(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7061042号
(P7061042)

(45)発行日　令和4年4月27日(2022.4.27)

(24)登録日　令和4年4月19日(2022.4.19)

(51)国際特許分類

		FI		
G09C 1/00 (2006.01)		G09C	1/00	660D
G06F 16/00 (2019.01)		G09C	1/00	650Z
		G06F	16/00	

請求項の数　10　（全21頁）

(21)出願番号　特願2018-135417(P2018-135417)
(22)出願日　平成30年7月19日(2018.7.19)
(65)公開番号　特開2019-35949(P2019-35949A)
(43)公開日　平成31年3月7日(2019.3.7)
審査請求日　令和3年7月13日(2021.7.13)
(31)優先権主張番号　15/675,055
(32)優先日　平成29年8月11日(2017.8.11)
(33)優先権主張国・地域又は機関
米国(US)
早期審査対象出願

(73)特許権者　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
(74)代理人　100109070

最終頁に続く

(54)【発明の名称】　暗号化データベースに対する解析を支援するシステムおよびアーキテクチャ

(57)【特許請求の範囲】
【請求項１】
暗号化データベースを生成するためのコンピュータで実行される方法であって、前記方法は、
（ａ）１つ以上の列に平文データエントリを有する平文データベースを受信することと、
（ｂ）拡張された平文データベースを生成するために前記受信した平文データベースを拡張することであって、前記拡張が１つ以上の列を前記受信した平文データベースに追加することを含み、前記追加された各列が条件付きクエリのために利用可能にされるべき属性に対応する、拡張することと、
（ｃ）前記拡張された平文データベースを暗号化して、暗号化データエントリを含む前記暗号化データベースを生成することと
を備え、
前記暗号化データベースが、前記追加された列に対応するそれらの属性についての少なくとも１つの形式の条件付きクエリをサポートし、前記少なくとも１つの形式の条件付きクエリが、暗号化結果を生成するように、その復号なしで前記暗号化データエントリについて計算される、
方法。
【請求項２】
前記暗号化データが、意味的にセキュアな暗号化によって暗号化される、請求項１に記載の方法。

【請求項３】
前記暗号化データが、準同型暗号化技術を使用して暗号化される、請求項２に記載の方法。
【請求項４】
前記準同型暗号化技術が、相加的準同型暗号化技術である、請求項３に記載の方法。
【請求項５】
前記準同型暗号化技術が、２－ＤＮＦ（論理和標準形）演算をサポートする、請求項３に記載の方法。
【請求項６】
前記少なくとも１つの形式の条件付きクエリが、ＷＨＥＲＥクエリまたはＧＲＯＵＰＢＹクエリのうちの１つである、請求項１に記載の方法。
【請求項７】
暗号化データベースを管理するシステムであって、前記システムは、
ハードウエアを介して実装された抽出、転送、および書き込み（ＥＴＬ）サーバであって、前記ＥＴＬサーバは、（ｉ）１つ以上の列において暗号化されていないデータエントリを有する平文データベースを入力として受信し、（ｉｉ）拡張された平文データベースを生成するように前記受信した平文データベースを拡張し、前記拡張された平文データベースが、前記入力された平文データベースに対する１つ以上の列の追加を含み、前記追加された各列が条件付きクエリに利用可能にされるべき属性に対応し、（ｉｉｉ）暗号化データエントリを含む前記暗号化データベースを生成するように前記拡張された平文データベースを暗号化するように動作する、ＥＴＬサーバと、
ハードウエアを介して実装されたデータベース（ＤＢ）サーバであって、前記ＤＢサーバは、（ｉ）前記ＥＴＬサーバから前記暗号化データベースを受信して保持し、（ｉｉ）前記ＤＢサーバに提出されたクエリに応答して暗号化データを返すように動作する、ＤＢサーバと、
ハードウエアを介して実装された計算サーバであって、（ｉ）クエリを前記ＤＢサーバに提出し、（ｉｉ）前記ＤＢサーバから返された暗号化データに関する計算を実行するように動作する、計算サーバと
を備え、前記計算が、前記暗号化データの復号なしに前記暗号化データベースからの前記暗号化データに対して実行され、前記計算から得られた結果が暗号化され、前記暗号化データベースが、前記暗号化データの基礎となる前記暗号化されていないデータのサンプルを明らかにすることなく、少なくとも１つの形式の条件付きクエリに応答して正しい暗号化結果を得ることをサポートするように構成されている、システム。
【請求項８】
前記暗号化データが、意味的にセキュアな暗号化によって暗号化される、請求項７に記載のシステム。
【請求項９】
前記暗号化データが、準同型暗号化技術を使用して暗号化される、請求項８に記載のシステム。
【請求項１０】
前記準同型暗号化技術が、相加的準同型暗号化技術である、請求項９に記載のシステム。
【発明の詳細な説明】
【技術分野】
【０００１】
本明細書の主題は、一般に、データセキュリティの技術に関する。本明細書に開示された例示的な実施形態は、構造化照会言語（ＳＱＬ）データベース（ＤＢ）および／またはデータベース管理システム（ＤＢＭＳ）（例えば、ＭｙＳＱＬなど）と共に特定のアプリケーションを見出し、それらは、ここでは特に参照しながら記載される。しかしながら、本明細書に開示されるものなどの様々な例示的実施形態はまた、他の種類のリレーショナルデータベースならびにグラフィックおよび／または他の種類のデータベースを含む他の同様のアプリケーションにも適用可能であることを理解されたい。

【背景技術】
【０００２】
データセキュリティの分野では、特定の種類の暗号化データベースは不明ではない。しかしながら、従来は、データ解析における設計によって十分なプライバシーを組み込んだ適切なアーキテクチャについての強力なコンセンサスはなかった。例えば、従来のデータベースクエリに適用可能なままであるように暗号化データをどのように配置または記憶されるべきか、どのようにキー管理が実行されるべきか、および／または秘密解析を効率的に実行するためにどのような対策を講じなければならないかなどについての強力なコンセンサスはなかった。
【０００３】
１つの提案されたアーキテクチャは、「ＣｒｙｐｔＤＢ」として知られている。 一般に、異なる暗号化がデータに連続的に適用される「暗号化の玉ねぎ（ｏｎｉｏｎｓ ｏｆ ｅｎｃｒｙｐｔｉｏｎ）」を使用して暗号化フォーマットでデータを記憶する必要がある。したがって、クエリに応じて、適切な暗号化フォーマットでデータにアクセスし、クエリに応答するために十分な計算を実行することができるまで、暗号化のレイヤは削除される（玉ねぎをはがすなど）ことができる。このアプローチは、ＳＱＬクエリのサブセットをサポートするように示されているが、これは、平文データベースの解析に既に熟練しているデータサイエンティストにとっては有益であるが、いくつかの潜在的な欠点がある。
【０００４】
上述したアプローチの潜在的な欠点の１つは、ＣｒｙｐｔＤＢにおいて使用される全ての暗号システムが十分に強力なセキュリティ特性を有するとは限らないということである。例えば、決定論的暗号化および順序保持暗号化は、特定のクエリに応答するのに十分な情報よりも多くの情報を漏洩することがある。特に、ＳＱＬ ＷＨＥＲＥクエリを実行するために、ＣｒｙｐｔＤＢは、決定論的暗号化レイヤにアクセスする必要があり、データの分布および濃度を知っている人に追加情報を漏洩することがある。
【発明の概要】
【０００５】
暗号化データベースを提供および／または利用するための新規なおよび／または改良されたシステムおよび／または方法が本明細書に開示される。
【課題を解決するための手段】
【０００６】
例示的な１つの実施形態によれば、暗号化データベースを生成する方法が提供される。本方法は、１つ以上の列に平文データエントリを有する平文データベースを受信することと、拡張された平文データベースを生成するために受信した平文データベースを拡張することであって、拡張が１つ以上の列を受信した平文データベースに追加することを含み、追加された各列が条件付きクエリのために利用可能にされるべき属性に対応する、拡張することと、拡張された平文データベースを暗号化して、暗号化データエントリを含む暗号化データベースを生成することとを含む。暗号化データベースは、追加された列に対応するそれらの属性についての少なくとも１つの形式の条件付きクエリをサポートし、少なくとも１つの形式の条件付きクエリは、暗号化結果を生成するように、その復号なしで暗号化データエントリについて計算される。
【０００７】
別の例示的な実施形態によれば、暗号化データベースを管理するシステムが提供される。システムは、１つ以上のコンピュータに実装された抽出、転送、および書き込み（ＥＴＬ）であって、ＥＴＬサーバは、（ｉ）１つ以上の列において暗号化されていないデータエントリを有する平文データベースを入力として受信し、（ｉｉ）拡張された平文データベースを生成するように受信した平文データベースを拡張し、前記拡張された平文データベースが、入力された平文データベースに対する１つ以上の列の追加を含み、前記追加された列が条件付きクエリに利用可能にされるべき属性に対応し、（ｉｉｉ）暗号化データエントリを含む暗号化データベースを生成するように拡張された平文データベースを暗号化

