

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2017 年 9 月 28 日 (28.09.2017)



(10) 国际公布号
WO 2017/162085 A1

- (51) 国际专利分类号:
G06F 19/00 (2011.01)
- (21) 国际申请号: PCT/CN2017/076876
- (22) 国际申请日: 2017 年 3 月 16 日 (16.03.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610180084.0 2016 年 3 月 25 日 (25.03.2016) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Cayman (KY)。
- (72) 发明人: 雷宗雄 (LEI, Zongxiong); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。
- (74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY

AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: DATA ANALYSIS METHOD AND APPARATUS

(54) 发明名称: 数据分析方法及装置

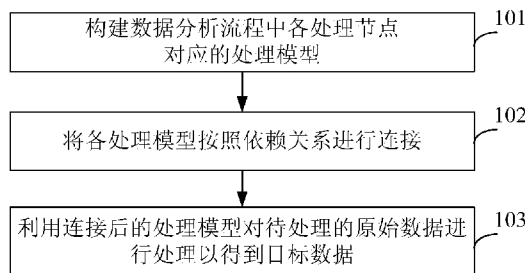


图 1

- 101 CONSTRUCTING PROCESSING MODELS CORRESPONDING TO VARIOUS PROCESSING NODES IN A DATA ANALYSIS PROCESS
- 102 CONNECTING THE PROCESSING MODELS ACCORDING TO A DEPENDENCE RELATIONSHIP
- 103 USING THE CONNECTED PROCESSING MODELS TO PROCESS ORIGINAL DATA TO BE PROCESSED SO AS TO OBTAIN TARGET DATA

(57) Abstract: Provided are a data analysis method and apparatus. The method comprises: constructing processing models corresponding to various processing nodes in a data analysis process; successively connecting the processing models according to a dependence relationship; and using the connected processing models to process original data to be processed so as to obtain target data. In the present invention, various processing models are directly connected according to a dependence relationship, an output of the previous processing model can be directly input into the next processing model, and intermediate data is no longer loaded, thereby saving resources, omitting a procedure of loading intermediate data and improving data analysis efficiency.

(57) 摘要: 本发明提供了一种数据分析方法及装置, 通过构建数据分析流程中各处理节点对应的处理模型, 将各处理模型按照依赖关系依次连接, 利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。本发明中各处理模型按照依赖关系直接连接, 上一个处理模型的输出可以直接输入到下一个处理模型中, 中间数据不再进行落地, 节省资源, 并且省略中间数据的加载过程, 提高了数据分析效率。

WO 2017/162085 A1



本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

数据分析方法及装置

本申请要求 2016 年 03 月 25 日递交的申请号为 201610180084.0、发明名称为“数据分析方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本发明属于数据处理领域，尤其涉及一种数据分析方法及装置。

背景技术

数据分析过程中往往包含大量的数据清洗、特征平滑、特征规范化、特征提炼、特征选择等处理节点。在传统数据分析中，对于每一个处理节点，都会产生对应的中间数据，生成的中间数据需要进行落地处理，由于中间数据用于作为下一个处理节点的上游数据依赖，下一个数据节点需要加载上一个数据节点落地后的中间数据作为其输入。

在传统数据分析中，上下两个处理节点通过落地的中间数据进行链接，数据每经过一个处理节点就需要做一次数据加载和落地，对于海量数据或者复杂的数据分析过程来说，中间数据的数据量巨大，不仅导致计算资源、输入输出（IO）的资源浪费，而且数据加载和落地也会严重影响数据分析的效率。

发明内容

本发明提供一种数据分析方法及装置，用于解决现有数据分析中数据每经过一个处理节点就需要做一次数据加载和落地，不仅导致资源费，且影响数据分析的效率的问题。

为了实现上述目的，本发明提供了一种数据分析方法，包括：

构建数据分析流程中各处理节点对应的处理模型；

将各处理模型按照依赖关系依次连接生成处理模块；

利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。

25 为了实现上述目的，本发明提供了一种数据分析装置，包括：

构建模块，用于构建数据分析流程中各处理节点对应的处理模型；

连接模块，用于将各处理模型按照依赖关系进行连接；

处理模块，用于利用连接后的处理模型对待处理的原始数据进行以处理以得到目标数据。

30 本发明提供的数据分析方法及装置，通过构建数据分析流程中各处理节点对应的处

理模型，将各处理模型按照依赖关系依次连接，利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。本发明中各处理模型按照依赖关系直接连接，上一个处理模型的输出可以直接输入到下一个处理模型中，中间数据不再进行落地，节省资源，并且省略中间数据的加载过程，提高了数据分析效率。

