



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21), (22) Заявка: 2005119376/09, 21.06.2005

(24) Дата начала отсчета срока действия патента:
21.06.2005(30) Конвенционный приоритет:
22.06.2004 US 10/873,676

(43) Дата публикации заявки: 27.12.2006

(45) Опубликовано: 27.08.2010 Бюл. № 24

(56) Список документов, цитированных в отчете о поиске: **Charran Eric, Introduction to Data Mining with SQL Server**, 08.01.2003 г. (найдено в Интернет по адресу: performance.com/articles/dba/data-mining-intro-p1.aspx). RU 2167450 C2, 27.12.2000. **Tang Z., Building Data Mining Solutions with SQL Server 2000, 2001** (найдено в Интернет по адресу: (см. прод.)

Адрес для переписки:

129090, Москва, ул. Б.Спасская, 25, стр.3,
ООО "Юридическая фирма Городиский и
Партнеры", пат.пов. Ю.Д.Кузнецову,
рег.№ 595

(72) Автор(ы):

**МАКЛЕННАН К. Джеймс (US),
КИМ Пюнгчул (US),
ТАН Чжаохой (US)**

(73) Патентообладатель(и):

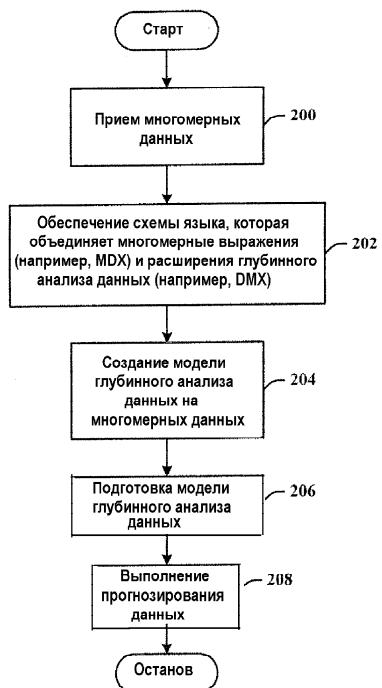
МАЙКРОСОФТ КОРПОРЕЙШН (US)

(54) ОБЪЕДИНЕНИЕ МНОГОМЕРНЫХ ВЫРАЖЕНИЙ И РАСШИРЕНИЙ ГЛУБИННОГО АНАЛИЗА ДАННЫХ ДЛЯ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ ИЗ КУБОВ OLAP

(57) Реферат:

Изобретение относится к способам поиска в базах данных и анализа баз данных. Изобретение позволяет облегчить операции глубинного анализа данных по многомерным кубам данных. Компонент глубинного анализа данных объединяет многомерные выражения (например, MDX) и расширения глубинного анализа данных (например, DMX) для выполнения операций глубинного анализа данных в отношении данных, находящихся в

многомерных кубах данных (например, в кубах интерактивной аналитической обработки (OLAP)) для создания модели глубинного анализа данных. При этом объединение многомерных выражений и расширений глубинного анализа данных выполняется посредством использования внутренней многомерной структуры многомерного куба данных для записи реляционного запроса в виде многомерного выражения. 4 н. и 16 з.п. ф-лы, 5 ил.



ФИГ.2

(56) (продолжение):

[http://csis.bits-](http://csis.bits-pilani.ac.in/faculty/goel/Data%20Mining/Assignemnts/Assignment%202/DM%20with%20SQL%20Server%20K.pdf)

[pilani.ac.in/faculty/goel/Data%20Mining/Assignemnts/Assignment%202/DM%20with%20SQL%20Server%20K.pdf](http://csis.bits-pilani.ac.in/faculty/goel/Data%20Mining/Assignemnts/Assignment%202/DM%20with%20SQL%20Server%20K.pdf). US 2003/236784 A1, 25.12.2003. US 2003/145000 A1, 31.07.2003.

RU 2398273 C2

RU 2398273 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(51) Int. Cl.
G06F 17/30 (2006.01)

(12) ABSTRACT OF INVENTION

(21), (22) Application: **2005119376/09, 21.06.2005**
 (24) Effective date for property rights:
21.06.2005
 (30) Priority:
22.06.2004 US 10/873,676
 (43) Application published: **27.12.2006**
 (45) Date of publication: **27.08.2010 Bull. 24**

Mail address:
**129090, Moskva, ul. B.Spasskaja, 25, str.3, OOO
 "Juridicheskaja firma Gorodisskij i Partnery",
 pat.pov. Ju.D.Kuznetsovu, reg.№ 595**

(72) Inventor(s):
**MAKLENNAN K. Dzhejms (US),
 KIM Pjungchul (US),
 TAN Chzhaokhoj (US)**
 (73) Proprietor(s):
MAJKROSOFT KORPOREJShN (US)

(54) COMBINATION OF MULTIDIMENSIONAL EXPRESSIONS AND DEEP DATA ANALYSIS EXTENSIONS FOR EXTRACTING INFORMATION FROM OLAP CUBES

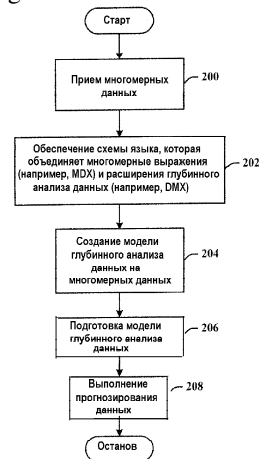
(57) Abstract:

FIELD: information technology.

SUBSTANCE: deep data analysis component combines multidimensional expressions (e.g. MDX) and deep data analysis extensions (e.g. DMX) for performing deep analysis of data in multidimensional data cubes (e.g. in online analytical processing (OLAP) cubes) to create a deep data analysis model. Multidimensional expressions and deep data analysis extensions are combined by using the internal multidimensional structure of a multidimensional data cube to record a relational request in form of a multidimensional expression.

EFFECT: invention simplifies deed data analysis in multidimensional data cubes.

20 cl, 5 dwg



ФИГ.2

RU 2 398 273 C2

RU 2 398 273 C2

Область техники, к которой относится изобретение

Настоящее изобретение относится к базам данных, в частности к способам для поиска и анализа таких баз данных.

