



(12) 发明专利申请

(10) 申请公布号 CN 118014641 A

(43) 申请公布日 2024. 05. 10

(21) 申请号 202410094192.0

(22) 申请日 2024.01.23

(71) 申请人 杭州富算科技有限公司

地址 310051 浙江省杭州市滨江区西兴街
道缤纷街615号4楼401室

(72) 发明人 尤志强 陈立峰 赵东 杜吉锋
蔡晓娟 王兆凯 杨云波 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.

G06Q 30/0203 (2023.01)

G06Q 30/0601 (2023.01)

G06F 16/2458 (2019.01)

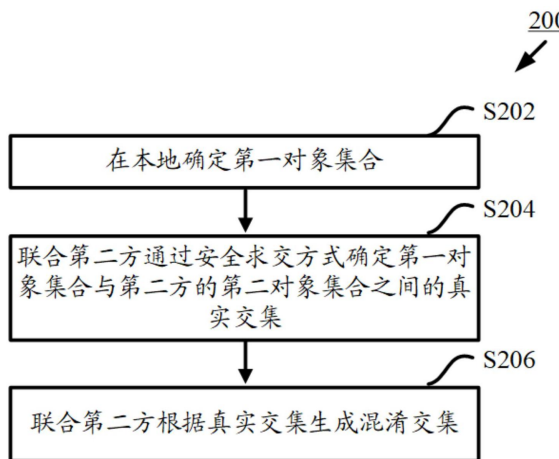
权利要求书5页 说明书16页 附图8页

(54) 发明名称

第一方联合第二方挖掘目标对象的方法及装置

(57) 摘要

本公开的实施例提供一种第一方联合第二方挖掘目标对象的方法及装置。该方法由第一方执行。该方法包括：在本地确定第一对象集合，第一对象集合中的对象是针对目标对象选出的；联合第二方通过安全求交方式确定第一对象集合与第二方的第二对象集合之间的真实交集，第二对象集合中的对象是针对目标对象选出的；以及联合第二方根据真实交集生成混淆交集，混淆交集包括真实交集和混淆对象，混淆对象是第一对象集合中不属于真实交集的对象中的一部分，混淆对象的数量与真实交集的大小之比是随机的，混淆交集集中的对象被确定为目标对象。



1. 一种第一方联合第二方挖掘目标对象的方法,所述方法由所述第一方执行,其特征在于,所述方法包括:

在本地确定第一对象集合,其中,所述第一对象集合中的对象是针对所述目标对象选出的;

联合所述第二方通过安全求交方式确定所述第一对象集合与所述第二方的第二对象集合之间的真实交集,其中,所述第二对象集合中的对象是针对所述目标对象选出的,所述第一对象集合中的每个对象对应一个标记值,所述标记值由第一标记碎片和第二标记碎片之和来确定,所述第一方持有所述第一标记碎片,所述第二方持有所述第二标记碎片,所述真实交集集中的每个对象所对应的标记值等于第一值,所述第一对象集合中不属于所述真实交集的每个对象所对应的标记值不等于所述第一值;以及

联合所述第二方根据所述真实交集生成混淆交集,其中,所述混淆交集包括所述真实交集和混淆对象,所述混淆对象是所述第一对象集合中不属于所述真实交集的对象中的一部分,所述混淆对象的数量与所述真实交集的大小之比是随机的,所述混淆交集集中的对象被确定为所述目标对象。

2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

通过布谷鸟哈希方式将所述第一对象集合中的对象的唯一标识符映射到第一哈希值集合中,其中,所述第一对象集合中的每个对象的唯一标识符对应一个哈希值,所述第一哈希值集合的大小大于所述第一对象集合的大小;

向所述第一哈希值集合中的空白位置填入随机值;

联合所述第二方将所述第一哈希值集合中的每个随机值与不等于所述第一值的标记值相关联,其中,所述标记值由第一标记碎片和第二标记碎片之和来确定,所述第一方持有所述第一标记碎片,所述第二方持有所述第二标记碎片;

联合所述第二方将所述第一哈希值集合中的每个元素的唯一标识符转换成第一编码碎片和第二编码碎片,其中,所述唯一标识符由相应的第一编码碎片和第二编码碎片之和来确定,所述第一方持有所述第一编码碎片,所述第二方持有所述第二编码碎片;以及

联合所述第二方将所述第一哈希值集合中的每个元素与一个缺失标志数相关联,其中,所述缺失标志数由第一缺失标志碎片和第二缺失标志碎片之和来确定,所述第一方持有所述第一缺失标志碎片,所述第二方持有所述第二缺失标志碎片,与所述哈希值相关联的缺失标志数被设置为第二值,与所述随机值相关联的缺失标志数被设置为不等于所述第二值,对应所述第一哈希值集合中的同一元素的第一编码碎片与第一缺失标志碎片相关联,对应所述第一哈希值集合中的同一元素的第二编码碎片与第二缺失标志碎片相关联。

3. 根据权利要求2所述的方法,其特征在于,联合所述第二方根据所述真实交集生成混淆交集包括:

联合所述第二方通过安全求交方式确定所述第一对象集合与所述第二对象集合之间的所述真实交集的大小,其中,所述真实交集的大小由第一交集量级碎片和第二交集量级碎片之和来确定,所述第一方持有所述第一交集量级碎片,所述第二方持有所述第二交集量级碎片;

将所述第一对象集合的大小减去所述第一交集量级碎片以获得第一非交集量级碎片;

与所述第二方联合确定随机比例,其中,所述随机比例由第一比例碎片和第二比例碎

片之和来确定,所述第一方持有所述第一比例碎片,所述第二方持有所述第二比例碎片;

与所述第二方联合将所述真实交集的大小乘以所述随机比例的积除以非交集大小以获得掺杂比例,其中,所述非交集大小等于所述第一非交集量级碎片与第二非交集量级碎片之和,所述第二非交集量级碎片由所述第二方将零减去所述第二交集量级碎片来获得,所述掺杂比例等于第一掺杂比例碎片与第二掺杂比例碎片之和,所述第一方持有所述第一掺杂比例碎片,所述第二方持有所述第二掺杂比例碎片;

与所述第二方联合针对所述第一哈希值集合中的每个元素确定一个参考随机数,其中,所述参考随机数符合均匀分布,所述参考随机数由第一参考随机数碎片和第二参考随机数碎片之和来确定,所述第一方持有所述第一参考随机数碎片,所述第二方持有所述第二参考随机数碎片;

与所述第二方联合比较所述第一哈希值集合中的每个元素所对应的参考随机数是否小于所述掺杂比例;

与所述第二方联合比较所述第一哈希值集合中的每个元素对应的标记值是否等于所述第一值;

与所述第二方联合比较所述第一哈希值集合中的每个元素对应的缺失标志数是否等于所述第二值;以及

根据以下条件中的一个,与所述第二方联合从所述第一对象集合中选出所述目标对象:

所述目标对象对应的标记值等于所述第一值;或者

所述目标对象对应的标记值不等于所述第一值、所述目标对象对应的缺失标志数等于所述第二值并且所述目标对象对应的参考随机数小于所述掺杂比例。

4. 根据权利要求3所述的方法,其特征在于,与所述第二方联合确定随机比例包括:

生成第一随机数,其中,所述第一随机数大于0且小于或者等于0.5;

将所述第一随机数碎片化成第一随机数碎片和第二随机数碎片;

接收来自所述第二方的第三随机数碎片,其中,所述第二方生成第二随机数并将所述第二随机数碎片化成第三随机数碎片和第四随机数碎片,所述第二随机数大于0且小于或者等于0.5;

将所述第一随机数碎片和所述第三随机数碎片相加以获得第一随机数碎片和;

生成第三随机数和第四随机数,其中,所述第三随机数大于或者等于0且小于1,所述第四随机数大于0且小于或者等于1,所述第三随机数小于所述第四随机数;

将所述第一随机数碎片和乘以所述第四随机数与所述第三随机数之差的积加上所述第三随机数以获得所述第一比例碎片;

向所述第二方发送所述第二随机数碎片、所述第三随机数和所述第四随机数,以便所述第二方将所述第二随机数碎片与所述第四随机数碎片之和乘以所述第四随机数与所述第三随机数之差的积加上所述第三随机数来获得所述第二比例碎片。

5. 根据权利要求3所述的方法,其特征在于,与所述第二方联合针对所述第一哈希值集合中的每个元素确定一个参考随机数包括:

生成第一随机数序列和第二随机数序列,其中,所述第一随机数序列和所述第二随机数序列中的随机数的数量等于所述第一哈希值集合的大小,所述第一随机数序列和所述第

二随机数序列中的每个随机数大于或者等于0且小于1；

将所述第一随机数序列碎片化成第一碎片序列和第二碎片序列；

将所述第二随机数序列碎片化成第三碎片序列和第四碎片序列；

接收来自所述第二方的第五碎片序列和第七碎片序列,其中,所述第二方生成第三随机数序列和第四随机数序列,将所述第三随机数序列碎片化成所述第五碎片序列和第六碎片序列,并且将所述第四随机数序列碎片化成所述第七碎片序列和第八碎片序列,所述第三随机数序列和所述第四随机数序列中的随机数数量等于所述第一哈希值集合的大小,所述第三随机数序列和所述第四随机数序列中的每个随机数大于或者等于0且小于1；

向所述第二方发送所述第二碎片序列和所述第四碎片序列；

利用所述第一碎片序列和所述第五碎片序列,与所述第二方联合比较所述第一随机数序列中的每个随机数是否小于所述第三随机数序列中的相应随机数,其中,所述比较结果由第一布尔结果碎片和第二布尔结果碎片进行异或的结果来确定,所述第一方持有所述第一布尔结果碎片,所述第二方持有所述第二布尔结果碎片；

与所述第二方联合将所述第一布尔结果碎片和所述第二布尔结果碎片分别转化为第一算术结果碎片和第二算术结果碎片,其中,所述第一方持有所述第一算术结果碎片,所述第二方持有所述第二算术结果碎片；

根据所述第一算术结果碎片来生成第一乘法因子碎片,其中,所述第一乘法因子碎片中的每个元素等于1与所述第一算术结果碎片中的相应元素之差；

与所述第二方联合计算所述第四随机数序列与乘法因子之积以获得掩膜序列,其中,所述第二方根据所述第二算术结果碎片来生成第二乘法因子碎片,所述第二乘法因子碎片中的每个元素等于0与所述第二算术结果碎片中的相应元素之差,所述乘法因子等于所述第一乘法因子碎片与所述第二乘法因子碎片之和,所述掩膜序列等于第一掩膜碎片序列和第二掩膜碎片序列之和,所述第一方持有所述第一掩膜碎片序列,所述第二方持有所述第二掩膜碎片序列；

利用所述第三碎片序列、所述第一算术结果碎片和所述第一掩膜碎片序列,与所述第二方联合计算所述第二随机数序列乘以所述第一算术结果碎片与所述第二算术结果碎片之和的积加上所述掩膜序列以获得参考随机数序列,其中,所述参考随机数序列等于第一参考随机数碎片序列和第二参考随机数碎片序列之和,所述第一方持有所述第一参考随机数碎片序列,所述第二方持有所述第二参考随机数碎片序列,所述第一参考随机数碎片序列包括针对每个第一数据碎片的第一参考随机数碎片,所述第二参考随机数碎片序列包括针对每个第一数据碎片的第二参考随机数碎片。

6. 根据权利要求5所述的方法,其特征在于,与所述第二方联合比较所述第一随机数序列中的每个随机数是否小于所述第三随机数序列中的相应随机数包括:

将所述第一碎片序列减去所述第五碎片序列以获得第九碎片序列；

获得第一布尔零碎片序列和第一算术零碎片序列,其中,所述第一布尔零碎片序列中的每个元素与第二布尔零碎片序列中的相应元素异或的结果为0,所述第一算术零碎片序列中的每个元素与第二算术零碎片序列中的相应元素相加的结果为0,所述第二方拥有所述第二布尔零碎片序列和所述第二算术零碎片序列；

计算所述第九碎片序列与所述第一算术零碎片序列之和与所述第一布尔零碎片序列

异或的结果,以获得第一运算碎片序列;