5

附图说明

图 1 为本发明实施例一的数据分析方法的流程示意图；

图 2 为本发明实施例二的数据分析方法的流程示意图；

图 3 为本发明实施例二的 DAG 图；

10 图 4 为本发明实施例二的数据分析方法的应用示例示意图之一；

图 5 为本发明实施例三的数据分析方法的流程示意图；

图 6 为本发明实施例四的数据分析装置的结构示意图；

图 7 为本发明实施例五的数据分析装置的结构示意图。

15 具体实施方式

下面结合附图对本发明实施例提供的评估指标获取方法及装置进行详细描述。

实施例一

如图 1 所示，其为本发明实施例一的数据分析方法的流程示意图。该数据分析方法包括以下步骤：

20 S101、构建数据分析流程中各处理节点对应的处理模型。

首先，对设定的数据分析流程进行分析，获取每个数据分析流程中各处理节点，在获取到各处理节点后，能够依据各处理节点的功能，构建每个处理节点相应的处理模型。

例如，该数据分析流程中包括有特征平滑、特征归一化、特征提取、特征选择等处理节点，这些处理节点都有特定的处理功能，这些处理节点能够对输入的数据进行处理以得到一个相应的结果，这个结果在数据分析流程中就是中间数据。例如特征归一化用于将原始数据基于每个特征的均值和标准差进行归一化处理，归一化处理后的数据就是该特征归一化的中间数据。本实施例中，为了避免产生中间数据，将特征归一化这个处理节点进行模型化，特征归一化处理模型对输入的数据具有数据转换功能，该处理模型能够记录每个特征均值和标准差，能够对原始数据进行转换。

30 S102、将各处理模型按照依赖关系进行连接。

在对数据分析流程进行分析时，需要获取到各处理模型的依赖关系，当构建出各处理节点对应的处理模型后，按照处理节点之间的依赖关系，将各处理模型依次有序地串联起。为了使各处理模型能够直接连接，需要设置有数据接口，本实施例中，各处理模型的数据接口是统一的，经过数据接口依据依赖关系将各处理模型进行顺次连接后，数据分析流程就可以转换成一个有序的执行逻辑。

S103、利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。

当各处理模型的连接后，将能够有序地执行逻辑，可以将待处理的原始数据输入到连接后的各处理模型中，原始数据首先进入处于执行逻辑顶端的处理模型中，然后经过处理模型后的数据依次进入下一个处理模型，直到进入到处于执行逻辑尾部的处理模型中，该处理模型最后输出的数据就是目标数据。

本实施例提供的数据分析方法，将处理模型按照依赖关系连接，上一个处理模型的输出可以直接通过数据接口输入到下一个处理模型中，中间数据不再进行落地，节省资源，而且由于上一个处理模型生成的中间数据直接进入下一个处理模型，避免了中间数据的加载过程，提高了数据分析效率。

15 实施例二

如图 2 所示，其为本发明实施例二的数据分析方法的流程示意图。该数据分析方法包括以下步骤：

S201、获取数据分析流程的无回路有向 DAG 图。

数据分析流程是由一系列的处理节点组成的，对数据分析流程进行特征分析，能够获取到该数据分析流程的无回路有向图（Directed Acyclic Graph，简称 DAG 图），DAG 图可以串联出一系列有序的处理节点。

S202、解析 DAG 图获取数据分析流程的相关信息。

其中，所述相关信息中包括：处理节点的逻辑功能、处理节点之间的依赖关系以及各处理模型存储地址。

对 DAG 图进行解析获取到该数据分析流程的相关信息，其中，该相关信息中包括数据分析流程中包括的处理节点的逻辑功能、处理节点之间的依赖关系和各处理模型存储地址。相关信息中还可以包括输入数据信息、输出数据信息以及用户配置参数等。这些相关信息可以生成一个具有节点依赖关系的可扩展标记语言（Extensible Markup Language，简称 XML）文件，将该 XML 文件保存到数据库备份中并提交到后台。

S203、根据所述相关信息中每个处理节点的逻辑功能构建对应的处理模型。

在获取到每个处理节点的逻辑功能后，根据处理节点的逻辑功能构建对应的处理模型。例如，一个数据缩放处理节点用于将大于设定范围的数据进行缩小，将小于预设范围的数据进行放大，根据这个处理节点的逻辑功能就可以构建相应的数据缩放模型。

S204、根据所述相关信息中处理节点之间的依赖关系将各处理模型通过数据接口连接。

在生成各处理节点对应的处理模型后，根据所述相关信息中处理节点之间的依赖关系，将各处理模型通过数据接口直接连接。具体地，后台接收到对 DAG 图解析得到的 XML 文件后，能够获取到分析 DAG 图中处理节点之间的依赖关系。后台根据处理节点之间的依赖关系对各处理模型的代码程序自动组装，即后台根据处理节点之间的依赖关系对各处理模型的代码程序进行 DAG 化，将组装后的代码程序保存并编译成可运行文件。然后基于已经设计好的数据接口，将各处理模块的代码依次组装，组成完成后，对每个处理模型进行初始化。为了实现处理模型的直接连接，对数据接口进行了统一化处理，从而可以方便地把一系列的代码组合起来，在根据依赖关系进行串联后，数据分析流程转换成程序层面的一个有序的执行逻辑。