Предшествующий уровень техники

5 Появление глобальной сети связи, такой как Интернет, привело к обмену огромными объемами информации. Кроме того, затраты на хранение и поддержку такой информации уменьшились, что привело к возникновению огромных структур хранения данных, к которым необходимо обеспечить доступ. Огромные объемы
10 данных могут храниться в виде хранилища данных, которое является базой данных, как правило, отражающей коммерческую историю организации. Исторические данные используются для анализа, который поддерживает бизнес-решения на многих уровнях, от стратегического планирования до оценки эффективности отдельного подразделения. Оно также может включать в себя получение данных, сохраненных в
15 реляционной базе данных, и обработку данных, для того чтобы сделать его более эффективным инструментом для запроса и анализа. Для того чтобы более эффективно управлять хранением данных в меньшем масштабе, применяется концепция специализированных хранилищ данных, согласно которой обрабатывается только
20 целевое подмножество данных.

Несмотря на то что было создано много языков, используемых для описания данных и для манипулирования данными в двух измерениях, например, SQL (структурированный язык запросов), многомерные данные, с другой стороны, могут
25 быть представлены структурами, имеющими более чем два измерения. Эти многомерные структуры называются кубами. Куб - это многомерная база данных, которая представляет данные подобно трехмерной электронной таблицы в отличие от реляционной базы данных. Куб позволяет быстро отобразить различные представления данных при помощи использования концепций измерений и
30 многомерных показателей. Измерения определяют структуру куба (например, географическое положение или тип продукта), в то время как многомерные показатели предоставляют интересующие количественные величины конечному пользователю (например, объем продаж, количество единиц на складе и суммарные расходы). Положения ячеек в кубе определяются пересечением элементов измерения, и
35 значения многомерного показателя агрегируются для предоставления значений в ячейках.

Информация в хранилище данных или в специализированном хранилище данных может быть обработана с использованием интерактивной аналитической
40 обработки (OLAP). OLAP представляет данные в виде кубов. OLAP делает возможным эффективное использование хранилищ данных и специализированных хранилищ данных для интерактивного анализа и обеспечивает быстрые ответы на интерактивные сложные запросы анализа. Системы OLAP обеспечивают
быстродействие и гибкость, необходимые для поддержки анализа в реальном времени.

45 Одной известной архитектурой, которая может облегчить использование OLAP для многомерного запроса и анализа, является язык MDX (язык многомерных выражений). MDX - это язык, который поддерживает определение и манипуляцию многомерными объектами и данными, тем самым делая доступ к данным из
50 множества измерений более легким и интуитивно понятным. Синтаксис языка MDX во многом похож на SQL (структурированный язык запросов) синтаксис (но он не является расширением языка SQL). Как и в запросе в SQL, каждый запрос в MDX требует запроса данных (предложение SELECT), начальную точку

(предложение FROM) и фильтр (предложение WHERE). Эти и другие ключевые слова обеспечивают инструменты, используемые для извлечения определенной части данных из куба для анализа. Язык MDX также предоставляет большой набор функций для манипуляции извлеченными данными, а также возможность расширения языка MDX при помощи функций, определяемых пользователем.

Глубинный анализ данных - это нахождение интересных структур в данных (например, образцов и правил), которые могут быть интерпретированы как знание о данных или могут быть использованы для прогнозирования событий, связанных с данными. Эти структуры могут принимать форму образцов, которые являются кратким описанием набора данных. Глубинный анализ данных делает исследование и эксплуатацию больших баз данных легкой, удобной и практичной для тех, кто имеет данные, а не годы подготовки в области статистики и анализа данных. «Знание», извлеченное алгоритмом глубинного анализа данных, может принимать разные формы и использоваться различным образом. Глубинный анализ данных может представлять набор правил, дерево решений, регрессионную модель или набор ассоциаций, наряду с другими возможностями. Он может быть использован для получения сводок данных или для получения представления о ранее неизвестных корреляциях. Глубинный анализ данных также может быть использован для прогнозирования событий, связанных с данными, например пропущенных значений, записей, для которых некоторая информация не известна, и так далее. Существует много различных методик глубинного анализа данных, большинство из которых происходит из областей обучения машины, статистики и программирования баз данных.

То, что необходимо, так это схема, облегчающая взаимодействие операций глубинного анализа по кубам OLAP.

Сущность изобретения

Нижеследующее представляет собой упрощенное краткое изложение сущности изобретения, для того чтобы обеспечить базовое понимание некоторых аспектов изобретения. Это краткое описание не является расширенным обзором изобретения. Оно не предназначено для определения ключевых/критических элементов изобретения или для определения объема изобретения. Его единственное предназначение заключается в представлении некоторых концепций изобретения в упрощенной форме, предваряя более детальное описание, представленное ниже.

Настоящее изобретение, раскрытое и заявленное здесь, в одном его аспекте, включает в себя формальный язык, который объединяет язык многомерных выражений (например, MDX) и расширения глубинного анализа данных (например, DMX) для выполнения операций глубинного анализа данных в отношении данных, находящихся в кубах OLAP. Как правило, операции глубинного анализа данных выполняют операции над набором исходных данных, обозначаемым как <source-data-query> (<источник-данные-запрос>). До настоящего времени элементы <source-data-query> ограничивались реляционными запросами, действующими непосредственно на реляционную базу данных, или оператором SHAPE, который принимает реляционный запрос и формирует из него вложенный набор строк. Настоящее изобретение обеспечивает то, что <source-data-query> может быть не только реляционным запросом, а и многомерным запросом, сформированным, например, при помощи языка MDX.

В другом аспекте настоящего изобретения, модели глубинного анализа данных используются для выполнения прогнозирования в отношении данных, содержащихся

в кубе OLAP.

В другом аспекте настоящего изобретения, что касается создания модели, настоящее изобретение основывается на том, что до создания тип исходных данных не известен и не установлен до фазы подготовки. В известных системах «тип» модели
5 определяется до создания, типом может быть модель с реляционным источником или источником OLAP.

Более того, модель глубинного анализа данных может быть подготовлена из произвольного источника данных, безотносительно к его реляционному или
10 многомерному происхождению. Связывание столбцов поддерживается согласно явному порядку столбцов как в многомерном, так и в реляционном источнике в отличие от известных систем, в которых связывание столбцов поддерживается согласно соответствию имен между моделью глубинного анализа данных и кубом OLAP.

Более того, модель глубинного анализа данных может принимать в качестве источника данных для прогнозирования произвольный источник данных, безотносительно к его реляционному или многомерному происхождению. Раскрытая
15 здесь архитектура позволяет осуществлять прогнозирование с использованием DMX и позволяет использовать куб OLAP в качестве источника для прогнозирования на основе любой модели, безотносительно к тому, каким образом она была создана или подготовлена. Традиционно, прогнозирование в отношении данных в кубе OLAP выполняется в MDX, и только с использованием модели глубинного анализа данных,
20 подготовленной для того же куба.

