(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117609621 A
(43)申请公布日 2024.02.27

(21)申请号 202311625862.9

(22)申请日 2023.11.30

(71)申请人 北京富算科技有限公司
地址 100070 北京市丰台区南四环西路188号十六区18号楼1至15层101内7层701-8

(72)发明人 尤志强 赵东 杨云波 蔡晓娟 王兆凯 杜吉锋 陈立峰 孙小超 赵华宇 卫骞 杜浩 卞阳 张伟奇

(74)专利代理机构 北京慧加伦知识产权代理有限公司 16035
专利代理师 李永敏

(51)Int.Cl.
G06F 16/9535(2019.01)
G06F 18/22(2023.01)
G06F 21/60(2013.01)

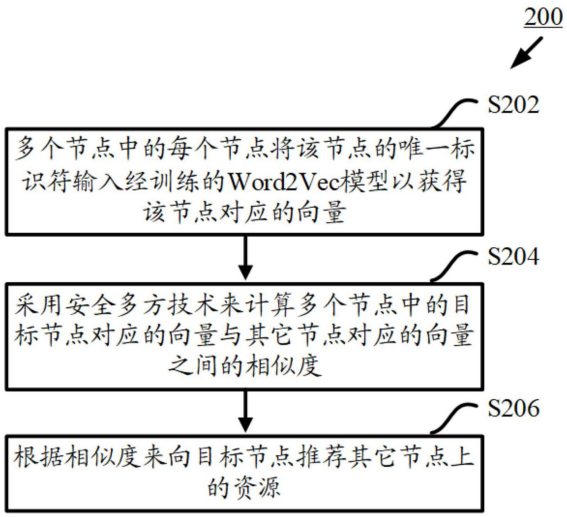
权利要求书3页 说明书13页 附图11页

(54)发明名称

用于在多个节点中进行资源推荐的方法

(57)摘要

本公开的实施例提供一种用于在多个节点中进行资源推荐的方法。该方法包括:多个节点中的每个节点将该节点的唯一标识符输入经训练的Word2Vec模型以获得该节点对应的向量,Word2Vec模型由多个节点采用安全多方技术进行联合训练并被训练成使得相似度越高的节点在向量空间中的距离越近;采用安全多方技术来计算多个节点中的目标节点对应的向量与其它节点对应的向量之间的相似度,其它节点是多个节点中除了目标节点之外的节点;以及根据相似度来向目标节点推荐其它节点上的资源。



CN 117609621 A

1.一种用于在多个节点中进行资源推荐的方法,其特征在于,所述方法包括:

所述多个节点中的每个节点将该节点的唯一标识符输入经训练的Word2Vec模型以获得该节点对应的向量,所述Word2Vec模型由所述多个节点采用安全多方技术进行联合训练并被训练成使得相似度越高的节点在向量空间中的距离越近;

采用安全多方技术来计算所述多个节点中的目标节点对应的向量与其它节点对应的向量之间的相似度,所述其它节点是所述多个节点中除了所述目标节点之外的节点;以及

根据所述相似度来向所述目标节点推荐所述其它节点上的资源。

2.根据权利要求1所述的方法,其特征在于,所述目标节点对应的向量与所述其它节点中的任一节点对应的向量之间的相似度根据下式来计算:

$$\cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

其中,cos(θ)表示所述相似度,A表示所述目标节点对应的第一向量,B表示所述任一节点对应的第二向量。

3.根据权利要求2所述的方法,其特征在于,通过以下操作来计算所述目标节点对应的向量与所述任一节点对应的向量之间的相似度:

所述目标节点与所述任一节点通过安全多方技术来计算A·B;

所述目标节点独立计算||A||;

所述任一节点独立计算||B||并向所述目标节点发送所计算的||B||;以及

所述目标节点根据A·B、||A||和||B||来计算所述相似度。

4.根据权利要求3所述的方法,其特征在于,所述目标节点与所述任一节点通过安全多方技术来计算A·B包括:

所述目标节点将所述第一向量碎片化成第一碎片矩阵和第二碎片矩阵并向所述任一节点提供所述第二碎片矩阵;

所述任一节点将所述第二向量碎片化成第三碎片矩阵和第四碎片矩阵并向所述目标节点提供所述第三碎片矩阵;

所述目标节点将所述第一碎片矩阵乘以噪声矩阵以生成加噪矩阵并向所述任一节点提供所述加噪矩阵,其中,所述噪声矩阵的大小与所述第一碎片矩阵相同且所述噪声矩阵中的每个元素为随机数;

所述目标节点将所述第一向量和所述第三碎片矩阵进行相乘以生成第一值;

所述任一节点将所述加噪矩阵乘以所述第四碎片矩阵的转置矩阵以生成第一加噪乘积矩阵,对所述第一加噪乘积矩阵进行逐项加密以生成密态第一加噪乘积矩阵,并向所述目标节点提供所述密态第一加噪乘积矩阵;

所述目标节点逐项计算所述密态第一加噪乘积矩阵与所述噪声矩阵之商以生成密态第一乘积矩阵,对所述密态第一乘积矩阵进行密态求和以生成第一密态和,将所述第一密态和减去随机数以生成密态第一中间值,并向所述目标节点提供所述密态第一中间值;

所述任一节点对所述第二碎片矩阵和所述第四碎片矩阵进行相乘以生成第二值,对所述密态第一中间值进行同态解密以获得第一中间值,将所述第二值与所述第一中间值相加以生成第二中间值,并向所述目标节点提供所述第二中间值;

所述目标节点将所述随机数、所述第一值与所述第二中间值相加以获得A•B的结果。

5.根据权利要求1至4中任一项所述的方法,其特征在于,所述方法还包括:

通过安全求交方式计算所述多个节点中的每两个节点之间的交集大小;

根据所述多个节点中的每两个节点之间的交集大小来计算该两个节点之间的交集比例;以及

根据所计算的交集比例来构建所述多个节点的虚拟数据价值网络,其中,在所述虚拟数据价值网络中,每个节点被表示为一个顶点,每两个节点之间的交集比例作为该两个节点之间的边的权重;

其中,所述Word2Vec模型基于所述虚拟数据价值网络来对所述多个节点进行联合训练。

6.根据权利要求5所述的方法,其特征在于,通过安全求交方式计算所述多个节点中的每两个节点之间的交集大小包括:

获得第一节点的第一原始数据矩阵中的唯一标识符向量;

将所述第一原始数据矩阵中的唯一标识符向量转换成第一哈希向量;

获得第二节点的第二原始数据矩阵中的唯一标识符向量;

将所述第二原始数据矩阵中的唯一标识符向量转换成第二哈希向量;

比较所述第一哈希向量中的每个第一哈希值与所述第二哈希向量中的每个第二哈希值以将所述第一哈希向量中与所述第二哈希值相等的第一哈希值的个数确定为所述第一节点与所述第二节点之间的交集大小;

其中,比较所述第一哈希值与所述第二哈希值包括:

由所述第一节点和所述第二节点联合确定所述第一哈希值是否小于所述第二哈希值;

响应于所述第一哈希值不小于所述第二哈希值,由所述第一节点和所述第二节点联合确定所述第二哈希值是否小于所述第一哈希值;

响应于所述第二哈希值不小于所述第一哈希值,确定所述第一哈希值等于所述第二哈希值;

其中,由所述第一节点和所述第二节点联合确定所述第一哈希值是否小于所述第二哈希值包括:

由所述第一节点将所述第一哈希值碎片化为第一碎片值和第二碎片值并向所述第二节点发送所述第二碎片值;

由所述第二节点将所述第二哈希值碎片化为第三碎片值和第四碎片值并向所述第一节点发送所述第三碎片值;

由所述第一节点将所述第一碎片值减去所述第三碎片值以获得第五碎片值;

由所述第二节点将所述第二碎片值减去所述第四碎片值以获得第六碎片值;

生成第一布尔零碎片、第二布尔零碎片、第一算术零碎片、第二算术零碎片,其中,所述第一布尔零碎片与所述第二布尔零碎片异或的结果为0,所述第一算术零碎片与所述第二算术零碎片相加的结果为0;

将所述第一布尔零碎片和所述第一算术零碎片分配给所述第一节点;

将所述第二布尔零碎片和所述第二算术零碎片分配给所述第二节点;

由所述第一节点计算所述第五碎片值与所述第一算术零碎片之和与所述第一布尔零

碎片异或的结果,以获得第一运算碎片;

由所述第二节点计算所述第六碎片值与所述第二算术零碎片之和,以获得第二运算碎片;

由所述第一节点和所述第二节点联合利用所述第一节点处的第一并行前缀加法器和所述第二节点处的第二并行前缀加法器在所述第一节点处获得第一符号位碎片并且在所述第二节点处获得第二符号位碎片,其中,所述第一并行前缀加法器的输入为所述第一运算碎片和第三运算碎片,所述第二并行前缀加法器的输入为所述第二运算碎片和第四运算碎片,所述第三运算碎片等于0,所述第四运算碎片等于所述第二布尔零碎片;

对所述第一符号位碎片与所述第二符号位碎片执行异或操作以获得比较值;

响应于所述比较值为真,确定所述第一哈希值小于所述第二哈希值;以及

响应于所述比较值不为真,确定所述第一哈希值不小于所述第二哈希值。

7.根据权利要求5所述的方法,其特征在于,根据所计算的交集比例来构建所述多个节点的虚拟数据价值网络包括:

根据目标评估指标来评估所述多个节点的数据质量;以及

将数据质量低于预设值的节点从所述虚拟数据价值网络中删除。

8.根据权利要求5所述的方法,其特征在于,根据所计算的交集比例来构建所述多个节点的虚拟数据价值网络包括:

获取所述多个节点中的目标节点的数据源领域偏好;

确定链接到所述目标节点的所有节点中符合所述数据源领域偏好的候选节点;以及

在所述虚拟数据价值网络中,根据增强因子来调整每个候选节点到所述目标节点之间的边的权重。

9.根据权利要求8所述的方法,其特征在于,根据增强因子来调整每个候选节点到所述目标节点之间的边的权重包括:将每个候选节点到所述目标节点之间的边的权重乘以所述增强因子;

其中,在所述数据源领域偏好为相似数据的情况下,所述增强因子为第一常数;

在所述数据源领域偏好为互补数据的情况下,所述增强因子为第二常数。

10.根据权利要求8所述的方法,其特征在于,根据增强因子来调整每个候选节点到所述目标节点之间的边的权重包括根据下式来调整所述权重:

$$w_{t+1} = e^{w_t} \times e^{p}$$

其中,w_{t+1}表示调整后的权重,w_t表示调整前的权重,p表示所述增强因子;

在数据源领域偏好为相似数据的情况下,所述增强因子为第三常数;

在数据源领域偏好为互补数据的情况下,所述增强因子为第四常数。

用于在多个节点中进行资源推荐的方法

技术领域

[0001] 本公开的实施例涉及计算机技术领域,具体地,涉及用于在多个节点中进行资源推荐的方法。

背景技术

[0002] 金融科技是近些年以来发展非常迅猛的行业之一。金融科技是指运用数字化技术和创新来支持或提供金融服务的行业与实践。金融科技主要包括数字支付、网络借贷、大数据风控、流动性管理、投资管理、保险科技、金融监管等业务场景。金融科技是重度依赖数据源的行业。但近些年,随着数据安全越来越被重视,数据已经不可能再像先前无序的交易模式,采用直接传输原始数据方式进行。数据隐私保护是金融科技行业必须要面对的重大转变。隐私保护是指在保护敏感数据的前提下开展联合建模、数据价值流通等工作。这种转变,不仅仅发生在金融科技行业,在医疗行业、教育行业、互联网行业、涉及数据价值共享的行业都面临数据安全保护的现实要求。

