(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2019-61233
(P2019-61233A)
(43)公開日　平成31年4月18日(2019.4.18)

| (51)Int.Cl. | | | F I | | | テーマコード（参考） |
|---|---|---|---|---|---|---|
| G09C | 1/00 | (2006.01) | G09C | 1/00 | 650Z | 5J104 |
| H04L | 9/30 | (2006.01) | H04L | 9/00 | 663Z | |

審査請求 未請求　請求項の数 10　ＯＬ　（全 31 頁）

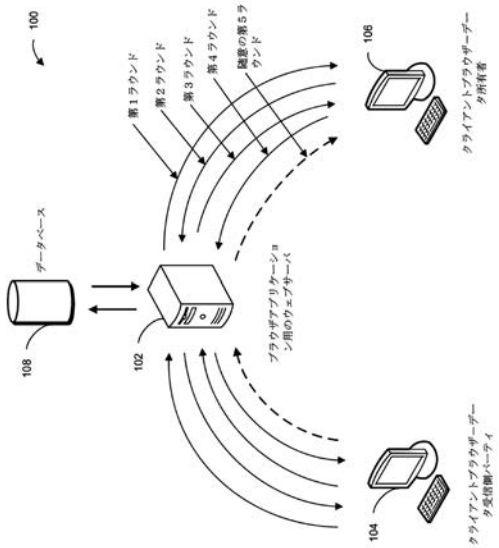
(21)出願番号　特願2018-165692 (P2018-165692)
(22)出願日　平成30年9月5日 (2018.9.5)
(31)優先権主張番号　15/714,803
(32)優先日　平成29年9月25日 (2017.9.25)
(33)優先権主張国　米国 (US)

(71)出願人　504407000
パロ　アルト　リサーチ　センター　インコーポレイテッド
アメリカ合衆国　カリフォルニア州　９４３０４　パロ　アルト　カイオーテ　ヒル　ロード　３３３３
(74)代理人　100094569
弁理士　田中　伸一郎
(74)代理人　100088694
弁理士　弟子丸　健
(74)代理人　100067013
弁理士　大塚　文昭
(74)代理人　100086771
弁理士　西島　孝喜

最終頁に続く

(54)【発明の名称】データを共有する有用性の安全な２パーティ評価のためのシステム及び方法

(57)【要約】　（修正有）
【課題】第２のパーティとデータを共有することの有用性に関する統計を計算することによって分類器を改善するためのシステムを提供する。
【解決手段】ラウンド１を行う役割を果すパーティ（クライアントコンピュータ１０４）は、計算を行いセッションＩＤと共に、結果をウェブサーバに提出する。ウェブサーバは、結果をデータベース１０８に記憶し、プロトコルの次のラウンドを行う必要がある旨の通知と共に、結果をもう一つのパーティに伝送する。第２パーティ（クライアントコンピュータ１０６）は、ラウンド２の計算を行い、次いでセッションＩＤを有するウェブサーバに結果を記録する。ウェブサーバ１０２は、他のクライアントから入手可能なラウンドの出力があるときに、クライアントへの通知を継続する。ラウンドが使用可能であると通知されたときに、各パーティは次の計算ラウンドを行い、ウェブサーバに結果を送信する。
【選択図】図１



【特許請求の範囲】
【請求項１】
２つ以上のパーティからのデータを組み合わせることの有用性をセキュアに評価するためのコンピュータ実装の方法であって、前記方法が、
多数のラウンドにおいて、暗号化されたデータ及び前記データについて計算された、暗号化された統計を交換することであって、統計が、前記データの１つまたは２つ以上の属性から導出される尺度である、交換することと、
秘密暗号解読鍵を共有することなく、暗号化されたドメインにおいて計算を使用して、暗号化された有用性統計を構築することと、
１つまたは２つ以上のパーティによって、それぞれの秘密暗号解読鍵を適用して、前記暗号化された有用性統計を解読し、有用性統計を取得することと、を含む、方法。
【請求項２】
前記有用性統計が所定の閾値を超えたと判定することと、
前記第２のコンピューティングデバイスから一組のデータを要求することと、
前記一組のデータを適用して、分類器を改善することと、をさらに含む、請求項１に記載の方法。
【請求項３】
有用性統計を取得するためのコンピュータ実装の方法であって、前記方法が、第１のコンピューティングデバイスを備え、
公開鍵／秘密鍵の対に基づいて一組のクラスラベルを暗号化して、一組の暗号化されたクラスラベルを取得することと、
第２のコンピューティングデバイスに公開鍵及び前記一組の暗号化されたクラスラベルを送信することと、
前記公開鍵に基づいて前記第２のコンピューティングデバイスによって計算された、暗号化された値を受信することと、
秘密鍵に基づいて、前記暗号化された値を解読して、解読された値を取得することと、
前記第２のコンピューティングデバイスに、前記解読された値に基づいて計算された一対の暗号化された値を送信することと、
前記第２のコンピューティングデバイスから暗号化された有用性統計を受信することと、
前記暗号化された有用性統計を解読して、解読された有用性統計を取得することと、を含む、方法。
【請求項４】
前記一対の暗号化された値を送信することが、
前記解読された値に基づいて一対の値を計算することと、
前記公開鍵に基づいて、前記一対の値を暗号化して、前記一対の暗号化された値を取得することと、をさらに含む、請求項３に記載の方法。
【請求項５】
前記解読された有用性統計が所定の閾値を超えると判定することと、
第２のコンピューティングデバイスから一組のデータを要求することと、
前記一組のデータを適用して、前記分類器を改善することと、をさらに含む、請求項３に記載の方法。
【請求項６】
前記クラスラベルが、バイナリ属性と関連付けられる、請求項３に記載の方法。
【請求項７】
前記第２のコンピューティングデバイスに要求を送信して、有用性統計を計算することをさらに含む、請求項３に記載の方法。
【請求項８】
有用性統計を取得するためのコンピューティングシステムであって、前記方法が、前記システムが、

１つまたは２つ以上のプロセッサと、
命令が記憶された、前記１つまたは２つ以上のプロセッサに結合された非一時的コンピュータ可読媒体と、を備え、該命令が、前記１つまたは２つ以上のプロセッサによって実行されたときに、前記１つまたは２つ以上のプロセッサに
公開鍵／秘密鍵の対に基づいて一組のクラスラベルを暗号化して、一組の暗号化されたクラスラベルを取得することと、
第２のコンピューティングデバイスに公開鍵及び前記一組の暗号化されたクラスラベルを送信することと、
前記公開鍵に基づいて前記第２のコンピューティングデバイスによって計算された、暗号化された値を受信することと、
秘密鍵に基づいて、前記暗号化された値を解読して、解読された値を取得することと、
前記第２のコンピューティングデバイスに、前記解読された値に基づいて計算された一対の暗号化された値を送信することと、
前記第２のコンピューティングデバイスから暗号化された有用性統計を受信することと、
前記暗号化された有用性統計を解読して、解読された有用性統計を取得することと、を含む、動作を行わせる、コンピューティングシステム。
【請求項９】
前記一対の暗号化された値を送信することが、
前記解読された値に基づいて一対の値を計算することと、
前記公開鍵に基づいて、前記一対の値を暗号化して、前記一対の暗号化された値を取得することと、をさらに含む、請求項８に記載のシステム。
【請求項１０】
前記解読された有用性統計が所定の閾値を超えると判定することと、
前記第２のコンピューティングデバイスから一組のデータを要求することと、
前記一組のデータを適用して、前記分類器を改善することと、をさらに含む、請求項８に記載のシステム。
【発明の詳細な説明】
【技術分野】
【０００１】
背景
政府出資研究に関する記述
本発明は、政府運輸省により授与された契約番号Ｇ０１２．３７８３．００（３７８３）ＦＨＷＡ－ＥＡＲＰ－Ｓａｆｅｔｙの下で、米国政府支援によって行われた。米国政府は、本発明において特定の権利を有する。
【０００２】
本開示は、概して、機械学習に関する。より具体的には、本開示は、特徴データを共有する有用性を評価するための方法及びシステムに関する。
【背景技術】
【０００３】
ＩＢＭは、同社が、２５０京バイトのデータを作成し、今日の世界におけるデータの９０％が、過去２年だけで作成されたことを報告した。種々の利害関係者へのデータの有用性を考慮すれば、データ取引は、重要な事業に成長するであろう。
【０００４】
システムの異なる態様に関するデータは、異なる利害関係者によって獲得され得る。システムの完全な観点を得るために、データは、利害関係者の間で取引することができる。
【０００５】
例えば、モノのインターネット（ＩｏＴ）エコシステムにおいて、概して、ＩｏＴデバイスは、異なる所有者（製造業者、サービスプロバイダ、消費者、その他）によって所有され、単一の所有者によって収集されたデータは、システムの部分的な視野だけしか提供

しない。
【０００６】
取引相手を識別すること、及びデータの値を決定することは、特に、値がデータの特性及びコンテンツに依存するので、厄介な課題であり得る。データを取得しようとしているパーティは、データの値または有用性を事前に決定することができる場合にだけ、データを望み得る。しかしながら、データ所有者は、データを転送するための条件について同意する前に、データのコンテンツを明らかにしたくない場合がある。したがって、実際のデータへのアクセスを伴わずに、データの有用性をセキュアに測定することが重要である。
【発明の概要】
【０００７】
本明細書で説明される実施形態は、２つ以上のパーティからのデータを組み合わせることの有用性をセキュアに評価するためのシステムを提供する。動作中に、本システムは、多数のラウンドにおいて、暗号化されたデータ及び該データについて計算された、暗号化された統計を交換することができる。統計は、データの１つまたは２つ以上の属性から導出される尺度である。本システムは、秘密暗号解読鍵を共有することなく、暗号化されたドメインにおいて計算を使用して、暗号化された有用性統計を構築することができる。次いで、１つまたは２つ以上のパーティは、それぞれの秘密暗号解読鍵を適用して、暗号化された有用性統計を解読し、有用性統計を取得することができる。
【０００８】
この実施形態の変形例において、本システムは、有用性統計が所定の閾値を超えたと判定し、第２のコンピューティングデバイスから一組のデータを要求することができる。システムは、一組のデータを適用して、分類器を改善することができる。
【０００９】
本明細書で説明される別の実施形態は、有用性統計を取得するためのシステムを含む。動作中に、本システムは、公開鍵／秘密鍵の対に基づいて一組のクラスラベルを暗号化して、暗号化されたクラスラベルを取得することができる。本システムは、第２のコンピューティングデバイスに、公開鍵及び一組の暗号化されたクラスラベルを送信することができる。本システムは、公開鍵に基づいて第２のコンピューティングデバイスによって計算された、暗号化された値を受信することができる。本システムは、秘密鍵に基づいて、暗号化された値を解読して、解読された値を取得することができる。本システムは、次いで、第２のコンピューティングデバイスに、解読された値に基づいて計算された一対の暗号化された値を送信することができる。本システムは、その後に、第２のコンピューティングデバイスから暗号化された有用性統計を受信し、暗号化された有用性統計を解読して、解読された有用性統計を取得することができる。
【００１０】
この実施形態の変形例において、一対の暗号化された値を送信することは、解読された値に基づいて一対の値を計算することをさらに含む。本システムはまた、公開鍵に基づいて、一対の値を暗号化して、一対の暗号化された値を取得することもできる。
【００１１】
この実施形態の変形例において、本システムは、解読された有用性統計が所定の閾値を超えると判定することができる。システムは、第２のコンピューティングデバイスから一組のデータを要求し、一組のデータを適用して、分類器を改善することができる。
【００１２】
この実施形態の変形例において、クラスラベルは、バイナリ属性と関連付けられる。
【００１３】
この実施形態の変形例において、本システムは、第２のコンピューティングデバイスに要求を送信して、有用性統計を計算することができる。
【００１４】
この実施形態の変形例において、公開鍵及び一組の暗号化されたクラスラベルを送信することは、一組のクラスラベルを含む特徴ベクトルの要素の合計に基づいて値を計算する

ことをさらに含む。本システムは、計算値を暗号化し、第２のコンピューティングデバイスに暗号化された値を送信することができる。

【００１５】

この実施形態の変形例において、本システムは、２ストリームモードで第２の暗号化された有用性統計を計算し、第２のコンピューティングデバイスに第２の暗号化された有用性統計を送信することができる。

【００１６】

さらなる変形例において、本システムは、第１の使い捨て鍵に基づいて、解読された有用性統計を再暗号化して、再暗号化された有用性統計を取得することができる。本システムは、第２のコンピューティングデバイスに再暗号化された有用性統計を送信することができる。本システムは、第２のコンピューティングデバイスから、第２の再暗号化された有用性統計値及び第２の使い捨て鍵を受信することができる。本システムは、次いで、第２のコンピューティングデバイスの第１の使い捨て鍵を送信することができる。本システムは、第２の使い捨て鍵に基づいて第２の再暗号化された有用性統計を解読して、第２の有用性統計を取得し、第２の有用性統計が、解読された有用性統計に等しいと判定することができる。

