде скан-копии документа. Пример документа представлен ниже:

ЛОКАЛЬНЫЙ СМЕТНЫЙ РАСЧЕТ (СМЕТА) № 02-01-											
№ п/п	Обоснование	Наименование работ и затрат	Единица измерения	Количество			Сметная стоимость в базисном уровне цен (в текущем уровне цен (г. 8) для ресурсов, отсутствующих в СНБ), руб.			Индекс	Сметная стоимость в текущем уровне цен, руб.
				на единицу	коэффициенты	всего с учетом коэффициентов	на единицу	коэффициенты	всего		
1	2	3	4	5	6	7	8	9	10	11	12
Раздел 1. Общестроительные работы											
33	ФССЦ-01.7.17.06-0083	Круг алмазный отрезной, диаметр 180 мм (Автомобильные дороги)	шт			0,237696	170,70		40,57	6,82	277
Раздел 2. Пожарная сигнализация и оповещение о пожаре											
12	ФССЦ-21.1.06.10-0236	Кабель силовой с медными жилами ВВГнг-FRLS 3х2,5-1000	1000 м			0,01	14 843,60		148,44	6,82	1 012
Раздел 1. Электромонтажные работы											
4	ФССЦ-82.1.01.09-0026	Выключатели автоматические: «ЕК» ВА47-29М 1Р 16А, характеристика С (Инженерное оборудование)	шт			5	12,00		60,00	4,98	299

Соответственно, после получения данных о продукте, например, упомянутого выше документа, модуль 11 извлечения сущностей известными методами выполняет распознавание текста документа и посредством семантического анализа, например, с помощью заранее обученной на соответствующей выборке данных нейронной сети, которой он может быть оснащен, извлекает из текста набор сущностей, содержащих информацию о наименовании продукта и его параметрах. Например, набор сущностей, извлеченных из упомянутого выше документа - сметы может включать: сущность "продукт", экземпляр сущности "Круг алмазный отрезной", атрибут сущности "диаметр 180 мм", атрибут сущности "стоимость 277".

Извлеченный набор сущностей может быть сохранен модулем 11 в БД 12, например, в следующем виде:

Наименование	Параметры
Круг алмазный отрезной	Материал, диаметр
Кабель силовой с медными жилами	Тип, материал, размерность (количество жил, сечение), длина
Выключатели автоматические	Функциональность, фирма производитель, значение тока, характеристика

Металл
 Размерность (количество жил, сечение)
 Стоимость
 Тип
 Фирма производитель
 Функциональность
 Характеристика

Также для каждой сущности, характеризующей продукт, могут быть сохранены атрибуты сущностей, характеризующие параметры продукта, например, в следующем виде:

Также набор сущностей в виде сигналов направляется модулем 11 в модуль 13 классификации, который в зависимости от типа продукта классифицирует параметры продукта на следующие типы (категории) параметров:

- фиксированные параметры, отклонение которых недопустимо (строгая категория);
- параметры с допустимыми отклонениями (допустимая категория);
- параметр, приоритет которого выше при наименьшем его значении;
- параметр, приоритет которого выше при наибольшем его значении.

Классификация параметров упомянутым модулем 13 выполняется посредством сравнения параметров продукта, в частности сигналов, характеризующих эти параметры, с заранее заданным для каждого типа продукта набором параметров, в частности также представленные в виде сигналов, для которых заранее определены типы параметров разработчиком модуля 13. Например, в качестве типа параметра "фиксированные параметры" (строгая категория) могут быть заданы параметры, характеризующие: материал, из которого изготовлен продукт; значение потребляемого тока; по меньшей мере одну функцию продукта; количество жил кабеля; значение сечения кабеля и пр. Для упомянутого выше примера продукта "Круг алмазный отрезной" фиксированным параметром может быть параметр, характеризующий материал.

В качестве типа параметра "параметры с допустимыми отклонениями" (допустимая категория) могут быть заданы параметры, характеризующие: диаметр; длина; фирма производитель, страна производитель и пр. Например, для упомянутого выше примера продукта "Круг алмазный отрезной" параметрами с допустимыми отклонениями могут быть параметры, характеризующие диаметр, ID фирмы производителя, ID страны изготовителя и пр.