[0003] 各行各业中的各类政务主体、行业主体、公司主体、机构主体可组合成网络,而此类网络中的主体数量可能非常大。这些主体在带来丰富的数据资源的同时,也大幅度提升了需求方用户找到其所需数据资源的难度,也就是需求方用户会面临信息过载问题。需求方用户很难从海量数据源信息中找到自己感兴趣或需要的数据资源内容。如何进行高效且快速的资源推荐成为需要解决的问题。

发明内容

[0004] 本文中描述的实施例提供了一种用于在多个节点中进行资源推荐的方法、装置以及存储有计算机程序的计算机可读存储介质。

[0005] 根据本公开的第一方面,提供了一种用于在多个节点中进行资源推荐的方法。该方法包括:多个节点中的每个节点将该节点的唯一标识符输入经训练的Word2Vec模型以获得该节点对应的向量,Word2Vec模型由多个节点采用安全多方技术进行联合训练并被训练成使得相似度越高的节点在向量空间中的距离越近;采用安全多方技术来计算多个节点中的目标节点对应的向量与其它节点对应的向量之间的相似度,其它节点是多个节点中除了目标节点之外的节点;以及根据相似度来向目标节点推荐其它节点上的资源。

[0006] 在本公开的一些实施例中,目标节点对应的向量与其它节点中的任一节点对应的向量之间的相似度根据下式来计算:

[0007] $$\cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

[0008] 其中,cos(θ)表示相似度,A表示目标节点对应的第一向量,B表示任一节点对应的第二向量。

[0009] 在本公开的一些实施例中,通过以下操作来计算目标节点对应的向量与任一节点对应的向量之间的相似度:目标节点与任一节点通过安全多方技术来计算A·B;目标节点

独立计算||A||;任一节点独立计算||B||并向目标节点发送所计算的||B||;以及目标节点根据A·B、||A||和||B||来计算相似度。

[0010] 在本公开的一些实施例中,目标节点与任一节点通过安全多方技术来计算A·B包括:目标节点将第一向量碎片化成第一碎片矩阵和第二碎片矩阵并向任一节点提供第二碎片矩阵;任一节点将第二向量碎片化成第三碎片矩阵和第四碎片矩阵并向目标节点提供第三碎片矩阵;目标节点将第一碎片矩阵乘以噪声矩阵以生成加噪矩阵并向任一节点提供加噪矩阵,其中,噪声矩阵的大小与第一碎片矩阵相同且噪声矩阵中的每个元素为随机数;目标节点将第一向量和第三碎片矩阵进行相乘以生成第一值;任一节点将加噪矩阵乘以第四碎片矩阵的转置矩阵以生成第一加噪乘积矩阵,对第一加噪乘积矩阵进行逐项加密以生成密态第一加噪乘积矩阵,并向目标节点提供密态第一加噪乘积矩阵;目标节点逐项计算密态第一加噪乘积矩阵与噪声矩阵之商以生成密态第一乘积矩阵,对密态第一乘积矩阵进行密态求和以生成第一密态和,将第一密态和减去随机数以生成密态第一中间值,并向目标节点提供密态第一中间值;任一节点对第二碎片矩阵和第四碎片矩阵进行相乘以生成第二值,对密态第一中间值进行同态解密以获得第一中间值,将第二值与第一中间值相加以生成第二中间值,并向目标节点提供第二中间值;目标节点将随机数、第一值与第二中间值相加以获得A·B的结果。

[0011] 在本公开的一些实施例中,方法还包括:通过安全求交方式计算多个节点中的每两个节点之间的交集大小;根据多个节点中的每两个节点之间的交集大小来计算该两个节点之间的交集比例;以及根据所计算的交集比例来构建多个节点的虚拟数据价值网络,其中,在虚拟数据价值网络中,每个节点被表示为一个顶点,每两个节点之间的交集比例作为该两个节点之间的边的权重;其中,Word2Vec模型基于虚拟数据价值网络来对多个节点进行联合训练。

[0012] 在本公开的一些实施例中,通过安全求交方式计算多个节点中的每两个节点之间的交集大小包括:获得第一节点的第一原始数据矩阵中的唯一标识符向量;将第一原始数据矩阵中的唯一标识符向量转换成第一哈希向量;获得第二节点的第二原始数据矩阵中的唯一标识符向量;将第二原始数据矩阵中的唯一标识符向量转换成第二哈希向量;比较第一哈希向量中的每个第一哈希值与第二哈希向量中的每个第二哈希值以将第一哈希向量中与第二哈希值相等的第一哈希值的个数确定为第一节点与第二节点之间的交集大小;其中,比较第一哈希值与第二哈希值包括:由第一节点和第二节点联合确定第一哈希值是否小于第二哈希值;响应于第一哈希值不小于第二哈希值,由第一节点和第二节点联合确定第二哈希值是否小于第一哈希值;响应于第二哈希值不小于第一哈希值,确定第一哈希值等于第二哈希值;其中,由第一节点和第二节点联合确定第一哈希值是否小于第二哈希值包括:由第一节点将第一哈希值碎片化为第一碎片值和第二碎片值并向第二节点发送第二碎片值;由第二节点将第二哈希值碎片化为第三碎片值和第四碎片值并向第一节点发送第三碎片值;由第一节点将第一碎片值减去第三碎片值以获得第五碎片值;由第二节点将第二碎片值减去第四碎片值以获得第六碎片值;生成第一布尔零碎片、第二布尔零碎片、第一算术零碎片、第二算术零碎片,其中,第一布尔零碎片与第二布尔零碎片异或的结果为0,第一算术零碎片与第二算术零碎片相加的结果为0;将第一布尔零碎片和第一算术零碎片分配给第一节点;将第二布尔零碎片和第二算术零碎片分配给第二节点;由第一节点计算第

五碎片值与第一算术零碎片之和与第一布尔零碎片异或的结果,以获得第一运算碎片;由第二节点计算第六碎片值与第二算术零碎片之和,以获得第二运算碎片;由第一节点和第二节点联合利用第一节点处的第一并行前缀加法器和第二节点处的第二并行前缀加法器在第一节点处获得第一符号位碎片并且在第二节点处获得第二符号位碎片,其中,第一并行前缀加法器的输入为第一运算碎片和第三运算碎片,第二并行前缀加法器的输入为第二运算碎片和第四运算碎片,第三运算碎片等于0,第四运算碎片等于第二布尔零碎片;对第一符号位碎片与第二符号位碎片执行异或操作以获得比较值;响应于比较值为真,确定第一哈希值小于第二哈希值;以及响应于比较值不为真,确定第一哈希值不小于第二哈希值。

[0013] 在本公开的一些实施例中,根据所计算的交集比例来构建多个节点的虚拟数据价值网络包括:根据目标评估指标来评估多个节点的数据质量;以及将数据质量低于预设值的节点从虚拟数据价值网络中删除。

[0014] 在本公开的一些实施例中,根据所计算的交集比例来构建多个节点的虚拟数据价值网络包括:获取多个节点中的目标节点的数据源领域偏好;确定链接到目标节点的所有节点中符合数据源领域偏好的候选节点;以及在虚拟数据价值网络中,根据增强因子来调整每个候选节点到目标节点之间的边的权重。

[0015] 在本公开的一些实施例中,根据增强因子来调整每个候选节点到目标节点之间的边的权重包括:将每个候选节点到目标节点之间的边的权重乘以增强因子。其中,在数据源领域偏好为相似数据的情况下,增强因子为第一常数。在数据源领域偏好为互补数据的情况下,增强因子为第二常数。

[0016] 在本公开的一些实施例中,根据增强因子来调整每个候选节点到目标节点之间的边的权重包括根据下式来调整权重:

[0017] $$w_{t+1} = e^{w_t} \times e^{p}$$

[0018] 其中,w_{t+1}表示调整后的权重,w_t表示调整前的权重,p表示增强因子。在数据源领域偏好为相似数据的情况下,增强因子为第三常数。在数据源领域偏好为互补数据的情况下,增强因子为第四常数。

[0019] 根据本公开的第二方面,提供了一种用于在多个节点中进行资源推荐的装置。该装置包括至少一个处理器;以及存储有计算机程序的至少一个存储器。当计算机程序由至少一个处理器执行时,使得装置执行根据本公开的第一方面所述的方法的步骤。

[0020] 根据本公开的第三方面,提供了一种存储有计算机程序的计算机可读存储介质,其中,计算机程序在由处理器执行时实现根据本公开的第一方面所述的方法的步骤。

附图说明

[0021] 为了更清楚地说明本公开的实施例的技术方案,下面将对实施例的附图进行简要说明,应当知道,以下描述的附图仅仅涉及本公开的一些实施例,而非对本公开的限制,其中:

[0022] 图1是数联网的示意性拓扑图;

[0023] 图2是根据本公开的实施例的用于在多个节点中进行资源推荐的方法的示意性流程图;

[0024] 图3是根据本公开的实施例的确定第一哈希值是否小于第二哈希值的步骤的示意

性流程图和信令方案；

[0025] 图4是图3中的动作311的示意性流程图和信令方案；

[0026] 图5是图4中的动作403的示意性流程图和信令方案；

[0027] 图6是图4中的动作404和405的示意性流程图和信令方案；

[0028] 图7是根据本公开的实施例的虚拟数据价值网络的一个示例图；

[0029] 图8是根据本公开的实施例的虚拟数据价值网络的另一个示例图；

[0030] 图9是根据本公开的实施例的虚拟数据价值网络的又一个示例图；

[0031] 图10是根据本公开的实施例的目标节点与任一节点通过安全多方技术来计算A·B的示意性流程图和信令方案；

[0032] 图11是根据本公开的实施例的用于在多个节点中进行资源推荐的装置的示意性框图。

[0033] 在附图中，最后两位数字相同的标记对应于相同的元素。需要注意的是，附图中的元素是示意性的，没有按比例绘制。

具体实施方式

[0034] 为了使本公开的实施例的目的、技术方案和优点更加清楚，下面将结合附图，对本公开的实施例的技术方案进行清楚、完整的描述。显然，所描述的实施例是本公开的一部分实施例，而不是全部的实施例。基于所描述的本公开的实施例，本领域技术人员在无需创造性劳动的前提下所获得的所有其它实施例，也都属于本公开保护的范围。

[0035] 除非另外定义，否则在此使用的所有术语(包括技术和科学术语)具有与本公开主题所属领域的技术人员所通常理解的相同含义。进一步将理解的是，诸如在通常使用的词典中定义的那些的术语应解释为具有与说明书上下文和相关技术中它们的含义一致的含义，并且将不以理想化或过于正式的形式来解释，除非在此另外明确定义。另外，诸如“第一”和“第二”的术语仅用于将一个部件(或部件的一部分)与另一个部件(或部件的另一部分)区分开。

[0036] 如上所述，各行各业中的各类政务主体、行业主体、公司主体、机构主体可组合成网络，而此类网络中的主体数量可能非常大。数联网就是此类网络的一个示例。数联网作为开放性的网络体系，可以支持各类主体通过被允许的方式接入数联网，接入数联网后就形成了相应数据节点。数据节点之间存在联合建模、联合统计等多方协同计算的业务需求。成熟的数联网中数据节点数量级可以达到万、十万甚至是百万量级，规模跨越多个数量级。

[0037] 图1示出数联网的示意性拓扑图。数联网可包括多个子网10。每个子网10包括枢纽节点11和与枢纽节点直接连接的多个参与节点12。该多个子网10中的枢纽节点11相互直接连接。枢纽节点11与枢纽节点11之间可以通过专网进行互联。枢纽节点11承担对参与节点12进行信息聚合、寻址导航等功能。参与节点12可以是各类政务主体、行业主体、公司主体、机构主体等。直接连接到同一个枢纽节点11的参与节点12通过该枢纽节点11进行通信。直接连接到不同枢纽节点11的参与节点12通过它们各自直接连接的枢纽节点11进行通信。也就是说，参与节点12只与其直接连接的枢纽节点11直接通信，枢纽节点11之间可直接通信，而参与节点12之间需经由相应的枢纽节点11进行通信。