与所述第二方联合利用所述第一方处的第一并行前缀加法器和所述第二方处的第二并行前缀加法器在所述第一方处获得第一符号位碎片序列并且在所述第二方处获得第二符号位碎片序列,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序列和第三运算碎片序列,所述第二并行前缀加法器的输入为第二运算碎片序列和第四运算碎片序列,所述第二运算碎片序列由所述第二方计算第十碎片序列与所述第二算术零碎片序列之和来获得,所述第十碎片序列由所述第二方将所述第二碎片序列减去所述第六碎片序列来获得,所述第三运算碎片序列中的每个元素等于0,所述第四运算碎片序列等于所述第二布尔零碎片序列;

接收来自所述第二方的所述第二符号位碎片序列;

对所述第一符号位碎片序列与所述第二符号位碎片序列执行异或操作以获得比较值序列;

响应于所述比较值序列中的第一比较值为真,确定所述第一随机数序列中与所述第一比较值相对应的随机数小于所述第三随机数序列中与所述第一比较值相对应的随机数;以及

响应于所述比较值序列中的所述第一比较值不为真,确定所述第一随机数序列中与所述第一比较值相对应的随机数不小于所述第三随机数序列中与所述第一比较值相对应的随机数。

7. 根据权利要求3所述的方法,其特征在于,与所述第二方联合比较所述第一哈希值集合中的每个元素所对应的参考随机数是否小于所述掺杂比例包括:

将所述第一参考随机数碎片减去所述第一比例碎片以获得第一比较碎片值;

获得第一布尔零碎片和第一算术零碎片,其中,所述第一布尔零碎片与第二布尔零碎片异或的结果为0,所述第一算术零碎片与第二算术零碎片相加的结果为0,所述第二方拥有所述第二布尔零碎片和所述第二算术零碎片;

计算所述第一比较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;

由所述第一方和所述第二方联合利用所述第一方处的第一并行前缀加法器和所述第二方处的第二并行前缀加法器在所述第一方处获得第一符号位碎片并且在所述第二方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第二运算碎片和第四运算碎片,所述第二运算碎片由所述第二方计算第二比较碎片值与所述第二算术零碎片之和来获得,所述第二比较碎片值由所述第二方将所述第二参考随机数碎片减去所述第二比例碎片来获得,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;

接收来自所述第二方的所述第二符号位碎片;

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第二比较值;

响应于所述第二比较值为真,确定所述参考随机数小于所述掺杂比例;以及

响应于所述第二比较值不为真,确定所述参考随机数不小于所述掺杂比例。

8. 根据权利要求2至7中任一项所述的方法,其特征在于,所述方法还包括:

获得所述第二对象集合的大小;

比较所述第一对象集合的大小与所述第二对象集合的大小；
将所述第一对象集合和所述第二对象集合中较大的对象集合的大小作为参考大小；以
及

将所述第一哈希值集合的大小设置为等于所述参考大小的 α 倍,其中, α 大于1。

9.一种第一方联合第二方挖掘目标对象的装置,所述装置位于所述第一方处,其特征在于,所述装置包括:

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;

其中,当所述计算机程序由所述至少一个处理器执行时,使得所述装置执行根据权利要求1至8中任一项所述的方法的步骤。

10.一种存储有计算机程序的计算机可读存储介质,其特征在于,所述计算机程序在由处理器执行时实现根据权利要求1至8中任一项所述的方法的步骤。

第一方联合第二方挖掘目标对象的方法及装置

技术领域

[0001] 本公开的实施例涉及数据处理技术领域,具体地,涉及第一方联合第二方挖掘目标对象的方法及装置。

背景技术

[0002] 在大多数系统中都会存在长尾问题。比如在电商平台,纸巾或者袜子购买量非常大,而高端相机、特色地方风味的美食、登山装备等商品销量较小。这是因为此类商品的受众比较小,只有一些地域特性强的或者是相应领域的发烧友,才会去购买小众商品。但这类小众商品相对于流行商品,往往价值更高。发掘这类小众商品及其潜在用户,具有显著的业务价值。

[0003] 在数据交易的场景中,同样会存在长尾问题。有些数据集单价低,交易量大。有一些特殊的数据集,由于私域性、独占性、用户特性强,使得这类小众数据的价值会很高,在某些场景下具有很高的业务价值。比如在一些应用场景下对高净值用户进行偏好识别,可以用于进一步挖掘扩展新的高价值用户,或者对存量高净值用户做活跃度提升营销。这些对于银行、电商、保险、证券等行业在新的价值用户增长方面的需求极具吸引力。

[0004] 因此,挖掘长尾用户群体会带来创新和应用的机会。通过深入挖掘长尾数据,可以发现隐藏在稀有事件中的有价值信息,从而可以实现更个性化、多样化的推荐服务、拉新业务、存量激活业务,提高用户满意度。而关于高价值的长尾数据集,各机构自身往往会非常看重,不希望暴露任何敏感信息,包括具体的交集个体信息。因此,希望能通过群体信息来掩盖个体信息,符合相关数据安全法的合规性。

[0005] 现阶段用于挖掘长尾用户群体的方案在合规性与功能性层面还无法做到同时兼顾。

发明内容

[0006] 本文中描述的实施例提供了一种第一方联合第二方挖掘目标对象的方法、装置以及存储有计算机程序的计算机可读存储介质。

[0007] 根据本公开的第一方面,提供了一种第一方联合第二方挖掘目标对象的方法。该方法由第一方执行。该方法包括:在本地确定第一对象集合,其中,第一对象集合中的对象是针对目标对象选出的;联合第二方通过安全求交方式确定第一对象集合与第二方的第二对象集合之间的真实交集,其中,第二对象集合中的对象是针对目标对象选出的,第一对象集合中的每个对象对应一个标记值,标记值由第一标记碎片和第二标记碎片之和来确定,第一方持有第一标记碎片,第二方持有第二标记碎片,真实交集中的每个对象所对应的标记值等于第一值,第一对象集合中不属于真实交集的每个对象所对应的标记值不等于第一值;以及联合第二方根据真实交集生成混淆交集,其中,混淆交集包括真实交集和混淆对象,混淆对象是第一对象集合中不属于真实交集的对象中的一部分,混淆对象的数量与真实交集的大小之比是随机的,混淆交集集中的对象被确定为目标对象。

[0008] 在本公开的一些实施例中,方法还包括:通过布谷鸟哈希方式将第一对象集合中的对象的唯一标识符映射到第一哈希值集合中,其中,第一对象集合中的每个对象的唯一标识符对应一个哈希值,第一哈希值集合的大小大于第一对象集合的大小;向第一哈希值集合中的空白位置填入随机值;联合第二方将第一哈希值集合中的每个随机值与不等于第一值的标记值相关联,其中,标记值由第一标记碎片和第二标记碎片之和来确定,第一方持有第一标记碎片,第二方持有第二标记碎片;联合第二方将第一哈希值集合中的每个元素的唯一标识符转换成第一编码碎片和第二编码碎片,其中,唯一标识符由相应的第一编码碎片和第二编码碎片之和来确定,第一方持有第一编码碎片,第二方持有第二编码碎片;以及联合第二方将第一哈希值集合中的每个元素与一个缺失标志数相关联,其中,缺失标志数由第一缺失标志碎片和第二缺失标志碎片之和来确定,第一方持有第一缺失标志碎片,第二方持有第二缺失标志碎片,与哈希值相关联的缺失标志数被设置为第二值,与随机值相关联的缺失标志数被设置为不等于第二值,对应第一哈希值集合中的同一元素的第一编码碎片与第一缺失标志碎片相关联,对应第一哈希值集合中的同一元素的第二编码碎片与第二缺失标志碎片相关联。

[0009] 在本公开的一些实施例中,联合第二方根据真实交集生成混淆交集包括:联合第二方通过安全求交方式确定第一对象集合与第二对象集合之间的真实交集的大小,其中,真实交集的大小由第一交集量级碎片和第二交集量级碎片之和来确定,第一方持有第一交集量级碎片,第二方持有第二交集量级碎片;将第一对象集合的大小减去第一交集量级碎片以获得第一非交集量级碎片;与第二方联合确定随机比例,其中,随机比例由第一比例碎片和第二比例碎片之和来确定,第一方持有第一比例碎片,第二方持有第二比例碎片;与第二方联合将真实交集的大小乘以随机比例的积除以非交集大小以获得掺杂比例,其中,非交集大小等于第一非交集量级碎片与第二非交集量级碎片之和,第二非交集量级碎片由第二方将零减去第二交集量级碎片来获得,掺杂比例等于第一掺杂比例碎片与第二掺杂比例碎片之和,第一方持有第一掺杂比例碎片,第二方持有第二掺杂比例碎片;与第二方联合针对第一哈希值集合中的每个元素确定一个参考随机数,其中,参考随机数符合均匀分布,参考随机数由第一参考随机数碎片和第二参考随机数碎片之和来确定,第一方持有第一参考随机数碎片,第二方持有第二参考随机数碎片;与第二方联合比较第一哈希值集合中的每个元素所对应的参考随机数是否小于掺杂比例;与第二方联合比较第一哈希值集合中的每个元素对应的标记值是否等于第一值;与第二方联合比较第一哈希值集合中的每个元素对应的缺失标志数是否等于第二值;以及根据以下条件中的一个,与第二方联合从第一对象集合中选出目标对象:目标对象对应的标记值等于第一值;或者目标对象对应的标记值不等于第一值、目标对象对应的缺失标志数等于第二值并且目标对象对应的参考随机数小于掺杂比例。

[0010] 在本公开的一些实施例中,与第二方联合确定随机比例包括:生成第一随机数,其中,第一随机数大于0且小于或者等于0.5;将第一随机数碎片化成第一随机数碎片和第二随机数碎片;接收来自第二方的第三随机数碎片,其中,第二方生成第二随机数并将第二随机数碎片化成第三随机数碎片和第四随机数碎片,第二随机数大于0且小于或者等于0.5;将第一随机数碎片和第三随机数碎片相加以获得第一随机数碎片和;生成第三随机数和第四随机数,其中,第三随机数大于或者等于0且小于1,第四随机数大于0且小于或者等于1,

第三随机数小于第四随机数;将第一随机数碎片和乘以第四随机数与第三随机数之差的积加上第三随机数以获得第一比例碎片;向第二方发送第二随机数碎片、第三随机数和第四随机数,以便第二方将第二随机数碎片与第四随机数碎片之和乘以第四随机数与第三随机数之差的积加上第三随机数来获得第二比例碎片。

[0011] 在本公开的一些实施例中,与第二方联合针对第一哈希值集合中的每个元素确定一个参考随机数包括:生成第一随机数序列和第二随机数序列,其中,第一随机数序列和第二随机数序列中的随机数的数量等于第一哈希值集合的大小,第一随机数序列和第二随机数序列中的每个随机数大于或者等于0且小于1;将第一随机数序列碎片化成第一碎片序列和第二碎片序列;将第二随机数序列碎片化成第三碎片序列和第四碎片序列;接收来自第二方的第五碎片序列和第七碎片序列,其中,第二方生成第三随机数序列和第四随机数序列,将第三随机数序列碎片化成第五碎片序列和第六碎片序列,并且将第四随机数序列碎片化成第七碎片序列和第八碎片序列,第三随机数序列和第四随机数序列中的随机数数量等于第一哈希值集合的大小,第三随机数序列和第四随机数序列中的每个随机数大于或者等于0且小于1;向第二方发送第二碎片序列和第四碎片序列;利用第一碎片序列和第五碎片序列,与第二方联合比较第一随机数序列中的每个随机数是否小于第三随机数序列中的相应随机数,其中,比较结果由第一布尔结果碎片和第二布尔结果碎片进行异或的结果来确定,第一方持有第一布尔结果碎片,第二方持有第二布尔结果碎片;与第二方联合将第一布尔结果碎片和第二布尔结果碎片分别转化为第一算术结果碎片和第二算术结果碎片,其中,第一方持有第一算术结果碎片,第二方持有第二算术结果碎片;根据第一算术结果碎片来生成第一乘法因子碎片,其中,第一乘法因子碎片中的每个元素等于1与第一算术结果碎片中的相应元素之差;与第二方联合计算第四随机数序列与乘法因子之积以获得掩膜序列,其中,第二方根据第二算术结果碎片来生成第二乘法因子碎片,第二乘法因子碎片中的每个元素等于0与第二算术结果碎片中的相应元素之差,乘法因子等于第一乘法因子碎片与第二乘法因子碎片之和,掩膜序列等于第一掩膜碎片序列和第二掩膜碎片序列之和,第一方持有第一掩膜碎片序列,第二方持有第二掩膜碎片序列;利用第三碎片序列、第一算术结果碎片和第一掩膜碎片序列,与第二方联合计算第二随机数序列乘以第一算术结果碎片与第二算术结果碎片之和的积加上掩膜序列以获得参考随机数序列,其中,参考随机数序列等于第一参考随机数碎片序列和第二参考随机数碎片序列之和,第一方持有第一参考随机数碎片序列,第二方持有第二参考随机数碎片序列,第一参考随机数碎片序列包括针对每个第一数据碎片的第一参考随机数碎片,第二参考随机数碎片序列包括针对每个第一数据碎片的第二参考随机数碎片。