するように動作する前記ＥＴＬサーバと、１つ以上のコンピュータに実装されたデータベース（ＤＢ）であって、（ｉ）ＥＴＬサーバから暗号化データベースを受信して保持し、（ｉｉ）ＤＢサーバに提出されたクエリに応答して暗号化データを返すように動作する前記ＤＢサーバと、１つ以上のコンピュータに実装された計算であって、（ｉ）クエリをＤＢサーバに提出し、（ｉｉ）ＤＢサーバから返された暗号化データに関する計算を実行するように動作する前記計算サーバとを含む。適切には、計算は、暗号化データの復号なしに暗号化データベースからの暗号化データに対して実行され、前記計算から得られた結果は暗号化される。暗号化データベースは、暗号化データの基礎となる暗号化されていないデータのサンプルを明らかにすることなく、少なくとも１つの形式の条件付きクエリに応答して正しい暗号化結果を得ることをサポートするように適切に構成されている。

【図面の簡単な説明】

【０００８】

【図１】図１は、本発明の主題の態様にかかる例示的なシステムおよび／またはアーキテクチャを示す概略図である。

【図２】図２は、本発明の主題の態様にかかるデータベースを前処理して暗号化する例示的なプロセスおよび／または方法を示すフローチャートである。

【発明を実施するための形態】

【０００９】

明瞭且つ簡潔にするために、本明細書は、本明細書に提示された好ましいおよび／または他の実施形態にしたがっておよび／またはそれらに適合して修正または変更された範囲を除き、それらの構成または動作に関するさらなる詳細な説明なしに当該技術分野において一般に知られている構造的および／または機能的要素、関連する標準、アルゴリズムおよび／またはプロトコル、および他の構成要素、方法および／またはプロセスを指すものとする。さらに、本明細書に開示された装置および方法は、例を介しておよび図面を参照して詳細に説明される。別段の指定がない限り、図中の同様の符号は、図面全体にわたって同一、類似または対応する要素の参照を示す。開示されて記載された例、配置、構成、構成要素、要素、装置、方法、材料などに対する変更が行われることができ、特定の用途のために所望されることができることが理解される。この開示において、特定の材料、技術、配置などのいかなる特定も、提示された特定の例に関連するか、またはそのような材料、技術、配置などの一般的な説明にすぎない。具体的な詳細または例の特定は、そのように具体的に指定されていない限り、義務的または限定的であると解釈されるように意図されておらず、そうであるべきではない。装置および方法の選択された例は、図面を参照して以下に詳細に開示されて記載される。

【００１０】

本発明の主題の態様によれば、相加的準同型暗号システムによって保護されたＤＢに対して実行可能な計算のセットが拡張される。１つの例示的な実施形態は、ＳＱＬのような言語を使用して照会されることができるリレーショナルＤＢに関する。例えば、学校のＤＢにおける子供の数を合計する以下のＳＱＬクエリを考える：

ＳＥＬＥＣＴ　ＡＶＧ（年齢）ＦＲＯＭ　ｈｏｇｗａｒｔｓ

【００１１】

各世帯における子供の数が、意味的にセキュアな相加的準同型暗号システムを使用して暗号化されている場合、このクエリは、暗号化されたドメインで処理され、暗号化された結果を返す。そして、この結果は、アナリストまたは適切な特権を有する他のユーザによって復号される。

【００１２】

僅かにより複雑なクエリを考える：

ＳＥＬＥＣＴ　ＡＶＧ（年齢）ＦＲＯＭ　ｈｏｇｗａｒｔｓ

ＷＨＥＲＥ　ｇｅｎｄｅｒ＝“女性”；

【００１３】

このクエリにおいて、暗号化されたドメインの合計は、指定された条件、すなわち、性別

属性が「女性」を満たすテーブルにおける特定の行にわたってのみ実行される。この小さな追加であっても、クエリの実行は、もはや簡単ではない。その理由は、意味的にセキュアな暗号化では、テーブルの暗号化された「性別」列における暗号文を調べることだけで、どの行が女性の子供に対応するのかを容易に実現可能に判断することができないからである。

【００１４】

代わりに、性別列のために決定論的暗号化を使用することもできる。しかしながら、これは、男性用および女性用の２種類の暗号文のみをもたらす。これは、意味的にセキュアな相加的準同型暗号システムを使用して暗号化された各生徒の年齢をなおも保持したまま、女性および男性の子供に対応する行を特定することを容易とする。上述した決定論的暗号化アプローチの問題点は、ＤＢサーバのみならず、暗号化されたドメインにおけるクエリを処理している任意の計算ノードにも情報を漏洩する可能性があるということである。実際には、記憶および計算は、クラウドまたはその他のものに基づくことができ、それゆえに、信頼できない当事者によって処理されることがある。例えば、そのようなＤＢにおける選択された平文攻撃（ＣＰＡ）は、データベースにおけるどのエントリが少女に対応し且つ少年に対応するのかを明白に明らかにする。本発明の主題の態様によれば、そのようなプライバシー漏洩を回避するためのシステムおよび／または方法が提案される。より具体的には、上記のような条件付きクエリをセマンティックセキュリティによって処理することができるシステムおよび／または方法が記載される。

【００１５】

本明細書は、プライバシーに敏感なデータに対して解析を実行するためのアーキテクチャを記載する。適切には、アーキテクチャは、ＤＢサーバ、例えば信頼できないＤＢサーバを含む。実際には、ＤＢサーバは、暗号化データを記憶し、ＭｙＳＱＬなどのリレーショナルデータベース管理システムによってバックアップされる。このアーキテクチャはまた、暗号化されたドメインで計算を実行する計算サーバ（例えば、信頼できない計算サーバ）と、ＤＢに問い合わせるウェブサーバまたはウェブサービスとを含む。適切には、計算サーバは、ウェブサーバ／サービスによって提出された全てのクエリについて暗号化結果を返す。実際には、許可されたユーザのみが返された結果を復号することができる。例示的な実施形態において、基礎データのサンプルを発見することなく、集計関数（例えば、総和、線形結合、基本分類器、カウントクエリおよびヒストグラムを含む）が計算される。適切には、本発明の主題の目的（例えば、ＳＱＬクエリのサブセットをサポートする一方で、改善されたプライバシー保証を提供する）は、記載されたシステムおよび／またはアーキテクチャ内で意味的にセキュアな準同型暗号化技術を使用して達成される。

【００１６】

一般に、本明細書において記載されるシステムおよび／またはアーキテクチャは、プライバシー保護データ解析を可能にし、それにより、データセットの所有者またはキュレータは、データセットの暗号化バージョンに対してクエリを実行する能力をユーザ（例えば、アナリストまたはデータサイエンティストなど）を与える。より正確には、開示されたシステムおよび／またはアーキテクチャは、データセットの所有者およびアナリストに以下の機能を適切に提供する：

