(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117493630 A
(43)申请公布日 2024.02.02

(21)申请号 202311595088.1

(22)申请日 2023.11.27

(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 王兆凯 赵东 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 16/9035(2019.01)
G06F 16/903(2019.01)
G06F 16/901(2019.01)

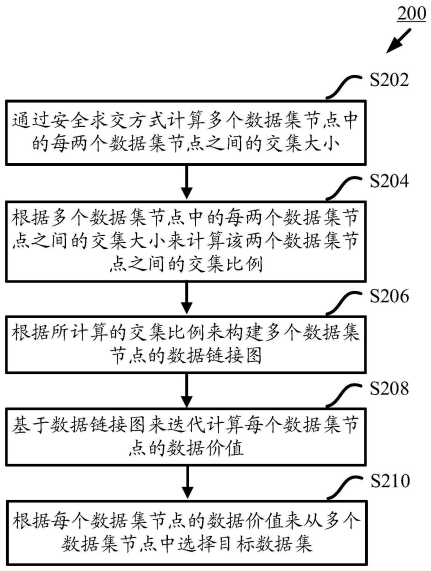
权利要求书3页 说明书11页 附图10页

(54)发明名称

从多个数据集节点中筛选数据集的方法

(57)摘要

本公开的实施例提供一种从多个数据集节点中筛选数据集的方法。该方法包括:通过安全求交方式计算每两个数据集节点之间的交集大小;根据每两个数据集节点之间的交集大小来计算该两个数据集节点之间的交集比例;根据所计算的交集比例来构建数据链接图;基于数据链接图来迭代计算每个数据集节点的数据价值;以及根据每个数据集节点的数据价值来从多个数据集节点中选择目标数据集。基于数据链接图来迭代计算每个数据集节点的数据价值包括:将每个数据集节点的数据价值初始化为预设常数值;以及在每一轮迭代过程中遍历数据链接图以将每个数据集节点的数据价值按照边的权重分配给其所链接的数据集节点,直至每个数据集节点的数据价值收敛。



CN 117493630 A

1.一种从多个数据集节点中筛选数据集的方法,其特征在于,所述方法包括:

通过安全求交方式计算所述多个数据集节点中的每两个数据集节点之间的交集大小;

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

根据所计算的交集比例来构建所述多个数据集节点的数据链接图,其中,在所述数据链接图中,每个数据集节点被表示为一个顶点,每两个数据集节点之间的交集比例作为该两个数据集节点之间的边的权重;

基于所述数据链接图来迭代计算每个数据集节点的数据价值;以及

根据每个数据集节点的所述数据价值来从所述多个数据集节点中选择目标数据集;

其中,基于所述数据链接图来迭代计算每个数据集节点的数据价值包括:

将每个数据集节点的数据价值初始化为预设常数值;以及

在每一轮迭代过程中遍历所述数据链接图以将每个数据集节点的数据价值按照边的权重分配给其所链接的数据集节点,直至每个数据集节点的所述数据价值收敛。

2.根据权利要求1所述的方法,其特征在于,所述数据链接图被构建成无向图,第i数据集节点与第j数据集节点之间的交集比例被计算为:

$P=(2\times C)/(A+B)$

其中,P表示所述第i数据集节点与所述第j数据集节点之间的交集比例,A表示所述第i数据集节点的数据集大小,B表示所述第j数据集节点的数据集大小,C表示所述第i数据集节点与所述第j数据集节点之间的交集大小。

3.根据权利要求2所述的方法,其特征在于,在每一轮迭代过程中,第i数据集节点的数据价值被计算为:

$$PR(i)=(1-d)+d\times\sum_{j\in In(i)}\frac{PR(j)\times Weight(j,i)}{\sum_{k\in Out(j)}Weight(j,k)}$$

其中,PR(i)表示所述第i数据集节点的数据价值,d是大于0且小于1的常数,In(i)表示链接到所述第i数据集节点的所有数据集节点,PR(j)表示链接到所述第i数据集节点的第j数据集节点的当前数据价值,Out(j)表示链接到所述第j数据集节点的所有数据集节点,Weight(j,i)表示所述第j数据集节点与所述第i数据集节点之间的边的权重,Weight(j,k)表示所述第j数据集节点与所述第k数据集节点之间的边的权重。

4.根据权利要求1所述的方法,其特征在于,所述数据链接图被构建成有向图,从第i数据集节点到第j数据集节点的交集比例被计算为:

$Pa=C/A$

其中,Pa表示从第i数据集节点到第j数据集节点的交集比例,A表示所述第i数据集节点的数据集大小,C表示所述第i数据集节点与所述第j数据集节点之间的交集大小。

5.根据权利要求4所述的方法,其特征在于,在每一轮迭代过程中,第i数据集节点的数据价值被计算为:

$$PR(i)=\frac{(1-d)}{N}+d\times\sum_{j\in In(i)}\frac{PR(j)\times Weight(j,i)}{\Sigma Weight(j)}$$

其中,PR(i)表示所述第i数据集节点的数据价值,d是大于0且小于1的常数,N表示所述数据链接图中的节点总数,In(i)表示链接到所述第i数据集节点的所有数据集节点,PR(j)表示链接到所述第i数据集节点的第j数据集节点的当前数据价值,Weight(j,i)表示从所述第j数据集节点到所述第i数据集节点的边的权重,ΣWeight(j)表示所述第j数据集节点的所有出链的权重之和。

6.根据权利要求1至5中任一项所述的方法,其特征在于,通过安全求交方式计算所述多个数据集节点中的每两个数据集节点之间的交集大小包括:

获得第一数据集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;

获得第二数据集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数据集节点与所述第二数据集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:

由所述第一数据集节点和所述第二数据集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,由所述第一数据集节点和所述第二数据集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,由所述第一数据集节点和所述第二数据集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:

由所述第一数据集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数据集节点发送所述第二碎片值;

由所述第二数据集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数据集节点发送所述第三碎片值;

由所述第一数据集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;

由所述第二数据集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数据集节点;

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数据集节点;

由所述第一数据集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;

由所述第二数据集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;

由所述第一数据集节点和所述第二数据集节点联合利用所述第一数据集节点处的第

一并行前缀加法器和所述第二数据集节点处的第二并行前缀加法器在所述第一数据集节点处获得第一符号位碎片并且在所述第二数据集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1至5中任一项所述的方法,其特征在于,根据所计算的交集比例来构建所述多个数据集节点的数据链接图包括:

根据目标评估指标来评估所述多个数据集节点的数据质量;以及

将数据质量低于预设值的数据集节点从所述数据链接图中删除。

8.根据权利要求1至5中任一项所述的方法,其特征在于,根据所计算的交集比例来构建所述多个数据集节点的数据链接图包括:

获取所述多个数据集节点中的目标数据集节点的数据源领域偏好;

确定链接到所述目标数据集节点的所有数据集节点中符合所述数据源领域偏好的候选数据集节点;以及

在所述数据链接图中,根据增强因子来调整每个候选数据集节点到所述目标数据集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数据集节点到所述目标数据集节点之间的边的权重包括:将每个候选数据集节点到所述目标数据集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数据集节点到所述目标数据集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。

从多个数据集节点中筛选数据集的方法

技术领域

[0001] 本公开的实施例涉及计算机技术领域,具体地,涉及从多个数据集节点中筛选数据集的方法。

背景技术

[0002] 随着互联网的发展,各类政务主体、行业主体、公司主体、机构主体可经由互联网被关联起来。每个主体可被看作一个节点。在这种场景下要执行安全的高价值数据探查,往往只能通过点对点的质量评估技术来进行单点识别,效率低且不能捕捉相关节点的数据价值影响力信息和共性关系,而且无法发挥网络拓扑能力,难以提供高效的批量化的高价值数据安全筛选方法。

发明内容

[0003] 本文中描述的实施例提供了一种从多个数据集节点中筛选数据集的方法、装置以及存储有计算机程序的计算机可读存储介质。

[0004] 根据本公开的第一方面,提供了一种从多个数据集节点中筛选数据集的方法。该方法包括:通过安全求交方式计算该多个数据集节点中的每两个数据集节点之间的交集大小;根据该多个数据集节点中的每两个数据集节点之间的交集大小来计算该两个数据集节点之间的交集比例;根据所计算的交集比例来构建多个数据集节点的数据链接图,其中,在数据链接图中,每个数据集节点被表示为一个顶点,每两个数据集节点之间的交集比例作为该两个数据集节点之间的边的权重;基于数据链接图来迭代计算每个数据集节点的数据价值;以及根据每个数据集节点的数据价值来从多个数据集节点中选择目标数据集。其中,基于数据链接图来迭代计算每个数据集节点的数据价值包括:将每个数据集节点的数据价值初始化为预设常数值;以及在每一轮迭代过程中遍历数据链接图以将每个数据集节点的数据价值按照边的权重分配给其所链接的数据集节点,直至每个数据集节点的数据价值收敛。