[0012] 在本公开的一些实施例中,与第二方联合比较第一随机数序列中的每个随机数是否小于第三随机数序列中的相应随机数包括:将第一碎片序列减去第五碎片序列以获得第九碎片序列;获得第一布尔零碎片序列和第一算术零碎片序列,其中,第一布尔零碎片序列中的每个元素与第二布尔零碎片序列中的相应元素异或的结果为0,第一算术零碎片序列中的每个元素与第二算术零碎片序列中的相应元素相加的结果为0,第二方拥有第二布尔零碎片序列和第二算术零碎片序列;计算第九碎片序列与第一算术零碎片序列之和与第一布尔零碎片序列异或的结果,以获得第一运算碎片序列;与第二方联合利用第一方处的第一并行前缀加法器和第二方处的第二并行前缀加法器在第一方处获得第一符号位碎片序

列并且在第二方处获得第二符号位碎片序列,其中,第一并行前缀加法器的输入为第一运算碎片序列和第三运算碎片序列,第二并行前缀加法器的输入为第二运算碎片序列和第四运算碎片序列,第二运算碎片序列由第二方计算第十碎片序列与第二算术零碎片序列之和来获得,第十碎片序列由第二方将第二碎片序列减去第六碎片序列来获得,第三运算碎片序列中的每个元素等于0,第四运算碎片序列等于第二布尔零碎片序列;接收来自第二方的第二符号位碎片序列;对第一符号位碎片序列与第二符号位碎片序列执行异或操作以获得比较值序列;响应于比较值序列中的第一比较值为真,确定第一随机数序列中与第一比较值相对应的随机数小于第三随机数序列中与第一比较值相对应的随机数;以及响应于比较值序列中的第一比较值不为真,确定第一随机数序列中与第一比较值相对应的随机数不小于第三随机数序列中与第一比较值相对应的随机数。

[0013] 在本公开的一些实施例中,与第二方联合比较第一哈希值集合中的每个元素所对应的参考随机数是否小于掺杂比例包括:将第一参考随机数碎片减去第一比例碎片以获得第一比较碎片值;获得第一布尔零碎片和第一算术零碎片,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,第二方拥有第二布尔零碎片和第二算术零碎片;计算第一比较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;由第一方和第二方联合利用第一方处的第一并行前缀加法器和第二方处的第二并行前缀加法器在第一方处获得第一符号位碎片并且在第二方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,第二运算碎片由第二方计算第二比较碎片值与第二算术零碎片之和来获得,第二比较碎片值由第二方将第二参考随机数碎片减去第二比例碎片来获得,第三运算碎片等于0,第四运算碎片等于第二布尔零碎片;接收来自第二方的第二符号位碎片;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第二比较值;响应于第二比较值为真,确定参考随机数小于掺杂比例;以及响应于第二比较值不为真,确定参考随机数不小于掺杂比例。

[0014] 在本公开的一些实施例中,方法还包括:获得第二对象集合的大小;比较第一对象集合的大小与第二对象集合的大小;将第一对象集合和第二对象集合中较大的对象集合的大小作为参考大小;以及将第一哈希值集合的大小设置为等于参考大小的 α 倍,其中, α 大于1。

[0015] 根据本公开的第二方面,提供了一种第一方联合第二方挖掘目标对象的装置。该装置位于第一方处。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。

[0016] 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

[0017] 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:

- [0018] 图1是数联网的示意性拓扑图；
- [0019] 图2是根据本公开的实施例的第一方联合第三方挖掘目标对象的方法的示意性流程图；
- [0020] 图3是根据本公开的实施例的第一方联合第三方通过安全求交方式确定第一对象集合与第二方的第二对象集合之间的真实交集的步骤的示意性流程图和信令方案；
- [0021] 图4是可编程的不经意伪随机函数 (OPPRF) 的执行步骤的示意性流程图和信令方案；
- [0022] 图5是根据本公开的实施例的第一方联合第三方根据真实交集生成混淆交集的步骤的示意性流程图和信令方案；
- [0023] 图6是根据本公开的实施例的确定随机比例的步骤的示意性流程图和信令方案；
- [0024] 图7是根据本公开的实施例的针对第一哈希值集中的每个元素确定一个参考随机数的步骤的示意性流程图和信令方案；
- [0025] 图8是根据本公开的实施例的确定第一值是否小于第二值的步骤的示意性流程图和信令方案；
- [0026] 图9是图8中的动作811的示意性流程图和信令方案；
- [0027] 图10是图9中的动作903的示意性流程图和信令方案；
- [0028] 图11是图9中的动作904和905的示意性流程图和信令方案；
- [0029] 图12是根据本公开的实施例的第一方联合第三方挖掘目标对象的装置的示意性框图。
- [0030] 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

[0031] 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。

[0032] 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。

[0033] 如上所述,在大多数系统中都会存在长尾问题。在数据交易的场景中,同样会存在长尾问题。如果能够安全地挖掘长尾用户群体,则有助于提高对高价值数据的利用率。特别地,如果在涉及大量数据交易的数联网中能够安全地挖掘长尾用户群体,则有助于促进数联网的数据交易和健康发展。

[0034] 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。

[0035] 在实践中,参与节点12之间可能需要联合挖掘目标对象。联合挖掘目标对象的双方在上下文中被称为“第一方”和“第二方”。目标对象是第一方中的对象。对象可以是用户、数据,或者其它对象。

[0036] 在挖掘目标对象的过程中,不应该定位到个体信息,否则会涉及潜在的隐私风险且不满足法规要求。为了确保数据的隐私与安全,本公开提出引入群体信息,以避免准确定位到具体个体。这样不仅可以符合合规性要求,还有助于维护用户的隐私权和数据安全。通过将焦点从个体信息转移到群体信息上,能够在数据分析和处理的过程中采用更为抽象和综合的视角。这样不仅有助于避免潜在的隐私泄露,还能够为数据分析提供更广泛、更全面的洞见。群体信息的引入有助于消除数据中的个体特征,使得任何尝试追踪或识别单个个体的企图都变得异常困难。

[0037] 这种策略的优势在于在保护个体隐私的同时,仍然能够提供有价值的群体趋势和模式。通过对群体信息的聚焦,我们可以更好地理解广泛的目标用户行为趋势,而无需涉及具体的个体身份。这不仅有助于遵循数据保护法规,同时也为数据分析提供了一种更安全、更负责任的方法。在当前数据保护意识日益增强的环境下,采用这样的方法能够确保数据的合规性和隐私性。

[0038] 基于上述思考,本公开提出一种第一方联合第二方挖掘目标对象的方法。图2示出根据本公开的实施例的第一方联合第二方挖掘目标对象的方法的示意性流程图。该方法由第一方执行。第一方可以是挖掘任务的发起方,也可以是挖掘任务的合作方。目标对象是第一方中的对象。对象可以是用户、数据,或者其它对象。

[0039] 在图2的框S202处,第一方在本地确定第一对象集合。第一对象集合中的对象是针对目标对象选出的。并行地,第二方在本地确定第二对象集合。第二对象集合中的对象也是针对目标对象选出的。例如,目标对象是具有高购买力且爱好滑雪的客群,那么第一对象集合可以是第一方中的具有高购买力的对象的集合,第二对象集合可以是第二方中的爱好滑雪的对象的集合。可替代地,第一对象集合可以是第一方中的爱好滑雪的对象的集合,第二对象集合可以是第二方中的具有高购买力的对象的集合。也就是说,第一对象集合与第二对象集合之间的交集能够满足对目标对象的要求。

[0040] 在图2的框S204处,第一方联合第二方通过安全求交方式确定第一对象集合与第二方的第二对象集合之间的真实交集。图3示出根据本公开的实施例的第一方联合第二方通过安全求交方式确定第一对象集合与第二方的第二对象集合之间的真实交集的步骤的示意性流程图和信令方案。

[0041] 如图3所示,第一对象集合中的对象的唯一标识符(ID)用ID_A来表示。ID_A是向量值,ID_A中的每个元素对应一个对象的ID。第一对象集合的大小为N。第二对象集合中的对象的唯一标识符(ID)用ID_B来表示。ID_B是向量值,ID_B中的每个元素对应一个对象的ID。第二对象集合的大小为M。第二方P2在动作304向第一方P1发送M。

[0042] 第一方P1在动作305比较N和M的大小。如果N大于M,则将N作为参考大小。如果M大于N,则将M作为参考大小。通过布谷鸟哈希方式将第一对象集合中的对象的唯一标识符ID_A映射到第一哈希值集合H_A中。其中,第一对象集合中的每个对象的唯一标识符被映射成三个哈希值。由于布谷鸟哈希采用鸠占鹊巢的机制,因此最终每个哈希索引桶中只保留一个哈希值。这样,第一对象集合中的每个对象的唯一标识符对应一个哈希值。第一哈希值集合的大小大于第一对象集合的大小,因此第一哈希值集合中存在一些空白位置。具体地,第一哈希值集合的大小被设置为等于参考大小的 α 倍。其中, α 大于1。在一个示例中, α 等于1.27。在另一个示例中, α 等于1.32。然后,可向第一哈希值集合中的空白位置(即未与第一对象集合建立映射关系的位置)填入随机值。第一方P1还在动作305将第一哈希值集合中的每个元素与一个缺失标志数BL相关联。与哈希值相关联的缺失标志数BL被设置为第二值(例如,0)。与随机值相关联的缺失标志数BL被设置为不等于第二值。

[0043] 第二方P2在动作306通过简单哈希方式将第二对象集合中的对象的唯一标识符映射到第二哈希值集合H_B中。其中,第二对象集合中的每个对象的唯一标识符被映射成三个哈希值。这三个哈希值可对应同一个哈希索引桶,也可对应不同的哈希索引桶。第二哈希值集合的大小也等于参考大小的 α 倍。也就是说,第二哈希值集合的大小等于第一哈希值集合的大小。

[0044] 第一方P1在动作307通过可编程的不经意伪随机函数(Oblivious Programmable Pseudo-Random Function,OPPRF)联合第二方P2将第一哈希值集合中的每个元素的唯一标识符转换成第一编码值Code₀,将第二哈希值集合中的每个元素的唯一标识符转换成第二编码值Code₁。图4示出可编程的不经意伪随机函数(OPPRF)的执行步骤的示意性流程图和信令方案。发送者将数据集 $Y = (Y_0, Y_1, \dots, Y_{n-1})$ 作为不经意伪随机函数(Oblivious Pseudo-Random Function,OPRF)的输入,接收者将数据集 $X = (X_0, X_1, \dots, X_{n-1})$ 作为OPRF的输入。在动作403处的OPRF被运行之后,发送者获得密钥K,接收者获得 $F_K(X)$ 。 $F_K(X)$ 表示在伪随机函数中使用密钥K对X计算之后的结果。从 $F_K(X)$ 无法还原出K。图4中的 $\{F_K(X) | X = (X_0, X_1, \dots, X_{n-1})\}$ 表示 $F_K(X)$ 中的每个元素和X中的每个元素是成对的。发送者在动作405利用K本地生成 $F_K(Y)$ 。发送者在动作406生成随机数向量 $Code_0 = (c_0, c_1, \dots, c_{n-1})$ 。发送者在动作407计算 $H_Y Q = Code_0 - F_K(Y)$,那么可在动作408得到 $Q = H_Y^{-1}(Code_0 - F_K(Y))$ 。H表示哈希函数。 H_Y 表示对Y的哈希函数。 H_Y^{-1} 表示对 H_Y 取逆运算。发送者在动作410将Q发送给接收者,接收者在动作411基于Q在本地针对X做H函数计算,得到 H_X 。然后计算 $Code_1 = H_X Q + F_K(X)$ 。如果 $X_i = Y_j$,则 $Code_1_i = Code_0_j$,否则, $Code_1_i$ 等于一个随机数。这样通过比较 $Code_1$ 和 $Code_0$ 就可以确定X与Y的交集。

