

(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
2021 年 7 月 8 日 (08.07.2021)



(10) 国际公布号
WO 2021/135540 A1

(51) 国际专利分类号:
G06F 16/9536 (2019.01)

安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(21) 国际申请号: PCT/CN2020/122828

(22) 国际申请日: 2020 年 10 月 22 日 (22.10.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010591171.1 2020年6月24日 (24.06.2020) CN

(71) 申请人: 平安科技 (深圳) 有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福

(72) 发明人: 敖琦 (AO, Qi); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 唐炳武 (TANG, Bingwu); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市世联合知识产权代理有限公司 (SL INTELLECTUAL PROPERTY CO., LTD.); 中国广东省深圳市福田区沙头街道天安社区泰然四路 29 号天安创新科技广场一期 A 座 805E, Guangdong 518000 (CN)。

(54) Title: NEO4J-BASED ANOMALOUS USER PROCESSING METHOD AND APPARATUS, COMPUTER DEVICE, AND MEDIUM

(54) 发明名称: 基于Neo4j的异常用户处理方法、装置、计算机设备和介质

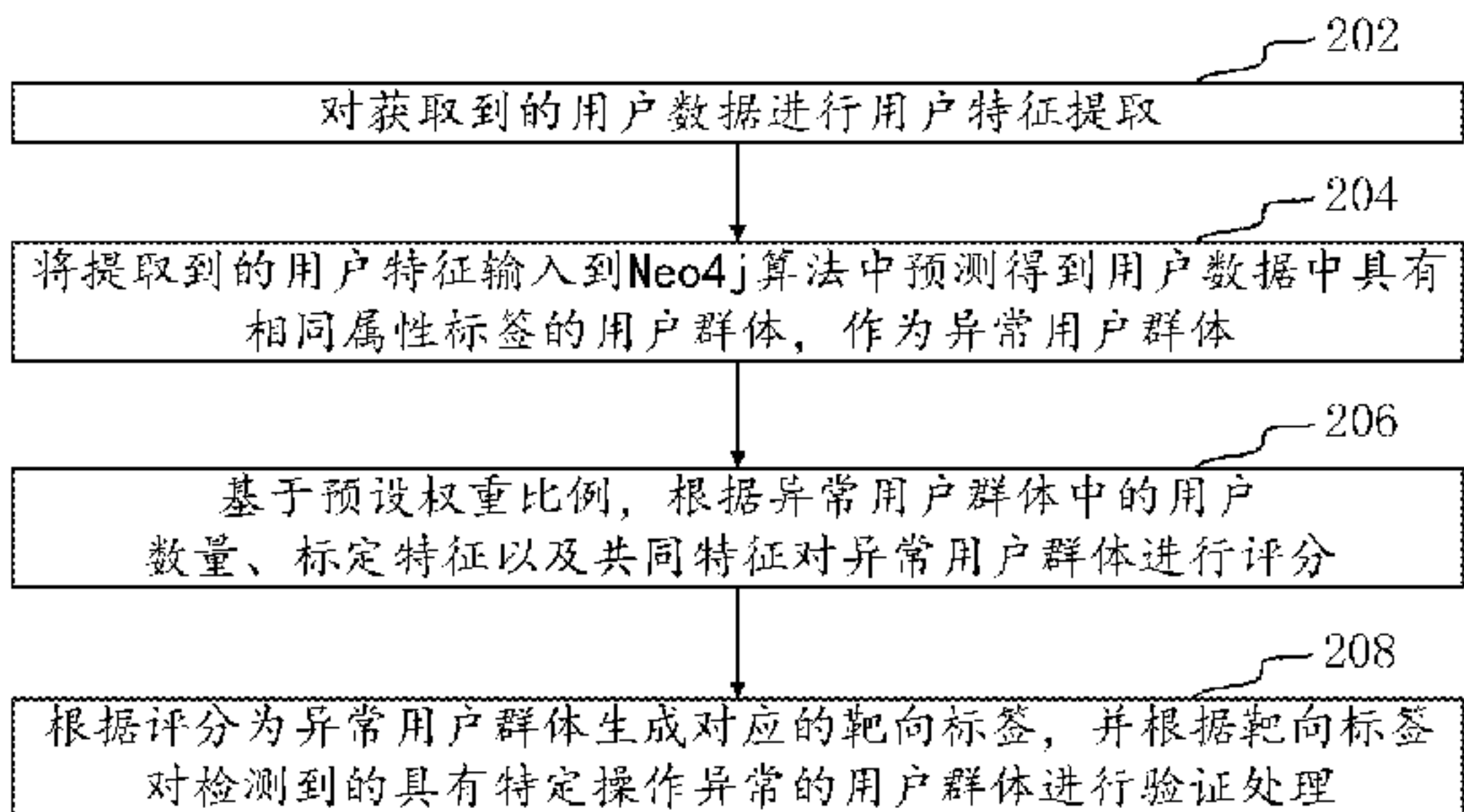


图 2

- 202 Perform user feature extraction on acquired user data
204 Input the extracted user features into a Neo4j algorithm to predict and obtain from the user data a user group having a same attribute label, and determine the user group is an anomalous user group
206 On the basis of a preset weight ratio, score the abnormal user group on the basis of the number of users in the anomalous user group, delineating features, and common features
208 Generate a corresponding targeting label for the anomalous user group on the basis of the score and, on the basis of the targeting label, perform authentication processing on a detected anomalous user group having a specific operation

(57) Abstract: A Neo4j-based anomalous user processing method and apparatus, a computer device, and a readable storage medium, relating to the fields of big data and data mining. The method comprises: performing user feature extraction on acquired user data (202); inputting the extracted user features into a Neo4j algorithm to predict and obtain from the user data a user group having a same attribute label, and determining the user group is an anomalous user group (204); on the basis of a preset weight ratio, scoring the abnormal user group on the basis of the number of users in the anomalous user group, delineating features, and common features (206); and generating a corresponding targeting label for the anomalous user group on the basis of the score and, on the basis of the targeting label, processing a detected anomalous user group having a specific operation (208). The invention solves the technical problem of inaccurate processing of anomalous users.

(57) 摘要: 一种基于Neo4j的异常用户处理方法、装置、计算机设备及可读存储介质, 涉及了大数据和数据挖掘领域。所述方法包括对获取到的用户数据进行用户特征提取 (202); 将提取到的用户特征输入到Neo4j算法中预测得到用户数据中具有相同属性标签的用户群体, 作为异常用户群体 (204); 基于预设权重比例, 根据异常用户群体中的用户数量、标定特征以及共同特征对异常用户群体进行评分 (206); 根据评分为异常用户群体生成对应的靶向标签, 根据靶向标签对检测到的具有特定操作的异常用户群体进行处理 (208)。解决了对异常用户的处理不准确的技术问题。

WO 2021/135540 A1

(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, IT, JO, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。
- 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布(细则48.2(h))。
- 根据申请人的请求, 在条约第21条(2)(a)所规定的期限届满之前进行。

基于 Neo4j 的异常用户处理方法、装置、计算机设备和介质

本申请以 2020 年 6 月 24 日提交的申请号为 202010591171.1，名称为“基于 Neo4j 的异常用户处理方法、装置、计算机设备和介质”的中国发明专利申请为基础，并要求其优先权。

5

技术领域

本申请涉及大数据领域，特别是涉及一种基于 Neo4j 的异常用户处理方法、装置、计算机设备和存储介质。

10 背景技术

目前，各大互联网平台为了争夺用户，往往会采用让利补贴的方式提高活跃度，但同时也滋养了一群特殊的用户——“羊毛党”。“羊毛党”逐渐从个体玩家发展为有组织、有规模的职业集体，这些异常用户不仅为企业带来了负担，还为后台服务器的数据处理带来了困难。现有技术中通常会将检测到的异常用户加入黑名单，但发明人意识到这种对异常用户验证处理的方式过于简单粗暴，这就会导致一类倾向于从异常用户转为正常用户的用户也无法享受正常的权限，给异常用户的验证处理造成极大的不准确性。

发明内容

基于此，有必要针对上述技术问题，本申请提供一种基于 Neo4j 的异常用户处理方法、装置、计算机设备及存储介质，以解决现有技术中对异常用户的处理不准确的技术问题。

一种基于 Neo4j 的异常用户处理方法，所述方法包括：

对获取到的用户数据进行用户特征提取；

将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

25 基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；并

根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

一种基于 Neo4j 的异常用户处理装置，所述装置包括：

30 提取模块，用于对获取到的用户数据进行用户特征提取；

检测模块，用于将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

评分模块，用于基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；以及

35 处理模块，用于根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

一种计算机设备，包括存储器和处理器，以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现下述基于 Neo4j 的异常用户处理方法的步骤：

40 对获取到的用户数据进行用户特征提取；

将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；以及

45 根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

一种计算机可读存储介质，所述计算机可读存储介质存储有计算机可读指令，所述计算机可读指令被处理器执行时实现下述基于 Neo4j 的异常用户处理方法的步骤：对获取到的用户数据进行用户特征提取；

5 将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；以及

根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

10 上述基于 Neo4j 的异常用户处理方法、装置、计算机设备和存储介质，通过根据异常用户群体中的用户数量、特征对异常用户进行评级分级，然后根据评级分级生成靶向标签，当异常用户触发特定操作后根据靶向标签触发不同的处理机制，对异常用户进行针对性的验证处理，解决了现有技术中对异常用户的验证处理不准确的技术问题。

15 附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的描述中所需要使用到的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

20 图 1 为基于 Neo4j 的异常用户处理方法的应用环境示意图；

图 2 为基于 Neo4j 的异常用户处理方法的流程示意图；

图 3 为图 2 中步骤 204 的流程示意图；

图 4a 为步骤 302 中一个关系图谱示意图；

图 4b 为步骤 302 中另一个关系图谱示意图；

25 图 4c 为步骤 302 中的另一关系图谱示意图；

图 5 为图 2 中步骤 206 的流程示意图；

图 6 为基于 Neo4j 的异常用户处理装置的示意图；

图 7 为一个实施例中计算机设备的示意图。

30 具体实施方式

除非另有定义，本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同；本文中在申请的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本申请；本申请的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形，意图在于覆盖不排他的包含。本申请的说明书和权利要求书或上述附图中的术语“第一”、“第二”等是用于区别不同对象，而不是用于描述特定顺序。

40 在本文中提及“实施例”意味着，结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例，也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是，本文所描述的实施例可以与其它实施例相结合。

为了使本申请的目的、技术方案及优点更加清楚明白，下面结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

45 本申请实施例提供的基于 Neo4j 的异常用户处理方法，可以应用于如图 1 所示的应用环境中。其中，该应用环境可以包括终端 102、网络以及服务端 104，网络用于在终端 102 和服

务端 104 之间提供通信链路介质，网络可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

5 用户可以使用终端 102 通过网络与服务端 104 交互，以接收或发送消息等。终端 102 上可以安装有各种通讯客户端应用，例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

终端 102 可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、电子书阅读器、MP3 播放器 (Moving Picture Experts Group Audio Layer III, 动态影像专家压缩标准音频层面 3)、MP4 (Moving Picture Experts Group Audio Layer IV, 动态影像专家压缩标准音频层面 4) 播放器、膝上型便携计算机和台式计算机等等。

10 服务端 104 可以是提供各种服务的服务器，例如对终端 102 上显示的页面提供支持的后台服务器。

需要说明的是，本申请实施例所提供的基于 Neo4j 的异常用户处理方法一般由服务端/终端执行，相应地，基于 Neo4j 的异常用户处理装置一般设置于服务端/终端设备中。