【００１７】

この実施形態の変形例において、第２のコンピューティングデバイスは、統計的プライバシー技法を一組のデータに適用して、共有すること／組み合わせることが考慮される第２の一組のデータを取得する。

【００１８】

この実施形態の変形例において、第２のコンピューティングデバイスは、一組の暗号化されたクラスラベル及び公開鍵に基づいて、暗号化された値を計算することを含む動作を行う。

【００１９】

この実施形態の変形例において、第２のコンピューティングデバイスは、加算準同型秘密鍵暗号化スキームと関連付けられたバイナリ動作、及びスカラー乗算準同型秘密鍵暗号化スキームと関連付けられたバイナリ動作、のうちの少なくとも１つを行う。

【００２０】

本明細書で説明される別の実施形態は、命令を記憶する非一時的コンピュータ可読記憶媒体を含み、該命令は、多数のコンピューティングデバイスを有するコンピューティングシステムによって実行されたときに、該システムに、２つ以上のパーティからのデータを組み合わせる有用性をセキュアに評価するための方法を行わせる。本方法の動作中に、本システムは、多数のラウンドにおいて、暗号化されたデータ及び該データについて計算された、暗号化された統計を交換することができる。統計は、データの１つまたは２つ以上の属性から導出される尺度である。本システムは、秘密暗号解読鍵を共有することなく、暗号化されたドメインにおいて、計算を使用して暗号化された有用性統計を構築することができる。次いで、１つまたは２つ以上のパーティは、それぞれの秘密暗号解読鍵を適用して、暗号化された有用性統計を解読し、有用性統計を取得することができる。

【００２１】

本明細書で説明される別の実施形態は、命令を記憶する非一時的コンピュータ可読記憶媒体を含み、該命令は、コンピューティングシステム（例えば、コンピュータ）によって実行されたときに、システムに、有用性統計を取得するための方法を行わせる。本方法の動作中に、本システムは、公開鍵／秘密鍵の対に基づいて一組のクラスラベルを暗号化して、一組の暗号化されたクラスラベルを取得することができる。本システムは、第２のコンピューティングデバイスに、公開鍵及び一組の暗号化されたクラスラベルを送信することができる。本システムは、公開鍵に基づいて第２のコンピューティングデバイスによって計算された、暗号化された値を受信することができる。本システムは、秘密鍵に基づいて、暗号化された値を解読して、解読された値を取得することができる。本システムは、次いで、第２のコンピューティングデバイスに、解読された値に基づいて計算された一対

の暗号化された値を送信することができる。本システムは、その後に、第２のコンピューティングデバイスから暗号化された有用性統計を受信し、暗号化された有用性統計を解読して、解読された有用性統計を取得することができる。

【００２２】

本明細書で説明される別の実施形態は、有用性統計を生成することを容易にするプロトコルを行うためのシステムを含む。動作中に、本システム（例えば、サーバ）は、２つのコンピューティングデバイスとのネットワーク接続を確立する要求を受信する。本システムは、ネットワーク接続をセットアップすることができ、また、セッションＩＤを割り当てることができる。システムは、２つのコンピューティングデバイスにセッションＩＤを送信することができる。システムは、セッションＩＤを有する第１のコンピューティングデバイスからデータを受信し、通知を有する第２のコンピューティングデバイスにデータを送信して、プロトコルの次のラウンドを行うことができる。

【００２３】

この実施形態の変形例において、本システムは、セッションＩＤを有する第２のコンピューティングデバイスからデータを受信し、第２の通知を有する第１のコンピューティングデバイスにデータを送信して、プロトコルの次のラウンドを行うことができる。

【００２４】

本明細書で開示される別の実施形態は、命令を記憶する非一時的コンピュータ可読記憶媒体を含み、該命令は、コンピュータによって実行されたときに、該コンピュータに、有用性統計を生成することを容易にするための方法を行わせる。本方法の動作中に、コンピュータは、２つのコンピューティングデバイスとのネットワーク接続を確立する要求を受信することができる。コンピュータは、ネットワーク接続をセットアップすることができ、また、セッションＩＤを割り当てることができる。コンピュータは、２つのコンピューティングデバイスにセッションＩＤを送信することができる。コンピュータは、セッションＩＤを有する第１のコンピューティングデバイスからデータを受信し、通知を有する第２のコンピューティングデバイスにデータを送信して、プロトコルの次のラウンドを行うことができる。

【図面の簡単な説明】

【００２５】

【図１】一実施形態に従う、分類器を強化するプロトコルを行ってカイ二乗値を計算するための例示的なシステムを例示するブロック図である。

【図２Ａ】一実施形態に従う、サーバによってプロトコルを行うための例示的な方法を例示するフローチャートである。

【図２Ｂ】一実施形態に従う、サーバによってプロトコルを行うための例示的な方法を例示するフローチャートである。

【図３】一実施形態に従う、データ受信側パーティによってプロトコルのラウンド１を行うための例示的な方法を例示するフローチャートである。

【図４】一実施形態に従う、データ所有者によってプロトコルのラウンド２を行うための例示的な方法を例示するフローチャートである。

【図５】一実施形態に従う、データ受信側パーティによってプロトコルのラウンド３を行うための例示的な方法を例示するフローチャートである。

【図６】一実施形態に従う、データ所有者によってプロトコルのラウンド４を行うための例示的な方法を例示するフローチャートである。

【図７】一実施形態に従う、データ所有者によって代替のプロトコルのラウンド２を行うための例示的な方法を例示するフローチャートである。

【図８】一実施形態に従う、データ受信側パーティによって代替のプロトコルのラウンド３を行うための例示的な方法を例示するフローチャートである。

【図９】一実施形態に従う、データ所有者によって代替のプロトコルのラウンド４を行うための例示的な方法を例示するフローチャートである。

【図１０】一実施形態に従う、分類器を改善するデータを適用するための例示的な方法を

例示するフローチャートである。
【図１１】一実施形態に従う、データ受信側パーティと関連付けられた例示的な装置を例示するブロック図である。
【図１２】一実施形態に従う、データ所有者と関連付けられた例示的な装置を例示するブロック図である。
【図１３】一実施形態に従う、分類器強化システムを容易にする例示的なコンピュータ及び通信システムを例示する図である。
【００２６】
図面中、同じ参照番号は、同じ図面要素を指す。
【発明を実施するための形態】
【００２７】
以下の説明は、任意の当業者が本実施形態を作製し、使用することを可能にするように提示され、また、特定の用途及び要件の文脈において提供される。開示された実施形態に対する種々の変更は、当業者には容易に明らかになるであろうし、また本明細書で定義された一般的な原理は、本開示の趣旨及び範囲から逸脱することなく、他の実施形態及び用途に適用され得る。したがって、開示される本システムは、示される実施形態に限定されるのではなく、本明細書で開示される原理及び特徴と一致する最も広い範囲を与えられるべきである。
【００２８】
概要
本明細書で説明される実施形態は、特徴データの有用性を定量化する統計をセキュアに計算するために２つのコンピューティングシステムが行うことができるプロトコルを導入することによって、分類器を改善するという問題を解決する。統計は、コンピューティングシステムの一方からの特徴データが、もう一方のコンピューティングシステムの分類器を改善するために有用であるという可能性を示す。いくつかの実施形態において、統計は、$\chi^2$値である。さらに、実施形態において、コンピューティングデバイスは、暗号化されたデータを使用して$\chi^2$計算を行うことができ、それによって、データのプライバシーを維持する。$\chi^2$値は、カイ二乗値とも称される場合があることに留意されたい。
【００２９】
下の開示は、特徴データを明らかにすることなく、バイナリ分類問題において、特徴の有用性をセキュアに計算するための効果的なプロトコルを説明する。プロトコルを行うパーティが該パーティのデータを使用した算出の結果をセキュアに暗号化し、暗号解読鍵を提供しないので、プロトコルは、信頼できるサードパーティの媒介を必要としない。具体的には、本開示は、２つのパーティの間の４ラウンドのプロトコルを提示する。例えば、２つのパーティは、特徴ベクトルのデータ所有者、及び特徴ベクトルを潜在的に取得することができる別のパーティ（例えば、データ受信側パーティ）とすることができる。データ所有者は、バイナリの特徴ベクトルを有することができ、データ受信側パーティは、バイナリクラスベクトルを有することができ、データ受信側パーティは、データ所有者の特徴ベクトルが、該データ所有者の分類の正確性を改善することができるかどうかを学習することを望む。考慮する有用性は、データ所有者によって共有されるデータが、データ受信側パーティの既存のデータセットの分類を改善することが予期されるかどうかである。データ所有者及びデータ受信側パーティは、データ受信側パーティの目の前のタスクに関するデータの値を決定することができる。
【００３０】
プロトコルは、潜在的データ受信側パーティに特徴ベクトル自体を明らかにすることなく、特徴ベクトルを使用して、データ受信側パーティの分類器を改善するための有用性の尺度を明らかにする。プロトコルはまた、データ受信側パーティのデータをデータ所有者に明らかにしない。２つのパーティは、実際のデータにアクセスすることなく、分類器を改善するためのデータの有用性をセキュアに測定することができる。データの潜在的受け側は、共有データ辞書内の仕様以外のデータに関してはそれ以上何も学習しない。本明細

書で説明される実施形態は、プライバシー保護計算のために、Ｐａｌｌｉｅｒ準同型暗号化を利用することができる。プロトコルは、正直であるが奇妙な（ｈｏｎｅｓｔ－ｂｕｔ－ｃｕｒｉｏｕｓ）攻撃者に対する証明可能なセキュリティを有する。

【００３１】

本明細書で開示される開示及び実施例は、利用可能な全ての特徴に基づく分類を有する構造データセットを伴うことができる。具体的には、実施例は、２つのパーティ、Ｃａｒｏｌ、及びＦｅｌｉｘを含むことができる。Ｃａｒｏｌは、利用可能な特徴$f^{(1)}$、・・・、$f^{(l)}$に従って生成される特定の特徴及びクラスベクトルを含むデータセットを有することができ、ここで、$f^{(j)}$は特徴ベクトルである。Ｆｅｌｉｘは、分類を改善する際にＣａｒｏｌに対して有用であり得る特徴列ｆを保有することができる。パーティは、特徴ベクトル及びクラスベクトルのための共通インデックス鍵を共有することができることに留意されたい。

【００３２】

表記法。$c=(c_1, c_2, \cdots, c_n)$をＣａｒｏｌによるクラスラベルベクトルとし、$f=(f_1, f_2, \cdots, f_n)$をＦｅｌｉｘによる特徴ベクトルとする。いくつかの実施形態において、クラスラベル及び特徴はどちらもバイナリ属性である。すなわち、全ての$1 \leqq i \leqq n$について、$c_i \in \{0, 1\}$及び$f_i \in \{0, 1\}$である。$c_i$は、Ｃａｒｏｌのデータセットにおけるｉ番目の記録のクラス変数を表すものとする。$f_i$は、Ｃａｒｏｌのデータセットのｉ番目の記録に対応する、Ｆｅｌｉｘの特徴ベクトルにおける特徴値とする。

【００３３】

$\chi^2$の特徴選択

特徴選択は、無情報特徴を取り除き、良好な予測変数を構築するのに有用である特徴のサブセットを選択するプロセスである。特徴選択の基準は、アプリケーションによって異なり得る。例えば、しばしば、ピアソン相関係数を使用して、直線回帰の依存関係を検出し、相互情報を共通に使用して、離散または正規特徴をランク付けする。

【００３４】

本開示は、バイナリ特徴の有用性を判定することに重点を置いている。いくつかの実施形態において、本システムは、その幅の広い適用性及びその暗号化ツールに対する従順性のため、有用性の尺度として、$\chi^2$統計を使用することができる。より具体的には、対数関数的計算を対象とする相互情報とは異なり、$\chi^2$統計の算出は、追加及び重複だけを対象とし、これらは、暗号化ツールに対してより従順である。種々の実施形態はまた、$\chi^2$統計と異なる有用性手段も使用することができ、また、マルチクラス分類器を一連のバイナリ分類器に分解し、次いで、開示されるプロトコル／方法を使用することができる。

【００３５】

クラスラベルベクトルｃ及び対応する特徴ベクトルｆを考える。Ａは、$f_i=0$及び$c_i=0$を有する行の数である。Ｂは、$f_i=0$及び$c_i=1$を有する行の数である。Ｃは、$f_i=1$及び$c_i=0$を有する行の数である。Ｄは、$f_i=1$及び$c_i=1$を有する行の数である。表１は、ｆ及びｃの双方向分割表を示す。ｆ及びｃの$\chi^2$統計は、次のように定義することができる。