Для реализации вышеизложенных и связанных с ними целей определенные иллюстративные аспекты настоящего изобретения описаны здесь в связи со
25 следующим описанием и сопроводительными чертежами. Эти аспекты являются показательными, тем не менее в отношении лишь нескольких из различных способов, которыми могут быть применены принципы настоящего изобретения, и
30 подразумевается, что настоящее изобретение включает в себя все такие способы и их эквиваленты. Другие преимущества и новые признаки настоящего изобретения станут очевидными из нижеследующего детального описания, рассматриваемого совместно с чертежами.

35 Перечень чертежей

Фиг.1 - иллюстрация системы, обеспечивающей интеграцию многомерных выражений и расширений глубинного анализа данных в соответствии с настоящим изобретением.

40 Фиг.2 - блок-схема последовательности операций согласно одной методике глубинного анализа многомерных данных в соответствии с настоящим изобретением.

Фиг.3 - блок-схема последовательности операций примерных способов использования многомерных выражений в качестве ввода для расширений глубинного анализа многомерных данных в соответствии с настоящим изобретением.

45 Фиг.4 - структурная схема компьютера, предназначенного для выполнения раскрытой архитектуры.

Фиг.5 - структурная схема примерной вычислительной среды в соответствии с настоящим изобретением.

50 Подробное описание изобретения

Настоящее изобретение описано со ссылкой на чертежи, на которых одинаковые ссылочные номера используются для указания на одинаковые элементы. В следующем описании, для целей объяснения, объясняются многочисленные специфические детали

для того, чтобы обеспечить исчерпывающее понимание настоящего изобретения. Тем не менее может быть очевидным, что настоящее изобретение может быть реализовано на практике без этих специфических деталей. В других случаях широко известные структуры и устройства показаны в форме блок-схем, для того чтобы упростить

описание настоящего изобретения.

В этой заявке термины «компонент» и «система» предназначены для обозначения относящейся к вычислительным системам сущности, являющейся или аппаратным обеспечением, или комбинацией аппаратного и программного обеспечения, или программным обеспечением, или программным обеспечением в момент выполнения. Например, компонент, может быть, но не в ограничительном смысле, процессом, выполняемым процессором, процессором, объектом, исполняемым файлом, потоком исполнения, программой и/или компьютером. В качестве иллюстрации, и приложение, выполняемое сервером, и сервер могут быть компонентами. Один или более компонентов могут находиться в пределах процесса и/или потока исполнения, и компонент может быть размещен на одном компьютере и/или распределен между двумя и более компьютерами.

Изобретение обеспечивает архитектуру для интеграции многомерных выражений и расширений глубинного анализа данных, для выполнения операций глубинного анализа данных над данными, находящимися в кубах OLAP. В настоящее время элементы <source-data-query> ограничиваются реляционными запросами, действующими непосредственно на реляционную базу данных, или оператором SHAPE, который принимает реляционный запрос и формирует из него вложенный набор строк. Настоящее изобретение обеспечивает то, что <source-data-query> может быть не только реляционным запросом, а и многомерным запросом, сформированным с использованием многомерных выражений.

Одним способом, при помощи которого это может быть выполнено, является использование MDX и DMX. MDX есть акроним от MultiDimensional eXpressions, как определено в спецификации OLE (связь и включение объекта) для баз данных для OLAP, разработанной корпорацией Microsoft, целиком включенной в настоящее описание посредством ссылки. DMX есть акроним от Data Mining eXtensions, как определено в спецификации OLE для баз данных для глубинного анализа данных, разработанной корпорацией Microsoft, целиком включенной в настоящее описание посредством ссылки. Для целей данного описания операции создания модели, подготовки модели и прогнозирования в отношении новых данных описаны ниже.

В DMX эти операции выполняются с использованием операторов CREATE MINING MODEL, INSERT INTO и SELECT... PREDICTION JOIN. Например:

```
CREATE MINING MODEL MyModel
```

```
(
```

CustomerID	LONG	KEY,
Age	LONG	CONTINUOUS,
Gender	TEXT	DISCRETE,
Occupation	TEXT	DISCRETE,
HomeOwner	BOOLEAN	DISCRETE,
TotalSales	DOUBLE	CONTINUOUS
MemberCard	TEXT	DISCRETE PREDICT
Products	TABLE	

```
(
```

```
ProductID TEXT KEY
```

)

) USING Microsoft_Decision_Trees

INSERT INTO MyModel(CustomerID, Age, Gender, Occupation,
Homeowner, TotalSales, MemberCard,

Products(Product))

<source-data-query>

SELECT Predict(MemberCard) FROM MyModel PREDICTION JOIN

<source-data-query> as t

ON MyModel.Age = t.Age, MyModel.Gender = t.Gender, ...,

MyModel.Products.ProductID=t.Products.ProductID

Теперь обратимся к Фиг.1, где проиллюстрирована система 100, обеспечивающая интеграцию многомерных выражений и расширений глубинного анализа данных в соответствии с настоящим изобретением. Здесь имеется источник 102 многомерных данных (например, куб OLAP (интерактивная обработка данных)) в котором осуществляется глубинный анализ данных. Компонент 104 глубинного анализа данных включает в себя компонент 106 многомерного выражения (например, MDX) и компонент 108 расширения глубинного анализа данных, которые совместно обеспечивают глубинный анализ данных куба 102 OLAP. Выводом для компонента глубинного анализа данных являются данные, которые используются в качестве ввода для создания модели глубинного анализа данных, подготовки и прогнозирования.

В альтернативной реализации принимается во внимание, что размещение компонента 106 многомерного выражения и компонента 108 расширения глубинного анализа данных в одном общем компоненте 104 глубинного анализа данных не требуется, и это может быть отдельная внешняя сущность по отношению к компоненту 104 глубинного анализа данных. Таким образом, эти компоненты (106 и 108) могут быть независимы, при условии, что один снабжает другой.

Теперь обратимся к Фиг.2, где проиллюстрирована блок-схема последовательности операций согласно одной методике глубинного анализа многомерных данных в соответствии с настоящим изобретением. Несмотря на то что в целях упрощения изложения одна или более методологий, показанных здесь, например, в форме блок-схемы последовательности операций, представлены и описаны как последовательность действий, должно быть понятно и оценено, что настоящее изобретение не ограничивается порядком следования действий, так как некоторые действия могут в соответствии с настоящим изобретением происходить в другом порядке и/или одновременно с другими действиями из числа представленных и описанных здесь. Например, для специалиста в данной области техники должно быть понятно, что методика может быть альтернативно представлена как серия взаимосвязанных состояний или событий, таких как на диаграмме состояния. Более того, не все проиллюстрированные действия могут требоваться для реализации методики в соответствии с настоящим изобретением.

