

(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
2022 年 3 月 24 日 (24.03.2022)



(10) 国际公布号
WO 2022/057631 A1

(51) 国际专利分类号:
G06F 21/71 (2013.01) **G06F 21/60** (2013.01)
G06F 21/62 (2013.01)

(21) 国际申请号: PCT/CN2021/116129

(22) 国际申请日: 2021 年 9 月 2 日 (02.09.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010966318.0 2020 年 9 月 15 日 (15.09.2020) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 王才华(WANG, Caihua); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。 刘洋(LIU, Yang); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房 (部位: 自编 01-03 和 08-12 单元) (仅限办公用途), Guangdong 510623 (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, 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,

(54) Title: DATA PROCESSING METHOD AND SYSTEM BASED ON NODE GROUP, AND DEVICE AND MEDIUM

(54) 发明名称: 基于节点群的数据处理方法、系统、设备和介质

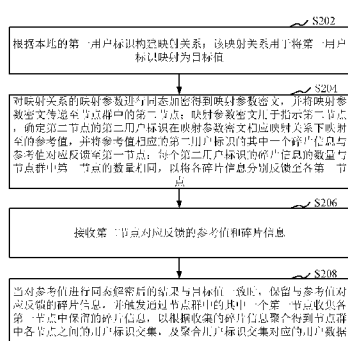


图 2

S202 Construct a mapping relationship according to a local first user identifier, wherein the mapping relationship is used for mapping the first user identifier into a target value

S204 Perform homomorphic encryption on a mapping parameter of the mapping relationship to obtain mapping parameter ciphertext, and transmit the mapping parameter ciphertext to a second node in the node group, wherein the mapping parameter ciphertext is used for indicating the second node, determining a reference value to which a second user identifier of the second node is mapped under the mapping relationship corresponding to the mapping parameter ciphertext, and feeding one piece of fragment information of the second user identifier corresponding to the reference value back to the first node

S206 Receive the reference value and the fragment information, which are correspondingly fed back by the second node

S208 When a homomorphic decryption result of the reference value is consistent with the target value, reserve the fragment information that is fed back correspondingly to the reference value, and trigger the collection of the fragment information reserved in each first node by means of one of the first nodes in the node group, so as to aggregate the collected fragment information to obtain a user identifier intersection between the nodes in the node group, and aggregate user data corresponding to the user identifier intersection

(57) Abstract: The present application relates to a data processing method and system based on a node group, and a device and a medium. The method is executed by one of first nodes in a node group, and comprises: constructing a mapping relationship according to a local first user identifier, wherein the mapping relationship is used for mapping the first user identifier into a target value; performing homomorphic encryption on a mapping parameter of the mapping relationship to obtain mapping parameter ciphertext, and transmitting the mapping parameter ciphertext to a second node in the node group; receiving a reference value and fragment information, which are correspondingly fed back by the second node; when a homomorphic decryption result of the reference value is consistent with the target value, reserving the fragment information that is fed back correspondingly to the reference value, and triggering the collection of the fragment information reserved in each first node by means of one of the first nodes in the node group, so as to aggregate the collected fragment information to obtain a user identifier intersection between the nodes in the node group, and aggregate user data corresponding to the user identifier intersection.

(57) 摘要: 本申请涉及一种基于节点群的数据处理方法、系统、设备和介质。该方法由节点群中的其中一个第一节点执行, 包括: 根据本地的第一用户标识构建映射关系; 映射关系用于将第一用户标识映射为目标值; 对映射关系的映射参数进行同态加密得到映射参数密文, 并将映射参数密文传递至节点群中的第二节点; 接收第二节点对应反馈的参考值和碎片信息; 当对参考值进行同态解密后的结果与目标值一致时, 保留与参考值对应反馈的碎片信息, 并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息, 以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集, 及聚合用户标识交集对应的用户数据。

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))。

说明书

发明名称：基于节点群的数据处理方法、系统、设备和介质

本申请要求于 2020 年 09 月 15 日提交中国专利局，申请号为 202010966318.0，申请名称为“基于节点群的数据处理方法、系统、设备和介质”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。

技术领域

本申请涉及计算机技术领域，特别是涉及一种基于节点群的数据处理方法、装置、设备和介质。

背景技术

随着计算机技术的发展，在互联网上处理数据越来越普遍，隐私集合求交是安全多方计算领域研究的基础问题之一，对于若干参与者，每一方拥有一个私有的数据集合，目标是让所有参与方获悉这些集合的公共交集，同时保证非交集的内容不泄露。

传统技术中，数据处理方对隐私数据的访问权限过大，很容易造成安全隐患。

发明内容

根据本申请提供的各种实施例，提供一种基于节点群的数据处理方法、系统、设备和介质。

一种基于节点群的数据处理方法，由所述节点群中的其中一个第一节点执行，所述方法包括：根据本地的第一用户标识构建映射关系；所述映射关系用于将所述第一用户标识映射为目标值；对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点；所述映射参数密文用于指示所述第二节点，确定所述第二节点的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息与所述参考值对应反馈至所述第一节点；接收所述第二节点对应反馈的参考值和碎片信息；以及，当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

一种基于节点群的数据处理方法，由所述节点群中的第二节点执行，所述方法包括：接收所述节点群中多个第一节点分别发送的映射参数密文；对各所述映射参数密文同态解密得到的映射参数明文相应的映射关系，用于将各所述映射参数密文所源自第一节点的第一用户标识映射为目标值；将本地的第二用户标识拆分为多个碎片信息；对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述映射参数密文所源自的第一节点；以及，对应反馈至所述第一节点的参考值和碎片信息，用于指示所述第一节点在对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

一种基于节点群的数据处理系统，所述系统所应用的节点群包括多个第一节点和一个第二节点；其中，所述第一节点用于根据本地的第一用户标识构建映射关系；对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点；所述映射参数用于将所述第一用户标识映射为目标值；所述第二节点用于接收所

述节点群中多个第一节点分别发送的映射参数密文；将本地的第二用户标识拆分为多个碎片信息；对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述映射参数密文所源自的第一节点；所述第一节点还用于接收所述第二节点对应反馈的参考值和碎片信息；以及，当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

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上述基于节点群的数据处理方法的步骤。

一个或多个存储有计算机可读指令的非易失性存储介质，所述计算机可读指令被一个或多个处理器执行时实现上述基于节点群的数据处理方法的步骤。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征、目的和优点将从说明书、附图以及权利要求书变得明显。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用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的前提下，还可以根据这些附图获得其他的附图。

图1为一个实施例中基于节点群的数据处理方法的应用环境图；

图2为一个实施例中基于节点群的数据处理方法的流程示意图；

图3为一个实施例中基于节点群的数据处理方法的应用示意图；

图4为另一个实施例中基于节点群的数据处理方法的流程示意图；

图5为另一个实施例中基于节点群的数据处理方法的流程示意图；

图6为一个实施例中基于节点群的数据处理系统中的数据流向图；

图7为一个实施例中基于节点群的数据处理中数据变化的示意图；

图8为一个实施例中基于节点群的数据处理装置的结构框图；

图9为另一个实施例中基于节点群的数据处理装置的结构框图；

图10为另一个实施例中基于节点群的数据处理装置的结构框图；以及

图11为一个实施例中计算机设备的内部结构图。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，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。

人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说，人工智能是计算机科学的一个综合技术，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法，使机器具有感知、推理与决策的功能。

人工智能技术是一门综合学科，涉及领域广泛，既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

机器学习(Machine Learning, ML)是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以

获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、式教学习等技术。

随着人工智能技术研究和进步，人工智能技术在多个领域展开研究和应用，例如常见的智能家居、智能穿戴设备、虚拟助理、智能音箱、智能营销、无人驾驶、自动驾驶、无人机、机器人、智能医疗、智能客服等，相信随着技术的发展，人工智能技术将在更多的领域得到应用，并发挥越来越重要的价值。

本申请实施例提供的方案涉及人工智能的机器学习等技术，具体通过如下实施例进行说明：

本申请提供的基于节点群的数据处理方法，可以应用于如图 1 所示的应用环境中。在一个实施例中，节点群是多个节点的集合，节点群中包括两种角色类型的计算机设备（以下简称节点）：第一节点 102 和第二节点 104，第一节点的数量可以多于一个，第二节点为一个；节点群中任一节点既可作为第一节点，也可作为第二节点。第一节点与第二节点之间可通过网络进行通信。节点群中的各节点之间相互独立。在一个实施例中，节点群中的各节点相互之间不具有信任关系。

具体地，第一节点可用于根据本地的第一用户标识构建映射关系；对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点；其中，该映射关系用于将第一用户标识映射为目标值。第二节点可用于接收节点群中多个第一节点分别发送的映射参数密文；将本地的第二用户标识拆分为多个碎片信息；对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至映射参数密文所源自的第一节点；其中，每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，以将各碎片信息分别反馈至各第一节点。第一节点还可接收第二节点对应反馈的参考值和碎片信息；当对参考值进行同态解密后的结果与目标值一致时，保留与参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，及聚合用户标识交集对应的用户数据。

在一个实施例中，节点群中还可以包括一种角色类型的节点：聚合节点，该聚合节点为节点群中的其中一个第一节点。聚合节点与其他第一节点之间可通过网络进行通信。聚合节点用于接收节点群中其他第一节点发送的被保留的碎片信息；当对应相同碎片标识的碎片信息的数量达到第一节点的数量时，聚合对应相同碎片标识的碎片信息，得到节点群中各节点之间的公共用户标识；根据公共用户标识得到节点群中各节点之间的用户标识交集，以聚合用户标识交集对应的用户数据。

在一个实施例中，聚合用户标识交集对应的用户数据可用作机器学习模型的训练样本。

其中，节点可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。

在一个实施例中，如图 2 所示，提供了一种基于节点群的数据处理方法，以该方法应用于图 1 所示节点群中的其中一个第一节点为例进行说明，该基于节点群的数据处理方法包括以下步骤：

步骤 202，根据本地的第一用户标识构建映射关系；该映射关系用于将第一用户标识映射为目标值。

其中，本地的第一用户标识是第一节点上存储的用户标识。用户标识用于唯一标识一个用户。可以理解，在本申请中，节点群中的各节点均存储有用户标识，以及与用户标识对应

存储的用户数据。节点群中的第一节点多于一个，每个第一节点上的用户标识都可以统称为第一用户标识，但是对应不同的第一节点。

举例说明，图3示出了一个实施例中基于节点群的数据处理方法的应用示意图。参考图3，可以看到节点群中至少包括第一节点F1（参与方1）和第一节点F2（参与方2），对于第一节点F1，其上包括第一用户标识（样本ID）如：0、1、3、6、11等。对于第一节点F2，其上包括第一用户标识（样本ID）如：1、2、4、5、7等。

其中，映射关系是两个元素集之间元素相互对应的关系。比如，两个元素集A与B之间存在着对应关系，而且对于元素集A中的每一个元素a，元素集B中总有唯一的一个元素b与它对应，这种对应关系即为从A到B的映射关系。本申请中所涉及的映射关系，用于将属于一个元素集的第一用户标识映射为属于另一元素集的目标值。目标值具体可以是特定的一个或者多个数值。

具体地，节点群中的任一第一节点可以获取本地的第一用户标识，以及待映射至的目标值，并直接基于该第一用户标识构建映射关系，或者，对第一用户标识进行处理后，再基于处理后得到的结果构建映射关系。

在一个实施例中，节点群中第一节点的数量为至少两个。

步骤204，对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点；映射参数密文用于指示第二节点，确定第二节点的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息与参考值对应反馈至第一节点；每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，以将各碎片信息分别反馈至各第一节点。

其中，映射参数是用于描述映射关系的参数。映射参数用于唯一确定一个映射关系。同态加密是基于数学难题的计算复杂性理论的密码学技术。对经过同态加密的数据进行处理得到一个输出，将这一输出进行同态解密，其解密结果与对未加密的原始数据进行相同的处理所得到的输出是相同的。映射参数密文是对映射关系的映射参数进行同态加密后的结果。映射参数密文也可用于唯一确定一个映射关系，该映射关系与映射参数所确定的映射关系不同。

第二用户标识是第二节点上存储的用户标识。第二节点的第二用户标识在映射参数密文相应映射关系下映射至的参考值，可以是第二节点上的第二用户标识直接在映射参数密文相应映射关系下映射至的数值；也可以是第二用户标识在进行处理后，处理后的结果在映射参数密文相应映射关系下映射至的参考值。碎片信息是第二用户标识的部分信息，可以通过对第二用户标识拆分得到，也可以通过对第二用户标识进行处理后得到的数据拆分得到。

举例说明，继续参考图3，可以看到节点群中包括一个第二节点S（参与方n），对于第二节点S，其上包括第一用户标识（样本ID）如：0、1、6、11、20等。

具体地，节点群中的任一第一节点可以在构建出映射关系后，将该映射关系的映射参数进行同态加密得到映射参数密文，再将映射参数密文传递至节点群中的第二节点。第二节点在接收到映射参数密文后，基于该映射参数密文确定出一个映射关系，确定本地的第二用户标识在该映射关系下映射至的参考值；第二节点再将参考值和参考值所源自的第二用户标识对应反馈至第一节点。

其中，每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，这样第二节点在接收到每个第一节点发送的映射参数密文，确定第二用户标识在每个映射参数密文相应映射关系下映射至的参考值后，对每个参考值分别对应一个碎片信息后反馈至每个第一节点。即参考值和碎片信息是成对反馈的。