【００３６】

【数１】

$$\chi^2(\mathbf{f}, \mathbf{c}) = \frac{n(\mathrm{AD} - \mathrm{BC})^2}{(\mathrm{A} + \mathrm{C})(\mathrm{A} + \mathrm{B})(\mathrm{C} + \mathrm{D})(\mathrm{B} + \mathrm{D})}$$

【００３７】

【表１】

| f \ c | 0 | 1 |
|---|---|---|
| 0 | A | B |
| 1 | C | D |

表１　ｆ及びｃの双方向分割表

【００３８】

システムは、$\chi^2$（ｆ，ｃ）を使用して、ｆ及びｃの独立性を試験することができる。表２は、異なる$\chi^2$値の下での独立性仮説を拒否する信頼度を示す。例えば、$\chi^2$（ｆ，ｃ）が１０．８３よりも大きいときに、独立性仮説は、９９．９％を超える信頼度で拒否することができ、これは、特徴ベクトルｆが、クラスラベルベクトルｃと相関している可能性が非常に高いことを示す。

【００３９】

【表２】

| $\chi^2$（f，c） | 信頼度 |
|---|---|
| １０．８３ | ９９．９％ |
| ７．８８ | ９９．５％ |
| ６．６３ | ９９％ |
| ３．８４ | ９５％ |
| ２．７１ | ９０％ |

表２　異なる$\chi^2$値の下での独立性仮説を拒否する信頼度

【００４０】

暗号化ツール

ＰＫＥスキーム及びＣＰＡセキュリティ。下記は、公開鍵暗号化（ＰＫＥ）スキーム及び選択された選択平文攻撃（ＣＰＡ）セキュリティの標準的な定義の説明であり、これらは、本開示において使用される。

【００４１】

ＰＫＥスキーム。メッセージ空間Ｍを有するスキームＰＫＥは、３つの確率的多項式時間（ＰＰＴ）技法Ｇｅｎ、Ｅｎｃ、Ｄｅｃを含むことができる。鍵生成技法Ｇｅｎ（$1^k$）は、公開鍵ｐｋ及び秘密鍵ｓｋを出力する。暗号化技法Ｅｎｃ（ｐｋ，ｍ）は、ｐｋ及びメッセージｍ∈Ｍをとり、暗号テキストｃを出力する。暗号解読技法Ｄｅｃ（ｓｋ，ｃ）は、ｓｋ及び暗号テキストｃをとり、メッセージｍを出力する。正確性のために、いくつかの実施形態は、全ての（ｐｋ，ｓｋ）についてＤｅｃ（ｓｋ，ｃ）＝ｍ←Ｇｅｎ（$1^k$）、全てのｃ←Ｅｎｃ（ｐｋ，ｍ）、及び全てがｍ∈Ｍであることを必要とし得る。

【００４２】

無視可能関数。あらゆる可能な整数ｃについて、全てｘ＞Ｎの場合に

【００４３】

【数２】

$$|f(x)| \leq \frac{1}{x^c}$$

【００４４】

となるような整数Ｎが存在する場合は、

【００４５】

【数３】

$$\text{関数} f : \mathbb{N} \rightarrow \mathbb{R}$$

【００４６】

を無視することができる。無視可能関数は、ｎｅｇｌ（・）によって表すことができる。

【００４７】

ＣＰＡ実験。ＰＫＥスキームＰＫＥに対する攻撃者Ａによる平文選択攻撃（ＣＰＡ）ゲームを下で説明する。

【００４８】

［表］

| |
|---|
| 技法１　$\mathrm{PubK}^{\mathrm{CPA}}_{A,\mathrm{PKE}}$ 実験 |
| 入力：セキュリティパラメータｋ |
| １：（ｐｋ，ｓｋ）←Ｇｅｎ（$1^k$） |
| ２：攻撃者Ａには、$1^k$、ｐｋ、及び$\mathrm{Enc}_{pk}$（・）へのオラクルアクセスが与えられる |
| Ａは、同じ長さの一対のメッセージ（$m_0$，$m_1$）を出力する |
| ３：一様なビットｂ∈｛０，１｝が選択され、ｃ←$\mathrm{Enc}_{pk}$（$m_b$）がＡに与えられる |
| ４：Ａは、引き続き$\mathrm{Enc}_{pk}$（・）へのアクセスを有し、ビットｂ’を出力する |
| 出力：ｂ’＝ｂの場合は１、それ以外の場合は０ |

【００４９】

ＣＰＡセキュリティ。全ての確率的多項式時間攻撃者Ａについて、次式のような無視可能関数ｎｅｇｌが存在する場合、ＰＫＥスキームＰＫＥ＝（Ｇｅｎ、Ｅｎｃ、Ｄｅｃ）は、平文選択攻撃下で、区別できない暗号化を有するか、またはＣＰＡセキュアである。

【００５０】

【数４】

$$\Pr\left[\mathrm{PubK}^{\mathrm{CPA}}_{A,\mathrm{PKE}}(k) = 1\right] \leq \frac{1}{2} + \mathrm{negl}(k)、$$

【００５１】

式中、実験

【００５２】

【数５】

$$\mathrm{PubK}^{\mathrm{CPA}}_{A,\mathrm{PKE}}$$

【００５３】

は、技法１において定義され、確率は、Ａ及び実験のランダム性に引き継がれる。

【００５４】

Ｐａｉｌｌｉｅｒ暗号化スキーム。ＧｅｎＭｏｄｕｌｕｓを多項式時間技法とし、これは、入力$1^k$に対して、（Ｎ、ｐ、ｑ）を出力し、ここで、Ｎ＝ｐｑであり、ｐ及びｑは、ｋビット素数である（ただし、ｐまたはｑが、ｋにおいて無視可能な確率で素数でないことを除く）。以下の暗号化スキームを定義する。

－Ｇｅｎ：$1^k$の入力に対して、ＧｅｎＭｏｄｕｌｕｓ（$1^k$）を動作させて、（Ｎ、ｐ、ｑ）を取得する。公開鍵は、ｐｋ＝Ｎであり、秘密鍵は、ｓｋ＝＜Ｎ，Φ（Ｎ）＞である。

－Ｅｎｃ：公開鍵Ｎ及びメッセージ

【００５５】

【数６】

$$m \in \mathbb{Z}_N$$

【００５６】

の入力に対して、一様な

【００５７】

【数７】

$$r \leftarrow \mathbb{Z}_N^*$$

【００５８】

を選択し、次の暗号テキストを出力する

【００５９】

ｃ：＝［（１＋Ｎ）$^{m}$・ｒ$^{N}$　ｍｏｄ　Ｎ$^{2}$］。

－秘密鍵＜Ｎ、Φ（Ｎ）＞及び暗号テキストｃに対して、次式を計算する。

【００６０】

【数８】

$$m := \left[ \frac{\left[ c^{\phi(N)} \bmod N^2 \right] - 1}{N} \cdot \phi(N)^{-1} \bmod N \right]$$

【００６１】

Ｐａｉｌｌｉｅｒ暗号化。本開示は、Ｐａｉｌｌｉｅｒ暗号化スキームを使用して、２パーティ特徴選択技法においてプライバシーを保護し、Ｐａｉｌｌｉｅｒ暗号化の付加準同型特性を用いて、$\chi^2$統計を算出する。Ｐａｉｌｌｉｅｒ暗号化スキームは、付加準同型及びスカラー乗算をサポートする。下記は、加法準同型及びスカラー乗算準同型の定義である。実施形態は、Ｐａｉｌｌｉｅｒ暗号化スキームに限定されず、さらに、種々の実施形態は、他の暗号化スキームも利用することができる。

【００６２】

スカラー乗算準同型。ＰＫＥスキームＰＫＥ＝（Ｇｅｎ、Ｅｎｃ、Ｄｅｃ）は、バイナリ演算

【００６３】

【数９】

$$\oplus$$

【００６４】

が存在する場合、加算準同型であると言われ、よって、全てのｋ∈Ｎについて、及び全てのｍ$_1$、ｍ$_2$∈Ｍについて、次式が成り立つ

【００６５】

【数１０】

$$\Pr\left[ m^* = m_1 + m_2 \;\middle|\; \begin{array}{c} (pk, sk) \leftarrow \mathsf{Gen}(1^k) \\ c_1 \leftarrow \mathsf{Enc}_{pk}(m_1), c_2 \leftarrow \mathsf{Enc}_{pk}(m_2) \\ c^* \leftarrow c_1 \oplus c_2 \\ m^* \leftarrow \mathsf{Dec}_{sk}(c^*) \end{array} \right] = 1 - \mathrm{negl}(k)。$$

【００６６】

スカラー乗算準同型。ＰＫＥスキームＰＫＥ＝（Ｇｅｎ、Ｅｎｃ、Ｄｅｃ）は、バイナリ演算

【００６７】

【数１１】

$$\otimes$$

【００６８】

が存在する場合、スカラー乗算準同型であると言われ、よって、全てのｋ∈Ｎについて、及び全てのｍ$_1$、ｍ$_2$∈Ｍについて、次式が成り立つ

【００６９】

【数１２】

$$\Pr\left[m^* = m_1 m_2 \middle| \begin{array}{c}(pk, sk) \leftarrow \mathsf{Gen}(1^k) \\ c \leftarrow \mathsf{Enc}_{pk}(m_2) \\ c^* \leftarrow m_1 \otimes c \\ m^* \leftarrow \mathsf{Dec}_{sk}(c^*)\end{array}\right] = 1 - \mathrm{negl}(k)。$$

【００７０】

分類器を改善するための例示的なシステム

図１は、一実施形態に従う、分類器を強化するプロトコルを行ってカイ二乗値を計算するための例示的なシステムを例示するブロック図を提示する。図１は、プロトコル計算のラウンドを行うために２つのパーティが使用することができる、システム１００を表す。ウェブサーバ１０２は、通信ネットワークを通じてクライアントコンピュータ１０４、１０６と通信することができる。ウェブサーバ１０２は、ウェブサービスを使用してχ$^2$を計算するためのプロトコルを管理することができる。クライアントコンピュータ１０４、１０６は、ウェブサーバにアクセスするブラウザを含むことができる。クライアントコンピュータは、暗号化計算を行うことができる。２つのパーティは、ブラウザセッションを開き、ウェブサーバ１０２に接続し、そして、該パーティが、プロトコルに従ってデータを共有する有用性を評価したいことに同意することができる。ウェブサーバ１０２は、セッションＩＤを割り当てて、該セッションＩＤを両ユーザと共有することができる。ウェブサービスは、セッションＩＤをデータベースに記憶して、プロトコルラウンドの進捗を追跡するために使用することができる。

【００７１】

ラウンド１を行う役割を果たすパーティ（例えば、クライアントコンピュータ１０４）は、例えばＨＴＴＰＳ　ＰＯＳＴを使用して、計算を行い、セッションＩＤと共に、結果をウェブサーバに提出することができる。ウェブサーバは、結果をデータベース１０８に記憶し、プロトコルの次のラウンドを行う必要がある旨の通知と共に、結果をもう一方のパーティに伝送することができる。

【００７２】

第２のパーティ（例えば、クライアントコンピュータ１０６）は、ラウンド２の計算を行い、次いで、セッションＩＤを有するウェブサーバに結果を記録することができる。上述のように、ウェブサーバは、結果をデータベース１０８に記憶し、次のラウンドを行うための通知と共に、結果を第１のパーティに伝送することができる。

【００７３】

ウェブサーバ１０２は、他のクライアントから入手可能なラウンドの出力があるときに、クライアントへの通知を継続することができる。ラウンドが使用可能であると通知されたときに、各パーティは、次の計算ラウンドを行い、ウェブサーバ１０２に結果を送信することができる。プロセスは、４つのラウンド全てが完了するまで継続する。

【００７４】

随意に、ウェブアプリケーションは、例えばＪａｖａＳｃｒｉｐｔ（登録商標）及びウェブソケットツールを使用して、クライアント側のラウンドの計算を含むことができ、よって、ユーザエクスペリエンスは、単にブラウザにデータへのアクセスを与え、次いで、全てのラウンドが完了するまで待機する。この事例において、システムは、ユーザのデータがサーバにアップロードされない旨をユーザに通知することができる。いくつかの実施形態において、クライアントは、ピアツーピアネットワークにおいて互いに直接通信することによってプロトコルを行うことができ、ウェブサーバは、不要である。

【００７５】

いくつかの実施形態において、分類器は、セキュリティデータを分類することができる

。本明細書で開示される技法は、分類器を改善することができる。例えば、開示される技法は、ネットワークイベントが部外者からのネットワークへの攻撃であるのか、または通常のネットワークアクティビティであるのか等の、ネットワークイベントの分類を改善することができる。分類器はまた、企業におけるユーザアクティビティ等のユーザアクティビティを分類して、ユーザが企業へのインサイダー攻撃を行っているのか、または通常のネットワークユーザーアクティビティを行っていのるかを判定することができる。