На этапе 200 многомерные данные предоставляют для глубинного анализа данных. На этапе 202 обеспечивают схему языка, которая объединяет многомерные выражения и расширения глубинного анализа данных и работает с многомерными данными. На этапе 204 из многомерных данных создают модель глубинного анализа данных. На этапе 206 модель глубинного анализа подготавливают на основе данных. На этапе 208 может быть выполнено прогнозирование в отношении новых данных при помощи расширений глубинного анализа. Затем процесс достигает точки останова.

Теперь обратимся к Фиг.3, где проиллюстрирована блок-схема последовательности

операций примерных способов использования многомерных выражений в качестве ввода для расширений глубинного анализа данных в соответствии с настоящим изобретением. Традиционно запросы расширения глубинного анализа данных в качестве источника данных ожидают реляционные таблицы. Эти запросы могут

5 принимать или плоские таблицы в качестве ввода, или вложенные таблицы, созданные с помощью директивы SHAPE. Связывание выполняется или по порядку столбцов - как в операторе INSERT INTO, или при помощи явного отображения с использованием предложения ON, как в операторе SELECT... PREDICTION JOIN.

10 Например:

```
INSERT INTO MyModel (CustomerID, Age, Gender, Occupation,
HomeOwner,
TotalSales, MemberCard,
Products{SKIP, ProductID})
```

15 SHAPE (SELECT CustomerID, Age, Gender, Occupation, HomeOwner, TotalSales, MemberCard

```
FROM MyTable)
```

```
APPEND {(SELECT CustomerID, ProductID From ProductFacts )
```

```
RELATE CustomerID to CustomerID} as Products
```

```
SELECT Predict(MemberCard) FROM MyModel PREDICTION JOIN
```

```
SHAPE (SELECT CustomerID, Age, Gender, Occupation, HomeOwner, TotalSales,
```

```
MemberCard
```

```
FROM MyTable)
```

```
APPEND {(SELECT CustomerID, ProductID From ProductFacts )
```

```
RELATE CustomerID to CustomerID} as Products as t
```

```
ON MyModel.Age = t.Age, MyModel.Gender = t.Gender, ...,
```

```
MyModel.Products.ProductID=t.Products.ProductID
```

30 Раскрытое изобретение предусматривает множество способов использования запросов многомерных выражений (например, MDX) в качестве ввода для операторов расширения глубинного анализа данных (например, DMX). В качестве примера, а не ограничения, в следующем описании используется MDX в качестве одного способа, которым может быть выполнено настоящее изобретение. На этапе 300 первым

35 способом является простая замена реляционного запроса запросом MDX. Например, запрос:

```
SELECT CustomerID, Age, Gender, Occupation, HomeOwner, TotalSales, MemberCard
FROM MyTable
```

40 может быть выражен в кубе MDX как

```
SELECT Measures.TotalSales ON COLUMNS,
```

```
Customers.Members DIMENSION PROPERTIES CustomerID, Age, Gender, Occupation,
HomeOwner, MemberCard ON ROWS
```

```
FROM MyCube
```

45 В другом примере, следующий запрос:

```
SELECT CustomerID, ProductID From ProductFacts
```

может быть выражен как

```
SELECT. ON COLUMNS,
```

```
NON EMPTY CROSSJOIN (Customers, Products)
```

```
DIMENSION PROPERTIES Customer.CustomerID, Products. ProductID ON ROWS
```

```
FROM MyCube
```

50 Таким образом, в этой первой форме выражения MDX могут быть подставлены

вместо реляционных запросов.

На этапе 302 вторая форма предусматривает удаление конструкции SHAPE (форма) при помощи использования преимущества внутренней многомерной структуры куба. Приведенные выше оформленные реляционные запросы могут быть затем записаны в MDX как:

```
SELECT Measures. TotalSales ON COLUMNS,
Customers. Members DIMENSION PROPERTIES CustomerID, Age, Gender, Occupation,
HomeOwner, MemberCard ON ROWS,
NON EMPTY Products. Members DIMENSION PROPERTIES ProductID on PAGES FROM
MyCube
```

Дополнительные вложенные таблицы могут быть упорядочены по дополнительным осям.

На этапе 304 третья форма изобретения включает в себя операторы, которые связаны только по имени, например SELECT, PREDICTION и JOIN. В этой форме процессор расширения глубинного анализа данных использует преимущество структуры куба, что позволяет упростить запросы, расширяя предложение ON для извлечения информации из запроса MDX. Например:

```
SELECT Predict(MemberCard) FROM MyModel PREDICTION JOIN
SELECT Measures.TotalSales ON COLUMN,
Customers.Members ON ROWS,
NON EMPTY Products ON PAGES
FROM MyCube as t
ON MyModel.Age = t.Customers.Age,
MyModel.Gender = t.Customers.Gender,
MyModel.Occupation = t.Customers.Occupation,
...
MyModel.TotalSales = t.TotalSales,
MyModel.Products.ProductID = t.Products.ProductID
```

По отношению к созданию модели настоящее изобретение утверждает, что до создания тип источника данных не известен и не установлен до фазы подготовки. В известных системах «тип» модели определяется до создания, причем типом может быть модель с реляционным источником или источником OLAP.

Модель глубинного анализа данных может быть подготовлена на основе произвольного источника данных, безотносительно к его реляционному или многомерному происхождению. Связывание столбцов поддерживается согласно явному порядку столбцов как в многомерном, так и в реляционном источнике в отличие от известных систем, в которых связывание столбцов поддерживается согласно соответствию имен между моделью глубинного анализа данных и кубом OLAP.

Модель глубинного анализа данных может принимать в качестве источника данных для прогнозирования произвольный источник данных, безотносительно к его реляционному или многомерному происхождению. Традиционно, прогнозирование в отношении данных в кубе OLAP выполняется в MDX и только с использованием модели глубинного анализа данных, подготовленной для того же куба. Раскрытая здесь архитектура позволяет прогнозировать с использованием DMX и позволяет использовать куб OLAP в качестве источника для прогнозирования на основе любой модели, безотносительно к тому, каким образом она была создана или подготовлена.

