(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 114595479 B
(45)授权公告日 2022.08.26

(21)申请号 202210500520.3
(22)申请日 2022.05.10
(65)同一申请的已公布的文献号
申请公布号 CN 114595479 A
(43)申请公布日 2022.06.07
(73)专利权人 富算科技(上海)有限公司
地址 200135 上海市浦东新区中国(上海)自由贸易试验区浦东大道1200号2层A区
(72)发明人 尤志强 卞阳 赵东 朱崇炳
(74)专利代理机构 北京超凡宏宇专利代理事务所(特殊普通合伙) 11463
专利代理师 唐正瑜
(51)Int.Cl.
G06F 21/60(2013.01)
G06F 16/22(2019.01)
(56)对比文件
CN 112235111 A,2021.01.15
CN 113392422 A,2021.09.14
EP 3089091 A1,2016.11.02
US 2020358594 A1,2020.11.12
CN 111737011 A,2020.10.02
韩姝敏等.一种基于隐私保护下的多方记录链接方法.《软件学报》.2017,(第09期),
Bohler.secure multi-party couputation of differentially private median.《29th USENIX Security Symposium》.2020,
审查员 岳孟果

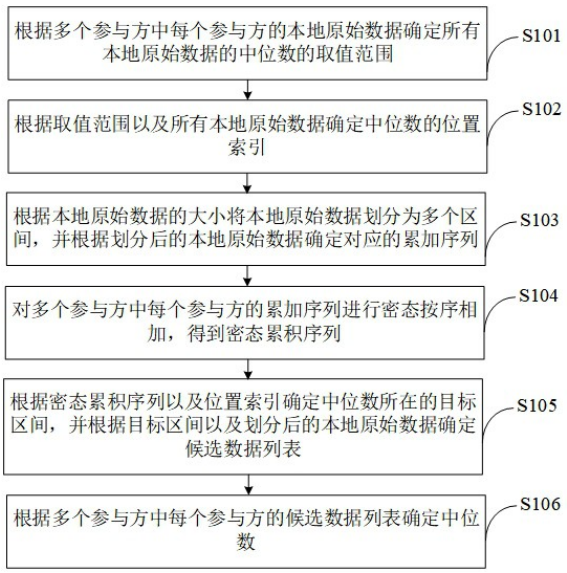
权利要求书3页 说明书15页 附图2页

(54)发明名称

一种数据中位数确定方法及装置

(57)摘要

本申请提供一种数据中位数确定方法及装置,应用于多方安全计算领域,方法包括:根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围;根据取值范围以及所有本地原始数据确定中位数的位置索引;根据本地原始数据的大小将本地原始数据划分为多个区间,并根据划分后的本地原始数据确定对应的累加序列;对多个参与方中每个参与方的累加序列进行密态按序相加,得到密态累积序列;根据密态累积序列以及位置索引确定中位数所在的目标区间,并根据目标区间以及划分后的本地原始数据确定候选数据列表;根据多个参与方中每个参与方的候选数据列表确定中位数。



CN 114595479 B

1.一种数据中位数确定方法,其特征在于,包括:

根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围;

根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引;

根据本地原始数据的大小将所述本地原始数据划分为多个区间,并根据划分后的本地原始数据确定对应的累加序列;其中,所述累加序列中的每一数值大小表征所述本地原始数据落入对应区间之前的所有数据的数量与落入对应区间中的所有数据的数量之和;

对多个参与方中每个参与方的累加序列进行密态按序相加,得到密态累积序列;

根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间,并根据所述目标区间以及划分后的本地原始数据确定候选数据列表;

根据多个参与方中每个参与方的候选数据列表确定所述中位数;

所述根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围,包括:

对所述本地原始数据进行求和得到对应的数据和,并确定所述本地原始数据中的第一数据数量;

对所述数据和进行密态化得到第一密态数据;

对多个参与方中每个参与方的第一数据数量进行求和得到所有参与方的第二数据数量;

根据所述第二数据数量以及多个参与方中每个参与方的第一密态数据,计算多个第一密态数据的均值得到第二密态数据,并根据所述第二密态数据执行多方安全计算标准差算子得到第三密态数据;

根据所述第二密态数据以及所述第三密态数据确定中位数的密态取值范围,并根据所述密态取值范围得到所述取值范围;

所述根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引,包括:

根据所述取值范围对所述本地原始数据进行过滤,得到在所述取值范围之内的数据集以及在所述取值范围之外的第三数据数量;

根据所述第二数据数量确定所述中位数的初始索引;

根据所述中位数的初始索引以及多个参与方中每个参与方的数据集中小于所述取值范围的第四数据数量确定所述位置索引;

所述根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间,包括:

将所述密态累积序列中的数据依次与所述位置索引的大小进行比较,直到所述密态累积序列中的数据大于等于所述位置索引,则将所述密态累积序列中对应的区间确定为所述目标区间;

所述根据多个参与方中每个参与方的候选数据列表确定所述中位数,包括:

确定所述候选数据列表的初始中位数;

对多个参与方中每个参与方的初始中位数进行密态排序,得到所述初始中位数的中间中位数;

对所述候选数据列表进行密态化得到密态数据列表;

将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较，并记录大于所述中间中位数的第六数据数量；

若所述第六数据数量小于所述位置索引，则随机挑选所有的密态数据列表中大于所述中间中位数的一个数据作为新的中间中位数，并重复执行所述将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较，并记录大于所述中间中位数的第六数据数量的步骤，直到确定所述中位数。

2.根据权利要求1所述的数据中位数确定方法，其特征在于，所述根据划分后的本地原始数据确定对应的累加序列，包括：

记录每个区间内的第五数据数量；

针对第i个区间，对第1个区间的第五数据数量至第i个区间的第五数据数量进行求和，得到所述累加序列中的第i个数值大小；其中，$1 \leqslant i \leqslant N$，N为区间数量且为正整数。

3.一种数据中位数确定装置，其特征在于，包括：

第一确定模块，用于根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围；

第二确定模块，用于根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引；

划分模块，用于根据本地原始数据的大小将所述本地原始数据划分为多个区间，并根据划分后的本地原始数据确定对应的累加序列；其中，所述累加序列中的每一数值大小表征所述本地原始数据落入对应区间之前的所有数据的数量与落入对应区间中的所有数据的数量之和；

相加模块，用于对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列；

第三确定模块，用于根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间，并根据所述目标区间以及划分后的本地原始数据确定候选数据列表；

第四确定模块，用于根据多个参与方中每个参与方的候选数据列表确定所述中位数；

所述第一确定模块具体用于：对所述本地原始数据进行求和得到对应的数据和，并确定所述本地原始数据中的第一数据数量；对所述数据和进行密态化得到第一密态数据；对多个参与方中每个参与方的第一数据数量进行求和得到所有参与方的第二数据数量；根据所述第二数据数量以及多个参与方中每个参与方的第一密态数据，计算多个第一密态数据的均值得到第二密态数据，并根据所述第二密态数据执行多方安全计算标准差算子得到第三密态数据；根据所述第二密态数据以及所述第三密态数据确定中位数的密态取值范围，并根据所述密态取值范围得到所述取值范围；