15 应该理解，图 1 中的终端、网络和服务端的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务端。

其中，终端 102 通过网络与服务端 104 进行通信。服务端 104 通过网络从终端 102 获取用户数据，并对用户数据进行特征提取，将用户特征输入 Neo4j 算法预测出异常用户群体，然后根据异常用户群体中的用户数量、标定特征以及共同特征对异常用户群体中进行分评分，并根据评分生成对应的靶向标签，最后根据靶向标签对异常用户群体中的异常用户进行对应的验证处理。其中，终端 102 和服务端 104 之间通过网络进行连接，该网络可以是有线网络或者无线网络，终端 102 可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备，服务端 104 可以用独立的服务器或者是多个组成的服务器集群来实现。

25 在一个实施例中，如图 2 所示，提供了一种基于 Neo4j 的异常用户处理方法，以该方法应用于图 1 中的服务端为例进行说明，包括以下步骤：

步骤 202，对获取到的用户数据进行用户特征提取。

30 用户数据包括用户的手机号码、IP 地址、设备号以及手机号码、IP 地址、设备号之间的联系、不同用户的用户特征之间的关系等等，另外，还包括不同的用户所参与的活动类型，比如每日签到、新人注册、邀新注册等等。其中，一个用户可以使用一个手机号、多个 IP 地址和设备号。

用户特征提取是为了提取出以用户的手机号码、IP 地址、设备号为维度的特征值，以及手机号码、IP 地址、设备号三种特征之间的关联关系的过程。

需要强调的是，为进一步保证上述用户数据的私密和安全性，上述用户数据还可以分布式存储于区块链的节点中。

35 步骤 204，将提取到的用户特征输入到 Neo4j 算法中进行预测，得到用户数据中具有相同属性标签的用户群体，作为异常用户群体。

Neo4j 是一个高性能的，NOSQL 图形数据库，它将结构化数据存储在网络上而不是表中。标签传播算法 (Label Propagation) 是 Neo4j 中通过节点之间的边传播标签的算法。

40 用户的手机号码、IP 地址以及设备号都有一个标注特征，标注该手机号码、IP 地址或者设备号是否在服务端的黑名单、白名单或者未确认名单中，其中，一个用户的手机号码、与该手机号码对应的设备号可以在白名单、但是其所对应的 IP 地址可以在黑名单中；也有可能是该用户的手机号码、IP 地址以及设备号都在黑名单中。

黑名单和白名单中的数据是根据历史经验获取的，黑名单数据有些是用户举报、拉黑得到的，白名单是根据用户在某 APP 中的活动参与量等特征得到的。

45 将提取得到的用户特征输入到标签传播算法中，标签传播算法会以手机号码、IP 以及设备号这三类作为关系节点，基于三类关系节点之间的联系构建关系网络。

其中，两关系节点之间的关系，在关系网络中表现为连接两关系节点的边。用户数据中包括实时用户数据，可以将实时用户数据直接应用实时采集到的用户数据，保证异常用户预测的及时性。然后根据构建的关系网络预测得到异常用户群体。

5 本实施例通过 Neo4j 算法进行异常用户群体的预测，无需根据既定的节点模板得到风险图，使得预测过程的灵活性更强，普适性更强。

步骤 206，基于预设权重比例，根据异常用户群体中的用户数量、标定特征以及共同特征对异常用户群体进行评分。

10 预设权重比例是根据历史经验得到的用于整合异常用户群体各特征值的比例，可以根据具体应用场景设定；标定特征包括但不限于用户的每日签到、新人注册、邀新注册、特定活动等维度的特征；特定活动包括但不限于双十一活动、618 活动等对用户来说具有获得高回报卡券、优惠券等虚拟对象的活动；共同特征指的是异常用户群体中的异常用户可能具有的相同特征维度的特征，比如一个异常用户群体中超过 50% 的异常用户具有相同的设备号、手机归属地、注册时间或相同的 IP。

15 步骤 208，根据评分为异常用户群体生成对应的靶向标签，并根据靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

特定操作一般包括：异常用户群体中异常用户的请求登录某 app 账号的操作、请求获取某 app 中某功能模块的操作等。不同范围的评分对应不同的靶向标签，不同的靶向标签会触发不同的验证处理机制。比如，当得到某一异常用户群体的评分为 91 时，可以为该异常用户群体生成一个靶向标签为拒绝登录和/或拒绝参与活动的标签，限制该异常用户群体的活动，
20 比如拒绝登录的验证处理，以实现对不同异常用户群体进行不同验证处理的目的，还可以实现降低异常用户群体产生的冗余数据，从而提高服务端数据处理效率的目的。

上述基于 Neo4j 的异常用户处理方法中，通过在服务端通过 Neo4j 算法预测出异常用户群体，然后根据异常用户群体的三类指定的特征对其进行评分，根据评分结果生成对应的靶向标签，根据不同的靶向标签对异常用户群体中异常用户的登录、活动等行为进行不同程度的处理，实现
25 对不同异常用户群体进行不同验证处理目的，还可以实现降低异常用户群体产生的冗余数据，从而提高服务端数据处理效率的目的。

在一个实施例中，如图 3 所示，步骤 204，包括：

步骤 302，通过 Neo4j 算法构建基于各用户特征之间关联关系的关系网络，其中，关系网络主要由关系节点为手机号码、IP 地址以及设备号组成。

30 如图 4a、4b 关系图谱示意图所示，三角形代表 IP 地址、圆形代表手机号码、五边形代表设备号，一个 IP 地址可以对应多个手机号、一个手机号码也可以对应多个设备号、一个设备号也可以对应多个 IP 地址，通常，设备号与 IP 地址通过手机号码进行关联，形成二度关系，而手机号码分别与设备号、IP 地址进行直接关联，形成一度关系。

其中，图 4a 是根据正常用户数据（白名单用户、非白非黑未确认节点的用户数据）构建出的关系网络，可以看出在单点特征上正常用户的关系网络较为分散；图 4b 是根据黑名单用户数据构建的关系网络，在一些单点特征上，同正常用户一样呈现出分散的特点，使得单点特征防御难以奏效，但是如果将用户特征用网络的形式建模展示，会发现在一些特殊的图形特征上，异常行为明显异于正常行为。其中，单点特征指的是单个特征。比如，用户使用的设备号数量、IP 数量这是两个特征，单单从一个特征上看，比如设备号数量这一个单点特征
40 看，无法区分异常用户和正常用户。

用户特征是指用户在某个时间（段）做了某一件事情，例如用户在凌晨 1 点到 3 点登陆了某 APP 10 次；用户 30 天内在某 APP 中签到了 20 次，这些都属于用户特征。

进一步地，通过 Neo4j 构建基于用户数据的关系网络如图 4c 中关系图谱示意图所示，其中，图 4c 只是示例性的。

45 步骤 304，获取关系节点的标注特征，并根据标注特征对关系节点进行分类，得到黑名单节点、白名单节点以及未确认节点，其中，标注特征为关系节点的初始类别。

标注特征表示用户的手机号码、IP 地址或者设备号是属于黑名单中的黑名单节点、白名单中的白名单节点或者未确认名单中的未确认节点中的哪一种。

刚获取到的关系网络都会有一个初始类别，初始类别包括该 IP 地址、手机号码或设备号属于黑名单、白名单还是未确认名单，其中，可为黑名单的关系节点生成的标签为 1，为白名单以及未确认名单的关系节点生成的标签为 0。

步骤 306，为关系节点生成属性标签，其中，属性标签包括异常、正常。

特别的，黑名单中的手机号码、设备号以及 IP 地址的标注特征，用于标注各关系节点的初始的属性标签属于异常，还有一个属性标签：black 标签；而白名单中的关系节点的标注特征，用于标注各关系节点的初始的属性标签属于正常，且还具有一个属性标签：white 标签，black 标签和 white 标签是为了防止黑、白名单中关系节点的属性标签在后续节点标签更新时被刷新；而未确认名单中的关系节点（手机号码、IP 地址以及设备号）的属性标签为：unidentified（意为未确认标签），在后续的标签预测中可以被刷新，其初始的标注特征为正常。

即，位于黑名单或者白名单中的 IP 地址、设备号或手机号码的属性标签最终不会被更新或改变。

步骤 308，分别以属性标签为异常的手机号码、IP 地址的关系节点为种子节点，从关系网络中获取路径长度不大于第一预设值的节点路径，得到节点关系图谱，其中，关系图谱包括基于异常手机号码节点的异常号码关系图谱和基于异常 IP 节点的异常 IP 关系图谱。

以属性标签为异常的 IP 地址为种子节点，获取路径长度不大于第一预设值的节点路径组成的基于异常 IP 节点的异常 IP 关系图谱，其中，第一预设值根据经验可以设置为 10。还需要以属性标签为异常的手机号码为种子节点，获取路径长度不大于 10 的节点路径组成的基于异常手机号码节点的异常号码关系图谱。

具体地，可以通过 Neo4j 的 cypher 语言查找路径长度在 1 到 10 之间的关系节点组成的节点路径。表达方式可以为 (a) -[*1..10]-> (b)，cypher 语言类似 SQL 查询语句。其中，a、b 表示的是关系节点，[*1..10] 表示的是 1 到 10 度的节点路径。

本实施例以两个不同类型的关系节点作为种子节点进行关系图谱的构建是为了使得预测的图谱更加全面、预测的过程更加高效；若以一个关系节点作为种子节点构建一个大的关系图谱，则需要获取不大于 N 的节点路径，N 为整数；而 N 需要远远大于 10 才能够保证大部分的节点连接情况能够被包含在关系图谱中，而这样就会导致在关系图谱构建过程中路径切分耗时过长，影响预测效率。所以使用两个节点构建两个关系图谱，即能够保证关系图谱的完整性，也能够保证切分的效率。

步骤 310，分别对异常号码关系图谱与异常 IP 关系图谱中的未确认节点进行标签预测，得到异常用户群体。

这一步也是在标签传播算法里实现的，在首轮标签更新前，对于未确认名单中的关系节点的标签默认为正常，而其他的关系节点属性标签要么确定是异常，要么确定是正常。

例如，关系节点 A 的属性标签已更新完，从正常变成了异常，而关系节点 B 与关系节点 A 互为邻居节点，关系节点 B 的标签更新结果就会受到关系节点 A 的标签影响。

具体地，对比未确认节点属性标签为异常、正常的邻居节点的数量，得到标签对比结果。

标签刷新时，对于未确认节点（unidentified 节点），根据其邻居节点的标签，对其标签进行更新。

根据标签对比结果更新未确认节点属性标签。

对于未确认名单中的某个未确认节点，结合该关系节点所有的邻居节点属性标签，若所有邻居节点属性标签为正常的大于属性标签为异常的，则为该关系节点生成正常的属性标签，反之，若则生成异常属性标签。如果，属性标签为正常的数量和为异常的数量邻居节点一样多，则给该关系节点随机生成标签异常或者正常，这时生成异常或正常的标签对

异常用户群体的预测并无不利影响。

按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正，并重复属性标签对比、更新的操作，直到每个未确认节点的属性标签不再改变或者更新次数达到阈值，将最后一次属性标签更新后关系图谱，作为待预测图谱。

5 其中，预设修正条件是在标签更新过程中对关系节点的标签进行调整的条件。具体地，在每一轮标签变更后，需要对一部分关系节点的标签结果进行修正：即与白名单中 IP 地址具有一度关系的未确认节点，其属性标签都置为正常。对两个关系图谱中的每个未确认节点，依次按这样的方式，进行标签刷新，直到每个关系节点的标签结果不变或刷新的轮数达到阈值 10000。