[0045] 如果第一方P1是发送者,第二方P2是接收者,则图3中的Code₀相当于图4中的Code₀,图3中的Code₁相当于图4中的Code₁。如果第一方P1是接收者,第二方P2是发送者,则图3中的Code₀相当于图4中的Code₁,图3中的Code₁相当于图4中的Code₀。

[0046] 回到图3,第一方P1在动作310联合第二方P2进行多方安全求交。在多方安全求交的过程中,第一编码值Code₀被碎片化成第一编码碎片和第二编码碎片,第二编码值Code₁被碎片化成第三编码碎片和第四编码碎片。因此,第一哈希值集合中的每个元素的唯一标识符由相应的第一编码碎片和第二编码碎片之和来确定。第二哈希值集合中的每个元素的唯一标识符由相应的第三编码碎片和第四编码碎片之和来确定。第一方P1持有第一编码碎

片和第三编码碎片。第二方P2持有第二编码碎片和第四编码碎片。缺失标志数BL由第一缺失标志碎片和第二缺失标志碎片之和来确定。第一方P1持有第一缺失标志碎片。第二方P2持有第二缺失标志碎片。对应第一哈希值集合中的同一元素的第一编码碎片与第一缺失标志碎片相关联。对应第一哈希值集合中的同一元素的第二编码碎片与第二缺失标志碎片相关联。

[0047] 在多方安全求交完成之后,第一方P1联合第二方P2将第一哈希值集合中的每个元素与一个标记值PSI相关联。该标记值PSI由第一标记碎片PSI0和第二标记碎片PSI1之和来确定。第一方P1持有第一标记碎片PSI0。第二方P2持有第二标记碎片PSI1。第一哈希值集合与第二哈希值集合中的相同哈希值对应的标记值PSI被设置为等于第一值(例如,等于1)。第一哈希值集合与第二哈希值集合中的不同哈希值对应的标记值PSI被设置为不等于第一值(例如,等于0)。应注意,第一哈希值集合中的每个随机值对应的标记值PSI也被设置为不等于第一值(例如,等于0)。对应第一哈希值集合中的同一元素的第一编码碎片H_A0与第一标记碎片PSI0相关联。对应第一哈希值集合中的同一元素的第二编码碎片H_A1与第二标记碎片PSI1相关联。

[0048] 相应地,第一对象集中的每个对象对应一个标记值PSI。标记值PSI由第一标记碎片PSI0和第二标记碎片PSI1之和来确定。第一方P1持有第一标记碎片PSI0。第二方P2持有第二标记碎片PSI1。第一对象集合与第二对象集合之间的真实交集中的每个对象所对应的标记值PSI等于第一值。第一对象集合中不属于真实交集的每个对象所对应的标记值PSI不等于第一值。标记值PSI可指示对应的元素是否属于真实交集。

[0049] 回到图2,在框S206处,第一方P1联合第二方P2根据真实交集生成混淆交集。其中,混淆交集包括真实交集和混淆对象。混淆对象是第一对象集合中不属于真实交集的对象中的一部分。混淆对象的数量与真实交集的大小之比是随机的。混淆交集集中的对象被确定为目标对象。图5示出根据本公开的实施例的第一方P1联合第二方P2根据真实交集生成混淆交集的步骤的示意性流程图和信令方案。

[0050] 参考图5,在图3的动作310进行多方安全求交之后,可获得第一对象集合与第二对象集合之间的真实交集的大小 s_i 。其中,真实交集的大小 s_i 由第一交集量级碎片 s_{i_0} 和第二交集量级碎片 s_{i_1} 之和来确定($s_i = s_{i_0} + s_{i_1}$)。第一方P1持有第一交集量级碎片 s_{i_0} 。第二方P2持有第二交集量级碎片 s_{i_1} 。在图5的动作503,第一方P1将第一对象集合的大小减去第一交集量级碎片 s_{i_0} 以获得第一非交集量级碎片 ni_0 ($ni_0 = N - s_{i_0}$)。第二方P2在动作504将0减去第二交集量级碎片 s_{i_1} 以获得第二非交集量级碎片 ni_1 ($ni_1 = 0 - s_{i_1}$)。第一方P1在动作505与第二方P2联合确定随机比例 P_s 。其中,随机比例 P_s 由第一比例碎片 P_{s_0} 和第二比例碎片 P_{s_1} 之和来确定。第一方P1持有第一比例碎片 P_{s_0} 。第二方P2持有第二比例碎片 P_{s_1} 。图6示出根据本公开的实施例的确定随机比例 P_s 的步骤的示意性流程图和信令方案。

[0051] 在图6的示例中,由第一方P1在动作601生成第一随机数 r_1 并将第一随机数 r_1 碎片化成第一随机数碎片 r_{1_0} 和第二随机数碎片 r_{1_1} 。其中,第一随机数 r_1 大于0且小于或者等于0.8。由第二方P2在动作602生成第二随机数 r_2 并将第二随机数 r_2 碎片化成第三随机数碎片 r_{2_0} 和第四随机数碎片 r_{2_1} 。其中,第二随机数 r_2 大于0且小于或者等于0.8。由第一方P1在动作603向第二方P2发送第二随机数碎片 r_{1_1} 。由第二方P2在动作604向第一方P1发送第

三随机数碎片 $r2_0$ 。由第一方P1在动作605将第一随机数碎片 $r1_0$ 和第三随机数碎片 $r2_0$ 相加以获得第一随机数碎片和 rs_0 。由第二方P2在动作606将第二随机数碎片 $r1_1$ 和第四随机数碎片 $r2_1$ 相加以获得第二随机数碎片和 rs_1 。由第一方P1在动作607生成第三随机数 $start$ 和第四随机数 end 。其中,第三随机数 $start$ 大于或者等于0且小于1。第四随机数 end 大于0且小于或者等于1。第三随机数 $start$ 小于第四随机数 end 。在一个示例中,可生成范围在0至1之间的2个随机数,将数值较小的随机数称为第三随机数 $start$,将数值较大的随机数称为第四随机数 end 。由第一方P1在动作608向第二方P2发送第三随机数 $start$ 和第四随机数 end 。由第一方P1在动作611将第一随机数碎片和 rs_0 乘以第四随机数 end 与第三随机数 $start$ 之差的积加上第三随机数 $start$ 以获得第一比例碎片 Ps_0 ,即 $Ps_0 = start + rs_0 \times (end - start)$ 。由第二方P2在动作612将第二随机数碎片和 rs_1 (即,第二随机数碎片 $r1_1$ 与第四随机数碎片 $r2_1$ 之和)乘以第四随机数 end 与第三随机数 $start$ 之差的积加上第三随机数 $start$ 来获得第二比例碎片 Ps_1 ,即 $Ps_1 = start + rs_1 \times (end - start)$ 。

[0052] 回到图5,第一方P1在动作509与第二方P2联合将真实交集的大小 si 乘以随机比例 Ps 的积除以非交集大小 ni 以获得掺杂比例 P 。 $\langle P \rangle = \langle si \rangle \odot \langle Ps \rangle / \langle ni \rangle$ 。其中, \odot 表示哈达玛积。 $\langle \rangle$ 表示碎片态。非交集大小 ni 等于第一非交集量级碎片 ni_0 与第二非交集量级碎片 ni_1 之和。掺杂比例 P 等于第一掺杂比例碎片 P_0 与第二掺杂比例碎片 P_1 之和。第一方P1持有第一掺杂比例碎片 P_0 。第二方P2持有第二掺杂比例碎片 P_1 。

[0053] 第一方P1在动作510与第二方P2联合针对第一哈希值集合中的每个元素确定一个参考随机数 R 。其中,参考随机数 R 符合均匀分布。参考随机数 R 由第一参考随机数碎片 $R0$ 和第二参考随机数碎片 $R1$ 之和来确定。第一方P1持有第一参考随机数碎片 $R0$ 。第二方P2持有第二参考随机数碎片 $R1$ 。图7示出根据本公开的实施例的针对第一哈希值集合中的每个元素确定一个参考随机数的步骤的示意性流程图和信令方案。

[0054] 在图7的示例中,由第一方P1在动作701生成第一随机数序列 $r11$ 和第二随机数序列 $r12$,将第一随机数序列 $r11$ 碎片化成第一碎片序列 $r11_0$ 和第二碎片序列 $r11_1$ ($r11 = r11_0 + r11_1$),并将第二随机数序列 $r12$ 碎片化成第三碎片序列 $r12_0$ 和第四碎片序列 $r12_1$ ($r12 = r12_0 + r12_1$)。第一随机数序列 $r11$ 和第二随机数序列 $r12$ 分别包括多个随机数。第一随机数序列 $r11$ 和第二随机数序列 $r12$ 中的随机数的数量等于第一数据碎片集中的第一数据碎片的数量。第一随机数序列 $r11$ 和第二随机数序列 $r12$ 中的每个随机数大于或者等于0且小于1。第一碎片序列 $r11_0$ 、第二碎片序列 $r11_1$ 、第三碎片序列 $r12_0$ 和第四碎片序列 $r12_1$ 的大小与第一随机数序列 $r11$ 的大小相同。这里的大小指的是所包含的元素数量。假设第一随机数序列 $r11$ 为 $[0.8, 0.11, 0.1]$,则可将第一随机数序列 $r11$ 碎片化成第一碎片序列 $r11_0 = [0.7, 0.1, 0.8]$ 和第二碎片序列 $r11_1 = [0.1, 0.10, -0.7]$ 。

[0055] 第二方P2在动作702生成第三随机数序列 $r21$ 和第四随机数序列 $r22$,将第三随机数序列 $r21$ 碎片化成第五碎片序列 $r21_0$ 和第六碎片序列 $r21_1$ ($r21 = r21_0 + r21_1$),并且将第四随机数序列 $r22$ 碎片化成第七碎片序列 $r22_0$ 和第八碎片序列 $r22_1$ ($r22 = r22_0 + r22_1$)。第三随机数序列 $r21$ 和第四随机数序列 $r22$ 分别包括多个随机数。第三随机数序列 $r21$ 和第四随机数序列 $r22$ 中的随机数数量等于第一数据碎片集中的第一数据碎片的数量。第三随机数序列 $r21$ 和第四随机数序列 $r22$ 中的每个随机数大于或者等于0且小于1。第五碎片序列 $r21_0$ 、第六碎片序列 $r21_1$ 、第七碎片序列 $r22_0$ 和第八碎片序列 $r22_1$ 的大小与第

一随机数序列r11的大小相同。

[0056] 由第一方P1在动作703向第二方P2发送第二碎片序列r11_1和第四碎片序列r12_1。由第二方P2在动作704向第一方P1发送第五碎片序列r21_0和第七碎片序列r22_0。

[0057] 由第一方P1在动作705利用第一碎片序列r11_0和第五碎片序列r21_0,与第二方P2联合比较第一随机数序列r11中的每个随机数是否小于第三随机数序列r21中的相应随机数。如图7所示,在动作705中,第二方P2利用第二碎片序列r11_1和第六碎片序列r21_1来联合比较。动作705的具体操作过程将在下文结合图8至图11进行介绍。动作705的比较结果Con由第一布尔结果碎片Con_0和第二布尔结果碎片Con_1进行异或的结果来确定。 $Con = Con_0 \oplus Con_1$ 。第一方P1持有第一布尔结果碎片Con_0。第二方P2持有第二布尔结果碎片Con_1。如果r11中的某个随机数小于r21中的相应随机数,则Con中的相应元素为真,否则为假。