【００７６】

いくつかの実施形態において、分類器の機能は、ネットワークルータに実装することができる。ルータは、ネットワークイベントを自動的に分類し、ネットワークルーティングを変更すること、警報を発すること、または別様には管理者に通知することによって、疑わしいネットワークイベントに応答することができる。ルータは、本明細書で説明される技法を使用して、他のネットワークデバイスと自動的に接続して、セキュリティデータを取引し、受信し、または転送し、そして、分類器を自動的に改善することができる。

【００７７】

プロトコルによって、パーティは、データを伝送する前に、データが分類器を改善する可能性があるかどうかを判定することができる。これは、データが分類器を改善する可能性がないとパーティが判定した場合に、データ転送を生じさせる必要がないので、ネットワークリソースの使用を低減させることができる。いくつかのシナリオにおいて、データは、ネットワーク帯域幅を消費し得る大量のデータであり得、したがって、データを伝送する前に、データが分類器を改善する可能性があるかどうかを最初に決定することが有益であり得る。

【００７８】

分類器はまた、ルールに基づく分類器とすることもできる。システムは、分類器のルールを追加または変更することによって、分類器の一組のルールを改善することができる。データを取得するプロトコルを使用し、分類器のルールを修正することで、システムは、分類器が以前に行うことができなかった機能を分類器が行うことができるように、分類器を改善することができる。

【００７９】

パーティは、サードパーティを雇うことなく、取引データの有用性を評価することができるが、雇う場合は、時間及び金を消費し、かつサードパーティが信用できることを必要とする。データプライバシーは、サードパーティから保護することができる。

【００８０】

有用性を計算するための迅速な方法によって、パーティは、データの提供者がデータの潜在的受け側の間で選択する、受け側が潜在的データ提供者の間で選択する、及び必要に応じて、パーティがデータの脱感作方法の間で選択する等の、より多くのオプションを効率的に考慮することができる。

【００８１】

サーバによって行われるプロトコル

図２Ａ～２Ｂは、一実施形態に従う、サーバによってプロトコルを行うための例示的な方法を例示するフローチャートである。サーバは、図２Ａ～２Ｂの動作を行うことができる。動作中に、サーバは、２つのコンピューティングデバイスとの接続を確立する要求を受信することができる（動作２０２）。サーバは、２つのコンピューティングデバイスとの接続をセットアップすることによって応答することができる。サーバは、セッションＩＤを２つのコンピューティングデバイスに割り当てて、該セッションＩＤを共有することができる（動作２０４）。サーバは、セッションＩＤをデータベースに記憶することができる。

【００８２】

サーバは、次いで、セッションＩＤを有するデータ受信側パーティからデータを受信することができる（動作２０６）。このデータは、暗号化された公開鍵、及び暗号化された特徴ベクトルのクラスラベル要素、ならびに他の暗号化されたデータを含むことができる

。サーバは、データをデータベースに記憶することができる。

【００８３】

サーバは、次いで、プロトコルの次のラウンドを行う必要がある旨の通知と共に、データ所有者に受信データを送信することができる（動作２０８）。サーバは、セッションＩＤを有するデータ所有者からデータを受信することができる。例えば、受信データは、$Enc_{pk}(rD)$を含むことができる。サーバは、データをデータベースに記憶することができる（動作２１０）。サーバは、その後に、プロトコルの次のラウンドを行う必要がある旨の通知と共に、データ受信側パーティに受信データを送信することができる（動作２１２）。

【００８４】

サーバは、次いで、セッションＩＤを有するデータ受信側パーティからデータを受信することができる。サーバは、データをデータベースに記憶することができる（動作２１４）。例えば、データは、一対の暗号化された値を含むことができる。

【００８５】

【数１３】

$$\left(\mathrm{Enc}_{pk}\left(\frac{r^2D^2}{(B+D)(A+C)}\right),\ \mathrm{Enc}_{pk}\left(\frac{rD}{A+C}\right)\right)$$

【００８６】

サーバは、次いで、プロトコルの次のラウンドを行う必要がある旨の通知と共に、データ所有者に受信データを送信することができる（動作２１６）。

【００８７】

サーバは、次いで、セッションＩＤを有するデータ所有者からデータ（例えば、暗号化されたカイ二乗値）を受信することができる（動作２１８）。サーバは、データをデータベースに記憶することができる。サーバは、データ受信側パーティに受信データを送信することができる（動作２２０）。サーバ及び２つのコンピューティングデバイスは、プロトコルの４つのラウンドを完了した。

【００８８】

例示的なシナリオ

下記は、本明細書で説明されるプロトコル及び図１に関して説明されるシステムを使用する、異なる使用事例シナリオの説明である。

【００８９】

実施例１：この実施例において、データ受信側パーティは、データ所有者からデータの列を取得することに関心がある。データ受信側パーティは、テーブルを使用して訓練される分類器を改善するために、データをそのデータテーブルに追加することを意図している。所有者は、データを提供する意思があるが、データ受信側パーティは、実際にデータが分類器を改善することを確信していない。データ受信側パーティは、所有者のデータのためのデータ辞書にアクセスし、共通インデックスキーを共有するが、データ受信側パーティは、データ要素の実際の値を知らない。パーティは、データ受信側パーティがラウンド１を行うことに関して、プロトコルを使用することに同意する。プロトコルに従った後に、データ受信側パーティは、$\chi^2$統計の値を有するが、該値は、所有者が提供する意思があるデータの有用性の尺度である。データ受信側パーティが重要であるとみなす任意の他の情報と共に、$\chi^2$値に基づいて、データ受信側パーティは、データを取得するかどうかを判定する。いくつかのシナリオにおいて、データ所有者は、データ受信側パーティにデータを販売するか、または他のデータについて取引することができる。

【００９０】

実施例２：この実施例において、パーティは、プロトコルに従うことに同意するが、２ストリームモードにおいてである。２ストリームモードは、ラウンド１において、両パーティが暗号化された値を計算し、送信することから始まり、ラウンド４において、両パーティが暗号化された$\chi^2$値を送信する。$\chi^2$に関する計算は、ｃ及びｆに関して対称的で

あるので、両パーティは、同じ$\chi^2$値にならなければならないので、両パーティが、両ストリームにおいて同じデータを使用する（すなわち、不正を行わない）ことを前提とする。この実施例の場合、第５のラウンドが存在してもよく、該ラウンドにおいて、各パーティは、パーティ自体の秘密鍵によって$\chi^2$を解読した後に、新しい使い捨て鍵によって$\chi^2$を再暗号化する。この使い捨て鍵は、対称的な暗号化キーとすることができる。次いで、各パーティは、もう一方のパーティに再暗号化された$\chi^2$を送信する。各パーティはまた、使い捨て鍵を送信することもできるが、もう一方のパーティから再暗号化された$\chi^2$メッセージを受信した後にだけである。各パーティは、再暗号化された$\chi^2$及び関連付けられた使い捨て鍵の交換を行って、各パーティが同じ$\chi^2$値に注目していることを確認する。

【００９１】

各パーティは、もう一方のパーティからの使い捨て鍵によって再暗号化された$\chi^2$メッセージを解読し、ラウンド４においてもう一方のパーティによって解読された$\chi^2$が、両パーティが受信し、ラウンド４において解読したものと同じであることを確認することができる。この追加のラウンドは、両パーティが、同じ$\chi^2$値に注目して、２つの暗号化されたドメイン計算ストリームにおいて異なるデータを使用することによって一方のパーティが不正を行ったという可能性を排除することを両パーティに保証する。そのような保証がなければ、データ所有者は、潜在的なデータ受信側パーティが、データ受信側パーティに対してデータの正確な値を隠蔽するために、データ所有者の$\chi^2$について偽の低い有用性のデータを使用したと懸念する場合がある。このプロトコルが使用された後に、$\chi^2$をセキュアなままにしておく必要がある場合、使い捨て鍵は、伝送する前に、もう一方のパーティの公開鍵によって暗号化することができ、よって、サードパーティの攻撃者は、使い捨て鍵にアクセスできない。

【００９２】

実施例３：この実施例において、データの所有者は、（例えば、販売、ライセンス付与、取引、及び／または転送することによって）特徴列へのアクセスを提供する意思があるが、所有者は、受け側（例えば、ライセンシー）の数をｎ＝５またはｎ＝２等の少ない数ｎに制限している。この実施例において、所有者は、データのうちの最も高い有用性を得る受信側パーティを選択したい場合がある。データ所有者及び／またはデータ受信側パーティは、データ受け側（例えば、ライセンシー）の数を制限して、機密データに対するリスクを最小化したい場合がある（受け側が多くなることは、より多くのことが間違った方向に進むことを意味する）。データ所有者及び／またはデータ受信側パーティはまた、それらの競争相手及び一般大衆からデータを遠ざけることを保証することによって、データ受信側パーティに対する値を最大にしたい場合もある。データは、該データへのアクセスが制限されるときに、しばしば、より有益及び／または有用である。この実施例において、パーティは、プロトコルに従うことに同意し得るが、先の実施例において説明した２ストリームモードを使用する。いくつかの使用事例シナリオにおいて、データ所有者は、より高い価格を支払う意思があるので、データから最も高い有用性を得る限られた数の受け側に、特徴列へのアクセスを販売することができる。

【００９３】

いくつかのシナリオにおいて、データ所有者は、潜在的ライセンシーのための$\chi^2$値を確立することができ、データ所有者は、そのグループからライセンシーを獲得すること、使用分野に関する任意の購入者条件を受諾すること、ライセンシーの最大数、購入者の競争者に販売しないこと、その他を選択することができる。価格は、固定価格または価格対有用性のスケジュールのように、予め設定することができ、またはパーティは、ライセンシーを獲得することを選択した後に、交渉することができる。潜在的ライセンシーは、それらをライセンシーとして考慮することから除外することができるので、真の値を低く評価する有用性計算プロトコルにおいて該データを使用しないように動機付けされる。データ所有者は、それらのデータがより有益であることを示すので、最良の有用性数をもたらすデータを使用するように動機付けされる。したがって、両パーティは、有用性計算プロ

トコルを「ゲーミング」しないことによって得るべきものを有する。

【００９４】

実施例４：この実施例において、データの所有者は、（例えば、アクセスを与える、または販売することによって）特徴列へのアクセスを提供する意思があるが、所有者は、統計的プライバシーをデータに適用して、機密情報を保護することが必要である。多くの異なる統計的プライバシー方法が存在するので、所有者及びデータ受信側パーティは、セキュリティと有用性との間の許容可能なトレードオフを提供する方法（ならびに対応する方法パラメータ及び設定）を選択したい場合がある。この実施例において、パーティは、プロトコルに従うことに同意し得るが、実施例２において説明される２ストリームモードを使用して、データ所有者が多数のデータ受け側の間の考慮事項を管理することを補助する。パーティは、互いに関心のある一組の方法、パラメータ値、及び設定にわたる、一組の統計的プライバシー変形例に関する$\chi^2$値を計算することができる。ＡＲＸ、オープンソースのデータ匿名化ツール等の、データ所有者が使用して統計的プライバシーをデータに適用するためのいくつかのツールが存在する。ＡＲＸは、機密個人データを匿名化するためのオープンソースソフトウェアである。結果に基づいて、データ所有者は、どの方法及びパラメータ値が許容可能なセキュリティを提供するのかを通信することができ、潜在的データ受信側パーティは、どの選択が許容可能な有用性を提供するのかを通信することができる。パーティは、次いで、使用する方法に関して合意に到達することができ、データ所有者は、その方法を適用し、データ受信側パーティに秘密化されたデータを提供することができる。統計的プライバシー方法及びそれらのパラメータの例は、下の表３に与えられる。

【００９５】

【表３】

| 統計的プライバシー方法 | パラメータ及び設定（角括弧の範囲） |
|---|---|
| ε－δ差分プライバシー | ε、δ、一般化レベル |
| ｋ－匿名 | Ｋ |
| ｋ－マップ | Ｋ、推定器の種類（ポアソン、ゼロ切り捨てポアソン、なし）、推定器の有意水準 |
| δ－存在 | $\delta_{最小}$、$\delta_{最大}$ |
| 平均再識別リスク | 閾値 |
| 母集団一意性 | 閾値、モデル（Ｄａｎｋａｒ、Ｐｉｔｍａｎ、Ｚａｙａｔｚ、ＳＮＢ） |
| サンプル一意性 | 閾値 |

