(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 105659249 A
(43)申请公布日 2016.06.08

(21)申请号 201380078968.X
(22)申请日 2013.11.21
(30)优先权数据
61/867,546 2013.08.19 US
(85)PCT国际申请进入国家阶段日
2016.02.18
(86)PCT国际申请的申请数据
PCT/US2013/071290 2013.11.21
(87)PCT国际申请的公布数据
WO2015/026386 EN 2015.02.26
(71)申请人 汤姆逊许可公司
地址 法国伊西莱穆利欧市
(72)发明人 纳蒂亚·法瓦兹
阿巴萨利·马克杜米·卡克哈基
(74)专利代理机构 北京东方亿思知识产权代理有限责任公司 11258
代理人 李晓冬
(51)Int.Cl.
G06F 21/62(2006.01)

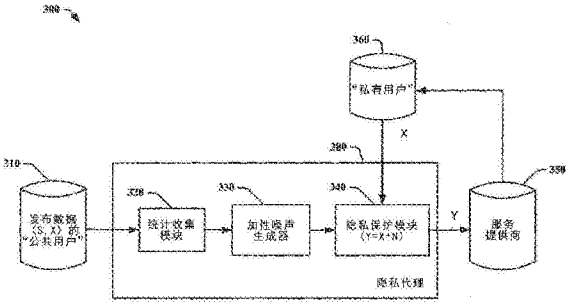
权利要求书2页 说明书9页 附图4页

(54)发明名称

用于通过加性噪声的效用感知隐私保护映射的方法和装置

(57)摘要

本实施例以用户遇到的隐私效用权衡为重点，该用户希望将一些公共数据（用X表示）发布至分析者以期得到一些效用，这些公共数据与他的私有数据（用S表示）相关。当噪声作为隐私保护机制被添加时（也就是说Y＝X+N，其中Y是实际发布至分析者的数据，并且N是噪声），我们示出针对连续数据X在L2范数失真下添加高斯噪声是最优的。我们用高斯机制表示添加最小化最坏情况下的信息泄露的高斯噪声的机制。高斯机制的参数基于X的协方差的特征向量和特征值来被确定。我们还开发了离散数据X的概率隐私保护映射机制，其中随机离散噪声遵循最大熵分布。



CN 105659249 A

1.一种用于处理用户的用户数据的方法，包括以下步骤：

访问所述用户数据，所述用户数据包括私有数据和公共数据，所述私有数据对应于第一类数据，并且所述公共数据对应于第二类数据；

确定(120)所述第一类数据的协方差矩阵；

响应于所述协方差矩阵生成(130)高斯噪声；

通过将所生成的高斯噪声添加至所述用户的所述公共数据来修改(140)所述公共数据；以及

将经修改的数据发布(150)至服务提供商和数据收集机构中的至少一个。

2.如权利要求1所述的方法，其中，所述公共数据包括所述用户已经指示可以被公开发布的数据，并且所述私有数据包括所述用户已经指示将不被公开发布的数据。

3.如权利要求1所述的方法，其中，生成高斯噪声的步骤包括以下步骤：

确定所述协方差矩阵的特征值和特征向量；以及

响应于所确定的特征值和特征向量分别确定另一特征值和另一特征向量，其中所述高斯噪声响应于所述另一特征值和另一特征向量被生成。

4.如权利要求1所述的方法，其中，所确定的另一特征向量实质上与所确定的所述协方差矩阵的特征向量相同。

5.如权利要求1所述的方法，其中，生成高斯噪声的步骤还响应于失真约束。

6.如权利要求1所述的方法，其中，生成高斯噪声的步骤包括独立于所述第二类数据的信息来生成。

7.如权利要求1所述的方法，还包括以下步骤：

基于发布的数据来接收服务。

8.一种用于处理用户的用户数据的方法，包括以下步骤：

访问所述用户数据，所述用户数据包括私有数据和公共数据；

访问(220)关于效用D的约束，所述效用响应于所述用户的所述公共数据和发布的数据；

响应于效用约束生成(230)随机噪声Z，所述随机噪声遵循所述效用约束下的最大熵概率分布；

将所生成的噪声添加(240)至所述用户的所述公共数据以生成所述用户的所述发布的数据；以及

将所述发布的数据发布(250)至服务提供商和数据收集机构中的至少一个。

9.如权利要求8所述的方法，其中，所述随机噪声遵循分布 $P[Z=i]=AB^{-|i|^p}$，其中A和B被选择为使得 $\sum_{i=-\infty}^{\infty} A B^{-|i|^p}=1$，其中p是整数。

10.如权利要求9所述的方法，其中，$\mathbb{E}[|Z|^p]^{\frac{1}{p}}=D$。

11.一种用于处理用户的用户数据的装置，包括：

统计收集模块(320)，所述统计收集模块被配置为确定所述用户数据的第一类数据的协方差矩阵，所述用户数据包括私有数据和公共数据，所述私有数据对应于所述第一类数据，并且所述公共数据对应于第二类数据；以及

加性噪声生成器(330),所述加性噪声生成器被配置为响应于所述协方差矩阵生成高斯噪声;以及

隐私保护模块(340),所述隐私保护模块被配置为

通过将所生成的高斯噪声添加至所述用户的所述公共数据来修改所述公共数据,以及

将所修改的数据发布至服务提供商和数据收集机构中的至少一个。

12.如权利要求11所述的装置,其中,所述公共数据包括所述用户已经指示可以被公开发布的数据,并且所述私有数据包括所述用户已经指示将不被公开发布的数据。

13.如权利要求11所述的装置,其中,加性噪声生成器(330)被配置为

确定所述协方差矩阵的特征值和特征向量,以及

响应于所确定的特征值和特征向量分别确定另一特征值和另一特征向量,其中所述高斯噪声响应于所述另一特征值和另一特征向量被生成。

14.如权利要求11所述的装置,其中,所确定的另一特征向量实质上与所确定的所述协方差矩阵的特征向量相同。

15.如权利要求11所述的装置,其中,所述加性噪声生成器被配置为响应于失真约束。

16.如权利要求11所述的装置,其中,所述加性噪声生成器独立于所述第二类数据的信息生成所述高斯噪声。

17.如权利要求11所述的装置,还包括处理器,所述处理器被配置为基于发布的数据接收服务。

18.一种用于处理用户的用户数据的装置,包括:

统计收集模块(320),所述统计收集模块被配置为访问关于效用D的约束,所述效用响应于所述用户的公共数据和发布的数据;

加性噪声生成器,所述加性噪声生成器被配置为响应于效用约束生成随机噪声Z,所述随机噪声遵循所述效用约束下的最大熵概率分布;

隐私保护模块(340),所述隐私保护模块被配置为

访问所述用户数据,所述用户数据包括私有数据和所述公共数据,

将所生成的噪声添加至所述用户的所述公共数据以生成所述用户的所述发布的数据,以及

将所述发布的数据发布至服务提供商和数据收集机构中的至少一个。

19.如权利要求18所述的装置,其中,所述随机噪声遵循分布$P[Z=i]=AB^{-|i|^p}$,其中A和B被选择为使得$\sum_{i=-\infty}^{\infty}AB^{-|i|^p}=1$,其中p是整数。