一般情况下，代码程序的重新组合会引入新的程序缺陷（bug），导致部署时风险较大，并且需要进行再次测试，导致资源的重复。本实施例中，将数据分析流程在代码层面进行了 DAG 化，能够减少程序缺陷的数量，可以一体化将各处理模型的可运行代码程序打包直接部署到线上环境，这将极大的减少线上部署的风险。

S205、根据所述相关信息中各处理模型存储地址将各处理模型进行落地存储。

在生成处理模型后，为了避免处理模型的重复计算，可以根据相关信息中各处理模型的存储地址，将各处理模型进行落地存储，提高处理模型的复用率。实际应用中，处理模型的大小规模远远小于中间数据的大小，不仅能够节省资源，而且有利于数据分析的效率。

S206、将待处理的原始数据输入到连接后的处理模型中进行处理以得到目标数据。

在生成将处理模型进行连接后，将待处理原始数据进行输入，经过连接后的处理模型的处理，得到最终的目标数据。本实施例中，对原始数据进行处理的过程在内存中完成，进而不用落地每个处理模型的中间数据。

为了更好地理解上述本实施例提供的数据分析方法，下面举例进行说明：

对原始数据的数据分析流程包括以下处理节点：数据规范、数据缩放和数据平滑。

对该数据分析流程进行分析，得到该数据分析流程的 DAG 图，如图 3 所示，在该 DAG 图

中每个处理节点输出的数据为中间数据，例如数据缩放处理节点输出的中间数据为缩放后数据，数据平滑处理节点输出的中间数据为平滑后数据。

对 DAG 图进行解析，可以获取到该数据分析流程中各处理节点之间的相关信息，其中相关信息中包括：处理节点的逻辑功能、处理节点之间的依赖关系以及各处理模型存储地址。在该示例中，处理节点之间的依赖关系为：数据规范依赖数据平滑，数据平滑依赖数据缩放。

为了在数据分析过程中不再产生中间数据，避免中间数据的落地和加载，根据数据分析流程中各处理节点的逻辑功能，为每个处理节点构建对应的处理模型，这些处理模型具有相应的数据转换功能。具体包括：数据缩放模型、数据平滑模型和数据规范模型。

进一步地，按照处理节点之间的依赖关系，将各处理模型通过数据接口连接起来，如图 4 所示。在将处理模型连接后，形成了一个有序的执行逻辑，将原始数据输入到连接后的处理模型中，这样上述连接后的处理模型构成的执行逻辑就开始对原始数据进行处理，得到最终的目标数据。各处理模型对原始数据的处理过程可以在内存中完成，避免产生中间数据。进一步地，可以将构建的这些处理模型进行落地，可以根据用户配置的存储地址进行存储，以便于这些处理模型的复用。

本实施例提供的数据分析方法，通过获取数据分析流程的 DAG 图，对 DAG 图进行解析，根据解析结果构建各处理模型，以及将各处理模型按照依赖关系连接，利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。本实施例中处理模型之间直接连接，上一个处理模型的输出可以直接输入到下一个处理模型中，中间数据不再进行落地，节省资源，并且省略中间数据的加载过程，提高了数据分析效率。

实施例三

如图 5 所示，其为本发明实施例三的数据分析方法的流程示意图。在上述实施例的基础之上，利用连接后的处理模型对待处理原始数据进行处理以得到目标数据包括以下步骤：

S301、对原始数据进行数据校验。

在获取到原始数据后，需要对原始数据进行数据校验，首先检测用户是否存有相应的处理模型，如果存储有相应的处理模型，判断该原始数据是否未做过更改，具体地，根据相关信息中的输入数据信息和输出数据信息，如果原始数据未做过更改，说明不需要对已存储的处理模型进行更新，只需要直接获取到存储的各处理模型，则执行步骤 S302，否则执行步骤 S303。

S302、当原始数据通过数据校验时，依据所述相关信息中各处理模型的存储地址获取各处理模型。

当原始数据通过数据校验时，说明存储的处理模型不需要更新，则可以根据相关信息中各处理模型的存储地址获取到各处理模型。

5 S303、当原始数据未通过数据校验时，重新构建各处理模型。

在对原始数据进行数据校验时，当判断出原始数据做过更改时，需要对存储的各处理模型进行更新，如果未检测到已存储的处理模型，则对原始数据的数据分析流程进行分析，构建相应的处理模型。