[0038] 在数据交易中，如果两个节点之间的数据交集比例越高，说明节点之间的共同用

户更多,因此可以互补很多用户的特征信息。比如运营商与银行之间,如果共同用户的交集比例高,通过融合运营商的用户特征与银行侧的用户特征,更全面地刻画个体用户,信息越全,为用户提供的服务就更加精准,进而提升广告点击转化率、交易成功率等,因此不同节点之间的共同用户交集的比例大小是衡量节点之间价值的有效指标。另外,场景的互补性也能够提升节点之间的价值,比如运营商场景的数据对于风控、保险等具备加成的能力,一般银行风控建模、保险建模中引入运营商的数据后,能够明显提升模型的识别能力。而电商场景,如果引入政务、银行等数据(如公积金、社保、水电费、资产等)后,可以更好得对用户的价格敏感度和消费水平进行捕捉,提升电商交易的转化率和成功率。上述所提到的交集比例、场景互补外,节点本身的数据质量对于价值计算也非常重要,有一句话形容数据质量的重要性,“垃圾进、垃圾出”,形容的就是如果输入的数据质量很差,那么基于该数据获得的模型或者统计结果也是垃圾。因此在衡量数据价值的同时,需要将数据质量也纳入考虑。

[0039] 在实践中,数联网中可能存在海量的子网10。单个子网10中可能存在海量的参与节点12。如果将每个参与节点12看作一个节点,那么数联网中的节点的数量可能是非常庞大的。在一些应用场景下,需要在海量的节点中安全且高效地进行资源推荐,从而在数联网中为目标节点快速匹配最适合的资源。

[0040] 本公开的实施例聚焦于为用户(例如,数联网用户)提供高效发现潜在感兴趣或者潜在价值数据源信息的方案,通过引入基于数联网的“安全推荐算法”的技术手段,根据数联网用户自身的需求和偏好,过滤和提供相关性更高的信息,帮助用户更快找到自己需要的内容。

[0041] “安全推荐算法”关注两个关键点:(1)个性化推荐;(2)安全计算。针对“个性化推荐”,数联网用户由于涉及的行业、业务、数据价值等存在差异性,因此会有不同的偏好和需求。推荐系统可以根据用户的独有特征,提供个性化的推荐内容,以满足用户的个性化的数据发现需求。这可以提高用户的满意度和用户体验,并增加用户的忠诚度。同时推荐系统可以帮助用户发现他们可能未曾了解或考虑过的内容。通过分析用户的历史交易行为,推荐系统可以向用户推荐类似或相关的内容,扩展用户发现新数据源的视野。推荐系统可以通过向用户推荐他们可能感兴趣的产品或服务,提高数据交易转化率。而“安全计算”主要体现在推荐算法计算的全流程需要保证原始数据的安全性。

[0042] 本公开主要是在拥有庞大数据资源的网络(例如,数联网)上提升发现数据资源的效率,采用基于大数据及安全计算融合的推荐算法,推荐潜在的对客户数据源需求有增益的候选数据源节点进行订阅,通过自动化方式来实现资源的探索和发现。

[0043] 图2示出根据本公开的实施例的用于在多个节点中进行资源推荐的方法的示意性流程图。该多个节点可以是数联网中的参与节点。

[0044] 在图2的框S202处,该多个节点中的每个节点将该节点的唯一标识符输入经训练的Word2Vec模型以获得该节点对应的向量。在这里,Word2Vec模型由多个节点采用安全多方技术进行联合训练并被训练成使得相似度越高的节点在向量空间中的距离越近。

[0045] 在本公开的一些实施例中,Word2Vec模型基于虚拟数据价值网络来对多个节点进行联合训练。虚拟数据价值网络是针对该多个节点构建的网络拓扑结构。Word2Vec模型可以从该虚拟数据价值网络来捕捉用户对数据源兴趣的偏好,得到数联网节点偏好细致表达的数学形式。

[0046] 下面先介绍如何构建虚拟数据价值网络。

[0047] 在本公开的一些实施例中,可通过安全求交方式计算该多个节点中的每两个节点之间的交集大小。安全求交方式指的是不会泄露节点的原始数据信息(隐私信息)并且不会泄露交集和非交集等敏感信息的方式。然后,根据该多个节点中的每两个节点之间的交集大小来计算该两个节点之间的交集比例。之后,根据所计算的交集比例来构建多个节点的虚拟数据价值网络。其中,在虚拟数据价值网络中,每个节点被表示为一个顶点,每两个节点之间的交集比例作为该两个节点之间的边的权重。如果两个节点之间的交集比例为0,则该两个节点不相连。

[0048] 在通过安全求交方式计算该多个节点中的每两个节点之间的交集大小的过程中,可通过相同的方式来计算该多个节点中的任意两个节点之间的交集大小。下面以计算第一节点与第二节点之间的交集大小为例来说明安全求交的计算过程。其中,第一节点和第二节点是该多个节点中的任意两个不同的节点。在本公开的实施例中,不需要计算节点与其自身的交集大小。

[0049] 在一个示例中,可获得第一节点的第一原始数据矩阵中的唯一标识符向量。该唯一标识符向量包括第一原始数据矩阵中的每个原始数据的唯一标识符(ID)。然后,利用哈希函数将第一原始数据矩阵中的唯一标识符向量转换成第一哈希向量。第一哈希向量包括多个第一哈希值,每个第一哈希值对应第一原始数据矩阵中的一个唯一标识符。

[0050] 并行地,可获得第二节点的第二原始数据矩阵中的唯一标识符向量。该唯一标识符向量包括第二原始数据矩阵中的每个原始数据的ID。然后,利用哈希函数将第二原始数据矩阵中的唯一标识符向量转换成第二哈希向量。第二哈希向量包括多个第二哈希值,每个第二哈希值对应第二原始数据矩阵中的一个唯一标识符。

[0051] 然后,比较第一哈希向量中的每个第一哈希值与第二哈希向量中的每个第二哈希值以将第一哈希向量中与第二哈希值相等的第一哈希值的个数确定为第一节点与第二节点之间的交集大小。

[0052] 可通过以下方式来安全地比较第一哈希值与第二哈希值:由第一节点和第二节点联合确定第一哈希值是否小于第二哈希值;如果第一哈希值不小于第二哈希值,则由第一节点和第二节点联合确定第二哈希值是否小于第一哈希值;如果第二哈希值不小于第一哈希值,确定第一哈希值等于第二哈希值。

[0053] 图3示出根据本公开的实施例的确定第一哈希值是否小于第二哈希值的步骤的示意性流程图和信令方案。由第一节点P1在动作301将第一哈希值x碎片化为第一碎片值x1和第二碎片值x2(x=x1+x2)并在动作303向第二节点P2发送第二碎片值x2。由第二节点P2在动作302将第二哈希值y碎片化为第三碎片值y1和第四碎片值y2(y=y1+y2)并在动作304向第一节点P1发送第三碎片值y1。

[0054] 在动作305,由第一节点P1将第一碎片值x1减去第三碎片值y1以获得第五碎片值z1,即z1=x1-y1。在动作306,由第二节点P2将第二碎片值x2减去第四碎片值y2以获得第六碎片值z2,即z2=x2-y2。

[0055] 可由第一节点P1和第二节点P2中的一者生成第一布尔零碎片a1、第二布尔零碎片a2、第一算术零碎片b1、第二算术零碎片b2。其中,第一布尔零碎片a1与第二布尔零碎片a2异或的结果为$0(a1\oplus a2=0)$,第一算术零碎片b1与第二算术零碎片b2相加的结果为0

(b1+b2＝0)。

[0056] 在动作307,将第一布尔零碎片a1和第一算术零碎片b1分配给第一节点P1。在动作308,将第二布尔零碎片a2和第二算术零碎片b2分配给第二节点P2。

[0057] 在动作309,由第一节点P1计算第五碎片值z1与第一算术零碎片b1之和与第一布尔零碎片a1异或的结果,以获得第一运算碎片op11,即,$op11=(z1+b1)\oplus a1$。第一节点P1还持有第三运算碎片op21,其中,op21＝0。

[0058] 在动作310,由第二节点P2计算第六碎片值z2与第二算术零碎片b2之和,以获得第二运算碎片op22,即,op22＝z2+b2。第二节点P2还持有第四运算碎片op12,其中,op12＝a2。

[0059] 在动作311,由第一节点P1和第二节点P2联合利用第一节点P1处的第一并行前缀加法器和第二节点P2处的第二并行前缀加法器在第一节点P1处获得第一符号位碎片B1并且在第二节点P2处获得第二符号位碎片B2,其中,第一并行前缀加法器的输入为第一运算碎片op11和第三运算碎片op21,第二并行前缀加法器的输入为第二运算碎片op22和第四运算碎片op12。

[0060] 在由第一节点P1来确定比较结果的示例中,第二节点P2在动作313向第一节点P1发送第二符号位碎片B2。由第一节点P1在动作314对第一符号位碎片B1与第二符号位碎片B2执行异或操作以获得比较值。如果比较值为真,则确定第一哈希值小于第二哈希值。如果比较值不为真,则确定第一哈希值不小于第二哈希值。类似地,也可以由第二节点P2来确定比较结果。

[0061] 图4示出图3中的动作311的具体过程。在动作403,由第一节点P1根据第一运算碎片op11和第三运算碎片op21并且由第二节点P2根据第二运算碎片op22和第四运算碎片op12来共同生成第一中间碎片G1和第二中间碎片G2。

[0062] 图5示出由第一节点P1和第二节点P2联合执行的与运算的示意性流程图和信令方案。在图5中以第一节点P1拥有第一输入碎片W1和第二输入碎片V1且第二节点P2拥有第三输入碎片W2和第四输入碎片V2为例来进行说明。当图4中的动作403使用图5所示的方案时,第一运算碎片op11相当于第一输入碎片W1,第三运算碎片op21相当于第二输入碎片V1,第二运算碎片op22相当于第三输入碎片W2,第四运算碎片op12相当于第四输入碎片V2。

[0063] 下面描述图5所示的过程。

[0064] 第一节点P1在动作501获得三元组碎片矩阵<R1,S1,T1>,第二节点P2在动作502获得三元组碎片矩阵<R2,S2,T2>。其中,

[0065] $(R1\oplus R2)\ \&\ (S1\oplus S2)=(T1\oplus T2)$。

[0066] 第一节点P1在动作503对W1和R1执行异或操作以获得第三中间碎片D1,对V1和S1执行异或操作以获得第四中间碎片E1。第二节点P2在动作504对W2和R2执行异或操作以获得第五中间碎片D2,对V2和S2执行异或操作以获得第六中间碎片E2。

[0067] 第二节点P2在动作505向第一节点P1发送D2和E2。第一节点P1在动作506向第二节点P2发送D1和E1。第一节点P1在动作507对D1和D2执行异或操作以获得第一合成碎片D,对E1和E2执行异或操作以获得第二合成碎片E。类似的,第二节点P2在动作508对D1和D2执行异或操作以获得第一合成碎片D,对E1和E2执行异或操作以获得第二合成碎片E。

[0068] 第一节点P1在动作509计算第一输出碎片$O1=T1\oplus(R1\&E)\oplus(S1\&D)$

[0069] ⊕(E&D)。第二节点P2在动作510计算第二输出碎片O2=T2⊕(R2&E)⊕(S2&D)。当图4中的动作403使用图5所示的方案时,第一中间碎片G1相当于第一输出碎片O1,第二中间碎片G2相当于第二输出碎片O2。