10 标签迭代过程中，仅仅针对未确认名单中的关系节点进行迭代，遍历时间短。在对异常号码关系图谱中的属性标签进行预测时，在每次迭代后，对与白名单中 IP 地址关联的一度节点（即手机号码节点），重新置为正常，在对异常 IP 关系图谱中的属性标签进行预测时，在每次迭代后，将与白名单手机号关联的一度节点（IP）的标签重置为 0，防止使用企业 wifi 的用户被置为黑名单。提高异常用户群体预测的准确率。

15 断开待预测图谱中的异常边，得到多个关系子图，其中，异常边为任意两关系节点直接相连组成的边中至少有一个关系节点的属性标签为正常的边。

具体地，将与白名单节点中的 IP 地址具有一度关系的手机号码的属性标签设为正常，其中，一度关系是指两个直接相连的关系节点的关联关系。

20 通过该方式生成的关系子图中，所有的关系节点的标签都为正常，以此保证每一个手机号节点的邻居节点的属性标签也都为正常。本实施例依照只有当同一用户的手机号节点、IP 节点以及设备号节点的标签全部都是异常的时候，才会将该手机号节点所对应的用户标记为异常用户。

统计关系子图中关系节点的属性标签，若属性标签为异常的手机号码或 IP 地址的数量大于第二预设值，则将关系子图中的关系节点对应的用户群体作为异常用户群体。

25 用户是根据手机号码/IP 地址来确认的。一般羊毛党的数量都是聚众出现，使用一个手机号码在不同 IP 地址下、不同设备号上刷数据，还有在同一 IP 地址下使用不同手机号、不同设备号；其用户欺诈类型又可以分为养号欺诈团伙、活动欺诈团伙，比如某个欺诈团伙，所有的手机号码的在某 app 上的签到次数为 0，参加活动次数为 0，则可以将该欺诈团伙定义为养号欺诈团伙；若某个欺诈团伙，所有的手机号参加活动的次数不为 0，则该欺诈团伙定义为活动欺诈团伙。

30 进一步地，还可以比较得到的异常用户群体对异常用户群体进行去重处理，具体地，若基于异常手机号码节点的异常用户群体与基于异常 IP 节点的异常用户群体相同，则合并异常用户群体，得到合并后的异常用户群体。因为基于不同节点为中心进行群体预测，可能会得到同一用户，即在异常手机号码为中心节点进行预测得到的 A 群体中，也在异常 IP 节点为中心得到的 A 群体中，所以需要查找得到的异常用户群体是否有相同标签，将相同标签的异常用户群体的同一用户，进行合并，以节省存储空间，并给合并后的异常用户群体生成相同的靶向标签，后续当检测到该异常用户群体中的异常用户的特定操作，比如登录操作后，会根据该异常用户的靶向标签，对其进行对应的验证处理。如果没有检测到，则说明预测得到的是不同的异常用户群体，则为不同的异常用户群体生成不同的靶向标签，并当检测到这些异常用户群体中的异常用户的特定操作后，比如登录操作后，根据其对应的靶向标签，调用对应的验证接口对该异常用户进行验证处理。通过本实施例提高了数据存储的效率。

40 本实施例直接通过关系节点之间互为邻居标签，对需要确定标签的关系节点进行标签更新，不局限于为其中某一指定的关系节点更新标签，除特殊标签（黑名单、白名单）外，所有节点的标签在预测中都在不断被刷新，保证在标签预测时能够充分利用邻居标签的影响力，最后断开存在标签为正常的节点所在的关系边生成子图，使得预测过程的灵活性更强，普适性更强，而且在标签迭代过程中，仅仅针对未确认节点进行迭代，遍历时间短。且迭代一次

45

后，将与白名单的 IP 地址关联的一度节点（即手机号码）的标签重新置为正常，防止使用企业 wifi 的用户被置为黑名单这类情况的错误预测。

在一个实施例中，如图 5 所示，步骤 206，包括：

步骤 502，获取异常用户群体的用户数量，并根据第一预设评级分数和用户数量设定异常用户群体的第一评级分数。

用户数量指的是一个异常用户群体中异常用户的个数，可以根据该用户数量确定该异常用户群体的第一评级分数；具体地：

异常用户群体的用户数量与第一预设评级分数如表 1：

用户数量/个	第一预设评级分数/分
<=10	10
(10, 30]	20
(30, 50]	40
(50, 70]	60
(70, 90]	80
>90	100

表 1

例如，当用户数量低于 10 时，对应的第一预设评级分数为 10 分，则该异常用户群体的第一评级分数为 10。

步骤 504，获取异常用户群体中具有标定特征的异常用户的标定数量，根据标定特征的运算优先级确定异常用户群体的第二评级分数。

标定特征的运算优先级是指示不同标定特征对第二评级分数影响的优先程度。比如每日签到 40 分，新人注册 60 分，邀新注册 80 分，特定活动 100 分，其中，特定活动的运算优先级大于日签到，新人注册以及邀新注册，邀新注册的运算优先级又大于每日签到和新人注册，而新人注册的又大于每日签到。当一个异常用户群体中超过 50%的异常用户参与了每日签到、邀新注册以及特定活动，则以特定活动对应的分数作为该异常用户群体的第二评级分数。其中超过 50%即是标定数量。比如，若标定数量低于 50%，则可以不为该异常用户群体进行评分，也就是第二评级分数为 0。

步骤 506，统计异常用户群体中具有共同特征的异常用户的共有数量，并根据共有数量以及第二预设评级分数确定异常用户群体的第三评级分数。

共有数量是异常用户群体中具有某共同特征的异常用户的人数，可以用于评断是否根据共有特征对应的分数确定异常用户群体的第三评级分数。

共同特征是指异常用户具有相同 IP、手机归属地、相同注册时间以及相同设备号。

相同 IP-40、同手机归属地-60、同注册时间-80、同设备号-100；

若一个异常用户群体有 50%异常用户使用相同 IP、同注册时间、同设备号，则该异常用户群体的得分依据分数最高的共有特征对应的分数，即同设备号 100 分。

步骤 508，按照预设权重比例联合第一评级分数、第二评级分数以及第三评级分数，得到异常用户群体的评分。

依据以上 3 种分级分数条件，并按照预设权重比例 1:2:2，给异常用户群体打分，其中，预设权重比例是根据经验得到的，以这种比例计算评分的准确率比较高：

比如，异常用户群体 A：

第一分级分数：20 分*0.2

第二分级分数：80 分*0.4

第三分级分数：100 分*0.4

则，分数为：

4+32+40=76 分

最后根据最终分数为该异常用户群体生成不同的靶向标签；

具体地：

<60 登录时短信验证

[60, 80] 语音验证

[81, 90] 人脸验证

5 >91 拒绝登录

不同的靶向标签的异常用户触发服务端不同的验证机制，登录时短信验证的靶向标签会在用户登录时触发短信验证机制，对用户进行验证等等，从而当异常用户登录 APP/网页时，采取不同的验证方式，增加异常用户的验证门槛，防止机器批量操作，降低服务端垃圾数据的处理量，提高数据处理效率。

10 进一步地，若检测到异常用户的特定操作，则获取异常用户所在异常用户群体的靶向标签；该异常用户的靶向标签是在检测到还用户的特定操作之前预先生成的，当检测到该异常用户群体中的某个或某些异常用户的特定操作，比如登录操作后，则会获取该异常用户的靶向标签，并根据靶向标签调用对应的验证处理接口，对该异常用户的活动权限进行验证处理；判断对异常用户的活动权限的验证处理次数是否超过权限值，并将超过权限值的异常用户设
15 为锁定用户。

具体地，针对异常用户，根据他们的得分可以调用短信接口，发送短信验证码。如果在有效时间内，用户使用正确短信登录则登录成功，那么在有效时间内用户下一次登录时，则不会再次调用该短信接口。如果登录失败，则会再一次调用短信接口，直到调用短信接口的次数，即权限值连续 3 次，则意味着在有效时间内用户都没有输入正确的验证码，则登录失
20 败，锁定账号，并将该异常用户设为锁定用户，并限制其其他操作。

通过本实施例采取不同的验证方式，增加异常用户的验证门槛，防止机器批量操作，降低服务端垃圾数据的处理量，提高了数据处理效率。

应该理解的是，虽然图 2-图 3、图 5 的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步
25 骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，图 2-图 3、图 5 中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些子步骤或者阶段的执行顺序也不必是依次进行，而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

30 在一个实施例中，如图 6 所示，提供了一种基于 Neo4j 的异常用户处理装置，该基于 Neo4j 的异常用户处理装置与上述实施例中基于 Neo4j 的异常用户处理方法一一对应。该基于 Neo4j 的异常用户处理装置包括：

提取模块 602，用于对获取到的用户数据进行用户特征提取。

检测模块 604，用于将提取到的用户特征输入到 Neo4j 算法中进行预测，得到用户数据
35 中具有相同属性标签的用户群体，作为异常用户群体。

评分模块 606，用于基于预设权重比例，根据异常用户群体中的用户数量、标定特征以及共同特征对异常用户群体进行评分。以及

处理模块 608，用于根据评分为异常用户群体生成对应的靶向标签，并根据靶向标签对检测到的具有特定操作异常的用户群体进行验证处理。

40 需要强调的是，为进一步保证上述用户数据的私密和安全性，上述用户数据还可以分布式存储于区块链的节点中。

进一步地，检测模块 604，包括：

构建子模块，用于通过 Neo4j 算法构建基于各用户特征之间关联关系的关系网络，其中，关系网络主要由关系节点为手机号码、IP 地址以及设备号组成。

45 分类子模块，用于获取关系节点的标注特征，并根据标注特征对关系节点进行分类，得到黑名单节点、白名单节点以及未确认节点，其中，标注特征为关系节点的初始类别。

标签子模块, 用于为关系节点生成属性标签, 其中, 属性标签包括异常、正常。

筛选子模块, 用于分别以属性标签为异常的手机号码、IP 地址的关系节点为种子节点, 从关系网络中获取路径长度不大于第一预设值的节点路径, 得到节点关系图谱, 其中, 关系图谱包括基于异常手机号码节点的异常号码关系图谱和基于异常 IP 节点的异常 IP 关系图谱。

5 预测子模块, 用于分别对异常号码关系图谱与异常 IP 关系图谱中的未确认节点进行标签预测, 得到异常用户群体。

进一步地, 预测子模块, 包括:

对比单元, 用于对比未确认节点的属性标签为异常、正常的邻居节点的数量, 得到标签对比结果。

10 更新单元, 用于根据标签对比结果更新未确认节点的属性标签。

修正单元, 用于按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正, 并重复属性标签对比、更新的操作, 直到每个未确认节点的属性标签不再改变或者更新次数达到阈值, 将最后一次属性标签更新后关系图谱, 作为待预测图谱。

15 断边单元, 用于断开待预测图谱中的异常边, 得到多个关系子图, 其中, 异常边为任意两关系节点直接相连组成的边中至少有一个关系节点的属性标签为正常的边。

统计单元, 用于统计关系子图中关系节点的属性标签, 若属性标签为异常的手机号码或 IP 地址的数量大于第二预设值, 则将关系子图中的关系节点对应的用户群体作为异常用户群体。

进一步地, 预测子模块, 还包括:

20 比较单元, 用于比较得到的异常用户群体。

合并单元, 用于若基于异常手机号码节点的异常用户群体与基于异常 IP 节点的异常用户群体相同, 则合并异常用户群体, 得到合并后的异常用户群体。

进一步地, 评分模块 606, 包括:

25 第一分数子模块, 用于获取异常用户群体的用户数量, 并根据第一预设评级分数和用户数量设定异常用户群体的第一评级分数;