举例说明，假设节点群中包括两个第一节点F1和F2，以及第二节点S。F1根据本地的第一样本标识构建出映射关系Y1，并对Y1的映射参数ym1同态加密得到ys1传递至S；F2根据本地的第一样本标识构建出映射关系Y2，并对Y2的映射参数ym2同态加密得到ys2传递至S。S在接收到ys1和ys2后，分别确定本地的第二用户标识在ys1和ys2相应映射关系下映射至的参考值。例如，对于第二用户标识SID1，第二节点可以基于SID1拆分得到2个

碎片信息 P11 和 P12, 确定 SID1 在 ys1 相应映射关系下映射至的参考值 C11, 将 C11 和 P11 (或 P12) 作为数据对发送给 F1; 确定 SID1 在 ys2 相应映射关系下映射至的参考值 C12, 将 C12 和 P12 (或 P11) 作为数据对发送给 F2。需要说明的是, 多个第一节点不限制同时向第二传递映射参数密文; 第二节点可以分别接收到各映射参数密文, 其在每接收到一个映射参数密文时, 即可进行后续的处理, 无需等待接收到所有第一节点传递的映射参数密文之后再进行处理。

步骤 206, 接收第二节点对应反馈的参考值和碎片信息。

具体地, 节点群中的任一第一节点可以接收第二节点反馈的多个数据对。数据对的数量与第二节点上第二用户标识的数量相同。每个数据对包括一个参考值和一个碎片信息。在一个数据对中的参考值是基于该数据对中碎片信息所属的第二用户标识映射得到。

在另外的实施例中, 在步骤 202 之前, 可先对节点群中各节点的用户标识进行快速的近似求交, 过滤掉一部分不可能属于交集的样本标识, 再基于过滤后剩余的用户标识执行本申请提供的基于节点群的数据处理方法进行精确求交。那么此时, 第一节点接收到的数据对的数量小于第二节点上第二用户标识的数量。

步骤 208, 当对参考值进行同态解密后的结果与目标值一致时, 保留与参考值对应反馈的碎片信息, 并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息, 以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集, 及聚合用户标识交集对应的用户数据。

可以理解, 由于对经过同态加密的数据进行处理得到一个输出, 将这一输出进行同态解密, 其解密结果与对未加密的原始数据进行相同的处理所得到的输出是相同的。那么, 第一节点在接收到第二节点对应反馈的参考值和碎片信息后, 可对该参考值进行同态解密, 经过同态解密的解密结果与参考值所源自的第二用户标识在映射参数相应映射关系下映射至的数值一致, 基于此可将解密结果与目标值进行比较, 以判断本地与第二节点是否存在共有的用户标识。

具体地, 在对该参考值进行同态解密的解密结果与目标值一致时, 判定该参考值所源自的用户标识是本地与第二节点共有的用户标识。

其中, 保留的碎片信息用于表示其所源自的用户标识为至少两个节点之间共有的用户标识。进一步地, 第一节点在判定接收到的参考值所源自的用户标识是本地与第二节点共有的用户标识时, 则保留与该参考值对应接收到的碎片信息。继而可以触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息, 以在集齐公共用户标识的所有碎片信息后, 根据这些碎片信息恢复出公共用户标识, 得到节点群中各节点之间的用户标识交集, 从而可以聚合用户标识交集对应的用户数据, 用于后续的处理步骤。比如用作模型训练数据等。在另外的实施例中, 第一节点在判定接收到的参考值所源自的用户标识是本地与第二节点共有的用户标识时, 也可将与该参考值对应接收到的碎片信息标记为有效碎片。

举例说明, 假设节点群中包括第一节点 F1 和第二节点 S。F1 根据本地的第一样本标识构建出映射关系 Y1, 并对 Y1 的映射参数 ym1 同态加密得到 ys1 传递至 S。S 在接收到 ys1 后, 确定本地的第二用户标识 SID1 在 ys1 相应映射关系下映射至的参考值 C11, 并将 C11 和 SID1 的一个碎片信息 P11 对应反馈给第一节点 F1。第一节点在接收到对应反馈的 C11 和 P11 后, 对 C11 进行同态解密得到 D11, 比较 D11 与目标值 M。在 D11=M 时, 判定 SID1 为第一节点 F1 和第二节点 S 共有的用户标识, 则保留 C11 对应反馈的 P11。

在一些实施例中, 在对该参考值进行同态解密的解密结果与目标值不一致时, 判定该参考值所源自的用户标识不属于本地的第一用户标识。此时, 第一节点可以将该参考值对应反馈的碎片信息丢弃、置零或者标记为无效碎片。

举例说明, 第一节点在接收到对应反馈的 C11 和 P11 后, 对 C11 进行同态解密得到 D11, 比较 D11 与目标值 M。在 D11 不等于 M 时, 判定 SID1 不是第一节点 F1 和第二节点 S 共有的用户标识, 则将 C11 对应反馈的 P11 置零。

上述基于节点群的数据处理方法，节点群中的第一节点根据本地的第一用户标识构建映射关系，该映射关系用于将第一用户标识映射为目标值。那么在第一节点将对映射关系的映射参数进行同态加密得到映射参数密文传递至节点群中的第二节点后，第二节点可以在密文空间下基于本地的第二用户标识进行同态运算，得到第二用户标识在映射参数密文相应映射关系下映射至的参考值，然后将参考值相应的第二用户标识的其中一个碎片信息与参考值一起反馈至第一节点。第一节点再在对接收到的参考值进行同态解密后的结果与目标值一致时，保留与接收到的参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，以及聚合用户标识交集对应的用户数据。这样一方面，由于第一节点发送给第二节点的是映射参数密文，第二节点无法通过密文反推第一节点的第一用户标识，而且同态加密提供了一种对加密数据进行处理的功能，可以在不获取明文的情况下对密文进行某种运算，解密后等价于对明文进行相应的运算，以保证第一节点处的参考值与目标值的校验效力；另一方面，由于每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，这样每个第二用户标识的各碎片信息会分别反馈至各第一节点，第一节点也不会得到完整的第二节点的第二用户标识，除最后获得的节点群中各节点之间的用户标识交集外，各节点均不能通过自己获取的信息来盗取其他节点的隐私信息，从而保证了各节点隐私数据的安全，而且还可以在无需引入第三方的情况下进行数据处理。

需要说明的是，第一节点根据本地的第一用户标识构建出的映射关系，有且只会将第一用户标识映射为目标值。这样，当对反馈的参考值进行同态解密后的结果与目标值一致时，则可判定第二节点上存在与本地相同的用户标识。

在一个实施例中，根据本地的第一用户标识构建映射关系，包括：将本地的第一用户标识转换为中间值；以中间值为多项式的解，构建目标多项式；目标多项式将第一用户标识映射为统一的目标值，目标值为零。对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点，包括：对目标多项式的多项式系数进行同态加密得到多项式系数密文，并将多项式系数密文传递至节点群中的第二节点。

其中，中间值是用于基于用户标识构建多项式的中间数值。多项式是由多个单项式相乘组成的代数式，其中每个单项式根据一个中间值得到。以中间值为多项式的解构建的目标多项式将第一用户标识映射为统一的目标值，即零。

具体地，第一节点可以中间值为一元多项式的解，构建目标一元多项式；目标一元多项式将第一用户标识映射为统一的目标值，且目标值为零。

举例说明，假设第一节点 F1 将本地的第一用户标识转换为的中间值包括：h1、h2、h3、...、hi、...、hn，其中，n 为第一节点 F1 上第一用户标识的数量。第一节点 F1 根据每个中间值得到一个单项式，如 $y_1 = (x-h_1)$ ， $y_2 = (x-h_2)$ 等。将这些单项式相乘即得到目标多项式：

$$P_{\text{loy}} = (x-h_1) * (x-h_2) * \dots * (x-h_n) = \prod_{i=1}^n (x-h_i)$$

在一个实施例中，将本地的第一用户标识转换为中间值，包括：获取节点群中各节点之间共享的转换函数；通过转换函数将第一用户标识转换为中间值，以使得节点群中各节点之间相同的用户标识转换为相同的中间值。

其中，节点群中各节点之间共享的转换函数，是节点群中各节点统一使用以将用户标识转化为自然数值的函数。转换函数具体可以是哈希函数或者加密哈希函数。哈希函数也可称为散列函数，可把任意长度的输入压缩为预设长度的输出，该输出为哈希值或称为散列值。

具体地，第一节点可将本地的第一用户标识输入节点群中各节点之间共享的转换函数，获得转换得到的中间值。或者，第一节点可将本地的第一用户标识进行偏移后，输入节点群中各节点之间共享的转换函数，获得转换得到的中间值。其中，对第一用户标识进行偏移的偏移量是节点群中各节点之间共享的偏移量。

在本实施例中，由于节点群中各节点之间共享转换函数，那么不同的节点对于相同的用

户标识通过转换函数转换得到的中间值是相同的。这样为后续基于碎片信息回溯用户标识提供了保证。

进一步地，第一节点构建出目标多项式后，可将目标多项式展开为多个单项式相加组成的代数式，再获取展开后的目标多项式的多项式系数，对该多项式系数进行同态加密得到多项式系数密文，并将多项式系数密文传递至所述节点群中的第二节点。由于多个单项式相加组成的代数式可由每个单项式的系数唯一确定，那么展开后的目标多项式的多项式系数，即可唯一确定出展开后的目标多项式，多项式系数密文也可唯一确定出另外一个多项式。

举例说明，第一节点具体可通过哈希函数对本地的第一样本标识进行运算，映射得到第一样本标识的哈希值，将这些哈希值作为跟生成一个一元 n 次多项式（其中 n 为第一样本标识的数量），该多项式由大整数运算生成，不会损失精度。第一节点再对该多项式展开得到多项式系数，并加密后发生给第二节点。

上述实施例中，利用不经意传输的思想，将用户标识转换得到的中间值作为多项式的解，构建目标多项式，然后将目标多项式的系数同态加密后传递给第二节点，第二节点在密文空间中基于不经意多项式求值后将求值结果反馈至第一节点，这样第一节点在不传递系数明文不泄露本地隐私数据的情况下，对第二节点的用户标识与本地的用户标识进行比较，以确定本地与第二节点共有的用户标识的碎片信息，进而可以对多方隐私数据求交。

在一个实施例中，对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点，包括：获取本地基于同态加密方式生成的密钥对；密钥对包括公钥和私钥；通过公钥对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点。步骤 208，包括：当通过私钥对参考值进行同态解密后的结果与目标值一致时，保留与参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，及聚合用户标识交集对应的用户数据。

具体地，节点群中的任一第一节点可以基于同态加密方式生成本地的密钥对，该密钥对包括公钥 PK 和私钥 SK 。其中，基于同态加密方式具体可以是指数型 $ElGamal$ 密码系统，指数型 $ElGamal$ 密码系统具有同态运算的功能，可在不获取明文的情况下对密文进行某种运算，解密后等价于对明文进行加法或数乘运算。

进一步地，第一节点在构建映射关系后，通过公钥对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点。第一节点后续在接收到第二节点反馈的参考值后，通过私钥对参考值进行同态解密，再将解密后的结果与目标值进行比较。

在本实施例中，每个第一节点各自生成自己的公私钥对，且加密过程具有随机性，第二节点无法通过密文空间下的参考值反推对应的明文是否为目标值，这样保证了第一节点的用户标识不会被泄露给第二节点，保护非交集数据隐私。

在一个实施例中，步骤 208，包括：对参考值进行同态解密得到解密结果；当解密结果与目标值一致时，则保留与参考值对应反馈的碎片信息标记为碎片信息，并将保留的碎片信息与表示验证通过的解密结果对应传递至聚合节点；传递的碎片信息用于在聚合节点中聚合得到节点群中各节点之间的用户标识交集，以聚合用户标识交集对应的用户数据；聚合节点为节点群中的其中一个第一节点。

在一个实施例中，聚合用户标识交集对应的用户数据可用作机器学习模型的训练样本。

可以理解，本申请的目的在于求取节点群中各节点之间的用户标识交集，在求取用户标识交集的过程中，节点群中的节点分为两类，一类为第一节点，第一节点的数量为多个；另一类为第二节点，第二节点的数量唯一。每个第一节点均会根据本地的第一用户标识生成一个映射关系，该映射关系可以将第一用户标识映射为目标值，而且第一节点会将映射关系的映射参数同态加密后传递给第二节点。第二节点会对每个第一节点传递过来的映射参数密文，确定本地的第二用户标识在该映射参数密文相应映射关系下映射至的参考值，并且会基于第

二用户标识拆分为多个碎片信息，碎片信息的数量与第一节点的数量相同。

这样，第二节点会在反馈参考值的时候同步反馈一个碎片信息，以将第二用户标识的碎片信息分散给每个第一节点，每个第一节点则根据参考值来判断第二节点上是否有与本地共有的用户标识，保留共有用户标识的碎片信息，但由于第一节点仅获取了碎片信息，故第一节点不知道完整的用户标识。

那么，此时需要一个节点来汇聚被保留下来的碎片信息。具体地，可以将节点群中的其中一个第一节点作为聚合节点，用于聚合被保留下来的碎片信息得到节点群中各节点之间的用户标识交集，以聚合用户标识交集对应的用户数据应用于数据处理，比如用作机器学习模型的训练样本等。

如果将第二节点作为聚合节点，那么第二节点可得到其与任何一个第一节点之间两两的用户标识交集，第一节点却无法得到其第二节点之间的用户标识交集，此时对于第一节点和第二节点是不公平的。在将第一节点作为聚合节点时，第一节点只有聚齐碎片信息后，才能得到用户标识交集，且得到的用户标识交集是各节点之间的用户标识交集。因此，聚合节点只能从第一节点中选取。

