(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2016-531513
(P2016-531513A)
(43)公表日　平成28年10月6日(2016.10.6)

(51)Int.Cl.			F I			テーマコード（参考）
H04L	9/20	(2006.01)	H04L	9/00	653	5J104
G09C	1/00	(2006.01)	G09C	1/00	660D	

審査請求 未請求　予備審査請求 未請求　（全 25 頁）

(21)出願番号　特願2016-536079 (P2016-536079)
(86)(22)出願日　平成25年11月21日 (2013.11.21)
(85)翻訳文提出日　平成28年4月19日 (2016.4.19)
(86)国際出願番号　PCT/US2013/071290
(87)国際公開番号　WO2015/026386
(87)国際公開日　平成27年2月26日 (2015.2.26)
(31)優先権主張番号　61/867,546
(32)優先日　平成25年8月19日 (2013.8.19)
(33)優先権主張国　米国 (US)

(71)出願人　501263810
トムソン　ライセンシング
Ｔｈｏｍｓｏｎ　Ｌｉｃｅｎｓｉｎｇ
フランス国，　９２１３０　イツシー　レ　ムーリノー，　ル　ジヤンヌ　ダルク，　１－５
１－５，　ｒｕｅ　Ｊｅａｎｎｅ　ｄ’Ａｒｃ，　９２１３０　ＩＳＳＹ　ＬＥＳ　ＭＯＵＬＩＮＥＡＵＸ，　Ｆｒａｎｃｅ
(74)代理人　110001243
特許業務法人　谷・阿部特許事務所
(72)発明者　ナディア　ファワズ
アメリカ合衆国　９５０５０　カリフォルニア州　サンタ　クララ　ベロミー　ストリート　１５３１

最終頁に続く

(54)【発明の名称】付加ノイズを用いる効用対応プライバシー保護写像のための方法および装置

(57)【要約】
本実施形態は、いくつかの効用を得ることを期待して、（Ｓにより示される）彼のプライベートデータと相関関係がある、（Ｘにより示される）いくつかのパブリックデータを分析者に公開することを願うユーザにより直面されるプライバシー効用トレードオフに焦点を当てる。ノイズがプライバシー保護機構として追加されるとき、すなわちＹが分析者に実際に公開されるデータでありＮがノイズであってＹ＝Ｘ＋Ｎであるとき、我々はガウスノイズを追加することがl_2ノルム歪み下で連続データＸに関して最適であることを示す。我々はガウスノイズを追加する機構がガウス機構による最悪の場合の情報漏出を最小化することを示す。ガウス機構のパラメータはＸの共分散の固有ベクトル及び固有値に基づき決定される。離散データＸに関する確率的プライバシー保護写像機構も開発し、ランダム離散ノイズは最大エントロピー分散に従う。

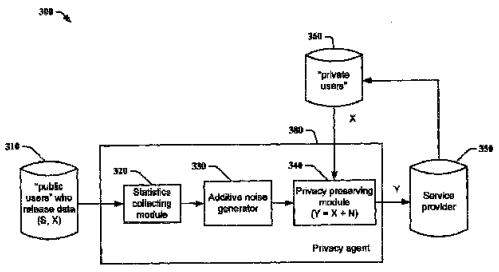


FIG. 3

【特許請求の範囲】
【請求項１】
ユーザに関するユーザデータを処理するための方法であって、
プライベートデータおよびパブリックデータを含む前記ユーザデータにアクセスするステップであって、前記プライベートデータは、第１のカテゴリのデータに対応し、前記パブリックデータは、第２のカテゴリのデータに対応する、前記ステップと、
前記第１のカテゴリのデータの共分散行列を決定するステップ（１２０）と、
前記共分散行列に応じてガウスノイズを生成するステップ（１３０）と、
前記生成されたガウスノイズを前記ユーザの前記パブリックデータに加えることによって、前記パブリックデータを修正するステップ（１４０）と、
前記修正されたデータを、サービスプロバイダとデータ収集エージェンシとの少なくとも一方へ公開するステップ（１５０）と、
を含む、前記方法。
【請求項２】
前記パブリックデータは、公然と公開されることを前記ユーザが示すデータを含み、前記プライベートデータは、公然と公開されるべきでないことを前記ユーザが示すデータを含む、請求項１に記載の方法。
【請求項３】
ガウスノイズを生成する前記ステップは、
前記共分散行列の固有値及び固有ベクトルを決定するステップと、
前記決定された固有値及び固有ベクトルに応じて他の固有値及び固有ベクトルをそれぞれ決定するステップであって、前記ガウスノイズは、前記他の固有値及び固有ベクトルに応じて生成される、前記ステップと、
を含む、請求項１に記載の方法。
【請求項４】
前記決定された他の固有ベクトルは、前記共分散行列の前記決定された固有ベクトルと実質的に同一である、請求項１に記載の方法。
【請求項５】
ガウスノイズを生成する前記ステップは、歪み制約にさらに応じる、請求項１に記載の方法。
【請求項６】
ガウスノイズを生成する前記ステップは、前記第２のカテゴリのデータの情報とは独立して生成するステップを含む、請求項１に記載の方法。
【請求項７】
前記公開されたデータに基づいてサービスを受信するステップをさらに含む、請求項１に記載の方法。
【請求項８】
ユーザに関するユーザデータを処理するための方法であって、
プライベートデータおよびパブリックデータを含む前記ユーザデータにアクセスするステップと、
効用Ｄに対する制約にアクセスするステップ（２２０）であって、前記効用は、前記ユーザの前記パブリックデータおよび公開されるデータに応じる、ステップと、
前記効用に対する制約に応じてランダムノイズＺを生成するステップ（２３０）であって、前記ランダムノイズは、前記効用に対する制約下の最大エントロピー確率分布に従う、前記ステップと、
前記生成されたノイズを前記ユーザの前記パブリックデータに加えて、前記ユーザに関する前記公開されるデータを生成するステップ（１４０）と、
前記公開されるデータを、サービスプロバイダとデータ収集エージェンシとの少なくとも一方へ公開するステップ（１５０）と、
を含む、前記方法。

【請求項９】
前記ランダムノイズは、分布
Ｐ[Ｚ＝ｉ]
に従い、Ａ及びＢは、
【数１】

$$\sum_{i=-\infty}^{\infty} A B^{-|i|^p} = 1$$

となるように選択され、ｐは整数である、請求項８に記載の方法。
【請求項１０】
【数２】

$$\mathbb{E}[|Z|^p]^{\frac{1}{p}} = D$$

である、請求項９に記載の方法。
【請求項１１】
ユーザに関するユーザデータを処理するための装置であって、
プライベートデータおよびパブリックデータを含む前記ユーザデータの第１のカテゴリのデータの共分散行列を決定するように構成された統計収集モジュール（３２０）であって、前記プライベートデータは、前記第１のカテゴリのデータに対応し、前記パブリックデータは、第２のカテゴリのデータに対応する、前記統計収集モジュールと、
前記共分散行列に応じてガウスノイズを生成するように構成された付加ノイズ生成器（３３０）と、
プライバシー保護モジュール（３４０）であって、
前記生成されたガウスノイズを前記ユーザの前記パブリックデータに加えることによって、前記パブリックデータを修正し、
前記修正されたデータを、サービスプロバイダとデータ収集エージェンシとの少なくとも一方へ公開する、
ように構成された、前記プライバシー保護モジュールと、
を含む、前記装置。
【請求項１２】
前記パブリックデータは、公然と公開されることを前記ユーザが示すデータを含み、前記プライベートデータは、公然と公開されるべきでないことを前記ユーザが示すデータを含む、請求項１１に記載の装置。
【請求項１３】
前記付加ノイズ生成器（３３０）は、
前記共分散行列の固有値及び固有ベクトルを決定し、
前記決定された固有値及び固有ベクトルに応じて他の固有値及び固有ベクトルをそれぞれ決定する、
ように構成され、前記ガウスノイズは、前記他の固有値及び固有ベクトルに応じて生成される、請求項１１に記載の装置。
【請求項１４】
前記決定された他の固有ベクトルは、前記共分散行列の前記決定された固有ベクトルと実質的に同一である、請求項１１に記載の装置。
【請求項１５】
前記付加ノイズ生成器は、歪み制約に応じるように構成されている、請求項１１に記載の装置。
【請求項１６】
前記付加ノイズ生成器は、前記ガウスノイズを前記第２のカテゴリのデータの情報とは独立して生成するように構成されている、請求項１１に記載の装置。
【請求項１７】
前記公開されたデータに基づいてサービスを受信するように構成されたプロセッサをさ