表３　統計的プライバシー方法及びそれらのパラメータの実施例

【００９６】

実施例５：この実施例において、データの所有者は、特徴列のサブセットへのアクセスを提供する意思があるが、データの拡散を制御するために、任意の１つの受け側に一組全てのデータへのアクセスを与えたくない。この実施例において、パーティは、プロトコルに従うことに同意するが、実施例２において説明される２ストリームモードを使用して、データ所有者が多数のデータ受け側の間の考慮事項を管理することを補助する。パーティは、所有者が公表する（または、ライセンスを供与する）意思があるデータの異なるサブセットの$\chi^2$値を計算する。パーティは、次いで、どのサブセットを公表する（またはライセンスを供与する）のかに関して合意に到達する。オプションとして、所有者が、サブセットを公表する前に、統計的プライバシーをサブセットに適用することを必要とする場合、パーティは、実施例４のプロセスに従って、統計的プライバシー方法を選択すること

ができる。
【００９７】
２ストリームバージョン
実施例６：先の実施例にあるような２ストリーム事例の場合、両パーティは、各ラウンドを完了することができ、そして、実施例２において説明されるような第５のラウンドを追加することができる。ウェブサーバ１０２が２つのパーティのためのセッションをセットアップするときに、ウェブサーバ１０２は、データベースに、セッションが２ストリーム方法によって行われることを示すデータを記憶することができる。
【００９８】
いくつかの実施形態において、ウェブアプリケーションは、例えばクライアントコンピュータ上で実行されるＪａｖａＳｃｒｉｐｔ（登録商標）または類似のコードによって、クライアント側の計算を行うので、ユーザは、ウェブアプリケーションの外側で各ラウンドを手動で実行する必要はない。この実施形態において、ユーザエクスペリエンスは、クライアント側環境にデータを一度ロードし、最終結果を待機することである。ユーザに対して、ウェブアプリケーションは、任意の必要な情報の計算、暗号化された結果の送信、及びもう一方のパーティからの新しい結果の準備ができたときの注意を取り扱う。計算は、クライアントにおいて行われるので、各パーティのデータは、もう一方のパーティ及びウェブサーバ自体からセキュアな状態を維持する。
【００９９】
いくつかの実施形態において、システムは、文書化されたアプリケーションプログラミングインタフェース（ＡＰＩ）を伴う中央サーバを有し、これは、クライアントシステムが、第２のパーティによるセッションをプログラム的に開始し、その妥当性を検査すること、計算のラウンドの結果を送達すること、ならびに、結果を受信すること、及び／またはもう一方のパーティからの結果を待機しているときにステータスをチェックすることを可能にする。次いで、カスタムまたはサードパーティのクライアント側ソフトウェアは、ブラウザに基づく手法に代わるものとして、これらのＡＰＩを直接使用することができる。
【０１００】
ウェブサーバは、各セッションの間、データベースを維持することができ、いつラウンドが開始されたのか、及びいつそのラウンドの結果が返されたのかを記録する。これは、ウェブサービスが、トランザクションを完了するようにリマインダを送信するようなアクションを行うこと、ラウンドが二度以上行われているかどうかを識別し、それが意図的であるかどうかを解明すること、及び随意に、各ラウンドにおいて交換された、暗号化されたデータを記憶することを可能にする。ウェブサービスはまた、ラウンドが完了したときに、情報の交換を終了させることもできる。
【０１０１】
いくつかの実施形態において、ウェブサービスは、マーケットプレースを含むことができ、そこにおいて、データの所有者は、データ辞書または同等物によって、（例えば、可能なライセンス付与、取引、または別用にはデータの提供に対して）データの所有者が利用できるデータをポストすることができる。データを取得したい（例えば、データまたは取引データに対するライセンスを購入したい）他の者は、要求をポストすることができ、または何が提供されているのかを検索することができる。例えば、異なるパーティは、セキュリティ関連のデータを取引して、それらの分類器を改善することができ、それによって、コンピュータネットワークにおけるセキュリティイベントの分類及び検出を改善する。パーティのコンピューティングデバイスは、計算を自動的に行って、互いからセキュリティ関連のデータを取得して、それらの分類器を改善することができる。これは、分類器を生成するための従来の手法よりも効率的である。ウェブサービスはまた、ユーザがアカウントを要求するときにユーザから収集されるデータを通して、ユーザ識別子の認証も提供することができる。
【０１０２】

実施例７：この実施例において、パーティは、（例えば、電子的な）オークションによって、データに対するライセンスを販売することができる。オークションの前に、潜在的入札者は、プロトコルに従うことができ、該潜在的入札者は、ラウンド１を行う。ラウンド４の後に、潜在的入札者は、自分の分類器アプリケーションの$\chi^2$値を有する。潜在的入札者の数、及び異なる$\chi^2$評価の数は、オークショニアによって制限することができる。これは、公開されるべきではないデータに関する情報をリバースエンジニアリングする目的で多数の$\chi^2$評価を行っている攻撃者のリスクを最小にする。オークションは、従来の「イングリッシュ」オークション、ダッチオークション、ヴィックリーオークションン、または別の種類のオークションとすることができる。いくつかの実施形態は、実施例６において説明されるウェブサービスを詳述する、ウェブに基づくオークションシステムを含むことができる。

【０１０３】

下の開示は、２パーティ設定の下で$\chi^2$統計を算出するための４ラウンドプロトコルを説明する。実施例において、一方のパーティは、Ｃａｒｏｌと命名され、特徴ベクトルｃを有する。もう一方のパーティは、Ｆｅｌｉｘであり、特徴ベクトルｆを有する。Ｃａｒｏｌの目的は、$\chi^2$（ｆ，ｃ）を学ぶことであり、Ｆｅｌｉｘの目的は、ｆに関する任意のさらなる情報を明らかにしないことである。いくつかの実施形態において、ｃのクラスラベル及びｆの特徴は、バイナリ属性である。すなわち、全ての$1 \leqq i \leqq n$について、$c_i \in \{0, 1\}$及び$f_i \in \{0, 1\}$である。Ａは、$f_i = 0$及び$c_i = 0$を有する行の数である。Ｂは、$f_i = 0$及び$c_i = 1$を有する行の数である。Ｃは、$f_i = 1$及び$c_i = 0$を有する行の数である。Ｄは、$f_i = 1$及び$c_i = 1$を有する行の数である。

【０１０４】

プロトコル

ラウンド１

図３は、一実施形態に従う、データ受信側パーティによってプロトコルのラウンド１を行うための例示的な方法を例示するフローチャートを例示する。データ受信側パーティＣａｒｏｌは、図３の動作を行うことができる。動作中に、データ受信側パーティは、下で説明される計算を行うことができる。

１．Ｐａｉｌｌｉｅｒ鍵の対（ｐｋ，ｓｋ）＝Ｇｅｎ（$1^k$）を生成する（動作３０２）。

２．ｐｋを有する全てのクラスラベルを暗号化する：$Enc_{pk}(c_1)$、$Enc_{pk}(c_2)$、・・・、$Enc_{pk}(c_n)$（動作３０４）。

３．中間値

【０１０５】

【数１４】

$$\frac{B+D}{A+C}$$

【０１０６】

を計算する。

【０１０７】

【数１５】

$$B + D = \sum_{i=1}^{n} c_i$$

【０１０８】

であり、かつＡ＋Ｃ＝ｎ－（Ｂ＋Ｄ）であるので、Ｃａｒｏｌは、分割表に基づいて、

【０１０９】

【数１６】

$$\frac{\sum_{i=1}^{n} c_i}{n-\sum_{i=1}^{n} c_i}$$

【０１１０】

を計算することによって、この値を取得することができることに留意されたい（動作３０６）。

４．ｐｋを有する中間値

【０１１１】

【数１７】

$$\frac{B+D}{A+C}$$

【０１１２】

を暗号化する：

【０１１３】

【数１８】

$$pk: \mathrm{Enc}_{pk}\left(\frac{B+D}{A+C}\right)$$

【０１１４】

（動作３０８）。

５．データ所有者Ｆｅｌｉｘに、非対称鍵の対（ｐｋ，ｓｋ）からの公開鍵ｐｋ、暗号化した特徴ベクトルの要素、及び暗号化した中間値を送信する：

【０１１５】

【数１９】

$$\left(pk, \mathrm{Enc}_{pk}(c_1), \mathrm{Enc}_{pk}(c_2), \ldots, \mathrm{Enc}_{pk}(c_n), \mathrm{Enc}_{pk}\left(\frac{B+D}{A+C}\right)\right)$$

【０１１６】

（動作３１０）。

【０１１７】

ラウンド２

図４は、一実施形態に従う、データ所有者によってプロトコルのラウンド２を行うための例示的な方法を例示するフローチャートを例示する。データ所有者（Ｆｅｌｉｘ）は、図４の動作を行うことができる。

【０１１８】

動作中に、データ所有者は、データ受信側パーティＣａｒｏｌから、公開鍵ｐｋ及び暗号化されたデータ（例えば、暗号化された特徴ベクトルの要素及び暗号化された中間値）を受信することができる（動作４０２）。

【０１１９】

データ所有者は、下で説明するように計算を行うことができる。

１．$\mathrm{Enc}_{pk}$（Ｄ）を計算する（動作４０４）。

データ所有者は、次式を計算することによってこの値を取得することができる。

【０１２０】

【数２０】

$$\oplus_{i=1}^{n}\left(f_i \otimes \mathrm{Enc}_{pk}(c_i)\right) = \oplus_{i=1}^{n} \mathrm{Enc}_{pk}(f_i c_i) = \mathrm{Enc}_{pk}\left(\sum_{i=1}^{n} f_i c_i\right)$$

ここで、$\sum_{i=1}^{n} f_i c_i = D$である。

【０１２１】

２．
【０１２２】
【数２１】

$$r \leftarrow \mathbb{Z}_N$$

【０１２３】
をサンプリングし、
【０１２４】
【数２２】

$$r \otimes \mathrm{Enc}_{pk}(\mathrm{D}) = \mathrm{Enc}_{pk}(r\mathrm{D})$$

【０１２５】
を計算する（動作４０６）。
３．データ受信側パーティＣａｒｏｌに以下の値を送信する：Ｅｎｃ$_{ｐｋ}$（ｒＤ）（動作４０８）。
【０１２６】
ラウンド３
図５は、一実施形態に従う、データ受信側パーティによってプロトコルのラウンド３を行うための例示的な方法を例示するフローチャートを提示する。データ受信側パーティＣａｒｏｌは、図５の動作を行うことができる。動作中に、データ受信側パーティは、暗号化された値Ｅｎｃ$_{ｐｋ}$（ｒＤ）を受信することができる（動作５０２）。
【０１２７】
データ受信側パーティは、以下の計算を行うことができる。
１．ｓｋを使用してＥｎｃ$_{ｐｋ}$（ｒｄ）を解読する（動作５０４）。
２．一対の値
【０１２８】
【数２３】

$$\frac{r^2\mathrm{D}^2}{(\mathrm{B}+\mathrm{D})(\mathrm{A}+\mathrm{C})} \text{及び} \frac{r\mathrm{D}}{\mathrm{A}+\mathrm{C}}$$

【０１２９】
を計算し、それらを暗号化する（動作５０６）：
【０１３０】
【数２４】

$$\mathrm{Enc}_{pk}\left(\frac{r^2\mathrm{D}^2}{(\mathrm{B}+\mathrm{D})(\mathrm{A}+\mathrm{C})}\right) \text{及び} \mathrm{Enc}_{pk}\left(\frac{r\mathrm{D}}{\mathrm{A}+\mathrm{C}}\right)$$

【０１３１】
３．データ所有者Ｆｅｌｉｘに、以下の一対の暗号化された値を送信する（動作５０８）：
【０１３２】
【数２５】

$$\left(\mathrm{Enc}_{pk}\left(\frac{r^2\mathrm{D}^2}{(\mathrm{B}+\mathrm{D})(\mathrm{A}+\mathrm{C})}\right),\ \mathrm{Enc}_{pk}\left(\frac{r\mathrm{D}}{\mathrm{A}+\mathrm{C}}\right)\right)$$

【０１３３】
ラウンド４
図６は、一実施形態に従う、データ所有者によってプロトコルのラウンド４を行うための例示的な方法を例示するフローチャートを提示する。動作中に、データ所有者Ｆｅｌｉｘは、データ受信側パーティから、一対の暗号化された値を受信することができる（動作６０２）。データ所有者Ｆｅｌｉｘは、特徴ベクトルｆ及び一対の暗号化された値に基づ

いて、暗号化されたχ２値を計算することができる（動作６０４）。
データ所有者Ｆｅｌｉｘは、以下の計算を行うことができる。
１．次式を計算することによってｒをキャンセルする
【０１３４】
【数２６】

$$r^{-2}\otimes Enc_{pk}\left(\frac{r^2D^2}{(B+D)(A+C)}\right)=Enc_{pk}\left(\frac{D^2}{(B+D)(A+C)}\right)$$

及び

$$r^{-1}\otimes Enc_{pk}\left(\frac{rD}{A+C}\right)=Enc_{pk}\left(\frac{D}{A+C}\right)$$ 。