所述第二确定模块具体用于：根据所述取值范围对所述本地原始数据进行过滤，得到在所述取值范围之内的数据集以及在所述取值范围之外的第三数据数量；根据所述第二数据数量确定所述中位数的初始索引；根据所述中位数的初始索引以及多个参与方中每个参与方的数据集中小于所述取值范围的第四数据数量确定所述位置索引；

所述第三确定模块具体用于：将所述密态累积序列中的数据依次与所述位置索引的大小进行比较，直到所述密态累积序列中的数据大于等于所述位置索引，则将所述密态累积序列中对应的区间确定为所述目标区间；

所述第四确定模块具体用于:确定所述候选数据列表的初始中位数;对多个参与方中每个参与方的初始中位数进行密态排序,得到所述初始中位数的中间中位数;对所述候选数据列表进行密态化得到密态数据列表;将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较,并记录大于所述中间中位数的第六数据数量;若所述第六数据数量小于所述位置索引,则随机挑选所有的密态数据列表中大于所述中间中位数的一个数据作为新的中间中位数,并重复执行所述将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较,并记录大于所述中间中位数的第六数据数量的步骤,直到确定所述中位数。

4.一种计算机程序产品,其特征在于,包括计算机程序指令,所述计算机程序指令被处理器读取并运行时,执行如权利要求1或2所述的方法。

5.一种电子设备,其特征在于,包括:处理器、存储器和总线;

所述处理器和所述存储器通过所述总线完成相互间的通信;

所述存储器存储有可被所述处理器执行的计算机程序指令,所述处理器调用所述计算机程序指令能够执行如权利要求1或2所述的方法。

6.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储计算机程序指令,所述计算机程序指令被计算机运行时,使所述计算机执行如权利要求1或2所述的方法。

一种数据中位数确定方法及装置

技术领域

[0001] 本申请涉及多方安全计算领域，具体而言，涉及一种数据中位数确定方法及装置。

背景技术

[0002] 随着企业对数据保护意识的增强，多方安全计算作为一种有效的数据保护技术逐步在企业业务中得到应用。在多方安全计算中，算子是最底层、最基础、也是最重要的计算单元，复杂的统计以及机器学习都需要建立在算子的基础之上。

[0003] 而中位数是其中一种非常重要的算子之一，它是按顺序排列的一组数据中居于中间位置的数，代表一个样本、种群或概率分布中的一个数值，其可将数值集合划分为相等的上下两部分，即在这组数据中，有一半的数据比他大，有一半的数据比他小，中位数是以它在所有标志值中所处的位置确定的全体单位标志值的代表值，不受分布数列的极大或极小值影响，从而在一定程度上提高了中位数对分布数列的代表性。比如人口统计、人均收入统计等都会使用到中位数。

[0004] 因此，在多个参与方进行数据联合计算的场景下，经常需要对多个参与方持有的数据计算中位数。目前业内多方安全计算中位数算子，其执行逻辑普遍是对多方数据进行秘密共享后，在全量的碎片数据状态下进行相关运算，因此计算复杂度很高。

发明内容

[0005] 本申请实施例的目的在于提供一种数据中位数确定方法及装置，用以解决现有技术中多方安全计算中位数算子的计算复杂度很高的技术问题。

[0006] 第一方面，本申请实施例提供一种数据中位数确定方法，包括：根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围；根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引；根据本地原始数据的大小将所述本地原始数据划分为多个区间，并根据划分后的本地原始数据确定对应的累加序列；其中，所述累加序列中的每一数值大小表征所述本地原始数据落入对应区间之前及落入对应区间中的所有数据的数量；对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列；根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间，并根据所述目标区间以及划分后的本地原始数据确定候选数据列表；根据多个参与方中每个参与方的候选数据列表确定所述中位数。在上述方案中，通过多次对数据的筛选，逐步缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外，将密文态中的部分中间计算转移至明文态进行处理，在保证数据安全的基础上，基于明文态以及密文态的混合使用，同样可以降低计算的复杂度。

[0007] 在可选的实施方式中，所述根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围，包括：对所述本地原始数据进行求和得到对应的数据和，并确定所述本地原始数据中的第一数据数量；对所述数据和进行密态化得到第一密态数据；对多个参与方中每个参与方的第一数据数量进行求和得到所有参与方的第二数据

数量;根据所述第二数据数量以及多个参与方中每个参与方的第一密态数据,计算多个第一密态数据的均值得到第二密态数据,并根据所述第二密态数据执行多方安全计算标准差算子得到第三密态数据;根据所述第二密态数据以及所述第三密态数据确定中位数的密态取值范围,并根据所述密态取值范围得到所述取值范围。在上述方案中,可以通过确定中位数的取值范围缩小中位数所在的范围,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外,由于第一数据数量的大小不涉及数据泄露,因此可以对明文态的第一数据数量进行处理;而由于数据和的大小涉及数据泄露,因此可以对密文态的数据和进行处理。因此,基于明文态以及密文态的混合使用,同样可以降低计算的复杂度。

[0008] 在可选的实施方式中,所述根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引,包括:根据所述取值范围对所述本地原始数据进行过滤,得到在所述取值范围之内的数据集以及在所述取值范围之外的第三数据数量;根据所述第二数据数量确定所述中位数的初始索引;根据所述中位数的初始索引以及多个参与方中每个参与方的数据集中小于所述取值范围的第四数据数量确定所述位置索引。在上述方案中,可以在中位数的取值范围的基础上,进一步的对中位数进行筛选过滤,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0009] 在可选的实施方式中,所述根据划分后的本地原始数据确定对应的累加序列,包括:记录每个区间内的第五数据数量;针对第i个区间,对第1个区间的第五数据数量至第i个区间的第五数据数量进行求和,得到所述累加序列中的第i个数值大小;其中,1≤i≤N,N为区间数量且为正整数。在上述方案中,通过对本地原始数据进行划分,可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数,从而可以通过将区间范围内的个数与位置索引进行比较,以进一步的对中位数进行筛选过滤,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0010] 在可选的实施方式中,所述根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间,包括:将所述密态累积序列中的数据依次与所述位置索引的大小进行比较,直到所述密态累积序列中的数据大于等于所述位置索引,则将所述密态累积序列中对应的区间确定为所述目标区间。在上述方案中,通过对本地原始数据进行划分,可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数,从而可以通过将区间范围内的个数与位置索引进行比较,以进一步的对中位数进行筛选过滤,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0011] 在可选的实施方式中,所述根据多个参与方中每个参与方的候选数据列表确定所述中位数,包括:确定所述候选数据列表的初始中位数;对多个参与方中每个参与方的初始中位数进行密态排序,得到所述初始中位数的中间中位数;对所述候选数据列表进行密态化得到密态数据列表;将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较,并记录大于所述中间中位数的第六数据数量;若所述第六数据数量小于所述位置索引,则随机挑选所有的密态数据列表中大于所述中间中位数的一个数据作为新的中间中位数,并重复执行所述将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较,并记录大于所述中间中位数的第六数据数量的步骤,直到确定所述中位数。在上述方案中,在确定范围较小的候选数据列表之后,便可以基于上述候选数据列表确定中位数的大小,其中,通过多次对数据的筛选,逐步缩小中位数所在的