20.如权利要求19所述的装置,其中,$\mathbb{E}[|Z|^p]^{\frac{1}{p}}=D$。

21.一种其上存储有指令的计算机可读存储介质,所述指令用于根据权利要求1-10处理用户的用户数据。

用于通过加性噪声的效用感知隐私保护映射的方法和装置

[0001] 相关申请的交叉引用

[0002] 本申请要求下面的于2013年8月19日提交的序列号为61/867,546、题为“Method and Apparatus for Utility-Aware Privacy Preserving Mapping through Additive Noise(用于通过加性噪声的效用感知隐私保护映射的方法和装置)”的美国临时申请的申请日的权益,该临时申请出于所有目的通过引用以其整体合并于此。

[0003] 本申请涉及于2012年8月20日提交的序列号为61/691,090、题为“A Framework for Privacy against Statistical Inference(用于针对统计推断的隐私的框架)”的美国临时专利申请(以下简称“Fawaz”)。该临时申请明确地通过引用以其整体合并于此。

[0004] 此外,本申请涉及下面的申请:(1)代理案号PU130120、题为“Method and Apparatus for Utility-Aware Privacy Preserving Mapping against Inference Attacks(用于针对推断攻击的效用感知隐私保护映射的方法和装置)”,以及(2)代理案号PU130121、题为“Method and Apparatus for Utility-Aware Privacy Preserving Mapping in View of Collusion and Composition(用于鉴于协同和组合的效用感知隐私保护映射的方法和装置)”,这两个申请通过引用以其整体被指定、合并于此,并且被同时提交申请。

技术领域

[0005] 本发明涉及用于保护隐私的方法和装置,并且更具体地,涉及用于向用户数据添加噪声以保护隐私的方法和装置。

背景技术

[0006] 在大数据时代中,对用户数据的收集和挖掘已经成为大量私有和公共机构的快速增长且常见的实践。例如,技术公司利用用户数据向它们的客户提供个性化的服务,政府机构依赖数据来处理各种挑战(例如,国家安全、国家卫生、预算和资金分配),或医疗机构分析数据以发现疾病的起源和潜在的治愈方法。在某些情况下,第三方对用户数据的收集、分析、或共享是在未经用户的同意或察觉的情况下执行的。在其他情况下,数据由用户自愿地发布至特定分析者,以便获得服务作为回报,例如,发布产品评价以获得推荐。这种服务、或用户从允许访问用户数据中得到的其他好处可以被称为效用。在上述两种情况中的任意一种情况下,隐私风险出现,这是由于一些收集的数据可能被用户认为是敏感的(例如,政治观点、健康状况、收入水平),或可能乍看之下似乎无恶意(例如,产品评价)但是导致推断出与其相关的更敏感的数据。后者的威胁指代推断攻击,推断攻击是一种通过利用隐私数据与公开发布的数据的关联来推断隐私数据的技术。

发明内容

[0007] 本原理提供了用于处理用户的用户数据的方法,包括以下步骤:访问用户数据,该用户数据包括私有数据和公共数据,私有数据对应于第一类数据,并且公共数据对应于第

二类数据；确定第一类数据的协方差矩阵；响应于协方差矩阵生成高斯噪声；通过将所生成的高斯噪声添加至用户的公共数据来修改公共数据；以及将所修改的数据发布至如下面所描述的服务提供商和数据收集机构中的至少一个。本原理还提供用于执行这些步骤的装置。

[0008] 本原理还提供了用于处理用户的用户数据的方法，包括以下步骤：访问用户数据，该用户数据包括私有数据和公共数据；访问关于效用D的约束，该效用响应于用户的公共数据和发布的数据；响应于效用约束生成随机噪声Z，该随机噪声遵循效用约束下的最大熵概率分布；以及将所生成的噪声添加至用户的公共数据以生成用户的发布的数据(如下面所描述的)。本原理还提供用于执行这些步骤的装置。

[0009] 本原理还提供其上存储有指令的计算机可读存储介质，所述指令用于根据上面所描述的方法处理用户的用户数据。

附图说明

[0010] 图1是描绘根据本原理的实施例的用于通过将高斯噪声添加至连续数据来保护隐私的示例性方法的流程图。

[0011] 图2是描绘根据本原理的实施例的用于通过将离散噪声添加至离散数据来保护隐私的示例性方法的流程图。

[0012] 图3是描绘根据本原理的实施例的示例性隐私代理的框图。

[0013] 图4是描绘根据本原理的实施例的具有多个隐私代理的示例性系统的框图。

具体实施方式

[0014] 我们考虑Fawaz中所描述的设置，其中用户具有相互关联的两种数据：用户想要保持私有的一些数据，以及用户愿意发布至分析者的一些非私有数据(用户可以从该分析者得到一些效用，例如，向服务提供商发布媒体偏好以接收更精确的内容推荐)。

[0015] 本申请中所使用的术语分析者(例如，分析者可以是服务提供商的系统的一部分)指代发布的数据的接收者，其表面上使用数据以便向用户提供效用。分析者是发布的数据的合法接收者。然而，分析者也可能非法地利用发布的数据并且推断与用户的私有数据有关的一些信息。这造成了隐私与效用要求之间的紧张关系。为了减少推断威胁同时维护效用，用户可以发布根据条件概率映射生成的数据的“失真版本”，该条件概率映射被称为“隐私保护映射”(其在效用约束下被设计)。

[0016] 在本申请中，我们将用户想要保持私有的数据称为“私有数据”，将用户愿意发布的数据称为“公共数据”，以及将用户实际发布的数据称为“发布的数据”。例如，用户可能想要保持他的政治观点私有，并且愿意在修改的情况下发布他的TV评价(例如，用户对节目的实际评价是4，但是他发布的评价是3)。在这种情况下，用户的政治观点被认为是该用户的私有数据，TV评价被认为是公共数据，并且发布的修改的TV评价被认为是发布的数据。注意，另一用户可能愿意在不做修改的情况下发布政治观点和TV评价两者，因此针对该另一用户，当仅考虑政治观点和TV评价时，私有数据、公共数据、和发布的数据之间不存在区别。如果许多人发布政治观点和TV评价，则分析者可能能够得到政治观点与TV评价之间的关联，并且因此可能能够推断想要保持政治观点私有的用户的政治观点。

[0017] 关于私有数据，其指的是用户不仅指示它不应当被公开发布而且用户不希望它从用户将发布的其他数据中被推断出来的数据。公共数据是用户将允许隐私代理发布(可能以失真的方式发布以防止对私有数据的推断)的数据。

[0018] 在一个实施例中，公共数据是服务提供商为了向用户提供服务而从用户请求的数据。然而，用户在将它发布至服务提供商之前将使它失真(即，修改)。在另一实施例中，公共数据是用户指示为“公共的”(在这意义上，只要发布采取防止推断私有数据的形式，用户将不介意发布它)的数据。