[0058] 由第一方P1在动作706与第二方P2联合将第一布尔结果碎片Con_0和第二布尔结果碎片Con_1分别转化为第一算术结果碎片sel_0和第二算术结果碎片sel_1。第一方持有第一算术结果碎片sel_0,第二方持有第二算术结果碎片sel_1。第一算术结果碎片sel_0和第二算术结果碎片sel_1之和sel($sel = sel_0 + sel_1$)是比较结果Con所对应的算术值。在比较结果Con为真的情况下, $sel = 1$ 。在比较结果Con不为真的情况下, $sel = 0$ 。该动作相当于将布尔值Con转化为算术值sel。

[0059] 由第一方P1在动作707根据第一算术结果碎片sel_0来生成第一乘法因子碎片sb_0。第一乘法因子碎片sb_0中的每个元素等于1与第一算术结果碎片sel_0中的相应元素之差($sb_0 = 1 - sel_0$)。在这里,相应元素指的是位置相同的元素。例如,第一算术结果碎片sel_0中的第i个元素是第一乘法因子碎片sb_0中的第i个元素的相应元素。由第二方P2在动作708根据第二算术结果碎片sel_1来生成第二乘法因子碎片sb_1。第二乘法因子碎片sb_1中的每个元素等于0与第二算术结果碎片sel_1中的相应元素之差。或者可以认为第二乘法因子碎片sb_1中的每个元素等于第二算术结果碎片sel_1中的相应元素的相反数。 $sb_1 = -sel_1$ 。

[0060] 由第一方P1在动作709与第二方P2联合计算第四随机数序列r22与乘法因子sb之积以获得掩膜序列r22_mask。其中,乘法因子sb等于第一乘法因子碎片sb_0与第二乘法因子碎片sb_1之和($sb = sb_0 + sb_1$)。掩膜序列r22_mask等于第一掩膜碎片序列r22_mask_0和第二掩膜碎片序列r22_mask_1之和。第一方P1持有第一掩膜碎片序列r22_mask_0。第二方P2持有第二掩膜碎片序列r22_mask_1。在图7中,<>表示碎片态。<r22_mask>表示掩膜序列r22_mask的碎片态。<r22>表示第四随机数序列r22的碎片态。<sb>表示乘法因子sb的碎片态。<r22_mask>= $\langle r22 \rangle \odot \langle sb \rangle$ 表示 $r22_mask = r22 \odot sb$ 的操作是在碎片态下完成的,不会泄露任何一方的原始数据。⊙表示哈达玛积。

[0061] 由第一方P1在动作710利用第三碎片序列r12_0、第一算术结果碎片sel_0和第一掩膜碎片序列r22_mask_0,与第二方P2联合计算第二随机数序列r12乘以sel(第一算术结果碎片sel_0与第二算术结果碎片sel_1之和)的积加上掩膜序列r22_mask以获得参考随机数序列R。 $R = r12 \odot sel + r22_mask$ 。如图7所示,在动作710中,第二方P2利用第四碎片序列r12_1、第二算术结果碎片sel_1和第二掩膜碎片序列r22_mask_1来联合计算。在动作710计算出的参考随机数序列R等于第一参考随机数碎片序列R0和第二参考随机数碎片序列R1之

和 $(R=R0+R1)$ 。第一方P1持有第一参考随机数碎片序列R0,第二方P2持有第二参考随机数碎片序列R1。第一参考随机数碎片序列R0包括针对每个第一数据碎片的第一参考随机数碎片。第二参考随机数碎片序列R1包括针对每个第一数据碎片的第二参考随机数碎片。 $\langle R \rangle = \langle r12 \rangle \odot \langle sel \rangle + \langle r22_mask \rangle$ 表示 $R=r12 \odot sel+r22_mask$ 的操作是在碎片态下完成的,不会泄露任何一方的原始数据。

[0062] 图8示出根据本公开的实施例的确定第一值是否小于第二值的步骤的示意性流程图和信令方案。由第一方P1在动作801将第一值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803向第二方P2发送第二碎片值x2。由第二方P2在动作802将第二值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804向第一方P1发送第三碎片值y1。在动作805,由第一方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806,由第二方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 $z2=x2-y2$ 。

[0063] 可由第一方P1和第二方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($b1+b2=0$)。在动作807,将第一布尔零碎片a1和第一算术零碎片b1分配给第一方P1。在动作808,将第二布尔零碎片a2和第二算术零碎片b2分配给第二方P2。

[0064] 在动作809,由第一方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, $op11=(z1+b1) \oplus a1$ 。第一方P1还持有第三运算碎片op21,其中, $op21=0$ 。

[0065] 在动作810,由第二方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, $op22=z2+b2$ 。第二方P2还持有第四运算碎片op12,其中, $op12=a2$ 。

[0066] 在动作811,由第一方P1和第二方P2联合利用第一方P1处的第一并行前缀加法器和第二方P2处的第二并行前缀加法器在第一方P1处获得第一符号位碎片B1并且在第二方P2处获得第二符号位碎片B2。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。

[0067] 在由第一方P1来确定比较结果的示例中,第二方P2在动作813向第一方P1发送第二符号位碎片B2。由第一方P1在动作814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,则确定第一值小于第二值。如果比较值不为真,则确定第一值不小于第二值。类似地,也可以由第二方P2来确定比较结果。

[0068] 图9示出图8中的动作811的具体过程。在动作903,由第一方P1根据第一运算碎片op11和第三运算碎片op21并且由第二方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。

[0069] 图10示出由第一方P1和第二方P2联合执行的与运算的示意性流程图和信令方案。在图10中以第一方P1拥有第一输入碎片W1和第二输入碎片V1且第二方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9中的动作903使用图10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。

[0070] 下面描述图10所示的过程。

[0071] 第一方P1在动作1001获得三元组碎片矩阵 $\langle R1, S1, T1 \rangle$, 第二方P2在动作1002获得三元组碎片矩阵 $\langle R2, S2, T2 \rangle$ 。其中, $(R1 \oplus R2) \& (S1 \oplus S2) = (T1 \oplus T2)$ 。

[0072] 第一方P1在动作1003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。第二方P2在动作1004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。

[0073] 第二方P2在动作1005向第一方P1发送D2和E2。第一方P1在动作1006向第二方P2发送D1和E1。第一方P1在动作1007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, 第二方P2在动作1008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。

[0074] 第一方P1在动作1009计算第一输出碎片 $O1 = T1 \oplus (R1 \& E) \oplus (S1 \& D) \oplus (E \& D)$ 。第二方P2在动作1010计算第二输出碎片 $O2 = T2 \oplus (R2 \& E) \oplus (S2 \& D)$ 。当图9中的动作903使用图10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。

[0075] 回到图9, 第一方P1在动作904根据第五中间碎片p1 ($p1 = op11 \oplus op21$) 对G1的每一位进行逐位循环计算。第二方P2在动作905根据第六中间碎片p2 ($p2 = op12 \oplus op22$) 对G2的每一位进行逐位循环计算。图11示出图9中的动作904和905的示意性流程图和信令方案。

[0076] 第一方P1在动作1101对G1执行左移 2^i 位的操作以得到第一临时碎片G11, 即 $G11 = G1 \ll 2^i$ 。第二方P2在动作1102对G2执行左移 2^i 位的操作以得到第二临时碎片G12, 即 $G12 = G2 \ll 2^i$ 。其中, i 表示当前循环的索引。

[0077] 在动作1103处, 由第一方P1和第二方P2联合执行图10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1103的操作, 第一方P1获得第七中间碎片F1, 第二方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。

[0078] 第一方P1在动作1104对p1执行左移 2^i 位的操作以获得第三临时碎片p11 (即 $p11 = p1 \ll 2^i$), 然后再将p11更新为p11与kmask异或的结果 (即, $p11 = p11 \oplus kmask$)。其中, kmask是大小与op11相同的矩阵且其每一个元素值均为 $2^i - 1$ 。

[0079] 第二方P2在动作1105对p2执行左移 2^i 位的操作以获得第四临时碎片p12 (即 $p12 = p2 \ll 2^i$), 然后再将p12更新为p12与kmask异或的结果 (即, $p12 = p12 \oplus kmask$)。

[0080] 在动作1106处, 由第一方P1和第二方P2联合执行图10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1106的操作, 第一方P1获得更新后的p1, 第二方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1103处。更新后的p2也会被代入下一循环的动作1103处。

[0081] 第一方P1在动作1107对G1和F1执行异或操作以获得更新后的G1 (即, $G1 = G1 \oplus F1$)。更新后的G1会被代入下一循环的动作1101处。第二方P2在动作1107对G2和F2执行异或操作以获得更新后的G2 (即, $G2 = G2 \oplus F2$)。更新后的G2会被代入下一循环的动作1102处。

[0082] 再次回到图9, 第一方P1在动作906对G1左移1位以获得第九中间碎片C1 (即, $C1 = G1 \ll 1$)。第二方P2在动作907对G2左移1位以获得第十中间碎片C2 (即, $C2 = G2 \ll 1$)。第一方P1在动作908对p1和C1执行异或操作以获得第十一中间碎片Z1 (即, $Z1 = p1 \oplus C1$)。第二方P2

在动作909对p2和C2执行异或操作以获得第十二中间碎片Z2(即, $Z2=p2 \oplus C2$)。

[0083] 第一方P1在动作910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, $mask=0x1\ll n-1$,n表示第一值x的位数。第二方P2在动作911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, $mask=0x1\ll n-1$,n表示第二值y的位数。第一方P1在动作912将Z1转换成布尔类型以获得第一符号位碎片B1。第二方P2在动作913将Z2转换成布尔类型以获得第二符号位碎片B2。

[0084] 在上述过程中,由于第一方P1没有获得第二值的完整信息,而第二方P2也没有获得第一值的完整信息,因此该计算过程是安全的,不会泄露任何原始信息。

[0085] 在本公开的一些实施例中,在图7的动作705中,在第一方P1与第二方P2联合比较第一随机数序列r11中的每个随机数是否小于第三随机数序列r21中的相应随机数的过程中,可使用图8至图11所示的方法来执行比较操作。可将第一碎片序列r11_0中的每个随机数当成图8中的x1,可将第五碎片序列r21_0中的每个随机数当成图8中的y1,可将第二碎片序列r11_1中的每个随机数当成图8中的x2,可将第六碎片序列r21_1中的每个随机数当成图8中的y2。然后从图8的动作805和806开始执行后续操作。

[0086] 可替代地,也可以将图8中x和y看成是序列。那么在将图8的方案应用于动作705的情况下,第一方P1将第一碎片序列r11_0减去第五碎片序列r21_0以获得第九碎片序列。第一方P1获得第一布尔零碎片序列和第一算术零碎片序列。其中,第一布尔零碎片序列中的每个元素与第二布尔零碎片序列中的相应元素异或的结果为0。第一算术零碎片序列中的每个元素与第二算术零碎片序列中的相应元素相加的结果为0。第二方P2拥有第二布尔零碎片序列和第二算术零碎片序列。第一方P1计算第九碎片序列与第一算术零碎片序列之和与第一布尔零碎片序列异或的结果,以获得第一运算碎片序列。第一方P1与第二方P2联合利用第一方P1处的第一并行前缀加法器和第二方P2处的第二并行前缀加法器在第一方P1处获得第一符号位碎片序列并且在第二方P2处获得第二符号位碎片序列。其中,第一并行前缀加法器的输入为第一运算碎片序列和第三运算碎片序列。第二并行前缀加法器的输入为第二运算碎片序列和第四运算碎片序列。第二运算碎片序列由第二方P2计算第十碎片序列与第二算术零碎片序列之和来获得。第十碎片序列由第二方P2将第二碎片序列r11_1减去第六碎片序列r21_1来获得。第三运算碎片序列中的每个元素等于0。第四运算碎片序列等于第二布尔零碎片序列。第一方P1接收来自第二方P2的第二符号位碎片序列。第一方P1对第一符号位碎片序列与第二符号位碎片序列执行异或操作以获得比较值序列。如果比较值序列中的第一比较值为真,第一方P1确定第一随机数序列r11中与第一比较值相对应的随机数小于第三随机数序列r21中与第一比较值相对应的随机数。如果比较值序列中的第一比较值不为真,第一方P1确定第一随机数序列r11中与第一比较值相对应的随机数不小于第三随机数序列r21中与第一比较值相对应的随机数。在这里,第一比较值指的是比较值序列中的任一个比较值。