В качестве типа параметра "параметр, приоритет которого выше при наименьшем его значении" могут быть заданы параметры, характеризующие: стоимость, вес, толщина, кредитная ставка и пр. Например, для упомянутого выше примера продукта "Круг алмазный отрезной" упомянутым параметром может быть параметр, характеризующий стоимость.

В качестве типа параметра "параметр, приоритет которого выше при наибольшем его значении" могут быть заданы параметры, характеризующие: частоту процессора; мощность зарядного устройства; процентную ставку по вкладу и пр.

Например, для упомянутой ранее документа - сметы набор сущностей может быть представлен в следующем виде:

Наименование	Параметр										
	Категория										
	Диаметр	Длина	Значение тока	Материал	Металл	Размерность (количество жил, сечение)	Стоимость	Тип	Фирма производитель	Функциональность	Характеристика
Допустимая	Допустимая	Строгая	Строгая	Строгая	Строгая	Высокая при наименьшем значении	Строгая	Допустимая	Строгая	Строгая	
Круг алмазный отрезной	180	-	-	Алмазный	-	-	277	-	-	-	-
Кабель силовой с медными жилами	-	1	-	-	Медь	3x2,5	148,44	ВВГнг-FRLS	-	-	-
Выключатели автоматические, ВА47-29М IP 16А	-	-	16А	-	-	-	299	-	ПЭК	Автоматический	С

Набор сущностей продукта и данные о типах (категориях) параметров направляются модулем 13 в модуль 14 поиска, который извлекает из набора сущностей данные о экземпляре сущности, например, "Круг алмазный отрезной" и обращается к БД, например, к БД 12, в которой может быть сохранена информация о различных продуктах, для поиска похожих продуктов. Данные о похожих товарах для БД могут быть собраны широко известными методами, например, при помощи средств для парсинга Интер-

нет-сайтов или извлечения данных из различных БД товаров и услуг, либо данные о продуктах, например, товарах могут быть собраны путем распознавания изображений, полученных с камер, установленных, например, в различных магазинах.

Поиск похожих продуктов выполняется упомянутым модулем 12 с применением широко известных методов семантического анализа текста, что позволяет учесть синонимы, содержащиеся в данных о экземпляре сущности. Например, для экземпляра сущности "Круг алмазный отрезной" при поиске с учетом синонимов могут быть найдены похожие продукты, содержащие следующие экземпляры сущностей: "Диск алмазный отрезной", "Круг отрезной абразивный", "Диск отрезной".

Для каждого найденного похожего продукта модуль 14 поиска извлекает описанным ранее способом набор сущностей, содержащий в том числе атрибуты сущностей, характеризующие параметры продукта, и направляет их вместе с полученным ранее набором сущностей и данными о типах (категориях) параметров, в виде сигналов, в модуль 15 сопоставления данных. Например, наборы сущностей продукта и похожих продуктов могут быть представлены в следующем виде:

Наименование	Параметр											
	Категория											
	Диаметр	Длина	Значение тока	Материал	Металл	Размерность (количество жил, сечение)	Стоимость	Тип	Фирма производитель	Функциональность	Характеристика	
Допустимая	Допустимая	Строгая	Строгая	Строгая	Строгая	Высшая при увеличении значения	Строгая	Допустимая	Строгая	Строгая		
Круг алмазный отрезной	180	-	-	Алмазный	-	-	277	-	-	-	-	
1. Диск алмазный отрезной STAYER	105	-	-	Алмазный	-	-	90	-	STAYER	-	-	
2. Круг отрезной абразивный "Луна"	172	-	-	Абразивный	-	-	77	-	Луна	-	-	
3. Диск отрезной BOSCH	180	-	-	Абразивный	-	-	133	-	BOSCH	-	-	
4. Диск алмазный отрезной ВИХРЬ	180	-	-	Алмазный	-	-	302	-	ВИХРЬ	-	-	
5. Диск алмазный отрезной FTT	180	-	-	Алмазный	-	-	190	-	FTT	-	-	
Кабель силовой с медными жилами	-	1	-	-	Медь	3x2,5	144,48	-	ВВГнг-FRLS	-	-	
1. Кабель силовой ВВГнг(А)-FRLS	-	1	-	-	Медь	3x2,5	156	-	ВВГнг(А)-FRLS	Конкорд	-	
2. Кабель силовой витнг Fb 3x2,5	-	1	-	-	Медь	3x2,5	77,5	-	витнг	-	-	
3. Кабель силовой ВВГнг 3x2,5	-	1	-	-	Алюминий	3x2,5	56	-	витнг fbs	-	-	
4. Кабель ВВГнг-FRLS 3x2,5 (1 ж)	-	1	-	-	Медь	3x2,5	100	-	ВВГнг-FRLS	-	-	
Выключатель автоматический, ВА47-29М 1P 16А	-	-	16А	-	-	-	299	-	-	ЛЕК	Автоматический	С
1. Авт. выкл. ВА47-29 1P 16А 4,5кА	-	-	16А	-	-	-	109	-	-	ТДМ	Автоматический	В
2. Автоматический выключатель модульный ВА47-29 1п16А	-	-	16А	-	-	-	299	-	-	ЛЕК	Автоматический	С
3. Авт. выключатель Энергия ВА 47-29 1P (С) 16 А	-	-	16А	-	-	-	177	-	-	Энергия	Автоматический	С
4. Автоматический выкл. КЭАЗ ВА47-29 1P16 А	-	-	16А	-	-	-	302	-	-	КЭАЗ	Автоматический	С