らに含む、請求項１１に記載の装置。
【請求項１８】
ユーザに関するユーザデータを処理するための装置であって、
効用Ｄに対する制約にアクセスするように構成された統計収集モジュール（３２０）であって、前記効用は、前記ユーザのパブリックデータおよび公開されるデータに応じる、前記統計収集モジュールと、
前記効用に対する制約に応じてランダムノイズＺを生成するように構成された付加ノイズ生成器であって、前記ランダムノイズは、前記効用に対する制約下の最大エントロピー確率分布に従う、前記付加ノイズ生成器と、
プライバシー保護モジュール（３４０）であって、
プライベートデータおよび前記パブリックデータを含む前記ユーザデータにアクセスし、
前記生成されたノイズを前記ユーザの前記パブリックデータに加えて、前記ユーザに関する前記公開されるデータを生成し、
前記公開されるデータを、サービスプロバイダとデータ収集エージェンシとの少なくとも一方へ公開する、
ように構成された、前記プライバシー保護モジュールと、
を含む、前記装置。
【請求項１９】
前記ランダムノイズは、分布
Ｐ[Ｚ＝ｉ]
に従い、Ａ及びＢは、
【数３】

$$\sum_{i=-\infty}^{\infty} A B^{-|i|^p} = 1$$

となるように選択され、ｐは整数である、請求項１８に記載の装置。
【請求項２０】
【数４】

$$\mathbb{E}[|Z|^p]^{\frac{1}{p}} = D$$

である、請求項１９に記載の装置。
【請求項２１】
請求項１から１０のいずれかに記載の方法による、ユーザに関するユーザデータを処理するための命令を格納したコンピュータ可読記憶媒体。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、プライバシーを保護するための方法および装置に関し、より詳細には、ユーザデータにノイズを加えてプライバシーを保護するための方法および装置に関する。
【背景技術】
【０００２】
関連出願の相互参照
本出願は、あらゆる目的で参照によりその全体が本明細書に組み込まれる、２０１３年８月１９日に出願された米国特許仮出願第６１／８６７，５４６号明細書、名称「Ｍｅｔｈｏｄ ａｎｄ Ａｐｐａｒａｔｕｓ ｆｏｒ Ｕｔｉｌｉｔｙ－Ａｗａｒｅ Ｐｒｉｖａｃｙ Ｐｒｅｓｅｒｖｉｎｇ Ｍａｐｐｉｎｇ ｔｈｒｏｕｇｈ Ａｄｄｉｔｉｖｅ Ｎｏｉｓｅ」の出願日の利益を主張する。
【０００３】
本出願は、２０１２年８月２０日に出願された米国特許仮出願第６１／６９１，０９０号明細書、名称「Ａ Ｆｒａｍｅｗｏｒｋ ｆｏｒ Ｐｒｉｖａｃｙ ａｇａｉｎｓｔ

Ｓｔａｔｉｓｔｉｃａｌ Ｉｎｆｅｒｅｎｃｅ」（以下、「Ｆａｗａｚ」）に関連付けられる。この仮出願は、明示的に参照によりその全体が本明細書に組み込まれる。

【０００４】

加えて、本出願は、以下の出願、（１）代理人整理番号第ＰＵ１３０１２０号、名称「Ｍｅｔｈｏｄ ａｎｄ Ａｐｐａｒａｔｕｓ ｆｏｒ Ｕｔｉｌｉｔｙ－Ａｗａｒｅ Ｐｒｉｖａｃｙ Ｐｒｅｓｅｒｖｉｎｇ Ｍａｐｐｉｎｇ ａｇａｉｎｓｔ Ｉｎｆｅｒｅｎｃｅ Ａｔｔａｃｋｓ」、および（２）代理人整理番号第ＰＵ１３０１２１号、名称「Ｍｅｔｈｏｄ ａｎｄ Ａｐｐａｒａｔｕｓ ｆｏｒ Ｕｔｉｌｉｔｙ－Ａｗａｒｅ Ｐｒｉｖａｃｙ Ｐｒｅｓｅｒｖｉｎｇ Ｍａｐｐｉｎｇ ｉｎ Ｖｉｅｗ ｏｆ Ｃｏｌｌｕｓｉｏｎ ａｎｄ Ｃｏｍｐｏｓｉｔｉｏｎ」に関連付けられ、これらは、同一譲受人に譲渡され、それら全体が参照により組み込まれ、本明細書と共に出願される。

【０００５】

ビッグデータの時代において、ユーザデータの収集およびマイニングは、多数の民間および公共機関により、急速に成長している一般的行為となっている。たとえば、技術会社は、ユーザデータを利用してそれらの顧客に個別化されたサービスを提供し、政府機関は、データに依拠して、様々な課題、たとえば、国家安全保障、国民健康、予算および基金配分に取り組み、または医療機関は、データを分析して病気の発端および潜在的治療法を発見する。場合によっては、ユーザのデータの収集、分析または第三者との共有は、ユーザの同意または認識なしに行われる。他の場合、データは、見返りにサービスを得るために、ユーザによって特定の分析者に自発的に公開され、たとえば、推奨を得るために製品評価が公開される。ユーザのデータにアクセスすることを許可することからユーザが得るこのサービスまたは他の利益は、効用（utility）と呼ばれることがある。いずれの場合も、プライバシーリスクが生じ、その理由は、収集されたデータの一部は、センシティブ（sensitive）であるとユーザがみなすことがある（たとえば、政治的意見、健康状態、所得レベル）、または、一見無害に見えることがある（たとえば、製品評価）が、それが相関付けられたよりセンシティブなデータの推定につながるからである。後者の脅威は、推論攻撃、すなわち、プライベートデータを、それの公然と公開されるデータ（released data）との相関を利用することによって推論する技法を指す。

【発明の概要】

【０００６】

本原理は、ユーザに関するユーザデータを処理するための方法であって、プライベートデータおよびパブリックデータを含む上記ユーザデータにアクセスするステップであって、上記プライベートデータは、第１のカテゴリのデータに対応し、上記パブリックデータは、第２のカテゴリのデータに対応する、ステップと、上記第１のカテゴリのデータの共分散行列を決定するステップと、上記共分散行列に応じてガウスノイズを生成するステップと、上記生成されたガウスノイズを上記ユーザの上記パブリックデータに加えることによって、上記パブリックデータを修正するステップと、後述されるように、上記修正されたデータを、サービスプロバイダとデータ収集エージェンシとの少なくとも一方へ公開するステップと、を含む、上記方法を提供する。本原理はまた、これらのステップを実施するための装置を提供する。

【０００７】

本原理はまた、ユーザに関するユーザデータを処理するための方法であって、プライベートデータおよびパブリックデータを含む上記ユーザデータにアクセスするステップと、効用Ｄに対する制約にアクセスするステップであって、上記効用は、上記ユーザの上記パブリックデータおよび公開されるデータに応じる、ステップと、上記効用制約に応じてランダムノイズＺを生成するステップであって、上記ランダムノイズは、上記効用制約下の最大エントロピー確率分布に従う、ステップと、後述されるように、上記生成されたノイズを上記ユーザの上記パブリックデータに加えて、上記ユーザに関する上記公開されるデータを生成するステップと、を含む、上記方法を提供する。本原理はまた、これらのステップを実行するための装置を提供する。