[0070] 回到图4,第一节点P1在动作404根据第五中间碎片p1(p1=op11⊕op21)对G1的每一位进行逐位循环计算。第二节点P2在动作405根据第六中间碎片p2(p2=op12⊕op22)对G2的每一位进行逐位循环计算。图6示出图4中的动作404和405的示意性流程图和信令方案。

[0071] 第一节点P1在动作601对G1执行左移2^i位的操作以得到第一临时碎片G11,即G11=G1<<2^i。第二节点P2在动作602对G2执行左移2^i位的操作以得到第二临时碎片G12,即G12=G2<<2^i。其中,i表示当前循环的索引。

[0072] 在动作603处,由第一节点P1和第二节点P2联合执行图5所示的与运算。G11相当于第一输入碎片W1,p1相当于第二输入碎片V1,G12相当于第三输入碎片W2,p2相当于第四输入碎片V2。经过动作603的操作,第一节点P1获得第七中间碎片F1,第二节点P2获得第八中间碎片F2。F1相当于第一输出碎片O1,F2相当于第二输出碎片O2。

[0073] 第一节点P1在动作604对p1执行左移2^i位的操作以获得第三临时碎片p11(即p11=p1<<2^i),然后再将p11更新为p11与kmask异或的结果(即,p11=p11⊕kmask)。其中,kmask是大小与op11相同的矩阵且其每一个元素值均为2^i-1。

[0074] 第二节点P2在动作605对p2执行左移2^i位的操作以获得第四临时碎片p12(即p12=p2<<2^i),然后再将p12更新为p12与kmask异或的结果(即,p12=p12⊕kmask)。

[0075] 在动作606处,由第一节点P1和第二节点P2联合执行图5所示的与运算。p1相当于第一输入碎片W1,p11相当于第二输入碎片V1,p2相当于第三输入碎片W2,p12相当于第四输入碎片V2。经过动作606的操作,第一节点P1获得更新后的p1,第二节点P2获得更新后的p2。更新后的p1相当于第一输出碎片O1,更新后的p2相当于第二输出碎片O2。更新后的p1会被代入下一循环的动作603处。更新后的p2也会被代入下一循环的动作603处。

[0076] 第一节点P1在动作607对G1和F1执行异或操作以获得更新后的G1(即,G1=G1⊕F1)。更新后的G1会被代入下一循环的动作601处。第二节点P2在动作608对G2和F2执行异或操作以获得更新后的G2(即,G2=G2⊕F2)。更新后的G2会被代入下一循环的动作602处。

[0077] 再次回到图4,第一节点P1在动作406对G1左移1位以获得第九中间碎片C1(即,C1=G1<<1)。第二节点P2在动作407对G2左移1位以获得第十中间碎片C2(即,C2=G2<<1)。

[0078] 第一节点P1在动作408对p1和C1执行异或操作以获得第十一中间碎片Z1(即,Z1=p1⊕C1)。第二节点P2在动作409对p2和C2执行异或操作以获得第十二中间碎片Z2(即,Z2=p2⊕C2)。

[0079] 第一节点P1在动作410对Z1和mask执行按位与操作以获得更新后的Z1(即,Z1=Z1&mask)。其中,mask=0x1<<n-1,n表示第一哈希值x的位数。第二节点P2在动作411对Z2和mask执行按位与操作以获得更新后的Z2(即,Z2=Z2&mask)。其中,mask=0x1<<n-1,n表示

第二哈希值y的位数。

[0080] 第一节点P1在动作412将Z1转换成布尔类型以获得第一符号位碎片B1。第二节点P2在动作413将Z2转换成布尔类型以获得第二符号位碎片B2。

[0081] 在上述过程中,由于第一节点P1没有获得第二哈希值的完整信息,而第二节点P2也没有获得第一哈希值的完整信息,因此该计算过程是安全的,除了交集大小外,不会泄露任何其他信息。

[0082] 确定第二哈希值是否小于第一哈希值的过程可与图3的过程类似,在此不再赘述。

[0083] 在根据该多个节点中的每两个节点之间的交集大小来计算该两个节点之间的交集比例的过程中,可以不考虑两个节点之间的方向性,也可以考虑两个节点之间的方向性。

[0084] 在不考虑两个节点之间的方向性的一些实施例中,第i节点与第j节点之间的交集比例被计算为:

[0085] $P = (2 \times C)/(A+B)$ (1)

[0086] 其中,P表示第i节点与第j节点之间的交集比例,A表示第i节点的数据集大小,B表示第j节点的数据集大小,C表示第i节点与第j节点之间的交集大小。在上下文中,第i节点和第j节点表示该多个节点中的任意两个不同的节点。

[0087] 在考虑两个节点之间的方向性的一些实施例中,从第i节点到第j节点的交集比例被计算为:

[0088] $Pa = C/A$ (2)

[0089] 其中,Pa表示从第i节点到第j节点的交集比例,A表示第i节点的数据集大小,C表示第i节点与第j节点之间的交集大小。

[0090] 从第j节点到第i节点的交集比例被计算为:

[0091] $Pb = C/B$ (3)

[0092] 其中,Pb表示从第j节点到第i节点的交集比例,B表示第j节点的数据集大小,C表示第i节点与第j节点之间的交集大小。

[0093] 如果不考虑两个节点之间的方向性,虚拟数据价值网络可被构建成无向图。任意两个不同的节点之间的交集比例按照式(1)来计算。图7示出被构建成无向图的虚拟数据价值网络的一个示例图。节点1至20被各自表示为一个顶点。两个顶点之间的边上标注的数字表示这两个节点之间的交集比例。

[0094] 如果考虑两个节点之间的方向性,虚拟数据价值网络可被构建成有向图。任意两个不同的节点之间的交集比例按照式(2)和式(3)来计算。图8示出被构建成有向图的虚拟数据价值网络的一个示例图。节点1至20被各自表示为一个顶点。两个顶点之间的边带有箭头,箭头表示方向。例如,从节点5到节点6的边上标注的数字0.34表示从节点5到节点6的交集比例为0.34。从节点6到节点5的边上标注的数字0.78表示从节点6到节点5的交集比例为0.78。

[0095] 在本公开的一些实施例中,在构建了图7或者图8所示的虚拟数据价值网络之后,可根据目标评估指标来评估虚拟数据价值网络中的多个节点的数据质量。然后将数据质量低于预设值的节点(可称为“低质量节点”)从虚拟数据价值网络中删除。这相当于一次初步筛选。目标评估指标可采用基于联邦学习的预处理算法来计算。目标评估指标可包括:数据缺失率指标(如果缺失信息过大不满足计算所需要求)、异常率指标(值域检测、数值合法性

检测、数值逻辑检测)、特征共线性方差膨胀因子(Variance Inflation Factor,简称VIF)检测、信息值(information value,简称IV)的检测(特征对于模型预测能力的贡献度)、数据重复性检测(重复数据比例)等。图9示出图7中的低质量节点被删除后的虚拟数据价值网络的一个示例图。

[0096] 进一步的,本公开的实施例提出可对虚拟数据价值网络中的边进行基于节点领域相似或者互补的增强。这里的虚拟数据价值网络可以是初始构建的虚拟数据价值网络,也可以是删除低质量节点之后的虚拟数据价值网络。数联网中的节点一般分属不同领域,比如运营商领域、保险领域、社保领域等。数联网中的节点可以明确提出其所需的数据源的领域偏好,比如保险公司节点提出需要社保政务数据,或者保险公司节点提出开展同类保险业务的数据源的需求。基于这类显式偏好的需求,通过增强边权重来体现偏好的强弱。

[0097] 在本公开的一些实施例中,可获取多个节点中的目标节点的数据源领域偏好。然后确定链接到目标节点的所有节点中符合数据源领域偏好的候选节点。接着,在虚拟数据价值网络中,根据增强因子来调整每个候选节点到目标节点之间的边的权重。边的权重可以通过线性方式来增强,也可以通过非线性方式来增强。

[0098] 在边的权重通过线性方式来增强的实施例中,可将每个候选节点到目标节点之间的边的权重乘以增强因子。其中,在数据源领域偏好为相似数据的情况下,增强因子为第一常数。在数据源领域偏好为互补数据的情况下,增强因子为第二常数。第一常数和第二常数都大于1。第一常数和第二常数可以根据需求进行调整,以达到合适的边权重增强效果。在一个示例中,第一常数等于1.1,第二常数等于1.3。

[0099] 在边的权重通过非线性方式来增强的实施例中,可根据下式来调整权重:

[0100] $$w_{t+1} = e^{w_t} \times e^{p} \qquad (4)$$

[0101] 其中,w_{t+1}表示调整后的权重,w_t表示调整前的权重,p表示增强因子。在数据源领域偏好为相似数据的情况下,增强因子为第三常数。在数据源领域偏好为互补数据的情况下,增强因子为第四常数。第三常数和第四常数都小于1。在一个示例中,第三常数等于0.2,第四常数等于0.5。

[0102] 回到图2,在框S204处,采用安全多方技术来计算多个节点中的目标节点对应的向量与其它节点对应的向量之间的相似度,其它节点是多个节点中除了目标节点之外的节点。进一步的,在本公开的一些实施例中,其它节点还可以被缩小为与目标节点之间的交集比例不为0的候选节点。

[0103] 在本公开的一些实施例中,目标节点对应的向量与其它节点中的任一节点对应的向量之间的相似度根据下式来计算:

[0104] $$\cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} \qquad (5)$$

[0105] 其中,cos(θ)表示相似度,A表示目标节点对应的第一向量,B表示该任一节点对应的第二向量。

[0106] 在本公开的一些实施例中,在基于式(5)计算目标节点对应的向量与该任一节点对应的向量之间的相似度的过程中,目标节点与该任一节点可通过安全多方技术来计算A·B。并行地,或者先后地,目标节点在本地独立计算||A||,该任一节点在本地独立计算||B||并向目标节点发送所计算的||B||。然后,目标节点根据A·B、||A||和||B||来计算相似

度cos(θ)。

[0107] 图10示出根据本公开的实施例的目标节点PA与该任一节点PB通过安全多方技术来计算A·B的示意性流程图和信令方案。

[0108] 目标节点PA在动作1001将第一向量A碎片化成第一碎片矩阵f0和第二碎片矩阵f1(A=f0+f1)并在动作1003向该任一节点PB提供第二碎片矩阵f1。该任一节点PB在动作1002将第二向量B碎片化成第三碎片矩阵lib0和第四碎片矩阵lib1(B=lib0+lib1)并在动作1004向目标节点PA提供第三碎片矩阵lib0。该任一节点PB还可生成用于半同态加密或者同态加密的一对公钥和私钥。

[0109] 目标节点PA在动作1005将第一碎片矩阵f0乘以(矩阵对应元素相乘)噪声矩阵r以生成加噪矩阵Nos(f0)(即,Nos(f0)=f0⊙r,Nos(f0)是f0与r的哈达玛积)并在动作1007向该任一节点PB提供加噪矩阵Nos(f0),其中,噪声矩阵r的大小与第一碎片矩阵f0相同且噪声矩阵r中的每个元素为随机数。

[0110] 目标节点PA在动作1009将第一向量A与第三碎片矩阵lib0进行相乘(矩阵相乘)以生成第一值vala(即,vala=A·lib0,vala是A与lib0的矩阵乘积)。

[0111] 该任一节点PB在动作1010将加噪矩阵Nos(f0)乘以(矩阵对应元素相乘)第四碎片矩阵lib1的转置矩阵$lib1^T$(这里的上标T表示矩阵转置)以生成第一加噪乘积矩阵Nos(valb1)(即,Nos(valb1)=Nos(f0)⊙$lib1^T$,Nos(valb1)是Nos(f0)与$lib1^T$的哈达玛积),对第一加噪乘积矩阵Nos(valb1)进行逐项加密以生成密态第一加噪乘积矩阵Enc(Nos(valb1)),并在动作1012向目标节点PA提供密态第一加噪乘积矩阵Enc(Nos(valb1))。