S304、将各处理模型按照依赖关系进行连接。

10 将各处理模型按照处理节点之间的依赖关系进行连接，形成一个有序地执行逻辑。

S305、将原始数据进行格式转换得到输入数据。

为了保证数据接口的统一性，需要对原始数据机械能格式转换，生成格式统一的输入数据，本实施例中，将原始数据统一转换成向量（Vector）或者矩阵（Matrix）格式的输入数据。

15 S306、将输入数据输入到连接后的处理模型中得到目标数据。

将处理模型按照依赖关系进行连接，形成了一个有序的执行逻辑，将原始数据输入到连接后的处理模型中，这样上述由连接的处理模型构成的执行逻辑就开始对原始数据进行处理，得到最终的目标数据。

S307、根据预设的效果验证条件对目标数据进行效果验证。

20 用户可以根据自身的需求，预设一定的效果验证条件，根据设定的效果验证条件对目标数据的处理效果进行验证。例如，提供数据处理前后的效果对比，用户能够非常直观地看到数据在处理前和处理后的效果对比。

实施例四

25 如图6所示，其为本发明实施例四的数据分析装置的结构示意图。该数据分析装置包括：构建模块11、连接模块12和处理模块13。

其中，构建模块11，用于构建数据分析流程中各处理节点对应的处理模型。

连接模块12，用于将各处理模型按照依赖关系进行连接。

处理模块13，用于利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。

30 本实施例提供的数据分析装置，将处理模型按照依赖关系连接，上一个处理模型的

输出可以直接通过数据接口输入到下一个处理模型中，中间数据不再进行落地，节省资源，而且由于上一个处理模型生成的中间数据直接进入下一个处理模型，避免了中间数据的加载过程，提高了数据分析效率。

实施例五

5 如图 7 所示，其为本发明实施例五的数据分析装置的结构示意图。该数据分析装置除了包括上述实施例四中的构建模块 11、连接模块 12 和处理模块 13 之外，还包括：获取模块 14、解析模块 15、效果验证模块 16 和落地模块 17。

获取模块 14，用于获取所述数据分析流程的无回路有向 DAG 图；

解析模块 15，用于解析所述 DAG 图获取数据分析流程的相关信息。

10 其中，所述相关信息中包括：处理节点的逻辑功能、处理节点之间的依赖关系以及各处理模型存储地址。

本实施例中处理模块一种可选的结构方式包括：数据校验单元 131、获取单元 132、格式转换单元 133 和处理单元 134。

其中，数据校验单元 131，用于对所述原始数据进行数据校验。

15 获取单元 132，用于当原始数据通过校验时依据所述相关信息中各处理模型的存储地址获取各处理模型。

格式转换单元 133，用于将所述原始数据进行格式转换得到输入数据。

处理单元 134，用于将输入数据依次输入到连接后的处理模型中进行处理以得到所述目标数据。

20 进一步地，数据分析装置还包括：效果验证模块 16。

效果验证模块 16，用于根据预设的效果验证条件对目标数据进行效果验证。

构建模块 11，具体用于根据所述相关信息中每个处理节点的所述逻辑功能构建对应的所述处理模型。

25 连接模块 12，具体用于根据所述相关信息中处理节点之间的依赖关系将各处理模型通过数据接口连接。

进一步地，数据分析装置还包括：落地模块 17。

落地模块 17，用于根据所述相关信息中各处理模型存储地址将各处理模型进行落地存储。

30 本实施例中，通过获取数据分析流程的 DAG 图，对 DAG 图进行解析，根据解析结果构建各处理模型，以及将各处理模型按照依赖关系连接，利用连接后的处理模型对待处

理的原始数据进行处理以得到目标数据。本实施例中处理模型之间直接连接，上一个处理模型的输出可以直接输入到下一个处理模型中，中间数据不再进行落地，节省资源，并且省略中间数据的加载过程，提高了数据分析效率。

进一步地，可以将构建的这些处理模型进行落地，可以根据用户配置的存储地址进行存储，以便于这些处理模型的复用。

本领域普通技术人员可以理解：实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时，执行包括上述各方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权利要求书

1. 一种数据分析方法，其特征在于，包括：

构建数据分析流程中各处理节点对应的处理模型；

将各处理模型按照处理节点之间的依赖关系进行连接；

5 利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。

2. 根据权利要求 1 所述的方法，其特征在于，所述构建数据分析流程中各处理节点的处理模型之前，包括：

获取所述数据分析流程的无回路有向 DAG 图；

解析所述 DAG 图获取所述数据分析流程的相关信息；

10 其中，所述相关信息中包括：处理节点的逻辑功能、处理节点之间的依赖关系以及各处理模型存储地址。

3. 根据权利要求 2 所述的方法，其特征在于，所述利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据，包括：