范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0012] 第二方面，本申请实施例提供一种数据中位数确定装置，包括：第一确定模块，用于根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围；第二确定模块，用于根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引；划分模块，用于根据本地原始数据的大小将所述本地原始数据划分为多个区间，并根据划分后的本地原始数据确定对应的累加序列；其中，所述累加序列中的每一数值大小表征所述本地原始数据落入对应区间之前及落入对应区间中的所有数据的数量；相加模块，用于对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列；第三确定模块，用于根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间，并根据所述目标区间以及划分后的本地原始数据确定候选数据列表；第四确定模块，用于根据多个参与方中每个参与方的候选数据列表确定所述中位数。在上述方案中，通过多次对数据的筛选，逐步缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外，将密文态中的部分中间计算转移至明文态进行处理，在保证数据安全的基础上，基于明文态以及密文态的混合使用，同样可以降低计算的复杂度。

[0013] 在可选的实施方式中，所述第一确定模块具体用于：对所述本地原始数据进行求和得到对应的数据和，并确定所述本地原始数据中的第一数据数量；对所述数据和进行密态化得到第一密态数据；对多个参与方中每个参与方的第一数据数量进行求和得到所有参与方的第二数据数量；根据所述第二数据数量以及多个参与方中每个参与方的第一密态数据，计算多个第一密态数据的均值得到第二密态数据，并根据所述第二密态数据执行多方安全计算标准差算子得到第三密态数据；根据所述第二密态数据以及所述第三密态数据确定中位数的密态取值范围，并根据所述密态取值范围得到所述取值范围。在上述方案中，可以通过确定中位数的取值范围缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外，由于第一数据数量的大小不涉及数据泄露，因此可以对明文态的第一数据数量进行处理；而由于数据和的大小涉及数据泄露，因此可以对密文态的数据和进行处理。因此，基于明文态以及密文态的混合使用，同样可以降低计算的复杂度。

[0014] 在可选的实施方式中，所述第二确定模块具体用于：根据所述取值范围对所述本地原始数据进行过滤，得到在所述取值范围之内的数据集以及在所述取值范围之外的第三数据数量；根据所述第二数据数量确定所述中位数的初始索引；根据所述中位数的初始索引以及多个参与方中每个参与方的数据集中小于所述取值范围的第四数据数量确定所述位置索引。在上述方案中，可以在中位数的取值范围的基础上，进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0015] 在可选的实施方式中，所述划分模块具体用于：记录每个区间内的第五数据数量；针对第i个区间，对第1个区间的第五数据数量至第i个区间的第五数据数量进行求和，得到所述累加序列中的第i个数值大小；其中，$1 \leqslant i \leqslant N$，N为区间数量且为正整数。在上述方案中，通过对本地原始数据进行划分，可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数，从而可以通过将区间范围内的个数与位置索引进行比较，以进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0016] 在可选的实施方式中，所述第三确定模块具体用于：将所述密态累积序列中的数

据依次与所述位置索引的大小进行比较，直到所述密态累积序列中的数据大于等于所述位置索引，则将所述密态累积序列中对应的区间确定为所述目标区间。在上述方案中，通过对本地原始数据进行划分，可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数，从而可以通过将区间范围内的个数与位置索引进行比较，以进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0017] 在可选的实施方式中，所述第四确定模块具体用于：确定所述候选数据列表的初始中位数；对多个参与方中每个参与方的初始中位数进行密态排序，得到所述初始中位数的中间中位数；对所述候选数据列表进行密态化得到密态数据列表；将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较，并记录大于所述中间中位数的第六数据数量；若所述第六数据数量小于所述位置索引，则随机挑选所有的密态数据列表中大于所述中间中位数的一个数据作为新的中间中位数，并重复执行所述将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较，并记录大于所述中间中位数的第六数据数量的步骤，直到确定所述中位数。在上述方案中，在确定范围较小的候选数据列表之后，便可以基于上述候选数据列表确定中位数的大小，其中，通过多次对数据的筛选，逐步缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0018] 第三方面，本申请实施例提供一种计算机程序产品，包括计算机程序指令，所述计算机程序指令被处理器读取并运行时，执行如第一方面所述的方法。

[0019] 第四方面，本申请实施例提供一种电子设备，包括：处理器、存储器和总线；所述处理器和所述存储器通过所述总线完成相互间的通信；所述存储器存储有可被所述处理器执行的计算机程序指令，所述处理器调用所述计算机程序指令能够执行如第一方面所述的方法。

[0020] 第五方面，本申请实施例提供一种计算机可读存储介质，所述计算机可读存储介质存储计算机程序指令，所述计算机程序指令被计算机运行时，使所述计算机执行如第一方面所述的方法。

[0021] 为使本申请的上述目的、特征和优点能更明显易懂，下文特举本申请实施例，并配合所附附图，作详细说明如下。

附图说明

[0022] 为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例中所需要使用的附图作简单地介绍，应当理解，以下附图仅示出了本申请的某些实施例，因此不应被看作是对范围的限定，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相关的附图。

[0023] 图1为本申请实施例提供的一种数据中位数确定方法的流程图；

[0024] 图2为本申请实施例提供的一种数据中位数确定装置的结构框图；

[0025] 图3为本申请实施例提供的一种电子设备的结构框图。

具体实施方式

[0026] 下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。

[0027] 请参照图1,图1为本申请实施例提供的一种数据中位数确定方法的流程图。在本申请实施例中,利用该数据中位数确定方法可以支持N个(N≥2)参与方联合进行中位数的查询计算,其中,N个参与方中的某一个参与方可以作为中位数查询计算的发起方。另外,对流程稍加改动,也可以支持无数据方作为发起方,对这N个参与方执行中位数查询计算,本文对此不做展开。

[0028] 可以理解的是,在上述数据中位数确定方法中,一部分步骤是每个参与方根据本地的数据独自执行的,而另一部分步骤是所有参与方协同执行的。针对所有参与方协同执行的部分步骤,作为一种实施方式,这部分步骤中的每一个步骤均可以由N个参与方中随机一个参与方执行;作为另一种实施方式,这部分步骤中的每一个步骤均可以由N个参与方中固定的一个参与方执行,例如:由发起方执行。

[0029] 在本申请实施例中,为了便于叙述,以发起方为执行主体对本申请实施例提供的数据中位数确定方法进行介绍。可以理解的是,在其他实施方式中,执行于其他参与方上的数据中位数确定方法可以包括比本申请实施例更少的步骤,或者执行于发起方上的数据中位数确定方法也可以包括比本申请实施例更少的步骤。本申请实施例对此不作具体的限定,本领域技术人员可以根据实际情况对数据中位数确定方法的具体步骤进行合适的调整。

[0030] 其中,本申请实施例提供的数据中位数确定方法可以包括如下内容:

[0031] 步骤S101:根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围。

[0032] 步骤S102:根据取值范围以及所有本地原始数据确定中位数的位置索引。

[0033] 步骤S103:根据本地原始数据的大小将本地原始数据划分为多个区间,并根据划分后的本地原始数据确定对应的累加序列。

[0034] 步骤S104:对多个参与方中每个参与方的累加序列进行密态按序相加,得到密态累积序列。

[0035] 步骤S105:根据密态累积序列以及位置索引确定中位数所在的目标区间,并根据目标区间以及划分后的本地原始数据确定候选数据列表。

[0036] 步骤S106:根据多个参与方中每个参与方的候选数据列表确定中位数。

[0037] 具体的,通过步骤S101-步骤S105可以实现多次对数据的筛选,从而逐步缩小中位数所在的范围。

[0038] 首先,在步骤S101中,每个参与方中本地均存储有本地原始数据,根据多个参与方中每个参与方的本地原始数据,可以确定上述所有的本地原始数据中的中位数的取值范围。其中,该取值范围是一个较大的取值范围。

[0039] 然后,在步骤S102中,根据上述取值范围以及上述所有的本地原始数据,可以确定中位数在所有的本地原始数据中的一个位置索引。其中,该位置索引为一个粗略的索引,只能确定出中位数大致的位置。

[0040] 接下来,在步骤S103中,通过将本地原始数据划分为多个区间,可以进一步的确定中位数具体在哪个区间中,并且可以根据划分后的本地原始数据确定对应的累加序列。其中,累加序列中的每一数值大小表征本地原始数据落入对应区间之前及落入对应区间中的所有数据的数量。

[0041] 接下来，在步骤S104中，由于上一步骤中每一个参与方均确定了一个对应的累加序列，因此可以对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列。其中，密态累积序列中的每一数值大小表征所有本地原始数据落入对应区间之前及落入对应区间中的所有数据的数量。

[0042] 接下来，在步骤S105中，根据上述密态累积序列以及上述位置索引可以确定中位数所在的目标区间。其中，该目标区间是一个较小的取值范围。根据上述目标区间以及划分后的本地原始数据，可以确定在目标区间中的候选数据列表，中位数位于所有候选数据列表中。

[0043] 最后，在步骤S106中，根据多个参与方中每个参与方的候选数据列表可以确定最终的中位数。

[0044] 可以理解的是，上述步骤S101-步骤S106均有多种实现方式，本申请实施例对此不作具体的限定，本领域技术人员可以根据实际情况选择合适的实现方式。其中，在后续实施例中，将举例对上述步骤的具体实施方式进行详细的介绍。

[0045] 在上述方案中，通过多次对数据的筛选，逐步缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外，将密文态中的部分中间计算转移至明文态进行处理，在保证数据安全的基础上，基于明文态以及密文态的混合使用，同样可以降低计算的复杂度。

[0046] 进一步的，在上述实施例的基础上，上述步骤S101具体可以包括如下内容：

[0047] 步骤1)，对本地原始数据进行求和得到对应的数据和，并确定本地原始数据中的第一数据数量。

[0048] 步骤2)，对数据和进行密态化得到第一密态数据。

[0049] 步骤3)，对多个参与方中每个参与方的第一数据数量进行求和得到所有参与方的第二数据数量。

[0050] 步骤4)，根据第二数据数量以及多个参与方中每个参与方的第一密态数据，计算多个第一密态数据的均值得到第二密态数据，并根据第二密态数据、各参与方的本地原始数据以及第二数据数量执行多方安全计算标准差算子得到标准差碎片化结果作为第三密态数据。

[0051] 步骤5)，根据第二密态数据以及第三密态数据确定中位数的密态取值范围，并根据密态取值范围得到取值范围。

[0052] 具体的，存在以下定理：如果X是由随机实数值组成的列表，其均值为μ，方差为σ^2，中位数为m，则$m \in [\mu - \sigma, \mu + \sigma]$。可以基于该定理确定中位数的取值范围。

[0053] 首先，在步骤1)中，各参与方各自在本地对本地原始数据进行求和，得到每个参与方对应的数据和；同时，各参与方各自在本地确定本地原始数据中数据的数量，作为第一数据数量。其中，在该步骤中各参与方是在明文态执行的。

[0054] 然后，在步骤2)中，各参与方通过多方安全计算的秘密共享机制，可以对各自的数据和进行密态化，每个参与方得到对应的第一密态数据。

[0055] 接下来，在步骤3)中，各参与方将各自的第一数据数量发送给发起方，发起方对接收到的所有的第一数据数量进行求和，得到所有参与方的总数据数量，即第二数据数量。

[0056] 其中，作为一种实施方式，由于第一数据数量的大小不会涉及数据泄密，因此，各

参与方可以通过明文通信的方式将各自的第一数据数量发送给发起方;作为另一种实施方式,在极其严苛的场景下,各参与方先分别在本地对各自的第一数据数量进行秘密共享碎片化,然后各参与节点执行密态的求和算子计算,得到求和结果碎片,最后将结果碎片发送给发起方进行结果的恢复,这样可以保证在得到结果的同时不暴露各节点持有的数据量。

[0057] 可以理解的是,本申请实施例对上述加密的方式也不作具体的限定,既可以采用上述秘密共享的方式,也可以采用其他方式,例如:半同态加密等。本领域技术人员可以根据实际情况进行合适的选择。

[0058] 作为一种实施方式,发起方在得到第二数据数量之后,还可以将第二数据数量广播告知其他参与方。

[0059] 接下来,在步骤4)中,通过多方安全计算加法算子,各参与方协同计算多个第一密态数据的和;通过多方安全计算除法算子,各参与方协同计算多个第一密态数据的和与第二数据数量的商,得到所有数据参与方全部数据对应的碎片化的均值结果,即第二密态数据;通过多方安全计算方差算子,各参与方协同计算得到碎片化的方差结果;通过多方安全计算开根号算子,各方协同计算得到碎片化的标准差结果,即为第三密态数据。

[0060] 最后,在步骤5)中,基于上述定理,可以确定中位数的密态取值范围为:大于等于第二密态数据与第三密态数据的差,小于等于第二密态数据与第三密态数据的和。

[0061] 其中,由于第二密态数据与第三密态数据这两个密文态数据涉及多个数值变量,因此恢复成明文态后,也不会被推测出各方具体的原始数值。因此,各参与方可以协同将密态取值范围进行恢复为明文态的取值范围。这样就实现了在多方安全计算场景下,在不泄露原始数据信息的前提下,各个参与方获得中位数的上下界的明文态取值范围。

[0062] 在上述方案中,可以通过确定中位数的取值范围缩小中位数所在的范围,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外,由于第一数据数量的大小不涉及数据泄露,因此可以对明文态的第一数据数量进行处理;而由于数据和的大小涉及数据泄露,因此可以对密文态的数据和进行处理。因此,基于明文态以及密文态的混合使用,同样可以降低计算的复杂度。

[0063] 进一步的,在上述实施例的基础上,上述步骤S102具体可以包括如下内容:

[0064] 步骤1),根据取值范围对本地原始数据进行过滤,得到在取值范围之内的数据集以及在取值范围之外的第三数据数量。

[0065] 步骤2),根据第二数据数量确定中位数的初始索引。

[0066] 步骤3),根据中位数的初始索引以及多个参与方中每个参与方的数据集中小于取值范围的第四数据数量确定位置索引。

[0067] 具体的,有了上述实施例中所求的中位数的取值范围,接下来可以进行筛选过滤的步骤,用于去除大量无效数据,进一步缩小中位数的取值范围。

[0068] 首先,在步骤1)中,由于已经确定中位数上下界的明文数值,因此每一个参与方可以对本地原始数据进行过滤。各参与方经过过滤处理后,可以得到取值范围内的数值集,以及取值范围之外的数据集中的数据数量,即第三数据数量。