[0005] 在本公开的一些实施例中,数据链接图被构建成无向图。第i数据集节点与第j数据集节点之间的交集比例被计算为:

[0006] $P=(2\times C)/(A+B)$

[0007] 其中,P表示第i数据集节点与第j数据集节点之间的交集比例,A表示第i数据集节点的数据集大小,B表示第j数据集节点的数据集大小,C表示第i数据集节点与第j数据集节点之间的交集大小。

[0008] 在本公开的一些实施例中,在每一轮迭代过程中,第i数据集节点的数据价值被计算为:

[0009] $$PR(i)=(1-d)+d\times\sum_{j\in In(i)}\frac{PR(j)\times Weight(j,i)}{\sum_{k\in Out(j)}Weight(j,k)}$$

[0010] 其中,PR(i)表示第i数据集节点的数据价值,d是大于0且小于1的常数,In(i)表示链接到第i数据集节点的所有数据集节点,PR(j)表示链接到第i数据集节点的第j数据集节点的当前数据价值,Out(j)表示链接到第j数据集节点的所有数据集节点,Weight(j,i)表示第j数据集节点与第i数据集节点之间的边的权重,Weight(j,k)表示第j数据集节点与第k数据集节点之间的边的权重。

[0011] 在本公开的一些实施例中,数据链接图被构建成有向图。从第i数据集节点到第j数据集节点的交集比例被计算为:

[0012] $Pa=C/A$

[0013] 其中,Pa表示从第i数据集节点到第j数据集节点的交集比例,A表示第i数据集节点的数据集大小,C表示第i数据集节点与第j数据集节点之间的交集大小。

[0014] 在本公开的一些实施例中,在每一轮迭代过程中,第i数据集节点的数据价值被计算为:

[0015] $$PR(i)=\frac{(1-d)}{N}+d\times\sum_{j\in In(i)}\frac{PR(j)\times Weight(j,i)}{\Sigma Weight(j)}$$

[0016] 其中,PR(i)表示第i数据集节点的数据价值,d是大于0且小于1的常数,N表示数据链接图中的节点总数,In(i)表示链接到第i数据集节点的所有数据集节点,PR(j)表示链接到第i数据集节点的第j数据集节点的当前数据价值,Weight(j,i)表示从第j数据集节点到第i数据集节点的边的权重,ΣWeight(j)表示第j数据集节点的所有出链的权重之和。

[0017] 在本公开的一些实施例中,通过安全求交方式计算多个数据集节点中的每两个数据集节点之间的交集大小包括:获得第一数据集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;获得第二数据集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数据集节点与第二数据集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:由第一数据集节点和第二数据集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,由第一数据集节点和第二数据集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,由第一数据集节点和第二数据集节点联合确定第一哈希值是否小于第二哈希值包括:由第一数据集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数据集节点发送第二碎片值;由第二数据集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数据集节点发送第三碎片值;由第一数据集节点将第一碎片值减去第三碎片值以获得第五碎片值;由第二数据集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数据集节点;将第二布尔零碎片和第二算术零碎片分配给第二数据集节点;由第一数据集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;由第二数据集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;由第一数据集节点和第二数据集节点联合利用第一数据集节点处的第一并行前缀加法器和第二数据集节点处的第二并行前缀加法器在第一数据集节点处获得第一符号位碎片并且在第二数据集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。

[0018] 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数据集节点的数据链接图包括:根据目标评估指标来评估多个数据集节点的数据质量;以及将数据质量低于预设值的数据集节点从数据链接图中删除。

[0019] 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数据集节点的数据链接图包括:获取多个数据集节点中的目标数据集节点的数据源领域偏好;确定链接到目标数据集节点的所有数据集节点中符合数据源领域偏好的候选数据集节点;以及在数据链接图中,根据增强因子来调整每个候选数据集节点到目标数据集节点之间的边的权重。

[0020] 在本公开的一些实施例中,根据增强因子来调整每个候选数据集节点到目标数据集节点之间的边的权重包括:将每个候选数据集节点到目标数据集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。

[0021] 在本公开的一些实施例中,根据增强因子来调整每个候选数据集节点到目标数据集节点之间的边的权重包括根据下式来调整权重:

[0022] $$w_{t+1} = e^{w_t} \times e^{p}$$

[0023] 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。

[0024] 根据本公开的第二方面,提供了一种从多个数据集节点中筛选数据集的装置。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。

[0025] 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

[0026] 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:

[0027] 图1是数联网的示意性拓扑图;

[0028] 图2是根据本公开的实施例的从多个数据集节点中筛选数据集的方法的示意性流程图;

[0029] 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；

[0030] 图4是图3中的动作311的示意性流程图和信令方案；

[0031] 图5是图4中的动作403的示意性流程图和信令方案；

[0032] 图6是图4中的动作404和405的示意性流程图和信令方案；

[0033] 图7是根据本公开的实施例的数据链接图的一个示例图；

[0034] 图8是根据本公开的实施例的数据链接图的另一个示例图；

[0035] 图9是根据本公开的实施例的数据链接图的又一个示例图；

[0036] 图10是根据本公开的实施例的从多个数据集节点中筛选数据集的装置的示意性框图。

[0037] 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

[0038] 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。

[0039] 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。

[0040] 本公开提出了一种从多个数据集节点中筛选数据集的方法,旨在高效且批量化地实现高价值数据的安全筛选。该多个数据集节点可以分布在数联网中。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。

[0041] 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看作一个数据集节点,那么数联网中的数据集节点的数量可能是非常庞大的。在一些应用场景下,需要在海量的数据集节点中快速筛选出高价值的数据集。

[0042] 根据本公开的实施例的从多个数据集节点中筛选数据集的方法基于安全计算技术实现自动化评估数据集价值,可以在不暴露数据的隐私信息的前提下,完成数联网全域的批量的数据集价值评估。图2示出根据本公开的实施例的从多个数据集节点中筛选数据

集的方法的示意性流程图。

[0043] 在图2的框S202处,通过安全求交方式计算该多个数据集节点中的每两个数据集节点之间的交集大小。安全求交方式指的是不会泄露数据集节点的原始数据信息(隐私信息)的方式并且不会泄露交集和非交集等敏感信息。该多个数据集节点中的任意两个数据集节点之间的交集大小可通过相同的方式来计算。下面以计算第一数据集节点与第二数据集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,第一数据集节点和第二数据集节点是该多个数据集节点中的任意两个不同的数据集节点。在本公开的实施例中,不需要计算数据集节点与其自身的交集大小。

[0044] 在一个示例中,可获得第一数据集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。

[0045] 并行地,可获得第二数据集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。

[0046] 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数据集节点与第二数据集节点之间的交集大小。

[0047] 可通过以下方式来安全地比较第一哈希值与第二哈希值:由第一数据集节点和第二数据集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,则由第一数据集节点和第二数据集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。

[0048] 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。由第一数据集节点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向第二数据集节点P2发送第二碎片值x2。由第二数据集节点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向第一数据集节点P1发送第三碎片值y1。

[0049] 在动作305,由第一数据集节点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,由第二数据集节点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。

[0050] 可由第一数据集节点P1和第二数据集节点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⊕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)。

[0051] 在动作307,将第一布尔零碎片a1和第一算术零碎片b1分配给第一数据集节点P1。在动作308,将第二布尔零碎片a2和第二算术零碎片b2分配给第二数据集节点P2。

[0052] 在动作309,由第一数据集节点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)⊕a1。第一数

据集节点P1还持有第三运算碎片op21,其中,op21＝0。

[0053] 在动作310,由第二数据集节点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。第二数据集节点P2还持有第四运算碎片op12,其中,op12＝a2。

[0054] 在动作311,由第一数据集节点P1和第二数据集节点P2联合利用第一数据集节点P1处的第一并行前缀加法器和第二数据集节点P2处的第二并行前缀加法器在第一数据集节点P1处获得第一符号位碎片B1并且在第二数据集节点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。

[0055] 在由第一数据集节点P1来确定比较结果的示例中,第二数据集节点P2在动作313向第一数据集节点P1发送第二符号位碎片B2。由第一数据集节点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,也可以由第二数据集节点P2来确定比较结果。