[0112] 目标节点PA在动作1013逐项计算密态第一加噪乘积矩阵Enc(Nos(valb1))与噪声矩阵r之商以生成密态第一乘积矩阵Enc(valb1)(即,Enc(Nos(valb1))/r=Enc(valb1))。在一个示例中,可通过将密态第一加噪乘积矩阵Enc(Nos(valb1))乘以噪声矩阵r的逆矩阵1/r来生成密态第一乘积矩阵Enc(valb1)(即,Enc(Nos(valb1))⊙(1/r)=Enc(valb1))。目标节点PA还在动作1013对密态第一乘积矩阵Enc(valb1)进行密态求和以生成第一密态和Enc(valb1f),将第一密态和Enc(valb1f)减去随机数R以生成密态第一中间值Enc(valb)(即,Enc(valb)=Enc(valb1f)-R),并在动作1015向目标节点PA提供密态第一中间值Enc(valb)。

[0113] 该任一节点PB在动作1014对第二碎片矩阵f1和第四碎片矩阵lib1进行相乘(矩阵相乘)以生成第二值valb2(即,valb2=f1·lib1,valb2是f1与lib1的矩阵乘积)。该任一节点PB在动作1016对密态第一中间值Enc(valb)进行同态解密以获得第一中间值valb,将第二值valb2与第一中间值valb相加以生成第二中间值valf2(即,valf2=valb+valb2),并在动作1017向目标节点PA提供第二中间值valf2。

[0114] 目标节点PA在动作1018将随机数R与第一值vala相加以获得临时中间值valf1。目标节点PA在动作1019将第二中间值valf2与临时中间值valf1相加以获得A·B的结果。A·B=valf1+valf2。

[0115] 下面以一个示例来更具体地说明图10的过程。图10的计算任务是对A和B进行内积计算,其中A的大小为1×K。B的大小为K×1,因此A·B的大小为1×1。

[0116] 第一向量A被碎片化为A=f0+f1。第二向量B被碎片化为B=lib0+lib1。经过动作1003和1004,目标节点PA持有f0、f1、lib0,该任一节点PB持有lib0、lib1、f1。f0和f1的大小

都为1×K。lib0和lib1的大小都为K×1。

[0117] 假设:第一向量A为[1,2,3],f0=[1,1,2],f1=[0,1,1]。第二向量B为$[3,4,5]^T$,$lib0=[0,0,1]^T$,$lib1=[3,4,4]^T$。目标节点PA在动作1005生成的噪声矩阵r为[0.5,0.6,0.7]。目标节点PA在动作1013生成的随机数R为100。

[0118] 那么目标节点PA在动作1005计算得到Nos(f0)=[1,1,2]⊙[0.5,0.6,0.7]=[0.5,0.6,1.4]。该任一节点PB在动作1010计算得到Nos(valb1)=Nos(f0)⊙$lib1^T$=[0.5,0.6,1.4]⊙[3,4,4]=[1.5,2.4,5.6]。经过公钥逐项加密得到Enc(Nos(valb1))=[Enc(1.5),Enc(2.4),Enc(5.6)]。在动作1012该任一节点PB将Enc(Nos(valb1))发送给目标节点PA进行去噪处理。

[0119] 目标节点PA在动作1013对Enc(Nos(valb1))逐项进行去除噪声。Enc(valb1)=Enc(Nos(valb1))⊙(1/r)=[Enc(3),Enc(4),Enc(8)]。进一步对Enc(valb1)进行求和,得到Enc(valb1f)=Enc(15)。那么Enc(valb)=Enc(valb1f)-100=Enc(-85)。目标节点PA在动作1015将Enc(valb)发送给该任一节点PB进行解密。该任一节点PB在动作1016得到valb=-85。

[0120] 该任一节点PB在动作1014执行valb2=f1·lib1=[0,1,1]·$[3,4,4]^T$=8。该任一节点PB在动作1016计算valf2=valb+valb2=-85+8=-77。

[0121] 目标节点PA在动作1009计算vala=A·lib0=[1,2,3]·$[0,0,1]^T$=3。目标节点PA在动作1018计算valf1=R+vala=100+3=103。目标节点PA在动作1019计算valf1+valf2=103-77=26。该结果与A·B=[1,2,3]·$[3,4,5]^T$=26的结果一致。可见,上述过程能够正确计算出A·B的结果,且不会泄露任何一方的数据。

[0122] 再回到图2,在框S206处,根据在框S204处计算的相似度来向目标节点PA推荐其它节点上的资源。cos(θ)的值越接近1表示越相似,cos(θ)的值越接近0表示越不相似。

[0123] 在本公开的一些实施例中,对于每个目标节点,计算其与所有候选节点的相似度。根据相似度的降序对候选节点进行排序。相似度越高的候选节点越可能被推荐给目标节点。在一个示例中,可以选择相似度最高的前几个候选节点作为对目标节点的推荐结果。这些候选节点可以显示在目标节点在数联网操作平台的界面上,通过目标用户的选择订阅来实现数据源的自动化探索。

[0124] 另外对于网络中的新加入节点,由于其未进行嵌入向量的表征学习(即,未参与Word2Vec模型的训练),其可以通过价值网络的构建方法,加入到虚拟数据价值网络中。例如,可使用相邻节点的向量表征,如采用邻接节点向量的均值来计算新加入节点的初始化向量表征,进而实现可被计算相似度的目的,解决新加入节点的冷启动问题。

[0125] 在图3至图6和图10的流程图中,动作编号的顺序不用于限定动作执行的先后顺序。除了具有输入输出关系(或者因果关系)的动作必须具有先后顺序之外,其他动作可以并行地执行,或者按照除图示之外的其他顺序来执行。

[0126] 图11示出根据本公开的实施例的用于在多个节点中进行资源推荐的装置1100的示意性框图。如图11所示,该装置1100可包括处理器1110和存储有计算机程序的存储器1120。当计算机程序由处理器1110执行时,使得装置1100可执行如图2所示的方法200的步骤。在一个示例中,装置1100可以是计算机设备或云计算节点。装置1100可使得多个节点中的每个节点将该节点的唯一标识符输入经训练的Word2Vec模型以获得该节点对应的向量。

Word2Vec模型由多个节点采用安全多方技术进行联合训练并被训练成使得相似度越高的节点在向量空间中的距离越近。装置1100可采用安全多方技术来计算多个节点中的目标节点对应的向量与其它节点对应的向量之间的相似度,其它节点是多个节点中除了目标节点之外的节点。装置1100可根据相似度来向目标节点推荐其它节点上的资源。

[0127] 在本公开的实施例中,处理器1110可以是例如中央处理单元(CPU)、微处理器、数字信号处理器(DSP)、基于多核的处理器架构的处理器等。存储器1120可以是使用数据存储技术实现的任何类型的存储器,包括但不限于随机存取存储器、只读存储器、基于半导体的存储器、闪存、磁盘存储器等。

[0128] 此外,在本公开的实施例中,装置1100也可包括输入设备1130,例如键盘、鼠标等,用于输入节点信息。另外,装置1100还可包括输出设备1140,例如显示器等,用于输出资源推荐结果。

[0129] 在本公开的其它实施例中,还提供了一种存储有计算机程序的计算机可读存储介质,其中,计算机程序在由处理器执行时能够实现如图2所示的方法的步骤。

[0130] 综上所述,根据本公开的实施例的用于在多个节点中进行资源推荐的方法能够实现安全且高效的批量资源推荐,在整个资源推荐过程中,不会泄露任一节点的原始数据信息(隐私信息)并且不会泄露节点之间的交集和非交集等敏感信息。

[0131] 附图中的流程图和框图显示了根据本公开的多个实施例的装置和方法的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分,所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的是,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0132] 除非上下文中另外明确地指出,否则在本文和所附权利要求中所使用的词语的单数形式包括复数,反之亦然。因而,当提及单数时,通常包括相应术语的复数。相似地,措辞“包含”和“包括”将解释为包含在内而不是独占性地。同样地,术语“包括”和“或”应当解释为包括在内的,除非本文中明确禁止这样的解释。在本文中使用术语“示例”之处,特别是当其位于一组术语之后时,所述“示例”仅仅是示例性的和阐述性的,且不应当被认为是独占性的或广泛性的。

[0133] 适应性的进一步的方面和范围从本文中提供的描述变得明显。应当理解,本申请的各个方面可以单独或者与一个或多个其它方面组合实施。还应当理解,本文中的描述和特定实施例旨在仅说明的目的并不旨在限制本申请的范围。

[0134] 以上对本公开的若干实施例进行了详细描述,但显然,本领域技术人员可以在不脱离本公开的精神和范围的情况下对本公开的实施例进行各种修改和变型。本公开的保护范围由所附的权利要求限定。

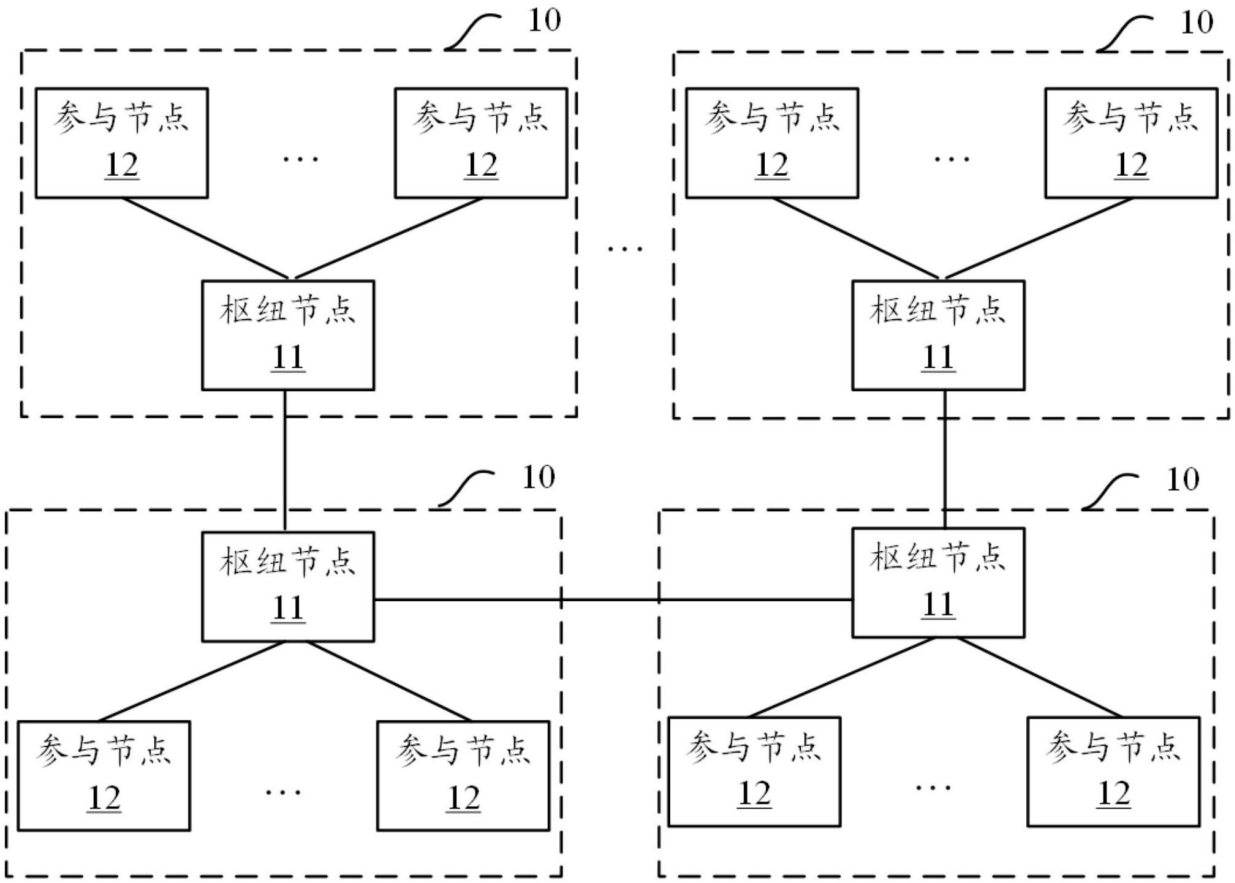

10
参与节点
12
…
参与节点
12
枢纽节点
11
10
参与节点
12
…
参与节点
12
枢纽节点
11
10
枢纽节点
11
10
枢纽节点
11
参与节点
12
…
参与节点
12
参与节点
12
…
参与节点
12