[0069] 然后,在步骤2)中,根据第二数据数量,可以初步确定中位数的初始索引。举例来说,假设第二数据数量为M,若M为奇数,则中位数的初始索引可以表示为$k = (M + 1)/2$;若M为偶数,则中位数的初始索引有两个,可以分别表示为$k_1 = (M + 1)/2$以及

$k_2 = (M+2)/2$。

[0070] 可以理解的是，在后续步骤中，为了方便表述，均以M为奇数进行叙述；其中，M为偶数的实现方式与M为奇数的实现方式类似，区别仅在于会比奇数场景下多几个重复的计算步骤。

[0071] 最后，在步骤3)中，根据上述初始索引以及多个参与方中每个参与方的数据集中小于取值范围的第四数据数量。举例来说，假设所有参与方中数据大小小于取值范围的数据数量，即第四数据数量为q个，则可以对初始索引进行更新，得到位置索引$k' = k - q$。

[0072] 可以理解的是，在本申请实施例中，上述三个步骤均是在明文态执行的。

[0073] 在上述方案中，可以在中位数的取值范围的基础上，进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0074] 进一步的，在上述实施例的基础上，上述步骤S103具体可以包括如下内容：

[0075] 步骤1)，根据本地原始数据的大小将本地原始数据划分为多个区间。

[0076] 步骤2)，记录每个区间内的第五数据数量。

[0077] 步骤3)，针对第i个区间，对第1个区间的第五数据数量至第i个区间的第五数据数量进行求和，得到累加序列中的第i个数值大小；其中，1≤i≤N，N为区间数量且为正整数。

[0078] 具体的，通过上述实施例中的筛选过滤以及调整中位数的索引，可以获得了范围更小的取值范围，接下来可以对各参与方的本地原始数据进行处理。

[0079] 首先，在上述步骤1)中，可以根据本地原始数据的大小将本地原始数据划分为多个区间。作为一种实施方式，可以首先对中位数的取值范围进行分箱处理，这个区间即为$[\mu-\sigma, \mu+\sigma]$；然后设置分箱数B，那么分箱间隔依次为：

[0080] $$\mu-\sigma+\left[\frac{i}{B}\times 2\sigma, \frac{i+1}{B}\times 2\sigma\right]。$$

[0081] 其中，由于每一个参与方，其分箱层面都是一致的，因此所分割的箱的数量以及区间也是一致的。各个参与方在明文态下，可以按照上述分箱的区间对本地原始数据进行分箱，将值对应到每一个区间中。

[0082] 然后，在上述步骤2)中，各个参与方可以记录落入每个区间的数据数量，即第五数据数量。作为一种实施方式，各个参与方还可以记录落入区间内数值的索引，构建区间对应的数值索引结合，以备后续计算使用。

[0083] 可以理解的是，上述过程即构建直方分布图的过程，当各参与方处理完后，各自就能获得相对应的直方图数值分布。

[0084] 最后，在上述步骤3)中，各个参与方可以针对第i个区间，对第1个区间的第五数据数量至第i个区间的第五数据数量进行求和，得到累加序列中的第i个数值大小。其中，累加的过程可以表示为如下公式：

[0085] $$Lsum1[i] = Lsum1[i-1] + L1[i]；$$

[0086] 其中，$Lsum1[i]$为累加序列中的第i个数值大小，$Lsum1[i-1]$为累加序列中的第i-1个数值大小，$L1[i]$为第i个区间的第五数据数量。

[0087] 举例来说，比如某一个参与方的本地原始数据包括[5,8,10,4,6,9,2,6]，而分箱

的区间分别为[2,5)，[5,8)，[8,11)，那么就能统计处各个区间中的第五数值数量：[2,5)中第五数值数量为2个，[5,8)中第五数值数量3个，[8,11)第五数值数量3个；累加序列为$Lsum1[i]=[2,5,8]$。

[0088] 通过以上计算，可以获得每一个参与方的直方图分布以及累加序列。由于累加序列在各个参与方上长度一致，且每一个分箱含义也是一致，因此接下来可以执行上述步骤S104，即对多个累加序列进行密态按序相加，得到密态累积序列。

[0089] 在上述方案中，通过对本地原始数据进行划分，可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数，从而可以通过将区间范围内的个数与位置索引进行比较，以进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0090] 进一步的，在上述实施例的基础上，上述步骤S105具体可以包括如下内容：

[0091] 步骤1)，将密态累积序列中的数据依次与位置索引的大小进行比较，直到密态累积序列中的数据大于等于位置索引，则将密态累积序列中对应的区间确定为目标区间。

[0092] 步骤2)，根据目标区间以及划分后的本地原始数据确定候选数据列表。

[0093] 具体的，在上述实施中，得到了密态累积序列，被划分为B个区间，每个区间的密文态数值表示的是所有参与方的本地原始数据中数值在该区间范围内的个数。而密态累积序列的区间代表的数值从左到右是有序的，因此可以通过逐步进行密态比较确定中位数落在哪个区间。

[0094] 首先，在上述步骤1)中，可以将密态累积序列E_Lsum中的数据依次与位置索引的大小进行比较，比较的公式可以为：

[0095] $k' \le E\_Lsum$；

[0096] 其中，将密态累积序列中的第一数据与位置索引的大小进行比较，若第一数据小于位置索引，则将密态累积序列中的第二数据与位置索引的大小进行比较；依次类推，直至密态累积序列中的数据大于等于位置索引，则将密态累积序列中对应的区间确定为目标区间。其中，可以将中位数的位置索引更新为：原位置索引减去目标区间的第五数据数量。

[0097] 举例来说，假设中位数的位置索引为23，而E_Lsum=[1,8,13,20,31,49,52]：将密态累积序列中的1与23进行比较，由于1小于23，则将密态累积序列中的8与23进行比较；由于8小于23，则将密态累积序列中的13与23进行比较；由于13小于23，则将密态累积序列中的20与23进行比较；由于20小于23，则将密态累积序列中的31与23进行比较；由于31大于23，则将密态累积序列中第五个区间确定为目标区间。

[0098] 然后，在上述步骤2)中，可以根据目标区间以及划分后的本地原始数据确定候选数据列表。

[0099] 其中，作为一种实施方式，发起方可以在确定中位数所在的目标区间之后，将目标区间同步给其他参与方，其他参与方可以根据目标区间从对应的区间中提取出候选数据列表。

[0100] 作为另一种实施方式，各参与方在得到候选数据列表之后，还可以对候选数据列表进行排序。其中，由于候选数据列表是明文态，排序算法可以使用堆排序等高效算法。

[0101] 在上述方案中，通过对本地原始数据进行划分，可以获得所有参与方的本地原始

数据中数值在某一个区间范围内的个数，从而可以通过将区间范围内的个数与位置索引进行比较，以进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0102] 进一步的，在上述实施例的基础上，上述步骤S106具体可以包括如下内容：