[0019] 如上面所讨论的，具体类别的数据是否被认为是私有数据或公共数据是基于具体用户的角度。为了便于表示，我们从当前用户的角度将具体类别的数据称为私有数据或公共数据。例如，当尝试为想要保持他的政治观点私有的当前用户设计隐私保护映射时，我们针对当前用户和愿意发布他的政治观点的另一用户两者将政治观点称为私有数据。

[0020] 在本原理中，我们使用发布的数据与公共数据之间的失真作为效用的度量。当失真较大时，发布的数据更不同于公共数据，并且更多的隐私被保护，用户从失真的数据中得到的效用可能更少。另一方面，当失真较小时，发布的数据是公共数据的更加精确的表示，并且用户可以接收更多的效用，例如，接收更精确的内容推荐。

[0021] 在一个实施例中，为了针对统计推断保护隐私，我们对隐私效用权衡进行建模并且通过解决最小化信息泄露的优化问题来设计隐私保护映射，信息泄露被定义为私有数据与发布的数据之间的交互信息，服从失真约束。

[0022] 在Fawaz中，找到隐私保护映射依赖基本假设(链接私有数据和发布的数据的先验联合分布是已知的并且可以被提供为对优化问题的输入)。在实践中，真正的先验分布可能不是已知的，而是可以从可以被观测到的一组样本数据中估计某些先验统计。例如，可以从不关心隐私并且公开地发布不同类别的数据(可能被关心他们的隐私的用户认为是私有数据或公共数据)的一组用户中估计先验联合分布。可替代地，当不能观测到私有数据时，可以从仅发布他们的公共数据的一组用户中估计将要发布的公共数据的边缘分布或仅仅它的二阶统计。基于这组样本估计的统计然后被用来设计将被应用至关心其隐私的新用户的隐私保护映射机制。在实践中，估计的先验统计与真正的先验统计之间也可能存在不匹配，例如，由于可观测的样本的数量过小，或可观测的数据不完全。

[0023] 为了用公式表示问题，公共数据由具有概率分布P_X的随机变量$X \in \mathcal{X}$表示。X与私有数据相关，私有数据由随机变量$S \in \mathcal{S}$表示。S和X的关联由联合分布$P_{S,X}$定义。由随机变量$Y \in \mathcal{Y}$表示的发布的数据是X的失真版本。Y经由将X传递通过核$P_{Y|X}$被获得。在本申请中，术语“核”指概率性地将数据X映射数据Y的条件概率。也就是说，核$P_{Y|X}$是我们希望设计的隐私保护映射。由于Y是只有X的概率函数，在本申请中，我们假定S→X→Y形成马尔可夫(Markov)链。因此，一旦我们定义$P_{Y|X}$，我们就具有联合分布$P_{S,X,Y}=P_{Y|X}P_{S,X}$以及具体的联合分布$P_{S,Y}$。

[0024] 下面，我们首先定义隐私概念，然后定义精确度概念。

[0025] 定义1.假定S→X→Y。如果从联合分布$P_{S,X,Y}=P_{Y|X}P_{S,X}$导出的分布$P_{S,Y}$满足以下式子，则核$P_{Y|X}$被称为∈散度私有

[0026]

$$D(P_{S,Y}||P_S P_Y) \triangleq \mathbb{E}_{S,Y}\left[\log \frac{P(S|Y)}{P(S)}\right] \triangleq I(S;Y) = \epsilon H(S), \quad (1)$$

[0027] 其中D(.)是K-L散度，$\mathbb{E}(.)$随机变量的期望，H(.)是熵，∈∈[0,1]被称为泄露因子，并且交互信息I(S;Y)表示信息泄露。

[0028] 我们假设如果∈＝0则机制具有完全的隐私。在极端情况下，∈＝0意味着发布的随机变量Y独立于私有随机变量S，并且∈＝1意味着S完全可从Y恢复(S是Y的确定性函数)。注意，可以假定Y完全独立于S以具有完全的隐私(∈＝0)，但是这可能导致不良的精确度水平。我们定义精确度如下。

[0029] 定义2.令$\mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+$为失真度量。如果$\mathbb{E}[d(X,Y)] \leq D$则核$P_{Y|X}$被称为D精确的。

[0030] 隐私保护映射的泄露因子∈与失真等级D之间存在权衡。

[0031] 本原理提出了当仅有部分先前的统计知识是可用的时设计效用感知隐私保护映射机制的方法。更具体地，本原理提供了在加性噪声机制(其中噪声在公共数据被发布之前被添加至公共数据)的情况下的隐私保护映射机制。在分析中，我们假定噪声的均值为零。当均值不为零时该机制也可以被应用。在一个实施例中，因为熵对均值不敏感，所以针对非零均值的结果是相同的。该机制仅要求知道针对连续数据和离散数据二者将被发布的数据的二阶矩。

[0032] 高斯机制

[0033] 在一个实施例中，我们考虑连续公共数据X和通过将噪声添加至信号(即，Y＝X+N)实现的隐私保护映射方案。示例性连续公共数据可以是用户的身高或血压。在不知道P_X和$P_{S,X}$的情况下通过知道VAR(X)(或多维X的情况下的协方差矩阵)获得映射。首先，我们示出当噪声被添加至公共数据以保护隐私时在所有隐私保护映射机制中添加高斯噪声是最优的。

[0034] 由于S→X→Y，我们有I(S;Y)≤I(X;Y)。为了界定信息泄露I(S;Y)的界限，我们界定I(X;Y)。如果x＝f(s)是S的确定性函数，则I(S;Y)＝I(X;Y)并且界限是紧的(这例如在当针对某些矩阵A的X＝AS时的线性回归中发生)。

[0035] 令$X \in \mathbb{R}^n$。通过C_X表示X的协方差矩阵。令Y＝X+N，其中N是独立于X的噪声，N具有均值0和协方差矩阵C_N。注意，我们当仅有一个随机变量时使用方差(σ_X^2)的记法，并且当存在多个随机变量时使用协方差(C_N)的记法。我们得到下列结果。

[0036] 命题2.假定在设计隐私保护映射中P_X是未知的并且我们仅知道针对某些σ_X的$\mathrm{VAR}(X) \leq \sigma_X^2$。同样考虑通过将独立噪声N添加至信号X所获得的隐私保护方案的类别。噪声具有零均值和不大于针对某些σ_X的σ_N^2的方差(l_2范数失真)。我们示出，高斯噪声在下面的情况下是最优的：

[0037] $$\max_{P_X:X\perp N_G,\mathrm{VAR}(X)\leq\sigma_X^2} I(X;X+N_G) \leq \max_{P_X:X\perp N,\mathrm{VAR}(X)\leq\sigma_X^2} I(X;X+N), \quad (15)$$

[0038] 其中N_G表示高斯噪声并且N是随机变量，从而使得$\mathbb{E}[N_G]=\mathbb{E}[N]=0$和$\mathrm{VAR}(N_G)=\mathrm{VAR}(N)=\sigma_N^2$。这意味着，使用$N_G$的最坏情况下的信息泄露不大于使用N的最