Теперь обратимся к Фиг.4, где проиллюстрирована структурная схема компьютера,

предназначенного для выполнения раскрытой архитектуры. Для того чтобы обеспечить дополнительный контекст для различных аспектов настоящего изобретения, Фиг.4 и нижеследующее описание предназначены для краткого общего описания соответствующей вычислительной среды 400, в которой могут быть реализованы различные аспекты настоящего изобретения. Несмотря на то что изобретение было описано выше в общем контексте машиноисполняемых инструкций, которые могут выполняться на одном или более компьютерах, специалист в данной области техники должен понимать, что изобретение может быть также реализовано в комбинации с другими программными модулями и/или как комбинация аппаратных и программных средств.

В общем случае программные модули включают в себя процедуры, программы, компоненты, структуры данных и так далее, которые выполняют определенные задачи или реализуют определенные абстрактные типы данных. Более того, специалисту в данной области техники должно быть понятно, что способы изобретения могут быть реализованы на практике с использованием других конфигураций компьютерной системы, включая однопроцессорные или многопроцессорные компьютерные системы, мини-компьютеры, универсальные вычислительные системы, а также персональные компьютеры, карманные вычислительные устройства, основанную на микропроцессорах или программируемую бытовую электронную аппаратуру, и тому подобное, каждая из которых может быть функционально связана с одним или более ассоциированными устройствами.

Проиллюстрированные аспекты изобретения также могут быть реализованы на практике в распределенной вычислительной среде, где определенные задачи выполняются удаленными устройствами обработки, которые связаны через сеть связи. В распределенной вычислительной среде программные модули могут быть расположены как на локальных, так и на удаленных запоминающих устройствах.

Компьютер, как правило, имеет набор машиночитаемых носителей. Машиночитаемые носители могут быть любыми имеющимися носителями, к которым компьютер может осуществить доступ, и включают в себя как энергозависимые, так и энергонезависимые носители, как съемные, так и несъемные носители. Для примера, но не для ограничения, машиночитаемые носители могут включать в себя компьютерные носители информации и среды связи. Компьютерные носители информации включают в себя как энергозависимые, так и энергонезависимые носители, как съемные, так и несъемные носители, реализованные с использованием любого способа или технологии для хранения информации, такой как машиночитаемые инструкции, структуры данных, программные модули или другие данные. Компьютерные носители информации включают в себя, но не ограничены перечисленным, оперативное запоминающее устройство (ОЗУ, RAM), постоянное запоминающее устройство (ПЗУ, ROM), электрически стираемое программируемое ПЗУ (EEPROM), флэш-память или память, реализованную по другой технологии, ПЗУ на компакт-диске (CD-ROM), универсальные цифровые диски (DVD) или другие оптические запоминающие устройства, магнитные кассеты, накопители на магнитных дисках или другие магнитные запоминающие устройства, или любой другой носитель, который может быть использован для хранения необходимой информации и к которой компьютер может осуществить доступ.

Среда связи, как правило, воплощает машиночитаемые инструкции, структуры данных, программные модули или другие данные в модулированном

информационном сигнале, таком как сигнал несущей или другой механизм передачи, и включает в себя любую среду доставки информации. Термин «модулированный информационный сигнал» означает сигнал, одна или более характеристик которого установлены или изменены так, чтобы закодировать информацию в этом сигнале. Для примера, но без ограничения общности, среды связи включают в себя проводные среды, например проводную сеть или прямое проводное соединение, и беспроводные среды, такие как акустические, радиочастотные (RF), инфракрасные (IR) и другие беспроводные среды связи. Комбинации любых выше приведенных носителей и сред также попадают под определение машиночитаемых носителей.

Снова обратимся к Фиг.4, где проиллюстрирована примерная вычислительная среда 400 для реализации различных аспектов настоящего изобретения, которая включает в себя компьютер 402, причем компьютер 402 включает в себя процессор 404, системную память 406 и системную шину 408. Системная шина 408 связывает системные компоненты, включая, но не в ограничительном смысле, системную память 406, с процессором 404. Процессор 404 может быть любым из различных имеющихся на рынке процессоров. Также сдвоенные микропроцессоры и другие многопроцессорные архитектуры могут быть использованы в качестве процессора 404.

Системная шина 408 может относиться к любому типу из различных типов шинных структур, которая может быть далее взаимно соединена с шиной памяти (с контроллером памяти или без него), периферийной шиной и локальной шиной, с использованием любой из возможных шинных архитектур. Системная память 406 включает в себя постоянное запоминающее устройство 410 (ПЗУ) и оперативное запоминающее устройство 412 (ОЗУ). Базовая система ввода/вывода (BIOS) сохранена в энергонезависимой памяти 410, такой как постоянное (ПЗУ), стираемое программируемое ПЗУ (EPROM), EEPROM, и содержит базовые процедуры, помогающие передавать информацию между элементами компьютера 402, например, в процессе запуска. ОЗУ 412 может также включать в себя высокоскоростное ОЗУ, такое как статическое ОЗУ для кэширования данных.

Компьютер 402 дополнительно включает в себя внутренний накопитель 414 на жестких дисках (HDD) (например, EIDE, SATA), причем внутренний накопитель 414 на жестких дисках (HDD) может быть также сконфигурирован для внешнего использования в соответствующем корпусе (не показан на чертеже), магнитный дисковод (FDD) 416 (например, для записи на сменную дискету 418 или чтения с нее) и оптический дисковод 420 (например, для чтения диска 422 CD-ROM, или для чтения другого оптического накопителя с высокой плотностью записи, например DVD, или записи на него). Накопитель 414 на жестких дисках, магнитный дисковод 416 и оптический дисковод 420 могут быть соединены с системной шиной 408 при помощи интерфейса 424 накопителя на жестких дисках, интерфейса 426 магнитного дисковода и интерфейса 428 оптического дисковода соответственно. Интерфейс 424 для реализаций внешнего накопителя или дисковода включает в себя, по меньшей мере, один или оба варианта из технологий интерфейса, соответствующих универсальной последовательной шине (USB) и шине IEEE 1394.

Дисководы и накопители и относящиеся к ним машиночитаемые носители обеспечивают энергонезависимое хранение данных, структур данных, машиночитаемых инструкций и тому подобного. Для компьютера 402 дисководы, и накопители, и носители обеспечивают хранение любых данных в подходящем цифровом формате. Хотя описание машиночитаемых носителей, приведенное выше,

ссылается на накопитель на жестких дисках, сменную магнитную дискету и сменный оптический диск, например CD или DVD, специалистам в данной области техники должно быть понятно, что другие носители, которые могут быть прочитаны компьютером, такие как накопители zip, магнитные кассеты, платы флэш-памяти, картриджи и тому подобное, также могут быть использованы в примерной вычислительной среде, и более того, любой из этих носителей может содержать машиноисполняемые инструкции для выполнения способов настоящего изобретения.