对所述原始数据进行数据校验；

15 当原始数据通过数据校验时，依据所述相关信息中各处理模型的存储地址获取各处理模型；

将所述原始数据进行格式转换得到输入数据；

将所述输入数据输入到连接后的处理模型中进行处理以得到所述目标数据。

4. 根据权利要求 3 所述的方法，其特征在于，所述将所述输入数据输入到连接后的处理模型中得到所述目标数据之后，还包括：

20 根据预设的效果验证条件对目标数据进行效果验证。

5. 根据权利要求 2 所述的方法，其特征在于，所述构建数据分析流程中各处理节点对应的处理模型，包括：

根据所述相关信息中每个处理节点的所述逻辑功能构建对应的所述处理模型。

25 6. 根据权利要求 2 所述的方法，其特征在于，所述将各处理模型按照处理节点之间的依赖关系进行连接，包括：

根据所述相关信息中处理节点之间的依赖关系将各处理模型通过数据接口连接。

7. 根据权利要求 2 所述的方法，其特征在于，所述构建数据分析流程中各处理节点对应的处理模型之后，包括：

30 根据所述相关信息中各处理模型存储地址将各处理模型进行落地存储。

8. 一种数据分析装置，其特征在于，包括：

构建模块，用于构建数据分析流程中各处理节点对应的处理模型；

连接模块，用于将各处理模型按照依赖关系进行连接；

处理模块，用于利用连接后的处理模型对待处理的原始数据进行处理以得到目标数据。

9. 根据权利要求 8 所述的装置，其特征在于，还包括：

获取模块，用于获取所述数据分析流程的无回路有向 DAG 图；

解析模块，用于解析所述 DAG 图获取所述数据分析流程的相关信息；

其中，所述相关信息中包括：处理节点的逻辑功能、处理节点之间的依赖关系以及各处理模型存储地址。

10. 根据权利要求 9 所述的装置，其特征在于，所述处理模块包括：

数据校验单元，用于对所述原始数据进行数据校验；

获取单元，用于当原始数据通过校验时依据所述相关信息中各处理模型的存储地址获取各处理模型；