[0056] 图4示出图3中的动作311的具体过程。在动作403,由第一数据集节点P1根据第一运算碎片op11和第三运算碎片op21并且由第二数据集节点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。

[0057] 图5示出由第一数据集节点P1和第二数据集节点P2联合执行的与运算的示意性流程图和信令方案。在图5中以第一数据集节点P1拥有第一输入碎片W1和第二输入碎片V1且第二数据集节点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。

[0058] 下面描述图5所示的过程。

[0059] 第一数据集节点P1在动作501获得三元组碎片矩阵<R1,S1,T1>,第二数据集节点P2在动作502获得三元组碎片矩阵<R2,S2,T2>。其中,(R1⊕R2)&(S1⊕S2)＝(T1⊕T2)。

[0060] 第一数据集节点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。第二数据集节点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。

[0061] 第二数据集节点P2在动作505向第一数据集节点P1发送D2和E2。第一数据集节点P1在动作506向第二数据集节点P2发送D1和E1。第一数据集节点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,第二数据集节点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。

[0062] 第一数据集节点P1在动作509计算第一输出碎片O1＝T1⊕(R1&E)⊕(S1&D)⊕(E&D)。第二数据集节点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。

[0063] 回到图4,第一数据集节点P1在动作404根据第五中间碎片p1(p1＝op11⊕op21)对

G1的每一位进行诸位循环计算。第二数据集节点P2在动作405根据第六中间碎片p2(p2＝op12⊕op22)对G2的每一位进行诸位循环计算。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。

[0064] 第一数据集节点P1在动作601对G1执行左移2^i位的操作,即G1＝G1<<2^i。其中,i表示当前循环的索引。第一数据集节点P1还在动作601对p1执行左移2^i位的操作(即p1＝p1<<2^i),然后再将p1更新为p1与kmask异或的结果(即,p1＝p1⊕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。

[0065] 第二数据集节点P2在动作602对G2执行左移2^i位的操作,即G2＝G2<<2^i。第二数据集节点P2还在动作602对p2执行左移2^i位的操作(即p2＝p2<<2^i),然后再将p2更新为p2与kmask异或的结果(即,p2＝p2⊕kmask)。

[0066] 在动作603处,由第一数据集节点P1和第二数据集节点P2联合执行图5所示的与运算。G1相当于第一输入碎片W1,p1相当于第二输入碎片V1,G2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,第一数据集节点P1获得第七中间碎片F1,第二数据集节点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。

[0067] 在动作604处,由第一数据集节点P1和第二数据集节点P2联合执行图5所示的与运算。p1相当于第一输入碎片W1和第二输入碎片V1,p2相当于第三输入碎片W2和第四输入碎片V2。经过动作604的操作,第一数据集节点P1获得更新后的p1,第二数据集节点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601处。更新后的p2会被代入下一循环的动作602处。

[0068] 第一数据集节点P1在动作606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。第二数据集节点P2在动作607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。

[0069] 再次回到图4,第一数据集节点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)。第二数据集节点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)。

[0070] 第一数据集节点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)。第二数据集节点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)。

[0071] 第一数据集节点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。第二数据集节点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。

[0072] 第一数据集节点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。第二数据集节点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。

[0073] 在上述过程中,由于第一数据集节点P1没有获得第二哈希值的完整信息,而第二数据集节点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。

[0074] 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。

[0075] 回到图2,在框S204处,根据该多个数据集节点中的每两个数据集节点之间的交集大小来计算该两个数据集节点之间的交集比例。

[0076] 在本公开的一些实施例中,不考虑两个数据集节点之间的方向性。第i数据集节点与第j数据集节点之间的交集比例被计算为:

[0077] $P = (2 \times C)/(A+B)$ (1)

[0078] 其中,P表示第i数据集节点与第j数据集节点之间的交集比例,A表示第i数据集节点的数据集大小,B表示第j数据集节点的数据集大小,C表示第i数据集节点与第j数据集节点之间的交集大小。在上下文中,第i数据集节点和第j数据集节点表示该多个数据集节点中的任意两个不同的数据集节点。

[0079] 在本公开的另一些实施例中,考虑两个数据集节点之间的方向性。从第i数据集节点到第j数据集节点的交集比例被计算为:

[0080] $Pa = C/A$ (2)

[0081] 其中,Pa表示从第i数据集节点到第j数据集节点的交集比例,A表示第i数据集节点的数据集大小,C表示第i数据集节点与第j数据集节点之间的交集大小。

[0082] 从第j数据集节点到第i数据集节点的交集比例被计算为:

[0083] $Pb = C/B$ (3)

[0084] 其中,Pb表示从第j数据集节点到第i数据集节点的交集比例,B表示第j数据集节点的数据集大小,C表示第i数据集节点与第j数据集节点之间的交集大小。

[0085] 在框S206处,根据所计算的交集比例来构建多个数据集节点的数据链接图。其中,在数据链接图中,每个数据集节点被表示为一个顶点,每两个数据集节点之间的交集比例作为该两个数据集节点之间的边的权重。如果两个数据集节点之间的交集比例为0,则该两个数据集节点不相连。

[0086] 在本公开的一些实施例中,数据链接图可被构建成无向图。任意两个不同的数据集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数据链接图的一个示例图。数据集节点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数据集节点之间的交集比例。

[0087] 在本公开的另一些实施例中,数据链接图可被构建成有向图。任意两个不同的数据集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数据链接图的一个示例图。数据集节点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,从数据集节点5到数据集节点6的边上标注的数字0.34表示从数据集节点5到数据集节点6的交集比例为0.34。从数据集节点6到数据集节点5的边上标注的数字0.78表示从数据集节点6到数据集节点5的交集比例为0.78。

[0088] 在本公开的一些实施例中,在构建了图7或者图8所示的数据链接图之后,可根据目标评估指标来评估数据链接图中的多个数据集节点的数据质量。然后将数据质量低于预设值的数据集节点(可称为“低质量数据集节点”)从数据链接图中删除。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中的低质量数据集节点被删除后的数据链接图的一个示例图。

[0089] 进一步的,本公开的实施例提出可对数据链接图中的边进行基于数据集节点领域相似或者互补的增强。这里的数据链接图可以是在框S206处构建的数据链接图,也可以是删除低质量数据集节点之后的数据链接图。数联网中的数据集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。数联网中的数据集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。

[0090] 在本公开的一些实施例中,可获取多个数据集节点中的目标数据集节点的数据源领域偏好。然后确定链接到目标数据集节点的所有数据集节点中符合数据源领域偏好的候选数据集节点。接着,在数据链接图中,根据增强因子来调整每个候选数据集节点到目标数据集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。

[0091] 在边的权重通过线性方式来增强的实施例中,可将每个候选数据集节点到目标数据集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。

[0092] 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:

[0093] $$w_{t+1} = e^{w_t} \times e^{p} \qquad (4)$$

[0094] 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。

[0095] 回到图2,在框S208处,基于数据链接图来迭代计算每个数据集节点的数据价值。这里的数据链接图可以是在框S206处构建的数据链接图,也可以是删除低质量数据集节点之后的数据链接图,还可以是增强边权重之后的数据链接图。

[0096] 在本公开的一些实施例中,可将每个数据集节点的数据价值初始化为预设常数值。然后在每一轮迭代过程中遍历数据链接图以将每个数据集节点的数据价值按照边的权重分配给其所链接的数据集节点,直至每个数据集节点的数据价值收敛。

[0097] 在数据链接图被构建成无向图的实施例中,在每一轮迭代过程中,第i数据集节点的数据价值被计算为:

[0098] $$PR(i) = (1 - d) + d \times \sum_{j \in In(i)} \frac{PR(j) \times Weight(j,i)}{\sum_{k \in Out(j)} Weight(j,k)} \qquad (5)$$

[0099] 其中,PR(i)表示第i数据集节点的数据价值,d是大于0且小于1的常数,In(i)表示链接到第i数据集节点的所有数据集节点,PR(j)表示链接到第i数据集节点的第j数据集节点的当前数据价值,Out(j)表示链接到第j数据集节点的所有数据集节点,Weight(j,i)表

示第j数据集节点与第i数据集节点之间的边的权重,Weight(j,k)表示第j数据集节点与第k数据集节点之间的边的权重。

[0100] 在数据链接图被构建成有向图的实施例中,在每一轮迭代过程中,第i数据集节点的数据价值被计算为:

[0101] $$PR(i)=\frac{(1-d)}{N}+d\times\sum_{j\in In(i)}\frac{PR(j)\times Weight(j,i)}{\Sigma Weight(j)} \quad (6)$$