【０１３５】
２．次式を計算することによってχ２（ｆ，ｃ）の暗号化を計算する：
【０１３６】
【数２７】

$$\left(\frac{n^3}{(A+B)(C+D)}\otimes Enc_{pk}\left(\frac{D^2}{(B+D)(A+C)}\right)\right)\oplus\left(\frac{n(C+D)}{A+B}\otimes Enc_{pk}\left(\frac{B+D}{A+C}\right)\right)\oplus$$
$$\left(\frac{-2n^2}{A+B}\otimes Enc_{pk}\left(\frac{D}{A+C}\right)\right),$$

【０１３７】
式中、Ｃ＋Ｄ及びＡ＋Ｂは、次式のように計算される。
【０１３８】
【数２８】

$$C+D=\sum_{i=1}^{n}f_i,$$

【０１３９】
及び
Ａ＋Ｂ＝ｎ－（Ｃ＋Ｄ）。
以下に示すように、上記の計算は、Ｅｎｃ$_{pk}$（χ２（ｆ，ｃ））を与える。
ＡＤ－ＢＣ＝（Ａ＋Ｂ＋Ｃ＋Ｄ）Ｄ－（Ｂ＋Ｄ）（Ｃ＋Ｄ）なので、χ２（ｆ，ｃ）は、次式のように分解することができる：
【０１４０】
【数２９】

$$\chi^2(f,c)=\frac{n(AD-BC)^2}{(A+C)(A+B)(C+D)(B+D)}$$
$$=\frac{n^3}{(A+B)(C+D)}\frac{D^2}{(B+D)(A+C)}+\frac{n(C+D)}{(A+B)}\frac{(B+D)}{(A+C)}-\frac{2n^2}{(A+B)}\frac{D}{(A+C)}$$ 。

【０１４１】
３．Ｃａｒｏｌに以下の値を送信する（動作６０６）：
Ｅｎｃ$_{pk}$（χ２（ｆ，ｃ））
【０１４２】
代替のプロトコル
いくつかの実施形態において、システムは、乗算的盲検化ではなく、加法的盲検化を使用して、乱数ｒをラウンド２に導入することができ、乗算的ではなく加法的加法準同型の長所を利用する。これらの実施形態の場合、システムは、上で説明したようなラウンド１、及び下で説明するようなラウンド２～ラウンド４を行うことができる。
【０１４３】
ラウンド２

図７は、一実施形態に従う、データ所有者によって代替のプロトコルのラウンド２を行うための例示的な方法を例示するフローチャートを例示する。データ所有者（Ｆｅｌｉｘ）は、図７の動作を行うことができる。

【０１４４】

動作中に、データ所有者は、データ受信側パーティＣａｒｏｌから、公開鍵ｐｋ及び暗号化されたデータ（例えば、暗号化された特徴ベクトルの要素及び暗号化された中間値）を受信することができる（動作７０２）。

【０１４５】

データ所有者は、下で説明するように計算を行うことができる。

１．$Enc_{pk}$（Ｄ）を計算する（動作７０４）。

データ所有者は、次式を計算することによってこの値を取得することができる。

【０１４６】

【数３０】

$$\oplus_{i=1}^{n}\left(f_i \otimes Enc_{pk}(c_i)\right) = \oplus_{i=1}^{n} Enc_{pk}(f_i c_i) = Enc_{pk}(\sum_{i=1}^{n} f_i c_i)、$$

ここで、$\sum_{i=1}^{n} f_i c_i = D$である。

【０１４７】

２．

【０１４８】

【数３１】

$$r \leftarrow \mathbb{Z}_N$$

【０１４９】

をサンプリングし、

【０１５０】

【数３２】

$$r \oplus Enc_{pk}(D) = Enc_{pk}(r + D)$$

【０１５１】

を計算する（動作７０６）。

３．データ受信側パーティＣａｒｏｌに以下の値を送信する：$Enc_{pk}$（ｒ＋Ｄ）（動作７０８）。

【０１５２】

ラウンド３

図８は、一実施形態に従う、データ受信側パーティによって代替のプロトコルのラウンド３を行うための例示的な方法を例示するフローチャートを提示する。データ受信側パーティＣａｒｏｌは、図８の動作を行うことができる。動作中に、データ受信側パーティは、暗号化された値$Enc_{pk}$（ｒ＋Ｄ）を受信することができる（動作８０２）。

【０１５３】

データ受信側パーティは、以下の計算を行うことができる。

１．鍵ｓｋを使用して$Enc_{pk}$（ｒ＋Ｄ）を解読する（動作８０４）。

２．５つ一組の値を計算し、

【０１５４】

【数３３】

$$\frac{(r+D)^2}{(B+D)(A+C)}、\ \frac{(r+D)}{(B+D)(A+C)}、\ \frac{(r+D)}{(A+C)}、\ \frac{1}{(B+D)(A+C)}、\ \frac{1}{(A+C)}$$

【０１５５】

公開鍵ｐｋを使用して５つ一組の値を暗号化する（動作８０６）：

【０１５６】
【数３４】

$$\mathrm{Enc}_{pk}\left(\frac{(r+D)^2}{(B+D)(A+C)}\right)$$

$$\mathrm{Enc}_{pk}\left(\frac{(r+D)}{(A+C)}\right)$$

$$\mathrm{Enc}_{pk}\left(\frac{(r+D)}{(B+D)(A+C)}\right)$$

$$\mathrm{Enc}_{pk}\left(\frac{1}{(B+D)(A+C)}\right)$$

$$\mathrm{Enc}_{pk}\left(\frac{1}{(A+C)}\right)$$

【０１５７】
３．データ所有者Ｆｅｌｉｘに、５つ一組の暗号化された値を送信する（動作８０８）。
【０１５８】
ラウンド４
図９は、一実施形態に従う、データ所有者によって代替のプロトコルのラウンド４を行うための例示的な方法を例示するフローチャートを提示する。動作中に、データ所有者Ｆｅｌｉｘは、データ受信側パーティから、５つ一組の暗号化された値を受信することができる（動作９０２）。データ所有者Ｆｅｌｉｘは、特徴ベクトルｆ及び５つ一組の暗号化された値の一組に基づいて、暗号化されたχ$^2$を計算することができる（動作９０４）。
【０１５９】
データ所有者Ｆｅｌｉｘは、以下の計算を行うことができる。
１．次式を計算することによって、最初の２つの暗号化された値からｒを除去する
【０１６０】
【数３５】

$$\mathrm{Enc}_{pk}\left(\frac{(r+D)^2}{(B+D)(A+C)}\right)\oplus\left(r^2\otimes\mathrm{Enc}_{pk}\left(\frac{1}{(B+D)(A+C)}\right)\right)$$

$$\oplus\left(-2r\otimes\mathrm{Enc}_{pk}\left(\frac{(r+D)}{(B+D)(A+C)}\right)\right)=\mathrm{Enc}_{pk}\left(\frac{D^2}{(B+D)(A+C)}\right)$$

及び

$$\mathrm{Enc}_{pk}\left(\frac{(r+D)}{(A+C)}\right)\oplus\left(-r\otimes\mathrm{Enc}_{pk}\left(\frac{1}{A+C}\right)\right)=\mathrm{Enc}_{pk}\left(\frac{D}{A+C}\right)$$

【０１６１】
２．次式を計算することによってχ$^2$（ｆ，ｃ）の暗号化を計算する：
【０１６２】
【数３６】

$$\left(\frac{n^3}{(A+B)(C+D)}\otimes\mathrm{Enc}_{pk}\left(\frac{D^2}{(B+D)(A+C)}\right)\right)\oplus\left(\frac{n(C+D)}{A+B}\otimes\mathrm{Enc}_{pk}\left(\frac{B+D}{A+C}\right)\right)\oplus$$

$$\left(\frac{-2n^2}{A+B}\otimes\mathrm{Enc}_{pk}\left(\frac{D}{A+C}\right)\right),$$

【０１６３】

式中、Ｃ＋Ｄ及びＡ＋Ｂは、次式のように計算される

【０１６４】

【数３７】

$$C + D = \sum_{i=1}^{n} f_i、$$

【０１６５】

及び

Ａ＋Ｂ＝ｎ－（Ｃ＋Ｄ）。

以下に示すように、上記の計算は、$Enc_{pk}(\chi^2(f, c))$を与える。

ＡＤ－ＢＣ＝（Ａ＋Ｂ＋Ｃ＋Ｄ）Ｄ－（Ｂ＋Ｄ）（Ｃ＋Ｄ）なので、$\chi^2(f, c)$は、次式のように分解することができる：

【０１６６】

【数３８】

$$\chi^2(\mathbf{f},\mathbf{c}) = \frac{n(AD - BC)^2}{(A + C)(A + B)(C + D)(B + D)}$$

$$= \frac{n^3}{(A+B)(C+D)}\frac{D^2}{(B+D)(A+C)} + \frac{n(C+D)}{(A+B)}\frac{(B+D)}{(A+C)} - \frac{2n^2}{(A+B)}\frac{D}{(A+C)}$$

【０１６７】

３．Ｃａｒｏｌに以下の値を送信する（動作９０６）：

$Enc_{pk}(\chi^2(f, c))$

【０１６８】

ローカル計算

データ受信側パーティＣａｒｏｌは、$Enc_{pk}(\chi^2(f, c))$を解読して、$\chi^2(f, c)$を取得する。Ｃａｒｏｌだけが値$\chi^2(f, c)$を受信することに留意されたい。アプリケーションに応じて、Ｆｅｌｉｘも$\chi^2(f, c)$の値を知っている必要がある場合、Ｃａｒｏｌは、プロトコルを動作させた後に、Ｆｅｌｉｘに$\chi^2(f, c)$の値を送信することができる。

【０１６９】

図１０は、一実施形態に従う、分類器を改善するデータを適用するための例示的な方法を例示するフローチャートを提示する。動作中に、データ受信側パーティは、データ所有者から、暗号化された$\chi^2$値を受信することができる（動作１００２）。データ受信側パーティは、暗号化された$\chi^2$値を解読して、$\chi^2$値を取得することができる（動作１００４）。データ受信側パーティは、$\chi^2$値が所定の閾値を超えたと判定することができる（動作１００６）。データ受信側パーティは、データを要求して、データ所有者から受信することができる（動作１００８）。データ受信側パーティは、次いで、データ所有者から受信したデータを適用して、分類器を改善することができる（動作１０１０）。

【０１７０】

例示的な装置

図１１は、一実施形態に従う、データ受信側パーティと関連付けられた例示的な装置１１００を例示するブロック図を提示する。装置１１００は、有線または無線通信チャネルを介して互いに通信することができる、複数のモジュールを備えることができる。装置１１００は、１つまたは２つ以上の集積回路を使用して実現することができ、また、図１１に示されるモジュールよりも少ない、または多いモジュールを含むことができる。さらに、装置１１００は、コンピュータシステムに統合することができ、または他のコンピュータシステム及び／もしくはデバイスと通信することができる別々のデバイスとして実現することができる。

【０１７１】

具体的には、装置１１００は、暗号化モジュール１１０２、計算モジュール１１０４、

接続モジュール１１０６、及びデータ記憶装置１１０８の任意の組み合わせを備えることができる。装置１１００また、図１１に表されない追加的なモジュール及びデータに含むこともでき、異なる実現形態は、異なる一組のモジュールに従って機能を配設することができることに留意されたい。本明細書で開示される実施形態は、任意の特定のモジュールの配設に限定されない。

【０１７２】

いくつかの実現形態は、データの暗号化及び解読、鍵の対の生成、及び特徴ベクトルデータ等のデータの暗号化を含む動作を行うことができる、暗号化モジュール１１０２を含むことができる。計算モジュール１１０４は、中間値の計算等のプロトコルと関連付けられた計算を行うことができる。接続モジュール１１０６は、サーバ及び他のコンピューティングデバイスとの接続を確立することができる。データ記憶装置１１０８は、特徴ベクトルのデータ等の、本明細書で説明されるデータを記憶することができる。

【０１７３】

例示的な装置

図１２は、一実施形態に従う、データ所有者と関連付けられた例示的な装置１２００を例示するブロック図を提示する。装置１２００は、有線または無線通信チャネルを介して互いに通信することができる、複数のモジュールを備えることができる。装置１２００は、１つまたは２つ以上の集積回路を使用して実現することができ、また、図１２に示されるモジュールよりも少ない、または多いモジュールを含むことができる。さらに、装置１２００は、コンピュータシステムに統合することができ、または他のコンピュータシステム及び／もしくはデバイスと通信することができる別々のデバイスとして実現することができる。