图1

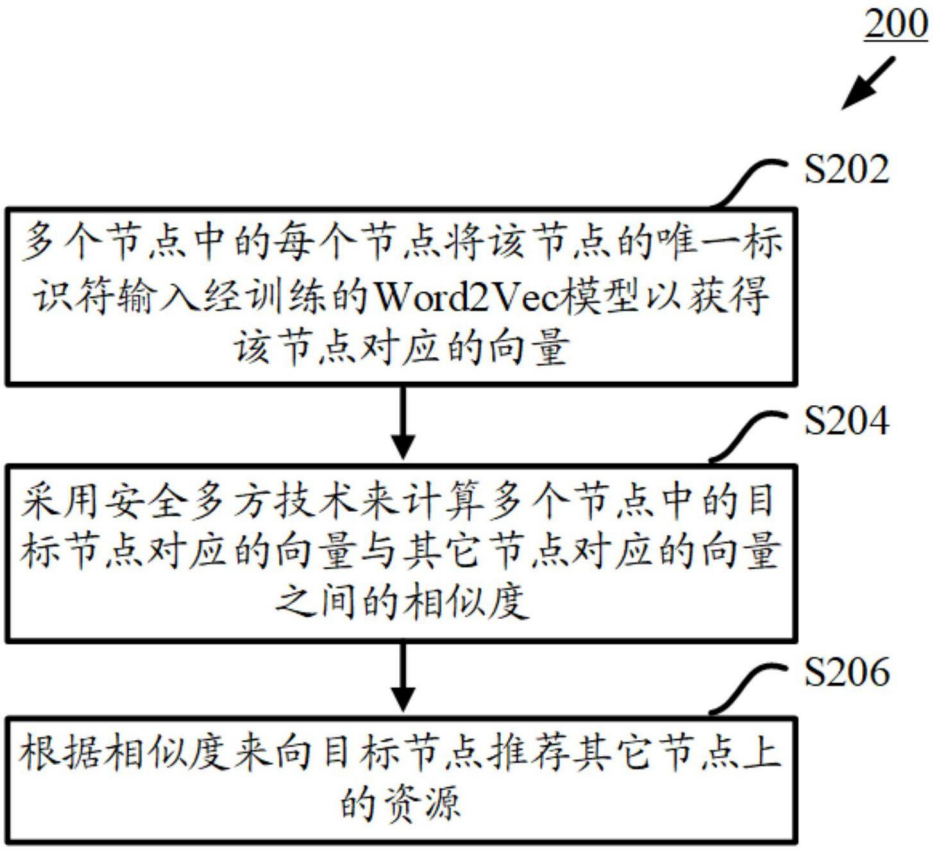

200
S202
多个节点中的每个节点将该节点的唯一标识符输入经训练的Word2Vec模型以获得该节点对应的向量
S204
采用安全多方技术来计算多个节点中的目标节点对应的向量与其它节点对应的向量之间的相似度
S206
根据相似度来向目标节点推荐其它节点上的资源

图2

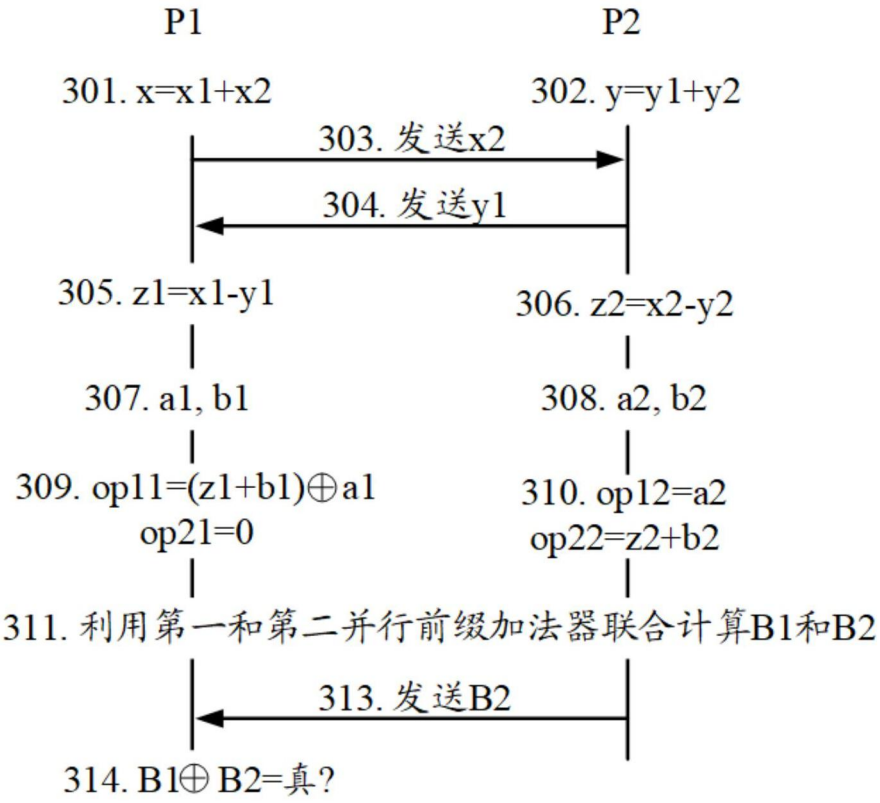

P1
P2
301. x=x1+x2
302. y=y1+y2
303. 发送x2
304. 发送y1
305. z1=x1-y1
306. z2=x2-y2
307. a1, b1
308. a2, b2
309. op11=(z1+b1)⊕a1
op21=0
310. op12=a2
op22=z2+b2
311. 利用第一和第二并行前缀加法器联合计算B1和B2
313. 发送B2
314. B1⊕ B2=真?

图3

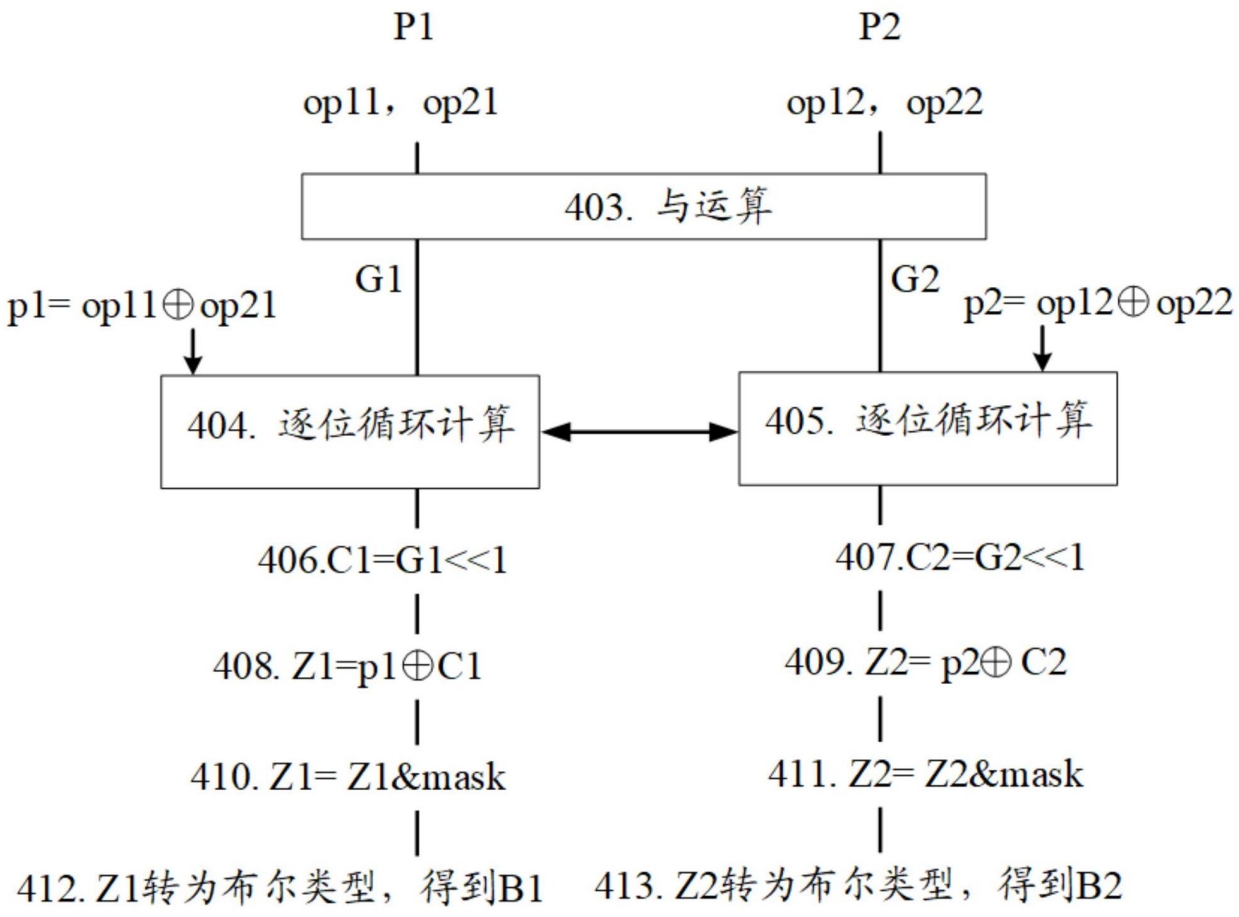

P1
P2
op11，op21
op12，op22
403. 与运算
G1
G2
p1= op11⊕op21
p2= op12⊕op22
404. 逐位循环计算
405. 逐位循环计算
406.C1=G1<<1
407.C2=G2<<1
408. Z1=p1⊕C1
409. Z2= p2⊕ C2
410. Z1= Z1&mask
411. Z2= Z2&mask
412. Z1转为布尔类型，得到B1
413. Z2转为布尔类型，得到B2