Некоторое количество программных модулей может сохраниться на накопителях и в ОЗУ 612, включая операционную систему 430, одну или более прикладных программ 432, другие программные модули 434 и данные 436 программ. Все или часть из операционной системы, прикладных программ, модулей и/или данных также могут быть кэшированы в ОЗУ 412.

Должно быть принято во внимание, что настоящее изобретение может быть реализовано с различными имеющимися на рынке операционными системами или с комбинацией операционных систем.

Пользователь может вводить команды и информацию в компьютер 402 посредством одного или более проводных/беспроводных устройств ввода, таких как клавиатура 438 и координатно-указательное устройство, например мышь 440. Другие устройства ввода (не показаны на чертеже) могут включать в себя микрофон, инфракрасный пульт дистанционного управления, джойстик, игровую панель, световое перо, сенсорный экран или тому подобное. Эти и другие устройства ввода часто соединены с процессором 404 посредством интерфейса 442 устройств ввода, который связан с системной шиной 408, но могут быть соединены при помощи других интерфейсов, таких как параллельный порт, последовательный порт IEEE 1394, игровой порт, порт универсальной последовательной шины (USB), инфракрасный интерфейс и так далее.

Монитор 444 или устройство отображения информации другого типа также подсоединен к системной шине 408 через интерфейс, например видеоадаптер 446. В дополнение к монитору 444 компьютер, как правило, включает в себя другие периферийные устройства вывода (не показаны на чертеже), такие как динамики, принтер и так далее.

Компьютер 402 может функционировать в сетевой среде посредством логических соединений через проводные и/или беспроводные линии связи с одним или более удаленными компьютерами, такими как удаленный компьютер 448. Удаленный компьютер 448 может быть рабочей станцией, сервером, маршрутизатором, персональным компьютером, портативным компьютером, игровым устройством с микропроцессором, одноранговым устройством или другим общим сетевым узлом и, как правило, включает в себя многие или все из элементов, описанных по отношению к компьютеру 402, хотя, в целях краткости, только запоминающее устройство 450 показано на чертеже. Изображенные логические соединения включают в себя проводные/беспроводные соединения к локальной сети (LAN) 452 и/или к сети большего размера, например глобальной сети (WAN) 454. Такие сетевые среды LAN и WAN широко распространены в офисах и компаниях и обеспечивают сети масштаба предприятия, такие как корпоративные локальные сети, все из которых могут быть соединены с глобальной сетью связи, например Интернет.

При работе в локальной сетевой среде компьютер 402 соединяется с локальной сетью 452 через проводной и/или беспроводной сетевой интерфейс или адаптер 456. Адаптер 456 может обеспечивать проводную или беспроводную связь с LAN 452,

которая может также включать в себя беспроводную точку доступа, размещенную в ней, для связи с беспроводным адаптером 456. При работе в глобальной сетевой среде компьютер 402 может включать в себя модем 458, или быть соединенным с сервером связи в LAN, или иметь другие средства для установления соединения через
5 глобальную сеть 454, например через Интернет. Модем 458, который может быть внешним или внутренним и проводным или беспроводным устройством, подсоединен к системной шине 408 через интерфейс последовательного порта 442. В сетевой среде программные модули, изображенные как относящиеся к компьютеру 402, или их
10 части, могут быть сохранены на удаленном запоминающем устройстве 450. Следует понимать, что описанные сетевые соединения являются примерными и что могут быть применены другие средства для установления линии связи между компьютерами.

Компьютер 402 выполнен с возможностью осуществления связи с любыми
15 беспроводными устройствами или сущностями, функционально расположенными в беспроводной сети связи, например принтером, сканером, настольным и/или портативным компьютером, портативным цифровым помощником, спутником связи, любой частью оборудования или местоположением, ассоциированным с дескриптором, обнаруживаемым по беспроводной сети (например, киоском, стендом
20 новостей, комнатой отдыха), и телефоном. Это включает в себя использование, по меньшей мере, беспроводных технологий Wi-Fi и Bluetooth™. Таким образом, связь может быть predetermined структурой, как в традиционных сетях, или просто устанавливаемым по необходимости соединением «точка-точка» между, по меньшей мере, двумя устройствами.

Wi-Fi или беспроводная сеть высокой точности воспроизведения позволяет
25 установить соединение с Интернет с дивана в доме, кровати в отеле или из конференц-зала на работе без проводов. Wi-Fi - это беспроводная технология, подобная сотовым телефонам, которая делает возможным отправку и получение данных в помещении
30 или на улице, такими устройствами, как компьютеры, в любом месте, находящемся в радиусе действия базовой станции. Сети Wi-Fi используют беспроводные технологии, называемые IEEE 802.11 (a, b, g и так далее), для обеспечения безопасной, надежной и быстрой беспроводной связи. Сеть Wi-Fi может быть использована для соединения
35 компьютеров между собой, с Интернет, с проводными сетями (которые используют IEEE 802.3 или Ethernet). Сети Wi-Fi функционируют в свободных от лицензии радиочастотных диапазонах 2,4 ГГц и 5 ГГц, со скоростью передачи данных в 11 Мбит/с (802.11b) или 54 Мбит/с (802.11a) или с использованием продуктов, которые содержат оба диапазона (двухдиапазонные), так что сети могут обеспечить
40 реальную производительность, сопоставимую с базовыми проводными сетями 10BaseT Ethernet, используемыми во многих офисах.

Обратимся теперь к Фиг.5, где проиллюстрирована схематичная структурная схема
примерной вычислительной среды 500 в соответствии с настоящим изобретением. Система 500 включает в себя один или более клиентов 502. Клиенты 502 могут быть
45 аппаратными и/или программными (например, потоки, процессы, вычислительные устройства). Клиенты 502 могут вмещать cookies (создаваемые серверами фрагменты данных о предыстории обращений к этим серверам) и/или ассоциированную контекстуальную информацию, применяя настоящее изобретение, для примера.
50 Система 500 также включает в себя один или более серверов 504. Серверы 504 могут также быть аппаратными или программными (например, потоки, процессы, вычислительные устройства). Серверы 504 могут вмещать потоки для выполнения преобразований, применяя настоящее изобретение, для примера. Один возможный

сеанс связи между клиентом 502 и сервером 504 может иметь форму пакета данных, адаптированного для передачи между двумя и более компьютерными процессами. Пакет данных может включать в себя cookies и/или ассоциированную контекстуальную информацию, для примера. Система 500 включает в себя сетевую инфраструктуру 506 (например, глобальную сеть связи, такую как Интернет), которая может быть использована для обеспечения связи между клиентом(и) 502 и сервером(ами) 504.