坏情况下的信息泄露。

[0039] 证明:使用高斯鞍点定理,我们有

$$\max_{P_X:\, X\perp\!\!\!\perp N_G,\, \mathrm{VAR}(X)\le\sigma_X^2} I(X;X+N_G)$$

[0040]
$$= I(X_G;X_G+N_G) \le I(X_G;X_G+N)$$
$$\le \max_{P_X:\, X\perp\!\!\!\perp N,\, \mathrm{VAR}(X)\le\sigma_X^2} I(X;X+N), \qquad (16)$$

[0041] 其中X_G服从具有零均值和方差σ_N^2的高斯分布。这样完成证明。

[0042] 现在我们知道当噪声被添加以保护隐私时,添加高斯噪声在加性噪声族(在l_2范数失真约束下)中是最优解决方案。在下文中,我们确定将被添加至公共数据的高斯噪声的最优参数。我们通过高斯机制表示在这样的参数处添加高斯噪声的机制。

[0043] 在一个示例性实施例中,针对给定C_X和失真等级D,高斯机制通过如图1所示出的步骤进行。

[0044] 方法100在105处开始。在步骤110处,基于由用户发布的公共数据来估计统计信息,该用户不关心他们的公共数据或私有数据的隐私。我们将这些用户表示为“公共用户”,并且将关心他们的私有数据的隐私的用户表示为“私有用户”。

[0045] 统计可以通过抓取web被收集、可以通过访问不同的数据库被收集、或可以由数据聚合器(例如,bluekai.com)提供。哪些统计信息可以被收集依赖于公共用户发布的内容。注意,表征方差比表征边缘分布P_X要求更少的数据。因此,我们可能处于我们可以精确地估计方差但是不可以精确地估计边缘分布的情况。在一个示例中,我们在步骤120处可能仅能够基于收集的统计信息来获得公共数据的均值和方差(或协方差)。

[0046] 在步骤130处,我们进行协方差矩阵C_X的特征值分解。高斯噪声N_G的协方差矩阵具有与C_X的特征向量相同的特征向量。此外,C_N的相应的特征值可以通过求解下面的优化问题来给出

[0047]
$$\min_{\sigma_i:\, 1\le i\le n} \prod_{i=1}^{n} \frac{\sigma_i+\lambda_i}{\sigma_i}$$
$$\text{s.t. } \textstyle\sum_{i=1}^{n}\sigma_i \le D, \qquad (17)$$

[0048] 其中λ_is和σ_is($1 \le \lambda_i \le n$)分别表示C_X和C_N的特征值。从所确定的特征向量和特征值中,我们然后可以通过它的特征分解确定高斯噪声的协方差矩阵C_N。随后,我们可以生成高斯噪声 $N_G \sim \mathcal{N}(0, C_N)$ 。失真由 $\sum_{i=1}^{n}\mathbb{E}[(Y_i - X_i)^2] = \mathrm{tr}(C_N) = \sum_{i=1}^{n}\sigma_i \le D$ 给出,其中tr()表示对角元素的和,并且n是向量X的维度。

[0049] 在步骤140处,高斯噪声被添加至公共数据,也就是说,$Y = X + N_G$。在步骤150处,失真数据然后被发布至例如服务提供商或数据收集机构。方法100在步骤199处结束。

[0050] 在下面的定理中我们证明所提出的高斯机制在l_2范数失真约束下是最优的。

[0051] 定理3.假定l_2范数失真和给定的失真等级D,高斯机制中最小化交互信息的最优高斯噪声满足:

[0052] 最优噪声N_G的协方差矩阵具有与C_X的特征向量相同的特征向量。同样,特征值在

(17)中给出。

[0053] 证明:我们有

[0054] $$I(X;X+N)\leq\frac{1}{2}\log\left(\frac{|C_X+C_N|}{|C_N|}\right), \quad (18)$$

[0055] 其中不等式来自T.M.Cover和J.A.Thomas的书(Wiley-interscience在2012年出版的“信息论原理(Elements of information theory)”)中的定理8.6.5。由于我们不知道X的分布,我们必须最小化上界$\frac{1}{2}\log\left(\frac{|C_X+C_N|}{|C_N|}\right)$,因为它可以用高斯X实现。考虑半定矩阵$C_X$的特征值分解,从而获得$C_X=Q\Lambda Q^t$,其中$QQ^t=I$,并且$\Lambda$是包括$C_X$的特征值的对角矩阵。我们有$\mathbb{E}[\sum_{i=1}^{k}(Y_i-X_i)^2]=tr(C_N)=tr(Q^tC_NQ)=\sum\sigma_i\leq D$,并且优化问题变为

[0056] $$\min_{C_N}\frac{|\Lambda+Q^tC_NQ|}{|Q^tC_NQ|},\ \text{s.t. } tr(Q^tC_NQ)\leq D.$$

[0057] 在不失一般性的情况下,假定$\lambda_1\geq\cdots\geq\lambda_n$。令$\sigma_1\geq\cdots\geq\sigma_n$为$Q^tC_NQ^t$的特征值。根据M.Fiedler的文章(美国数学学会论文集(Proceedings of the American Mathematical Society)中的“厄米特矩阵的总和的行列式的界限(Bounds for the determinant of the sum of Hermitian matrices)”)的定理1,我们有$|\Lambda+Q^tC_NQ|\geq\Pi(\lambda_i+\sigma_i)$,并且如果$Q^tC_NQ$是对角矩阵则等式成立。因此,使用具有相同特征值$\sigma_i$的对角矩阵,我们实现相同失真等级和更小的泄露,这违背最优。因此,Q^tC_NQ是对角矩阵。

[0058] 示例3.X是S的确定性实值函数X=f(S),并且$\mathbf{VAR}(X)=\sigma_X^2$。因为S→X→Y,我们有I(X;Y)=I(S;Y)。令$N\sim\mathcal{N}(0,\sigma_N^2)$并且Y=X+N。对于任意∈,我们可以实现(∈,D)散度失真隐私,其中$D=\frac{\sigma_X^2}{e^{2\epsilon H(S)}-1}$。

[0059] 注意1.这个分析工作仅针对∈>0。一旦我们想要有完美的隐私,即,∈=0,则该方案选择$\sigma_N^2=\infty$。在实践中,这意味着,Y独立于X。如果我们假定$Y=\mathbb{E}[X]$(确定性值),则I(Y;S)=0并且$\mathbb{E}[d(X,Y)]=\mathbf{VAR}(X)$。因此,对于大于或等于VAR(X)的失真等级,设置$Y=\mathbb{E}[X]$的确定性机制实现∈=0。