【０００８】
本原理はまた、上述された上記方法によるユーザに関するユーザデータを処理するための命令を格納したコンピュータ可読記憶媒体を提供する。
【図面の簡単な説明】
【０００９】
【図１】本原理の実施形態による、ガウスノイズを連続データに加えることによってプライバシーを保護するための例示的方法を示すフロー図である。
【図２】本原理の実施形態による、離散ノイズを離散データに加えることによってプライバシーを保護するための例示的方法を示すフロー図である。
【図３】本原理の実施形態による、例示的プライバシーエージェントを示すブロック図である。
【図４】本原理の実施形態による、複数のプライバシーエージェントを有する例示的システムを示すブロック図である。
【発明を実施するための形態】
【００１０】
我々は、Ｆａｗａｚに説明された設定を検討し、この設定では、ユーザが、相関付けられた２種類のデータを有し、すなわち、ユーザがプライベートのままにしたい一部のデータと、ユーザが分析者に積極的に公開し、そこからユーザがある種の効用を得られ得る、たとえば、メディア選好をサービスプロバイダに公開して、より精密なコンテンツ推奨を受け取る、一部の非プライベートデータとを有する。
【００１１】
本出願で使用されるとき、分析者という用語は、たとえば、サービスプロバイダのシステムの一部であってよく、ユーザへの効用を提供するためにデータを表向き使用する、公開されるデータの受信者を指す。分析者は、公開されるデータの正当な受信者である。しかしながら、分析者が、公開されるデータを不法に利用し、ユーザのプライベートデータに関するある種の情報を推論する可能性もあり得る。これは、プライバシーと効用の必要条件の間の緊張をもたらす。効用を保持しながら推論脅威を低減するために、ユーザは、効用制約下で設計された「プライバシー保護写像（privacy preserving mapping）」と呼ばれる条件付き確率的写像によって生成されたデータの「歪められたバージョン」を公開することができる。
【００１２】
本出願において、我々は、ユーザがプライベートのままにしたいデータを「プライベートデータ」と呼び、ユーザが積極的に公開しようとするデータを「パブリックデータ」と呼び、ユーザが実際に公開するデータを「公開されるデータ」と呼ぶ。たとえば、ユーザは、ユーザの政治的意見をプライベートに維持したいことがあり、修正を伴うユーザのＴＶ評価を積極的に公開する（たとえば、番組に関するユーザの実際の評価は４であるが、ユーザは評価を３として公開する）。この場合、ユーザの政治的意見は、このユーザのプライベートデータとみなされ、ＴＶ評価は、パブリックデータとみなされ、公開された修正されたＴＶ評価は、公開されるデータとみなされる。別のユーザが政治的意見とＴＶ評価の両方を修正なしに積極的に公開しようとすることもあり、したがって、この他のユーザの場合、政治的意見とＴＶ評価のみが考慮されるとき、プライベートデータ、パブリックデータ、および公開されるデータの間の区別がないことに留意されたい。多くの人々が政治的意見およびＴＶ評価を公開する場合、分析者は、政治的意見とＴＶ評価との間の相関を得ることができ、したがって、それをプライベートに維持したいユーザの政治的意見を推論できることがある。
【００１３】
プライベートデータに関して、これは、公然と公開されるべきでないことをユーザが示すのみならず、ユーザが公開することになる他のデータから推論されたくもないデータを指す。パブリックデータは、場合によってはプライベートデータの推論を防止するために歪められた方法で、プライバシーエージェントが公開することをユーザが許可するデータ

である。

【００１４】

一実施形態では、パブリックデータは、ユーザにサービスを提供するためにサービスプロバイダがユーザに要求するデータである。しかしながら、ユーザは、それをサービスプロバイダに公開する前に、それを歪める（すなわち修正する）。別の実施形態では、パブリックデータは、公開がプライベートデータの推論を防止する形態をとる限りはユーザがそれを公開するのを気にしないという意味で「パブリック」であるとユーザが示すデータである。

【００１５】

上述されたように、特定のカテゴリのデータがプライベートデータとみなされるかそれともパブリックデータとみなされるかは、特定のユーザの視点に基づく。表記を簡単にするために、我々は、特定のカテゴリのデータを、現在のユーザの視点から、プライベートデータまたはパブリックデータと呼ぶ。たとえば、自身の政治的意見をプライベートに維持したい現在のユーザに関するプライバシー保護写像を設計しようとするとき、我々は、現在のユーザと、自身の政治的意見を積極的に公開しようとする別のユーザとの両方に関して、政治的意見をプライベートデータと呼ぶ。

【００１６】

本原理では、我々は、公開されるデータとパブリックデータとの間の歪みを、効用の測度として使用する。歪みがより大きいとき、公開されるデータは、よりいっそうパブリックデータとは異なり、よりプライバシーが保護されるが、歪められたデータから得られる効用は、ユーザにとって、より小さくなり得る。他方で、歪みがより小さいとき、公開されるデータは、パブリックデータのより精密な表現となり、ユーザは、より大きい効用、たとえば、より精密なコンテンツ推奨を受け取る可能性がある。

【００１７】

一実施形態では、統計的推論に対してプライバシーを保護するために、我々は、歪み制約を受けるプライベートデータと公開されるデータとの間の相互情報量として定義される情報漏出を最小化する最適化問題を解くことにより、プライバシー効用トレードオフをモデリングし、プライバシー保護写像を設計する。

【００１８】

Ｆａｗａｚでは、プライバシー保護写像を求めることは、プライベートデータと公開されるデータとをリンクする事前同時分布（prior joint distribution）が知られていて最適化問題に対する入力として提供され得るという基本的仮定に依拠する。実際には、真の事前分布は知られていなくてよく、いくつかの事前統計値が、観測され得るサンプルデータのセットから推定されてよい。たとえば、事前同時分布は、プライバシーの懸念を持たず、異なるカテゴリのデータを公然と公開するユーザのセットから推定されてよく、それらのデータは、彼らのプライバシーを懸念するユーザによってプライベートデータまたはパブリックデータとみなされ得る。あるいは、プライベートデータが観測できないとき、公開されるべきパブリックデータの周辺分布、または単純にその２次統計値が、彼らのパブリックデータのみを公開するユーザのセットから推定され得る。このサンプルのセットに基づいて推定された統計値は、次いで、彼らのプライバシーを懸念する新しいユーザに適用されるプライバシー保護写像機構を設計するために使用される。実際には、たとえば、観測可能サンプルが少数のため、または観測可能データの不完全性のため、推定された事前統計値と真の事前統計値との間の不一致が存在することもある。