[0087] 回到图5,在动作513,第一方P1与第二方P2联合比较第一哈希值集合中的每个元素所对应的参考随机数R是否小于掺杂比例P,第一方P1与第二方P2联合比较第一哈希值集合中的每个元素对应的标记值PSI是否等于第一值,第一方P1与第二方P2联合比较第一哈希值集合中的每个元素对应的缺失标志数BL是否等于第二值。在动作513的这些比较操作可使用图8至图11所示的方法来执行。第一方P1还在动作513根据以下条件中的一个与第二

方P2联合从第一对象集合中选出目标对象:目标对象对应的标记值PSI等于第一值;或者目标对象对应的标记值PSI不等于第一值、目标对象对应的缺失标志数BL等于第二值并且目标对象对应的参考随机数R小于掺杂比例P。

[0088] 经过动作513, 第一方P1和第三方P2可分别得到目标对象的ID的一个碎片PSIDiffusion_0和PSIDiffusion_1。第三方P2在动作514向第一方P1发送PSIDiffusion_1。第一方P1在动作515计算目标对象的ID:PSIDiffusion=PSIDiffusion_0+PSIDiffusion_1。这样, 第一方P1可获得混淆交集集中的每个对象的ID。

[0089] 在本公开的一些实施例中, 第一对象集合和/或第二对象集合中的对象的特征可与该对象的ID一起参与运算, 这样, 在动作515, 可获得混淆交集集中的每个对象的ID及其对应的特征。

[0090] 根据本公开的实施例的第一方P1联合第三方P2挖掘目标对象的方法能够面向两种业务场景: (1) 高价值客群的新客拉新; (2) 高价值存量客群的活跃度营销。

[0091] 在高价值客群的新客拉新的业务场景下, 第一方P1可以是业务合作方(例如, 广告方), 第三方P2可以是发起方(例如, 银行节点)。在一个示例中, 银行节点可以筛选出可投资金额高于第一阈值且购买金额大于第二阈值的理财产品的客群(即, 具有高购买力的客群, 高价值客群)。广告方可筛选出能够投放广告的候选用户集合, 该候选用户集合中的每个用户可带有特征标签。银行节点和广告方可联合挖掘广告方中的目标对象, 由广告方获得高价值客群与候选用户集合之间的混淆交集的明文信息。广告方可按照用户的特征标签对混淆交集进行聚类处理以获得一个或多个聚类簇, 并在本地用户中寻找与混淆交集的各个聚类簇相似度较高的扩展用户。然后, 可向每个扩展用户投放银行节点的相应营销广告(例如, 理财产品信息)。不同聚类簇可被投放银行节点的不同营销广告。这样, 在广告方的帮助下, 银行节点能够根据已知的高价值客户群来开发新的高价值客户。

[0092] 在高价值存量客群的活跃度营销的业务场景下, 第一方P1可以是发起方(例如, 银行节点), 第三方P2可以是业务合作方(例如, 广告方)。在一个示例中, 银行节点的信用卡业务需要针对高价值客户投放“滑雪”相关的运营活动, 希望能够筛选出有滑雪偏好的存量高价值客群。在筛选过程中, 广告方由于存有大量的用户行为数据, 那么通过“关注滑雪大v的客群”、“某某电商购买过滑雪装备的客群”、“经常观看滑雪短视频的客群”、“去过滑雪场有定位的客群”等条件, 筛选与“滑雪”这一特性强相关的候选用户客群。银行节点和广告方可联合挖掘银行节点中的目标对象, 由银行节点可获得存量高价值客群与广告方的“滑雪”特性的客群之间的混淆交集的明文信息。银行节点可针对混淆交集投放“滑雪”相关的运营活动。

[0093] 图12示出根据本公开的实施例的第一方联合第三方挖掘目标对象的装置1200的示意性框图。装置1200位于第一方处。如图12所示, 该装置1200可包括处理器1210和存储有计算机程序的存储器1220。当计算机程序由处理器1210执行时, 使得装置1200可执行如图2所示的方法200的步骤。在一个示例中, 装置1200可以是计算机设备或云计算节点。装置1200可在本地确定第一对象集合。其中, 第一对象集合中的对象是针对目标对象选出的。装置1200可联合第三方通过安全求交方式确定第一对象集合与第三方的第二对象集合之间的真实交集。其中, 第二对象集合中的对象是针对目标对象选出的。第一对象集合中的每个对象对应一个标记值。标记值由第一标记碎片和第二标记碎片之和来确定。第一方持有第

一标记碎片。第二方持有第二标记碎片。真实交集中的每个对象所对应的标记值等于第一值。第一对象集合中不属于真实交集的每个对象所对应的标记值不等于第一值。装置1200可联合第二方根据真实交集生成混淆交集。其中,混淆交集包括真实交集和混淆对象。混淆对象是第一对象集合中不属于真实交集的对象中的一部分。混淆对象的数量与真实交集的大小之比是随机的。混淆交集集中的对象被确定为目标对象。

[0094] 在本公开的实施例中,处理器1210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1220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。

[0095] 此外,在本公开的实施例中,装置1200也可包括输入设备1230,例如键盘、鼠标等,用于输入用于挖掘目标对象的指令。另外,装置1200还可包括输出设备1240,例如显示器等,用于输出混淆交集集中的对象的ID。

[0096] 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。

[0097] 综上所述,根据本公开的实施例的第一方联合第二方挖掘目标对象的方法通过掺杂求交的方式借助于群体信息来挖掘目标对象,避免准确定位到具体个体。这样不仅可以符合合规性要求,还有助于维护用户的隐私权和数据安全。该方法通过OPRF来对第一对象集合和第二对象集合进行安全求交,能够在获得交集集中的对象的ID的同时获得相应对象的特征,这样该方法可以应用于新客拉新以及存量客户运营两种业务形态的推广,有助于更好地利用长尾高价值数据。

[0098] 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,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。