・データセット所有者には、暗号鍵の生成、データセットの前処理および暗号化の能力が提供される；および

・アナリストには、暗号化形式でデータセットに対してクエリを実行し、クエリ結果を復号するためにデータ所有者から必要な鍵を取得する能力が提供される。

【００１７】

適切には、上述した機能を提供するために、システムは、例えば、図１に示されるような３層アーキテクチャによって設計される。より具体的には、実際には、システムは、（ｉ）例えばウェブサービスおよび／またはウェブサーバ１０などを介して実装するいわゆる「フロントエンド」、（ｉｉ）計算サーバ（ＣＳ）２０を含むいわゆる「バックエンド」、および（ｉｉｉ）例えば、ＭｙＳＱＬまたは別のリレーショナルもしくは他の適切な種

類のＤＢを介して実装されたＤＢサーバを含むＤＢＭＳ３０を含むことができる。
【００１８】
実際には、基本的な実施形態は、以下の要素を含む：（１）複数の暗号化フォーマットのうちの１つ以上でデータにデータを記憶するＤＢサーバ（例えば、ＤＢサーバは、ＭｙＳＱＬ　ＤＢサーバなど）；（２）（例えば、適切なプライバシー保護プロトコルを介して）ＤＢによってサービスされる暗号化データを使用してユーザ／アナリストによって提供されるクエリを実行するＣＳ；および（３）ＤＢに対して行われたクエリをサポートし、作成されたクエリに対応する暗号化結果を受信するウェブサービス／サーバ。
【００１９】
拡張された実施形態において、鍵管理機能を提供する鍵認証局（ＫＡ）４０が含まれる。特に、ＫＡ４０は、ＤＢサーバに記憶された暗号化ＤＢを生成するための公開暗号鍵を提供する。さらに、ＫＡ４０は、プライバシー保護プロトコルを実行するために公開暗号鍵をＣＳ２０に提供する。最後に、ＫＡ４０は、ＣＳ２０から受信したクエリの暗号化結果をユーザが復号するのを可能にするウェブサービス／サーバの許可されたユーザに復号鍵を提供する。
【００２０】
さらに拡張された実施形態において、暗号化ＤＢ（ここではＤｅとして示される）を準備するために、追加の要素および／または要素が提供される。実際には、Ｄｅの準備は、暗号化されていないＤＢ（ここではＤｐとして示される）を入力として開始する。すなわち、Ｄｐにおけるデータ要素は、暗号化されていない形式でまたは平文として最初に記憶および／または保持される。暗号化の前に、Ｄｐのスキーマは、拡張されたスキーマが所望のＳＱＬクエリのサブセットをサポートするように拡張される。拡張されたＤＢ（ここではＤａとして示される）において、データ要素はまた、暗号化されていない形式でまたは平文として記憶／保持される。（結果のデータによる拡張スキーマを有する）Ｄａは、１つ以上の暗号化フォーマットを使用して暗号化されてＤｅを達成する。最後に、Ｄｅは、ＤＢサーバに送信される。
【００２１】
適切には、フロントエンドは、特定のタスクまたは複数のタスクに関する特定のビューおよび／またはユーザインターフェース（ＵＩ）を提供する。実際には、例えば、暗号化データセット３２ｅに対して構造化照会言語（ＳＱＬ）および／またはＳＱＬ様クエリを書き込んで実行するためにアナリストによって使用されるように、アナリストビュー１２および／または適切なＵＩが第１の場合に提供されるのに対して、第２の場合には、平文データセット３２を前処理および暗号化するためにデータ所有者によって使用されるように、抽出、転送および書き込み（ＥＴＬ）ビュー１４および／または適切なＵＩが提供される。実際には、フロントエンドは、ＣＳ２０、アナリストビュー１２、およびＥＴＬビュー１４のＥＴＬサーバ５０と直接対話することができる。適切な実施形態において、ＥＴＬサーバ５０は、データ所有者の制御下で信頼できる「ヘルパー」であってもよい。対照的に、ＣＳ２０は、別個のエンティティとして機能する信頼できないヘルパーであってもよい。適切には、ＣＳ２０は、受信したアナリストのクエリを処理し、ＤＢＭＳ３０と対話する責を負う。データセット３２ｅは、ＤＢＭＳ３０において暗号化されることから、ＣＳ２０は、暗号化データに対するクエリを実行し、フロントエンドを介してアナリストに（暗号化された）結果を返す。例えば鍵認証局（ＫＡ）４０から必要な復号鍵を予め取得したアナリストは、ＣＳ２０からフロントエンドによって受信したデータを復号することによってクエリの結果を復元することができる。
【００２２】
例示的な１つの実施形態において、ＤＢＭＳ３０は、標準的なＭｙＳＱＬ　ＤＢサーバによって実装されると共に、ＣＳ２０およびＥＴＬサーバ５０は、双方とも、（例えば、同じコードベースを使用する）Ｊａｖａ（登録商標）ベースのプログラムによって実装される。適切には、フロントエンドは、Ｊａｖａ（登録商標）ｓｃｒｉｐｔ、ハイパーテキストマークアップ言語（ＨＴＭＬ）およびカスケードスタイルシート（ＣＳＳ）の組み合わ

せとして実装されてもよい。これは、実際には、本明細書に記載された機能、目的、動作および／または目的を達成するのに適した上述した構成要素のために他の実装が使用されることができると述べた。

【００２３】

本明細書において記載される例示的な実施形態によれば、回避することが求められる１つの脅威は、システムにインポートされたデータセット３２に関する情報を学習するデータ所有者およびアナリスト以外の当事者である。これは、ＣＳ２０およびＤＢＭＳ３０における偶発的および意図的な漏洩の双方を含み、双方とも信頼できない可能性がある。適切には、インポートされたデータセット３２は、ＤＢＭＳ３０およびＣＳ２０のいずれにもアクセスできない鍵のもとで暗号化されることから、そのような漏洩は、設計によって防止される。さらに、コンテキストに応じて、システムのアーキテクチャは、例えば以下のようなさらなる脅威に対する保護を提供することができる。

（１）インポートされたデータセット３２にレコードが含まれている個人に関する機密情報を学習するアナリスト；および／または、

（２）アナリストのクエリに関して学習するＣＳ２０、ＤＢＭＳ３０および／またはデータ所有者。

【００２４】

例えば、上記項目（１）に対するガードは、個人のプライバシーを保持することと称され、差分プライバシーなどの技術を使用して達成されることができ、上記項目（２）に対するガードは、クエリプライバシーを保持することと称され、「現実の」クエリを難読化する「特別な」クエリを使用して達成されることができる。

【００２５】

ここで図２をさらに参照すると、例えばＥＴＬサーバ５０を介して平文ＤＢ　D_pが前処理されて暗号化されるプロセスおよび／または方法１００が示されている。実際には、ＤＢ所有者などは、ＥＴＬサーバ５０にアクセスして平文データセット３２を前処理および暗号化するためにＥＴＬビュー１４または他の適切なＵＩを使用するフロントエンドウェブサービス／サーバ１０を使用することができる。

【００２６】

適切には、プロセスまたは方法１００は、入力された平文ＤＢ　D_pを、暗号化されたドメインにおいて条件付きクエリをサポートする暗号化ＤＢ　D_eに変換する。もちろん、適切な意味的にセキュアな準同型暗号システムを使用してD_pの個々の各エントリを単に暗号化するだけで単純にD_eを構築することができるが、上記説明されたように、このアプローチは、所望なように条件付きクエリを効率的にサポートしない。代わりに、本明細書において記載されるように、D_aが最初に構築され、これは、D_pの拡張された平文バージョンであり、次いでD_aは暗号化されてD_eを得る。