Модуль 15 сопоставления данных, после получения набора сущностей продукта, данных о типах (категориях) параметров и набора сущностей похожих продуктов, извлекает из набора сущностей похожих продуктов атрибуты сущностей, характеризующие параметры похожих продуктов, сравнивает их с аналогичными атрибутами сущностей, извлеченных из набора сущностей продукта, и с учетом данных о типах (категориях) параметров в результате сравнения назначает каждому атрибуту сущностей коэффициент релевантности, например, выбранный из интервала от 1 до 5.

Например, если тип (категория) параметра указывает на то, что атрибут сущности является фиксированным параметром (строгая категория), при этом в результате сравнения атрибута сущности похожего продукта с атрибутом сущности продукта модуль 15 определил, что значения параметров совпадают, то модуль 15 назначает атрибуту сущности похожего продукта значение коэффициента релевантности (например, максимальное значение "1" для данного типа параметра), указывающее на то, что значения атрибутов сущностей совпадают. Если значение атрибута сущности похожего продукта не соответствует значению атрибута сущности продукта, то модуль 15 назначает атрибуту сущности похожего продукта значение коэффициента релевантности (например, минимальное значение "5" для данного типа параметра), указывающее на то, что значения атрибутов сущностей строгой категории не совпадают.

Если тип (категория) параметра указывает на то, что атрибут сущности является параметром с допустимыми отклонениями (допустимая категория), то модуль 15 может запросить, например, у БД 12 допустимое значение отклонений для данного типа (категории) параметра, после чего модуль 15 сравнивает атрибут сущности похожего продукта с атрибутом сущности продукта, относящиеся к упомянутому типу (категории). Упомянутое допустимое значение отклонений, например, может характеризовать: допустимую величину, например, в процентном соотношении, на которое может отклоняться значение атрибута сущности похожего продукта от атрибута сущностей продукта; допустимый диапазон значений атрибута сущности; или упомянутое допустимое значение отклонений может указывать на то, что атрибут сущности может принимать любое значение. Если значения упомянутых атрибутов сущностей совпадают, то модуль 15 назначает атрибуту сущности похожего продукта значение коэффициента релевантности (например, значение "2"), указывающее на то, что значения атрибутов сущностей совпадают.

Если значения упомянутых атрибутов сущностей не совпадают, то модуль 15 определяет значение отклонения атрибута сущности похожего продукта от атрибута сущности продукта и сравнивает полученное значение отклонения с допустимым значением отклонений для данного типа (категории) параметра, причем:

если значение отклонения находится в пределах допустимого значения отклонений, то модуль 15 назначает атрибуту сущности похожего продукта значение коэффициента релевантности (например, значение "3"), указывающее на то, что значение отклонения атрибута сущности похожего продукта нахо-

дится в пределах допустимого значения отклонений;

если значение отклонения не находится в пределах допустимого значения отклонений, то модуль 15 назначает атрибуту сущности похожего продукта значение коэффициента релевантности (например, значение "4"), указывающее на то, что значение отклонения атрибута сущности похожего продукта не находится в пределах допустимого значения отклонений.