[0102] 其中,PR(i)表示第i数据集节点的数据价值,d是大于0且小于1的常数,N表示数据链接图中的节点总数,In(i)表示链接到第i数据集节点的所有数据集节点,PR(j)表示链接到第i数据集节点的第j数据集节点的当前数据价值,Weight(j,i)表示从第j数据集节点到第i数据集节点的边的权重,ΣWeight(j)表示第j数据集节点的所有出链的权重之和。

[0103] 在迭代计算每个数据集节点的数据价值的过程中,如果一个数据集节点由一个高数据价值的数据集节点链接,那么它会得到更多的数据价值。如果一个数据集节点链接到很多数据集节点,那么它传递的数据价值就会分散得更多。数据链接图中的边的权重反应了两个数据集节点之间的链接质量,从而使得链接质量成为影响数据价值的因素。如果一个数据集节点链接到目标数据集节点的权重较高,那么这个链接将在数据价值计算中具有更大的影响力,因此会增加目标数据集节点的数据价值。

[0104] 借助于边的权重,可以更精细地控制数据集节点的重要性,因为链接的权重可以根据各种因素进行调整,如数据来源的权威性、链接的相关性等。这有助于在数联网中更准确地评估数据集的重要性,提高数据集筛选结果的质量。

[0105] 此外,为了使算法更具稳定性和收敛性,在式(5)和(6)中都引入了常数d。d可以被看作一个阻尼因子,有助于各个数据集节点的数据价值的收敛。

[0106] 在框S210处,根据每个数据集节点的数据价值来从多个数据集节点中选择目标数据集。在一个示例中,可对数据集节点的数据价值进行排序,按照排序结果来选择目标数据集。在另一个示例中,可将每个数据集节点的数据价值与数据价值阈值进行比较,从数据价值超过数据价值阈值的数据集节点中选择目标数据集。

[0107] 图10示出根据本公开的实施例的从多个数据集节点中筛选数据集的装置的示意性框图。如图10所示,该装置700可包括处理器710和存储有计算机程序的存储器720。当计算机程序由处理器710执行时,使得装置700可执行如图2所示的方法200的步骤。在一个示例中,装置700可以是计算机设备或云计算节点。装置700可通过安全求交方式计算该多个数据集节点中的每两个数据集节点之间的交集大小。装置700可根据该多个数据集节点中的每两个数据集节点之间的交集大小来计算该两个数据集节点之间的交集比例。装置700可根据所计算的交集比例来构建多个数据集节点的数据链接图。其中,在数据链接图中,每个数据集节点被表示为一个顶点。每两个数据集节点之间的交集比例作为该两个数据集节点之间的边的权重。装置700可基于数据链接图来迭代计算每个数据集节点的数据价值。装置700可根据每个数据集节点的数据价值来从多个数据集节点中选择目标数据集。其中,基于数据链接图来迭代计算每个数据集节点的数据价值包括:将每个数据集节点的数据价值初始化为预设常数值;以及在每一轮迭代过程中遍历数据链接图以将每个数据集节点的数据价值按照边的权重分配给其所链接的数据集节点,直至每个数据集节点的数据价值收

敛。

[0108] 在本公开的实施例中,处理器710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720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。

[0109] 此外,在本公开的实施例中,装置700也可包括输入设备730,例如键盘、鼠标等,用于输入各个数据集节点的输入数据。另外,装置700还可包括输出设备740,例如显示器等,用于输出筛选结果。

[0110] 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。

[0111] 综上所述,根据本公开的实施例的从多个数据集节点中筛选数据集的方法能够实现自动化评估数据集价值,可以在不暴露数据的隐私信息的前提下,完成批量的数据集价值评估。借助于数据链接图中边的权重,可以更精细地控制数据集节点的重要性,提高数据集筛选结果的质量。

[0112] 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。

[0113] 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。

[0114] 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。

[0115] 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。

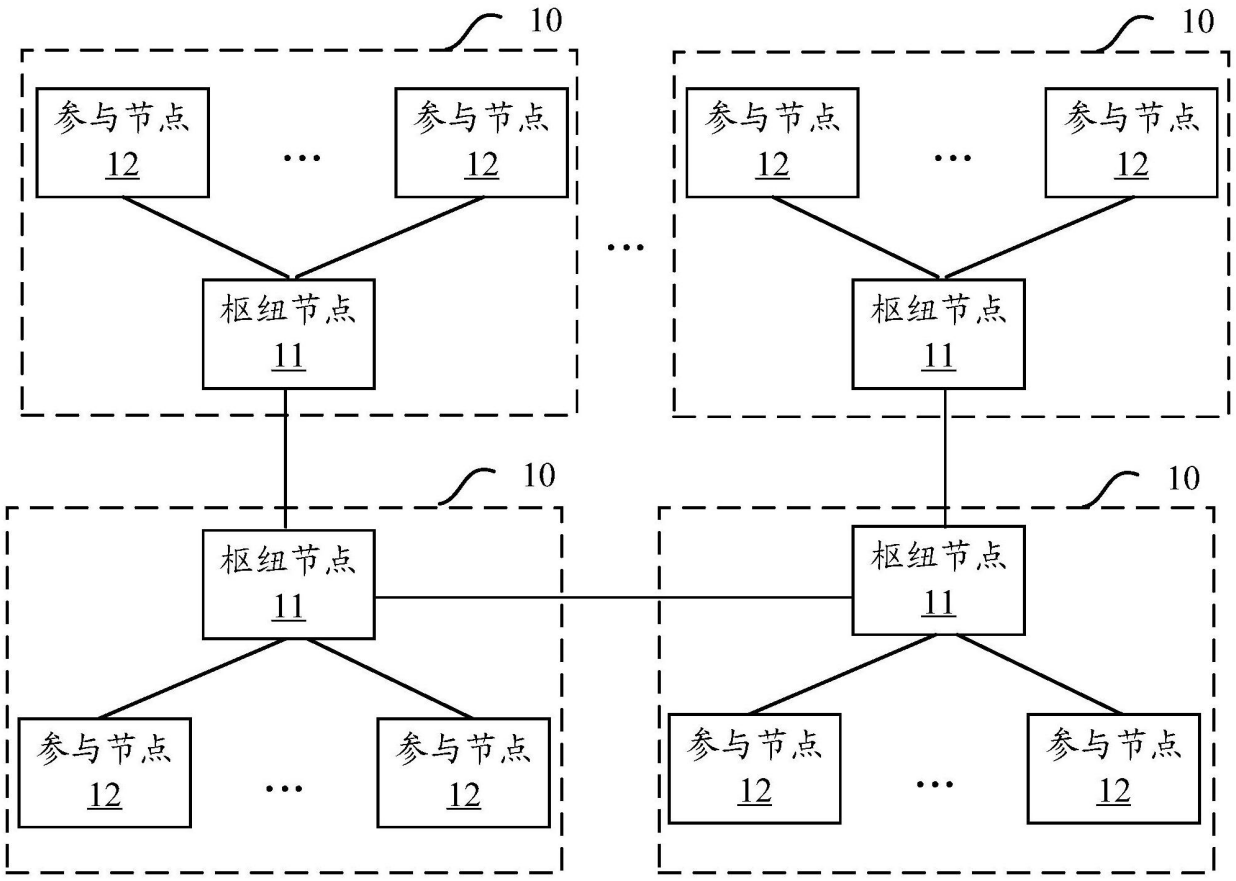

10
参与节点
12
参与节点
12
枢纽节点
11
10
参与节点
12
参与节点
12
枢纽节点
11
10
枢纽节点
11
参与节点
12
参与节点
12
10
枢纽节点
11
参与节点
12
参与节点
12