【００２７】

示されるように、ステップ１１０において、D_p（平文データセット３２を含む）が、例えばＥＴＬプロセスを使用してＥＴＬサーバ５０に入力される。D_pのデータ要素は、適切に暗号化されず、および／または平文形式で表される。適切には、D_pと共に、条件付きクエリなどに関連して使用するために利用可能とされる属性を特定する仕様が入力される。例えば、上記シナリオにおいて、性別属性は、２つの異なる値、すなわち男性または女性をとるような指定された条件付き属性であってもよい。条件付き属性は、暗号化されたドメインにおいて計算された条件付きクエリおよび／または同様のものの実行に利用可能とされるＤＢの属性である（例えば、入力D_pと共に指定および／または特定される）。

【００２８】

ステップ１２０において、入力D_pのスキーマは、例えば、ＤＢが暗号化された後に拡張スキーマが所望のクエリのセットをサポートするように、D_aを達成するように拡張される。

【００２９】

より具体的には、D_aの拡張および／または生成は、以下のように行われることができる

。結果の拡張平文ＤＢ　Ｄａは、入力Ｄｐと比較していくつかの追加の列を含む。追加の列の数は

【００３０】

【数１】

$$\bar{v} = \prod_{m=1}^{M} v_m$$

【００３１】

であり、Ｍ個の指定された条件付き属性があり、それらの属性のそれぞれがｖｍをとる場合、値ｍを可能とする。ここで、ｍ＝１、２、．．．、Ｍである。実際には、いくつかの属性は、特定のＤＢエントリについて１つの値のみをとることができる。例えば、上述した性別属性は、男性または女性の値をとることができる。他の属性は、特定のデータベースエントリについて複数の値をとることができる。例えば、特定の生徒の趣味属性は、読書、サイクリング、ハーブの研究などの複数の値をとることができる。

【００３２】

１つの適切な実施形態において、ＤｐからＤａを構築するために、以下のサブステップがＭ個の条件付き属性のそれぞれに適用される。以下のサブステップにおいて、ｖｍ個の異なる値をとる条件付き属性Ａｍを考える。ここで、ｍ＝１、２、．．．、Ｍである。そして、各属性Ａｍについて以下とする：

１）Ｄａ＝Ｄｐに設定する。

２）Ｄａにおいてｖｍ個の追加の列を作成する。便宜上、ｖｍ個の追加の列が行列Ｓに属している結果、データベースＤａに追加された追加のｖｍ×ｎ個のエントリを考える。この表記法は、ｉ行目且つｊ列目の追加要素をＳ（ｉ、ｊ）と指すのを可能とする。ここで、１≦ｉ≦ｎおよび１≦ｊ≦ｖｍである。

３）全ての行インデックスｉについて、Ｓ（ｉ、ｊ）をＡｍのｊ番目の可能な値のバイナリインジケータ変数に設定する。それゆえに、特定の行ｉ∈｛１，２、．．．、ｎ｝について、属性値がｂ∈｛１，２、．．．、ｖｍ｝である場合、全てのｊ≠ｂについてＳ（ｉ、ｂ）＝１およびＳ（ｉ、ｊ）＝０である。

【００３３】

拡張された平文ＤＢ　Ｄａを適切に取得した後、ステップ１３０において、例えば、ＫＡ４０から取得した公開鍵（ｐｋ）を使用して、Ｄａが暗号化されてＤｅを取得する。実際には、この暗号化は、付加的な秘密の共有と意味的にセキュアな準同型暗号システムとの組み合わせを使用することによって適切に達成される。特定の環境、計算、記憶および／または他の適用可能な考慮事項に応じて、他の変形例も同様に適切であるおよび／または望ましいかもしれないが、以下は、Ｄｅを構築するために使用される準同型暗号システムの種類に応じて使用されることができる２つの適切な実施形態を記載する。

【００３４】

第１の変形例または実施形態において、暗号化手順は、単一の乗算の後に無制限数の加算が続く暗号化されたドメイン計算をサポートする例えばＢＧＮ（Ｂｏｎｅｈ、ＧｏｈおよびＮｉｓｓｉｍ）暗号システムなどの２－ＤＮＦ（論理和標準形）演算をサポートする暗号システムを使用する。ＢＧＮ暗号システムでは、照会された属性毎に単一の暗号化された列を追加して（例えば、上記取得したような）データベースＤａが暗号化される。ＢＧＮ変形例の使用は、特に、暗号化結果を復号するために利用可能な計算能力は比較的高いが、ＤＢサーバにおいて利用可能な記憶能力が比較的低い場合に適用可能である。

【００３５】

第２の変形例または実施形態において、暗号化手順は、暗号化されたドメインの加算のみをサポートする例えばＰａｉｌｌｉｅｒ暗号システムなどの相加的準同型暗号システムを使用する。２つの変形例は、暗号システムの能力とＤｅの記憶オーバーヘッドとのトレー

ドオフに対処する。Ｐａｉｌｌｉｅｒ暗号システムでは、（照会された属性、条件付き属性）のペア毎に単一の暗号化された列を追加して（例えば、上記取得したような）データベースＤａが暗号化される。Ｐａｉｌｌｉｅｒバリアントの使用は、特に、暗号化結果を復号するために利用可能な計算能力は比較的低いが、ＤＢサーバにおいて利用可能な記憶能力は比較的高くすることができる場合に適用可能である。

【００３６】

ここで、上述した第１の変形例を参照すると、暗号化プロセスまたはステップ１３０への入力は、（１）

【００３７】

【数２】

$$\bar{v} = \prod_{j=i}^{m} v_i$$

【００３８】

の追加の列を含む拡張平文データベースＤａ、および（２）２－ＤＮＦ準同型暗号システムの公開鍵ｐｋを含む。結果の出力は、暗号化データベースＤｅであり、

【００３９】

【数３】

$$\bar{\bar{v}}$$

【００４０】

の追加の列のエントリは、２－ＤＮＦ準同型暗号システムを使用して暗号化され、照会されるべき属性は、追加的にブラインドされる。

【００４１】

（以下のサブステップの適用を含む）暗号化プロトコルを説明するために、ここでは、条件付き属性、すなわちＷＨＥＲＥ句に続く属性と、クエリ属性、すなわちＳＥＬＥＣＴ文に続く属性とに別個に焦点をあてる。

１）上記のように、条件付き属性をＡｍとする。ここで、ｖｍ個の異なる値をとるｍ＝１，２、．．．、Ｍである。変数ｊ＝１，２、．．．、ｖｍを使用してこれらの値をインデキシングする。そして、各属性Ａｍについて、ｉ行目におけるＡｍのｊ番目の可能な値についてのバイナリインジケータ変数Ｓ（ｉ、ｊ）が構成される。ここで、ｉ∈｛１，２、．．．、ｎ｝である。Ｓ（ｉ、ｊ）は、意味的にセキュアな２－ＤＮＦ準同型暗号システムを使用して暗号化され、Ｅ（ｐｋ、Ｓ（ｉ、ｊ））を取得する。

２）照会された属性をＱｋ、ｋ＝１，２、．．．、Ｌとする。そして、行ｉ∈｛１，２、．．．、ｎ｝における属性Ｑｋの各平文値Ｑｋ（ｉ）の代わりに、データ所有者の公開鍵ｐｋの下でのクエリ属性値の２－ＤＮＦ準同型暗号化であるＥ（ｐｋ、Ｑｋ（ｉ））を記憶する。