[0099] 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。

[0100] 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。

[0101] 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。

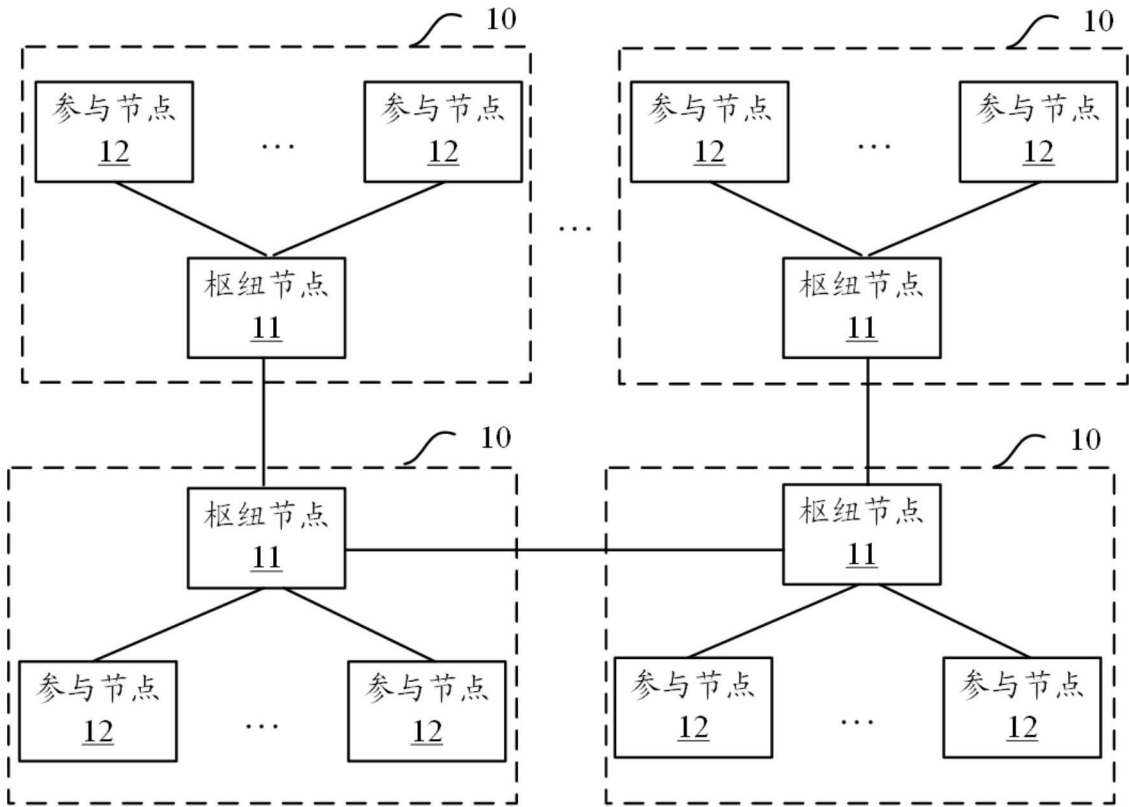


图1

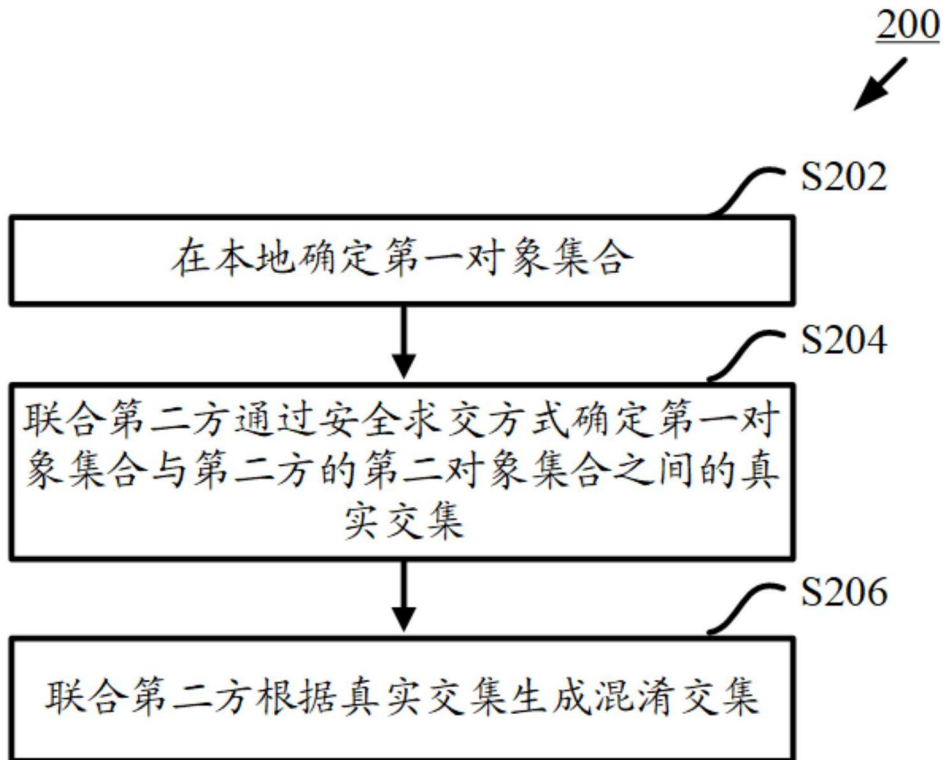


图2

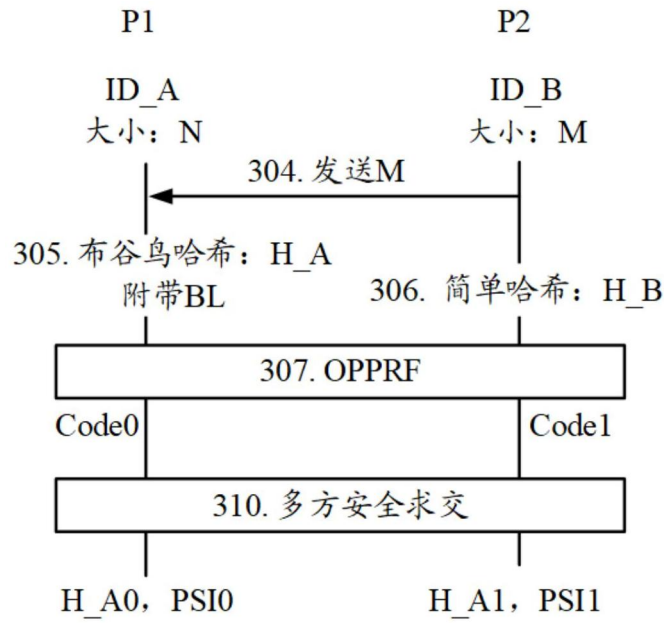


图3

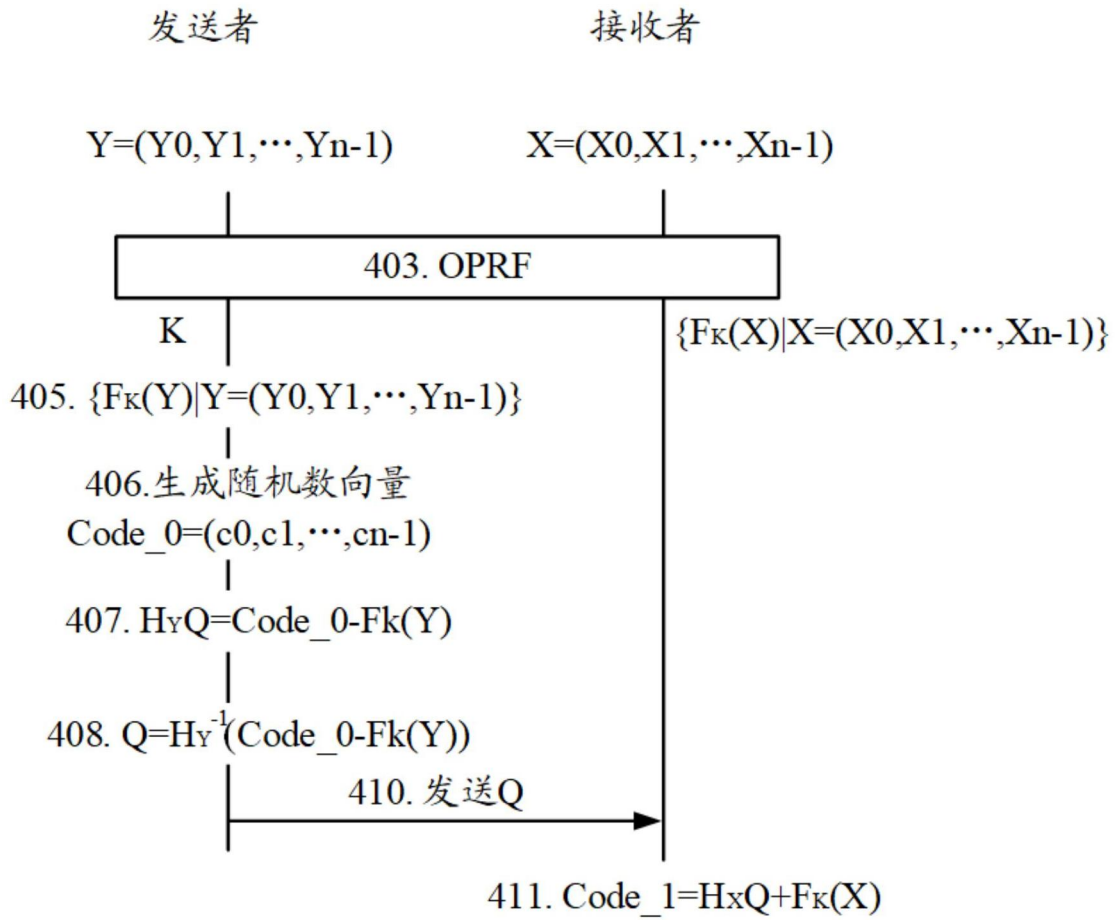


图4