格式转换单元，用于将所述原始数据进行格式转换得到输入数据；

处理单元，用于将所述输入数据依次输入到连接后的处理模型中进行处理以得到所述目标数据。

11. 根据权利要求 10 所述的装置，其特征在于，还包括：

效果验证模块，用于根据预设的效果验证条件对目标数据进行效果验证。

12. 根据权利要求 9 所述的装置，其特征在于，所述构建模块，具体用于根据所述相关信息中每个处理节点的所述逻辑功能构建对应的所述处理模型。

13. 根据权利要求 9 所述的装置，其特征在于，所述将各处理模型按照处理模型之间的依赖关系进行连接，包括：

根据所述相关信息中处理节点之间的依赖关系将各处理模型通过数据接口连接。

14. 根据权利要求 9 所述的装置，其特征在于，还包括：

落地模块，用于根据所述相关信息中各处理模型存储地址将各处理模型进行落地存储。

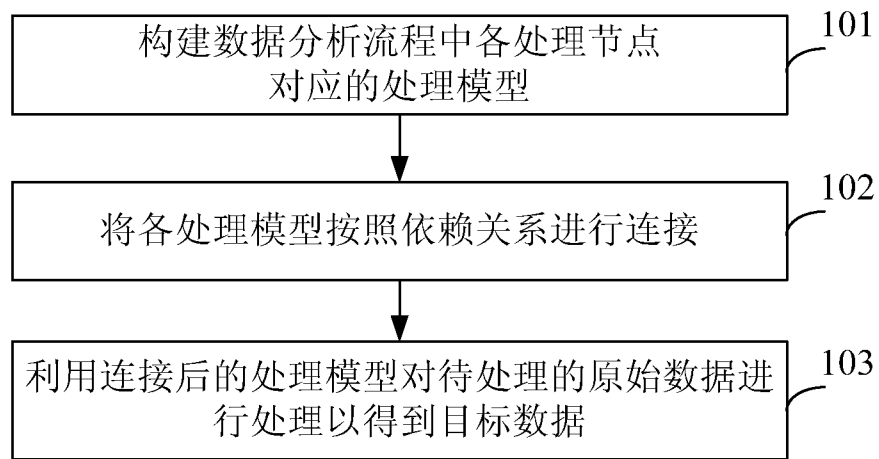


图 1

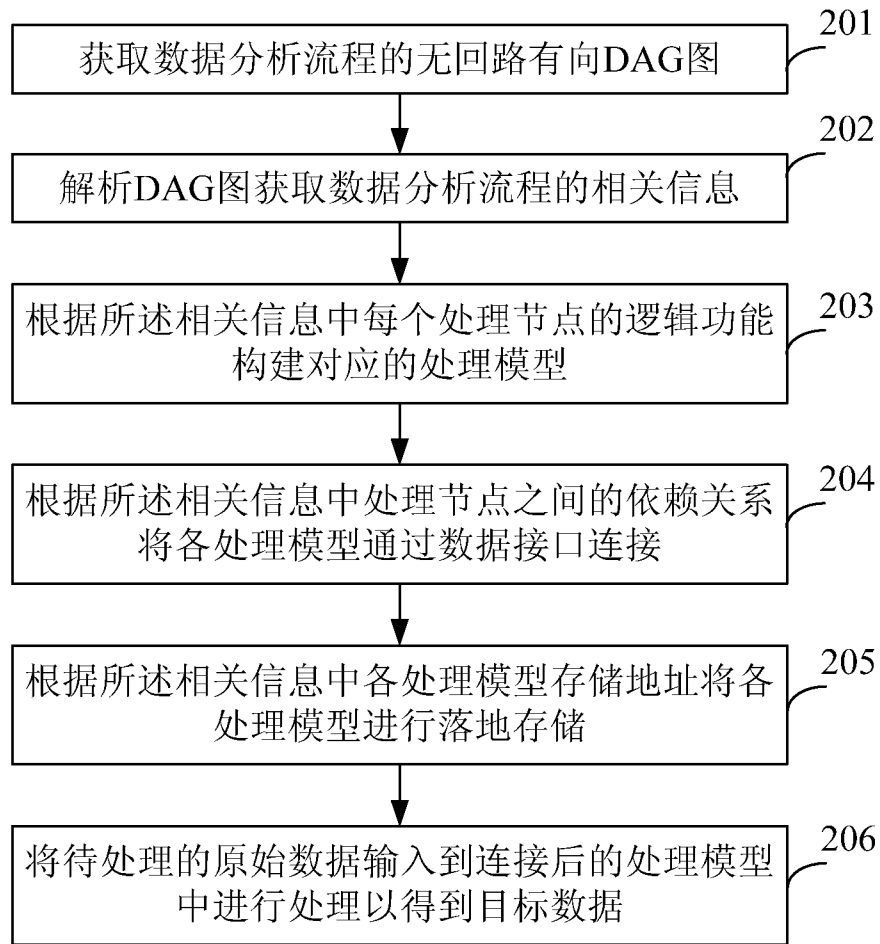


图 2

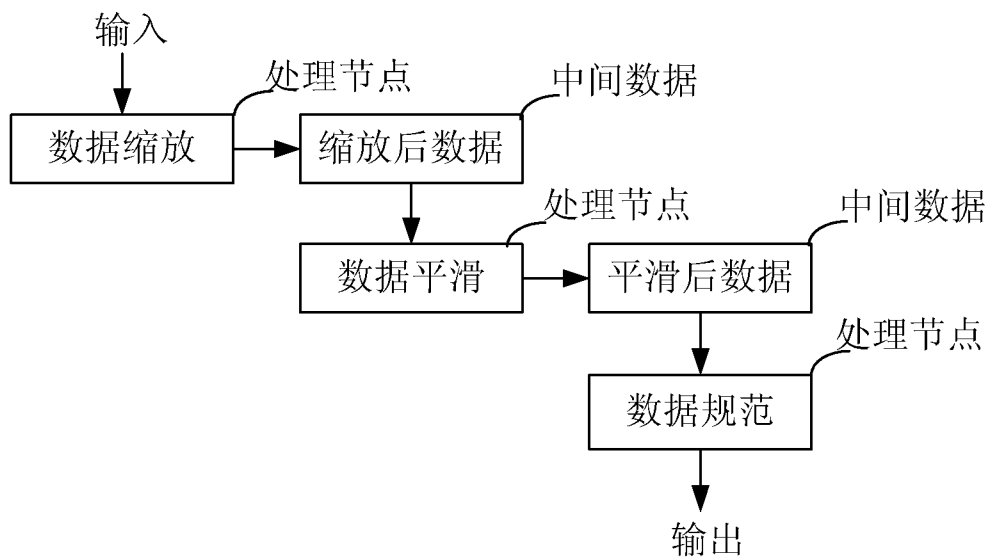


图 3

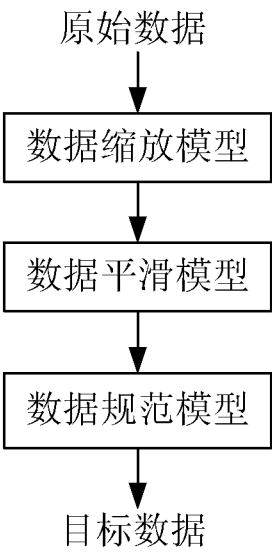


图 4

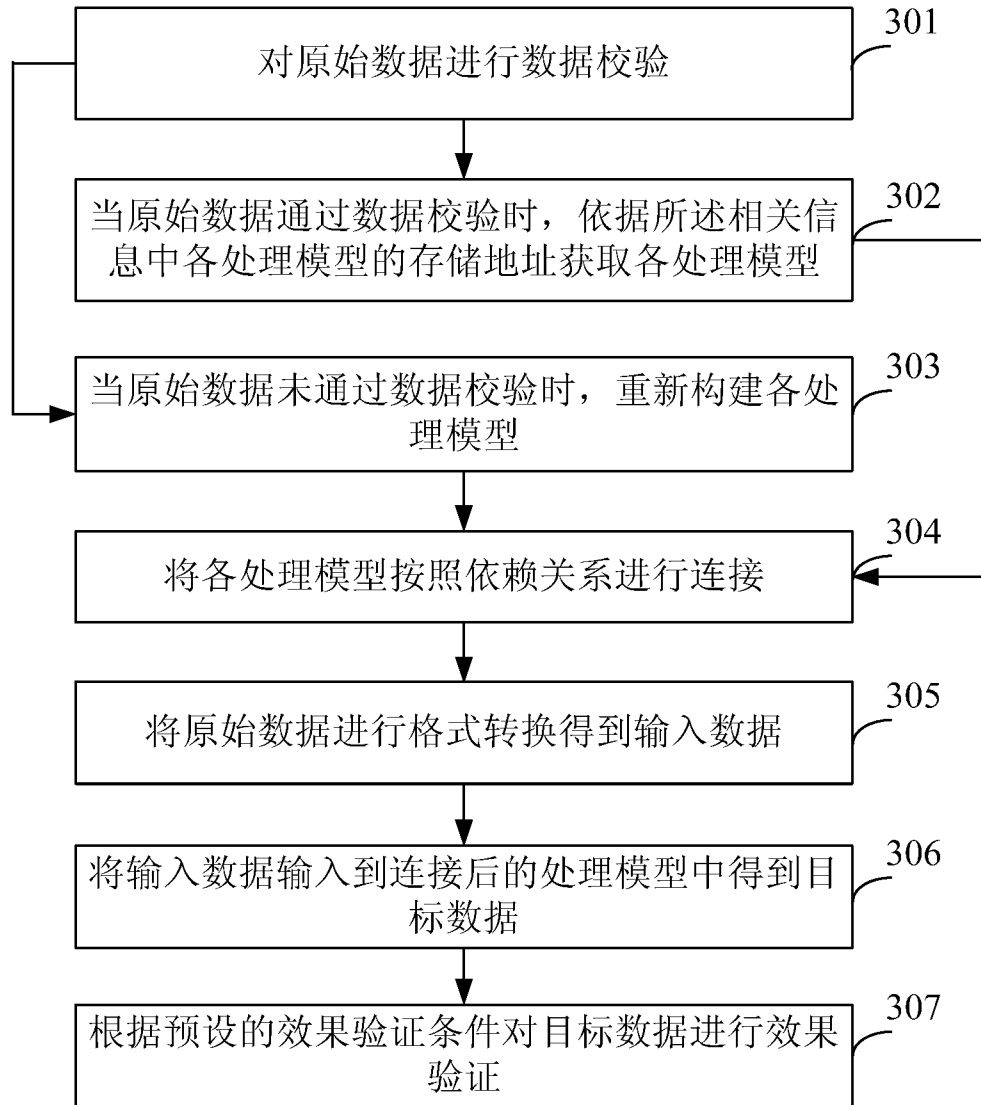


图 5

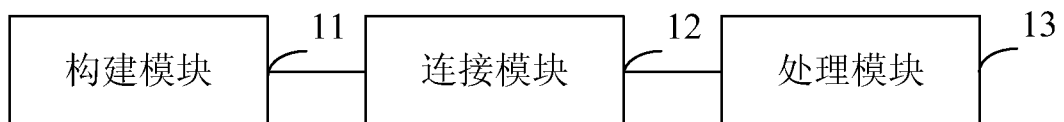


图 6

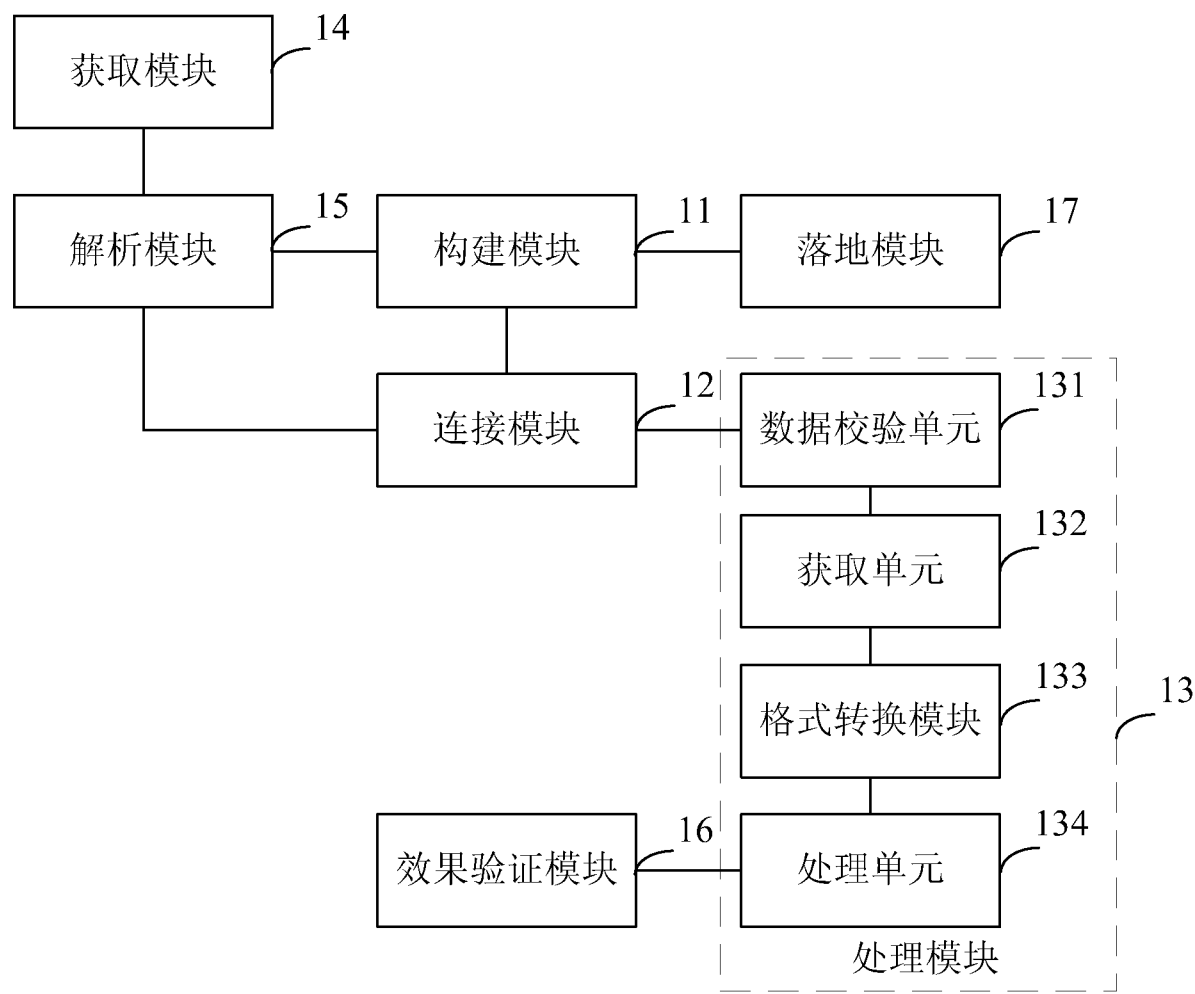


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/076876**A. CLASSIFICATION OF SUBJECT MATTER**