【００４２】

ＤａにおけるＭ個の条件付き属性および照会済み属性のそれぞれの値に上記ステップを適用した結果、暗号化データベースＤｅが得られる。

【００４３】

ここで、上述した第２の変形例を参照すると、暗号化プロセスまたはステップ１３０への入力は、（１）

【００４４】

【数４】

$$\bar{v} = \prod_{j=i}^{m} v_i$$

【００４５】
の追加の列を含む拡張平文データベースＤａ、および（２）相加的準同型暗号システムの公開鍵ｐｋを含む。結果の出力は、暗号化データベースＤｅであり、
【００４６】
【数５】

$$\bar{\bar{v}}$$

【００４７】
の追加の列のエントリは、相加的準同型暗号システムを使用して暗号化され、照会されるべき属性は、追加的にブラインドされる。
【００４８】
（以下のサブステップの適用を含む）暗号化プロトコルを説明するために、ここでは再度、条件付き属性、すなわちＷＨＥＲＥ句に続く属性と、クエリ属性、すなわちＳＥＬＥＣＴ文に続く属性とに別個に焦点をあてる。
１）上記のように、条件付き属性をＡｍとする。ここで、ｖｍ個の異なる値をとるｍ＝１，２、．．．、Ｍである。上記のように、変数ｊ＝１，２、．．．、ｖｍを使用してこれらの値をインデキシングする。そして、各属性Ａｍについて、ｉ行目におけるＡｍのｊ番目の可能な値についてのバイナリインジケータ変数Ｓ（ｉ、ｊ）が構成される。ここで、ｉ∈｛１，２、．．．、ｎ｝である。Ｓ（ｉ、ｊ）は、意味的にセキュアな相加的準同型暗号システムを使用して暗号化され、Ｅ（ｐｋ、Ｓ（ｉ、ｊ））を取得する。
２）照会された属性をＱｋ、ｋ＝１，２、．．．、Ｌとする。そして、行ｉ∈｛１，２、．．．、ｎ｝における属性Ｑｋの値Ｑｋ（ｉ）について、間隔∈［－Ｒｋ、Ｒｋ］から無作為に均一に整数ｒｋ（ｉ）を選択する。Ｒｋは正の整数であり、λ（Ｒｋ）によって示されるＲｋにおけるビット数はセキュリティパラメータである。平文属性値Ｑｋ（ｉ）の代わりに、加法的にブラインドされた値
【００４９】
【数６】

$$\tilde{Q}_k(i) = Q_k(i) + r_k(i)$$

【００５０】
を記憶する。これらのブラインドされた値を含む列を新たな属性
【００５１】
【数７】

$$\tilde{Q}_k$$

【００５２】
とみなす。
３）照会された各属性Ｑｋ、および条件付き属性Ａｍのｖｍ個の可能な値のそれぞれにつ

いて、$R_{k,m}$によって示される追加の列が導入される。この新たな列のｉ行目のエントリは、ｊ＝１，２、．．．、v_mについて、$R_{k,j}(i)=E(p_k, r_k(i)...S(i,j))$によって与えられる。

【００５３】

D_aにおけるＭ個の条件付き属性およびＬ個の照会された属性のそれぞれの値に上記ステップを適用した結果、暗号化データベースD_eを得る。上記検討した２－ＤＮＦ準同型の場合とは異なり、上記サブステップ３の結果として追加された列の総数は、

【００５４】

【数８】

$$L = \sum_{m=1}^{M} v_m$$

【００５５】

であり、このアプローチでは記憶効率が低下させる。

【００５６】

いずれの場合にも、このようにして暗号化ＤＢ　D_eを取得した後、ステップ１４０において、暗号化ＤＢ　D_eは、送信され、ロードされ、および／またはＤＢＭＳ３０に送られる。実際には、ＤＢＭＳ３０において、全てのテーブル名は平文で利用可能であり、全ての列属性名は平文で利用可能であるが、テーブルにおけるエントリは、上記説明されたように暗号化される。すなわち、ＤＢサーバにおいて平文で利用可能な項目は、属性名のみである。例えば、リレーショナルＤＢにおいて、テーブルおよび列の名称は既知であるが、それらの列におけるデータエントリは暗号化される。

【００５７】

このようにして暗号化ＤＢ　D_eをＤＢＭＳ３０にロードすると、それは、例えばアナリストビュー１２を使用してフロントエンドウェブサービス／サーバ１０を介してユーザ／アナリストによって提出されたクエリに応答してＣＳ２０によってアクセスされることができる。適切には、クエリは、平文形式で提出されることができる。ＤＢＭＳ３０のＤＢサーバは、順次、ＣＳ２０にデータを提供し、ＣＳ２０によって実行される計算は、暗号化されたドメインにおいて適切に実行される。提出されたクエリに応答して、ＣＳ２０によって実行された計算の結果は、暗号化された形式で返される。例えば、ＫＡ４０から利用可能にされたまたは取得された対応する復号鍵または秘密鍵（暗号化を実行するために使用される公開鍵に対応する）を使用して、アナリストは、受信した暗号化結果を復号し、それらを平文でみることができる。実際には、そのような復号／秘密鍵は、情報またはデータ漏洩の脅威の可能性を制限するように、ＣＳ２０またはＤＢＭＳ３０から利用可能とされずおよび／または保持されない。

【００５８】

１つの適切な実施形態によれば、アナリストによって供給される平文クエリに応答して、例えば、ＣＳ２０および／またはＤＢサーバにおいて実行される暗号化ドメイン計算プロトコルがここで記載される。適切には、アナリストまたは他の同様のユーザは、アナリストビュー１２を介して平文クエリを提出することができる。適切には、ＤＢＭＳ３０のＤＢサーバにおいて、テーブル名が平文で利用可能であり、列属性名が平文で利用可能であるが、テーブルにおけるエントリは上記説明されたように暗号化されることに留意されたい。

【００５９】

最初に、本明細書では、例えば、ＷＨＥＲＥ句のない単一のクエリ属性に対する合計および平均クエリなどの単純な集計クエリが考えられる。実際には、これらのクエリを実行するために、ＣＳ２０は、データを暗号化するために使用される暗号システムの相加的準同型特性を単に利用するにすぎない。集計プロトコルを介して、より複雑な条件付きクエリ

において繰り返し使用される演算のいくつかがここで示されている。

【００６０】

例えば、アナリストは、入力として、以下の形式の平文クエリを提供することができる：

ＳＥＬＥＣＴ　ＳＵＭ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；

【００６１】

適切には、ＣＳ２０は、このクエリを解析し、クエリ＿属性に属する暗号化データにアクセスしなければならないことを直ちに認識する。以前の表記に続いて、この属性は、Ｑによって示され、ｉ行目の個々の値は、Ｑ（ｉ）、ｉ＝１，２、．．．、ｎとして示される。

【００６２】

上述したクエリに応答して、ＣＳ２０は、以下によって与えられる暗号化された総和をアナリストに適切に返す：

【００６３】

【数９】

$$E\left(\mathrm{pk}, \sum_{i=1}^{n} Q(i)\right)$$

【００６４】

ここで、ｐｋは、ＤＢ　Ｄｅを暗号化するために使用された公開鍵である。アナリストがデータ所有者および／またはＫＡ４０から適切な許可および／または鍵を受信したと仮定すると、彼はこの結果を復号することができる。

【００６５】

より具体的には、例えば、以下のプロトコルにしたがうことができる：

１）ＣＳ２０は、属性Ｑに対応するテーブルｔａｂｌｅ＿ｎａｍｅにおけるｎ個のエントリを検索する。

２）それは、以下のように所望の結果を得るために相加的準同型特性を使用する：

【００６６】

【数１０】

$$\prod_{i=1}^{n} E(\mathrm{pk}, Q(i)) = E\left(\mathrm{pk}, \sum_{i=1}^{n} Q(i)\right)$$

【００６７】

３）ＣＳ２０は、例えばウェブサービス／サーバ１０のアナリストビュー１２を介してアナリストに結果を返す。

【００６８】

暗号化ＤＢ　Ｄｅの構築に基づいて、ＤＢにおける行の総数は秘密ではないことに留意されたい。それゆえに、例えば、アナリストは、以下の形式のクエリを評価することもできる：

ＳＥＬＥＣＴ　ＡＶＧ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；

【００６９】

注目すべきことに、実際には、ＣＳ２０は、平文総和にアクセスすることができないため、平均化を実行しない。したがって、このＡＶＧクエリに応答して、まず、例えば上記示されたような単純なＳＵＭプロトコルを実行し、さらに、行数ｎをアナリストに返す。アナリストは、例えばアナリストビュー１２を介して、総和を復号してｎで除算し、彼のクエリに対する回答を得ることができる。