图4

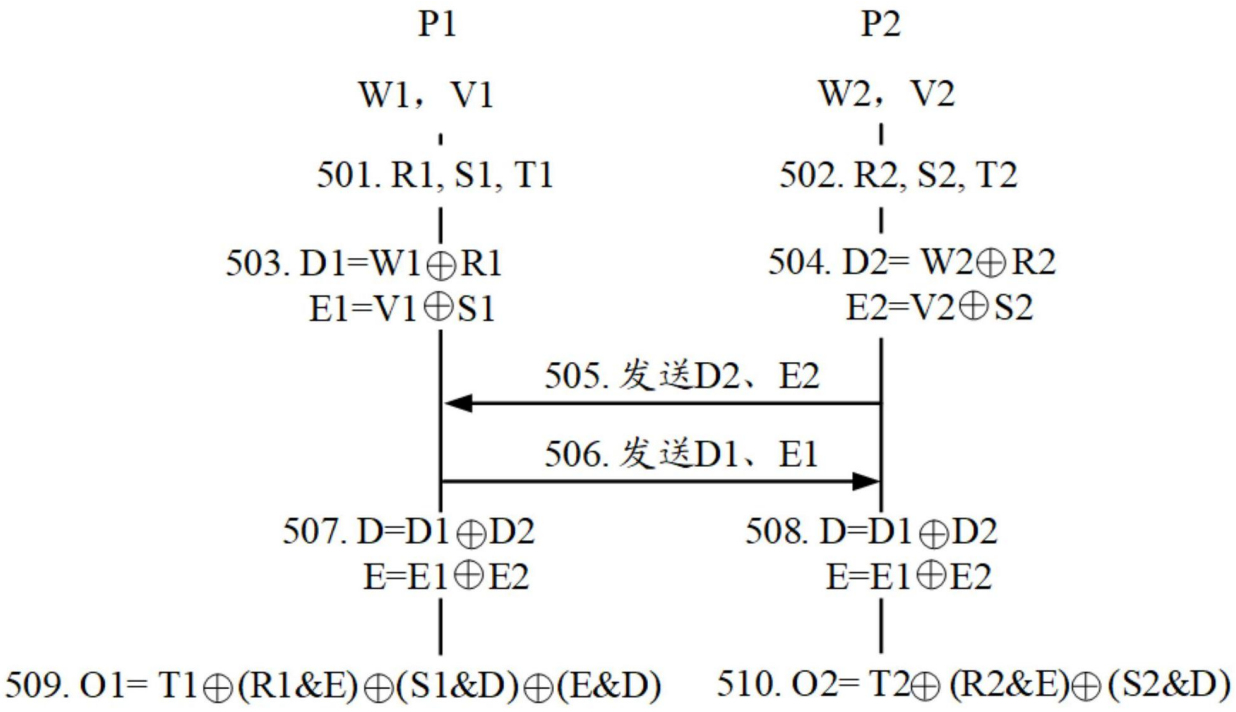

P1
P2
W1，V1
W2，V2
501. R1, S1, T1
502. R2, S2, T2
503. D1=W1⊕R1
E1=V1⊕S1
504. D2= W2⊕R2
E2=V2⊕S2
505. 发送D2、E2
506. 发送D1、E1
507. D=D1⊕D2
E=E1⊕E2
508. D=D1⊕D2
E=E1⊕E2
509. O1= T1⊕(R1&E)⊕(S1&D)⊕(E&D)
510. O2= T2⊕ (R2&E)⊕(S2&D)

图5

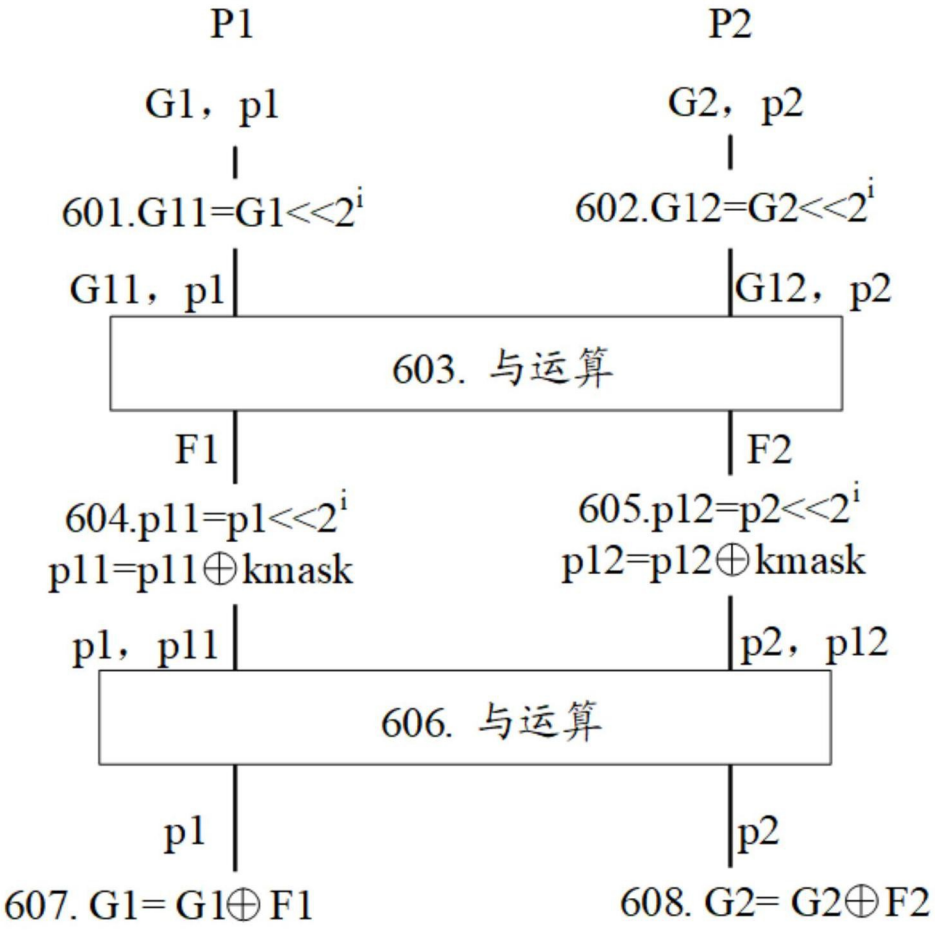

P1
P2
G1，p1
G2，p2
601.G11=G1<<2^i
602.G12=G2<<2^i
G11，p1
G12，p2
603. 与运算
F1
F2
604.p11=p1<<2^i
p11=p11⊕kmask
605.p12=p2<<2^i
p12=p12⊕kmask
p1，p11
p2，p12
606. 与运算
p1
p2
607. G1= G1⊕F1
608. G2= G2⊕F2

图6

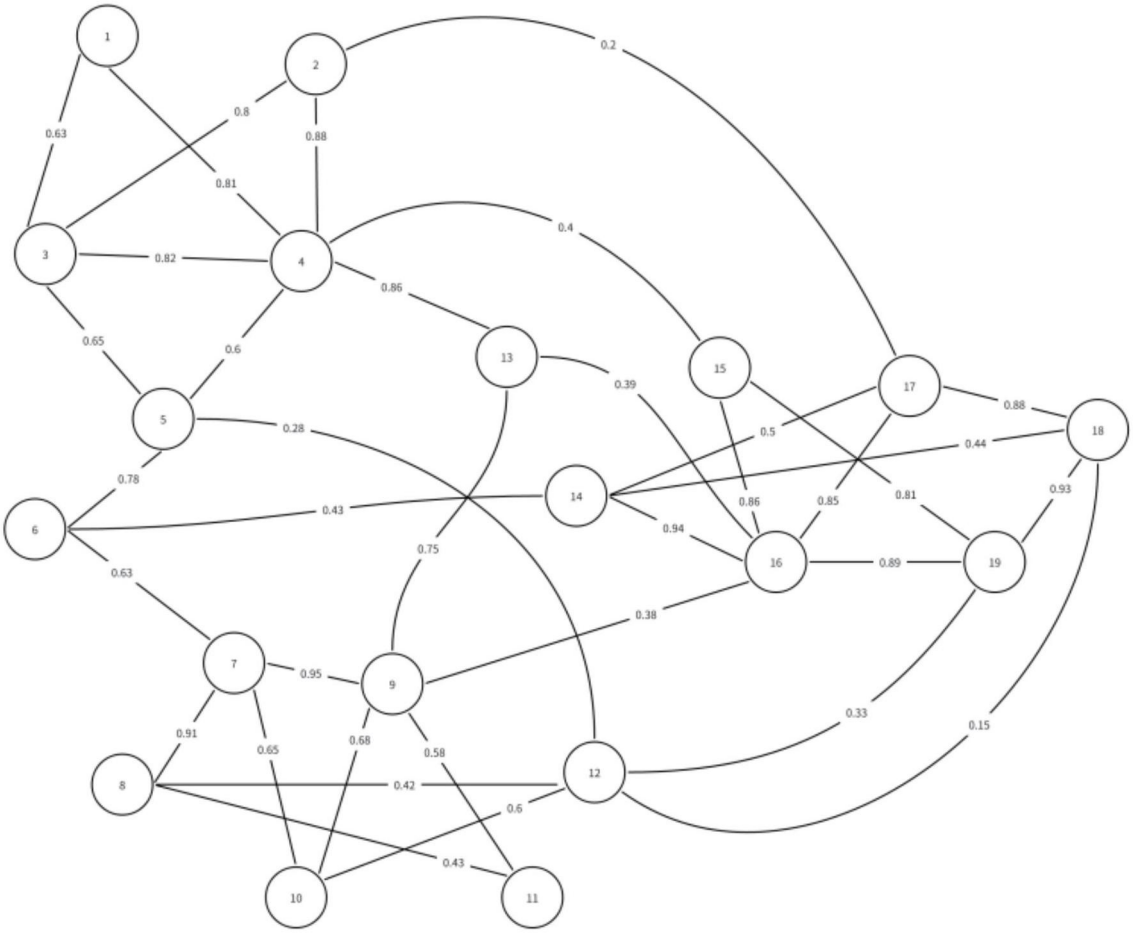

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
0.2
0.8
0.63
0.88
0.81
0.4
0.82
0.86
0.65
0.6
0.39
0.88
0.5
0.44
0.28
0.78
0.43
0.94
0.86
0.85
0.81
0.93
0.75
0.89
0.63
0.38
0.95
0.91
0.65
0.68
0.58
0.33
0.15
0.42
0.6
0.43

图7

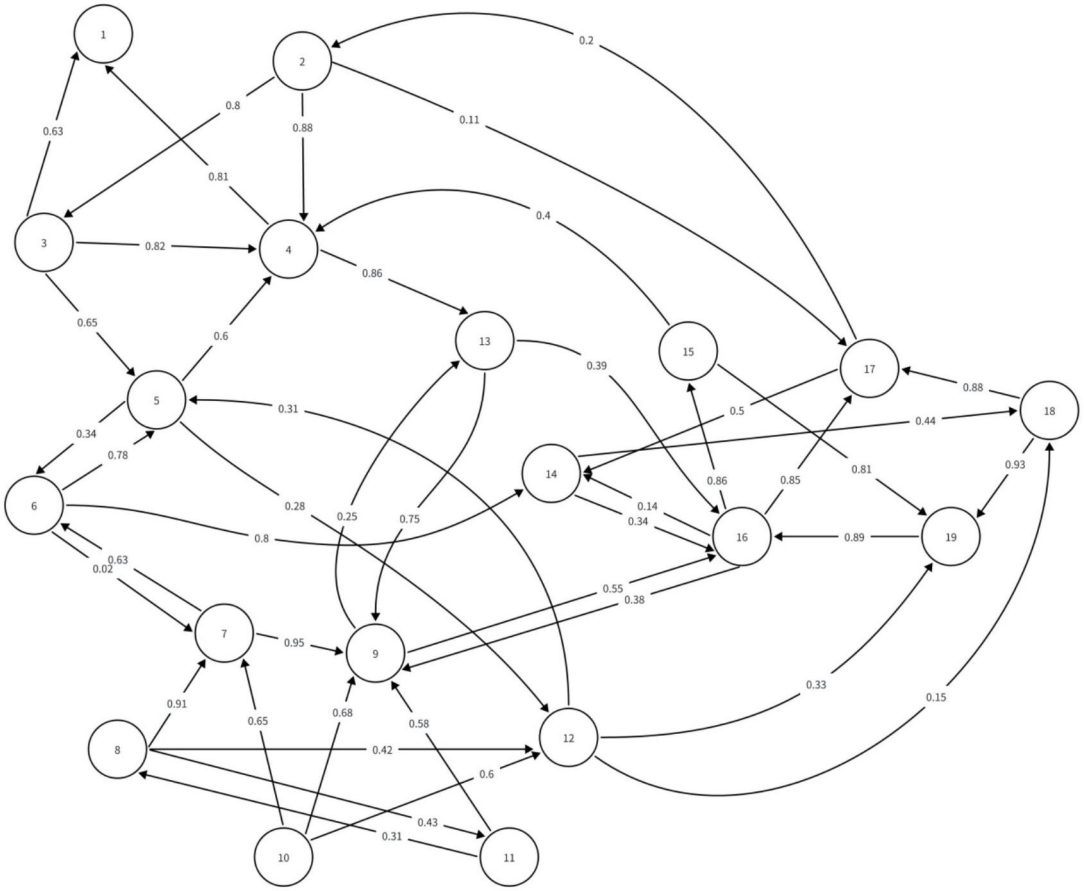

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
0.2
0.8
0.11
0.63
0.88
0.81
0.4
0.82
0.86
0.65
0.6
0.39
0.31
0.5
0.44
0.34
0.78
0.28
0.25
0.75
0.14
0.34
0.86
0.85
0.81
0.88
0.93
0.89
0.8
0.63
0.02
0.55
0.38
0.95
0.91
0.65
0.68
0.58
0.33
0.15
0.42
0.6
0.43
0.31