Если тип (категория) параметра указывает на то, что атрибут сущности является параметром, приоритет которого выше при наименьшем его значении, то в результате сравнения атрибута сущности похожего продукта и атрибута сущности продукта модулем 15 может быть определено следующее:

если значения атрибута сущности похожего продукта и атрибута сущности продукта совпадают, то модуль 15 назначает атрибуту сущности похожего продукта среднее значение коэффициента релевантности (например, значение "2");

если значения атрибута сущности похожего продукта и атрибута сущности продукта не совпадают, то модуль 15 определяет значение отклонения (в меньшую или большую сторону) атрибута сущности похожего продукта от атрибута продукта. Если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в меньшую сторону, то такому атрибуту сущности похожего продукта модуль 15 назначает максимально значение коэффициента релевантности для данного типа параметра, например, значение "1". Если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в большую сторону, то такому атрибуту сущности похожего продукта модуль 15 назначает минимальное значение коэффициента релевантности для данного типа параметра, например, значение "3".

Если тип параметра указывает на то, что атрибут сущности является параметром, приоритет которого выше при наибольшем его значении, то в результате сравнения атрибута сущности похожего продукта и атрибута сущности продукта модулем 15 может быть определено следующее:

если значения атрибута сущности похожего продукта и атрибута сущности продукта совпадают, то модуль 15 назначает атрибуту сущности похожего продукта среднее значение коэффициента релевантности, например, значение "2";

если значения атрибута сущности похожего продукта и атрибута сущности продукта не совпадают, то модуль 15 определяет значение отклонения (в меньшую или большую сторону) атрибута сущности похожего продукта от атрибута продукта. Если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в большую сторону, то такому атрибуту сущности похожего продукта модуль 15 назначает максимальное значение коэффициента релевантности для данного типа параметра, например, значение "1". Если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в меньшую сторону, то такому атрибуту сущности похожего продукта модуль 15 назначает минимальное значение коэффициента релевантности, например, значение "3".

Таким образом, модуль 15 определяет значения коэффициента релевантности для каждого атрибута сущности похожего продукта для каждого набора сущностей похожего продукта, полученного от модуля 14. Например, наборы сущностей продукта и похожих продуктов вместе со значениями коэффициентов релевантности могут быть представлены в следующем виде:

Наименование	Параметр										

	Категория										
Диаметр	Длина	Значение тока	Материал	Металл	Размерность (количество жил, сечение)	Стоимость	Тип	Фирма-производитель	Функциональность	Характеристика	
Допустимая	Допустимая	Строгая	Строгая	Строгая	Строгая	Высокая при наименьшем значении	Строгая	Допустимая	Строгая	Строгая	
Круг алмазный отрезной	180	-	-	Алмазный	-	-	277	-	-	-	-
1. Диск алмазный отрезной STAYER	105 = 4			Алмазный = 1			90 = 1		STAYER		
2. Круг отрезной абразивный "Луца"	172 = 3			Абразивный = 5			77 = 1		Луца		
3. Диск отрезной BOSCH	180 = 2			Абразивный = 5			133 = 1		BOSCH		
4. Диск алмазный отрезной ВИХРЬ	180 = 2			Алмазный = 1			302 = 3		ВИХРЬ		
5. Диск алмазный отрезной FTT	180 = 2			Алмазный = 1			190 = 1		FTT		
Кабель силовой с медными жилами	-	1	-	-	Медь	3x2,5	148,44	-	BBGm-FRLS	-	-
1. Кабель силовой BBGm(A)-FRLS		1			Медь = 1	3x2,5 = 1	156 = 3	BBGm(A)-FRLS = 1	Конкорд		
2. Кабель силовой вигит Fb 3x2,5		1			Медь = 1	3x2,5 = 1	77,5 = 1	высота = 5			
3. Кабель силовой BBGm 3x2,5		1			Алюминий = 5	3x2,5 = 1	58 = 1	высота FRLS = 1			
4. Кабель BBGm-FRLS 3x2,5 (1 м)		1			Медь = 1	3x2,5 = 1	100 = 1	BBGm-FRLS = 1			
Выключатели автоматические, ВА47-29M 1P 16A	-	-	16A	-	-	-	299	-	ТЕК	Автоматический	С
1. Авт. выкл. ВА47-29 1P 16A 4,5кА			16A = 1				109 = 1		ТДМ	Автоматический = 1	В = 5
2. Автоматический выключатель модульный ВА47-29 1P16A			16A = 1				299 = 2		ТЕК	Автоматический = 1	С = 1
3. Авт. выключатель Энергия ВА 47-29 1P (С) 16 А			16A = 1				177 = 1		Энергия	Автоматический = 1	С = 1
4. Автоматический выкл. КЭАЗ ВА47-29 1P16 А			16A = 1				302 = 3		КЭАЗ	Автоматический = 1	С = 1