【００７０】
次に、本明細書は、プライバシー保護のＷＨＥＲＥクエリ、すなわち、ＷＨＥＲＥ句を利用するクエリを含むプロトコルを取り扱う。ここでは、平文ＤＢ　Ｄｐが Ｄａに拡張されたときに導入された追加の列が利用される。以下は、２つの例示的なプロトコル、２－ＤＮＦ準同型暗号化に基づく実装用のものと、Ｐａｉｌｌｉｅｒ準同型暗号化に基づく実装用のものとを示す。いずれの場合にも、クエリ属性列におけるブラインドされた値と、関連するインジケータ属性列における暗号化値との内積は、例えば、以下に説明されるようにセキュアに計算される。
【００７１】
例えば、アナリストは、入力として、以下の形式の平文クエリを提供することができる：
ＳＥＬＥＣＴ　ＳＵＭ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；
ＷＨＥＲＥ　条件＿属性＝“ｓｏｍｅ＿ｖａｌｕｅ”；
【００７２】
適切には、ＣＳ２０は、このクエリを解析し、クエリ＿属性に属するブラインドデータにアクセスしなければならないことを認識する。以前の表記に続いて、この属性は、Ｑによって示される。さらにまた、条件＿属性に対応する暗号化データにアクセスしなければならない。以前の表記に続いて、この属性は、Ａによって示される。ｓｏｍｅ＿ｖａｌｕｅは、属性Ａによってとり得るｖ個の可能な値のうちのｊ番目のものとする。以前の表記に続いて、インジケータ変数は、データベースを拡張するために使用された行列Ｓの列Ｓ（・、ｊ）において処理され、行列における各値は、意味的にセキュアな準同型暗号化方式を使用して暗号化される。
【００７３】
上述したクエリに応答して、ＣＳ２０は、以下によって与えられる暗号化された総和をアナリストに適切に返す：
【００７４】
【数１１】

$$E\left(pk, \sum_{i=1}^{n} Q(i)\mathbb{I}_{\{S(i,j)=1\}}\right)$$

【００７５】
ここで、ｐｋは、ＤＢ　Ｄｅを暗号化するために使用された公開鍵である。アナリストがデータ所有者および／またはＫＡ４０から適切な許可および／または鍵を受信したと仮定すると、彼はこの結果を復号することができる。上述した構成により、インジケータ関数は、条件＿属性が値ｓｏｍｅ＿ｖａｌｕｅを有する値のもののみをクエリが抽出することを示す。
【００７６】
より具体的には、例えば、２－ＤＮＦ準同型方式の場合、以下のプロトコルにしたがうことができる。この場合、暗号化関数Ｅ（ｐｋ、・）は、データ所有者の公開鍵の下での２－ＤＮＦ準同型暗号化である。適切には、暗号化ＤＢ　Ｄｅは、上述した手順にしたがって準備される。
１）ＣＳ２０は、属性Ｑに対応するテーブルｔａｂｌｅ＿ｎａｍｅにおけるｎ個の暗号化エントリを検索する。これらは、Ｅ（ｐｋ，Ｑ（ｉ）），ｉ＝１，２、．．．、ｎとして表される。
２）ＣＳ２０は、ｎ個の暗号化インジケータ変数Ｓ（・、ｊ）、すなわち行列Ｓのｊ列目を検索する。これらは、Ｅ（ｐｋ、Ｓ（ｉ、ｊ））、ｉ＝１、．．．、ｎとして表される。
３）ＣＳ２０は、２－ＤＮＦ準同形特性を使用して以下を計算する：
【００７７】

【数１２】

$$\prod_{i=1}^{n} E\left(pk, S(i,j)\right)E\left(pk, Q(i)\right)$$

$$= \prod_{i=1}^{n} E\left(pk, S(i,j)Q(i)\right)$$

$$= \prod_{i=1}^{n} E\left(pk, Q(i)\mathbb{I}_{\{S(i,j)=1\}}\right)$$

$$= E\left(pk, \sum_{i=1}^{n} Q(i)\mathbb{I}_{\{S(i,j)=1\}}\right)$$

【００７８】

４）ＣＳ２０は、例えばウェブサービス／サーバ１０のアナリストビュー１２を介してアナリストに結果を返す。

【００７９】

より具体的には、例えば、Ｐａｉｌｌｉｅｒ準同型方式の場合、以下のプロトコルにしたがうことができる。この場合、暗号化関数Ｅ（ｐｋ、・）は、データ所有者の公開鍵の下での相加的準同型暗号化である。適切には、暗号化ＤＢ　Ｄｅは、上述した手順にしたがって準備される。

１）ＣＳ２０は、属性Ｑに対応するテーブルｔａｂｌｅ＿ｎａｍｅにおけるｎ個のブラインドエントリを検索する。これらは、Ｑ（ｉ），ｉ＝１，２、．．．、ｎとして表される。

２）ＣＳ２０は、ｎ個の暗号化インジケータ変数Ｓ（・、ｊ）、すなわち行列Ｓのｊ列目を検索する。これらは、Ｅ（ｐｋ、Ｓ（ｉ、ｊ））、ｉ＝１、．．．、ｎとして表される。

３）各属性Ｑｋおよび条件付き属性Ａのｊ番目の値に対応して、暗号化データベースＤｅは、暗号化ブラインドエントリの列Ｒｋ，ｊを含むことに留意されたい。最初の接尾辞ｋを削除すると、属性

【００８０】

【数１３】

$$\tilde{Q}$$

【００８１】

に対応するブラインドエントリの列を属性Ｒｊによって示す。ＣＳ２０はまた、Ｒ｜（ｉ）＝Ｅ（ｐｋ、ｒ（ｉ）．．．Ｓ（ｉ、ｊ））、ｉ＝１、．．．、ｎとして表される列Ｒｊからｎ個の暗号化されたブラインド項を検索する。

４）ＣＳ２０は、相加的準同型特性を使用して以下を計算する：

【００８２】

【数１４】

$$\prod_{i=1}^{n} E\left(pk, S(i,j)\right)^{\tilde{Q}(i)} R_j(i)^{-1}$$

$$= \prod_{i=1}^{n} E\left(pk, S(i,j)\tilde{Q}(i)\mathbb{I}_{\{S(i,j)=1\}}\right) E\left(pk, -r(i)\mathbb{I}_{\{S(i,j)=1\}}\right)$$

$$= \prod_{i=1}^{n} E\left(pk, (\tilde{Q} - r(i))\mathbb{I}_{\{S(i,j)=1\}}\right)$$

$$= \prod_{i=1}^{n} E\left(pk, Q(i)\mathbb{I}_{\{S(i,j)=1\}}\right)$$

$$= E\left(pk, \sum_{i=1}^{n} Q(i)\mathbb{I}_{\{S(i,j)=1\}}\right)$$

【００８３】

５）ＣＳ２０は、例えばウェブサービス／サーバ１０のアナリストビュー１２を介してアナリストに結果を返す。

【００８４】

データベースプライバシーのために所望されるように、プロトコルは、計算を実行しているＣＳ２０に対してさえも、上記合計計算のためにどの行が選択されたかを明らかにしない。これらは、値Ｓ（ｉ、ｊ）＝１の行である。アナリストは、実行された拡張と共に暗号化ＤＢ　Ｄｅのスキーマを既に知っていることに留意されたい。そのため、条件＿属性がｊ番目の可能な値をとった行数をみつけたい場合、例えば、以下のようなクエリを送ることができる：

ＳＥＬＥＣＴ　ＳＵＭ（条件＿属性＿の＿ｊ番目の＿値）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；

【００８５】

Ｓ（ｉ、ｊ）は、インジケータ変数であることから、これは、本質的にカウントクエリである。このカウントクエリに応答して、ＣＳ２０は、以下を使用して（２－ＤＮＦおよび相加的準同型方式の双方についての）行数の暗号化を返す：