Связь может быть обеспечена с использованием проводных (включая оптоволоконные) и беспроводных технологий. Клиенты 502 функционально связаны с одним или более клиентскими устройствами 508 хранения данных, которые могут быть использованы для хранения информации, локальной по отношению к клиенту(ам) 502 (например, cookies и/или ассоциированную контекстуальную информацию). Аналогичным образом, серверы 504 функционально связаны с одним или более серверными устройствами 510 хранения данных, которые могут быть использованы для хранения информации, локальной по отношению к серверу(ам) 504.

Описание, приведенное выше, включает в себя примеры настоящего изобретения. Конечно, невозможно описать каждую мыслимую комбинацию компонент или методик в целях описания настоящего изобретения, но специалист в данной области техники может понять, что возможно много дополнительных комбинаций и вариаций настоящего изобретения. Соответственно, настоящее изобретение предназначено для того, чтобы охватить все такие альтернативы, модификации и вариации, которые попадают в пределы объема и сущности, определяемых нижеследующей формулой изобретения. Более того, отметим, что термин «включает в себя» в той мере, как он используется и в детальном описании и в формуле изобретения, предназначен для обозначения включения, подобно термину «содержит» или «содержащий», в случае использования его в качестве вводного слова в формуле.

Формула изобретения

1. Система, обеспечивающая глубинный анализ данных в отношении многомерного куба данных, содержащая многомерный куб данных; и

компонент глубинного анализа данных для обеспечения глубинного анализа данных в отношении многомерного куба данных, причем компонент глубинного анализа данных выполнен с возможностью объединять многомерные выражения и расширения глубинного анализа данных во ввод для создания модели глубинного анализа данных, согласно которой должен выполняться глубинный анализ данных в отношении многомерного куба данных,

при этом компонент глубинного анализа данных объединяет многомерные выражения и расширения глубинного анализа данных посредством использования внутренней многомерной структуры многомерного куба данных для записи реляционного запроса в виде многомерного выражения.

2. Система по п.1, в которой многомерным кубом данных является куб интерактивной аналитической обработки (OLAP).

3. Система по п.1, в которой тип исходных данных модели глубинного анализа данных может быть задан либо во время создания, либо во время подготовки.

4. Система по п.1, в которой модель глубинного анализа данных независима от любой структуры источника, по которой она подготовлена.

5. Система по п.4, дополнительно содержащая средство для подготовки модели

глубинного анализа данных на основе произвольного источника данных, который является одним из реляционного и многомерного.

6. Система по п.1, в которой модель глубинного анализа данных подготовлена на основе произвольного источника данных.

7. Система по п.6, в которой произвольный источник данных является одним из реляционного и многомерного.

8. Система по п.1, в которой модель глубинного анализа данных подготовлена с использованием связывания столбцов, которое поддерживается согласно явному порядку столбцов в по меньшей мере одном из реляционного источника данных и многомерного источника данных.

9. Система по п.1, в которой модель глубинного анализа данных принимает произвольный источник данных для процесса прогнозирования.

10. Система по п.9, в которой произвольный источник данных является одним из реляционного и многомерного.

11. Система по п.9, в которой процесс прогнозирования осуществляется с использованием глубинного анализа данных.

12. Система по п.1, в которой ввод является перезаписью оформленного запроса в виде многомерного выражения для вложенной таблицы, причем оформленный запрос оформляется с помощью директивы SHAPE.

13. Система по п.1, в которой модель глубинного анализа данных подготовлена на основе реляционного источника данных, и эта модель применяется для прогнозирования в отношении куба OLAP.

14. Система по п.1, в которой модель глубинного анализа данных подготовлена на основе источника данных OLAP, и эта модель применяется для прогнозирования реляционных данных.

15. Система по п.1, в которой модель глубинного анализа данных принимает произвольный источник данных для процесса прогнозирования с использованием DMX (расширений глубинного анализа данных), причем этот произвольный источник является одним из реляционного и многомерного.

16. Считываемый компьютером носитель, на котором сохранены исполняемые компьютером инструкции для реализации системы по п.1.

17. Компьютер, который использует систему по п.1.

18. Реализуемый компьютером способ глубинного анализа данных в отношении куба интерактивной аналитической обработки (OLAP), включающий в себя этапы, на которых

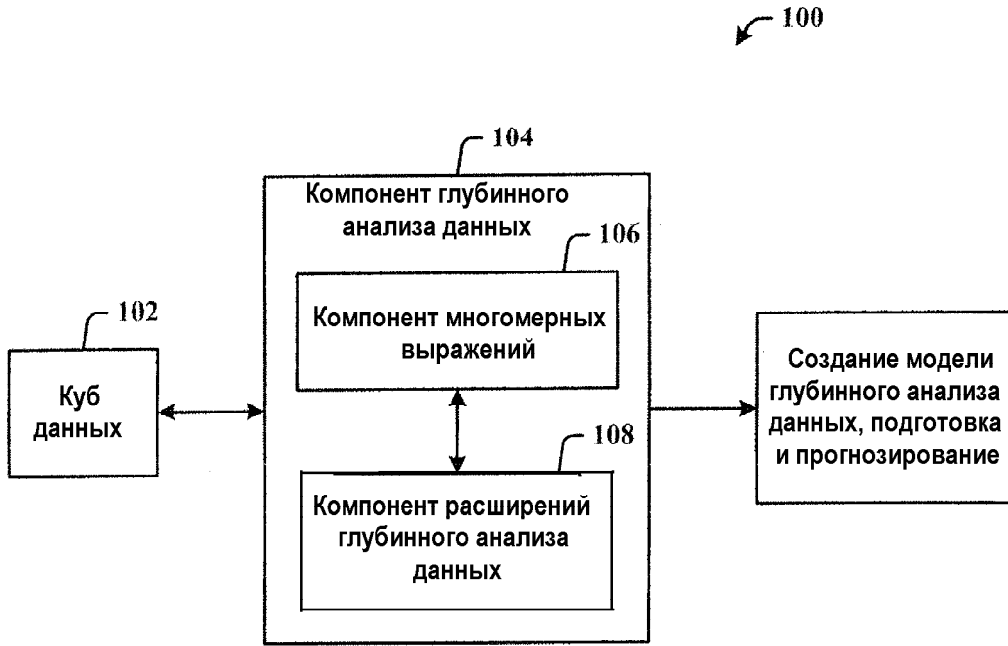
принимают куб OLAP;