第二分数子模块, 用于获取异常用户群体中具有标定特征的异常用户的标定数量, 根据标定特征的运算优先级确定异常用户群体的第二评级分数;

第三分数子模块, 用于统计异常用户群体中具有共同特征的异常用户的共有数量, 并根据共有数量以及第二预设评级分数确定异常用户群体的第三评级分数;

30 综合分数子模块, 用于按照预设权重比例联合第一评级分数、第二评级分数以及第三评级分数, 得到异常用户群体的评分。

进一步地, 处理模块 608, 包括:

获取子模块, 用于若检测到异常用户的特定操作, 则获取异常用户所在异常用户群体的靶向标签; 以及

35 调用子模块, 用于根据接口调用类型调用对应的接口对异常用户的活动权限进行处理; 以及

锁定子模块, 用于判断对异常用户的活动权限的处理次数是否超过权限值, 并将超过权限值的异常用户设为锁定用户。

40 上述基于 Neo4j 的异常用户处理装置, 通过根据异常用户群体中的用户数量、特征对异常用户进行评级分级, 然后根据评级分级生成靶向标签, 当异常用户触发特定操作后根据靶向标签触发不同的处理机制, 对异常用户进行针对性的验证处理, 解决了现有技术中对异常用户的处理不准确的技术问题。

45 在一个实施例中, 提供了一种计算机设备, 该计算机设备可以是服务器, 其内部结构图可以如图 7 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中, 该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据

库。该内存存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储用户数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种基于 Neo4j 的异常用户处理方法。

其中，本技术领域技术人员可以理解，这里的计算机设备是一种能够按照事先设定或存储的指令，自动进行数值计算和/或信息处理的设备，其硬件包括但不限于微处理器、专用集成电路 (Application Specific Integrated Circuit, ASIC)、可编程门阵列 (Field-Programmable Gate Array, FPGA)、数字处理器 (Digital Signal Processor, DSP)、嵌入式设备等。通过根据异常用户群体中的用户数量、特征对异常用户进行评级分级，然后根据评级分级生成靶向标签，当异常用户触发特定操作后根据靶向标签触发不同的处理机制，对异常用户进行针对性的验证处理，解决了现有技术中对异常用户的验证处理不准确的技术问题。

在一个实施例中，提供了一种计算机可读存储介质，其上存储有计算机可读指令，计算机可读指令被处理器执行时实现上述实施例中基于 Neo4j 的异常用户处理方法的步骤，例如图 2 所示的步骤 202 至步骤 208，或者，处理器执行计算机可读指令时实现上述实施例中基于 Neo4j 的异常用户处理装置的各模块/单元的功能，例如图 6 所示模块 602 至模块 608 的功能。为避免重复，此处不再赘述。通过根据异常用户群体中的用户数量、特征对异常用户进行评级分级，然后根据评级分级生成靶向标签，当异常用户触发特定操作后根据靶向标签触发不同的处理机制，对异常用户进行针对性的验证处理，解决了现有技术中对异常用户的验证处理不准确的技术问题。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于

一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器 (ROM)、可编程 ROM (PROM)、电可编程 ROM (EPROM)、电可擦除可编程 ROM (EEPROM) 或闪存。易失性存储器可包括随机存取存储器 (RAM) 或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM (SRAM)、动态 RAM (DRAM)、同步 DRAM (SDRAM)、双数据率 SDRAM (DDRSDRAM)、增强型 SDRAM (ESDRAM)、同步链路 (Synchlink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接 RAM (RDRAM)、直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。

以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形、改进或者对部分技术特征进行等同替换，而这些修改或者替换，并不使相同技术方案的本质脱离本申请个实施例技术方案地精神和范

畴，都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权利要求书

1. 一种基于 Neo4j 的异常用户处理方法，所述方法包括：

对获取到的用户数据进行用户特征提取；

将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

5 基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；以及

根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

10 2. 根据权利要求 1 所述的方法，其中，所述用户特征包括用户的手机号码、IP 地址、设备号以及各所述用户特征之间的关联关系，所述将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体，包括：

通过 Neo4j 算法构建基于各所述用户特征之间关联关系的关系网络，其中，所述关系网络主要由关系节点为所述手机号码、所述 IP 地址以及所述设备号组成；

15 获取所述关系节点的标注特征，并根据所述标注特征对所述关系节点进行分类，得到黑名单节点、白名单节点以及未确认节点，其中，所述标注特征为所述关系节点的初始类别；

为所述关系节点生成属性标签，其中，所述属性标签包括异常、正常；

分别以所述属性标签为异常的手机号码、IP 地址的关系节点为种子节点，从所述关系网络中获取路径长度不大于第一预设值的节点路径，得到节点关系图谱，其中，所述关系图谱包括基于异常手机号码节点的异常号码关系图谱和基于异常 IP 节点的异常 IP 关系图谱；

20 分别对所述异常号码关系图谱与所述异常 IP 关系图谱中的未确认节点进行标签预测，得到所述异常用户群体。

3. 根据权利要求 2 所述的方法，其中，所述分别对所述异常号码关系图谱与所述异常 IP 关系图谱中的未确认节点进行标签预测，得到所述异常用户群体，包括：

对比所述未确认节点的属性标签为异常、正常的邻居节点的数量，得到标签对比结果；

25 根据所述标签对比结果更新所述未确认节点的属性标签；

按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正，并重复属性标签对比、更新的操作，直到每个所述未确认节点的属性标签不再改变或者更新次数达到阈值，将最后一次属性标签更新后关系图谱，作为待预测图谱；

30 断开所述待预测图谱中的异常边，得到多个关系子图，其中，所述异常边为任意两关系节点直接相连组成的边中至少有一个关系节点的属性标签为正常的边；

统计所述关系子图中关系节点的属性标签，若属性标签为异常的手机号码或 IP 地址的数量大于第二预设值，则将所述关系子图中的关系节点对应的用户群体作为所述异常用户群体。

4. 根据权利要求 3 所述的方法，其中，所述按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正，包括：

35 将与所述白名单节点中的 IP 地址具有一度关系的手机号码的属性标签设为正常，其中，所述一度关系是指两个直接相连的关系节点的关联关系。

5. 根据权利要求 2 所述的方法，其中，在所述将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体之后，还包括：

40 比较得到的异常用户群体；

若基于异常手机号码节点的异常用户群体与基于异常 IP 节点的异常用户群体相同，则合并所述异常用户群体，得到合并后的异常用户群体。

45 6. 根据权利要求 1 所述的方法，其中，所述标定特征包括至少一个用户行为特征，不同所述用户行为特征具有不同的运算优先级，所述基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分，包括：

获取所述异常用户群体的用户数量，并根据第一预设评级分数和所述用户数量设定所述

异常用户群体的第一评级分数；

获取所述异常用户群体中具有所述标定特征的异常用户的标定数量，根据所述标定特征的运算优先级确定所述异常用户群体的第二评级分数；

5 统计所述异常用户群体中具有所述共同特征的异常用户的共有数量，并根据所述共有数量以及第二预设评级分数确定所述异常用户群体的第三评级分数；

按照所述预设权重比例联合所述第一评级分数、所述第二评级分数以及所述第三评级分数，得到所述异常用户群体的评分。

7. 根据权利要求1所述的方法，其中，所述靶向标签包括接口调用类型、权限值，所述根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理，包括：

10 若检测到异常用户的所述特定操作，则获取所述异常用户所在异常用户群体的靶向标签；并

根据所述接口调用类型调用对应的接口对所述异常用户的活动权限进行处理；并

判断对所述异常用户的活动权限的处理次数是否超过所述权限值，并将超过所述权限值的异常用户设为锁定用户。

15 8. 一种基于 Neo4j 的异常用户处理装置，包括：

提取模块，用于对获取到的用户数据进行用户特征提取；

检测模块，用于将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

20 评分模块，用于基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；以及

处理模块，用于根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

9. 一种计算机设备，包括存储器和处理器，所述存储器存储有计算机可读指令，所述处理器执行所述计算机可读指令时实现如下基于 Neo4j 的异常用户处理方法的步骤：

25 对获取到的用户数据进行用户特征提取；

将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体；

基于预设权重比例，根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分；以及

30 根据评分为异常用户群体生成对应的靶向标签，并根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理。

10. 根据权利要求9所述的计算机设备，其中，所述用户特征包括用户的手机号码、IP 地址、设备号以及各所述用户特征之间的关联关系，所述将提取到的用户特征输入到 Neo4j 算法中进行预测，得到所述用户数据中具有相同属性标签的用户群体，作为异常用户群体，
35 包括：

通过 Neo4j 算法构建基于各所述用户特征之间关联关系的关系网络，其中，所述关系网络主要由关系节点为所述手机号码、所述 IP 地址以及所述设备号组成；

获取所述关系节点的标注特征，并根据所述标注特征对所述关系节点进行分类，得到黑名单节点、白名单节点以及未确认节点，其中，所述标注特征为所述关系节点的初始类别；

40 为所述关系节点生成属性标签，其中，所述属性标签包括异常、正常；

分别以所述属性标签为异常的手机号码、IP 地址的关系节点为种子节点，从所述关系网络中获取路径长度不大于第一预设值的节点路径，得到节点关系图谱，其中，所述关系图谱包括基于异常手机号码节点的异常号码关系图谱和基于异常 IP 节点的异常 IP 关系图谱；

45 分别对所述异常号码关系图谱与所述异常 IP 关系图谱中的未确认节点进行标签预测，得到所述异常用户群体。

11. 根据权利要求10所述的计算机设备，其中，所述分别对所述异常号码关系图谱与所

述异常 IP 关系图谱中的未确认节点进行标签预测, 得到所述异常用户群体, 包括:

对比所述未确认节点的属性标签为异常、正常的邻居节点的数量, 得到标签对比结果;

根据所述标签对比结果更新所述未确认节点的属性标签;

5 按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正, 并重复属性标签对比、更新的操作, 直到每个所述未确认节点的属性标签不再改变或者更新次数达到阈值, 将最后一次属性标签更新后关系图谱, 作为待预测图谱;

断开所述待预测图谱中的异常边, 得到多个关系子图, 其中, 所述异常边为任意两关系节点直接相连组成的边中至少有一个关系节点的属性标签为正常的边;

10 统计所述关系子图中关系节点的属性标签, 若属性标签为异常的手机号码或 IP 地址的数量大于第二预设值, 则将所述关系子图中的关系节点对应的用户群体作为所述异常用户群体。

12. 根据权利要求 11 所述的计算机设备, 其中, 所述按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正, 包括:

将与所述白名单节点中的 IP 地址具有一度关系的手机号码的属性标签设为正常, 其中, 所述一度关系是指两个直接相连的关系节点的关联关系。

15 13. 根据权利要求 10 所述的计算机设备, 其中, 在所述将提取到的用户特征输入到 Neo4j 算法中进行预测, 得到所述用户数据中具有相同属性标签的用户群体, 作为异常用户群体之后, 还包括:

比较得到的异常用户群体;

20 若基于异常手机号码节点的异常用户群体与基于异常 IP 节点的异常用户群体相同, 则合并所述异常用户群体, 得到合并后的异常用户群体。

14. 根据权利要求 9 所述的计算机设备, 其中, 所述标定特征包括至少一个用户行为特征, 不同所述用户行为特征具有不同的运算优先级, 所述基于预设权重比例, 根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分, 包括:

25 获取所述异常用户群体的用户数量, 并根据第一预设评级分数和所述用户数量设定所述异常用户群体的第一评级分数;