[0060] 示例5.可以表明,通过添加具有方差$\sigma_N{}^2\geq\frac{1}{\epsilon^2}2\log(2/\delta)$的高斯噪声我们可以实现(∈,δ)差分隐私。这个方案导致失真$D\geq\frac{1}{\epsilon^2}2\log(2/\delta)$,并且信息的泄露$L\leq\frac{1}{2}\log\left(1+\frac{\sigma_X^2}{\frac{1}{\epsilon^2}2\log(2/\delta)}\right)$。用于比较的定性方式指出:使用(∈,δ)差分隐私高斯机制,我们将要求大的失真以实现小的泄露。另一方面,根据本原理使用散度隐私高斯机制,泄露L位的方案具有最小失真D,实现任意(∈,δ)差分隐私,其中$\frac{1}{\epsilon^2}2\log\left(\frac{2}{\delta}\right)=\frac{\sigma_X^2}{e^{2L}-1}$。

[0061] 离散机制

[0062] 在另一实施例中,我们考虑离散随机变量X,其中$\mathcal{X}=\mathbb{Z}$。再次,我们界定I(X;Y)以便界定I(S;Y)。令失真度量为l_p范数,即,X和Y之间的失真对于某些$1\leq p\leq\infty$为$\mathbb{E}[|X-Y|^p]^{\frac{1}{p}}$。

[0063] 定义5.对于给定$1 \le p \le \infty$,在具有小于或等于给定D的l_p范数的所有随机变量中,用$P^*_{p,D}$表示具有最大熵的分布。更正式地,$P^*_{p,D}$是在下面的优化中实现最大目标函数的概率度量

[0064] $$\max_{P_Z: Z\sim P_Z} H(Z),\ \text{s.t.}\ \mathbb{E}[|Z|^p]^{\frac{1}{p}} \le D.$$

[0065] 也就是说,优化问题是找到最大熵离散概率分布$P^*_{p,D}$,服从p^{th}矩的约束。最大熵由$H^*(p,D)$表示。

[0066] 接着,我们表征$P^*_{p,D}$和它的熵。

[0067] 命题3.对于任意$1 \le p \le \infty$,$P^*_{p,D}$由$P^*_{p,D}[Z=i] = AB^{-|i|^p}$给出,其中A和B被选择为使得$\sum_{i=-\infty}^{\infty} AB^{-|i|^p} = 1$并且$\mathbb{E}[|Z|^p]^{\frac{1}{p}} = D$。此外,我们有$H^*(p,D) = -\log A + (\log B)D^p$。

[0068] 证明:令$Z \sim AB^{-|i|^p}$和$W \sim P_W$,从而使得$\mathbb{E}[|W|^p]^{\frac{1}{p}} \le D$。由于$\mathbb{E}_{P_W}[|i|^p] \le D^p$,我们有

$$0 \le D(P_W||P_Z) = \mathbb{E}_{P_W}[\log \frac{P_W}{P_Z}]$$

[0069]
$$= -H(W) - \mathbb{E}_{P_W}[\log A - (\log B)|i|^p]$$
$$\le -H(W) - \log A + (\log B)\mathbb{E}_{P_Z}[|i|^p]$$
$$= -H(W) + H(Z).$$

[0070] 因此,$H(Z) \ge H(W)$,并且$H^*(p,D) = -\log A + (\log B)D^p$。

[0071] 我们用离散机制表示将噪声$Z \sim P^*_{p,D}$添加至离散公共数据的机制。在一个示例性实施例中,离散机制通过如图2所示出的步骤进行。

[0072] 方法200在205处开始。在步骤210处,访问参数(例如,p和D),从而定义失真度量。对于给定失真度量l_p($1 \le p \le \infty$)和失真等级D,在步骤220处计算如命题3中给出的概率度量$P^*_{p,D}$。注意,分布$P^*_{p,D}$仅由p和D确定,但是因为失真约束将隐私与精确度耦合,所以产生的隐私精确度权衡将依赖于X。

[0073] 在步骤230处,噪声在它在步骤240处被添加至公共数据(也就是说,Y=X+Z,其中$Z \sim P^*_{p,D}$)之前根据概率度量被生成。我们有$(X,Y) = \mathbb{E}[|Y-X|^p]^{\frac{1}{p}} = \mathbb{E}[|Z|^p]^{\frac{1}{p}} \le D$。方法200在步骤299处结束。

[0074] 接着,我们分析交互信息I(X;Y)。令$||X||_p = \mathbb{E}[|X|^p]^{\frac{1}{p}}$表示X的$l_p$范数。使用闵可夫斯基(Minkowski)不等式,我们得到$\mathbb{E}[|Y|^p]^{\frac{1}{p}} = \mathbb{E}[|X+Z|^p]^{\frac{1}{p}} \le \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Z|^p]^{\frac{1}{p}} = ||X||_p + D$。因此,我们获得

[0075] $I(S;Y) \le I(X;Y) = H(X+Z) - H(Z) \le H^*(p, ||X||_p + D) - H^*(p,D)$.

[0076] 也就是说,当使用离散机制时我们获得的隐私保证(即,信息泄露)由右项(其依赖

于D和X的平均l_p范数二者)界定上界。

[0077] 加性噪声技术的优点是,它不仅不要求许多关于X的统计的信息(更不必说关于S的信息),而且它还将优化问题简化为我们需要设计的是噪声的参数而不是必须指定完整核$P_{Y|X}$的简单问题。这大大简化了优化的尺寸,以及它的复杂性和用于解决它的计算/存储要求。

[0078] 有利的是,在不知道联合概率分布$P_{S,X}$并且只知道公共数据X的一阶矩和二阶矩的情况下,本原理提供通过将噪声添加至公共数据(针对连续数据和离散数据两者)的保护隐私的隐私保护映射机制。

[0079] 隐私代理是向用户提供隐私服务的实体。隐私代理可以执行下面各项的任意项:

[0080] -从用户接收他将何种数据视为私有的、他将何种数据视为公共的、以及他想要何种等级的隐私;

[0081] -计算隐私保护映射;

[0082] -实现针对用户的隐私保护映射(即,根据映射使他的数据失真);以及

[0083] -发布失真的数据(例如,向服务提供商或数据收集机构发布)。

[0084] 可以在保护用户数据的隐私的隐私代理中使用本原理。图3描绘了隐私代理在其中可以被使用的示例性系统300的框图。公共用户310发布他们的私有数据(S)和/或公共数据(X)。如前面所讨论的,公共用户可以按原样发布公共数据,也就是说,Y=X。由公共用户发布的信息变为对隐私代理有用的统计信息。

[0085] 隐私代理380包括统计收集模块320、加性噪声生成器330、和隐私保护模块340。统计收集模块320可以被用来收集公共数据的协方差。统计收集模块320还可以从数据聚合器(例如bluekai.com)接收统计。根据可用的统计信息,加性噪声生成器330例如基于高斯机制或离散机制来设计噪声。隐私保护模块340通过添加生成的噪声使私有用户360的公共数据在它被发布之前失真。在一个实施例中,统计收集模块320、加性噪声生成器330、和隐私保护模块340可以分别被用来执行方法100中的步骤110、130、和140。