【００１９】

問題を定式化するために、パブリックデータは、確率分布P_xを有する確率変数

【００２０】

【数１】

$$X \in \mathcal{X}$$

【００２１】

によって示される。Ｘは、確率変数

【００２２】
【数２】
$$S \in \mathcal{S}$$
【００２３】
によって示されるプライベートデータと相関付けられる。ＳとＸの相関は、同時分布$P_{S,X}$によって定義される。確率変数
【００２４】
【数３】
$$Y \in \mathcal{Y}$$
【００２５】
によって示される公開されるデータは、Ｘの歪められたバージョンである。Ｙは、Ｘをカーネル$P_{Y|X}$に通すことによって実現される。本出願では、用語「カーネル」は、データＸをデータＹに確率的に写像する条件付き確率を指す。すなわち、カーネル$P_{Y|X}$は、我々が設計するのを望むプライバシー保護写像である。Ｙは、Ｘのみの確率的関数であるので、本出願では、Ｓ→Ｘ→ＹがマルコフＳ連鎖を形成すると仮定する。したがって、$P_{Y|X}$を定義すると、同時分布$P_{S,X,Y}=P_{Y|X}P_{S,X}$、特に同時分布$P_{S,Y}$が得られる。
【００２６】
以下では、我々は、まずプライバシー概念を定義し、次いで精度概念を定義する。
定義１．Ｓ→Ｘ→Ｙを仮定する。カーネル$P_{Y|X}$は、同時分布$P_{S,X,Y}=P_{Y|X}P_{S,X}$から得られる分布$P_{S,Y}$が以下の式を満たす場合、ε－発散プライベートと呼ばれる。
【００２７】
【数４】
$$D(P_{S,Y}||P_S P_Y) \triangleq \mathbb{E}_{S,Y}\left[\log \frac{P(S|Y)}{P(S)}\right] \triangleq I(S;Y) = \epsilon H(S), \qquad (1)$$
【００２８】
上式において、Ｄ（．）はＫ－Ｌ発散であり、
【００２９】
【数５】
$$\mathbb{E}(.)$$
【００３０】
は確率変数の期待値であり、Ｈ（．）はエントロピーであり、ε∈［０，１］は漏出係数と呼ばれ、相互情報量Ｉ（Ｓ；Ｙ）は情報漏出を表す。
【００３１】
我々は、ε＝０の場合に、機構が完全プライバシーを有すると言う。極値の場合、ε＝０は、公開された確率変数Ｙが、プライベート確率変数Ｓから独立していることを暗示し、ε＝１は、ＳがＹから完全に復元可能である（ＳはＹの確定関数である）ことを暗示する。Ｙは、完全にＳから独立して完全プライバシーを有する（ε＝０）と仮定することができるが、これは、低い精度レベルをもたらす可能性があることに留意されたい。我々は、精度を以下のように定義する。
【００３２】
定義２．
【００３３】
【数６】
$$d: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+$$
【００３４】
を歪み測度とする。カーネル$P_{Y|X}$は、
【００３５】

【数７】

$$\mathbb{E}[d(X,Y)] \leq D$$

【００３６】

であるならば、Ｄ－精度があると呼ばれる。

【００３７】

プライバシー保護写像の漏出係数εと歪みレベルＤとの間にトレードオフがある。

【００３８】

本発明は、事前の一部の統計的知識のみが入手可能であるときに効用対応プライバシー保護写像機構を設計する方法を提案する。より詳細には、本原理は、付加ノイズ機構の部類におけるプライバシー保護写像機構を提供し、ノイズはパブリックデータに対してそれが公開される前に加えられる。分析において、我々は、ノイズの平均値をゼロであると仮定する。この機構は、平均がゼロでないときにも適用され得る。一例では、エントロピーが平均にセンシティブでないため、結果は非ゼロ平均に対して同じである。この機構は、データの２次モーメントの知識のみが、連続データと離散データの両方に関して公開されることを必要とする。

【００３９】

ガウス機構（Gaussian Mechanism）

一実施形態では、我々は、連続パブリックデータＸ、およびノイズを信号に加えることによって達成されるプライバシー保護写像方式、すなわちＹ＝Ｘ＋Ｎを考える。例示的連続パブリックデータは、ユーザの身長または血圧であり得る。写像は、P_Xおよび$P_{S,X}$を知ることなく、ＶＡＲ（Ｘ）（または多次元Ｘの場合の共分散行列）を知ることによって得られる。まず、我々は、プライバシーを保護するためにノイズをパブリックデータに加える場合に、すべてのプライバシー保護写像機構のうち、ガウスノイズを加えることが最適であることを示す。

【００４０】

Ｓ→Ｘ→Ｙであるので、Ｉ（Ｓ；Ｙ）≦Ｉ（Ｘ；Ｙ）が得られる。情報漏出Ｉ（Ｓ；Ｙ）を境界付けるために、我々はＩ（Ｘ；Ｙ）を境界付ける。Ｘ＝ｆ（Ｓ）がＳの確定関数である場合、Ｉ（Ｓ；Ｙ）＝Ｉ（Ｘ；Ｙ）であり、境界はタイトである（これは、たとえば、ある行列ＡについてＸ＝ＡＳであるときに線形回帰で発生する）。

【００４１】

【数８】

$$X \in \mathbb{R}^n$$

【００４２】

とする。Ｘの共分散行列をC_Xで示す。Ｙ＝Ｘ＋Ｎとし、式中、Ｎは、Ｘから独立したノイズであり、平均０および共分散行列C_Nを有する。我々は、１つの確率変数のみがあるときに分散

【００４３】

【数９】

$$(\sigma_N^2)$$

【００４４】

の表記を使用し、複数の確率変数があるときに共分散（C_N）の表記を使用することに留意されたい。我々は、以下の結果を有する。

命題２．　P_Xが、プライバシー保護写像の設計において知られておらず、我々は、あるσ_Xについての

【００４５】

【数１０】

$$\mathrm{VAR}(X) \leq \sigma_X^2$$

【００４６】

のみを知っていると仮定する。また、独立したノイズＮを信号Ｘに加えることによって得られたプライバシー保護方式のクラスを考える。このノイズは、ゼロ平均、およびあるσ_Nについての

【００４７】

【数１１】

$$\sigma_N^2$$

【００４８】

よりも大きくない分散（l_2ノルム歪み）を有する。我々は、ガウスノイズが下記の意味で最も良いことを示す：

【００４９】

【数１２】

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

【００５０】

式中、N_Gは、ガウスノイズを表し、Ｎは、

【００５１】

【数１３】

$$\mathbb{E}[N_G] = \mathbb{E}[N] = 0$$

【００５２】

かつ

【００５３】

【数１４】

$$\mathrm{VAR}(N_G) = \mathrm{VAR}(N) = \sigma_N^2$$

【００５４】

となるような確率変数である。これは、N_Gを使用した最悪の場合の情報漏出が、Ｎを使用した最悪の場合の情報漏出よりも大きくないことを暗示する。

【００５５】

証明：　ガウス鞍点定理を使用すると、

【００５６】

【数１５】

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

$$= I(X_G;X_G+N_G) \leq I(X_G;X_G+N)$$

$$\leq \max_{P_X:\,X\perp\!\!\!\perp N,\,\mathrm{VAR}(X)\leq\sigma_X^2} I(X;X+N), \qquad (16)$$

【００５７】

が得られ、式中、X_Gは、ゼロ平均および分散

【００５８】

【数１６】

$$\sigma_X^2$$

【００５９】
を有するガウス分布を持つ。これで証明を完了する。□
【００６０】
ここで、我々は、プライバシーを保護するためにノイズを加える場合、l_2ノルム歪み制約下で、ガウスノイズを加えることが付加ノイズの系統群における最適解であることが分かっている。以下では、我々は、ガウスノイズをパブリックデータに加えるための最適パラメータを決定する。我々は、ガウス機構によって、そのようなパラメータでガウスノイズを加える機構を表す。
【００６１】
１つの例示的実施形態では、所与のC_Xおよび歪みレベルＤについて、ガウス機構は、図１に示されるようなステップで進行する。
【００６２】
方法１００は、１０５から開始する。ステップ１１０で、それは、彼らのパブリックデータまたはプライベートデータのプライバシーについて懸念しないユーザによって公開されたパブリックデータに基づいて統計情報を推定する。我々は、これらのユーザを「パブリックユーザ」として表し、彼らのプライベートデータのプライバシーについて懸念するユーザを「プライベートユーザ」として表す。
【００６３】
これらの統計は、ウェブを巡回して異なるデータベースにアクセスすることによって収集されてもよく、またはデータアグリゲータ、たとえば、ｂｌｕｅｋａｉ．ｃｏｍによって提供されてもよい。どの統計情報が収集され得るかは、パブリックユーザが何を公開するかに応じて変わる。分散を特徴付けることは、周辺分布P_Xを特徴付けるよりも必要とされるデータが少ないことに留意されたい。したがって、我々は分散を推定できるが周辺分布を精密に推定できないという状況に我々はあり得る。一例では、我々は、ステップ１２０で、収集された統計情報に基づいてパブリックデータの平均および分散（または共分散）のみを取得できることがある。
【００６４】
ステップ１３０で、我々は共分散行列C_Xの固有値分解を行う。ガウスノイズN_Gの共分散行列は、C_Xの固有ベクトルと同じ固有ベクトルを有する。さらに、C_Nの対応する固有値が以下の最適化問題を解くことによって与えられる。
【００６５】
【数１７】