那么，对于未选为聚合节点的第一节点，第一节点可将其保留的碎片信息发送给聚合节点。再由聚合节点进行聚合操作。

举例说明，假设节点群中的其中一个第一节点 F1 接收到第二节点反馈的三个数据对：(C31, P32)、(C11, P11) 和 (C21, P13)。那么第一节点 F1 可对这三个数据对中的参考值进行同态解密，然后将解密结果与目标值 M 比较。C31 的同态解密为 J31，C11 的同态解密为 J11，C21 的同态解密为 J21。假设 J31=J11=M，J21 与 M 不相等，那么第一节点 F1 可将 P32 和 P11，并将保留的 P32 和 P11 发送至聚合节点。

上述实施例中，将隐私数据的加密和解密工作分摊在各节点，并通过其中一个第一节点来汇聚碎片信息，最终获取完整交集，由于碎片信息不可反推完整信息，且只有通过验证的碎片信息可以发送给聚合节点，因此各节点均不能通过自己获取的信息来盗取其他节点的隐私信息，保护非交集数据隐私。

在一个实施例中，如图 4 所示，提供了一种基于节点群的数据处理方法，以该方法应用于图 1 所示节点群中的选为聚合节点的第一节点为例进行说明，该方法包括以下步骤：

步骤 402，根据本地的第一用户标识构建映射关系；该映射关系用于将第一用户标识映射为目标值。

步骤 404，对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点；映射参数密文用于指示第二节点，确定第二节点的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息与参考值对应反馈至第一节点；每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，以将各碎片信息分别反馈至各第一节点；属于相同第二用户标识的碎片信息对应相同的碎片标识。

步骤 406，接收第二节点对应反馈的参考值和碎片信息。

步骤 408，当对参考值进行同态解密后的结果与目标值一致时，保留与参考值对应反馈的碎片信息。

步骤 410，接收节点群中其他第一节点发送的被保留的碎片信息。

步骤 412，当对应相同碎片标识的碎片信息的数量达到第一节点的数量时，聚合对应相同碎片标识的碎片信息，得到节点群中各节点之间的公共用户标识；根据公共用户标识得到节点群中各节点之间的用户标识交集，以聚合用户标识交集对应的用户数据用作机器学习模型的训练样本。

其中，属于相同第二用户标识的碎片信息对应相同的碎片标识，这样聚合节点可以根据碎片标识来聚合属于同一个第二用户标识的碎片信息，从而完成从碎片信息到完整用户标识的回溯。

具体地，聚合节点在接收到节点群中其他第一节点发送的被保留的碎片信息，然后按照碎片标识对碎片信息进行聚类，再统计每组碎片信息的数量是否达到第一节点的数量。在对应相同碎片标识的碎片信息的数量达到第一节点的数量时，判定该碎片信息所述的用户标识为各节点共有的用户标识，这样聚合对应相同碎片标识的碎片信息，就能得到节点群中各节点之间的公共用户标识。这样，后续就能根据公共用户标识得到节点群中各节点之间的用户标识交集，以及聚合用户标识交集对应的用户数据用作机器学习模型的训练样本。

举例说明，假设节点群中包括三个第一节点，那么第二节点可基于第二用户标识 **SID1** 拆分得到三个碎片信息 **P11**、**P12** 和 **P13**，这三个碎片信息对应相同的碎片标识 **P1**；第二节点还可基于第二用户标识 **SID2** 拆分得到三个碎片信息 **P21**、**P22** 和 **P23**，这三个碎片信息对应相同的碎片标识 **P2**；第二节点还可基于第二用户标识 **SID3** 拆分得到三个碎片信息 **P31**、**P32** 和 **P33**，这三个碎片信息对应相同的碎片标识 **P3**。各第一节点保留的碎片信息聚合到聚合节点时包括 **P11**、**P12**、**P13**、**P31**、**P32** 和 **P23**。那么，按照碎片标识对碎片信息进行聚类分为三组：**(P11、P12、P13)**、**(P31、P32)** 和 **(P23)**。该组数据 **(P11、P12、P13)** 包括 3 个碎片信息，表示 **P11**、**P12** 和 **P13** 所源自的 **SID1** 是节点群中各节点所共有的用户标识；该组数据 **(P31、P32)** 包括 2 个碎片信息，没有聚齐，表示存在一个第一节点上没有 **P31** 和 **P32** 所源自的用户标识；该组数据 **(P23)** 包括 1 个碎片信息，没有聚齐，表示存在两个第一节点上没有 **P23** 所源自的用户标识。

上述实施例中，将隐私数据的加密和解密工作分摊在各节点，并通过其中一个第一节点来汇聚碎片信息，最终获取完整交集，由于碎片信息不可反推完整信息，且只有通过验证的碎片信息可以发送给聚合节点，因此各节点均不能通过自己获取的信息来盗取其他节点的隐私信息，保护非交集数据隐私。

在一个实施例中，节点群中不同的节点中用户数据所属的维度各不相同。该基于节点群的数据处理方法还包括：获取节点群中各节点之间的用户标识交集；在本地查询用户标识交集中包括的公共用户标识所对应的用户数据；发送查询到的用户数据，发送的用户数据用于与节点群中其他节点发送的用户数据聚合后用作机器学习模型的训练样本。

其中，节点群中不同的节点中用户数据所属的维度各不相同。比如，某市一家银行拥有居民的财务信息，同时该市一家医院拥有居民的就医记录等。联邦机器学习(**Federated Learning**)，是指参与多方在数据不出库且保护数据隐私的前提下，合作完成机器学习模型训练任务。在纵向联邦学习中，参与多方各自拥有同一类型样本的不同维度的特征，对于各方共同拥有的样本，可以作为将其作为训练数据。纵向联邦学习的目标是汇总这些信息训练机器学习模型，同时保证隐私数据不出库。为了实现这一目标，需要进行样本对齐，即各参与方提供属于样本标识交集部分的用户数据用于联邦训练（比如哪些居民既注册了银行账户又注册了医院病历），该需求的本质是对各方样本标识进行隐私集合求交。

具体地，在基于前面的实施例得到节点群中各节点之间的用户标识交集后，可将用户标识交集通知给各用户标识。后续节点群中的节点或者节点群外的计算机设备需要节点群中各节点提供用户数据时，节点群中各节点可在本地查询用户标识交集中包括的公共用户标识所对应的用户数据，将查找到的用户数据发送给请求数据的节点或者计算机设备。该节点或者计算机设备则可将节点群中各节点发送的用户数据聚合后用作机器学习模型的训练样本。

继续参考图 3，节点群中的各节点上分别存储有用户不同特征维度的用户数据，比如第一节点 **F1**（参与方 1）存储有第 1 部分维度的用户数据，第一节点 **F2**（参与方 2）存储有第 2 部分维度的用户数据，第二节点 **S**（参与方 n）存储有第 n 部分维度的用户数据。对节点群中的各节点的用户标识经过本申请提供的基于节点群的数据处理方法进行处理后，可到的各节点之间的用户标识交集：用户标识（样本 ID）1、11 和 20。这样，后续再从节点群中的各节点获取用户数据进行联邦机器学习时，则仅获取各节点上用户标识交集对应的用户数据。

在本实施例中，得到节点群中各节点之间的用户标识交集后，可基于用户标识交集获取各节点的不同特征维度的用户数据，在用户数据对齐后，用于机器学习模型的训练，并保护

了各节点之间非交集数据隐私。

需要说明的是，上述实施例中聚合用户标识交集对应的用户数据用作机器学习模型的训练样本，仅是多方隐私集合求交的其中一个应用场景举例，多方隐私集合求交还可以应用于其他场景，比如，预测推广内容的实际效果或者寻找联系人等。本申请对多方隐私集合求交的应用场景不作限定。

在一个实施例中，如图 5 所示，提供了一种基于节点群的数据处理方法，以该方法应用于图 1 所示节点群中的第二节点为例进行说明，该方法包括以下步骤：

步骤 502，接收节点群中多个第一节点分别发送的映射参数密文；对各映射参数密文同态解密得到的映射参数明文相应的映射关系，用于将各映射参数密文所源自第一节点的第一用户标识映射为目标值。

需要说明的是，本实施例为在前实施例相应的由第二节点执行的技术方案，具有与前述实施例相应的技术构思和相应的技术特征。在本实施例和后续实施例中，由第一节点执行的操作可具体参考前述实施例中的具体描述，由第二节点执行的操作也可参考前述实施例中的具体描述。当然，前述实施例中由第一节点执行的操作也可参考本实施例以及后续实施例中的具体描述，前述实施例中由第二节点执行的操作也可参考本实施例以及后续实施例中的具体描述。

步骤 504，将本地的第二用户标识拆分为多个碎片信息；碎片信息的数量与节点群中第一节点的数量相同。

具体地，第二节点可直接将第二用户标识拆分为多个碎片信息，也可对第二用户标识进行处理后，将处理的结果拆分为多个碎片信息；且拆分所得的碎片信息的数量与节点群中第一节点的数量相同，这样可以将第二用户标识的碎片信息分散发布到每个第一节点，每个第一节点无法根据碎片信息得到完整的用户标识。而且，对于某个第二用户标识，当且仅当各第一节点均包括该第二用户标识时，该第二用户标识的所有碎片信息才会全部被保留，也就是说该第二用户标识是节点群中各节点共有的用户标识，进而可以也允许回溯出用户标识。

步骤 506，对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至映射参数密文所源自的第一节点；对应反馈至第一节点的参考值和碎片信息，用于指示第一节点在对参考值进行同态解密后的结果与目标值一致时，保留与参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，及聚合用户标识交集对应的用户数据。

具体地，第二节点可接收各第一节点分别发送的映射参数密文，对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值组成数据对，反馈至映射参数密文所源自的第一节点。

举例说明，假设节点群中包括两个第一节点 F1 和 F2，以及第二节点 S。F1 根据本地的第一样本标识构建出映射关系 Y1，并对 Y1 的映射参数 ym1 同态加密得到 ys1 传递至 S；F2 根据本地的第一样本标识构建出映射关系 Y2，并对 Y2 的映射参数 ym2 同态加密得到 ys2 传递至 S。S 在接收到 ys1 和 ys2 后，分别确定本地的第二用户标识在 ys1 和 ys2 相应映射关系下映射至的参考值。例如，对于第二用户标识 SID1，第二节点可以基于 SID1 拆分得到 2 个碎片信息 P11 和 P12，确定 SID1 在 ys1 相应映射关系下映射至的参考值 C11，将 C11 和 P11（或 P12）作为数据对发送给 F1；确定 SID1 在 ys2 相应映射关系下映射至的参考值 C12，将 C12 和 P12（或 P11）作为数据对发送给 F2。对于第二用户标识 SID2，第二节点可以基于 SID2 拆分得到 2 个碎片信息 P21 和 P22，确定 SID2 在 ys1 相应映射关系下映射至的参考值 C21，将 C21 和 P21（或 P22）作为数据对发送给 F1；确定 SID2 在 ys2 相应映射关系下映射至的参考值 C22，将 C22 和 P22（或 P21）作为数据对发送给 F2。需要说明的是，多个第一节点不

限制同时向第二传递映射参数密文；第二节点可以分别接收到各映射参数密文，其在每接收到一个映射参数密文时，即可进行后续的处理，无需等待接收到所有第一节点传递的映射参数密文之后再进行后续的处理。

后续，第一节点可对接收到的参考值进行同态解密，将同态解密后的结果与目标值进行比较，当两者一致时，保留与参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，及聚合用户标识交集对应的用户数据。

上述基于节点群的数据处理方法，节点群中的第一节点根据本地的第一用户标识构建映射关系，该映射关系用于将第一用户标识映射为目标值。那么在第一节点将对映射关系的映射参数进行同态加密得到映射参数密文传递至节点群中的第二节点后，第二节点可以在密文空间下基于本地的第二用户标识进行同态运算，得到第二用户标识在映射参数密文相应映射关系下映射至的参考值，然后将参考值相应的第二用户标识的其中一个碎片信息与参考值一起反馈至第一节点。第一节点再在对接收到的参考值进行同态解密后的结果与目标值一致时，保留与接收到的参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，以及聚合用户标识交集对应的用户数据。这样一方面，由于第一节点发送给第二节点的是映射参数密文，第二节点无法通过密文反推第一节点的第一用户标识，而且同态加密提供了一种对加密数据进行处理的功能，可以在不获取明文的情况下对密文进行某种运算，解密后等价于对明文进行相应的运算，以保证第一节点处的参考值与目标值的校验效力；另一方面，由于每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，这样每个第二用户标识的各碎片信息会分别反馈至各第一节点，第一节点也不会得到完整的第二节点的第二用户标识，除最后获得的节点群中各节点之间的用户标识交集外，各节点均不能通过自己获取的信息来盗取其他节点的隐私信息，从而保证了各节点隐私数据的安全，而且无需引入第三方。

在一个实施例中，接收节点群中多个第一节点分别发送的映射参数密文，包括：接收节点群中多个第一节点分别发送的多项式系数密文；各多项式系数密文所对应多项式系数明文相应的多项式，用于将各多项式系数密文所源自第一节点的第一用户标识映射为统一的目标值，目标值为零。对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至映射参数密文所源自的第一节点，包括：对于每个多项式系数密文，分别计算本地的第二用户标识在多项式系数密文相应多项式下求解得到的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至多项式系数密文所源自的第一节点。