图1

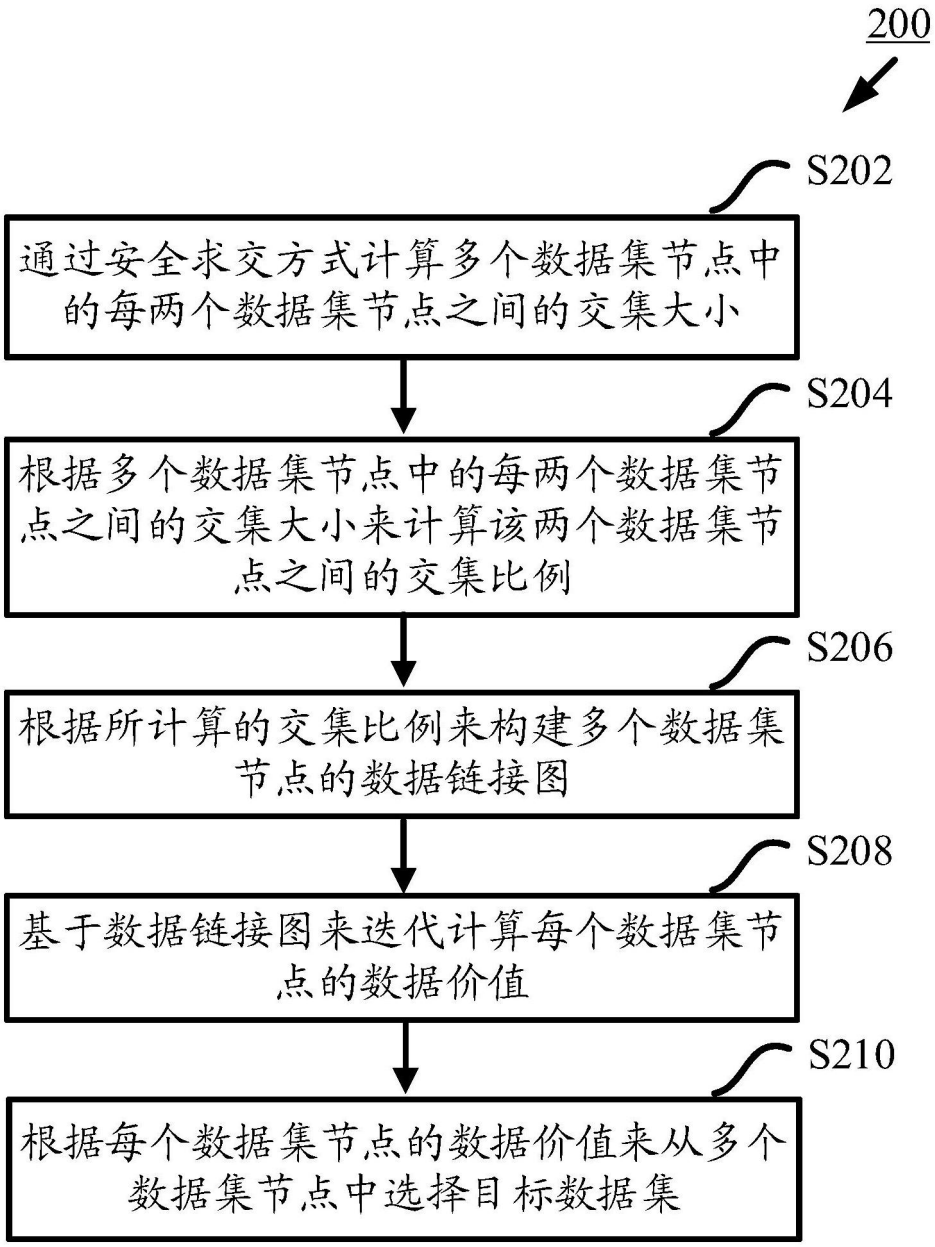

200
S202
通过安全求交方式计算多个数据集节点中的每两个数据集节点之间的交集大小
S204
根据多个数据集节点中的每两个数据集节点之间的交集大小来计算该两个数据集节点之间的交集比例
S206
根据所计算的交集比例来构建多个数据集节点的数据链接图
S208
基于数据链接图来迭代计算每个数据集节点的数据价值
S210
根据每个数据集节点的数据价值来从多个数据集节点中选择目标数据集

图2

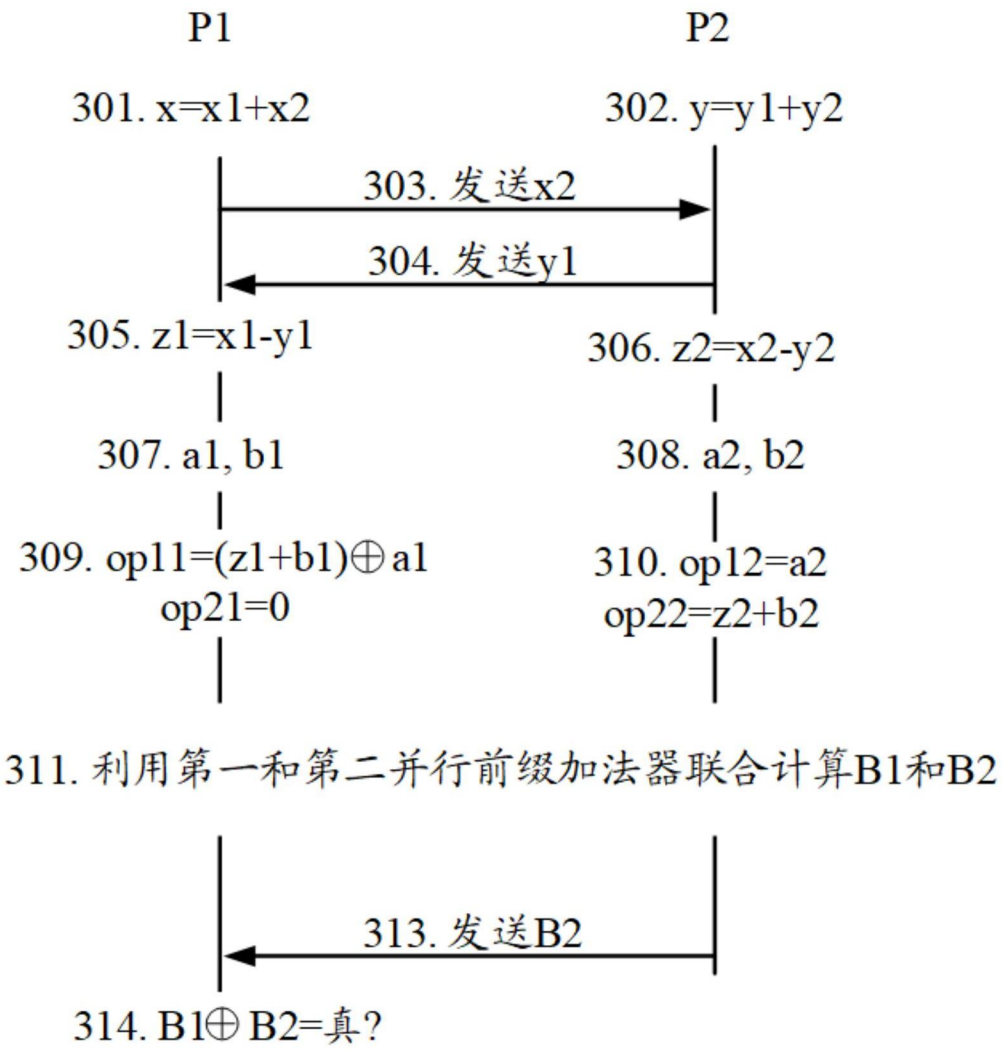

P1
P2
301. x=x1+x2
302. y=y1+y2
303. 发送x2
304. 发送y1
305. z1=x1-y1
306. z2=x2-y2
307. a1, b1
308. a2, b2
309. op11=(z1+b1)⊕a1
op21=0
310. op12=a2
op22=z2+b2
311. 利用第一和第二并行前缀加法器联合计算B1和B2
313. 发送B2
314. B1⊕B2=真?