[0086] 注意,隐私代理仅需要统计以进行工作而不需要知道数据收集模块中被收集的全部的数据。因此,在另一实施例中,数据收集模块可以是收集数据、并且然后计算统计、并且不需要是隐私代理的一部分的独立模块。数据收集模块与隐私代理共享统计。在一个实施例中,加性噪声生成器330和隐私保护模块340可以分别被用来执行方法200中的步骤220和230。

[0087] 隐私代理位于用户与用户数据的接收者(例如,服务提供商)之间。例如,隐私代理可以位于用户设备(例如,计算机、或机顶盒(STB))处。在另一示例中,隐私代理可以是单独的实体。

[0088] 隐私代理的所有模块可以位于一个设备处,或可以被分布在不同设备上,例如,统计收集模块320可以位于数据聚合器处,该数据聚合器仅向模块330发布统计,加性噪声生成器330可以位于“隐私服务提供商”处或位于连接至模块320的用户设备上的用户端处,并且隐私保护模块340可以位于隐私服务提供商处(然后该隐私保护模块340用作用户与服务提供商之间的中介,其中用户希望将数据发布至该服务提供商)或位于用户设备上的用户端处。

[0089] 隐私代理可以向服务提供商(例如,康卡斯特或Netflix)提供发布的数据,以便私

有用户360基于发布的数据改善接收到的服务，例如，推荐系统基于用户发布的电影评价向用户提供电影推荐。

[0090] 在图4中，我们示出了在系统中存在多个隐私代理。在不同变型中，不需要到处存在隐私代理，因为这不是对于隐私系统进行工作的要求。例如，可以仅在用户设备处存在隐私代理、或仅在服务提供商处存在、或在两者处都存在。在图4中，我们示出了针对Netflix和Facebook二者的相同隐私代理“C”。在另一实施例中，Facebook和Netflix处的隐私代理可以是但不必要是相同的。

[0091] 例如，可以以方法或过程、装置、软件程序、数据流、或信号来实现本文所描述的实现方式。虽然仅在单一形式的实现方式的上下文中讨论(例如，仅用方法进行讨论)，但是也可以以其他的形式(例如，装置或程序)实现所讨论的特征的实现方式。例如，可以以适当的硬件、软件、和固件实现装置。例如，可以在装置中实现方法，装置例如是处理器，其一般地指处理设备，包括例如，计算机、微处理器、集成电路、或可编程逻辑设备。处理器还可以包括通信设备，例如，计算机、蜂窝电话、便携式/个人数字助理(“PDA”)、以及促进终端用户之间的信息的传送的其他设备。

[0092] 对本原理的“一个实施例”或“实施例”或“一个实现方式”或“实现方式”以及它们的其他变型的提及表示结合实施例所描述的特定特征、结构、特性等等被包括在本原理的至少一个实施例中。因此，短语“在一个实施例中”或“在实施例中”或“在一个实现方式中”或“在实现方式中”的出现以及在贯穿说明书在各个地方出现的任意其他变型不一定都指同一实施例。

[0093] 此外，本申请或它的权利要求书可以指“确定”各种信息。确定信息可以包括以下各项中的一个或多个：例如，估计信息、计算信息、预测信息、或从存储器取回信息。

[0094] 此外，本申请或它的权利要求书可以指“访问”各种信息。访问信息可以包括以下各项中的一个或多个：例如，接收信息、取回信息(例如，从存储器)、存储信息、处理信息、发送信息、移动信息、复制信息、擦除信息、计算信息、确定信息、预测信息、或估计信息。

[0095] 此外，本申请或它的权利要求书可以指“接收”各种信息。接收以及“访问”旨在是广义的术语。接收信息可以包括以下各项中的一个或多个：例如，访问信息、或取回信息(例如，从存储器)。此外，“接收”通常在操作期间以一种方式或另一种方式涉及：例如，存储信息、处理信息、发送信息、移动信息、复制信息、擦除信息、计算信息、确定信息、预测信息、或估计信息。

[0096] 对本领域的技术人员将是明显的是，实现方式可以产生被格式化以携带信息(该信息例如可以被存储或被发送)的各种信号。信息可以包括例如用于执行方法的指令或由所描述的实现方式中的一个产生的数据。例如，信号可以被格式化以携带所描述的实施例的比特流。这样的信号例如可以被格式化为电磁波(例如，使用频谱的射频部分)或基带信号。例如，格式化可以包括编码数据流和用编码的数据流调制载波。信号携带的信息例如可以是模拟信息或数字信息。如公知的，可以通过各种不同的有线的链路或无线链路来发送信号。信号可以被存储在处理器可读介质上。