具体地，第二节点可将本地的第二用户标识输入节点群中各节点之间共享的转换函数，获得转换得到的中间值。或者，第二节点可将本地的第二用户标识进行偏移后，输入节点群中各节点之间共享的转换函数，获得转换得到的中间值。其中，对第一用户标识进行偏移的偏移量是节点群中各节点之间共享的偏移量。此后，第二节点再对于接收到的每个多项式系数密文，分别计算本地的第二用户标识的中间值在多项式系数密文相应多项式下求解得到的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至多项式系数密文所源自的第一节点。其中，节点群中各节点之间共享的转换函数，是节点群中各节点统一使用以将用户标识转化为自然数值的函数。转换函数具体可以是哈希函数或者加密哈希函数。哈希函数也可称为散列函数，可把任意长度的输入压缩为预设长度的输出，该输出为哈希值或称为散列值。

在本实施例中，第一节点巧妙地将用户标识转换得到的中间值作为多项式的解，构建目标多项式，这样可以将目标多项式的系数同态加密后传递给第二节点，第二节点在密文空间进行计算后将结果反馈至第一节点，这样第一节点在不传递系数明文不泄露本地用户标识的

情况下,对第二节点的用户标识与本地的用户标识进行比较,以确定本地与第二节点共有的用户标识的碎片信息。

在一个实施例中,将本地的第二用户标识拆分为多个碎片信息,包括:获取节点群中各节点之间共享的转换函数;通过转换函数将本地的第二用户标识转换为中间值,以使得节点群中各节点之间相同的用户标识转换为相同的中间值;将中间值拆分为多个随机数用作碎片信息;多个随机数之和为中间值;随机数的数量与节点群中第一节点的数量相同。

具体地,第二节点可将本地的第二用户标识输入节点群中各节点之间共享的转换函数,获得转换得到的中间值。或者,第二节点可将本地的第二用户标识进行偏移后,输入节点群中各节点之间共享的转换函数,获得转换得到的中间值。此后,第二节点再按照节点群中第一节点的数量,将中间值拆分为该数量的随机数,这多个随机数之和为中间值。每个随机数看作第二用户标识的一个碎片信息,这样第二节点在向第一节点反馈参考值时,可将参考值与其中一个随机数组合为数据对,这也第二用户标识的碎片信息分散至节点群中的各第一节点。后续,在节点群中的各第一节点中均存在与第二用户标识相同的第一用户标识时,第二用户标识的碎片信息会聚合至聚合节点,将这些碎片信息相加即可还原出第二用户标识所转换得到的中间值,继而找到该中间值对应的第二用户标识,也就是节点群中各节点公共的用户标识。

举例说明,假设节点群中包括三个第一节点,第二节点可接收到这三个第一节点分别发送的映射参数密文 $ys1$ (第一节点 $F1$)、 $ys2$ (第一节点 $F2$)和 $ys3$ (第一节点 $F3$)。对于第二节点上的第二用户标识 $SID1$,一方面第二节点可将 $SID1$ 输入节点群中各节点之间共享的转换函数 $H(x)$,获得转换得到的中间值 $H(SID1)$,再将 $H(SID1)$ 拆分为三个随机数的和: $H(SID1)=P11+P12+P13$;一方面第二节点分别确定出 $SID1$ 在 $ys1$ 相应映射关系下映射至的参考值 $C11$, $SID1$ 在 $ys2$ 相应映射关系下映射至的参考值 $C12$,以及 $SID1$ 在 $ys3$ 相应映射关系下映射至的参考值 $C13$ 。此后,第二节点可将每个参考值分别与一个拆分出的随机数组成数据对,反馈至该参考值对应的映射参数密文所源自的第一节点,比如将 $C11$ 和 $P11$ 组成数据对反馈至第一节点 $F1$,将 $C12$ 和 $P12$ 组成数据对反馈至第一节点 $F2$,将 $C13$ 和 $P13$ 组成数据对反馈至第一节点 $F3$ 。

在本实施例中,第二节点仅将用户标识的碎片信息发送给其他节点,保护第一节点与第二节点非交集数据隐私,提高数据安全性。

在一个实施例中,对于每个映射参数密文,分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值,并将参考值相应的第二用户标识的其中一个碎片信息和参考值,对应反馈至映射参数密文所源自的第一节点,包括:对于每个映射参数密文,分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值,并对参考值按照符合同态加密条件的方式添加噪声数据后,和参考值相应的第二用户标识的其中一个碎片信息,对应反馈至映射参数密文所源自的第一节点。

具体地,第二节点可先对每个映射参数密文,分别确定本地的第二用户标识在该映射参数密文相应映射关系下映射至的参考值,然后对该参考值按照符合同态加密条件的方式添加噪声数据。此后,第二节点可再获取该第二用户标识的其中一个碎片信息,将该碎片信息和添加噪声数据后的参考值组成数据对,共同反馈至映射参数密文所源自的第一节点。

其中,对参考值按照符合同态加密条件的方式添加噪声数据,是指对参考值添加噪声数据前后进行同态解密的结果不发生变化。

举例说明,假设节点群中包括三个第一节点,第二节点可接收到这三个第一节点分别发送的映射参数密文 $ys1$ (第一节点 $F1$)、 $ys2$ (第一节点 $F2$)和 $ys3$ (第一节点 $F3$)。对于第二节点上的第二用户标识 $SID1$,第二节点基于 $SID1$ 拆分得到三个碎片信息 $P11$ 、 $P12$ 和 $P13$;第二节点在确定出 $SID1$ 在 $ys1$ 相应映射关系下映射至的参考值 $C11$ 后,对该参考值 $C11$ 按照符合同态加密条件的方式添加噪声数据后得到 $D11$ 。第二节点再获取基于 $SID1$ 拆分得到的一个碎片信息 $P11$,将 $D11$ 和 $P11$ 组成数据对反馈至第一节点 $F1$ 。

在一个具体的实施例中，对参考值按照符合同态加密条件的方式添加噪声数据，具体是指将参考值在同态数乘运算下乘以一个随机数。这里加入随机数的作用是掩藏非零结果，避免参考值直接发送给第一节点后，第一节点根据参考值反推得到参考值所源自的用户标识，保护非交集数据隐私。

在本实施例中，对参考值按照符合同态加密条件的方式添加噪声数据再反馈至第一节点，避免了将参考值直接发送给第一节点后，第一节点根据参考值反推得到参考值所源自的用户标识，保护第一节点与第二节点非交集数据隐私。

本申请还提供一种基于节点群的数据处理系统，其特征在于，系统所应用的节点群包括多个第一节点和一个第二节点；其中，第一节点用于根据本地的第一用户标识构建映射关系；对映射关系的映射参数进行同态加密得到映射参数密文，并将映射参数密文传递至节点群中的第二节点；映射参数用于将第一用户标识映射为目标值。第二节点用于接收节点群中多个第一节点分别发送的映射参数密文；将本地的第二用户标识拆分为多个碎片信息；对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至映射参数密文所源自的第一节点。第一节点还用于接收第二节点对应反馈的参考值和碎片信息；当对参考值进行同态解密后的结果与目标值一致时，保留与参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，及聚合用户标识交集对应的用户数据。

在一个实施例中，第一节点还用于将本地的第一用户标识映射为中间值；以中间值为多项式的解，构建目标多项式；对目标多项式的多项式系数进行同态加密得到多项式系数密文，并将多项式系数密文传递至节点群中的第二节点；目标多项式用于将第一用户标识映射为统一的目标值，目标值为零。第二节点还用于接收节点群中多个第一节点分别发送的多项式系数密文；对于每个多项式系数密文，分别计算本地的第二用户标识在多项式系数密文相应多项式下求解得到的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至多项式系数密文所源自的第一节点；各多项式系数密文所对应多项式系数明文相应的多项式，用于将各多项式系数密文所源自第一节点的第一用户标识映射为统一的目标值，目标值为零。

在另外的实施例中，第一节点还可用于执行前述实施例中由第一节点执行的步骤，第二节点还可用于执行前述实施例中由第二节点执行的步骤。

本申请还提供一种应用场景，该应用场景应用上述的基于节点群的数据处理方法，以及基于节点群的数据处理系统。具体地，在该应用场景的应用如下：

本场景中包括三种角色：①多个主参与方（第一节点）：将本地隐私数据转化为多项式；②一个客参与方（第二节点）：基于本地数据求解各主参与方的多项式，将求解结果交给主参与方；③一个聚合参与方：从主参与方中选出一个作为聚合参与方用于聚合得到各参与方之间的数据交集。参与多方隐私求交的各参与方共同遵循基于不经意多项式求值的多方隐私集合求交协议，在该协议下，所有参与方可得多方隐私数据交集，同时保证非交集的内容不泄露。

具体地，参考图 6，该图示出了一个实施例中基于节点群的数据处理系统中的数据流向图。1) 每一个主参与方（聚合参与方可以看作一种特殊的主参与方）生成自己的指数型 ElGamal 密码系统公私钥对 PK 和 SK。其中，指数型 ElGamal 密码系统具有同态运算的功能，可以在不获取明文的情况下对密文进行某种运算，解密后等价于对明文进行加法或数乘运算。主参与方对本地样本标识 SID 进行加密哈希函数映射获得 SID 的哈希值，然后将哈希值作为根生成一个一元 n 次多项式(其中 n 为本地样本 SID 的数量)，由于该多项式由大整数运算生成，不会损失精度。主参与方再对多项式展开，将展开式的多项式系数使用 PK 进行加密，再将加密后的多项式系数（多项式系数密文）发送给客参与方。

比如，对于主参与方 H，SID 的哈希值： $H\_HID = \{H(id) | id \in H\_SID\}$ ；其中，H_HID

为主参与方 H 上的 SID 映射得到的哈希值集合, $H()$ 为哈希函数, H_SID 为主参与方 H 的 SID 集合。

主参与方 H 构建的一元高次多项式为: $Poly = \prod (x - hid), hid \in H\_SID$ 。

将上述多项式的展开式的多项式系数使用 PK 进行加密, 具体可以是将多项式系数组成列表并进行 ElGamal 得到多项式系数密文:

$$H\_Poly\_Co = [Encr(c_0), Encr(c_1), \dots, Encr(c_i), \dots, Encr(c_n)]$$

其中, c_i 为展开式的第 i 次项的系数, $Encr()$ 为加密函数。

2) 客参与方收取每一个主参与方的多项式系数密文, 对自己的每一个样本标识 SID 作如下操作:

a) 将 SID 通过加密哈希函数映射得到哈希值;

比如, 对于客参与方 G, SID 的哈希值为: $G\_HID = \{H(id) | id \in G\_SID\}$; 其中, G_HID 为客参与方 G 上的 SID 映射得到的哈希值集合, $H()$ 为哈希函数, G_SID 为客参与方的 SID 集合。

b) 将哈希值随机拆分成 p 个随机数之和 (其中 p 为主参与方的数量);

$$R = \{[\sigma_i_H, \dots, \sigma_i_A] | \sigma_i_H + \dots + \sigma_i_A = hid_i, hid_i \in G\_HID\}$$

其中, $[\sigma_i_H, \dots, \sigma_i_A]$ 为 p 个随机数列表, hid_i 为客参与方 SID 的哈希值。

c) 对每一个多项式系数密文确定一个新的多项式, 并执行下面的操作: 通过同态加密将哈希值代入新的多项式求值, 并将求值结果在同态数乘运算下乘以一个随机数 (非拆分哈希值得到的随机数)。这里加入随机数的效果是掩藏非零结果, 保护非交集数据隐私。由于 ElGamal 加密过程具有随机性 (同样的明文几次加密得到的结果不同), 因此客参与方不能通过对比密文的求值结果来推测明文的求值结果是否为 0。对于该多项式系数密文, 将求值结果和 b) 中拆分得到的随机数之一组成数据对, 并发送给这个多项式系数密文源自的主参与方。

比如, 主参与方 H 发送的多项式系数密文为 H_Poly_Co , 从 H_Poly_Co 恢复出的多项式 (同态加密密文空间) 为 $Encr\_H\_Poly()$ 。客参与方将本地的哈希值代入 $Encr\_H\_Poly()$ 求值, 得到求值结果 $Encr\_H\_Poly(hid_i)$, 将求值结果在同态数乘运算下乘以一个随机数 $rh_i * Encr\_H\_Poly(hid_i)$, rh_i 为随机数。客参与方再将 $rh_i * Encr\_H\_Poly(hid_i)$ 和拆分哈希值得到的其中一个随机数 σ_i_H 组成数据对 $V\_H = [rh_i * Encr\_H\_Poly(hid_i), \sigma_i_H]$ 反馈至主参与方 H。

再比如, 主参与方 A 发送的多项式系数密文为 A_Poly_Co , 从 A_Poly_Co 恢复出的多项式 (同态加密密文空间) 为 $Encr\_A\_Poly()$ 。客参与方将本地的哈希值代入 $Encr\_A\_Poly()$ 求值, 得到求值结果 $Encr\_A\_Poly(hid_i)$, 将求值结果在同态数乘运算下乘以一个随机数 $ra_i * Encr\_A\_Poly(hid_i)$, ra_i 为随机数。客参与方再将 $ra_i * Encr\_A\_Poly(hid_i)$ 和拆分哈希值得到的其中一个随机数 σ_i_A 组成数据对 $V\_A = [ra_i * Encr\_A\_Poly(hid_i), \sigma_i_A]$ 反馈至主参与方 A。