【０１７４】

具体的には、装置１２００は、暗号化モジュール１２０２、計算モジュール１２０４、接続モジュール１２０６、及びデータ記憶装置１２０８の任意の組み合わせを備えることができる。装置１２００また、図１２に表されない追加的なモジュール及びデータに含むこともでき、異なる実現形態は、異なる一組のモジュールに従って機能を配設することができることに留意されたい。本明細書で開示される実施形態は、任意の特定のモジュールの配設に限定されない。

【０１７５】

いくつかの実現形態は、データを暗号化及び解読することができる、暗号化モジュール１２０２を含むことができる。計算モジュール１２０４は、プロトコルと関連付けられた計算を行うことができる。接続モジュール１２０６は、サーバ及び他のコンピューティングデバイスとの接続を確立することができる。データ記憶装置１２０８は、本明細書で説明されるデータを記憶することができる。

【０１７６】

例示的なコンピュータ及び通信システム

図１３は、一実施形態に従う、分類器強化システムを容易にする例示的なコンピュータ及び通信システムを例示する。コンピュータ及び通信システム１３０２は、プロセッサ１３０４と、メモリ１３０６と、記憶デバイス１３０８と、を含む。メモリ１３０６は、管理されたメモリとしての役割を果たす揮発性メモリ（例えば、ＲＡＭ）を含むことができ、また、１つまたは２つ以上のメモリプールを記憶するために使用することができる。さらに、コンピュータ及び通信システム１３０２は、ディスプレイデバイス１３１０、キーボード１３１２、及びポインティングデバイス１３１４に結合することができる。

【０１７７】

記憶デバイス１３０８は、アプリケーション１３１６及び１３１８ならびにオペレーティングシステム１３２０等の、いくつかのアプリケーションを記憶することができる。記憶デバイス１３０８はまた、通信モジュール１３２２、データ管理モジュール１３２４、及びセッション管理モジュール１３２６のコードも記憶することができる。通信モジュール１３２２は、クライアントコンピューティングデバイスとのネットワーク接続を確立す

ることができる。データ管理モジュール１３２４は、データを記憶するためのメモリ及び／またはディスク記憶空間を割り当てることを含む、コンピューティングデバイスから受信されるデータを管理することができる。セッション管理モジュール１３２６は、セッションＩＤの確立、セッションＩＤの共有、セッションＩＤの記憶、及びセッションの終了を含む、クライアントコンピューティングデバイスとのセッションを管理することができる。

【０１７８】

動作中に、通信モジュール１３２２等の１つまたは２つ以上のアプリケーションを、記憶デバイス１３０８からメモリ１３０６にロードし、次いで、プロセッサ１３０４によって実行する。プログラムを実行している間、プロセッサ１３０４は、上述した関数を行う。

【０１７９】

この「発明を実施するための形態」において説明されるデータ構造及びコードは、典型的に、コンピュータ可読記憶媒体に記憶され、該コンピュータ可読記憶媒体は、コンピュータシステムによって使用するためのコード及び／またはデータを記憶することができる、任意のデバイスまたは媒体とすることができる。コンピュータ可読記憶媒体としては、揮発性メモリ、不揮発性メモリ、ディスクドライブ、磁気テープ、ＣＤ（コンパクトディスク）、ＤＶＤ（デジタル多用途ディスクまたはデジタルビデオディスク）等の磁気及び光記憶デバイス、または現在知られている、または今後開発されるコンピュータ可読媒体を記憶することが可能な他の媒体が挙げられるが、これらに限定されない。

【０１８０】

「発明を実施するための形態」の節において説明される方法及びプロセスは、上で説明したようなコンピュータ可読記憶媒体に記憶することができるコード及び／またはデータとして具現化することができる。コンピュータシステムが、コンピュータ可読記憶媒体に記憶されたコード及び／またはデータを読み取り、実行したときに、コンピュータシステムは、データ構造及びコードとして具体化され、かつコンピュータ可読記憶媒体内に記憶された方法及びプロセスを実行する。

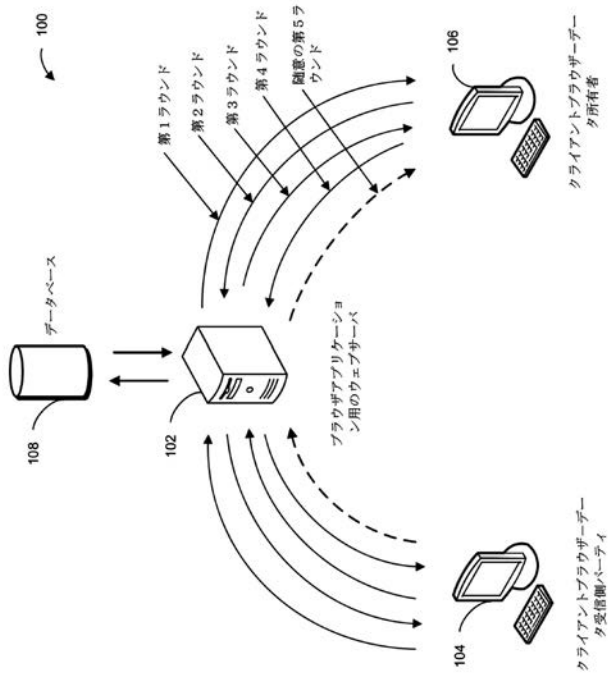
【０１８１】

さらに、上で説明した方法及びプロセスは、ハードウェアモジュールまたは装置に含めることができる。ハードウェアモジュールまたは装置としては、特定用途向け集積回路（ＡＳＩＣ）チップ、フィールドプログラマブルゲートアレイ（ＦＰＧＡ）、特定の時間に特定のソフトウェアモジュールまたはコードの一部を実行する専用または共有プロセッサ、及び現在知られている、または今後開発される他のプログラマブル論理デバイスを挙げることができるが、これらに限定されない。ハードウェアモジュールまたは装置を起動させると、それらの中に含まれる方法及びプロセスを実行する。

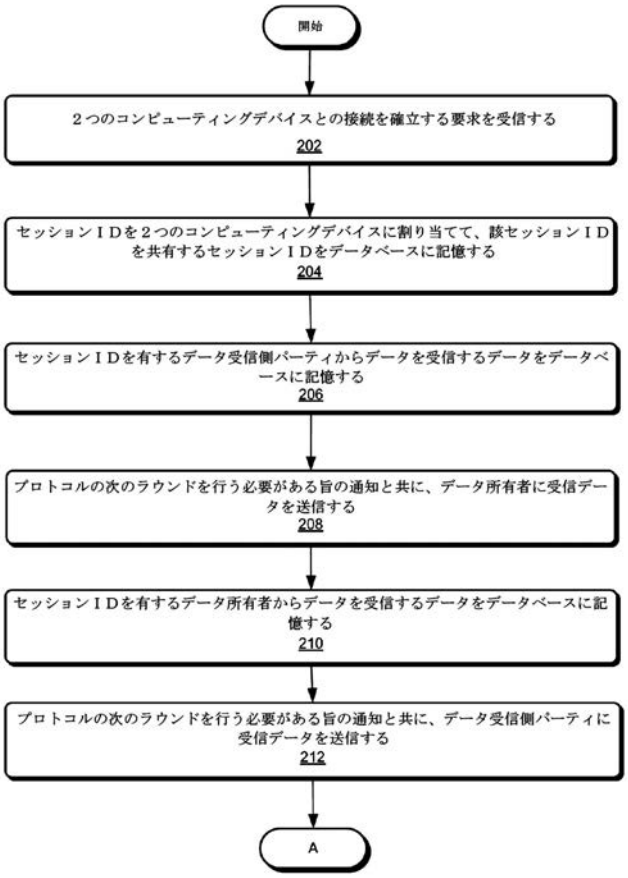
【０１８２】

本明細書で開示される実施形態の上述の説明は、例示及び説明のみを目的として提示されている。本記述は、網羅的であること、または本明細書で開示される実施形態を開示された形態に限定することを意図したものではない。したがって、多くの修正及び変形が、当業者に明白になるであろう。加えて、上述の開示は、本明細書で開示される実施形態を限定することを意図したものではない。本明細書で開示される実施形態の範囲は、添付の特許請求の範囲によって定義される。

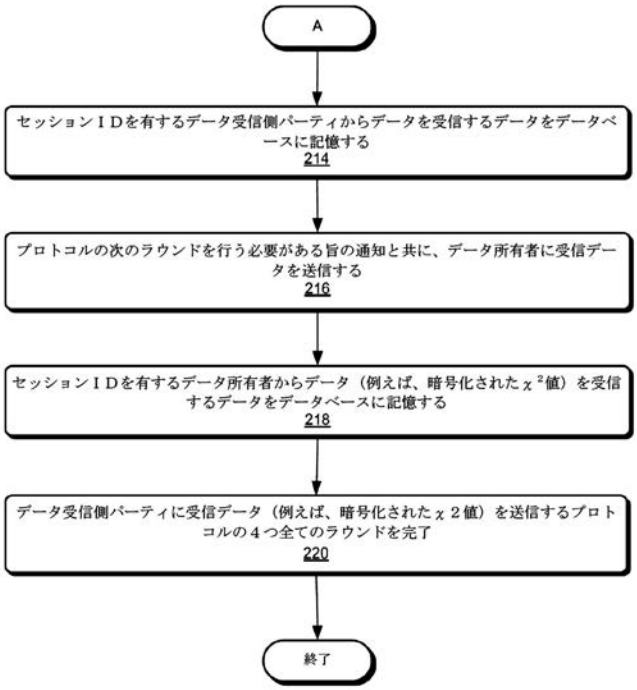
【図１】


100
第１ラウンド
第２ラウンド
第３ラウンド
第４ラウンド
随意の第５ラウンド
106
クライアントブラウザーデータ所有者
データベース
108
102
ブラウザアプリケーション用のウェブサーバ
104
クライアントブラウザーデータ受信側パーティ

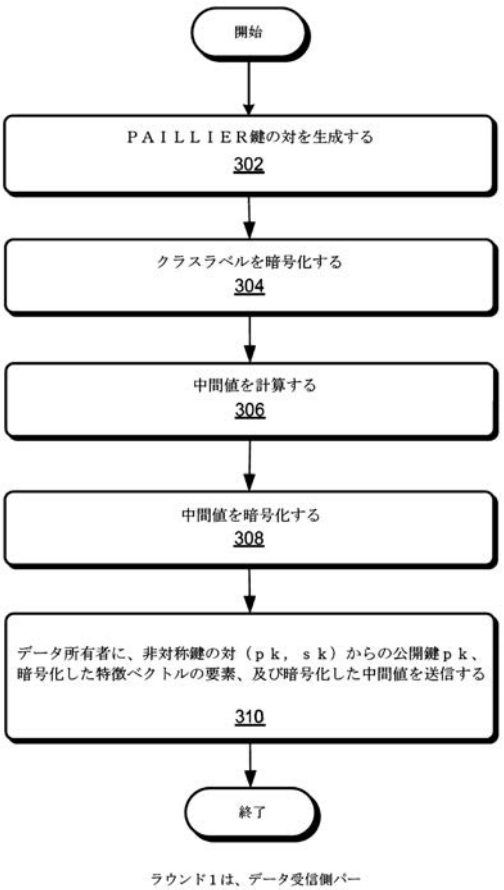
【図２Ａ】


開始
２つのコンピューティングデバイスとの接続を確立する要求を受信する
202
セッションＩＤを２つのコンピューティングデバイスに割り当てて、該セッションＩＤを共有するセッションＩＤをデータベースに記憶する
204
セッションＩＤを有するデータ受信側パーティからデータを受信するデータをデータベースに記憶する
206
プロトコルの次のラウンドを行う必要がある旨の通知と共に、データ所有者に受信データを送信する
208
セッションＩＤを有するデータ所有者からデータを受信するデータをデータベースに記憶する
210
プロトコルの次のラウンドを行う必要がある旨の通知と共に、データ受信側パーティに受信データを送信する
212
A

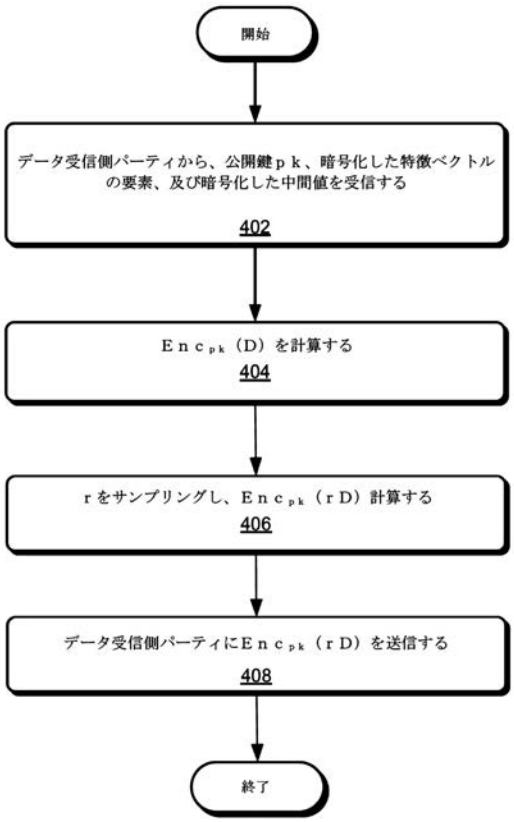
【図２Ｂ】


A
セッションＩＤを有するデータ受信側パーティからデータを受信するデータをデータベースに記憶する
214
プロトコルの次のラウンドを行う必要がある旨の通知と共に、データ所有者に受信データを送信する
216
セッションＩＤを有するデータ所有者からデータ（例えば、暗号化されたχ²値）を受信するデータをデータベースに記憶する
218
データ受信側パーティに受信データ（例えば、暗号化されたχ２値）を送信するプロトコルの４つ全てのラウンドを完了
220
終了