$$\min_{\sigma_i:\,1\le i\le n}\prod_{i=1}^{n}\frac{\sigma_i+\lambda_i}{\sigma_i}$$

$$\text{s.t. } \sum_{i=1}^{n}\sigma_i \le D \qquad (17)$$

【００６６】
式中、λ_iおよびσ_i（$1 \leqq \lambda_i \leqq n$）はそれぞれ、$C_X$および$C_N$の固有値を表す。決定された固有ベクトルおよび固有値から、我々は次いで、ガウスノイズについての共分散行列C_Nをその固有値分解を介して決定することができる。続いて、我々は、ガウスノイズ
【００６７】
【数１８】

$$N_G \sim \mathcal{N}(0, C_N)$$

【００６８】
を生成することができる。歪みは、
【００６９】

【数１９】

$$\sum_{i=1}^{n} \mathbb{E}\left[(Y_i - X_i)^2\right] = \mathrm{tr}(C_N) = \sum_{i=1}^{n} \sigma_i \leq D$$

【００７０】

によって与えられ、式中、ｔｒ（）は対角要素の和を表し、ｎはベクトルＸの次元である。

【００７１】

ステップ１４０で、ガウスノイズがパブリックデータに加えられ、すなわち、Ｙ＝Ｘ＋N_Gとなる。ステップ１５０で、歪められたデータは、次いで、たとえば、サービスプロバイダまたはデータ収集エージェンシに公開される。方法１００は、ステップ１９９で終了する。

【００７２】

以下の定理において、我々は、提示されるガウス機構がl_2ノルム歪み制約下で最適であることを証明する。

【００７３】

定理３．　l_2ノルム歪みおよび所与の歪みレベルＤを仮定して、相互情報量を最小化するガウス機構における最適ガウスノイズは、

最適ノイズN_Gの共分散行列がC_Xの固有ベクトルと同じ固有ベクトルを有することを満たす。また、固有値は、（１７）で与えられる。

【００７４】

証明：

【００７５】

【数２０】

$$I(X; X+N) \leq \frac{1}{2}\log\left(\frac{|C_X + C_N|}{|C_N|}\right) \qquad (18)$$

【００７６】

が得られ、この不等式は、Ｔ．Ｍ．Ｃｏｖｅｒ ａｎｄ Ｊ．Ａ．Ｔｈｏｍａｓによる書籍”Ｅｌｅｍｅｎｔｓ ｏｆ ｉｎｆｏｒｍａｔｉｏｎ ｔｈｅｏｒｙ，” Ｗｉｌｅｙ－ｉｎｔｅｒｓｃｉｅｎｃｅ，２０１２の定理８．６．５に由来するものである。我々はＸの分布を知らないので、我々は上限境界

【００７７】

【数２１】

$$\frac{1}{2}\log\left(\frac{|C_X+C_N|}{|C_N|}\right)$$

【００７８】

を最小化しなければならない。それがガウスＸを用いて達成可能であるからである。半正定値行列C_Xの固有値分解を考慮して、$C_X = Q \Lambda Q^t$を得る。ここで$QQ^t = I$であり、Λは、C_Xの固有値を含む対角行列である。

【００７９】

【数２２】

$$\mathbb{E}\left[\sum_{i=1}^{k}(Y_i - X_i)^2\right] = tr(C_N) = tr(Q^t C_N Q) = \sum \sigma_i \leq D$$

【００８０】

が得られ、最適化問題は、

【００８１】

【数２３】

$$\min_{C_N} \frac{|\Lambda+Q^t C_N Q|}{|Q^t C_N Q|}$$

【００８２】
となり、ただし、ｔｒ（$Q^t C_N Q$）≦Ｄである。
【００８３】
一般性を失うことなく、$\lambda_1 \geqq \cdots \geqq \lambda_n$と仮定する。$\sigma_1 \geqq \cdots \geqq \sigma_n$を$Q^t C_N Q^t$の固有値とする。Ｍ．　Ｆｉｅｄｌｅｒによる論文”Ｂｏｕｎｄｓ　ｆｏｒ　ｔｈｅ　ｄｅｔｅｒｍｉｎａｎｔ　ｏｆ　ｔｈｅ　ｓｕｍ　ｏｆ　Ｈｅｒｍｉｔｉａｎ　ｍａｔｒｉｃｅｓ，”　Ｐｒｏｃｅｅｄｉｎｇｓ　ｏｆ　ｔｈｅ　Ａｍｅｒｉｃａｎ　Ｍａｔｈｅｍａｔｉｃａｌ　Ｓｏｃｉｅｔｙの定理１によると、我々は、$|\Lambda + Q^t C_N Q| \geqq \Pi(\lambda_i + \sigma_i)$を有することになり、等式は、$Q^t C_N Q$が対角行列の場合に成立する。したがって、同じ固有値σ_iを有する対角行列を使用して、我々は、同じ歪みレベル、およびより小さい漏出を達成し、それは最適性と相反する。よって、$Q^t C_N Q$は対角行列である。□
【００８４】
例３．　ＸはＳの確定実数値関数、すなわちＸ＝ｆ（Ｓ）であり、
【００８５】
【数２４】

$$\mathrm{VAR}(X) = \sigma_X^2$$

【００８６】
であると仮定する。Ｓ→Ｘ→Ｙであるため、我々は、Ｉ（Ｘ；Ｙ）＝Ｉ（Ｓ；Ｙ）を有する。
【００８７】
【数２５】

$$N \sim \mathcal{N}(0, \sigma_N^2)$$

【００８８】
かつＹ＝Ｘ＋Ｎとする。任意のεについて、我々は、（ε，Ｄ）－発散歪みプライバシー、ただし、
【００８９】
【数２６】

$$D = \frac{\sigma_X^2}{e^{2\varepsilon H(S)} - 1}$$

【００９０】
を達成することができる。
【００９１】
注記１．　この分析は、ε＞０の場合のみ機能する。完全なプライバシーすなわちε＝０を得たいとき、この方式は、
【００９２】
【数２７】

$$\sigma_N^2 = \infty$$

【００９３】
を選択する。実際、これは、ＹがＸから独立していることを意味する。
【００９４】

【数２８】

$$Y = \mathbb{E}[X]$$

【００９５】

（確定値）を仮定する場合、Ｉ（Ｙ；Ｓ）＝０、かつ

【００９６】

【数２９】

$$\mathbb{E}[d(X,Y)] = \mathrm{VAR}(X)$$

【００９７】

である。したがって、ＶＡＲ（Ｘ）以上の歪みレベルの場合、

【００９８】

【数３０】

$$Y = \mathbb{E}[X]$$

【００９９】

をセットする決定機構は、ε＝０を達成する。

【０１００】

例５．　分散

【０１０１】

【数３１】

$$\sigma_N{}^2 \geq \frac{1}{\epsilon^2} 2\log(2/\delta)$$

【０１０２】

を有するガウスノイズを加えることによって我々は（ε，δ）－差分プライバシーを達成できることが示され得る。この方式は結果として、歪み

【０１０３】

【数３２】

$$D \geq \frac{1}{\epsilon^2} 2\log(2/\delta)$$

【０１０４】

および情報の漏出

【０１０５】

【数３３】

$$L \leq \frac{1}{2}\log\left(1 + \frac{\sigma_X^2}{\frac{1}{\epsilon^2}2\log(2/\delta)}\right)$$