获取所述异常用户群体中具有所述标定特征的异常用户的标定数量, 根据所述标定特征的运算优先级确定所述异常用户群体的第二评级分数;

统计所述异常用户群体中具有所述共同特征的异常用户的共有数量, 并根据所述共有数量以及第二预设评级分数确定所述异常用户群体的第三评级分数;

30 按照所述预设权重比例联合所述第一评级分数、所述第二评级分数以及所述第三评级分数, 得到所述异常用户群体的评分。

15. 根据权利要求 9 所述的计算机设备, 其中, 所述靶向标签包括接口调用类型、权限值, 所述根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理, 包括:

35 若检测到异常用户的所述特定操作, 则获取所述异常用户所在异常用户群体的靶向标签; 并

根据所述接口调用类型调用对应的接口对所述异常用户的活动权限进行处理; 并

判断对所述异常用户的活动权限的处理次数是否超过所述权限值, 并将超过所述权限值的异常用户设为锁定用户。

40 16. 一种计算机可读存储介质, 其上存储有计算机可读指令, 所述计算机可读指令被处理器执行时实现如下基于 Neo4j 的异常用户处理方法的步骤:

对获取到的用户数据进行用户特征提取;

将提取到的用户特征输入到 Neo4j 算法中进行预测, 得到所述用户数据中具有相同属性标签的用户群体, 作为异常用户群体;

45 基于预设权重比例, 根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分; 以及

根据评分为异常用户群体生成对应的靶向标签, 并根据所述靶向标签对检测到的具有特

定操作的异常用户群体进行验证处理。

17. 根据权利要求 16 所述的计算机可读存储介质, 其中, 所述用户特征包括用户的手机号码、IP 地址、设备号以及各所述用户特征之间的关联关系, 所述将提取到的用户特征输入到 Neo4j 算法中进行预测, 得到所述用户数据中具有相同属性标签的用户群体, 作为异常用户群体, 包括:

通过 Neo4j 算法构建基于各所述用户特征之间关联关系的关系网络, 其中, 所述关系网络主要由关系节点为所述手机号码、所述 IP 地址以及所述设备号组成;

获取所述关系节点的标注特征, 并根据所述标注特征对所述关系节点进行分类, 得到黑名单节点、白名单节点以及未确认节点, 其中, 所述标注特征为所述关系节点的初始类别;

为所述关系节点生成属性标签, 其中, 所述属性标签包括异常、正常;

分别以所述属性标签为异常的手手机号码、IP 地址的关系节点为种子节点, 从所述关系网络中获取路径长度不大于第一预设值的节点路径, 得到节点关系图谱, 其中, 所述关系图谱包括基于异常手机号码节点的异常号码关系图谱和基于异常 IP 节点的异常 IP 关系图谱;

分别对所述异常号码关系图谱与所述异常 IP 关系图谱中的未确认节点进行标签预测, 得到所述异常用户群体。

18. 根据权利要求 17 所述的计算机可读存储介质, 其中, 所述分别对所述异常号码关系图谱与所述异常 IP 关系图谱中的未确认节点进行标签预测, 得到所述异常用户群体, 包括:

对比所述未确认节点的属性标签为异常、正常的邻居节点的数量, 得到标签对比结果;

根据所述标签对比结果更新所述未确认节点的属性标签;

按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正, 并重复属性标签对比、更新的操作, 直到每个所述未确认节点的属性标签不再改变或者更新次数达到阈值, 将最后一次属性标签更新后关系图谱, 作为待预测图谱;

断开所述待预测图谱中的异常边, 得到多个关系子图, 其中, 所述异常边为任意两关系节点直接相连组成的边中至少有一个关系节点的属性标签为正常的边;

统计所述关系子图中关系节点的属性标签, 若属性标签为异常的手手机号码或 IP 地址的数量大于第二预设值, 则将所述关系子图中的关系节点对应的用户群体作为所述异常用户群体。

19. 根据权利要求 18 所述的计算机可读存储介质, 其中, 所述按照预设修正条件对属性标签更新后的关系节点进行属性标签的修正, 包括:

将与所述白名单节点中的 IP 地址具有一度关系的手机号码的属性标签设为正常, 其中, 所述一度关系是指两个直接相连的关系节点的关联关系。

20. 根据权利要求 17 所述的计算机可读存储介质, 其中, 在所述将提取到的用户特征输入到 Neo4j 算法中进行预测, 得到所述用户数据中具有相同属性标签的用户群体, 作为异常用户群体之后, 还包括:

比较得到的异常用户群体;

若基于异常手机号码节点的异常用户群体与基于异常 IP 节点的异常用户群体相同, 则合并所述异常用户群体, 得到合并后的异常用户群体。

21. 根据权利要求 16 所述的计算机可读存储介质, 其中, 所述标定特征包括至少一个用户行为特征, 不同所述用户行为特征具有不同的运算优先级, 所述基于预设权重比例, 根据所述异常用户群体中的用户数量、标定特征以及共同特征对所述异常用户群体进行评分, 包括:

获取所述异常用户群体的用户数量, 并根据第一预设评级分数和所述用户数量设定所述异常用户群体的第一评级分数;

获取所述异常用户群体中具有所述标定特征的异常用户的标定数量, 根据所述标定特征的运算优先级确定所述异常用户群体的第二评级分数;

统计所述异常用户群体中具有所述共同特征的异常用户的共有数量, 并根据所述共有数量以及第二预设评级分数确定所述异常用户群体的第三评级分数;

按照所述预设权重比例联合所述第一评级分数、所述第二评级分数以及所述第三评级分数，得到所述异常用户群体的评分。

22. 根据权利要求 16 所述的计算机可读存储介质，其中，所述靶向标签包括接口调用类型、权限值，所述根据所述靶向标签对检测到的具有特定操作的异常用户群体进行验证处理，
5 包括：