图3

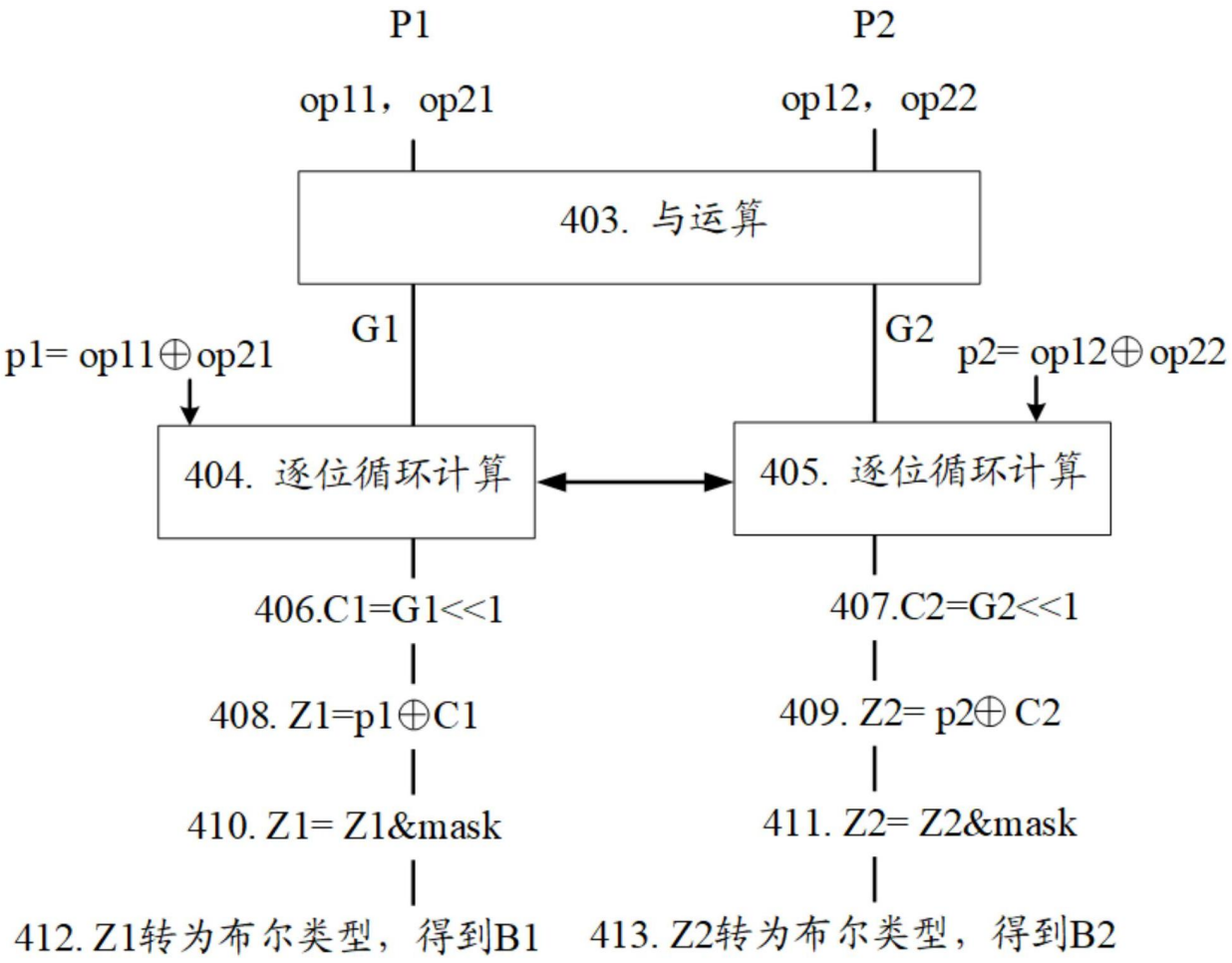

P1
P2
op11，op21
op12，op22
403. 与运算
G1
G2
p1= op11⊕op21
p2= op12⊕op22
404. 逐位循环计算
405. 逐位循环计算
406.C1=G1<<1
407.C2=G2<<1
408. Z1=p1⊕C1
409. Z2= p2⊕ C2
410. Z1= Z1&mask
411. Z2= Z2&mask
412. Z1转为布尔类型，得到B1
413. Z2转为布尔类型，得到B2

图4

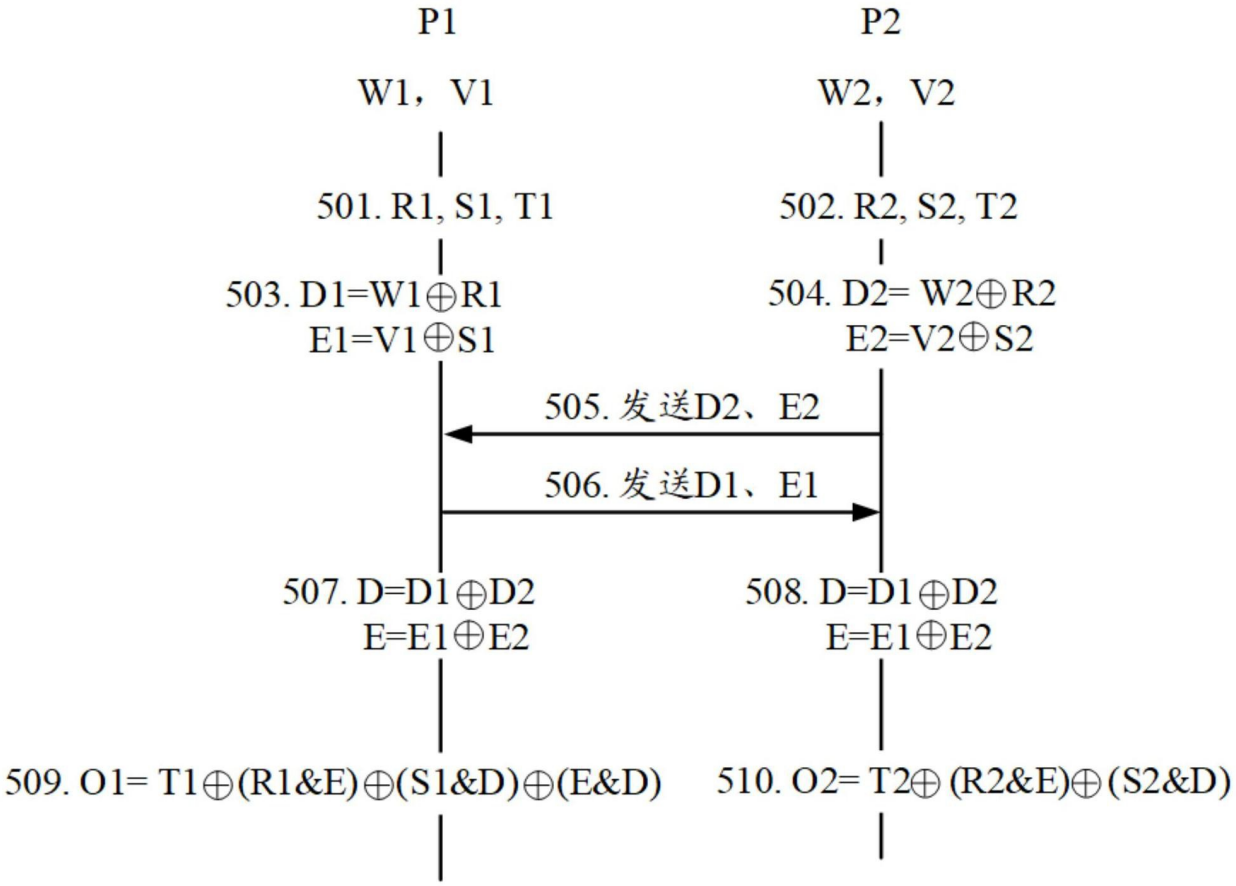

P1
P2
W1，V1
W2，V2
501. R1, S1, T1
502. R2, S2, T2
503. D1=W1⊕R1
E1=V1⊕S1
504. D2= W2⊕R2
E2=V2⊕S2
505. 发送D2、E2
506. 发送D1、E1
507. D=D1⊕D2
E=E1⊕E2
508. D=D1⊕D2
E=E1⊕E2
509. O1= T1⊕(R1&E)⊕(S1&D)⊕(E&D)
510. O2= T2⊕ (R2&E)⊕(S2&D)