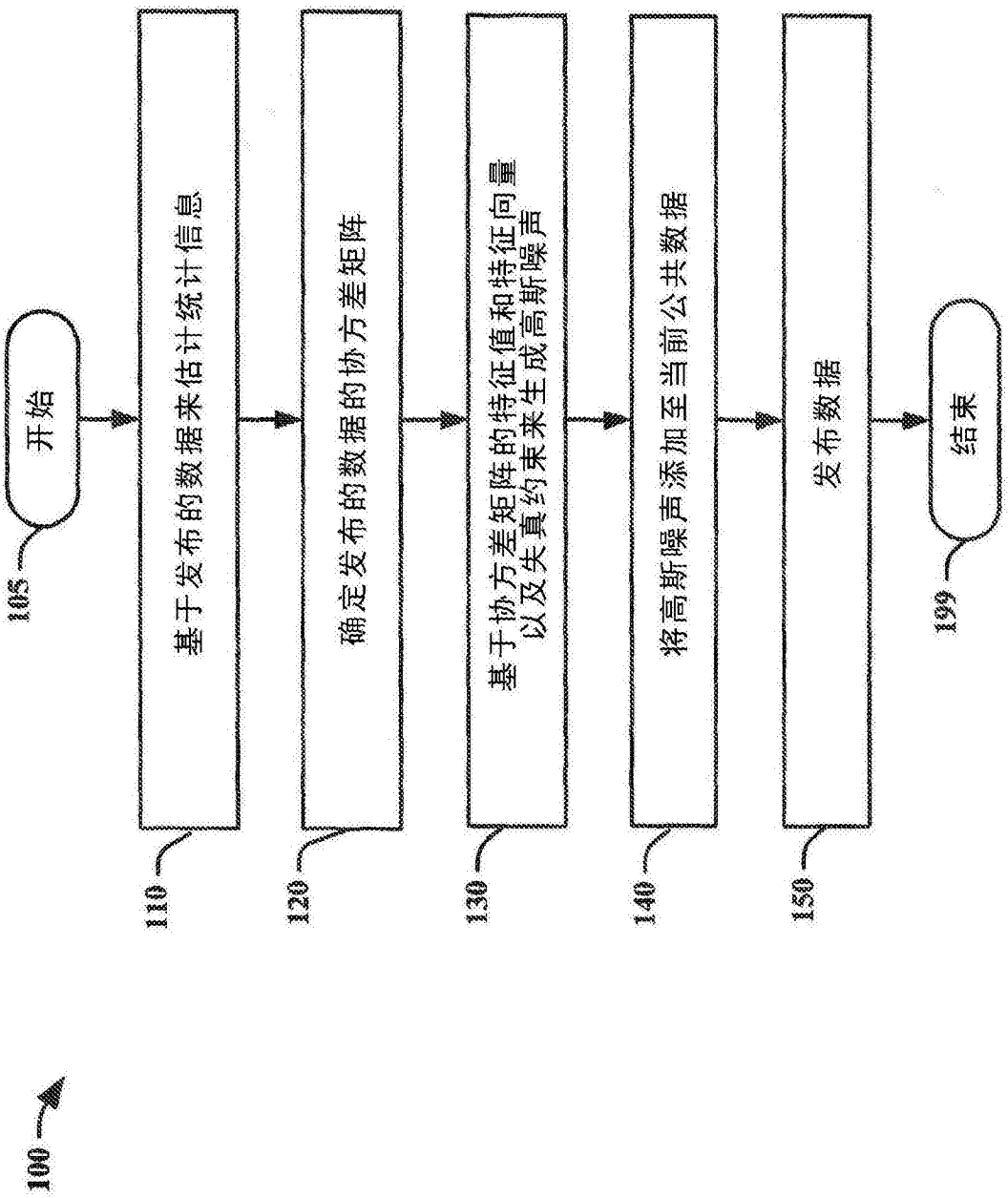

100
105
开始
110
基于发布的数据来估计统计信息
120
确定发布的数据的协方差矩阵
130
基于协方差矩阵的特征值和特征向量
以及失真约束来生成高斯噪声
140
将高斯噪声添加至当前公共数据
150
发布数据
199
结束

图1

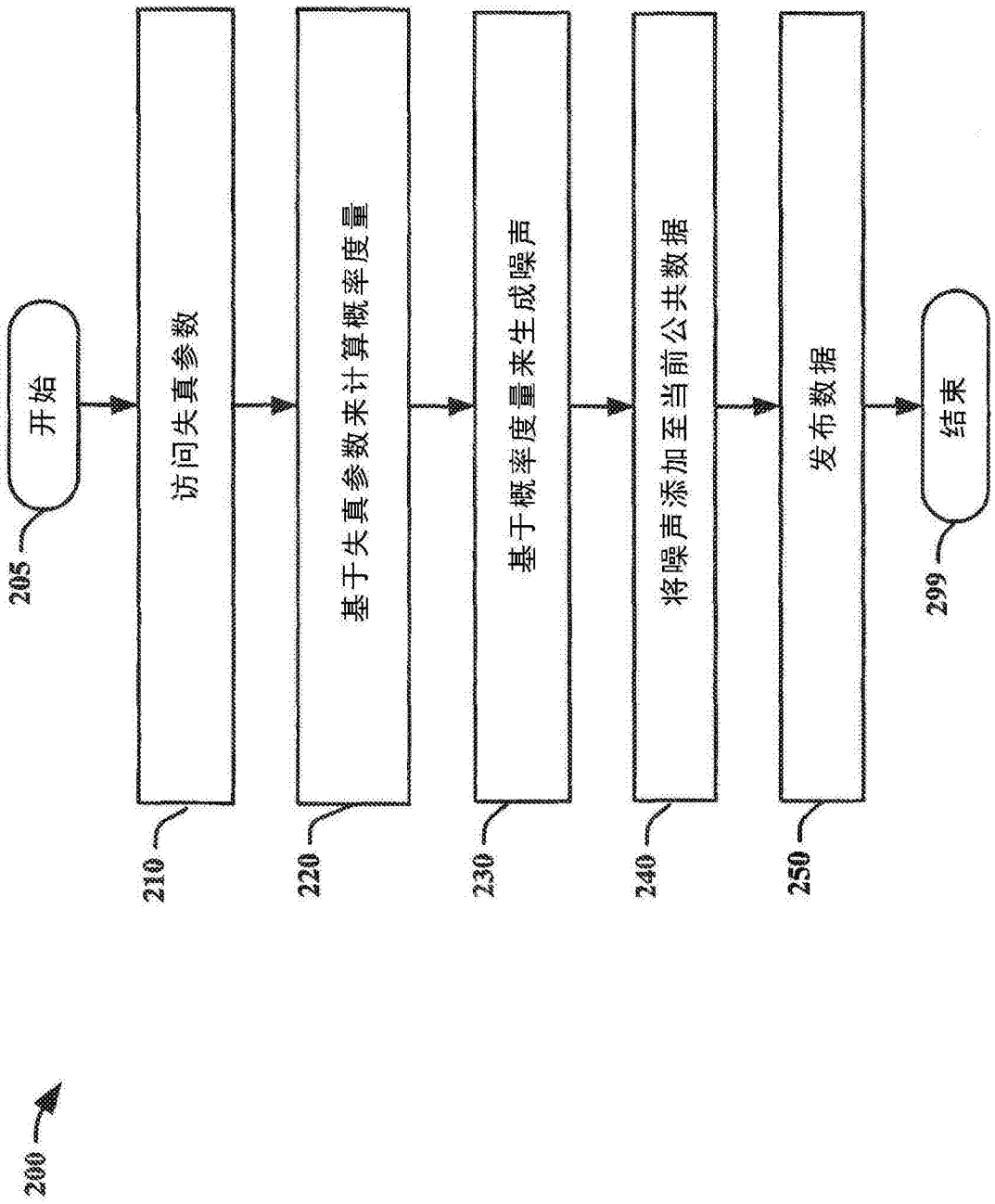

开始
205
访问失真参数
210
基于失真参数来计算概率度量
220
基于概率度量来生成噪声
230
将噪声添加至当前公共数据
240
发布数据
250
结束
299
200