3) 主参与方收到客参与方反馈的数据对, 该数据对中包括客参与方 SID 在本地多项式中密文的求值结果和该客参与方 SID 的碎片信息。主参与方使用 SK 对密文的求值结果解密, 如果解密结果为 0, 则保留碎片信息, 否则将碎片信息置为 0。主参与方再将验证结果明文和碎片信息发送给聚合参与方。

比如, 主参与方 H 接收到数据对 $V\_H = [rh_i * Encr\_H\_Poly(hid_i), \sigma_i_H]$ 后, 使用 SK 对数据对中的 $rh_i * Encr\_H\_Poly(hid_i)$ 进行解密, 得到解密后的数据对 $V\_H\_Decr = [rh_i *$

$\text{Poly}(\text{hid}_i), \sigma_i_H]$ 。如果 $\text{rh}_i * \text{Poly}(\text{hid}_i)$ 为 0, 则保留碎片信息 σ_i_H , 否则将碎片信息 σ_i_H 置为 0, 得到 $V\_H\_Res = [\text{rh}_i * \text{Poly}(\text{hid}_i), \sigma_i_H]$ 传递给聚合节点, 其中如果 $\text{rh}_i * \text{Poly}(\text{hid}_i)$ 为 0 则 $\sigma_i_H' = \sigma_i_H$, 否则 $\sigma_i_H' = 0$ 。

4) 聚合参与方收到所有客参与方 SID 在所有主参与方的验证结果和碎片信息, 如果在所有主参与方处的验证结果均为 0, 则将碎片信息相加得到交集 SID 的哈希值。将交集 SID 的哈希值发送给客参与方和其他主参与方。

假设主参与方 A 被选为聚合参与方, 那么对于聚合参与方 A,

SID 的哈希值: $A\_HID = \{H(\text{id}) | \text{id} \in A\_SID\}$, 其中, A_HID 为聚合参与方 A 上的 SID 映射得到的哈希值集合, $H()$ 为哈希函数, A_SID 为聚合参与方 A 的 SID 集合。

聚合参与方 A 构建的一元高次多项式为: $\text{Poly} = \prod (x - \text{hid}), \text{hid} \in A\_HID$ 。

将上述多项式的展开式的多项式系数使用 PK 进行加密, 具体可以是多项式系数组成列表并进行 ElGamal 得到多项式系数密文:

$$A\_Poly\_Co = [\text{Encr}(c_0), \text{Encr}(c_1), \dots, \text{Encr}(c_i), \dots, \text{Encr}(c_n)]$$

其中, c_i 为展开式的第 i 次项的系数, $\text{Encr}()$ 为加密函数。

聚合参与方 A 接收到数据对 $V\_A = [\text{ra}_i * \text{Encr}_A\_Poly(\text{hid}_i), \sigma_i_A]$ 后, 使用 SK 对数据对中的 $\text{ra}_i * \text{Encr}_A\_Poly(\text{hid}_i)$ 进行解密, 得到解密后的数据对 $V\_A\_Decr = [\text{ra}_i * \text{Poly}(\text{hid}_i), \sigma_i_A]$ 。如果 $\text{ra}_i * \text{Poly}(\text{hid}_i)$ 为 0, 则保留碎片信息 σ_i_A , 否则将碎片信息 σ_i_A 置为 0。

聚合参与方还会接收其他主参与方发送的验证结果和碎片信息, 比如, 主参与方 H 发送的 $V_{H\_Res} = [\text{rh}_i * \text{Poly}(\text{hid}_i), \sigma_i_H]$ 等。然后结合本地的信息, 得到交集 SID 的哈希值集合

$H\_I = \{H(\text{id}) | \text{id} \in \text{交集}\}$ 发送给客参与方和其他主参与方。

5) 各参与方得到了交集 SID 的哈希值, 通过对比本地 SID 和哈希值的映射表, 哈希值的映射来源 SID, 得到最终各参与方的交集 SID。

比如, 主参与方 H 通过 H_I 以及 H_HID 与 H_SID 之间的对应关系得到 SID 交集, 客参与方 G 通过 H_I 以及 G_HID 与 G_SID 之间的对应关系得到 SID 交集, 聚合参与方 A 通过 H_I 以及 A_HID 与 A_SID 之间的对应关系得到 SID 交集。

这样, 在本实施例中结合加密哈希函数、噪声混淆、同态密码和不经意多项式求值, 实现对多方隐私数据的多重信息保护。将交集隐私数据的加密和解密工作分摊在参与各方: 主参与方将隐私数据转化为一元高次多项式的根进行隐藏并提供加密哈希函数和同态加密的保护。客参与方在密文空间中, 利用同态密码做多项式求值, 然后利用人工噪声掩藏非零结果, 同时对隐私数据进行随机拆分, 将信息碎片和多项式求值结果分别发送给对应的主参与方。主参与方解密多项式求值结果进行根验证, 并将通过验证的信息碎片发送到其中一个主参与方 (也称聚合参与方), 由聚合参与方整合在主参与方通过验证的交集碎片信息, 最终获取完整交集。由于碎片信息不可反推完整信息, 且只有通过验证的碎片信息可以发送给聚合参与方, 因此各参与方均不能通过自己获取的信息来盗取其他参与方隐私信息。只有经过各方合作才能得到交集明文数据, 保证了在不引入可信第三方的基础上实现隐私集合求交, 且不会泄露任意两方之间的交集。而且通过本申请提供的实施例可以将纵向联邦学习样本对齐算法的参与方从两方扩展到多方, 扩大了纵向联邦学习的应用场景、提高了可用性。

图 7 示出了一个实施例中基于节点群的数据处理中数据变化的示意图。参考该图, 图中左侧为目前已有的数据处理技术方案, 可以看到求取三方交集的过程中是先分别求取主参与方和客参与方的两方交集, 再根据两个两方交集求取三方交集, 这样会泄露两方交集。图中右侧为本申请所提出的数据处理技术方案, 可以看到主参与方以本地 SID 的哈希值为根构建

多项式，客参与方采用本地 SID 的哈希值在密文空间进行多项式求值得到求值密文，主参与方再对求值密文解密后验证得到验证结果，再基于验证结果直接得到三方交集，此时只有三方交集才会以明文呈现，提高了各方隐私数据的安全性。

其中，一个主参与方构建的多项式如：

$$(x - H(1))(x - H(3))(x - H(6))(x - H(7))(x - H(9))(x - H(12))(x - H(13))(x - H(15))(x - H(18))(x - H(20))$$

另一个主参与方（聚合参与方）构建的多项式如：

$$(x - H(2))(x - H(4))(x - H(6))(x - H(8))(x - H(10))(x - H(12))(x - H(13))(x - H(16))(x - H(18))(x - H(24))$$

基于该多项式对客参与方的 SID 进行验证的验证结果为：

3	$\sigma_1 \rightarrow \text{Encr}(0)$	$\sigma_2 \rightarrow \text{Encr}(r)$
6	$\sigma_1 \rightarrow \text{Encr}(0)$	$\sigma_2 \rightarrow \text{Encr}(0)$
9	$\sigma_1 \rightarrow \text{Encr}(0)$	$\sigma_2 \rightarrow \text{Encr}(r)$
12	$\sigma_1 \rightarrow \text{Encr}(0)$	$\sigma_2 \rightarrow \text{Encr}(0)$
15	$\sigma_1 \rightarrow \text{Encr}(0)$	$\sigma_2 \rightarrow \text{Encr}(r)$
18	$\sigma_1 \rightarrow \text{Encr}(0)$	$\sigma_2 \rightarrow \text{Encr}(0)$
21	$\sigma_1 \rightarrow \text{Encr}(r)$	$\sigma_2 \rightarrow \text{Encr}(r)$
24	$\sigma_1 \rightarrow \text{Encr}(r)$	$\sigma_2 \rightarrow \text{Encr}(0)$
27	$\sigma_1 \rightarrow \text{Encr}(r)$	$\sigma_2 \rightarrow \text{Encr}(r)$
30	$\sigma_1 \rightarrow \text{Encr}(r)$	$\sigma_2 \rightarrow \text{Encr}(r)$

其中， $a \rightarrow b$ 表示键值对，比如 $\sigma_1 \rightarrow \text{Encr}(0)$ 表示验证结果 σ_1 的密文为 $\text{Encr}(0)$ ，明文为 0，代表 $\text{SID}=3$ 是主参与方和客参与方共有的 SID。再比如 $\sigma_2 \rightarrow \text{Encr}(r)$ 表示验证结果 σ_2 的密文为 $\text{Encr}(r)$ ，明文为 r ，代表 $\text{SID}=3$ 不是聚合参与方和客参与方共有的 SID。最终得到的三方交集为：6、12 和 18。

应该理解的是，虽然上述实施例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，上述实施例的流程图中的至少一部分步骤可以包括多个步骤或者多个阶段，这些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，这些步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

在一个实施例中，如图 8 所示，提供了一种基于节点群的数据处理装置，可应用于节点群中的其中一个第一节点。该装置可以采用软件模块或硬件模块，或者是二者的结合成为计算机设备的一部分，该装置具体包括：构建模块 801、传递模块 802、接收模块 803 和验证模

块 804, 其中:

构建模块 801, 用于根据本地的第一用户标识构建映射关系; 映射关系用于将第一用户标识映射为目标值;

传递模块 802, 用于对映射关系的映射参数进行同态加密得到映射参数密文, 并将映射参数密文传递至节点群中的第二节点; 映射参数密文用于指示第二节点, 确定第二节点的第二用户标识在映射参数密文相应映射关系下映射至的参考值, 并将参考值相应的第二用户标识的其中一个碎片信息与参考值对应反馈至第一节点; 每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同, 以将各碎片信息分别反馈至各第一节点;

接收模块 803, 接收第二节点对应反馈的参考值和碎片信息;

验证模块 804, 当对参考值进行同态解密后的结果与目标值一致时, 保留与参考值对应反馈的碎片信息, 并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息, 以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集, 及聚合用户标识交集对应的用户数据。

在一个实例中, 构建模块 801 还用于将本地的第一用户标识转换为中间值; 以中间值为多项式的解, 构建目标多项式; 目标多项式将第一用户标识映射为统一的目标值, 目标值为零。传递模块 802 还用于对目标多项式的多项式系数进行同态加密得到多项式系数密文, 并将多项式系数密文传递至节点群中的第二节点。

在一个实施例中, 构建模块 801 还用于获取节点群中各节点之间共享的转换函数; 通过转换函数将第一用户标识转换为中间值, 以使得节点群中各节点之间相同的用户标识转换为相同的中间值。

在一个实施例中, 传递模块 802 还用于获取本地基于同态加密方式生成的密钥对; 密钥对包括公钥和私钥; 通过公钥对映射关系的映射参数进行同态加密得到映射参数密文, 并将映射参数密文传递至节点群中的第二节点。验证模块 804 还用于当通过私钥对参考值进行同态解密后的结果与目标值一致时, 保留与参考值对应反馈的碎片信息, 并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息, 以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集, 及聚合用户标识交集对应的用户数据。

在一个实施例中, 验证模块 804 还用于对参考值进行同态解密得到解密结果; 当解密结果与目标值一致时, 则保留与参考值对应反馈的碎片信息, 并将保留的碎片信息与表示验证通过的解密结果对应传递至聚合节点; 传递的碎片信息用于在聚合节点中聚合得到节点群中各节点之间的用户标识交集, 以聚合用户标识交集对应的用户数据用作机器学习模型的训练样本; 聚合节点为节点群中的其中一个第一节点。

在一个实施例中, 如图 9 所示, 基于节点群的数据处理装置还包括聚合模块 805。

验证模块 804 还用于当对参考值进行同态解密后的结果与目标值一致时, 保留与参考值对应反馈的碎片信息。

接收模块 803 还用于接收节点群中其他第一节点发送的被保留的碎片信息。

聚合模块 805, 用于当对应相同碎片标识的碎片信息的数量达到第一节点的数量时, 聚合对应相同碎片标识的碎片信息, 得到节点群中各节点之间的公共用户标识; 根据公共用户标识得到节点群中各节点之间的用户标识交集, 以聚合用户标识交集对应的用户数据用作机器学习模型的训练样本。

在一个实施例中, 节点群中不同的节点中用户数据所属的维度各不相同。聚合模块 805 还用于获取节点群中各节点之间的用户标识交集; 在本地查询用户标识交集中包括的公共用户标识所对应的用户数据; 发送查询到的用户数据, 发送的用户数据用于与节点群中其他节点发送的用户数据聚合后用作机器学习模型的训练样本。

在一个实施例中, 如图 10 所示, 提供了一种基于节点群的数据处理装置, 可应用于节点群中的第二节点。该装置可以采用软件模块或硬件模块, 或者是二者的结合成为计算机设备的一部分, 该装置具体包括: 接收模块 1001、拆分模块 1002 和反馈模块 1003, 其中:

接收模块 1001，用于接收节点群中多个第一节点分别发送的映射参数密文；对各映射参数密文同态解密得到的映射参数明文相应的映射关系，用于将各映射参数密文所源自第一节点的第一用户标识映射为目标值；

拆分模块 1002，用于将本地的第二用户标识拆分为多个碎片信息；碎片信息的数量与节点群中第一节点的数量相同；

反馈模块 1003，用于对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至映射参数密文所源自的第一节点；对应反馈至第一节点的参考值和碎片信息，用于指示第一节点在对参考值进行同态解密后的结果与目标值一致时，保留与参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，及聚合用户标识交集对应的用户数据。