Далее на основе полученных значений коэффициентов релевантности для каждого похожего продукта, в частности для каждого набора сущностей похожего продукта, модуль 15 сопоставления данных определяет индекс релевантности. Например, индекс релевантности может быть определен по формуле:

$$I_{(R)} = \sum «1» \times 1 + \sum «2» \times 0,8 + \sum «3» \times 0,6 + \sum «4» \times 0 + \sum «5» \times 0;$$

где $\Sigma_{«1»}$ - сумма коэффициентов релевантности, которым было присвоено значение "1"; $\Sigma_{«2»}$ - сумма коэффициентов релевантности, которым было присвоено значение "2"; $\Sigma_{«3»}$ - сумма коэффициентов реле-

вантности, которым было присвоено значение "3"; Σ_{n4} - сумма коэффициентов релевантности, которым было присвоено значение "3"; Σ_{n5} - сумма коэффициентов релевантности, которым было присвоено значение "5".

Полученные индексы релевантности вместе с наборами сущностей похожих продуктов направляются упомянутым модулем 15 в модуль 16 разметки данных, который на основе наборов сущностей похожих продуктов формирует список продуктов и размечает похожие продукты в списке в соответствии с индексами релевантности похожих продуктов, например, в порядке убывания значения индекса релевантности.

Наименование	Параметр												Индекс релевантности
	Категория												
	Диаметр	Длина	Значение тока	Материал	Металл	Размерность (количество жид. сечений)	Стоимость	Тип	Фирма производитель	Функциональность	Характеристика		
Допустимая	Допустимая	Строгая	Строгая	Строгая	Строгая	Высокая при наименьшем значении	Строгая	Допустимая	Строгая	Строгая			
Крут алмазный отрезной	180	-	-	Алмазный	-	-	277	-	-	-	-	-	
5. Диск алмазный отрезной FTT	180 = 2			Алмазный = 1			190 = 1		FTT			2,8	
4. Диск алмазный отрезной ВИХРЬ	180 = 2			Алмазный = 1			302 = 3		ВИХРЬ			2,4	
1. Диск алмазный отрезной STAYER	193 = 4			Алмазный = 1			90 = 1		STAYER			2,0	
3. Диск отрезной BOSCH	180 = 2			Алмазный = 3			133 = 1		BOSCH			1,8	
2. Крут отрезной абразивный "Луга"	173 = 3			Абразивный = 5			77 = 1		Луга			1,6	
Кабель силовой с медными жилами	-	1	-	-	Медь	3x2,5	148,44	-	ВВГнг-FRLS	-	-	-	
4. Кабель ВВГнг-FRLS 3x2,5 (1 м)		1			Медь = 1	3x2,5 = 1	100 = 1		ВВГнг-FRLS = 1			4,0	
1. Кабель силовой ВВГнг(A)-FRLS		1			Медь = 1	3x2,5 = 1	156 = 3		ВВГнг(A)-FRLS = 1	Конкорд		3,6	
2. Кабель силовой витое б/б 3x2,5		1			Медь = 1	3x2,5 = 1	77,5 = 1		витое б/б = 5			3,0	
3. Кабель силовой ВВГнг 3x2,5		1			Алюминий = 5	3x2,5 = 1	56 = 1		витое б/б = 1			3,0	
Выключатели автоматические, ВА47-29М 1P 16А	-	-	16А	-	-	-	299	-	ТЭК	Автоматический	С	-	
3. Авт. выключатель Энергия ВА 47-29 1P (С) 16 А			16А = 1				177 = 1		Энергия	Автоматический = 1	С = 1	4	
2. Автоматический выключатель модульный ВА47-29 1P16А			16А = 1				259 = 2		ТЭК	Автоматический = 1	С = 1	3,8	
4. Автоматический выкл. КЭАЗ ВА47-29 1P16 А			16А = 1				302 = 3		КЭАЗ	Автоматический = 1	С = 1	3,6	
1. Авт. выкл. ВА47-29 1P 16А 4,5кА			16А = 1				109 = 1		ТДМ	Автоматический = 1	В = 5	3	