图5

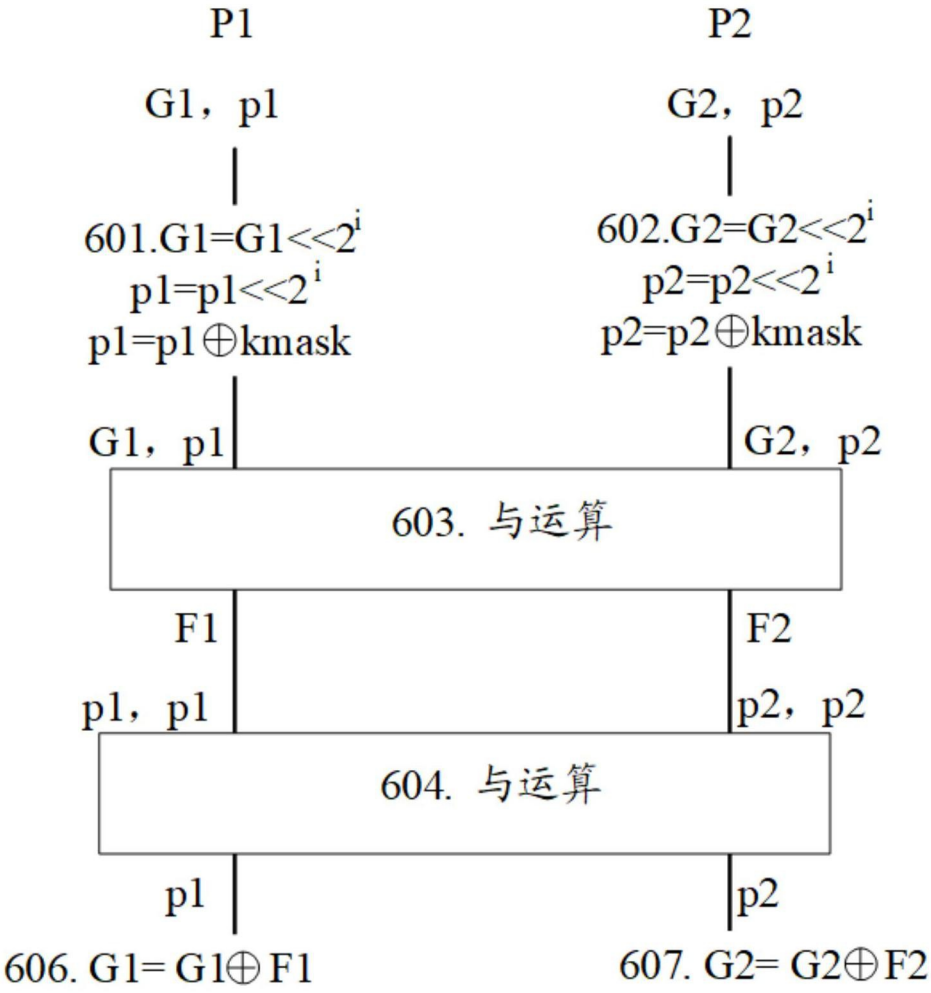

P1
P2
G1，p1
G2，p2
601.G1=G1<<2^i
p1=p1<<2^i
p1=p1⊕kmask
602.G2=G2<<2^i
p2=p2<<2^i
p2=p2⊕kmask
G1，p1
G2，p2
603. 与运算
F1
F2
p1，p1
p2，p2
604. 与运算
p1
p2
606. G1= G1⊕F1
607. G2= G2⊕F2

图6

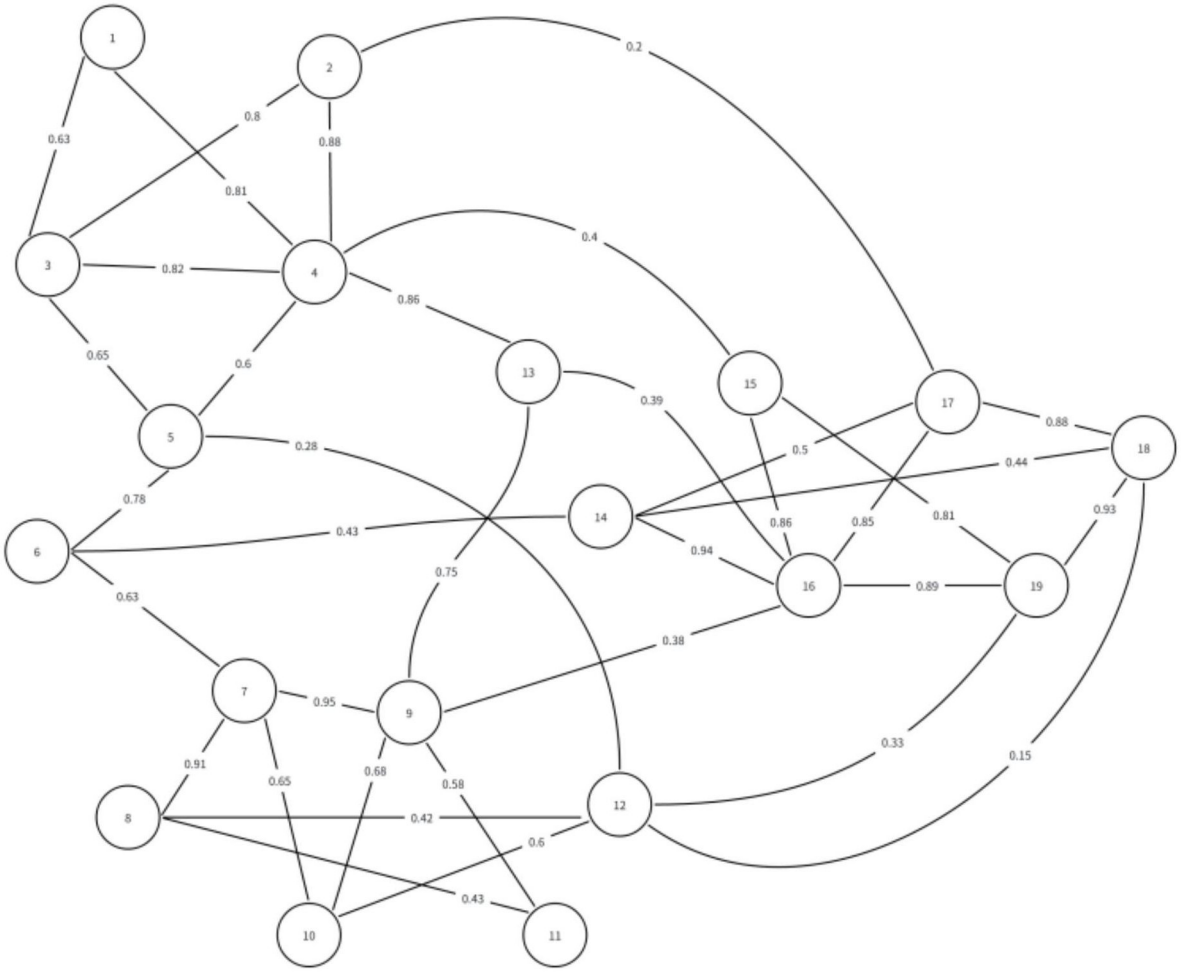

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
0.2
0.8
0.63
0.88
0.81
0.4
0.82
0.86
0.65
0.6
0.39
0.28
0.5
0.88
0.44
0.78
0.43
0.86
0.85
0.81
0.93
0.94
0.75
0.89
0.63
0.38
0.95
0.91
0.65
0.68
0.58
0.33
0.15
0.42
0.6
0.43