在一个实施例中，接收模块 1001 还用于接收节点群中多个第一节点分别发送的多项式系数密文；各多项式系数密文所对应多项式系数明文相应的多项式，用于将各多项式系数密文所源自第一节点的第一用户标识映射为统一的目标值，目标值为零。反馈模块 1003 还用于对于每个多项式系数密文，分别计算本地的第二用户标识在多项式系数密文相应多项式下求解得到的参考值，并将参考值相应的第二用户标识的其中一个碎片信息和参考值，对应反馈至多项式系数密文所源自的第一节点。

在一个实施例中，拆分模块 1002 还用于获取节点群中各节点之间共享的转换函数；通过转换函数将本地的第二用户标识转换为中间值，以使得节点群中各节点之间相同的用户标识转换为相同的中间值；将中间值拆分为多个随机数用作碎片信息；多个随机数之和为中间值；随机数的数量与节点群中第一节点的数量相同。

在一个实施例中，反馈模块 1003 还用于对于每个映射参数密文，分别确定本地的第二用户标识在映射参数密文相应映射关系下映射至的参考值，并对参考值按照符合同态加密条件的方式添加噪声数据后，和参考值相应的第二用户标识的其中一个碎片信息，对应反馈至映射参数密文所源自的第一节点。

上述基于节点群的数据处理装置，节点群中的第一节点根据本地的第一用户标识构建映射关系，该映射关系用于将第一用户标识映射为目标值。那么在第一节点将对映射关系的映射参数进行同态加密得到映射参数密文传递至节点群中的第二节点后，第二节点可以在密文空间下基于本地的第二用户标识进行同态运算，得到第二用户标识在映射参数密文相应映射关系下映射至的参考值，然后将参考值相应的第二用户标识的其中一个碎片信息与参考值一起反馈至第一节点。第一节点再在对接收到的参考值进行同态解密后的结果与目标值一致时，保留与接收到的参考值对应反馈的碎片信息，并触发通过节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到节点群中各节点之间的用户标识交集，以及聚合用户标识交集对应的用户数据。这样一方面，由于第一节点发送给第二节点的是映射参数密文，第二节点无法通过密文反推第一节点的第一用户标识，而且同态加密提供了一种对加密数据进行处理的功能，可以在不获取明文的情况下对密文进行某种运算，解密后等价于对明文进行相应的运算，以保证第一节点处的参考值与目标值的校验效力；另一方面，由于每个第二用户标识的碎片信息的数量与节点群中第一节点的数量相同，这样每个第二用户标识的各碎片信息会分别反馈至各第一节点，第一节点也不会得到完整的第二节点的第二用户标识，除最后获得的节点群中各节点之间的用户标识交集外，各节点均不能通过自己获取的信息来盗取其他节点的隐私信息，从而保证了各节点隐私数据的安全，而且无需引入第三方。

关于基于节点群的数据处理装置的具体限定可以参见上文中对于基于节点群的数据处理方法的限定，在此不再赘述。上述基于节点群的数据处理装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处

理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

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 11 所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。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基于节点群的数据处理数据。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节点群的数据处理方法。

本领域技术人员可以理解，图 11 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

在一个实施例中，还提供了一种计算机设备，包括存储器和处理器，存储器中存储有计算机可读指令，该处理器执行计算机可读指令时实现上述各方法实施例中的步骤。

在一个实施例中，提供了一种计算机可读存储介质，存储有计算机可读指令，该计算机可读指令被处理器执行时实现上述各方法实施例中的步骤。

在一个实施例中，提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该计算机设备执行上述各方法实施例中的步骤。

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易失性存储器中的至少一种。非易失性存储器可包括只读存储器（Read-Only Memory, ROM）、磁带、软盘、闪存或光存储器等。易失性存储器可包括随机存取存储器（Random Access Memory, RAM）或外部高速缓冲存储器。作为说明而非局限，RAM 可以是多种形式，比如静态随机存取存储器（Static Random Access Memory, SRAM）或动态随机存取存储器（Dynamic Random Access Memory, DRAM）等。

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和改进，这些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、一种基于节点群的数据处理方法，由所述节点群中的其中一个第一节点执行，所述方法包括：

根据本地的第一用户标识构建映射关系；所述映射关系用于将所述第一用户标识映射为目标值；

对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点；所述映射参数密文用于指示所述第二节点，确定所述第二节点的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息与所述参考值对应反馈至所述第一节点；

接收所述第二节点对应反馈的参考值和碎片信息；以及

当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

2、根据权利要求1所述的方法，其特征在于，所述根据本地的第一用户标识构建映射关系，包括：

将所述本地的第一用户标识转换为中间值；

以所述中间值为多项式的解，构建目标多项式；所述目标多项式将所述第一用户标识映射为统一的目标值；

所述对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点，包括：

对所述目标多项式的多项式系数进行同态加密得到多项式系数密文，并将所述多项式系数密文传递至所述节点群中的第二节点。

3、根据权利要求2所述的方法，其特征在于，所述将所述本地的第一用户标识转换为中间值，包括：

获取所述节点群中各节点之间共享的转换函数；

通过所述转换函数将所述第一用户标识转换为所述中间值，以使得所述节点群中各节点之间相同的用户标识转换为相同的中间值。

4、根据权利要求1所述的方法，其特征在于，所述对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点，包括：

获取本地基于同态加密方式生成的密钥对；所述密钥对包括公钥和私钥；

通过所述公钥对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点；

所述当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据，包括：

当通过所述私钥对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

5、根据权利要求1所述的方法，其特征在于，所述当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据，包括：

对所述参考值进行同态解密得到解密结果；

当所述解密结果与所述目标值一致时，则保留与所述参考值对应反馈的碎片信息，并将保留的所述碎片信息与表示验证通过的解密结果对应传递至聚合节点；传递的所述碎片信息用于在所述聚合节点中聚合得到所述节点群中各节点之间的用户标识交集，以聚合所述用户标识交集对应的用户数据用作机器学习模型的训练样本；所述聚合节点为所述节点群中的其中一个第一节点。

6、根据权利要求1所述的方法，其特征在于，属于相同第二用户标识的碎片信息对应相同的碎片标识；所述当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据，包括：

当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息；

接收所述节点群中其他第一节点发送的被保留的碎片信息；

当对应相同碎片标识的碎片信息的数量达到所述第一节点的数量时，聚合对应相同碎片标识的碎片信息，得到所述节点群中各节点之间的公共用户标识；以及

根据所述公共用户标识得到所述节点群中各节点之间的用户标识交集，以聚合所述用户标识交集对应的用户数据用作机器学习模型的训练样本。

7、根据权利要求1-6中任一项所述的方法，其特征在于，所述节点群中不同的节点中用户数据所属的维度各不相同；所述方法还包括：

获取所述节点群中各节点之间的用户标识交集；

在本地查询所述用户标识交集中包括的公共用户标识所对应的用户数据；以及

发送查询到的所述用户数据，发送的所述用户数据用于与所述节点群中其他节点发送的用户数据聚合后用作机器学习模型的训练样本。

8、一种基于节点群的数据处理方法，由所述节点群中的第二节点执行，所述方法包括：

接收所述节点群中多个第一节点分别发送的映射参数密文；对各所述映射参数密文同态解密得到的映射参数明文相应的映射关系，用于将各所述映射参数密文所源自第一节点的第一用户标识映射为目标值；

将本地的第二用户标识拆分为多个碎片信息；以及

对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述映射参数密文所源自的第一节点；对应反馈至所述第一节点的参考值和碎片信息，用于指示所述第一节点在对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

9、根据权利要求8所述的方法，其特征在于，所述接收所述节点群中多个第一节点分别发送的映射参数密文，包括：

接收所述节点群中多个第一节点分别发送的多项式系数密文；各所述多项式系数密文所对应多项式系数明文相应的多项式，用于将各所述多项式系数密文所源自第一节点的第一用户标识映射为统一的目标值；

所述对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述映射参数密文所源自的第一节点，包括：

对于每个所述多项式系数密文，分别计算所述本地的第二用户标识在所述多项式系数密文相应多项式下求解得到的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述多项式系数密文所源自的第一节点。

10、根据权利要求 8 所述的方法，其特征在于，所述将本地的第二用户标识拆分为多个碎片信息，包括：

获取所述节点群中各节点之间共享的转换函数；

通过所述转换函数将所述本地的第二用户标识转换为中间值，以使得所述节点群中各节点之间相同的用户标识转换为相同的中间值；以及

将所述中间值拆分为多个随机数用作碎片信息；所述多个随机数之和为所述中间值；所述随机数的数量与所述节点群中第一节点的数量相同。

11、根据权利要求 8 所述的方法，其特征在于，所述对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述映射参数密文所源自的第一节点，包括：

对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并对所述参考值按照符合同态加密条件的方式添加噪声数据后，和所述参考值相应的第二用户标识的其中一个碎片信息，对应反馈至所述映射参数密文所源自的第一节点。

12、一种基于节点群的数据处理系统，所述系统所应用的节点群包括多个第一节点和一个第二节点；其中，

所述第一节点用于根据本地的第一用户标识构建映射关系；对所述映射关系的映射参数进行同态加密得到映射参数密文，并将所述映射参数密文传递至所述节点群中的第二节点；所述映射参数用于将所述第一用户标识映射为目标值；

所述第二节点用于接收所述节点群中多个第一节点分别发送的映射参数密文；将本地的第二用户标识拆分为多个碎片信息；对于每个映射参数密文，分别确定本地的第二用户标识在所述映射参数密文相应映射关系下映射至的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述映射参数密文所源自的第一节点；以及

所述第一节点还用于接收所述第二节点对应反馈的参考值和碎片信息；当对所述参考值进行同态解密后的结果与所述目标值一致时，保留与所述参考值对应反馈的碎片信息，并触发通过所述节点群中的其中一个第一节点收集各第一节点中保留的碎片信息，以根据收集的碎片信息聚合得到所述节点群中各节点之间的用户标识交集，及聚合所述用户标识交集对应的用户数据。

13、根据权利要求 12 所述的系统，其特征在于，所述第一节点还用于将本地的第一用户标识映射为中间值；以所述中间值为多项式的解，构建目标多项式；对所述目标多项式的多项式系数进行同态加密得到多项式系数密文，并将所述多项式系数密文传递至所述节点群中的第二节点；所述目标多项式用于将所述第一用户标识映射为统一的目标值；

所述第二节点还用于接收所述节点群中多个第一节点分别发送的多项式系数密文；对于每个多项式系数密文，分别计算本地的第二用户标识在所述多项式系数密文相应多项式下求解得到的参考值，并将所述参考值相应的第二用户标识的其中一个碎片信息和所述参考值，对应反馈至所述多项式系数密文所源自的第一节点；各所述多项式系数密文所对应多项式系数明文相应的多项式，用于将各所述多项式系数密文所源自第一节点的第一用户标识映射为统一的目标值。

14、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，其特征在于，所述处理器执行所述计算机可读指令时实现权利要求 1 至 11 中任一项所述的方法的步骤。