Для продукта по предоставлению услуги по переводу документа упомянутые выше данные могут быть представлены в следующем виде:

Наименование услуги	Параметр						Индекс релевантности
	Категория						
	Срок исполнения, час	Стоимость перевода слова, руб.	Языки перевода	Рейтинг	Компания		
Высокая при наименьшем значении	Высокая при наименьшем значении	Строгая	Высокая при наибольшем значении	Допустимая			
Перевод документа с иранского языка на японский язык	10		иранский, японский				
	6 = 1	1,8 = 2	Русский, Английский, Монгольский, Иранский, Итальянский, Китайский, Японский, Французский, Немецкий = 1	4,5 = 2	Организация С	3,6	
	24 = 3	1,75 = 2	Русский, Английский, Монгольский, Иранский, Китайский, Японский, Французский, Немецкий = 1	4 = 2	Организация В	3,2	
	12 = 3	1,5 = 1	Русский, Английский, Монгольский, Иранский = 5	4,8 = 1	Организация А	2,6	

Соответственно, похожий продукт с наибольшим значением индекса релевантности определяется устройством 10 подбора продукта как наиболее похожий продукт.

Полученный список может быть сохранен в БД 12, направлен в устройство 1 пользователя для отображения его на экране, либо передан во внешние/внутренние системы, выполненные с возможностью автоматизированной замены в документе данных о продукте данными о похожем продукте, формирования в автоматизированном режиме альтернативного документа (например, локального сметного расчета) и возможностью направления команды для автоматизированного заказа товара согласно данным о продукте, представленным в документе. При этом в режиме замены в документе данных о продукте данными

о похожем продукте могут быть учтены заданные пользователем критерии атрибутов сущностей похожих товаров. Например, в упомянутой системе может быть определено, чтобы похожие продукты заменяются системой в документе только в том случае, если они имеют заданное значение атрибута сущности, например, значение сущности, характеризующее конкретную фирму, страну изготовления или пр.

Таким образом, за счет того, что поиск похожих продуктов осуществляется только лишь на основе экземпляра сущности продукта, т.е. без учета параметров продукта, расширяется выборка продуктов, а за счет того, что при сравнении атрибутов сущностей продукта с атрибутами сущностей похожих продуктов учитывается тип параметров, повышается точность при подборе похожего продукта.

В общем виде (см. фиг. 2) вычислительное устройство (200) содержит объединенные общей шиной информационного обмена один или несколько процессоров (201), средства памяти, такие как ОЗУ (202) и ПЗУ (203), интерфейсы ввода/вывода (204), устройства (средства) ввода/вывода (205), и устройство (средство) для сетевого взаимодействия (206).

Процессор (201) (или несколько процессоров, многоядерный процессор и т.п.) может выбираться из ассортимента устройств, широко применяемых в настоящее время, например, таких производителей, как: Intel™, AMD™, Apple™, Samsung Exynos™, MediaTEK™, Qualcomm Snapdragon™ и т.п. Под процессором или одним из используемых процессоров в устройстве (200) также необходимо учитывать графический процессор, например, GPU NVIDIA с программной моделью, совместимой с CUDA, или Graphcore, тип которых также является пригодным для полного или частичного выполнения способа, а также может применяться для обучения и применения моделей машинного обучения в различных информационных системах.

ОЗУ (202) представляет собой оперативную память и предназначено для хранения исполняемых процессором (201) машиночитаемых инструкций для выполнения необходимых операций по логической обработке данных. ОЗУ (202), как правило, содержит исполняемые инструкции операционной системы и соответствующих программных компонент (приложения, программные модули и т.п.). При этом, в качестве ОЗУ (202) может выступать доступный объем памяти графической карты или графического процессора.

ПЗУ (203) представляет собой одно или более устройств постоянного хранения данных, например, жесткий диск (HDD), твердотельный накопитель данных (SSD), флэш-память (EEPROM, NAND и т.п.), оптические носители информации (CD-R/RW, DVD-R/RW, BlueRay Disc, MD) и др.