[0103] 步骤1)，确定候选数据列表的初始中位数。

[0104] 步骤2)，对多个参与方中每个参与方的初始中位数进行密态排序，得到初始中位数的中间中位数。

[0105] 步骤3)，对候选数据列表进行密态化得到密态数据列表。

[0106] 步骤4)，将多个参与方中每个参与方的密态数据列表中的数据与中间中位数的大小进行比较，并记录大于中间中位数的第六数据数量。

[0107] 步骤5)，若第六数据数量小于位置索引，则随机挑选所有的密态数据列表中大于中间中位数的一个数据作为新的中间中位数，并重复执行将多个参与方中每个参与方的密态数据列表中的数据与中间中位数的大小进行比较，并记录大于中间中位数的第六数据数量的步骤，直到确定中位数。

[0108] 具体的，首先，在上述步骤1)中，各参与方可以在本地计算其候选数据列表的中位数，即初始中位数。

[0109] 然后，在上述步骤2)中，可以对多个参与方中每个参与方的初始中位数进行密态排序，得到多个初始中位数的中位数，即中间中位数。

[0110] 接下来，在上述步骤3)中，各参与方可以对候选数据列表中的数据进行密态化，得到密态数据列表。

[0111] 接下来，在上述步骤4)中，可以将所有密态数据列表中的数据一一与上述中间中位数进行比较。若密态数据列表中的数据小于等于中间中位数，则可以将该数据划分至低数值列表中；若密态数据列表中的数据大于中间中位数，则可以将该数据划分至高数值列表中，其中，高数值列表中的数据数量即为第六数据数量。

[0112] 接下来，在上述步骤5)中，将第六数据数量与位置索引进行比较，若第六数据数量小于位置索引，则表示中位数在高数值列表中；若第六数据数量大于位置索引，则表示中位数在低数值列表中；若第六数据数量等于位置索引，则表示找到中位数。

[0113] 如果中位数在高数值列表中，则更新中位数的索引为：原位置索引减去低数值列表中的数据数量。然后随机挑选高数值列表中的一个数作为新的中间中位数，其他数再与它进行比较，小于等于的划分至新的低数值列表，大于的划分至新的高数值列表，重复执行上述步骤，直到找到中位数。

[0114] 如果中位数在低数值列表中，其实现方式与上述高数值列表的实施方式类似，此处不再赘述。

[0115] 在上述方案中，在确定范围较小的候选数据列表之后，便可以基于上述候选数据列表确定中位数的大小，其中，通过多次对数据的筛选，逐步缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0116] 进一步的，在上述实施例的基础上，上述步骤S106具体可以包括如下内容：

[0117] 步骤1)，确定候选数据列表的初始中位数。

[0118] 步骤2)，对多个参与方中每个参与方的初始中位数进行密态排序，得到初始中位

数的中位数，该中位数即为目标中位数。

[0119] 在上述方案中，有些场景下对于中位数的精确性可以容忍一定的误差，因此可以采用上述近似计算方法，更快的计算出中位数。

[0120] 进一步的，在某些特殊场景中，本申请实施例提供的数据中位数确定方法可以包括如下内容：

[0121] 第一种特殊场景，各参与方的原始本地数据没有交叉。此时，本申请实施例提供的数据中位数确定方法可以包括如下内容：

[0122] 步骤1)，根据多个参与方中每个参与方的数据数量，确定中位数的索引值。

[0123] 步骤2)，对本地原始数据进行本地排序，得到有序列表。

[0124] 步骤3)，对多个参与方中每个参与方的最小值与最大值进行密文比较，确定是否满足互补交叉的条件，如果满足，则可以通过索引直接得到中位数。

[0125] 第二种特殊场景，N各参与方中，只有1个参与方的数据量大于1。此时，本申请实施例提供的数据中位数确定方法可以包括如下内容：

[0126] 步骤1)，根据多个参与方中每个参与方的数据数量，确定中位数的索引值k。

[0127] 步骤2)，对拥有多个数值的参与方的本地原始数据进行本地排序。分为以下两种情况：

[0128] 第一种，当上述本地原始数据中的数据数量大于中位数的索引k，则提取索引值$k-1$和k两个数值，其中$k-1$索引值标记为v_1，k索引值标记为v_2。

[0129] 步骤3)，参与方中的本地原始数据标记为d_i，将d_i与v_2进行密文比较，如果数值$d_i>v_2$，则直接排除；如果$d_i<v_2$，则继续将d_i与v_1进行密文比较，如果$d_i>v_1$，则将v_1更新为d_i、v_2更新为v_1。最后迭代结束，v_2即为所求中位数。

[0130] 第二种，当上述本地原始数据中的数据数量小于等于中位数的索引k，则选择本地原始数据中尾部的两个值，分别标记为v_1、v_2，其中$v_1<v_2$。

[0131] 步骤3)，参与方中的本地原始数据标记为d_i，将d_i与v_2进行密文比较，如果数值$d_i>v_2$，且本地原始数据中的数据数量加1不超过k，则将v_2的值更新为d_i；如果本地原始数据中的数据数量加1超过k，则直接排除di；如果$d_i<v_2$，且本地原始数据中的数据数量加1不超过k，则继续将d_i与v_1进行密文比较，如果$d_i<v_1$，则将k更新为$k-1$；如果$d_i>v_1$，则将v_1更新为d_i。最后迭代结束，v_2即为所求中位数。

[0132] 请参照图2，图2为本申请实施例提供的一种数据中位数确定装置的结构框图，该数据中位数确定装置200可以包括：第一确定模块201，用于根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围；第二确定模块202，用于根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引；划分模块203，用于根据本地原始数据的大小将所述本地原始数据划分为多个区间，并根据划分后的本地原始数据

确定对应的累加序列;其中,所述累加序列中的每一数值大小表征所述本地原始数据落入对应区间之前及落入对应区间中的所有数据的数量;相加模块204,用于对多个参与方中每个参与方的累加序列进行密态按序相加,得到密态累积序列;第三确定模块205,用于根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间,并根据所述目标区间以及划分后的本地原始数据确定候选数据列表;第四确定模块206,用于根据多个参与方中每个参与方的候选数据列表确定所述中位数。

[0133] 在本申请实施例中,通过多次对数据的筛选,逐步缩小中位数所在的范围,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外,将密文态中的部分中间计算转移至明文态进行处理,在保证数据安全的基础上,基于明文态以及密文态的混合使用,同样可以降低计算的复杂度。

[0134] 进一步的,所述第一确定模块201具体用于:对所述本地原始数据进行求和得到对应的数据和,并确定所述本地原始数据中的第一数据数量;对所述数据和进行密态化得到第一密态数据;对多个参与方中每个参与方的第一数据数量进行求和得到所有参与方的第二数据数量;根据所述第二数据数量以及多个参与方中每个参与方的第一密态数据,计算多个第一密态数据的均值得到第二密态数据,并根据所述第二密态数据执行多方安全计算标准差算子得到第三密态数据;根据所述第二密态数据以及所述第三密态数据确定中位数的密态取值范围,并根据所述密态取值范围得到所述取值范围。

[0135] 在本申请实施例中,可以通过确定中位数的取值范围缩小中位数所在的范围,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。此外,由于第一数据数量的大小不涉及数据泄露,因此可以对明文态的第一数据数量进行处理;而由于数据和的大小涉及数据泄露,因此可以对密文态的数据和进行处理。因此,基于明文态以及密文态的混合使用,同样可以降低计算的复杂度。