【０１０６】

をもたらす。比較のための定性的方法では、（ε，δ）差分プライバシーガウス機構を使用し、我々は少ない漏出を実現するために大きな歪みを要することを述べる。他方で、本原理による発散プライバシーガウス機構を使用して、最小限の歪みＤを用いてＬビットを漏出する方式は、任意の（ε，δ）－差分プライバシー、ただし、

【０１０７】

【数３４】

$$\frac{1}{\epsilon^2} 2\log\left(\frac{2}{\delta}\right) = \frac{\sigma_X^2}{e^{2L}-1}$$

【０１０８】

を達成する。

【０１０９】

離散機構

別の実施形態では、我々は、離散確率変数Ｘ、ただし、

【０１１０】

【数３５】

$$\mathcal{X} = \mathbb{Z}$$

【０１１１】

を考える。ここでも、Ｉ（Ｓ；Ｙ）を境界付けるためにＩ（Ｘ；Ｙ）を境界付ける。歪み測度をl_pノルムとする、すなわち、ＸとＹとの間の歪みは、ある$1 \leqq p \leqq \infty$において

【０１１２】

【数３６】

$$\mathbb{E}[|X-Y|^p]^{\frac{1}{p}}$$

【０１１３】

とする。

【０１１４】

定義５．　所与の$1 \leqq p \leqq \infty$について、所与のＤ以下のl_pノルムを有するすべての確率変数において、$P^*_{p,D}$によって最大エントロピーを有する分布を表す。より形式的には、$P^*_{p,D}$は、以下の最適化における最大目的関数を実現する確率測度である。

【０１１５】

【数３７】

$$\max_{P_Z : Z \sim P_Z} H(Z)$$

【０１１６】

ただし

【０１１７】

【数３８】

$$\mathbb{E}[|Z|^p]^{\frac{1}{p}} \leq D$$

【０１１８】

すなわち、最適化問題は、ｐ次モーメントに対する制約を受ける、最大エントロピー離散確率分布$P^*_{p,D}$を求めることである。最大エントロピーは、H^*（ｐ，Ｄ）で表される。

【０１１９】

次に、我々は、$P^*_{p,D}$およびそのエントロピーを特徴付ける。

【０１２０】

命題３．　任意の$1 \leqq p \leqq \infty$について、$P^*_{p,D}$は、

【０１２１】

【数３９】

$$P^*_{p,D}[Z=i] = AB^{-|i|^p}$$

【０１２２】

によって与えられ、式中、ＡおよびＢは、

【０１２３】

【数４０】

$$\sum_{i=-\infty}^{\infty} A\,B^{-|i|^p} = 1$$

【０１２４】

かつ

【０１２５】

【数４１】

$$\mathbb{E}[|Z|^p]^{\frac{1}{p}} = D$$

【０１２６】

となるように選択される。さらに、H^*（ｐ，Ｄ）＝－ｌｏｇＡ＋（ｌｏｇＢ）D^pが得られることになる。

【０１２７】

証明：

【０１２８】

【数４２】

$$\mathbb{E}[|W|^p]^{\frac{1}{p}} \leq D$$

【０１２９】

となるように

【０１３０】

【数４３】

$$Z \sim AB^{-|i|^p}$$

【０１３１】

および$W \sim P_W$とする。

【０１３２】

【数４４】

$$\mathbb{E}_{P_W}[|i|^p] \leq D^p$$

【０１３３】

であるので、

【０１３４】

【数４５】

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

$$= -H(W) - \mathbb{E}_{P_W}[\log A - (\log B)|i|^p]$$

$$\leq -H(W) - \log A + (\log B)\mathbb{E}_{P_Z}[|i|^p]$$

$$= -H(W) + H(Z)$$

【０１３５】

が得られる。

【０１３６】

したがって、Ｈ（Ｚ）≧Ｈ（Ｗ）であり、H^*（ｐ，Ｄ）＝－ｌｏｇＡ＋（ｌｏｇＢ）D^pである。□

我々は、離散機構によって、ノイズ$Z \sim P^*_{p,D}$を離散パブリックデータに加える機構を示す。１つの例示的実施形態では、離散機構は、図２に示されるようなステップで進行する。

【０１３７】

方法２００は、２０５から開始する。ステップ２１０で、それは、パラメータ、たとえば、ｐおよびＤにアクセスして、歪み測度を定義する。所与の歪み測度l_p（$1 \leq p \leq \infty$

）および歪みレベルＤについて、それは、ステップ２２０で、命題３で与えられるように確率測度$P^*_{p,D}$を計算する。分布$P^*_{p,D}$はｐおよびＤだけによって決定されるが、結果のプライバシー精度トレードオフは、歪み制約がプライバシーおよび精度を結合するので、Ｘに応じて変わることに留意されたい。

【０１３８】

ステップ２３０で、ノイズが確率測度によって生成され、その後、ステップ２４０で、それがパブリックデータに加えられる、すなわち、Ｙ＝Ｘ＋Ｚであり、ただしＺ～$P^*_{p,D}$である。

【０１３９】

【数４６】

$$d(X,Y) = \mathbb{E}[|Y-X|^p]^{\frac{1}{p}} = \mathbb{E}[|Z|^p]^{\frac{1}{p}} \leq D$$

【０１４０】

を得ることになる。方法２００は、ステップ２９９で終了する。

【０１４１】

次に、我々は、相互情報量Ｉ（Ｘ；Ｙ）を分析する。

【０１４２】

【数４７】

$$||X||_p = \mathbb{E}[|X|^p]^{\frac{1}{p}}$$

【０１４３】

がXのl_pノルムを表すとする。ミンコフスキーの不等式を使用すると、

【０１４４】

【数４８】

$$\mathbb{E}[|Y|^p]^{\frac{1}{p}} = \mathbb{E}[|X+Z|^p]^{\frac{1}{p}} \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Z|^p]^{\frac{1}{p}} = ||X||_p + D$$

【０１４５】

を得ることになる。したがって、

Ｉ（Ｓ；Ｙ）≦Ｉ（Ｘ；Ｙ）＝Ｈ（Ｘ＋Ｚ）－Ｈ（Ｚ）≦H^*（ｐ，｜｜Ｘ｜｜$_p$＋Ｄ）－H^*（ｐ，Ｄ）

を得る。すなわち、離散機構を使用するときに我々が得るプライバシー保証（すなわち情報漏出）は、ＤとＸの平均l_pノルムとの両方に依存する右項によって上限境界を付けられる。

【０１４６】

付加ノイズ技法の長所は、それが、Ｓに関する情報はもちろんＸの統計に関する多くの情報を必要としないことだけではなく、それが、最適化問題を単純な問題に軽減し、その場合、完全なカーネル$P_{Y|X}$を指定する必要がある代わりに、我々が設計を必要とするのがノイズのパラメータだけであることである。これにより、最適化のサイズがかなり縮小され、したがって、それを解決するためのその複雑性および計算／メモリ必要条件がかなり軽減される。

【０１４７】

有利なことに、同時確率分布$P_{S,X}$の知識なしに、パブリックデータＸの１次および２次モーメントの知識のみで、本原理は、連続データと離散データの両方について、ノイズをパブリックデータに追加することによってプライバシーを保護するプライバシー保護写像機構を提供する。