若检测到异常用户的所述特定操作，则获取所述异常用户所在异常用户群体的靶向标签；
并

根据所述接口调用类型调用对应的接口对所述异常用户的活动权限进行处理；并

判断对所述异常用户的活动权限的处理次数是否超过所述权限值，并将超过所述权限值
10 的异常用户设为锁定用户。

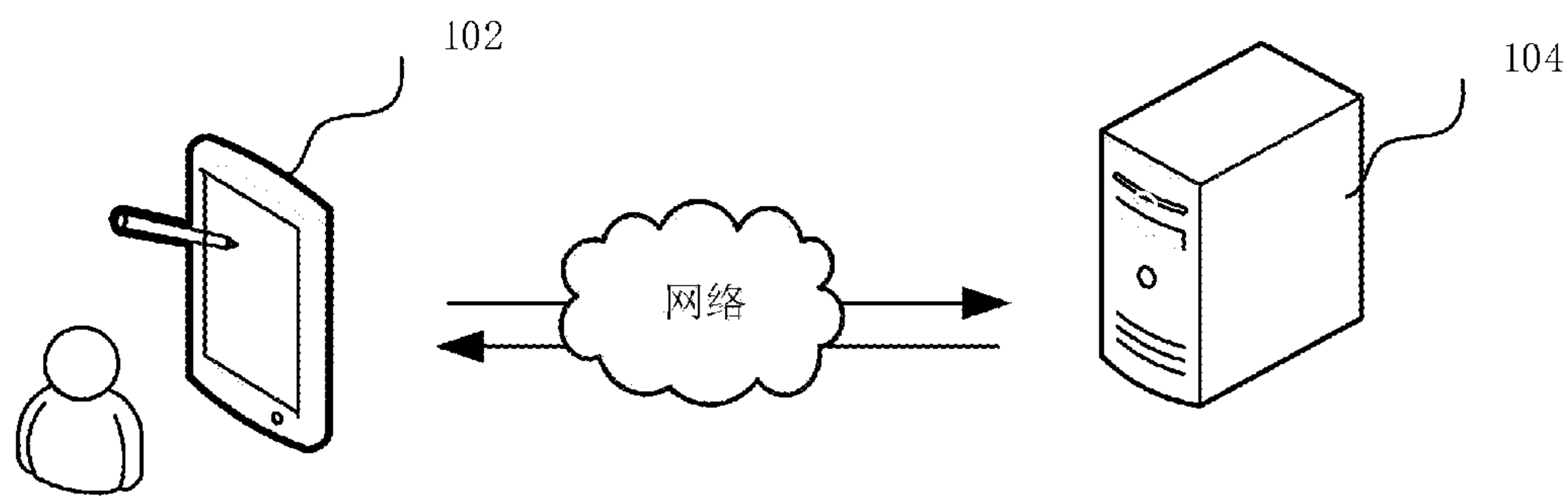


图 1

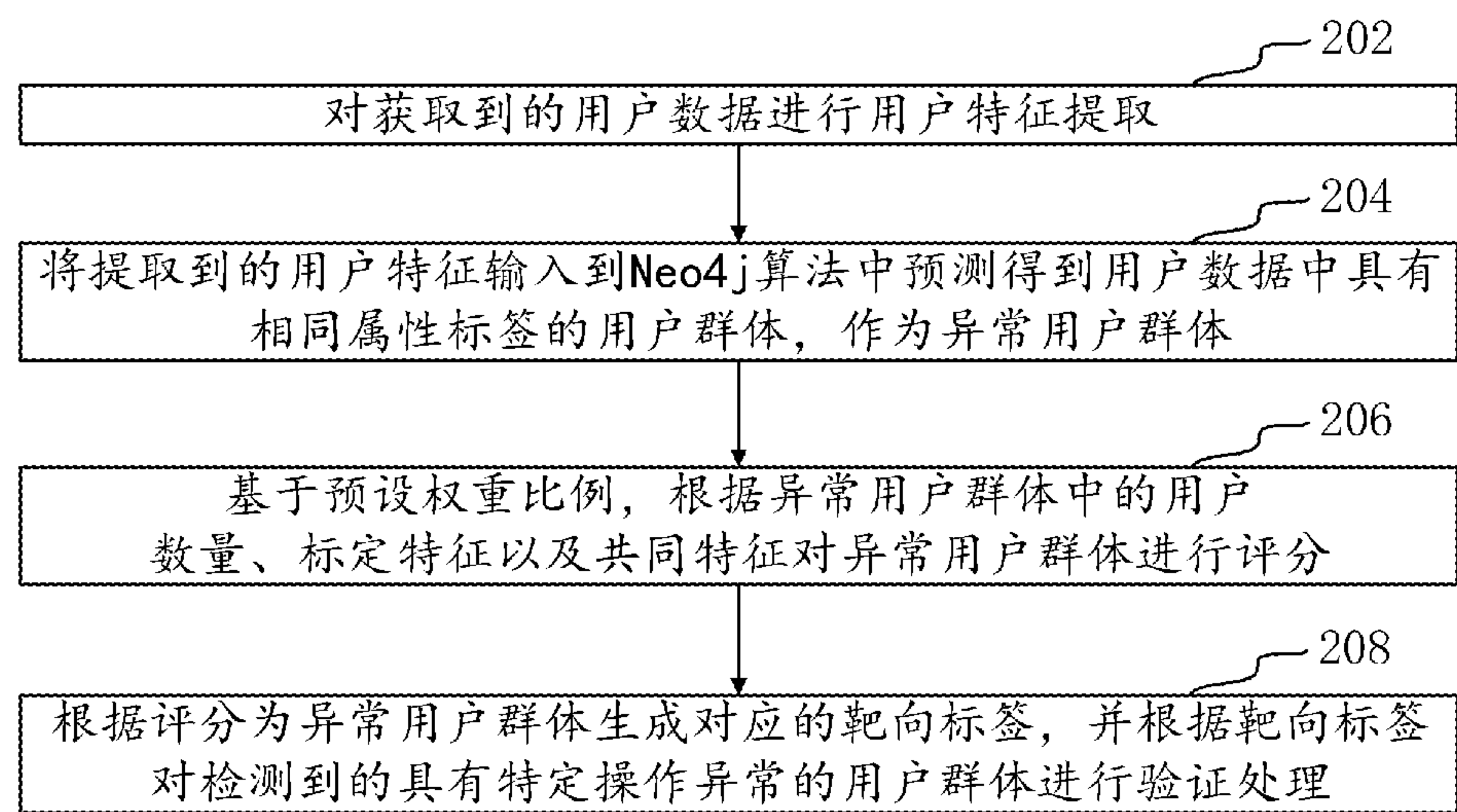


图 2

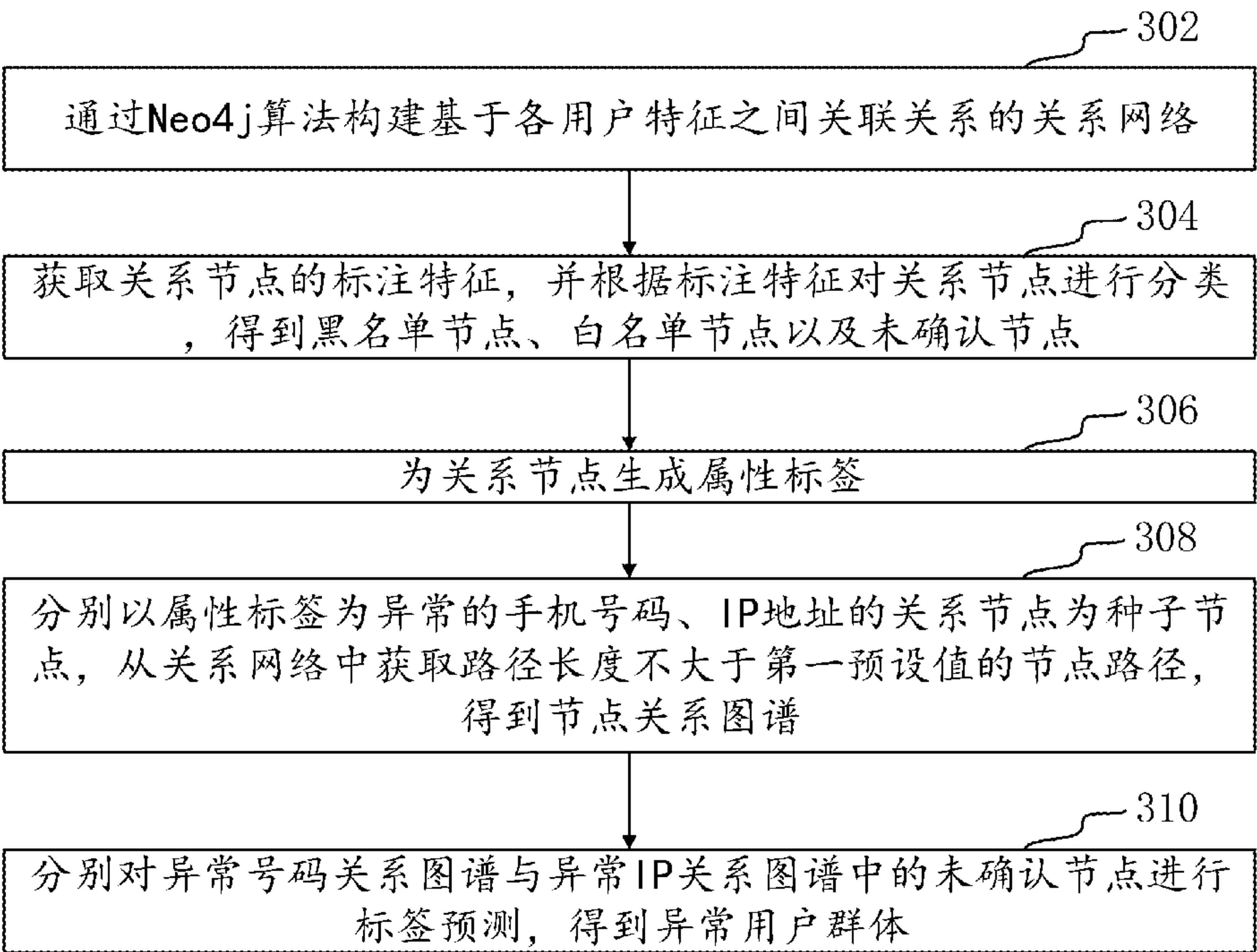


图 3

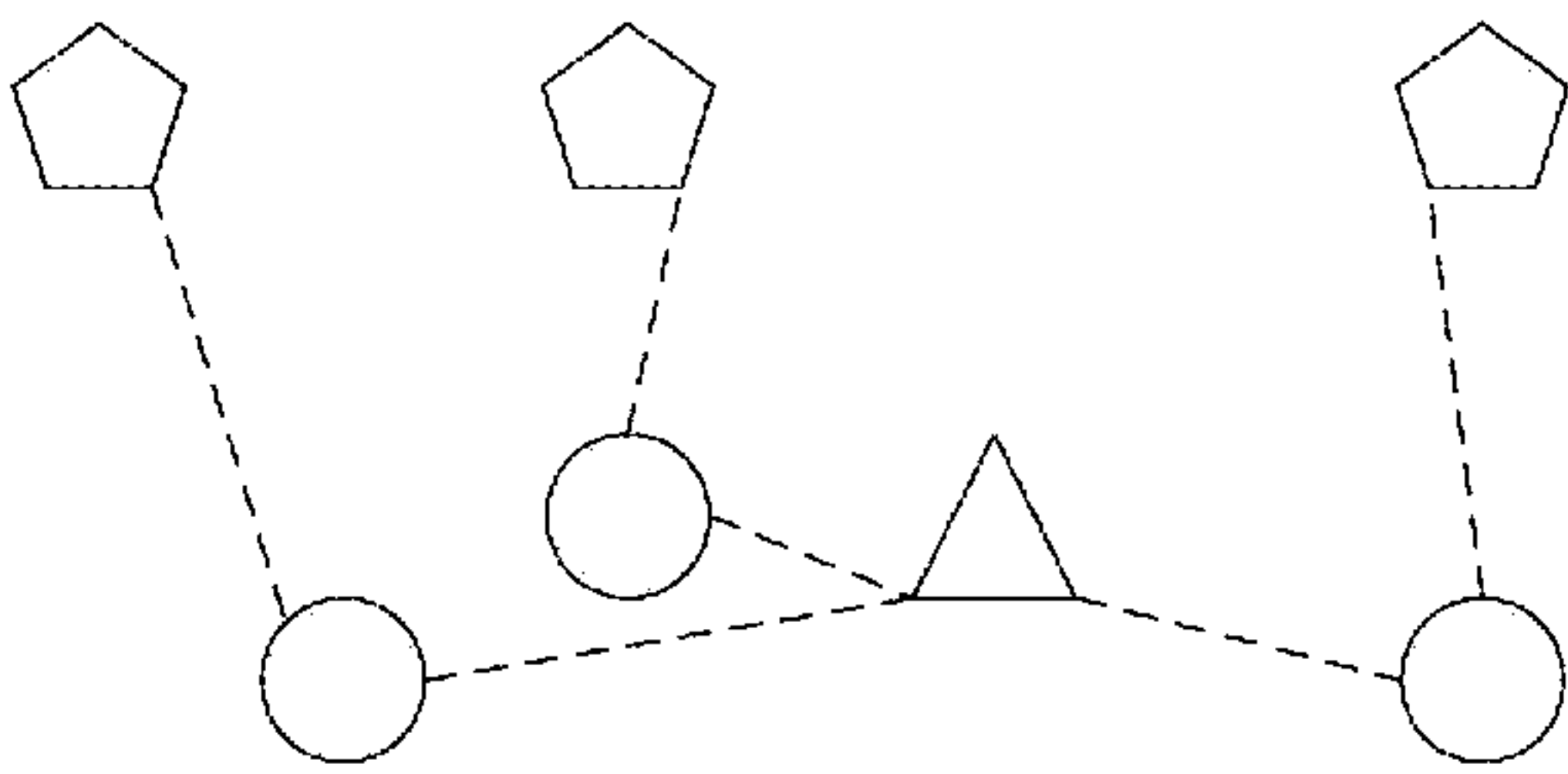
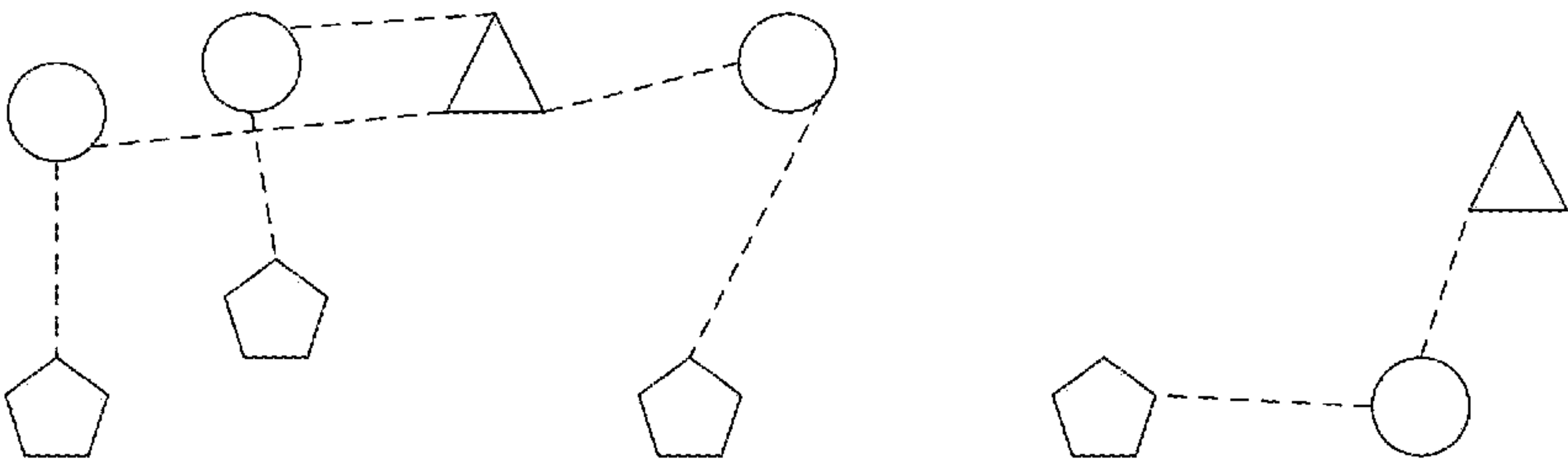