создают модель глубинного анализа данных на основе ввода, где многомерные выражения и расширения глубинного анализа данных объединены посредством использования внутренней многомерной структуры куба OLAP для записи реляционного запроса в виде многомерного выражения; и

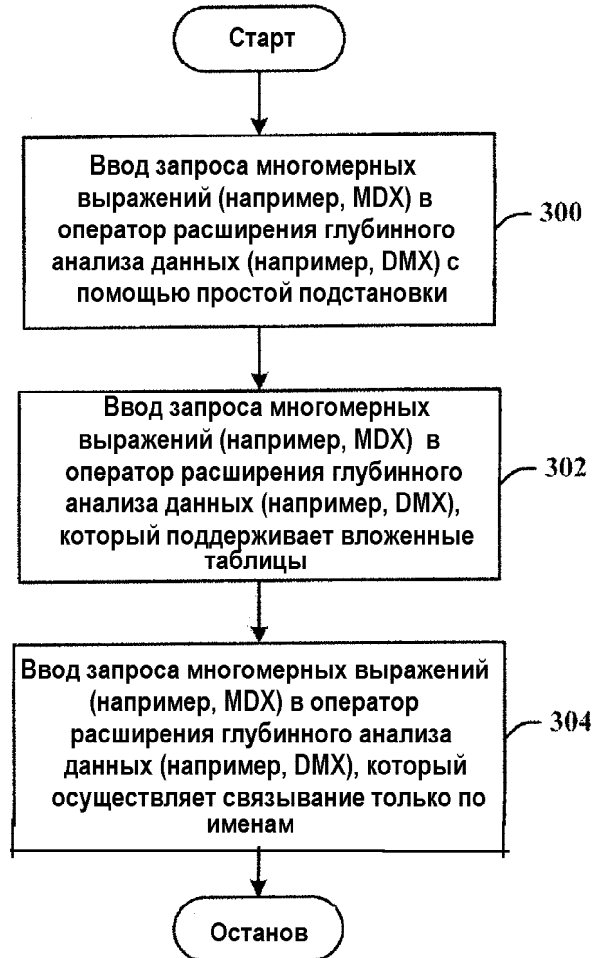
выполняют прогнозирования данных по отношению к данным, содержащимся в кубе OLAP посредством глубинного анализа данных в отношении куба OLAP согласно модели глубинного анализа данных.

19. Способ по п.18, в котором тип исходный данных модели глубинного анализа определяется на фазе ее подготовки.

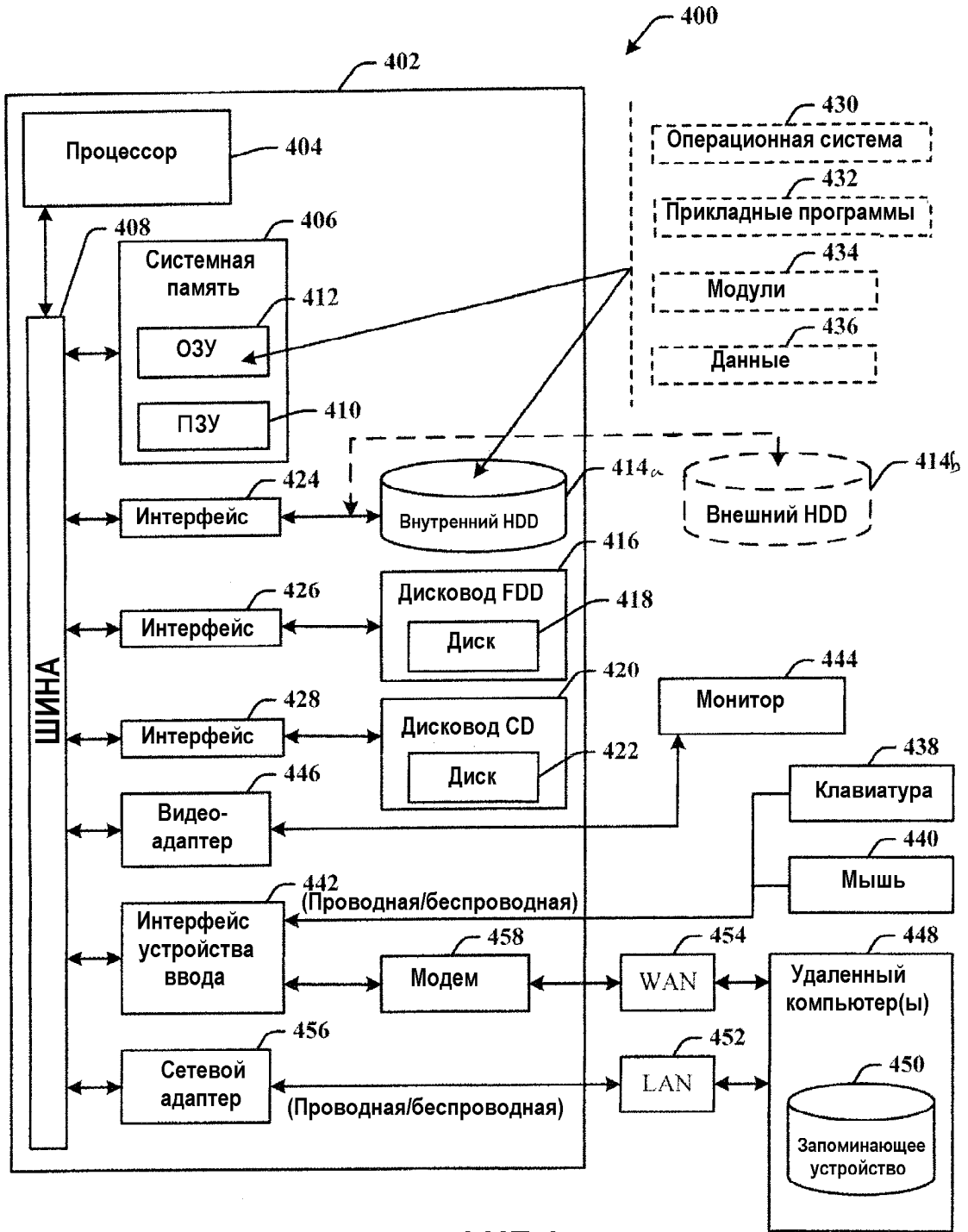
20. Способ по п.18, дополнительно включающий в себя этап, на котором подготавливают модель глубинного анализа на основе произвольного источника данных, при этом источник данных является одним из реляционного и многомерного.



ФИГ.1



ФИГ.3



ФИГ.4



ФИГ.5