图7

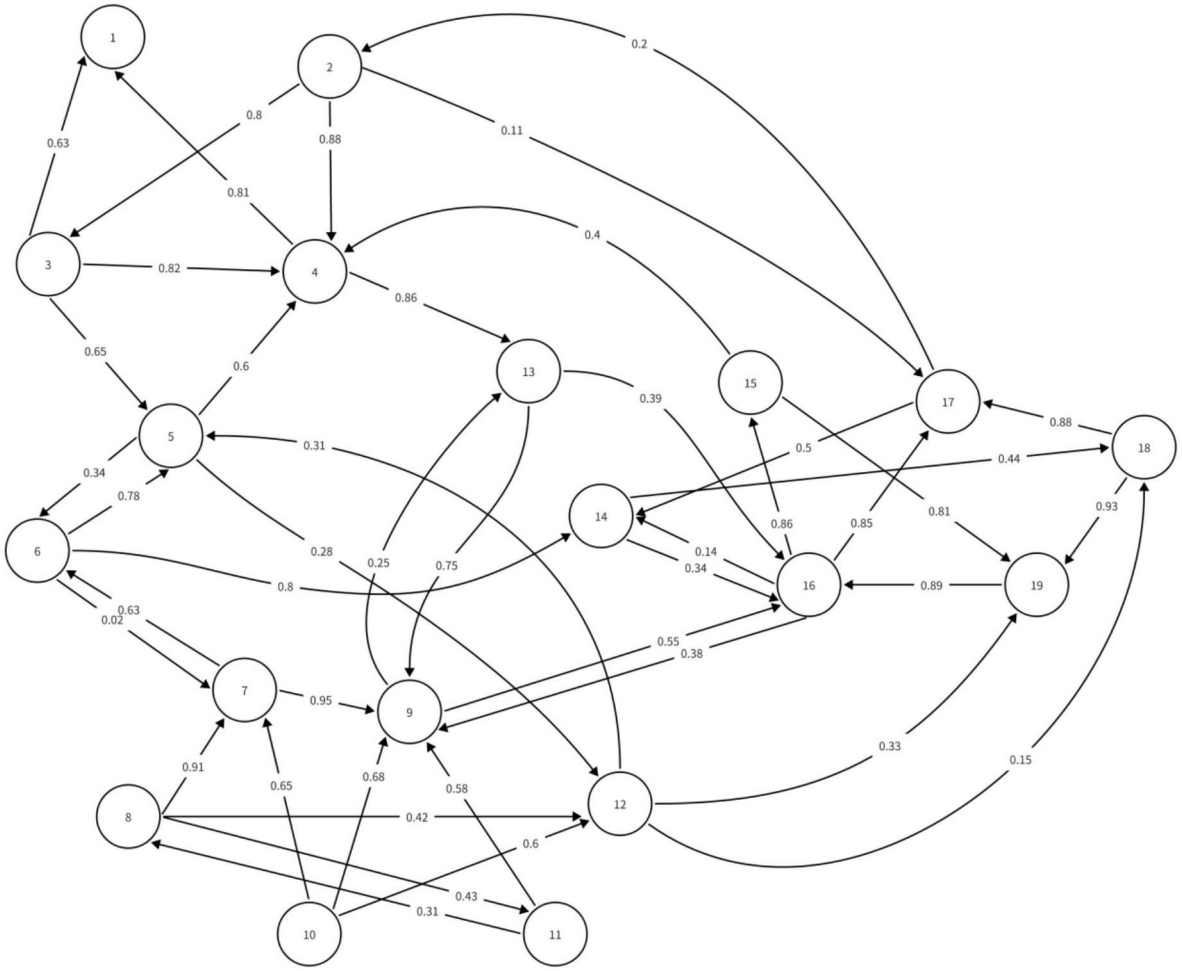

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
0.2
0.8
0.11
0.63
0.88
0.81
0.4
0.82
0.86
0.65
0.6
0.39
0.31
0.34
0.78
0.28
0.25
0.75
0.8
0.14
0.34
0.86
0.85
0.81
0.93
0.5
0.44
0.88
0.89
0.63
0.02
0.55
0.38
0.95
0.91
0.65
0.68
0.58
0.42
0.6
0.43
0.31
0.33
0.15

图8

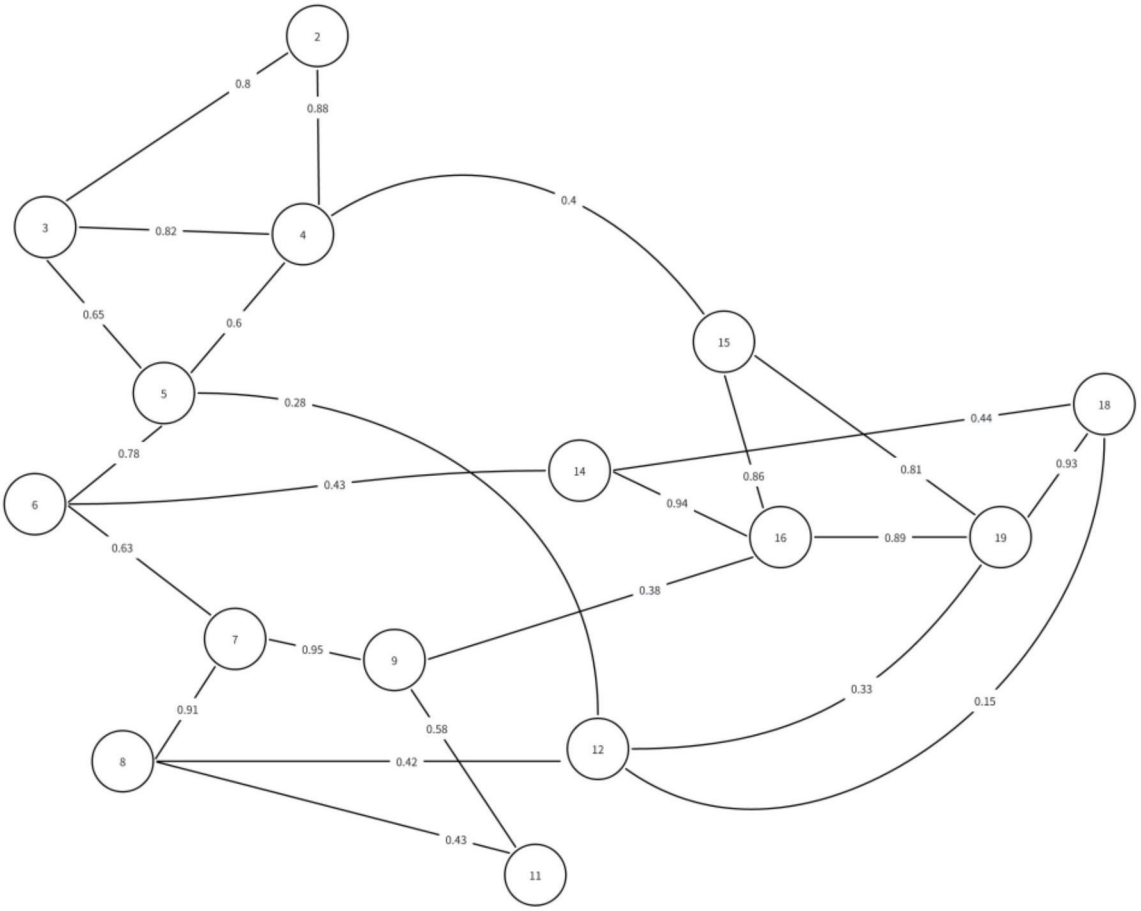

2
3
4
5
6
7
8
9
11
12
14
15
16
18
19
0.8
0.88
0.82
0.4
0.65
0.6
0.28
0.78
0.43
0.44
0.86
0.81
0.93
0.94
0.89
0.63
0.38
0.95
0.91
0.58
0.42
0.33
0.15
0.43

图9

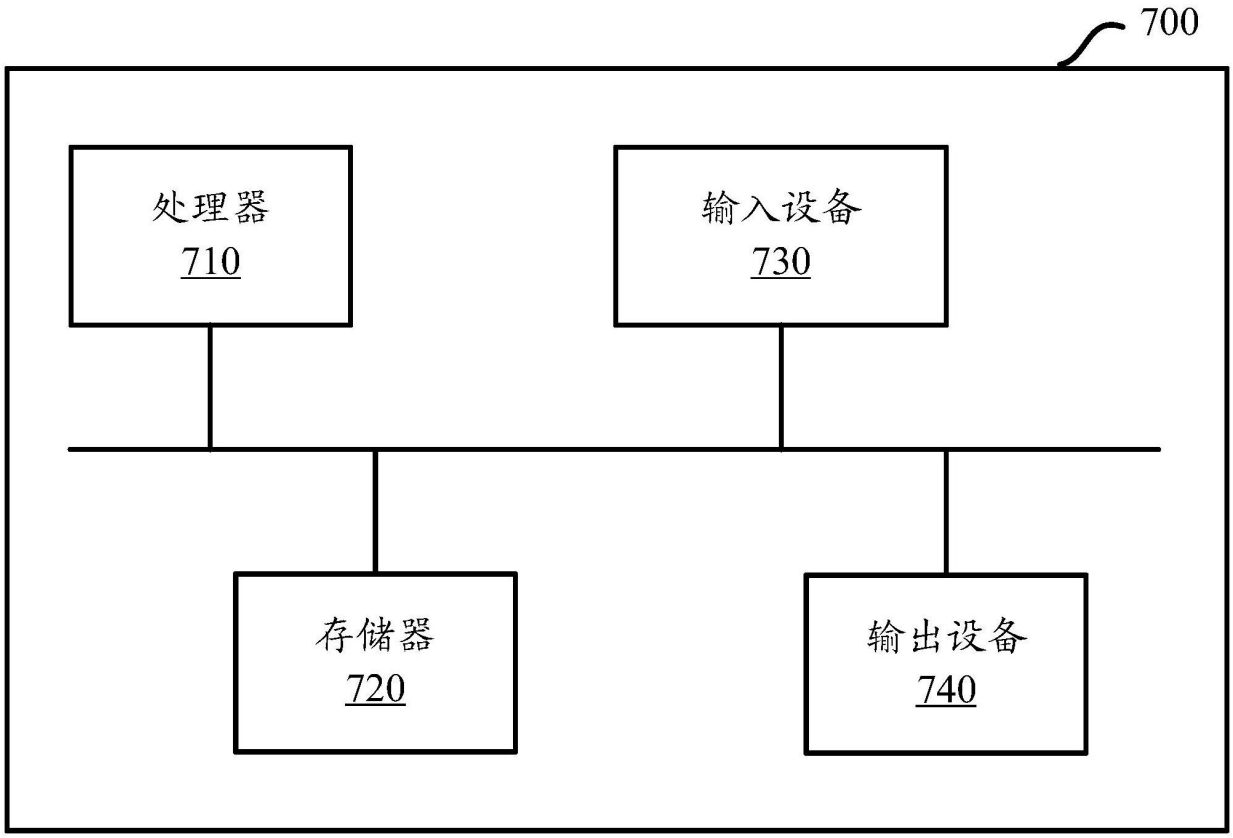


图10