图 4a

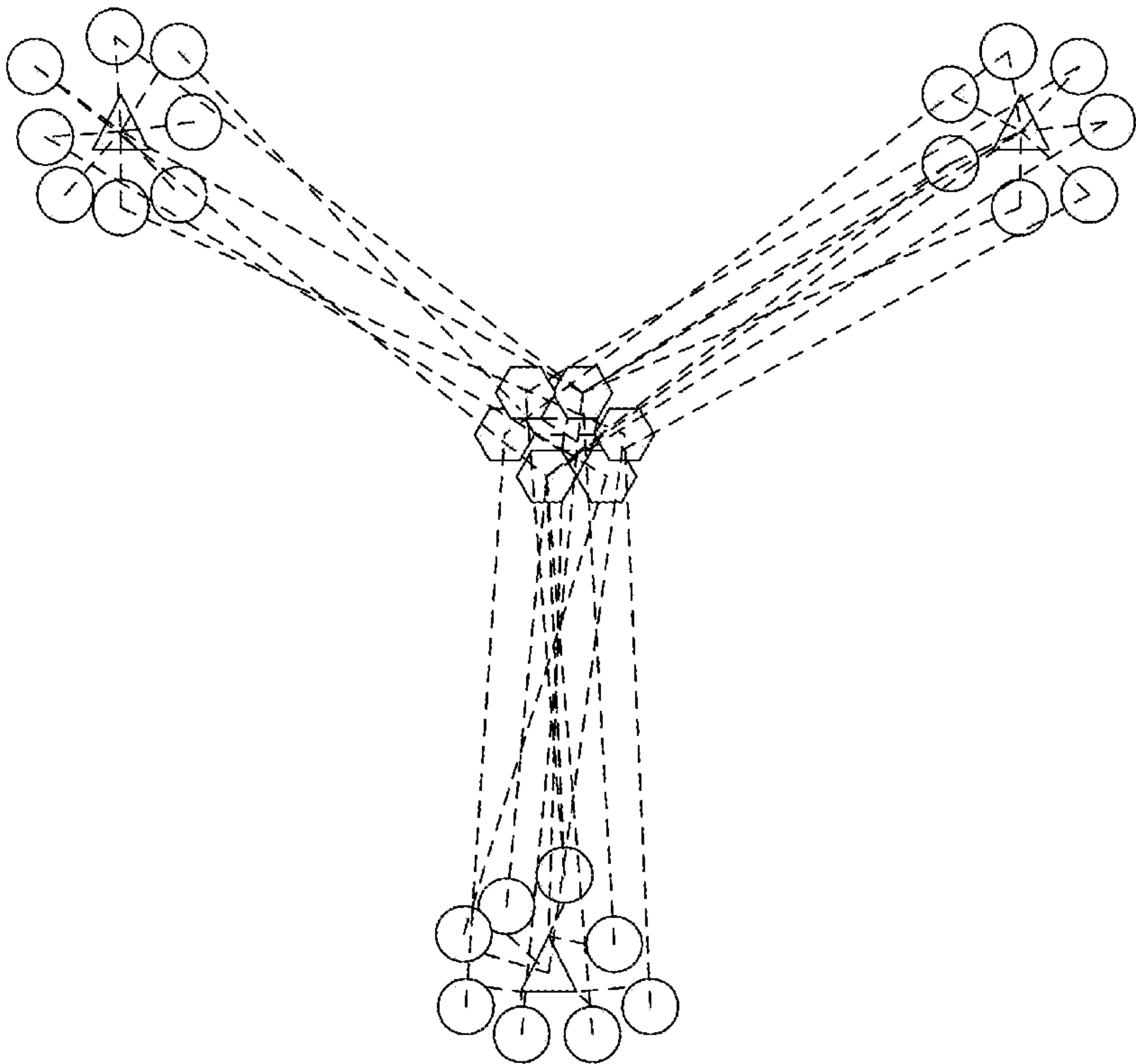


图 4b

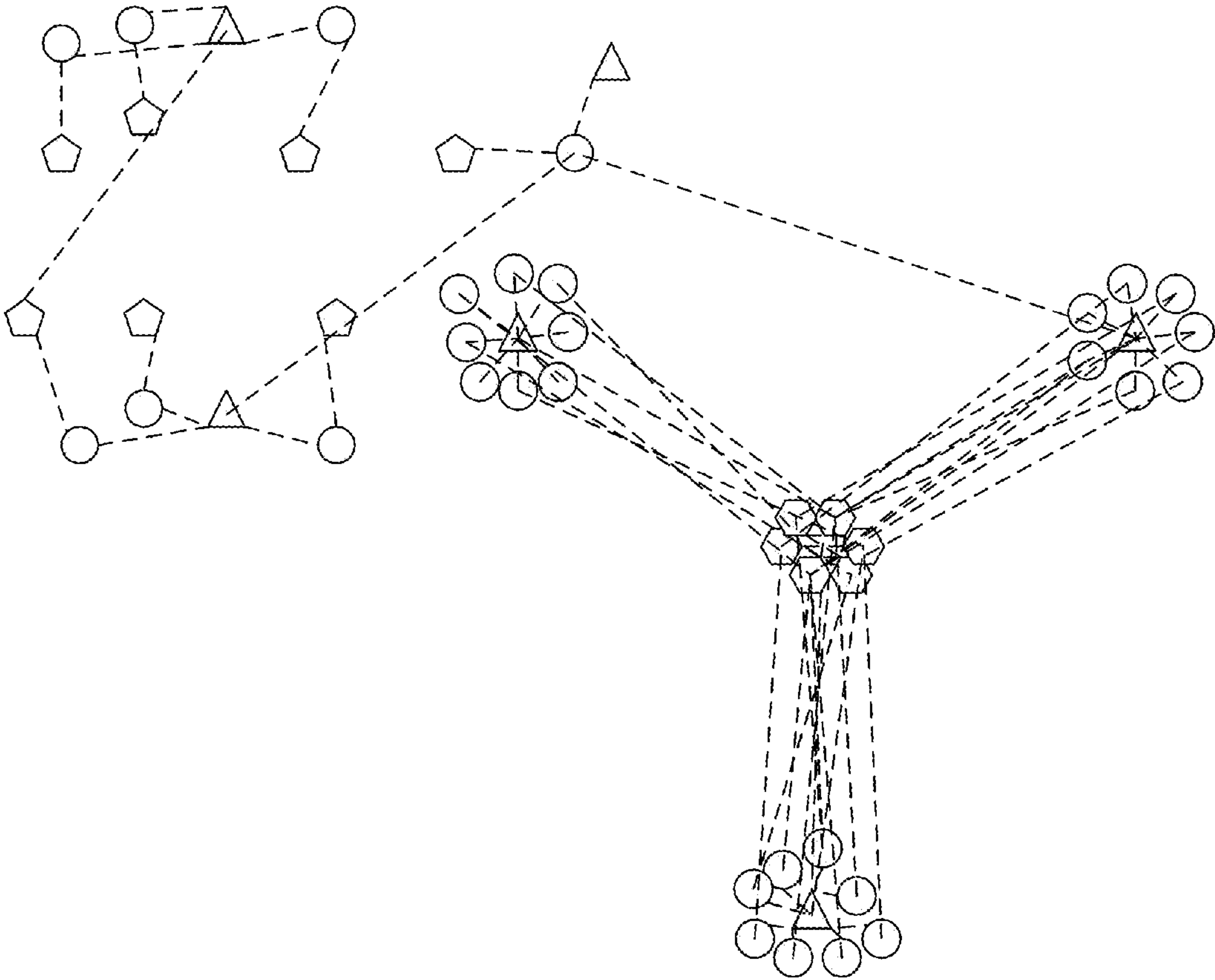


图 4c

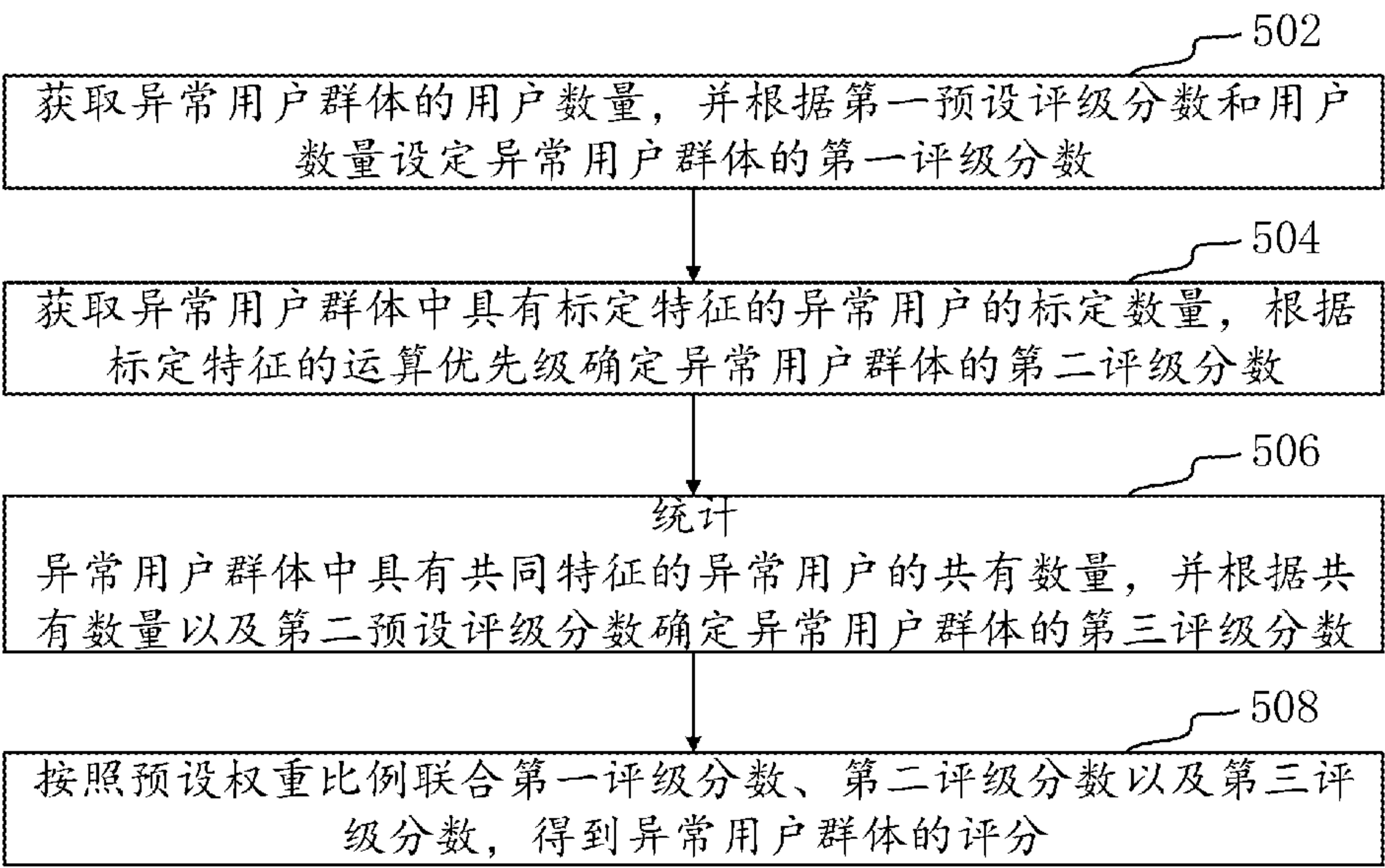


图 5

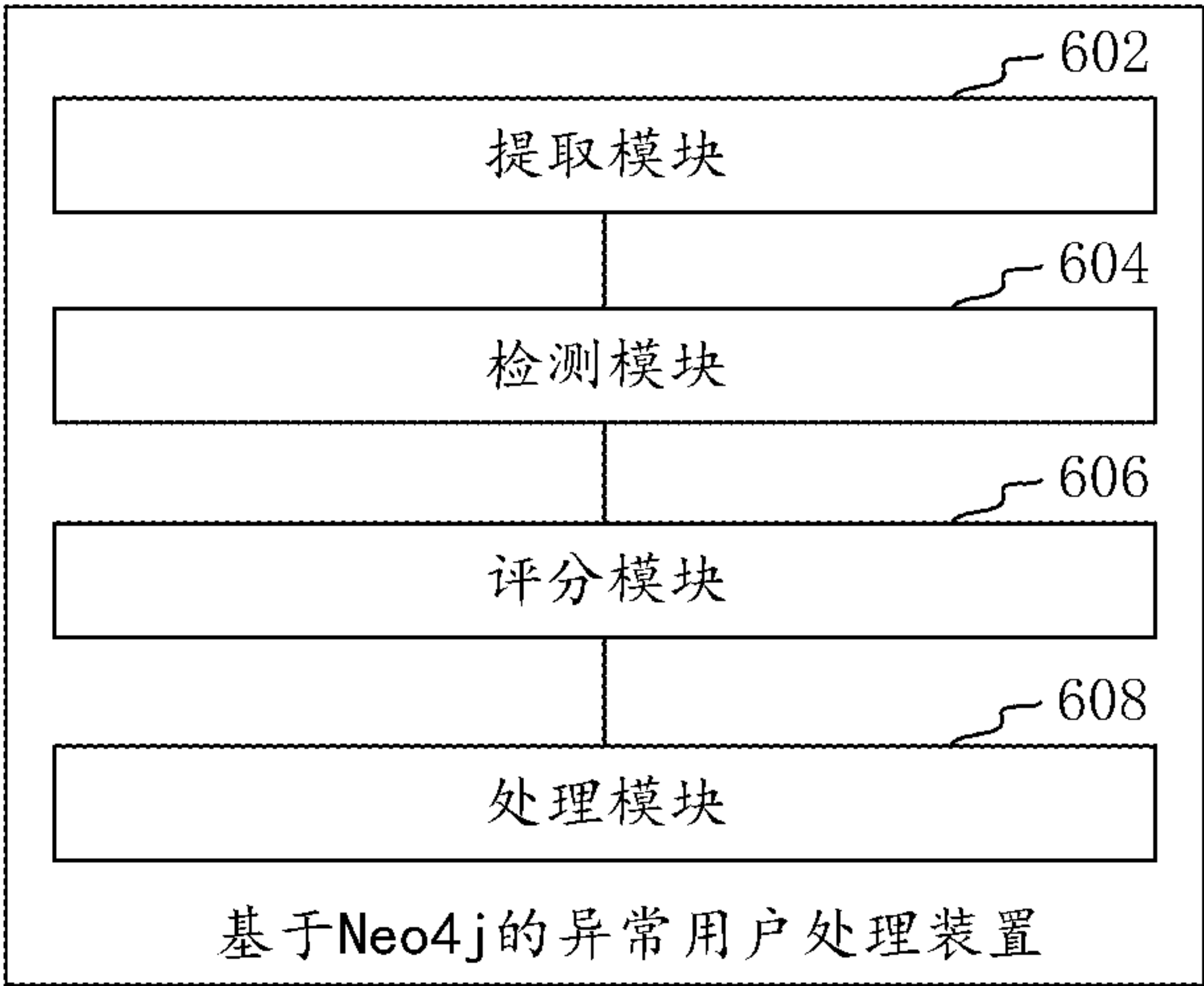


图 6

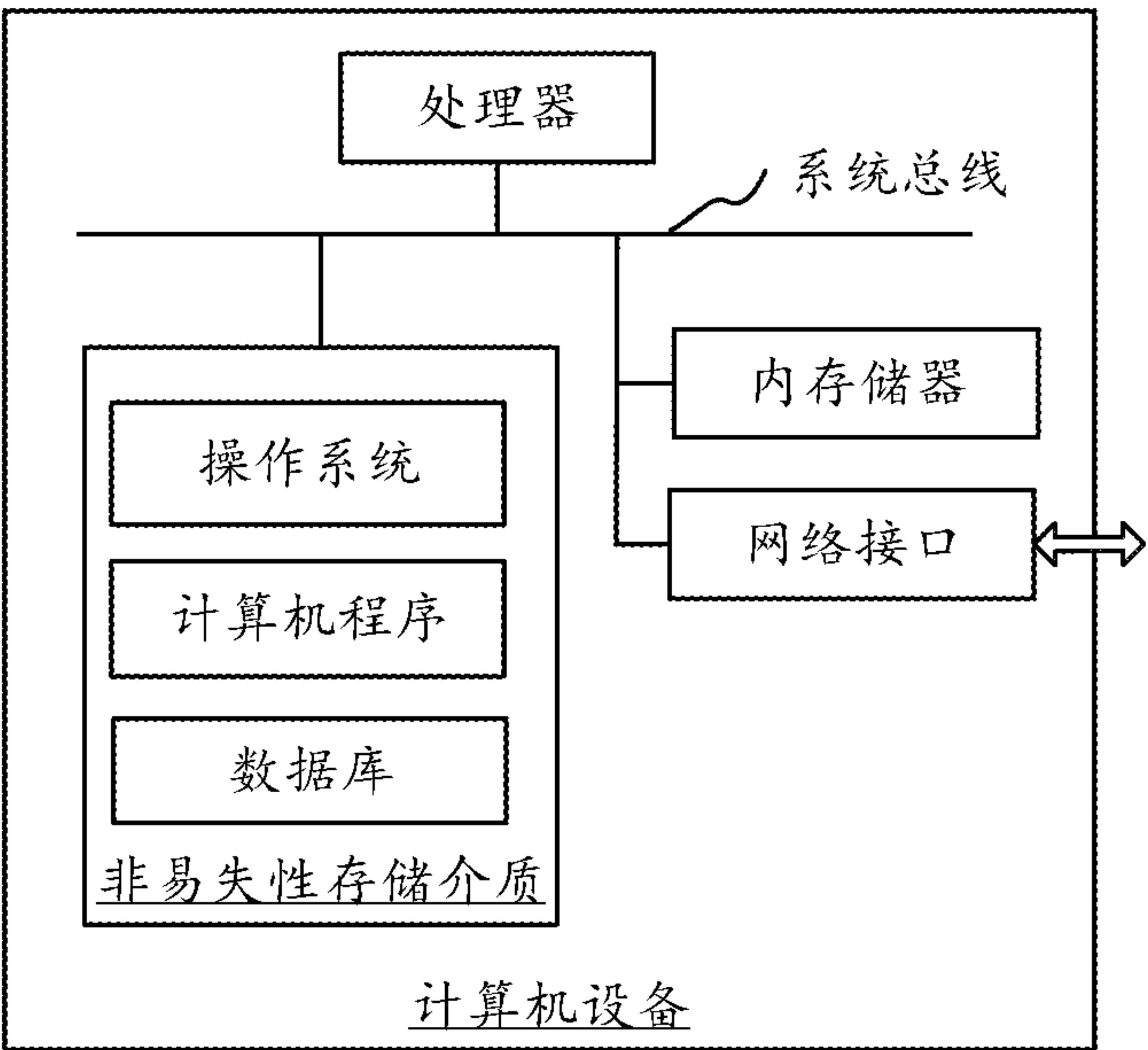


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/122828

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/9536(2019.01)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)

G06F; H04L

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)

CNPAT, EPODOC, WPI, CNKI, IEEE: Neo4j, 标签传播, 异常, 不可信, 非法, 可疑, 欺诈, 舞弊, 作弊, 风险, 羊毛党, 用户, 账号, 账户, 团, 群, 聚类, 簇, 评分, 值, 等级, 认证, 验证; tag, suspicious, malicious, exception, abnormal, illegal, fraud, risk, dangerous, user, account, group, cluster, community, partisan, gang, feature, value, level, security, verification, policy

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 110297912 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 01 October 2019 (2019-10-01) description, paragraphs [0032]-[0166], and figures 1-10	1, 6-9, 14-16, 21-22
Y	CN 107104973 A (BEIJING PANDA MUTUAL ENTERTAINMENT TECHNOLOGY CO., LTD.) 29 August 2017 (2017-08-29) description, paragraphs [0038]-[0103]	1, 6-9, 14-16, 21-22
Y	CN 106301978 A (ALIBABA GROUP HOLDING LIMITED) 04 January 2017 (2017-01-04) description, paragraphs [0032]-[0119], and figures 1-6	1, 6-9, 14-16, 21-22
Y	CN 108446821 A (PING AN LIFE INSURANCE COMPANY OF CHINA, LTD.) 24 August 2018 (2018-08-24) description paragraphs [0044]-[0064]	1, 6-9, 14-16, 21-22
A	US 20200195693 A1 (FORCEPOINT, LLC) 18 June 2020 (2020-06-18) entire document	1-22
A	CN 109284380 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 29 January 2019 (2019-01-29) entire document	1-22

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
“E” earlier application or patent but published on or after the international filing date
“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)
“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

“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
“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
“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
“&” document member of the same patent family

Date of the actual completion of the international search

11 March 2021

Date of mailing of the international search report