【図３】

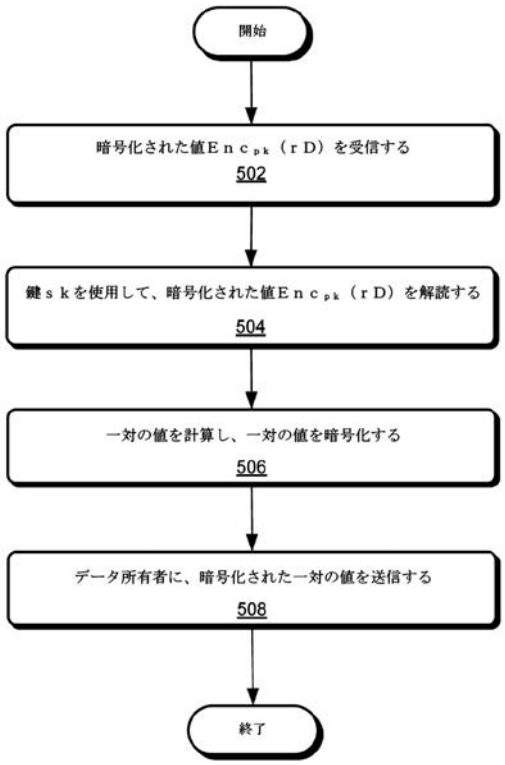

開始
ＰＡＩＬＬＩＥＲ鍵の対を生成する
302
クラスラベルを暗号化する
304
中間値を計算する
306
中間値を暗号化する
308
データ所有者に、非対称鍵の対（ｐｋ，ｓｋ）からの公開鍵ｐｋ、暗号化した特徴ベクトルの要素、及び暗号化した中間値を送信する
310
終了
ラウンド１は、データ受信側パーティによって行われる

【図４】



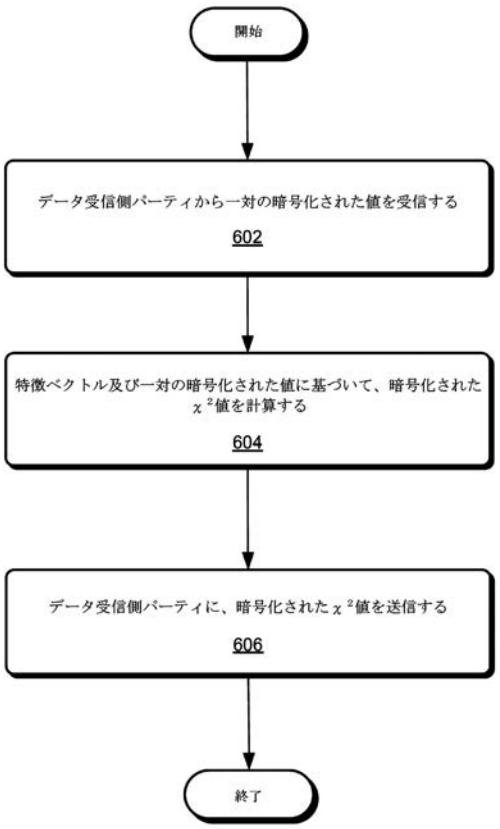

ラウンド２は、データ所有者によって行われる

【図５】



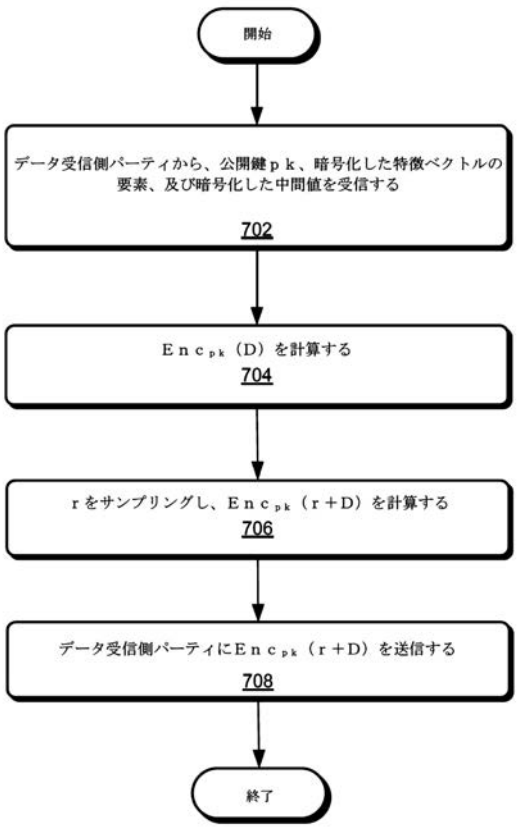

ラウンド３は、データ受信側パーティによって行われる

【図６】



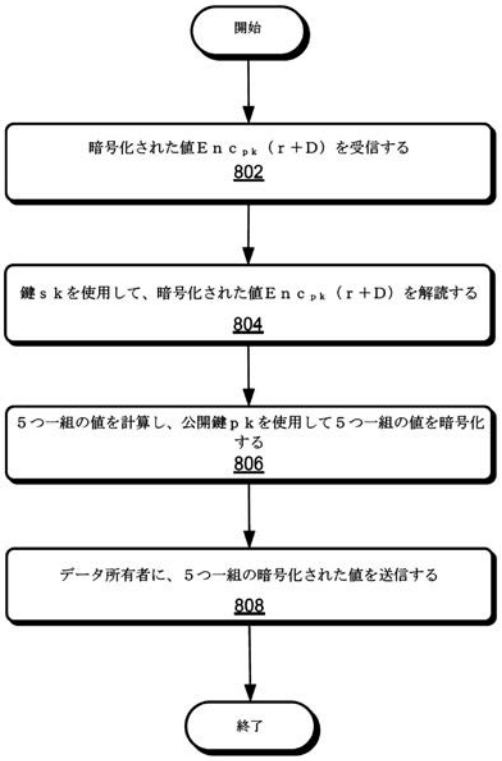

ラウンド４は、データ所有者によって行われる

【図７】



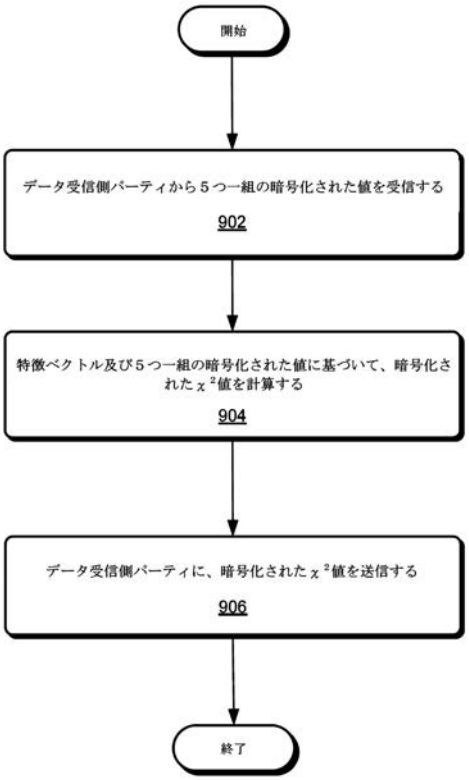

ラウンド２は、データ所有者によって行われる
（代替のプロトコル）

【図８】



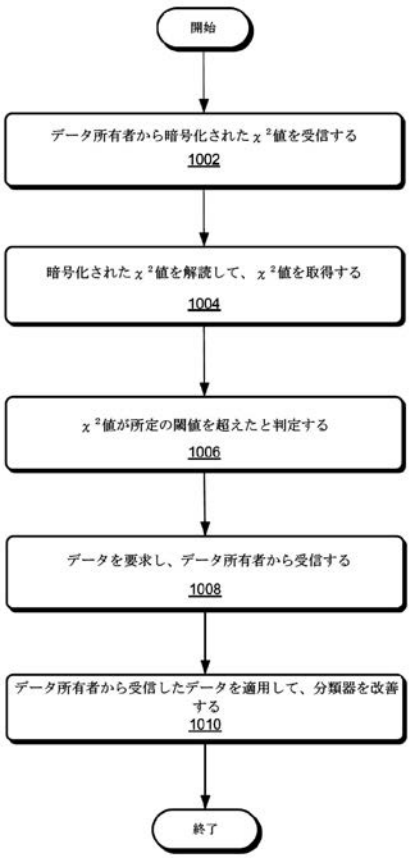

ラウンド３は、データ受信側パーティによって行われる（代替のプロトコル）

【図９】



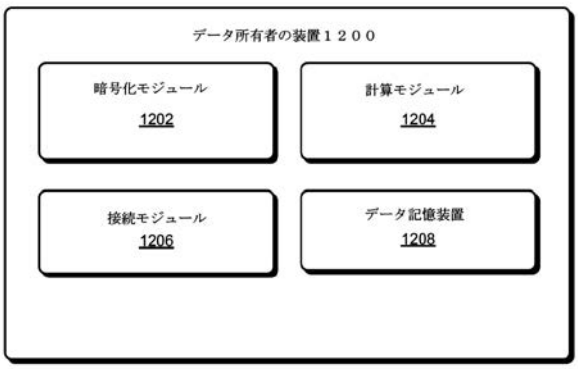

ラウンド４は、データ所有者によって行われる（代替のプロトコル）

【図１０】

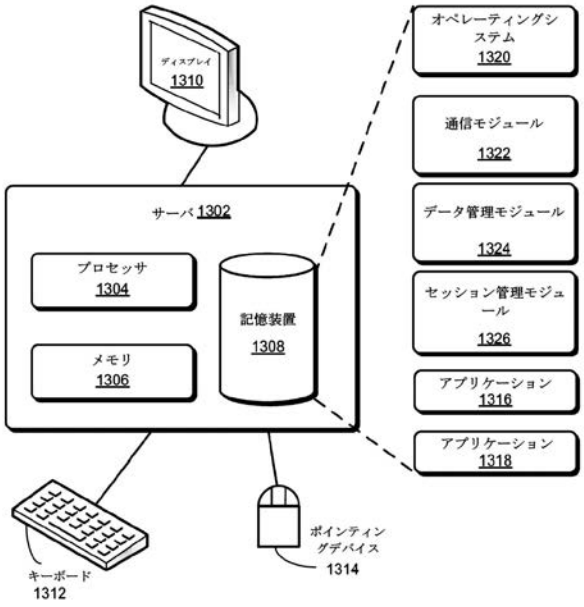



【図１１】


【図１２】



【図１３】


ディスプレイ
1310
サーバ 1302
プロセッサ
1304
メモリ
1306
記憶装置
1308
オペレーティングシステム
1320
通信モジュール
1322
データ管理モジュール
1324
セッション管理モジュール
1326
アプリケーション
1316
アプリケーション
1318
キーボード
1312
ポインティングデバイス
1314

フロントページの続き

(74)代理人 100109070
弁理士 須田 洋之
(74)代理人 100109335
弁理士 上杉 浩
(74)代理人 100120525
弁理士 近藤 直樹
(74)代理人 100139712
弁理士 那須 威夫
(74)代理人 100158551
弁理士 山崎 貴明
(72)発明者 フランシスコ・イー・トレス
アメリカ合衆国 カリフォルニア州 ９５１２０ サンノゼ ミアンダー・ドライブ ５８５７
(72)発明者 ヴァニシュリー・ハヌマンサ・ラオ
アメリカ合衆国 カリフォルニア州 ９４４０２ サンマテオ サウス・ビー・ストリート １３０５
(72)発明者 シャンタヌ・レイン
アメリカ合衆国 カリフォルニア州 ９４０２５ メンロー・パーク シャロン・パーク・ドライブ ６７５
(72)発明者 ユンフイ・ロン
アメリカ合衆国 イリノイ州 ６１８０１ アーバナ ノース・リンカーン・アベニュー ３０５

Ｆターム(参考) 5J104 AA12 AA16 EA04 EA19 JA21 NA02 NA36 NA37 PA07