[0136] 进一步的,所述第二确定模块202具体用于:根据所述取值范围对所述本地原始数据进行过滤,得到在所述取值范围之内的数据集以及在所述取值范围之外的第三数据数量;根据所述第二数据数量确定所述中位数的初始索引;根据所述中位数的初始索引以及多个参与方中每个参与方的数据集中小于所述取值范围的第四数据数量确定所述位置索引。

[0137] 在本申请实施例中,可以在中位数的取值范围的基础上,进一步的对中位数进行筛选过滤,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0138] 进一步的,所述划分模块203具体用于:记录每个区间内的第五数据数量;针对第i个区间,对第1个区间的第五数据数量至第i个区间的第五数据数量进行求和,得到所述累加序列中的第i个数值大小;其中,$1 \leqslant i \leqslant N$,N为区间数量且为正整数。

[0139] 在本申请实施例中,通过对本地原始数据进行划分,可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数,从而可以通过将区间范围内的个数与位置索引进行比较,以进一步的对中位数进行筛选过滤,这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0140] 进一步的,所述第三确定模块205具体用于:将所述密态累积序列中的数据依次与所述位置索引的大小进行比较,直到所述密态累积序列中的数据大于等于所述位置索引,则将所述密态累积序列中对应的区间确定为所述目标区间。

[0141] 在本申请实施例中，通过对本地原始数据进行划分，可以获得所有参与方的本地原始数据中数值在某一个区间范围内的个数，从而可以通过将区间范围内的个数与位置索引进行比较，以进一步的对中位数进行筛选过滤，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0142] 进一步的，所述第四确定模块206具体用于：确定所述候选数据列表的初始中位数；对多个参与方中每个参与方的初始中位数进行密态排序，得到所述初始中位数的中间中位数；对所述候选数据列表进行密态化得到密态数据列表；将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较，并记录大于所述中间中位数的第六数据数量；若所述第六数据数量小于所述位置索引，则随机挑选所有的密态数据列表中大于所述中间中位数的一个数据作为新的中间中位数，并重复执行所述将多个参与方中每个参与方的密态数据列表中的数据与所述中间中位数的大小进行比较，并记录大于所述中间中位数的第六数据数量的步骤，直到确定所述中位数。

[0143] 在本申请实施例中，在确定范围较小的候选数据列表之后，便可以基于上述候选数据列表确定中位数的大小，其中，通过多次对数据的筛选，逐步缩小中位数所在的范围，这样最后在密文态计算中位数的过程中可以降低计算的复杂度。

[0144] 请参照图3，图3为本申请实施例提供的一种电子设备的结构框图，该电子设备300包括：至少一个处理器301，至少一个通信接口302，至少一个存储器303和至少一个通信总线304。其中，通信总线304用于实现这些组件直接的连接通信，通信接口302用于与其他节点设备进行信令或数据的通信，存储器303存储有处理器301可执行的机器可读指令。当电子设备300运行时，处理器301与存储器303之间通过通信总线304通信，机器可读指令被处理器301调用时执行上述数据中位数确定方法。

[0145] 例如，本申请实施例的处理器301通过通信总线304从存储器303读取计算机程序并执行该计算机程序可以实现如下方法：步骤S101：根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围。步骤S102：根据取值范围以及所有本地原始数据确定中位数的位置索引。步骤S103：根据本地原始数据的大小将本地原始数据划分为多个区间，并根据划分后的本地原始数据确定对应的累加序列。步骤S104：对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列。步骤S105：根据密态累积序列以及位置索引确定中位数所在的目标区间，并根据目标区间以及划分后的本地原始数据确定候选数据列表。步骤S106：根据多个参与方中每个参与方的候选数据列表确定中位数。

[0146] 其中，处理器301包括一个或多个，其可以是一种集成电路芯片，具有信号的处理能力。上述的处理器301可以是通用处理器，包括中央处理器(Central Processing Unit，简称CPU)、微控制单元(Micro Controller Unit，简称MCU)、网络处理器(Network Processor，简称NP)或者其他常规处理器；还可以是专用处理器，包括神经网络处理器(Neural-network Processing Unit，简称NPU)、图形处理器(Graphics Processing Unit，简称GPU)、数字信号处理器(Digital Signal Processor，简称DSP)、专用集成电路(Application Specific Integrated Circuits，简称ASIC)、现场可编程门阵列(Field Programmable Gate Array，简称FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。并且，在处理器301为多个时，其中的一部分可以是通用处理器，另一

部分可以是专用处理器。

[0147] 存储器303包括一个或多个，其可以是，但不限于，随机存取存储器(Random Access Memory，简称RAM)，只读存储器(Read Only Memory，简称ROM)，可编程只读存储器(Programmable Read-Only Memory，简称PROM)，可擦除可编程只读存储器(Erasable Programmable Read-Only Memory，简称EPROM)，电可擦除可编程只读存储器(Electric Erasable Programmable Read-Only Memory，简称EEPROM)等。

[0148] 可以理解，图3所示的结构仅为示意，电子设备300还可包括比图3中所示更多或者更少的组件，或者具有与图3所示不同的配置。图3中所示的各组件可以采用硬件、软件或其组合实现。于本申请实施例中，电子设备300可以是，但不限于台式机、笔记本电脑、智能手机、智能穿戴设备、车载设备等实体设备，还可以是虚拟机等虚拟设备。另外，电子设备300也不一定是单台设备，还可以是多台设备的组合，例如服务器集群，等等。

[0149] 本申请实施例还提供一种计算机程序产品，包括存储在计算机可读存储介质上的计算机程序，计算机程序包括计算机程序指令，当计算机程序指令被计算机执行时，计算机能够执行上述实施例中数据中位数确定方法的步骤，例如包括：根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围；根据所述取值范围以及所有本地原始数据确定所述中位数的位置索引；根据本地原始数据的大小将所述本地原始数据划分为多个区间，并根据划分后的本地原始数据确定对应的累加序列；其中，所述累加序列中的每一数值大小表征所述本地原始数据落入对应区间之前及落入对应区间中的所有数据的数量；对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列；根据所述密态累积序列以及所述位置索引确定所述中位数所在的目标区间，并根据所述目标区间以及划分后的本地原始数据确定候选数据列表；根据多个参与方中每个参与方的候选数据列表确定所述中位数。

[0150] 在本申请所提供的实施例中，应该理解到，所揭露装置和方法，可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，又例如，多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

[0151] 另外，作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0152] 再者，在本申请各个实施例中的各功能模块可以集成在一起形成一个独立的部分，也可以是各个模块单独存在，也可以两个或两个以上模块集成形成一个独立的部分。

[0153] 需要说明的是，功能如果以软件功能模块的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机，服务器，或者网络设备等)执行本申请各个实施例所述方法的全部或

部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(Read-Only Memory,ROM)随机存取存储器(Random Access Memory,RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0154] 在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。

[0155] 以上所述仅为本申请的实施例而已,并不用于限制本申请的保护范围,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