Для организации работы компонентов устройства (200) и организации работы внешних подключаемых устройств применяются различные виды интерфейсов В/В (204). Выбор соответствующих интерфейсов зависит от конкретного исполнения вычислительного устройства, которые могут представлять собой, не ограничиваясь: PCI, AGP, PS/2, IrDa, FireWire, LPT, COM, SATA, IDE, Lightning, USB (2.0, 3.0, 3.1, micro, mini, type C), TRS/Audio jack (2.5, 3.5, 6.35), HDMI, DVI, VGA, Display Port, RJ45, RS232 и т.п.

Для обеспечения взаимодействия пользователя с вычислительным устройством (200) применяются различные средства (205) В/В информации, например, клавиатура, дисплей (монитор), сенсорный дисплей, тач-пад, джойстик, манипулятор мышь, световое перо, стилус, сенсорная панель, трекбол, динамики, микрофон, средства дополненной реальности, оптические сенсоры, планшет, световые индикаторы, проектор, камера, средства биометрической идентификации (сканер сетчатки глаза, сканер отпечатков пальцев, модуль распознавания голоса) и т.п.

Средство сетевого взаимодействия (206) обеспечивает передачу данных посредством внутренней или внешней вычислительной сети, например, Интранет, Интернет, ЛВС и т.п. В качестве одного или более средств (206) может использоваться, но не ограничиваясь: Ethernet карта, GSM модем, GPRS модем, LTE модем, 5G модем, модуль спутниковой связи, NFC модуль, Bluetooth и/или BLE модуль, Wi-Fi модуль и др.

Дополнительно могут применяться также средства спутниковой навигации в составе устройства (200), например, GPS, ГЛОНАСС, BeiDou, Galileo.

Конкретный выбор элементов устройства (200) для реализации различных программно-аппаратных архитектурных решений может варьироваться с сохранением обеспечиваемого требуемого функционала.

Модификации и улучшения вышеописанных вариантов осуществления настоящего технического решения будут ясны специалистам в данной области техники. Предшествующее описание представлено только в качестве примера и не несет никаких ограничений. Таким образом, объем настоящего технического решения ограничен только объемом прилагаемой формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ автоматизированного подбора похожего продукта, выполняемый по меньшей мере одним вычислительным устройством, содержащий этапы, на которых:

получают данные о по меньшей мере одном продукте;

извлекают из полученных данных набор сущностей продукта, содержащий по меньшей мере одну сущность, экземпляр сущности, характеризующий название продукта, и атрибуты сущности, характеризующие параметры продукта;

определяют типы параметров продукта посредством их классификации, причем типы параметров

включают: фиксированные параметры, отклонение которых недопустимо; параметры с допустимыми отклонениями; параметр, приоритет которого выше при наименьшем его значении, или параметр, приоритет которого выше при наибольшем его значении;

извлекают из набора сущностей продукта данные о экземпляре сущности и на его основе осуществляют поиск похожих продуктов;

для каждого найденного похожего продукта извлекают набор сущностей, содержащий атрибуты сущностей, характеризующие параметры продукта;

назначают атрибутам сущностей похожих продуктов коэффициенты релевантности посредством сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров;

на основе полученных коэффициентов релевантности определяют индексы релевантности похожих продуктов;

формируют список похожих продуктов и размечают похожие продукты в списке в соответствии с индексами релевантности похожих продуктов;

определяют на основе индекса релевантности похожих продуктов наиболее похожий продукт.

2. Способ по п.1, характеризующийся тем, что данные о продукте могут представлять собой информацию о товаре или услуги, причем данные о продукте извлекаются из сметы.

3. Способ по п.1, характеризующийся тем, что поиск похожих продуктов на основе экземпляра сущности осуществляется с применением семантического анализа текста.

4. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

определяют, что атрибут сущности относится к типу фиксированных параметров, причем если значения атрибутов сущностей совпадают, то назначают атрибуту сущности похожего продукта значение коэффициента релевантности, указывающее на то, что значения атрибутов сущностей совпадают или назначают атрибуту сущности максимальное значение коэффициента релевантности для данного типа параметра, а если значение атрибута сущности похожего продукта не соответствует значению атрибута сущности продукта, то назначают атрибуту сущности похожего продукта минимальное значение коэффициента релевантности для данного типа параметра, указывающее на то, что значения атрибутов сущностей не совпадают.

5. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

определяют, что атрибут сущности относится к типу параметров с допустимыми отклонениями, причем если значения атрибутов сущностей совпадают, то назначают атрибуту сущности похожего продукта значение коэффициента релевантности, указывающее на то, что значения атрибутов сущностей совпадают или назначают атрибуту сущности максимальное значение коэффициента релевантности для данного типа параметра.

6. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

определяют, что атрибут сущности относится к типу параметров с допустимыми отклонениями;

определяют, что значения атрибутов сущностей не совпадают;

определяют значение отклонения атрибута сущности похожего продукта от атрибута сущности продукта;

сравнивают полученное значение отклонения с допустимым значением отклонений для данного типа параметра, причем если значение отклонения находится в пределах допустимого значения отклонений, то назначают атрибуту сущности похожего продукта значение коэффициента релевантности, указывающее на то, что значение отклонения атрибута сущности похожего продукта находится в пределах допустимого значения отклонений, а если значение отклонения не находится в пределах допустимого значения отклонений, то назначают атрибуту сущности похожего продукта значение коэффициента релевантности, указывающее на то, что значение отклонения атрибута сущности похожего продукта не находится в пределах допустимого значения отклонений.

7. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

определяют, что атрибут сущности относится к типу параметра, приоритет которого выше при наименьшем его значении, причем если значения атрибута сущности похожего продукта и атрибута сущности продукта совпадают, то назначают атрибуту сущности похожего продукта среднее значение коэффициента релевантности.

8. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит

этапы, на которых:

определяют, что атрибут сущности относится к типу параметра, приоритет которого выше при наименьшем его значении;

определяют, что значения атрибутов сущностей не совпадают;

определяют значение отклонения (в меньшую или большую сторону) атрибута сущности похожего продукта от атрибута продукта, причем если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в меньшую сторону, то такому атрибуту сущности похожего продукта назначают максимально значение коэффициента релевантности для данного типа параметра, а если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в большую сторону, то такому атрибуту сущности похожего продукта назначают минимальное значение коэффициента релевантности для данного типа параметра.

9. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

определяют, что атрибут сущности относится к типу параметра, приоритет которого выше при наибольшем его значении, причем если значения атрибута сущности похожего продукта и атрибута сущности продукта совпадают, то назначают атрибуту сущности похожего продукта среднее значение коэффициента релевантности.

10. Способ по п.1, характеризующийся тем, что этап сравнения атрибутов сущностей похожих продуктов с аналогичными атрибутами сущностей продукта с учетом данных о типах параметров содержит этапы, на которых:

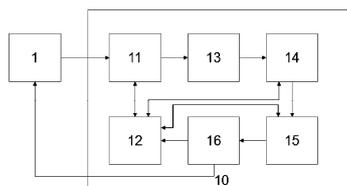
определяют, что атрибут сущности относится к типу параметра, приоритет которого выше при наибольшем его значении;

определяют, что значения атрибутов сущностей не совпадают;

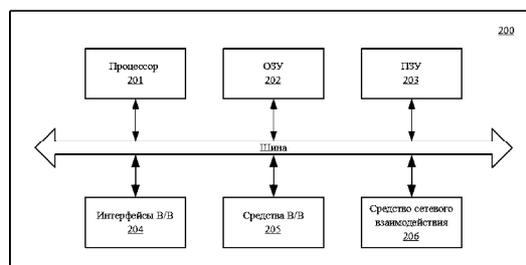
определяют значение отклонения (в меньшую или большую сторону) атрибута сущности похожего продукта от атрибута продукта, причем если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в меньшую сторону, то такому атрибуту сущности похожего продукта назначают минимальное значение коэффициента релевантности для данного типа параметра, а если значение атрибута сущности похожего продукта отклоняется от значения атрибута продукта в большую сторону, то такому атрибуту сущности похожего продукта назначают максимальное значение коэффициента релевантности для данного типа параметра.

11. Способ по п.1, характеризующийся тем, что дополнительно выполняют этап автоматизированной замены в документе данных о продукте данными о похожем продукте.

12. Устройство подбора продукта, содержащее по меньшей мере одно вычислительное устройство и по меньшей мере одно устройство памяти, содержащее машиночитаемые инструкции, которые при их исполнении по меньшей мере одним вычислительным устройством выполняют способ по любому из пп.1-11.



Фиг. 1



Фиг. 2