【００８６】

【数１５】

$$\text{行数} = \prod_{i=1}^{n} E(pk, S(i,j))$$

【００８７】
上記引数はまた、以下の形式の平均クエリを実行する方法も提供する：
ＳＥＬＥＣＴ　ＡＶＧ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ　ｎａｍｅ；
ＷＨＥＲＥ　条件＿属性＝“ｓｏｍｅ＿ｖａｌｕｅ”；
【００８８】
本質的に、ＣＳ２０が上記形式の平均クエリを解析する場合、内部的に、暗号化データベース上で以下の２つの関連するクエリを生成する：そのプロトコルが上述された条件付き総和クエリと、それに続く上述したカウントクエリ。それは、双方のクエリの結果をアナリストに返す。条件付き総和クエリの結果をカウントクエリの結果によって除算することにより、アナリストは、平均クエリの結果を取得する。
【００８９】
次に、本明細書は、プライバシー保護のＧＲＯＵＰＢＹクエリ、すなわち、ＧＲＯＵＰＢＹ句を利用するクエリを含むプロトコルを取り扱う。これらのプロトコルはまた、平文データベースＤｐがＤａに拡張されたときに導入された追加の列も利用する。基本的に、プロトコルは、上記プロトコルを使用して等価条件で複数のＷＨＥＲＥクエリを実行することによってＧＲＯＵＰＢＹ機能を実現する。
【００９０】
例えば、アナリストは、入力として、以下の形式の平文クエリを提供することができる：
ＳＥＬＥＣＴ　ＳＵＭ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；
ＧＲＯＵＰＢＹ　条件＿属性；
【００９１】
適切には、ＣＳ２０は、このクエリを解析し、クエリ＿属性に属するブラインドデータにアクセスしなければならないことを認識する。以前の表記に続いて、この属性は、Ｑによって示される。さらにまた、条件＿属性に対応する暗号化データにアクセスしなければならない。前述のように、条件付き属性は、Ａによって示され、ｖａｌｕｅ＿１、ｖａｌｕｅ＿２、．．．、ｖａｌｕｅ＿ｖによって示されるｖ個の可能な値をとることができると仮定する。暗号化データベースＤａのスキーマは、ＣＳ２０、ＤＢサーバおよびアナリストに知られているものとする。
【００９２】
適切には、ＣＳ２０は、アナリストに（例えばアナリストビュー１２を介して）暗号化されたｖ長ベクトルを返す。ここで、ベクトルのｊ番目の要素は、以下によって与えられる：
ｊ＝１，２、．．．、ｖについて、
【００９３】
【数１６】

$$E\left(pk, \sum_{i=1}^{n} Q(i)\mathbb{I}_{\{S(i,j)=1\}}\right) \qquad (1)$$

【００９４】
前述のように、ｐｋは、ＤＢ　Ｄｅを暗号化するために使用された公開鍵である。アナリストがデータ所有者および／またはＫＡ４０から適切な許可および／または鍵を受信したと仮定すると、彼は、この暗号化値のベクトルを復号することができる。使用される暗号化スキームに応じて、Ｅ（ｐｋ、・）は、それぞれ、相加的準同型または２－ＤＮＦ準同型暗号システムについての暗号文である。
【００９５】
より具体的には、各ｊ＝１，２、．．．、ｖについて、以下のステップが実行されること

ができる：
１）計算サーバは、以下によって与えられるＷＨＥＲＥクエリを作成する。
ＳＥＬＥＣＴ　ＳＵＭ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；
条件＿属性＝値ｊ；
２）上述したプロトコルを使用して、ＣＳ２０は、式（１）によって与えられる暗号化結果を取得する。
【００９６】
適切には、ＣＳ２０は、ＷＨＥＲＥプロトコルをｖ長ベクトルに繰り返し呼び出すことによって得られた暗号化結果を収集し、アナリストに送る。
【００９７】
明らかに、プライバシー保護のＧＲＯＵＰＢＹプロトコルの正確さは、ＷＨＥＲＥプロトコルのものにしたがう。前述のように集計クエリをカウントクエリと組み合わせることにより、ＣＳ２０はまた、以下の形式の平均クエリを実行することもできる：
ＳＥＬＥＣＴ　ＡＶＧ（クエリ＿属性）ＦＲＯＭ　ｔａｂｌｅ＿ｎａｍｅ；
ＧＲＯＵＰＢＹ　条件＿属性；
【００９８】
記載されたプライバシー保護クエリアーキテクチャの１つの適切な実施形態において、平文データベースは、１つ以上のテキストファイルの形式で読み取られる。データベース準備段階を容易にするために、データスキーマ、異なる属性の値の範囲および各属性のデータタイプをデータ所有者に表示するインターフェースが適切に提供される。例えば、インターフェースは、データ所有者に以下の能力を提供する：
１）カンマ区切り値（＊．ｃｓｖ）ファイルなどのテキストファイルからデータベースをインポートする。
２）データベースのスキーマを表示する。
３）各属性について、その属性を無視するか、平文で属性をインポートするか、または暗号化形式で属性をインポートするかを選択する。
４）条件付き属性について、暗号化されたドメインのＷＨＥＲＥまたはＧＲＯＵＰＢＹクエリをサポートするか否かを選択する。
５）各属性について、属性に関する統計分布に対するクエリをサポートするか否かを選択する。
６）Ｐａｉｌｌｉｅｒ準同型暗号化の鍵をロードするか、または新たな鍵ペアを生成する。
【００９９】
データ所有者のインターフェースから提供されたコマンドに応答して、暗号化サーバによって暗号化が平文データベースＤｐ上で実行され、暗号化データベースＤｅがＤＢＭＳ３０のＤＢサーバに記憶される。適切には、暗号化データベースは、ＭｙＳＱＬによってバックアップされる。データ所有者のインターフェースは、ブラウザにおけるＪａｖａ（登録商標）ｓｃｒｉｐｔで実装されることができ、ＲＥＳＴｆｕｌインターフェースを介して暗号化サーバおよびデータベースサーバと通信することができる。適切には、暗号化サーバによって実行される暗号化ルーチンは、Ｊａｖａ（登録商標）プログラミング言語で実装される。
【０１００】
データ所有者の場合と同様に、ブラウザにおけるＪａｖａ（登録商標）Ｓｃｒｉｐｔを使用して同様に、データアナリスト用のクエリインターフェースが提供される。これは、ＣＳ２０からのデータを送受信するように、コンピュータ、タブレットまたはスマートフォンからアクセスされることができるＲＥＳＴｆｕｌインターフェースとすることができる。照会を容易にするために、データアナリストには、以下の能力が提供されることができる：
１）ＭｙＳＱＬデータベースから暗号化データベースをインポートする。
２）ＣＳ２０によって返されたクエリの暗号文の結果を復号するための鍵をロードする。
３）データベーススキーマおよび属性がどのように暗号化されているかについての情報を