【０１４８】

プライバシーエージェントは、プライバシーサービスをユーザに提供するエンティティ

である。プライバシーエージェントは、以下の任意のものを行うことができる：
－　ユーザから、どんなデータをユーザがプライベートであるとみなすか、どんなデータをユーザがパブリックであるとみなすか、およびユーザが求めるプライバシーのレベルを受け取る；
－　プライバシー保護写像を計算する；
－　ユーザのためのプライバシー保護写像を実装する（すなわち、写像によりユーザのデータを歪める）；および
－　歪められたデータを、たとえば、サービスプロバイダまたはデータ収集エージェンシに公開する。

【０１４９】
本原理は、ユーザデータのプライバシーを保護するプライバシーエージェントにおいて使用され得る。図３は、プライバシーエージェントが使用され得る例示的システム３００のブロック図を示す。パブリックユーザ３１０は、彼らのプライベートデータ（Ｓ）および／またはパブリックデータ（Ｘ）を公開する。前述されたように、パブリックユーザは、パブリックデータをそのままで公開する、すなわちＹ＝Ｘである。パブリックユーザによって公開された情報は、プライバシーエージェントに有用な統計情報となる。

【０１５０】
プライバシーエージェント３８０は、統計収集モジュール３２０、付加ノイズ生成器３３０、およびプライバシー保護モジュール３４０を含む。統計収集モジュール３２０は、パブリックデータの共分散を収集するために使用され得る。統計収集モジュール３２０は、ｂｌｕｅｋａｉ．ｃｏｍなどのデータアグリゲータから統計値を受け取ってもよい。入手可能な統計情報に応じて、付加ノイズ生成器３３０は、たとえば、ガウス機構または離散機構に基づいて、ノイズを設計する。プライバシー保護モジュール３４０は、生成されたノイズを加えることによって、プライベートユーザ３６０のパブリックデータをそれが公開される前に歪める。一実施形態では、統計収集モジュール３２０、付加ノイズ生成器３３０、およびプライバシー保護モジュール３４０はそれぞれ、方法１００におけるステップ１１０、１３０、および１４０を実施するために使用され得る。

【０１５１】
プライバシーエージェントは、データ収集モジュールに収集された、データ全体の知識なしに作用するための統計値のみを必要とすることに留意されたい。したがって、別の実施形態では、データ収集モジュールは、データを収集し次いで統計値を計算する独立したモジュールであってもよく、プライバシーエージェントの一部である必要はない。データ収集モジュールは、プライバシーエージェントと統計値を共有する。一実施形態では、付加ノイズ生成器３３０およびプライバシー保護モジュール３４０はそれぞれ、方法２００におけるステップ２２０および２３０を実施するために使用され得る。

【０１５２】
プライバシーエージェントは、ユーザとユーザデータの受信者（たとえばサービスプロバイダ）との間に存在する。たとえば、プライバシーエージェントは、ユーザデバイス、たとえば、コンピュータまたはセットトップボックス（ＳＴＢ）に配置されてよい。別の例では、プライバシーエージェントは別個のエンティティであってよい。

【０１５３】
プライバシーエージェントのすべてのモジュールは、１つのデバイスに配置されてもよく、または異なるデバイスにわたって分散されてもよく、たとえば、統計収集モジュール３２０は、統計値をモジュール３３０に単に公開するデータアグリゲータに配置されてもよく、付加ノイズ生成器３３０は、「プライバシーサービスプロバイダ」に、またはモジュール３２０に接続されたユーザデバイス上のユーザ端に配置されてもよく、プライバシー保護モジュール３４０は、ユーザ間の媒介として働くプライバシーサービスプロバイダ、およびユーザがデータを公開したい先のサービスプロバイダ、またはユーザデバイス上のユーザ端に配置されてもよい。

【０１５４】

プライベートユーザ３６０が受け取るサービスを公開されるデータに基づいて改善するために、プライバシーエージェントは、公開されるデータを、サービスプロバイダ、たとえば、コムキャストまたはネットフリックスに提供することができ、たとえば、推奨システムが、その公開された映画ランキングに基づいてユーザに映画の推奨を提供する。

【０１５５】

図４では、我々は、システム内に複数のプライバシーエージェントがあることを示す。異なる変形形態では、プライバシーエージェントはプライバシーシステムが作用するための必要条件でないので、それはすべての場所にある必要がない。たとえば、プライバシーエージェントは、ユーザデバイスもしくはサービスプロバイダのみ、または両方にあってもよい。図４では、我々は、ネットフリックスとフェイスブック（Facebook）の両方に対して同じプライバシーエージェント「Ｃ」を示す。別の実施形態では、フェイスブックとネットフリックスにおけるプライバシーエージェントは、同じにすることができるが必ずしも同じでなくてよい。

【０１５６】

本明細書に説明される実装形態は、たとえば、方法もしくはプロセス、装置、ソフトウェアプログラム、データストリーム、または信号において実装されてよい。単一形態の実装形態の文脈でのみ論じられた（たとえば、方法としてのみ論じられた）場合でも、論じられた特徴の実装形態は、他の形態（たとえば、装置またはプログラム）で実装されてもよい。装置は、たとえば、適切なハードウェア、ソフトウェア、およびファームウェアで実装されてよい。方法は、たとえばコンピュータ、マイクロプロセッサ、集積回路、またはプログラマブル論理デバイスを含む、一般に処理デバイスを指すたとえばプロセッサなどの、たとえば装置において実装されてよい。プロセッサはまた、エンドユーザ間の情報の通信を促進する、たとえば、コンピュータ、セルフォン、ポータブル／パーソナルディジタルアシスタント（「ＰＤＡ」）、および他のデバイスなどの、通信デバイスを含む。

【０１５７】

本原理の「一実施形態」または「実施形態」または「一実装形態」または「実装形態」、およびそれらの他の変形形態への参照は、その実施形態に関連して説明される特定の特徴、構造、および特性などが本原理の少なくとも１つの実施形態に含まれることを意味する。したがって、本明細書を通して様々な箇所で出現する表現「一実施形態では」または「実施形態では」または「一実装形態では」または「実装形態では」、および任意の他の変形形態の出現は、必ずしもすべて同じ実施形態を参照するものではない。

【０１５８】

加えて、本出願またはその特許請求の範囲は、様々な情報片を「決定すること」を参照することがある。情報を決定することは、たとえば、情報を推定すること、情報を計算すること、情報を予測すること、またはメモリから情報を取り出すことのうちの１または複数を含み得る。

【０１５９】

さらに、本出願またはその特許請求の範囲は、様々な情報片に「アクセスすること」を参照することがある。情報にアクセスすることは、たとえば、情報を受け取ること、（たとえばメモリから）情報を取り出すこと、情報を記憶すること、情報を処理すること、情報を送信すること、情報を移動すること、情報をコピーすること、情報を消去すること、情報を計算すること、情報を決定すること、情報を予測すること、または情報を推定することのうちの１または複数を含み得る。

【０１６０】

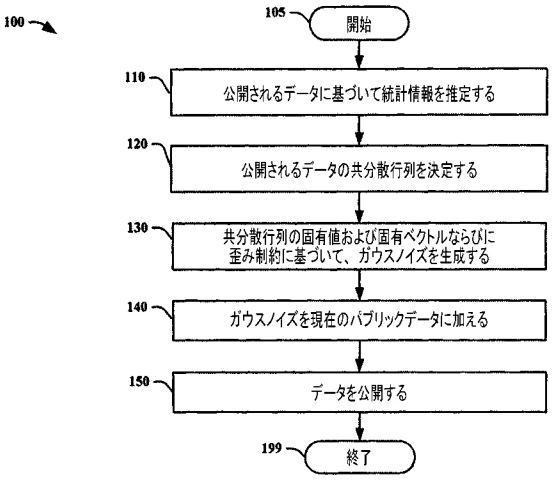
加えて、本出願またはその特許請求の範囲は、様々な情報片を「受け取る」ことを参照することがある。受け取ることは、「アクセスすること」と同様に広義の用語になるように意図されている。情報を受け取ることは、たとえば、情報にアクセスすること、または（たとえばメモリから）情報を取り出すことのうちの１または複数を含み得る。さらに、「受け取ること」は、通常は、何らかの形で、たとえば、情報を記憶すること、情報を処理すること、情報を送信すること、情報を移動すること、情報をコピーすること、情報を