图2

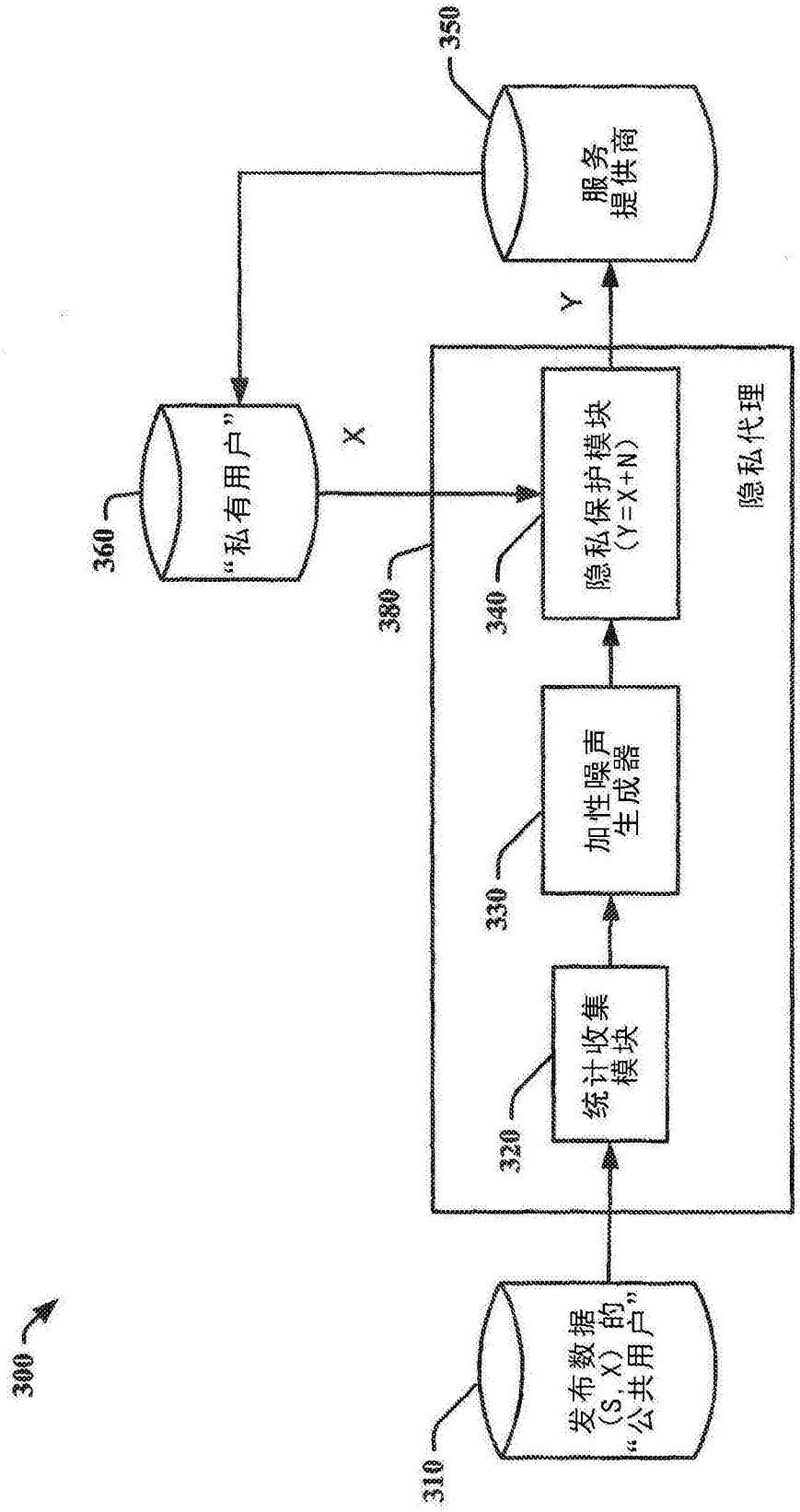

300
310
发布数据
(S, X)
"公共用户"
320
统计收集
模块
330
加性噪声
生成器
340
隐私保护模块
(Y=X+N)
隐私代理
350
服务
提供商
360
"私有用户"
380
X
Y

图3

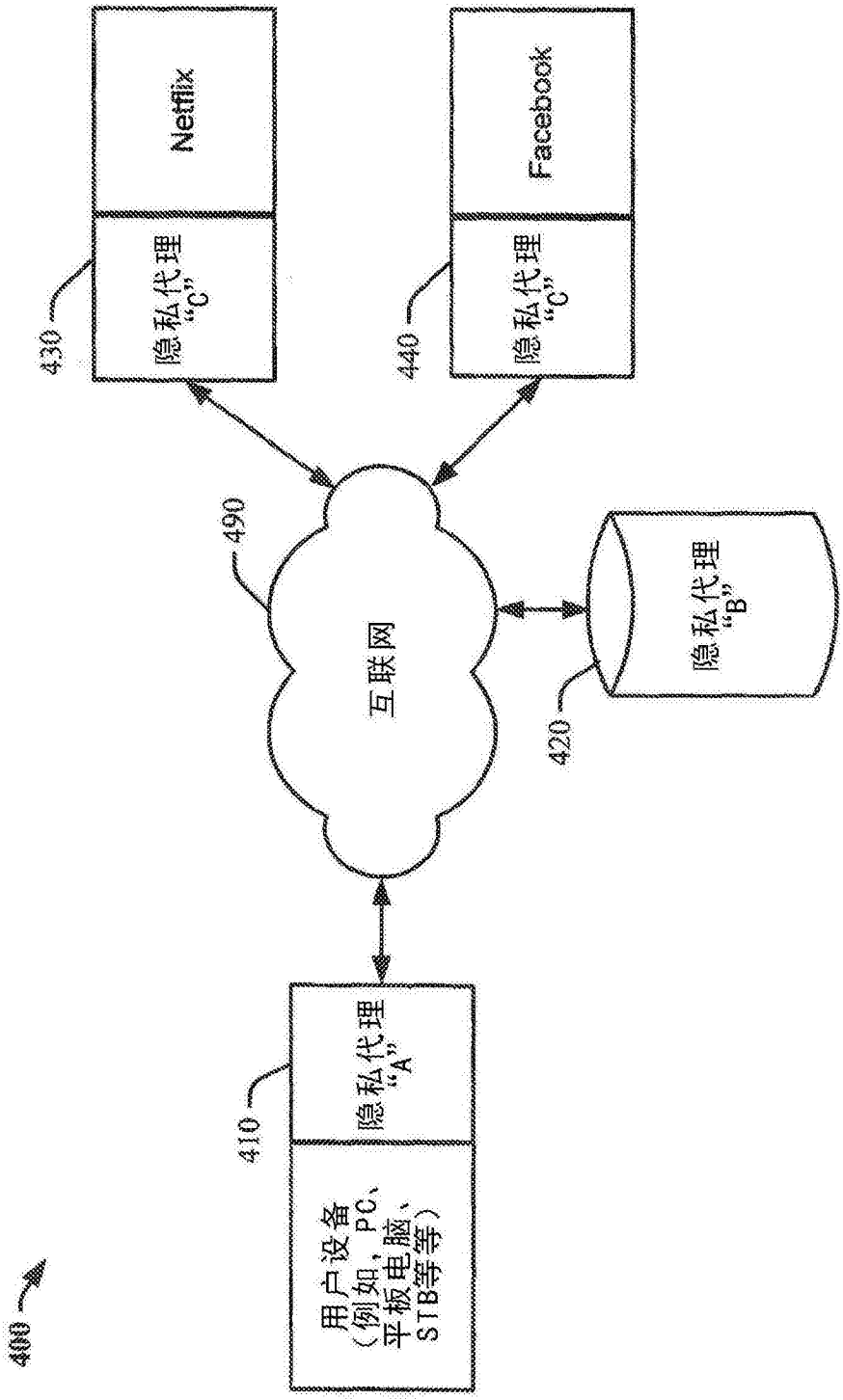

Netflix
隐私代理“C”
430
Facebook
隐私代理“C”
440
490
互联网
隐私代理“B”
420
410
隐私代理“A”
用户设备（例如，PC、平板电脑、STB等等）
400

图4