图8

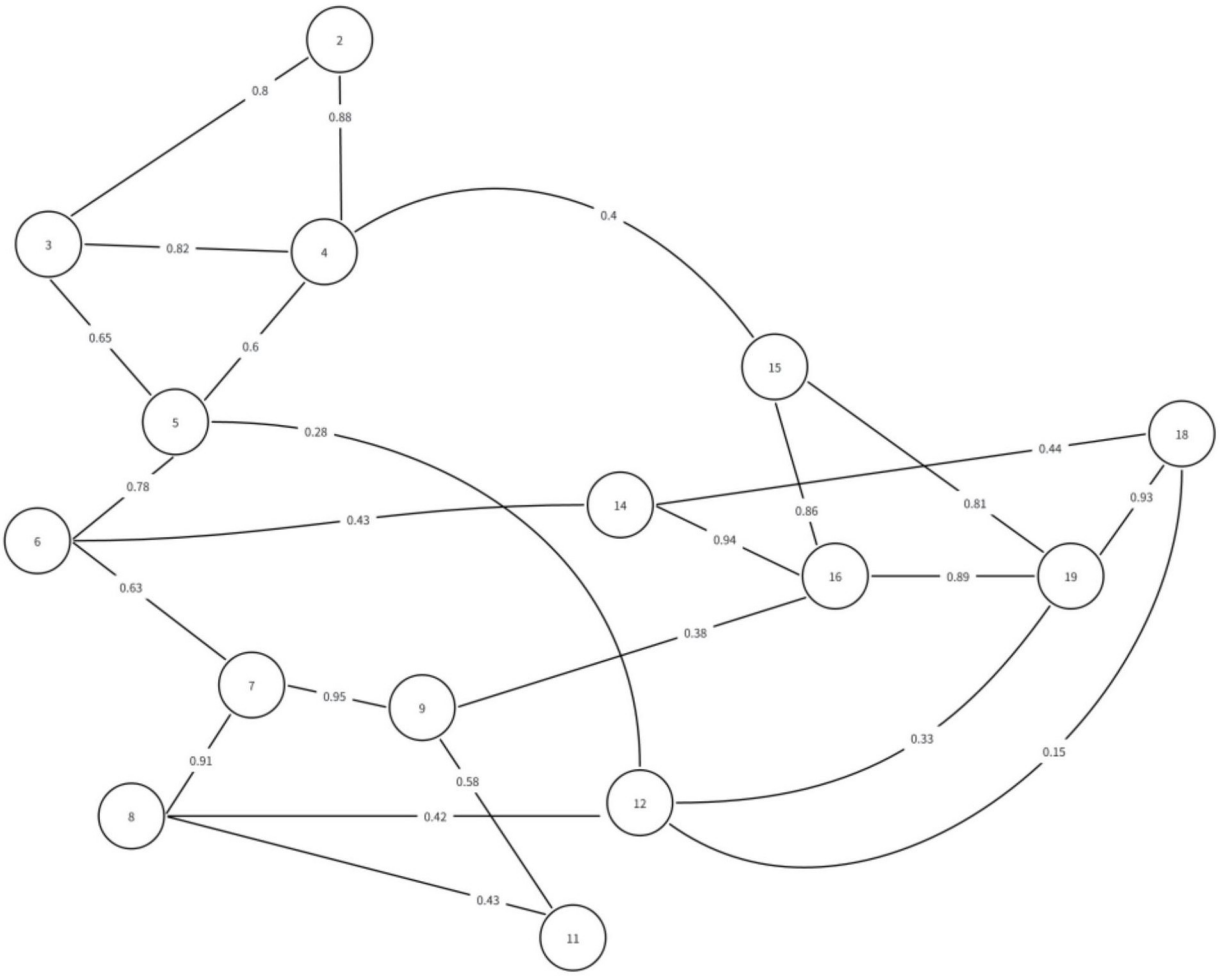

2
0.8
0.88
3
0.82
4
0.4
0.65
0.6
15
5
0.28
18
0.44
0.78
14
0.86
0.81
0.93
6
0.43
0.94
16
0.89
19
0.63
0.38
7
0.95
9
0.33
0.15
0.91
0.58
8
0.42
12
0.43
11

图9

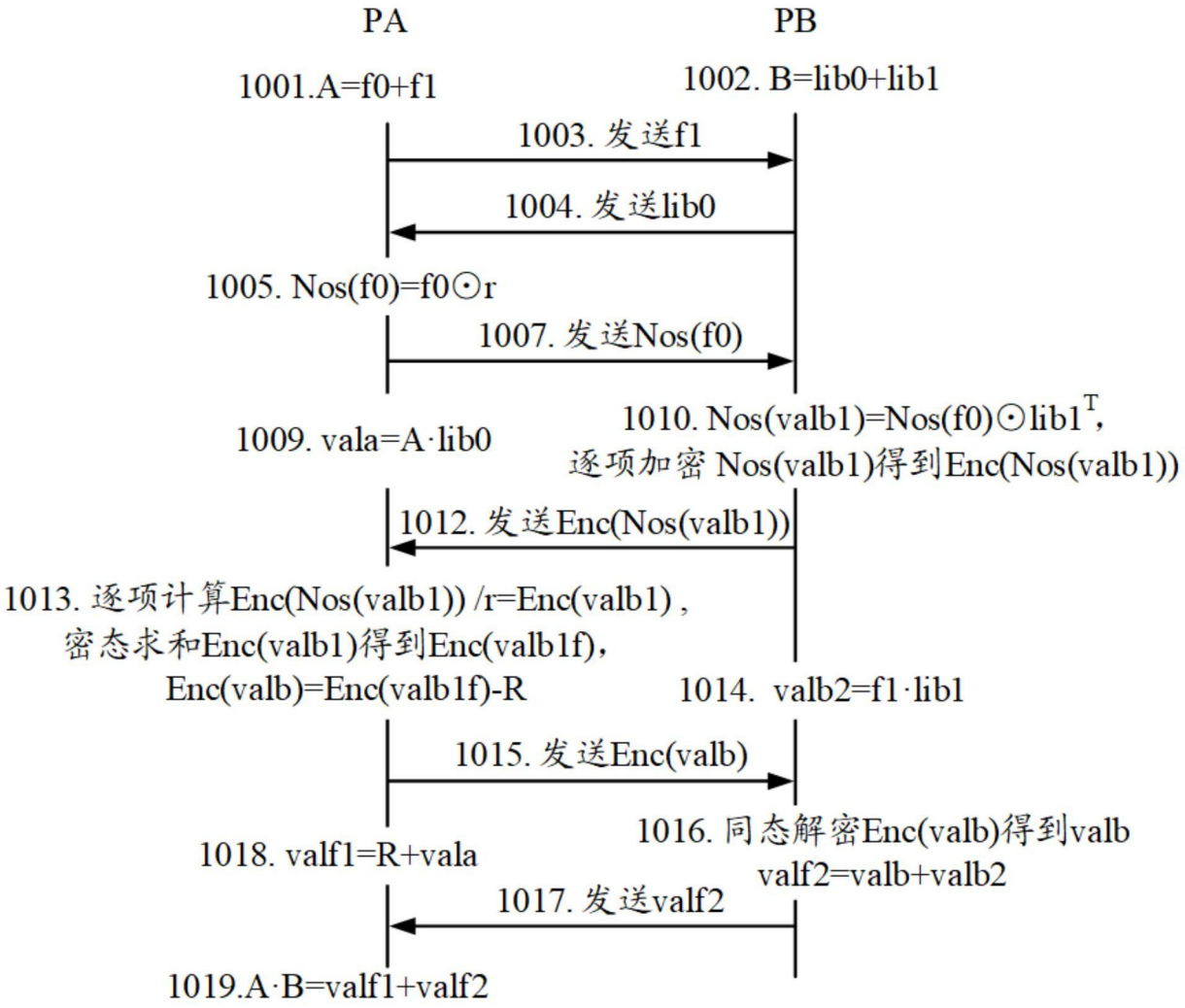

PA
PB
1001.A=f0+f1
1002. B=lib0+lib1
1003. 发送f1
1004. 发送lib0
1005. Nos(f0)=f0⊙r
1007. 发送Nos(f0)
1009. vala=A·lib0
1010. Nos(valb1)=Nos(f0)⊙lib1^T,
逐项加密 Nos(valb1)得到Enc(Nos(valb1))
1012. 发送Enc(Nos(valb1))
1013. 逐项计算Enc(Nos(valb1)) /r=Enc(valb1) ,
密态求和Enc(valb1)得到Enc(valb1f),
Enc(valb)=Enc(valb1f)-R
1014. valb2=f1·lib1
1015. 发送Enc(valb)
1016. 同态解密Enc(valb)得到valb
valf2=valb+valb2
1017. 发送valf2
1018. valf1=R+vala
1019.A·B=valf1+valf2

图10

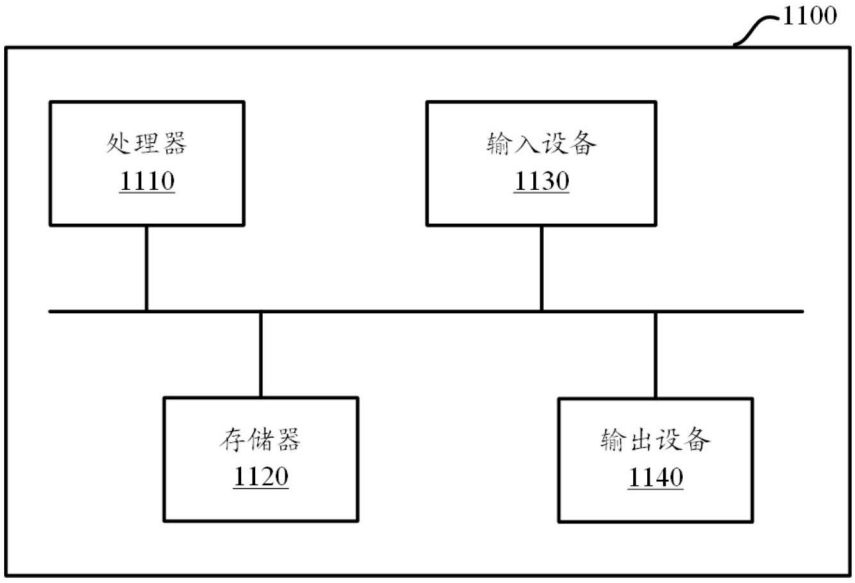


图11