15、一种计算机可读存储介质，存储有计算机可读指令，其特征在于，所述计算机可读指令被处理器执行时实现权利要求 1 至 11 中任一项所述的方法的步骤。

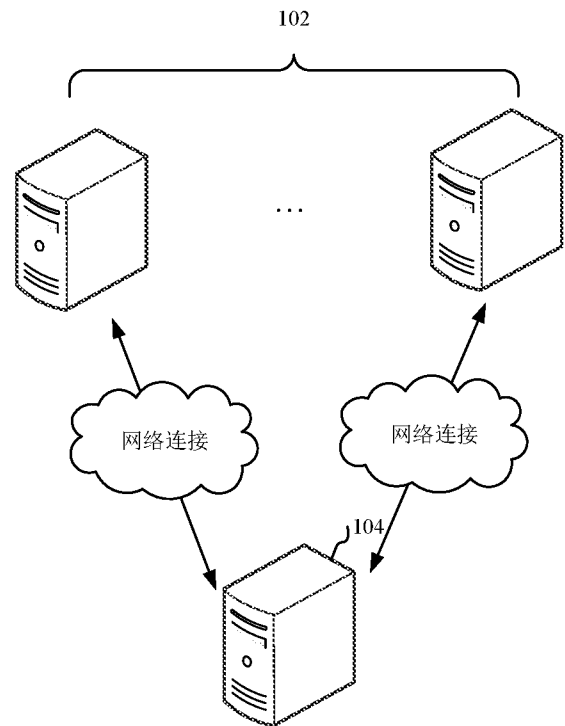


图 1

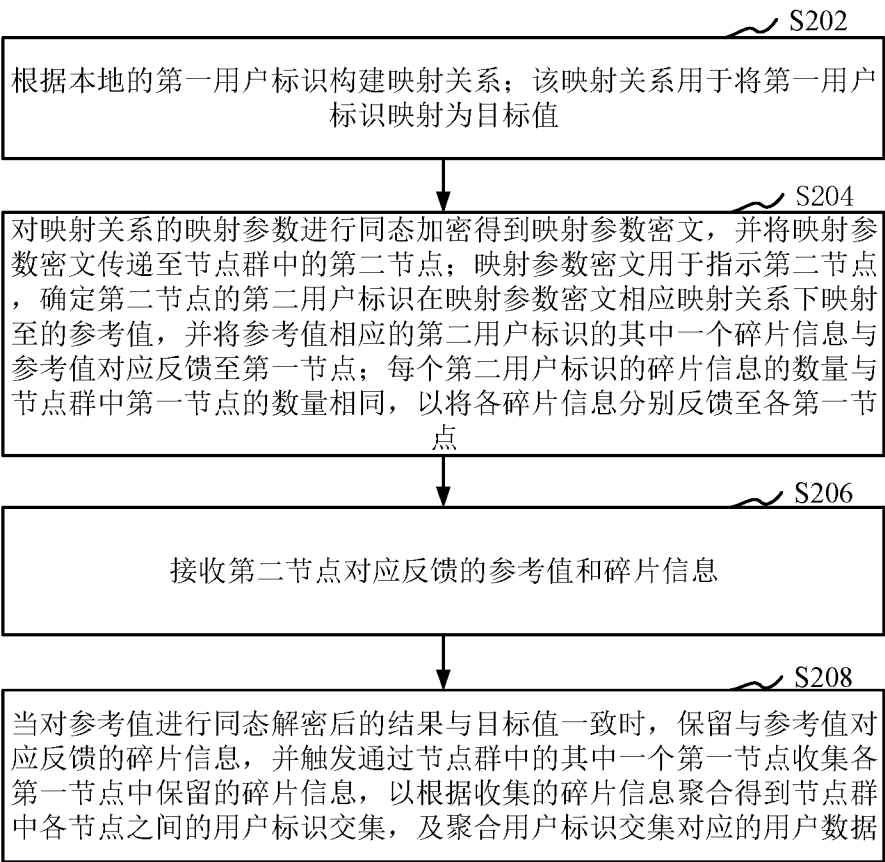


图 2

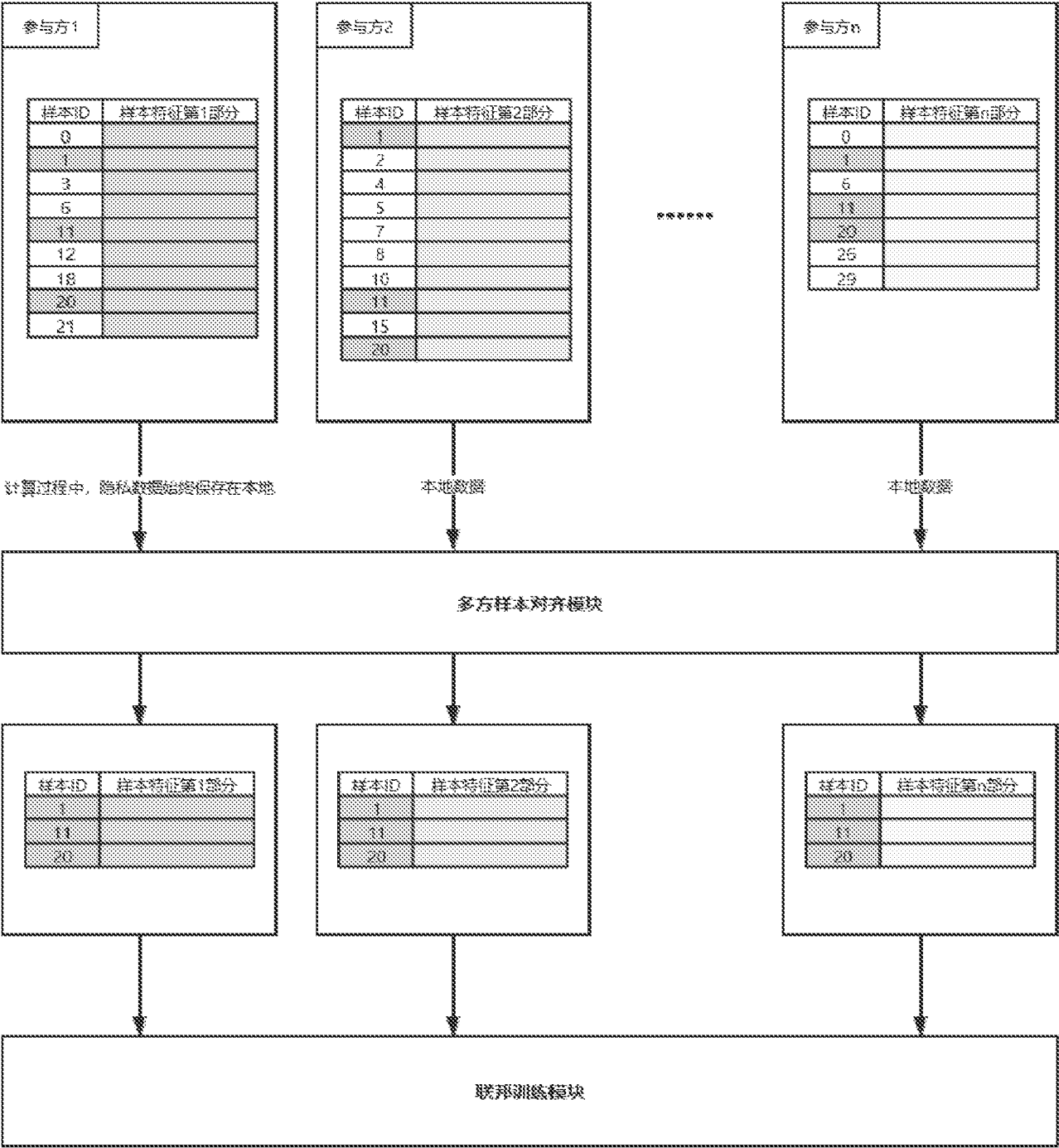


图 3

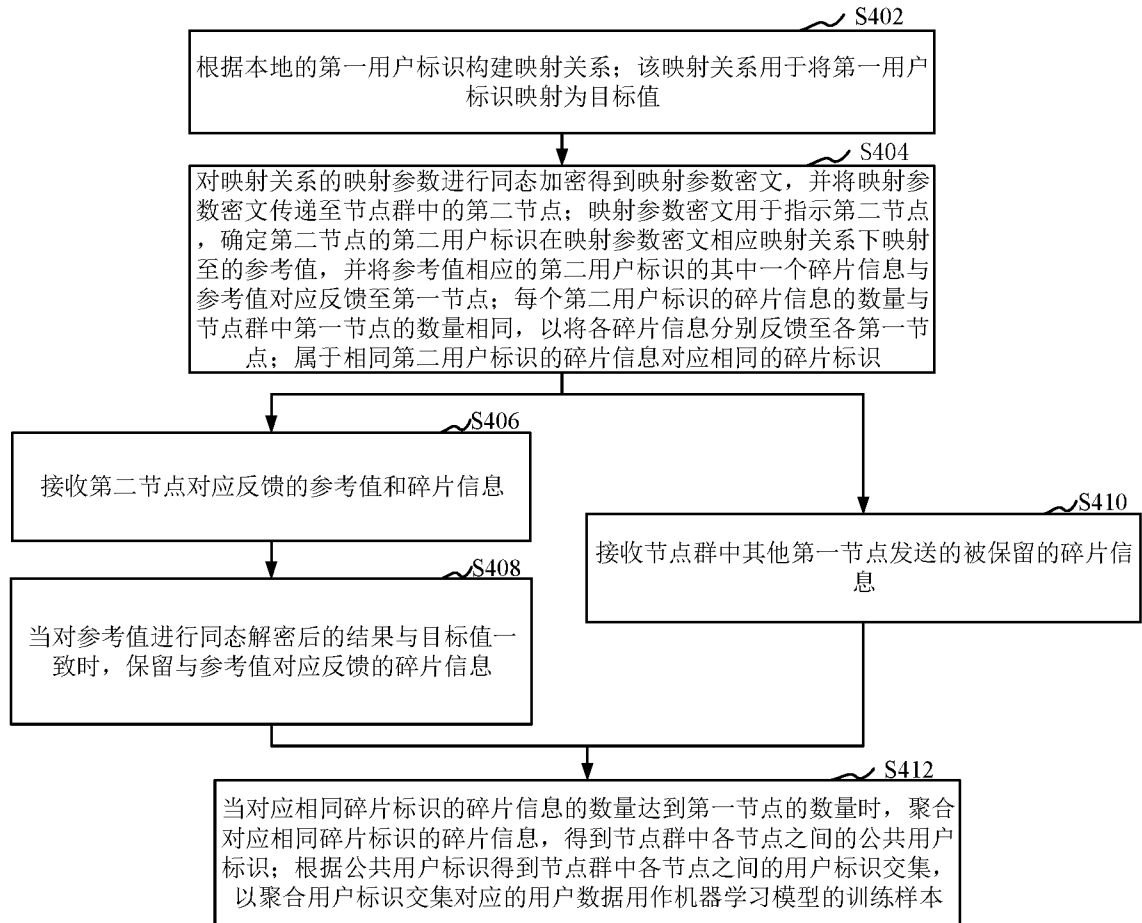


图 4

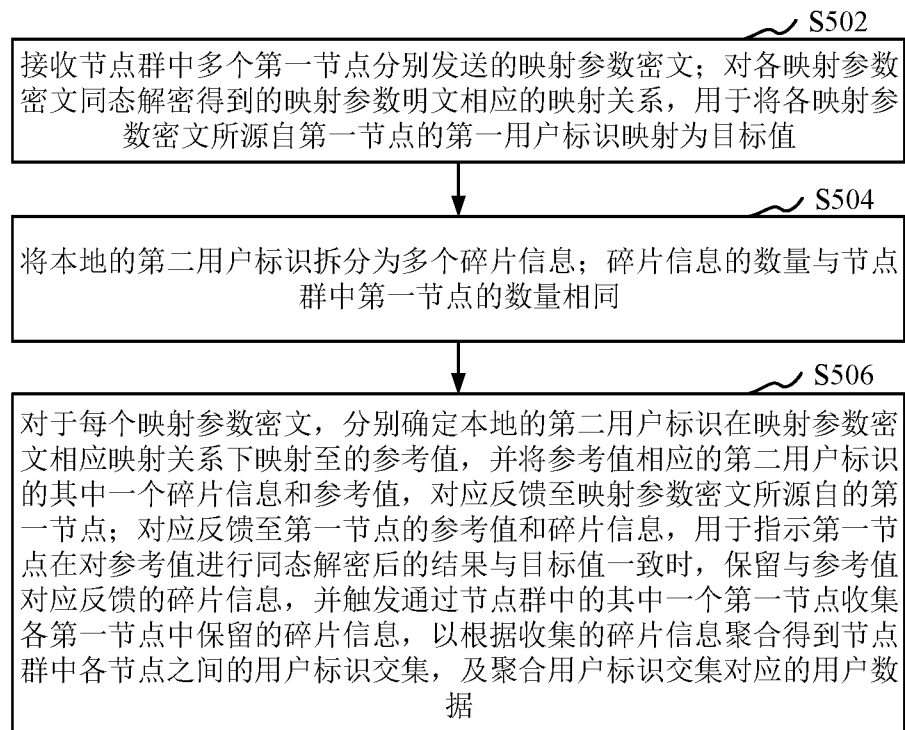


图 5

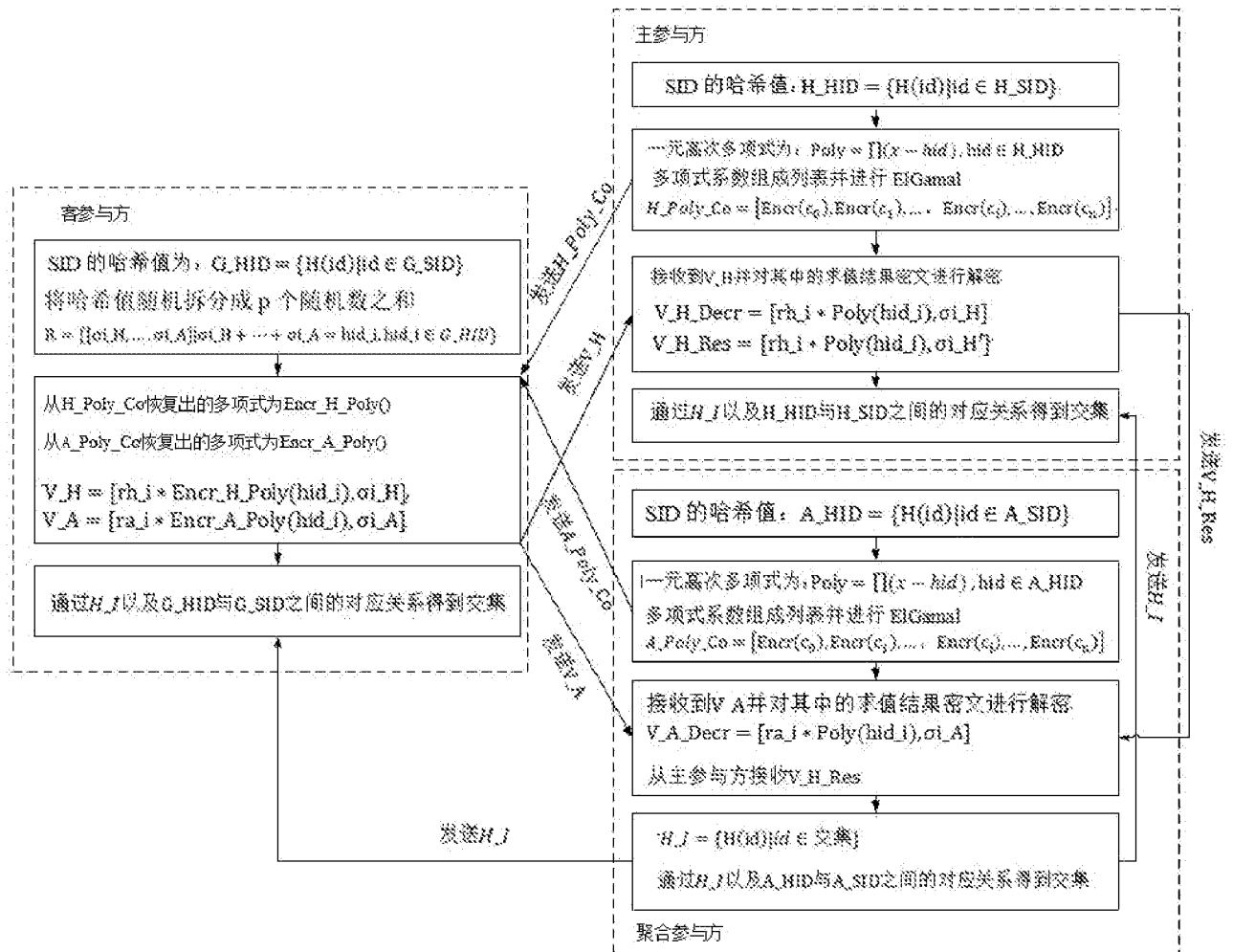


图 6

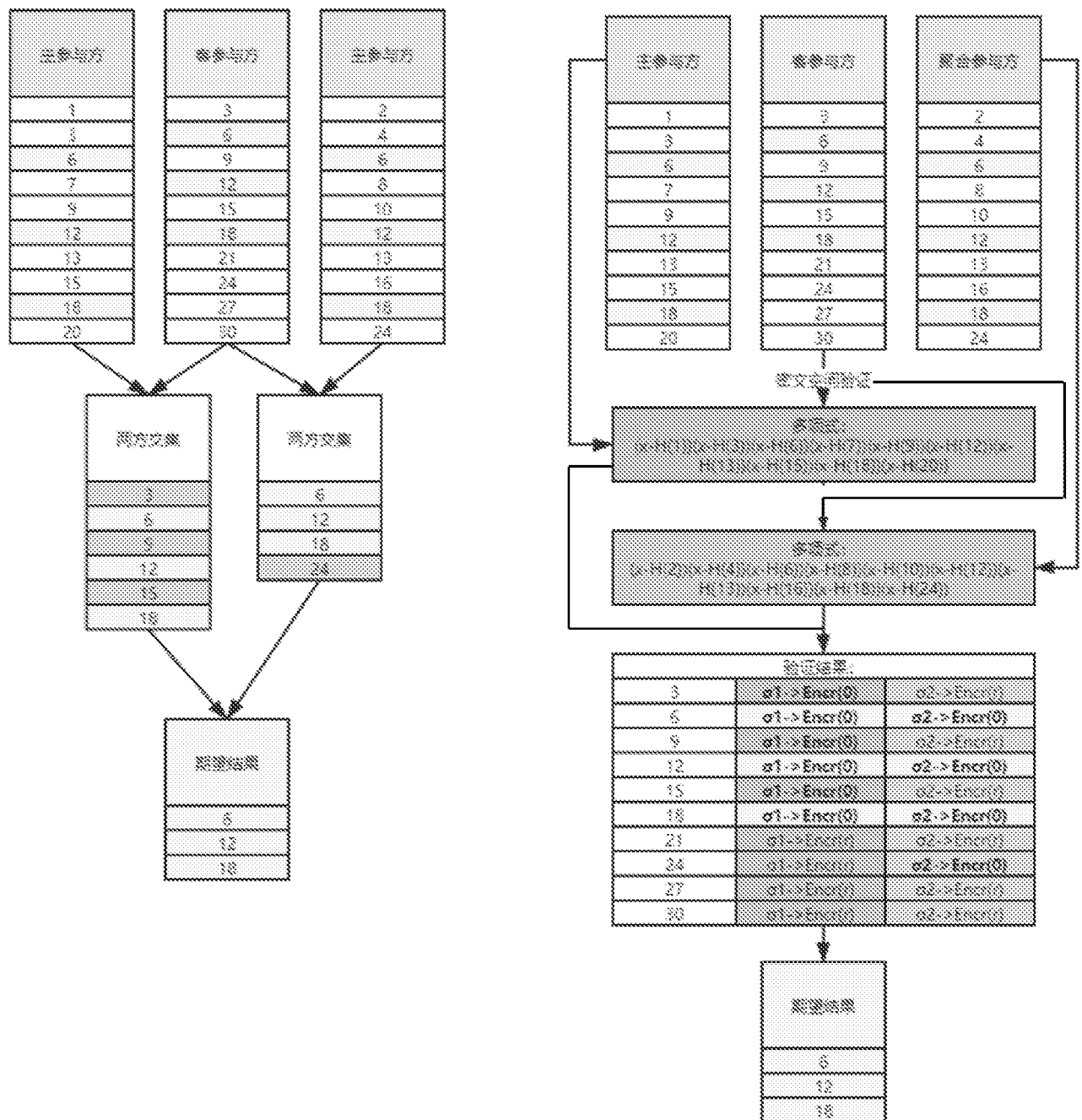


图 7

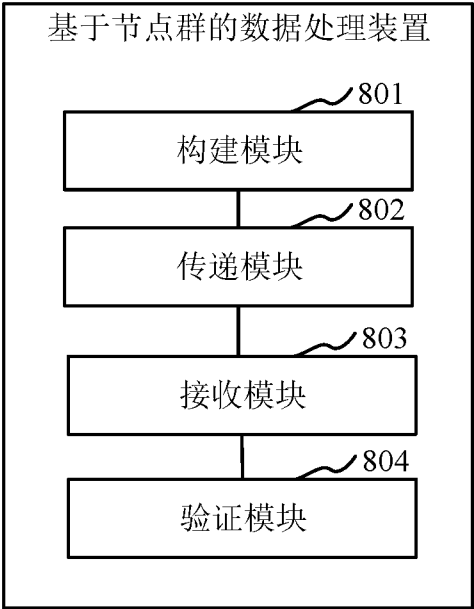


图 8

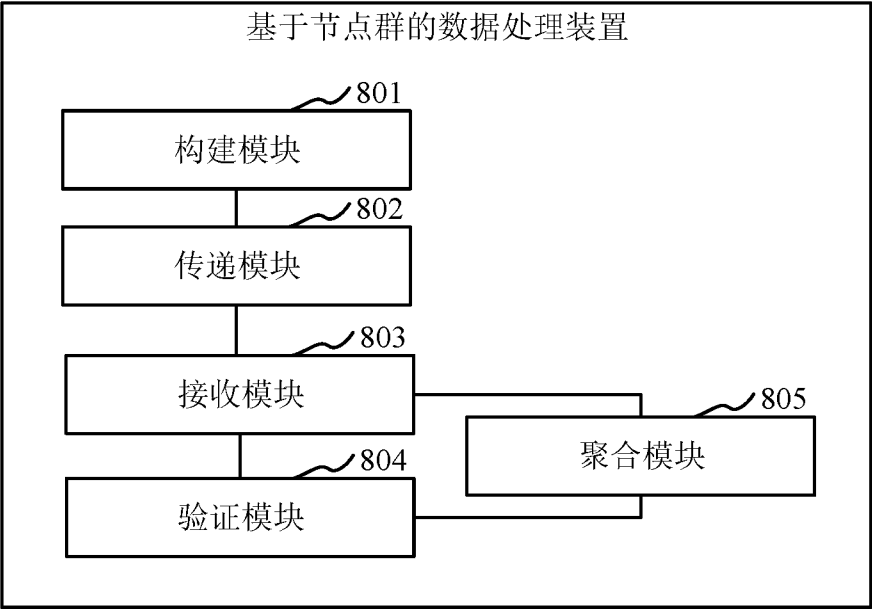


图 9

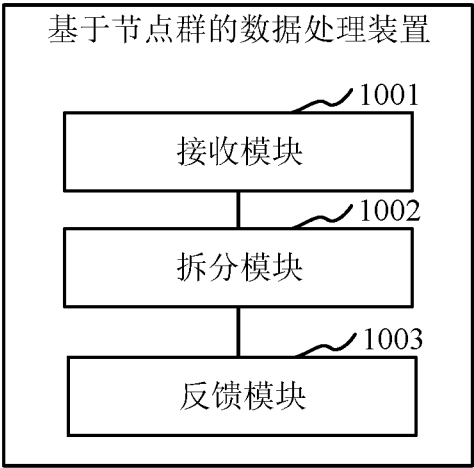


图 10

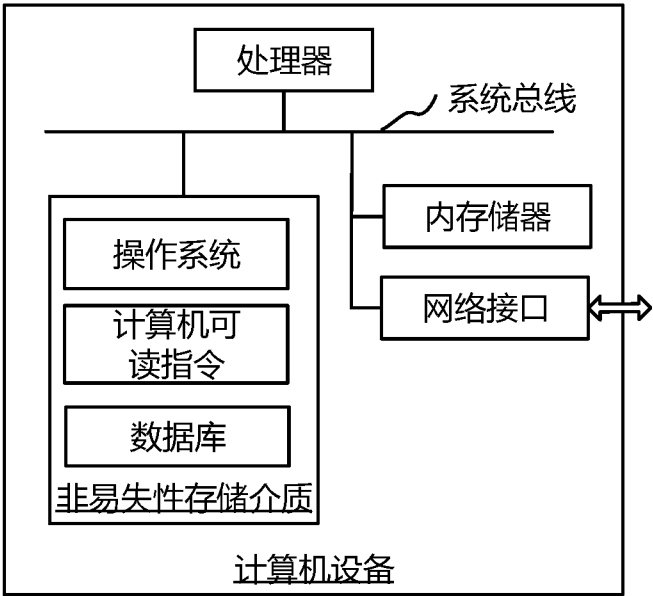


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/116129

A. CLASSIFICATION OF SUBJECT MATTER

G06F 21/71(2013.01)i; G06F 21/62(2013.01)i; G06F 21/60(2013.01)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

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; CNTXT; CNKI; SIPOABS; DWPI; USTXT; WOTXT; EPTXT: 映射, 加密, 转换, 第一, 第二, 节点, 用户, 标识, 信息, 碎片, 分割, 聚合, 整合, map, encrypt, transform, first, second, node, user, identification, information, fragment, segment, aggregate, integrate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111931253 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 13 November 2020 (2020-11-13) claims 1-15, description, paragraphs [0013]-[0155]	1-15
A	CN 108768642 A (ANHUI UNIVERSITY) 06 November 2018 (2018-11-06) description, paragraphs [0027]-[0092]	1-15
A	CN 110536263 A (ZTE CORPORATION) 03 December 2019 (2019-12-03) entire document	1-15
A	CN 111401558 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 10 July 2020 (2020-07-10) entire document	1-15
A	CN 104519071 A (UNIVERSITY OF SCIENCE AND TECHNOLOGY BEIJING) 15 April 2015 (2015-04-15) entire document	1-15



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

09 November 2021

Date of mailing of the international search report

22 November 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
Information on patent family members

International application No.

PCT/CN2021/116129