26 March 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/122828

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 109831459 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 31 May 2019 (2019-05-31) entire document	1-22

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2020/122828

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110297912	A	01 October 2019	None			
CN	107104973	A	29 August 2017	None			
CN	106301978	A	04 January 2017	None			
CN	108446821	A	24 August 2018	None			
US	20200195693	A1	18 June 2020	None			
CN	109284380	A	29 January 2019	WO	2020062690	A1	02 April 2020
CN	109831459	A	31 May 2019	None			

国际检索报告		国际申请号 PCT/CN2020/122828
A. 主题的分类 G06F 16/9536 (2019.01) 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) G06F; H04L 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)) CNPAT, EPDOC, WPI, CNKI, IEEE: Neo4j, 标签传播, 异常, 不可信, 非法, 可疑, 欺诈, 舞弊, 作弊, 风险, 羊毛党, 用户, 账号, 账户, 团, 群, 聚类, 簇, 评分, 值, 等级, 认证, 验证; tag, suspicious, malicious, exception, abnormal, illegal, fraud, risk, dangerous, user, account, group, cluster, community, partisan, gang, feature, value, level, security, verification, policy		
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
Y	CN 110297912 A (平安科技深圳有限公司) 2019年 10月 1日 (2019 - 10 - 01) 说明书第[0032]-[0166]段、图1-10	1, 6-9, 14-16, 21-22
Y	CN 107104973 A (北京潘达互娱科技有限公司) 2017年 8月 29日 (2017 - 08 - 29) 说明书第[0038]-[0103]段	1, 6-9, 14-16, 21-22
Y	CN 106301978 A (阿里巴巴集团控股有限公司) 2017年 1月 4日 (2017 - 01 - 04) 说明书第[0032]-[0119]段、图1-6	1, 6-9, 14-16, 21-22
Y	CN 108446821 A (中国平安人寿保险股份有限公司) 2018年 8月 24日 (2018 - 08 - 24) 说明书第[0044]-[0064]段	1, 6-9, 14-16, 21-22
A	US 20200195693 A1 (FORCEPOINT, LLC) 2020年 6月 18日 (2020 - 06 - 18) 全文	1-22
A	CN 109284380 A (平安科技深圳有限公司) 2019年 1月 29日 (2019 - 01 - 29) 全文	1-22
A	CN 109831459 A (百度在线网络技术北京有限公司) 2019年 5月 31日 (2019 - 05 - 31) 全文	1-22
☐ 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 2021年 3月 11日		国际检索报告邮寄日期 2021年 3月 26日
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 董洪梅 电话号码 86-(10)-53961528

国际检索报告 关于同族专利的信息				国际申请号 PCT/CN2020/122828			
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	110297912	A	2019年 10月 1日		无		
CN	107104973	A	2017年 8月 29日		无		
CN	106301978	A	2017年 1月 4日		无		
CN	108446821	A	2018年 8月 24日		无		
US	20200195693	A1	2020年 6月 18日		无		
CN	109284380	A	2019年 1月 29日	W0	2020062690	A1	2020年 4月 2日
CN	109831459	A	2019年 5月 31日		无		