G06F 19/00 (2011.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, DWPI: node, site, connect, model, pipeline, graph, dag, data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102831571 A (VIVANTE TECHNOLOGY (SHANGHAI) CO., LTD.), 19 December 2012 (19.12.2012), description, paragraphs 6-30, and figure 1	1, 8
A	CN 104239137 A (NEUSOFT CORPORATION), 24 December 2014 (24.12.2014), the whole document	1-14

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 20 April 2017 (20.04.2017)	Date of mailing of the international search report 27 April 2017 (27.04.2017)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer GAO, Ke Telephone No.: (86-10) 62412067

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2017/076876

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102831571 A	19 December 2012	CN 102831571 B	22 October 2014
CN 104239137 A	24 December 2014	None	

A. 主题的分类		
G06F 19/00 (2011.01) i		
按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献 (标明分类系统和分类号)		
G06		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))		
CNABS, DWPI: 节点, 连接, 模型, 流水线, 图, 数据, site, connect, model, pipeline, graph, dag, data		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 102831571 A (图芯芯片技术上海有限公司) 2012年 12月 19日 (2012 - 12 - 19) 说明书第6-30段, 附图1	1, 8
A	CN 104239137 A (东软集团股份有限公司) 2014年 12月 24日 (2014 - 12 - 24) 全文	1-14
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2017年 4月 20日		2017年 4月 27日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088		高可
传真号 (86-10) 62019451		电话号码 (86-10) 62412067

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/076876

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	102831571	A	2012年 12月 19日	CN 102831571 B	2014年 10月 22日
CN	104239137	A	2014年 12月 24日	无	