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111931253	A	13 November 2020	CN	111931253	B	15 January 2021
CN	108768642	A	06 November 2018	CN	108768642	B	01 June 2021
CN	110536263	A	03 December 2019	WO	2021004182	A1	14 January 2021
CN	111401558	A	10 July 2020	CN	111401558	B	09 October 2020
CN	104519071	A	15 April 2015	US	2018006811	A1	04 January 2018
				US	10411885	B2	10 September 2019
				WO	2016112734	A1	21 July 2016
				CN	104519071	B	11 August 2017

国际检索报告

国际申请号

PCT/CN2021/116129

A. 主题的分类

G06F 21/71(2013.01)i; G06F 21/62(2013.01)i; G06F 21/60(2013.01)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

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;CNTXT;CNKI;SIPOABS;DWPI;USTXT;WOTXT;EPTXT: 映射, 加密, 转换, 第一, 第二, 节点, 用户, 标识, 信息, 碎片, 分割, 聚合, 整合, map, encrypt, transform, first, second, node, user, identification, information, fragment, segment, aggregate, integrate

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
PX	CN 111931253 A (腾讯科技深圳有限公司) 2020年 11月 13日 (2020 - 11 - 13) 权利要求第1-15项, 说明书第[0013]-[0155]段	1-15
A	CN 108768642 A (安徽大学) 2018年 11月 6日 (2018 - 11 - 06) 说明书第[0027]-[0092]段	1-15
A	CN 110536263 A (中兴通讯股份有限公司) 2019年 12月 3日 (2019 - 12 - 03) 全文	1-15
A	CN 111401558 A (腾讯科技深圳有限公司) 2020年 7月 10日 (2020 - 07 - 10) 全文	1-15
A	CN 104519071 A (北京科技大学) 2015年 4月 15日 (2015 - 04 - 15) 全文	1-15

☐ 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

2021年 11月 9日

国际检索报告邮寄日期

2021年 11月 22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

受权官员

张秋芳

电话号码 (86-512) 88995959

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/116129

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	111931253	A	2020年 11月 13日	CN	111931253	B	2021年 1月 15日
CN	108768642	A	2018年 11月 6日	CN	108768642	B	2021年 6月 1日
CN	110536263	A	2019年 12月 3日	WO	2021004182	A1	2021年 1月 14日
CN	111401558	A	2020年 7月 10日	CN	111401558	B	2020年 10月 9日
CN	104519071	A	2015年 4月 15日	US	2018006811	A1	2018年 1月 4日
				US	10411885	B2	2019年 9月 10日
				WO	2016112734	A1	2016年 7月 21日
				CN	104519071	B	2017年 8月 11日