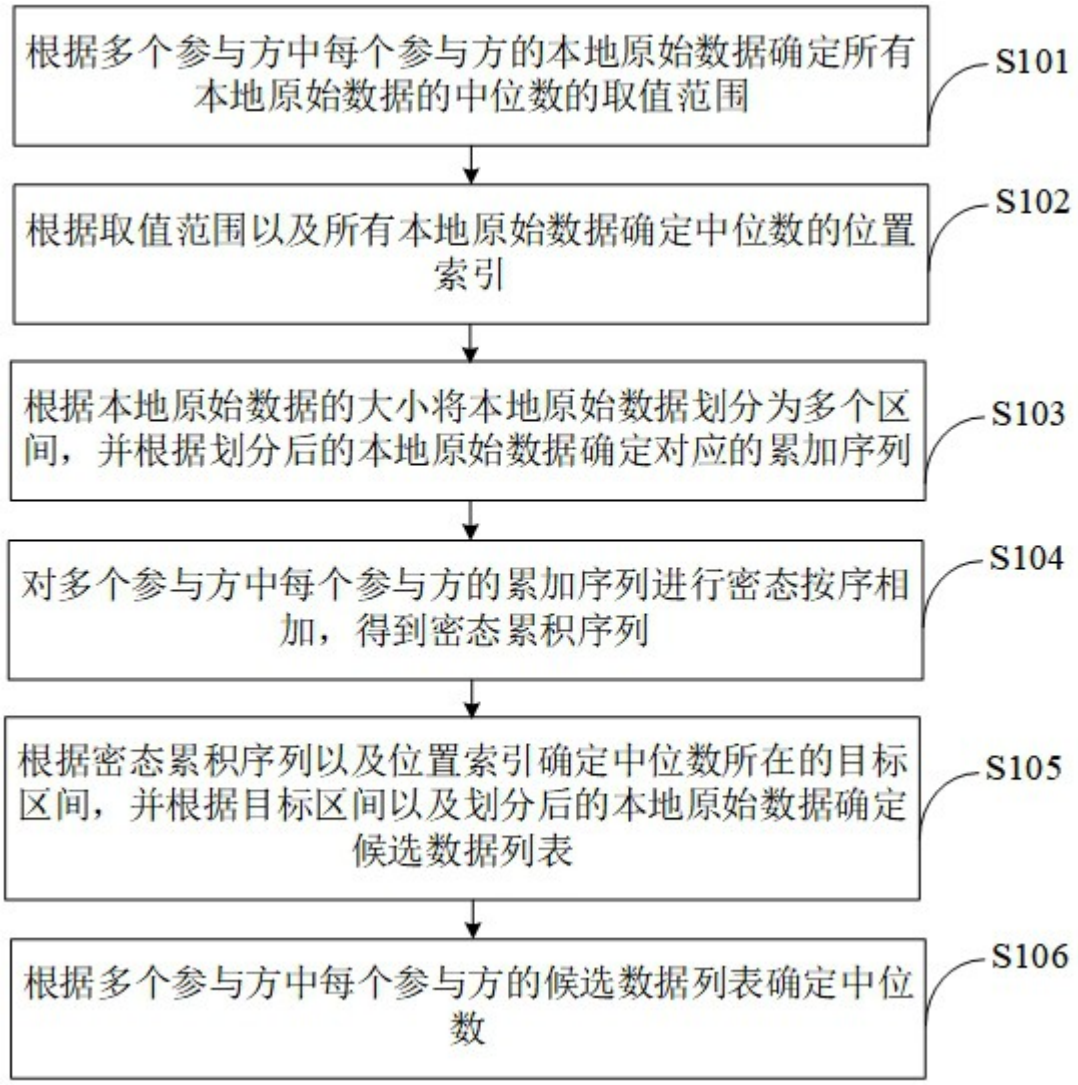

根据多个参与方中每个参与方的本地原始数据确定所有本地原始数据的中位数的取值范围
S101
根据取值范围以及所有本地原始数据确定中位数的位置索引
S102
根据本地原始数据的大小将本地原始数据划分为多个区间，并根据划分后的本地原始数据确定对应的累加序列
S103
对多个参与方中每个参与方的累加序列进行密态按序相加，得到密态累积序列
S104
根据密态累积序列以及位置索引确定中位数所在的目标区间，并根据目标区间以及划分后的本地原始数据确定候选数据列表
S105
根据多个参与方中每个参与方的候选数据列表确定中位数
S106

图1

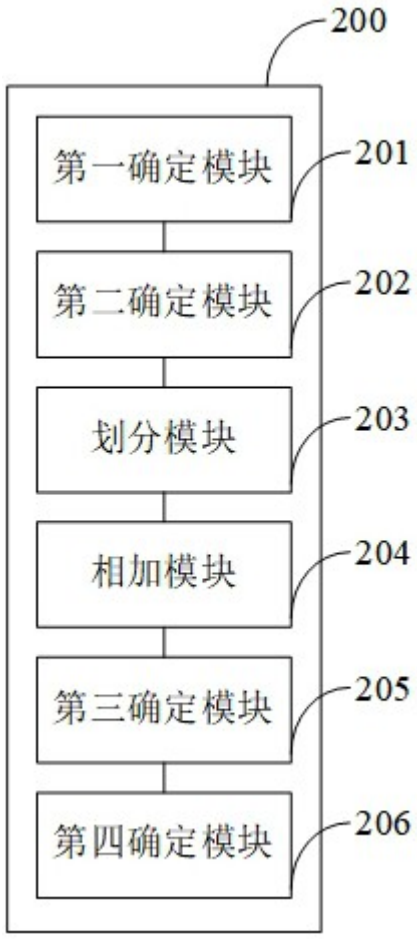

200
第一确定模块
201
第二确定模块
202
划分模块
203
相加模块
204
第三确定模块
205
第四确定模块
206

图2

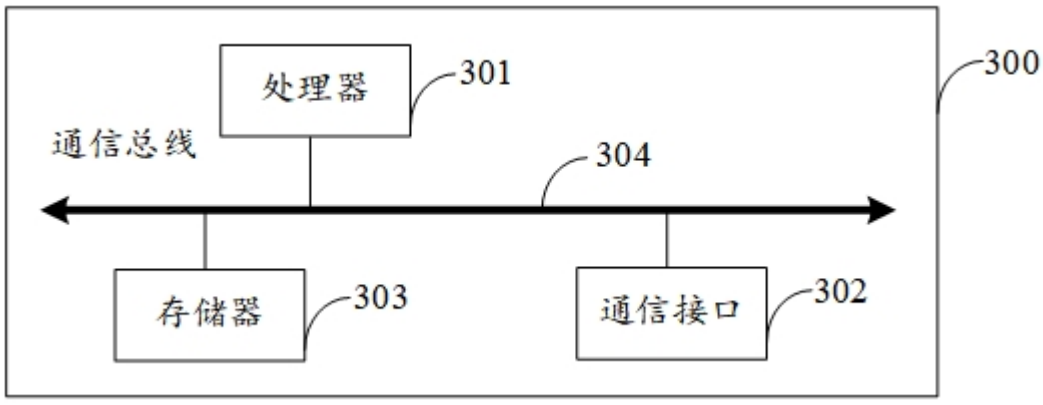

处理器
301
300
通信总线
304
存储器
303
通信接口
302

图3