消去すること、情報を計算すること、情報を決定すること、情報を予測すること、または情報を推定することなどの動作の際に行われる。

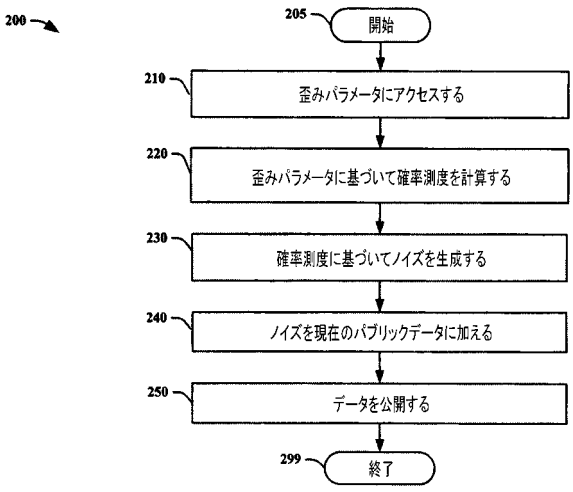
【０１６１】

当業者には明らかなように、実装形態は、たとえば、記憶または送信され得る情報を搬送するようにフォーマットされた様々な信号を生成し得る。情報は、たとえば、方法を実施するための命令、または説明された実装形態の１つによって生成されたデータを含み得る。たとえば、信号は、説明された実施形態のビットストリームを搬送するようにフォーマット化されてよい。そのような信号は、たとえば、電磁波として（たとえば、スペクトルの無線周波数部分を使用する）、またはベースバンド信号としてフォーマットされてよい。フォーマッティングは、たとえば、データストリームを符号化し、符号化されたデータストリームを用いて搬送波を変調することを含むことができる。信号が搬送する情報は、たとえば、アナログまたはディジタル情報であってよい。信号は、知られている様々な異なる有線またはワイヤレスリンクを介して送信されてよい。信号は、プロセッサ可読媒体に記憶されてよい。

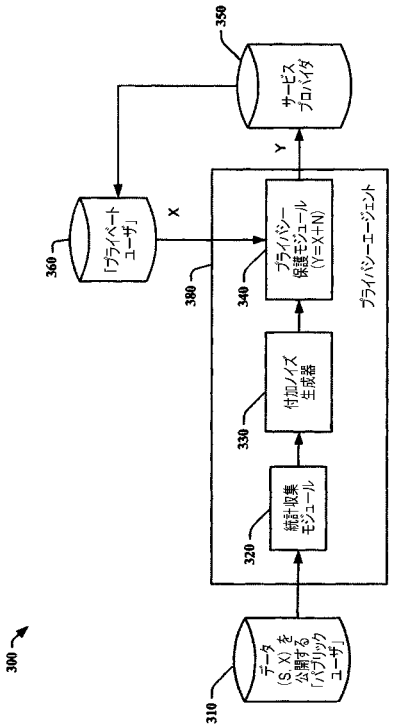
【図１】



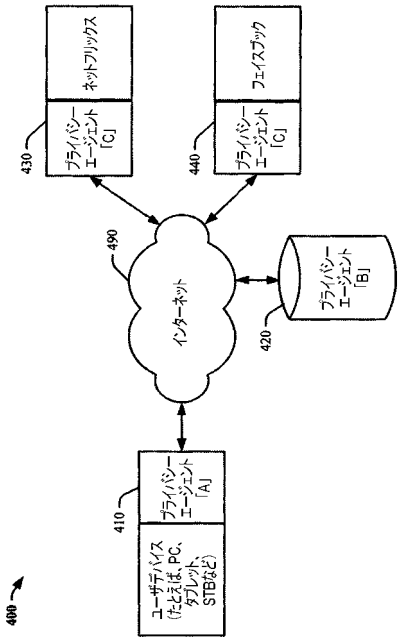
【図２】



【図３】


300
310
データ（S、X）を公開する「パブリックユーザ」
320
統計収集モジュール
330
付加ノイズ生成器
340
プライバシー保護モジュール（Y=X+N）
プライバシーエージェント
350
サービスプロバイダ
Y
360
「プライベートユーザ」
X
380

【図４】


400
410
プライバシーエージェント「A」
ユーザデバイス（たとえば、PC、タブレット、STBなど）
490
インターネット
420
プライバシーエージェント「B」
430
プライバシーエージェント「C」
ネットフリックス
440
プライバシーエージェント「C」
フェイスブック

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/071290

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F21/62
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	IBRAHIM YAKUT ET AL: "Privacy-Preserving Eigentaste-Based Collaborative Filtering", 29 October 2007 (2007-10-29), ADVANCES IN INFORMATION AND COMPUTER SECURITY; [LECTURE NOTES IN COMPUTER SCIENCE], SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 169 - 184, XP019073150, ISBN: 978-3-540-75650-7 page 172 - page 175 ----- -/--	1-21

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 April 2014	12/05/2014

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Koblitz, Birger

2

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/071290

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	FLAVIO DU PIN CALMON ET AL: "Privacy against statistical inference", COMMUNICATION, CONTROL, AND COMPUTING (ALLERTON), 2012 50TH ANNUAL ALLERTON CONFERENCE ON, IEEE, 1 October 2012 (2012-10-01), pages 1401-1408, XP032345161, DOI: 10.1109/ALLERTON.2012.6483382 ISBN: 978-1-4673-4537-8 the whole document	1-21
X,P	ALI MAKHDOUMI ET AL: "Privacy-utility tradeoff under statistical uncertainty", 2013 51ST ANNUAL ALLERTON CONFERENCE ON COMMUNICATION, CONTROL, AND COMPUTING (ALLERTON), 1 October 2013 (2013-10-01), pages 1627-1634, XP055103156, DOI: 10.1109/Allerton.2013.6736724 ISBN: 978-1-47-993409-6 page 7	1-21
A	JEREMIAH BLOCKI ET AL: "The Johnson-Lindenstrauss Transform Itself Preserves Differential Privacy", FOUNDATIONS OF COMPUTER SCIENCE (FOCS), 2012 IEEE 53RD ANNUAL SYMPOSIUM ON, IEEE, 20 October 2012 (2012-10-20), pages 410-419, XP032276913, DOI: 10.1109/FOCS.2012.67 ISBN: 978-1-4673-4383-1 page 2 - page 4	1-21
A	CYNTHIA DWORK ET AL: "Our Data, Ourselves: Privacy Via Distributed Noise Generation", 28 May 2006 (2006-05-28), ADVANCES IN CRYPTOLOGY - EUROCRYPT 2006 LECTURE NOTES IN COMPUTER SCIENCE;;LNCS, SPRINGER, BERLIN, DE, PAGE(S) 486 - 503, XP047029518, ISBN: 978-3-540-34546-6 the whole document	1-21
A	US 2008/209568 A1 (CHANG YUAN-CHI [US] ET AL) 28 August 2008 (2008-08-28) the whole document	1-21

2

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2013/071290

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008209568 A1	28-08-2008	NONE	

Form PCT/ISA/210 (patent family annex) (April 2005)

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,JP,KE,KG,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US

（特許庁注：以下のものは登録商標）
１．ＦＡＣＥＢＯＯＫ

(72)発明者 アバサリ マクドウミ カカーキ
アメリカ合衆国 ０２１４３ マサチューセッツ州 サマービル オーク ストリート ３５ アパートメント ２

Ｆターム(参考) 5J104 AA12 AA16 AA32 AA41 EA08 EA31 FA10 GA01 JA04 LA04
NA10 PA14