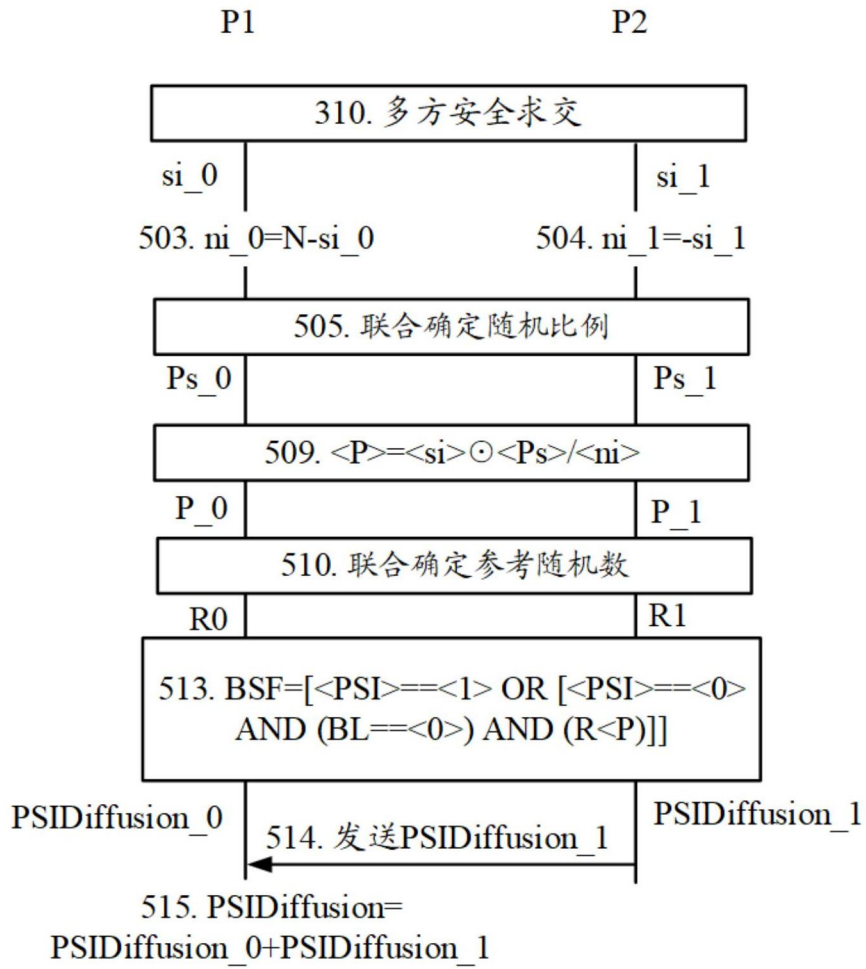


图5

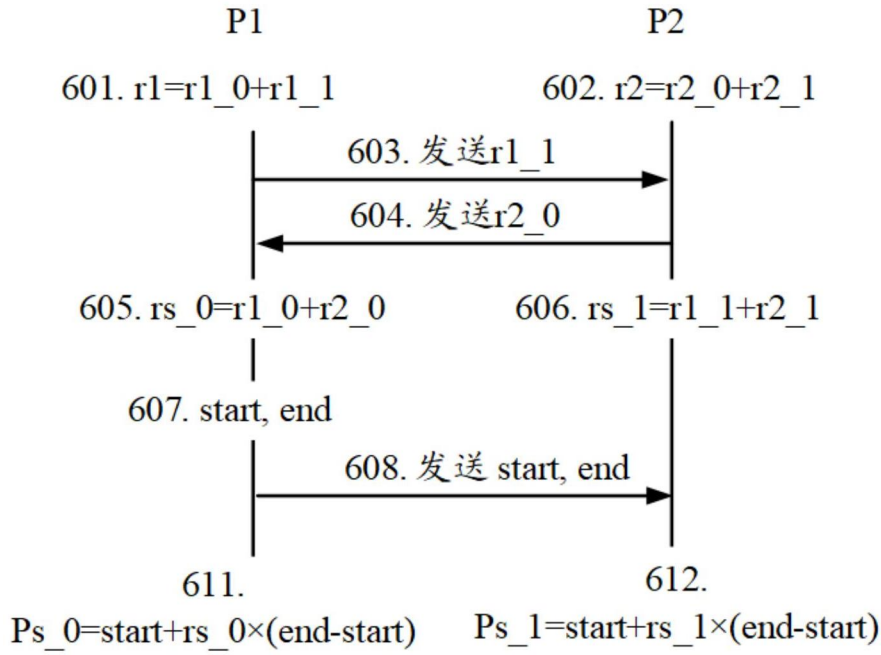


图6

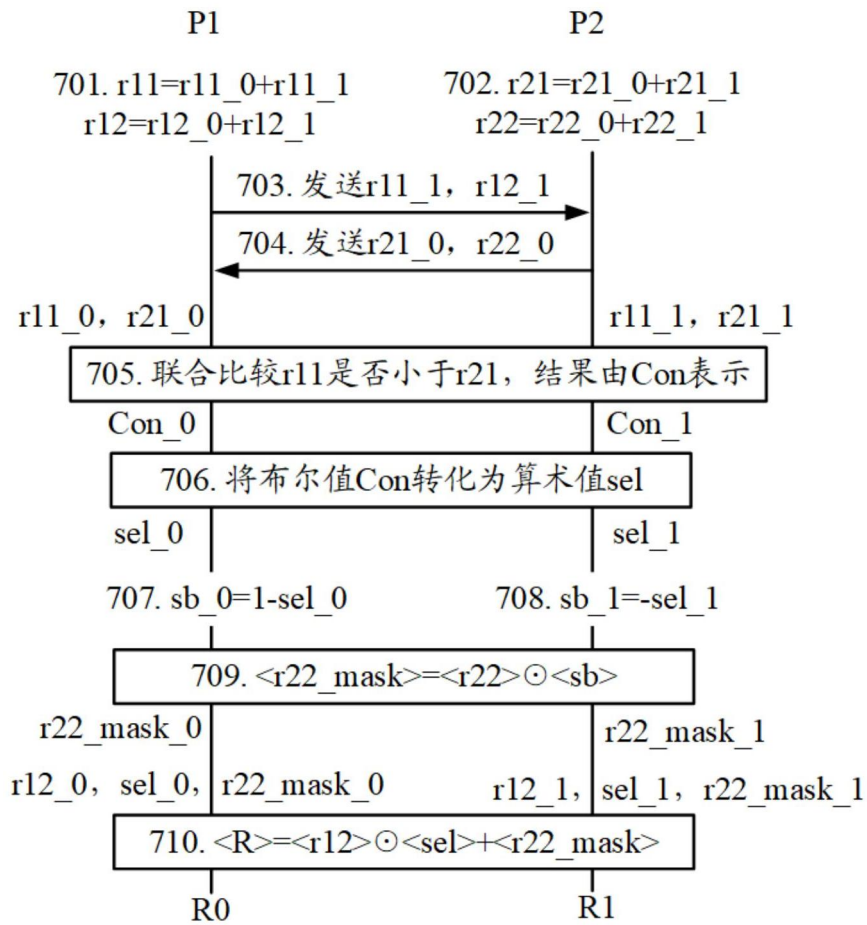


图7

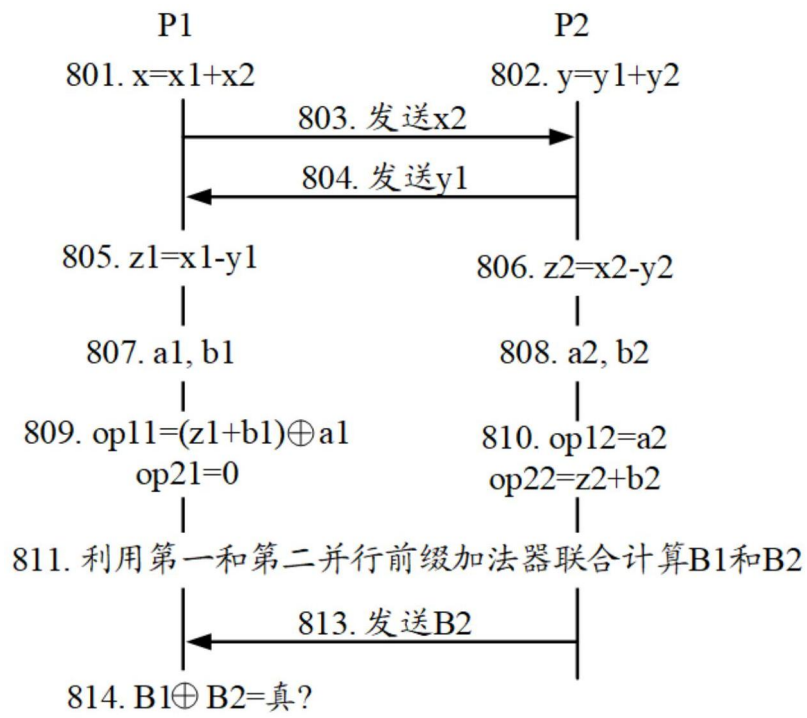


图8

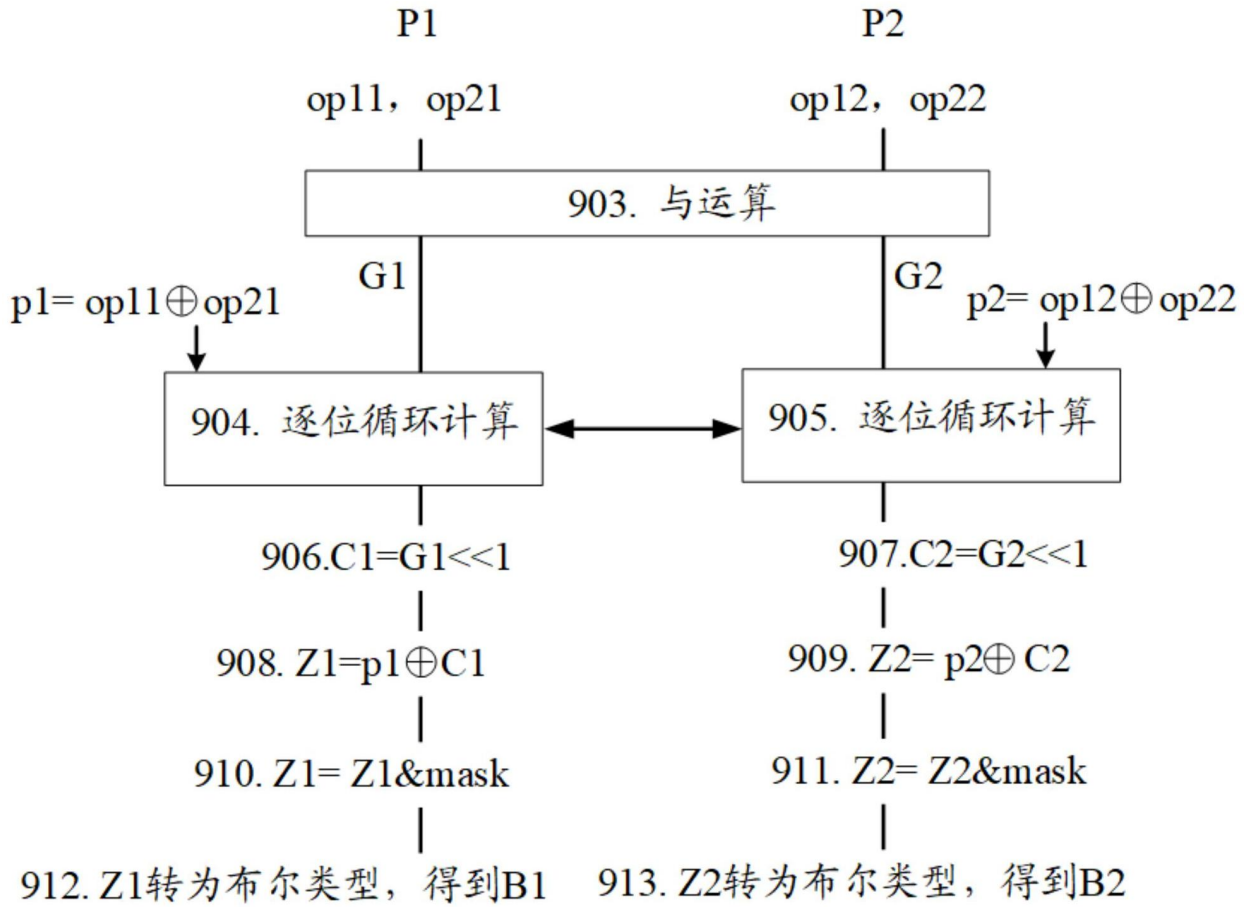


图9

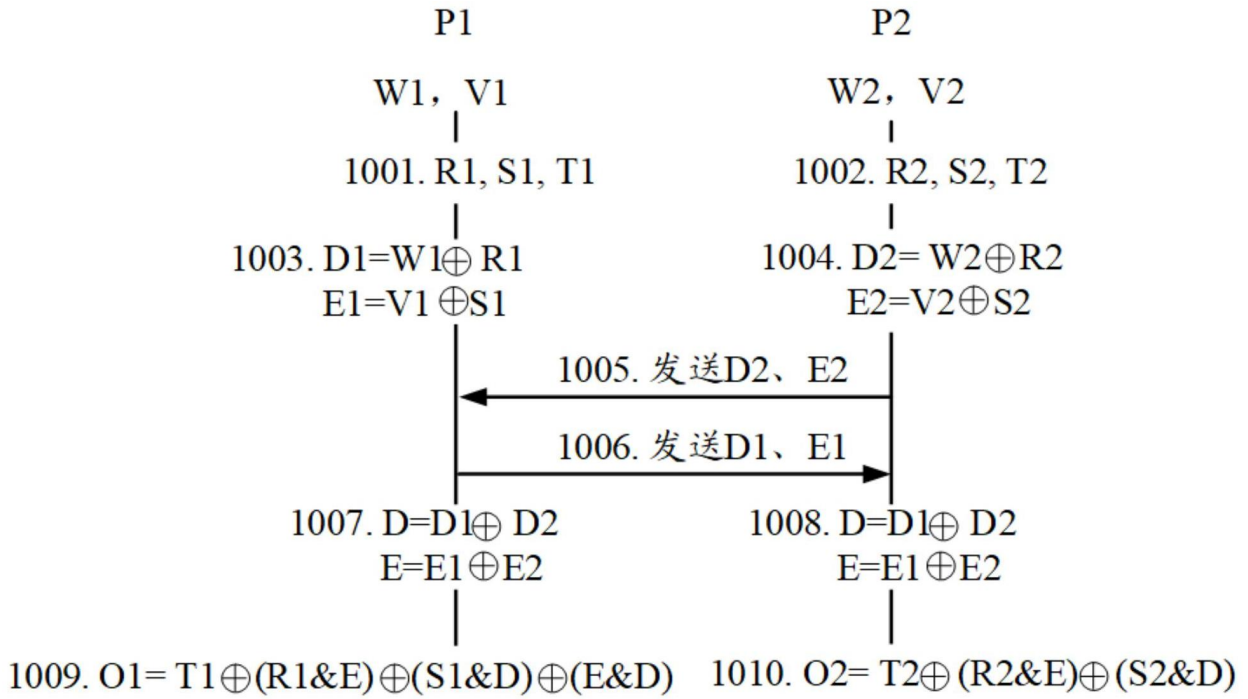


图10

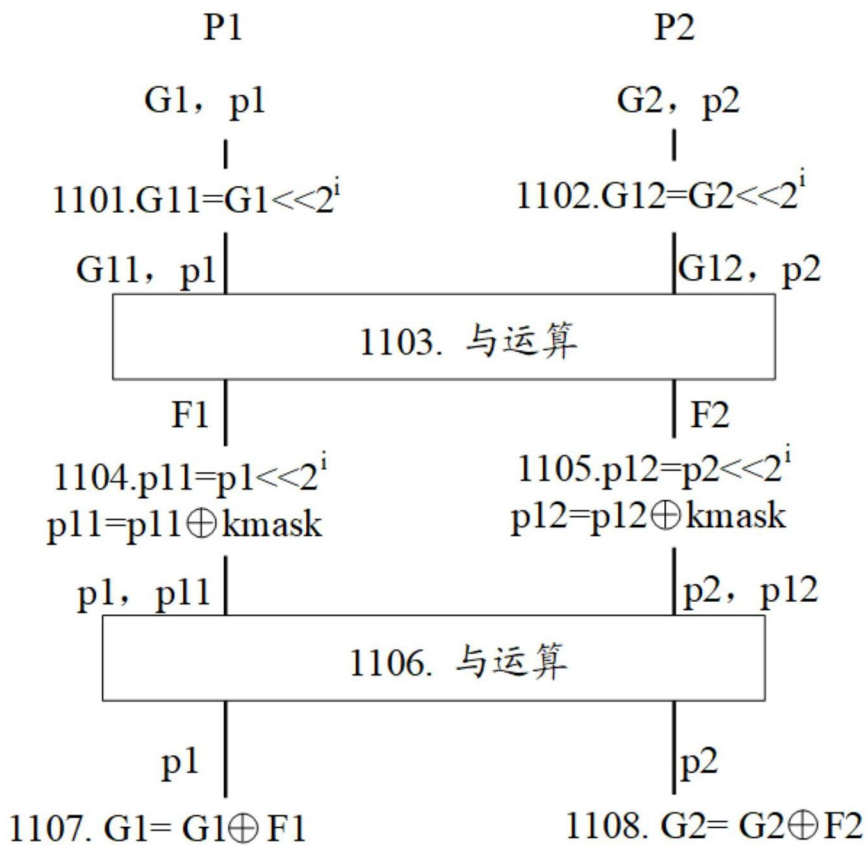


图11

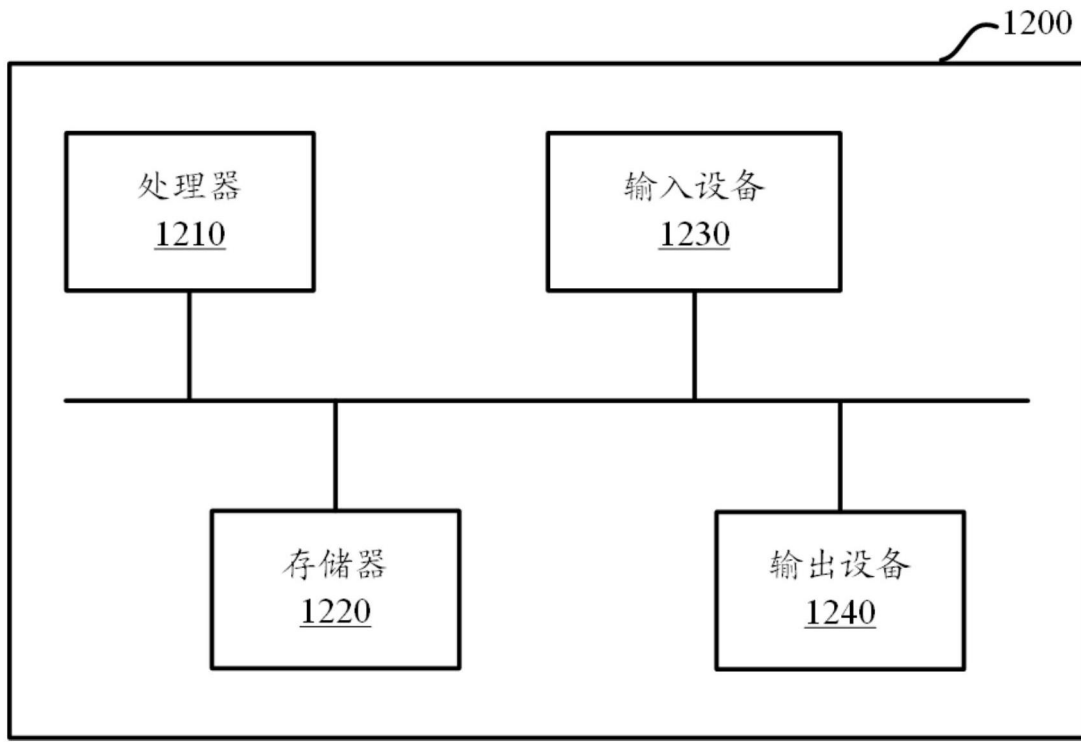


图12