表示する。
４）サポートされている一連のクエリを表示する。
５）暗号化されたおよび／または平文の属性を含むＳＱＬ様クエリを入力する。
６）照会結果を数値的に（または、要求される場合にはグラフィカルに）表示する。
【０１０１】
データベースサーバは、ＭｙＳＱＬによってバックアップされることから、暗号化されていない属性を含むいかなるクエリも、従来の方法で処理されることができる。暗号化データの処理を含むクエリは、本明細書に記載されたプロトコルを適切に使用する。
【０１０２】
上記方法、システム、プラットフォーム、モジュール、プロセス、アルゴリズムおよび／または装置は、特定の実施形態に関して記載された。しかしながら、特定の修正および／または変更もまた想定されることを理解されたい。
【０１０３】
本明細書に提示される特定の例示的な実施形態に関連して、特定の構造的および／または機能的特徴は、定義された要素および／または構成要素に組み込まれるものとして記載されることが理解されたい。しかしながら、これらの特徴はまた、同一または同様の利益のために、必要に応じて同様に他の要素および／または構成要素に組み込まれてもよいことが想定される。例示的な実施形態の異なる態様は、所望の用途に適した他の代替実施形態を達成するために必要に応じて選択的に使用されてもよく、それにより、他の代替実施形態は、それに組み込まれる態様の各利点を実現することも理解されたい。
【０１０４】
本明細書において記載された特定のタスク、ステップ、プロセス、方法、機能、要素および／または構成要素のうちの任意の１つ以上は、ハードウェア、ソフトウェア、ファームウェアまたはそれらの組み合わせによって適切に実装されることができることも理解されたい。特に、様々なモジュール、構成要素および／または要素は、本明細書に記載されたタスク、ステップ、プロセス、方法および／または機能のうちの１つ以上を実行するように構成および／またはセットアップされたプロセッサ、電気回路、コンピュータおよび／または他の電子データ処理装置によって具現化されることができる。例えば、特定の要素を具現化するプロセッサ、コンピュータまたは他の電子データ処理装置は、コンピュータまたは他の電子データ処理装置によって実施および／または実行されたときに、本明細書に記載されたタスク、ステップ、プロセス、方法および／または機能のうちの１つ以上が完了または実行されるように、（例えば、ソースコード、解釈コード、オブジェクトコード、直接実行可能コードなどの）コードの適切なリストもしくは他の同様の命令またはソフトウェアもしくはファームウェアによって提供され、供給されおよび／またはプログラミングされることができる。適切には、コードもしくは他の同様の命令またはソフトウェアもしくはファームウェアのリストは、コンピュータまたは他の電子データ処理装置によって提供可能および／または実行可能であるように、持続性コンピュータおよび／または機械可読記憶媒体または媒体として実装され、および／またはその内部におよび／またはその上に記録、記憶、包含、または含有される。例えば、適切な記憶媒体および／または媒体は、限定されるものではないが、フロッピーディスク、フレキシブルディスク、ハードディスク、磁気テープ、もしくは任意の他の磁気記憶媒体または媒体、ＣＤ－ＲＯＭ、ＤＶＤ、光ディスク、もしくは任意の他の光媒体または媒体、ＲＡＭ、ＲＯＭ、ＰＲＯＭ、ＥＰＲＯＭ、フラッシュＥＰＲＯＭ、もしくは他のメモリもしくはチップもしくはカートリッジ、またはコンピュータもしくは機械もしくは電子データ処理装置が読み出して使用することができる任意の他の有形媒体または媒体を含むことができる。本質的に、本明細書において使用される場合、持続性コンピュータ可読媒体および／または機械可読媒体および／または媒体は、一時的な伝搬信号を除き、全てのコンピュータ可読媒体および／または機械可読媒体および／または媒体を含む。
【０１０５】
必要に応じて、本明細書において記載された特定のタスク、ステップ、プロセス、方法、

機能、要素および／または構成要素のうちの任意の１つ以上は、１つ以上の汎用コンピュータ、専用コンピュータ、プログラミングされたマイクロプロセッサもしくはマイクロコントローラおよび周辺集積回路要素、ＡＳＩＣもしくは他の集積回路、デジタル信号プロセッサ、ディスクリート素子回路などのハードワイヤード電子回路もしくは論理回路、ＰＬＤ、ＰＬＡ、ＦＰＧＡ、グラフィックカードＣＰＵ（ＧＰＵ）、もしくはＰＡＬなどのプログラマブル論理素子に実装および／または具現化されることができる。一般に、本明細書に記載された各タスク、ステップ、プロセス、方法および／または機能を順次実装することができる有限状態機械を実装することができる任意の装置が使用可能である。

【０１０６】

さらに、本明細書に記載されている特定の要素は、適切な状況下で、独立した要素であってもよくまたは別の方法で分割されてもよいことを理解されたい。同様に、１つの特定の要素によって実行されるように記載された複数の特定の機能は、個々の機能を実行するために独立して動作する複数の別個の要素によって実行されてもよく、または、特定の個々の機能は、協調して動作する複数の別個の要素によって実行されてもよい。あるいは、互いに区別されて本明細書に記載および／または示されたいくつかの要素または構成要素は、必要に応じて物理的にまたは機能的に組み合わされてもよい。

【０１０７】

要するに、本明細書は、好ましい実施形態を参照して記載されている。明らかに、本明細書を読んで理解すると、修正および変更が他人に生じるであろう。本発明の主題は、添付の特許請求の範囲またはその均等物の範囲内に含まれる限りにおいて、そのような全ての修正および変更を含むと解釈されることが意図される。

【図面】

【図１】

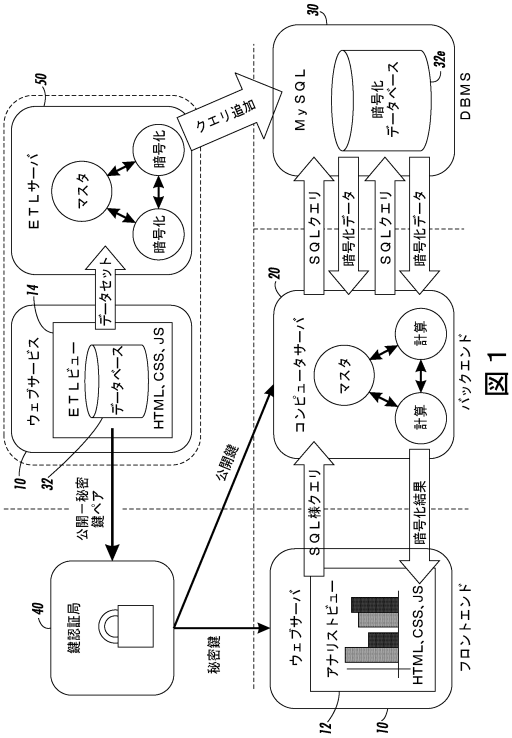
ETLサーバ
マスタ
暗号化
暗号化
クエリ追加
MySQL
暗号化データベース
DBMS
データセット
ウェブサービス
ETLビュー
データベース
HTML、CSS、JS
SQLクエリ
暗号化データ
SQLクエリ
暗号化データ
コンピュータサーバ
マスタ
計算
計算
バックエンド
公開鍵
公開－秘密鍵ペア
SQL様クエリ
暗号化結果
鍵認証局
秘密鍵
ウェブサーバ
アナリストビュー
HTML、CSS、JS
フロントエンド

図１

【図２】

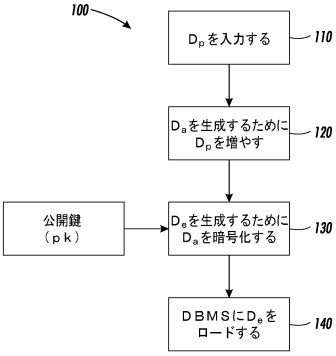
100
D_pを入力する
110
D_aを生成するためにD_pを増やす
120
公開鍵（pk）
D_aを生成するためにD_aを暗号化する
130
DBMSにD_aをロードする
140

図２

フロントページの続き

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　シャンタヌ・レイン
アメリカ合衆国 カリフォルニア州 ９４０２５ マウンテン・ビュー シャロン・パーク・ドライブ ６７５ アパートメント ２０１
(72)発明者　ヴィンセント・ビンシェドラー
アメリカ合衆国 イリノイ州 ６１８０１ アーバナ サウス・グッドウィン・アベニュー ３００ アパートメント ３１４
(72)発明者　アレハンドロ・イー・ブリトー
アメリカ合衆国 カリフォルニア州 ９４０４０ マウンテン・ビュー オルテガ・アベニュー １６３
(72)発明者　アーシン・ウズン
アメリカ合衆国 カリフォルニア州 ９５００８ キャンベル カプリ・ドライブ １１８６
(72)発明者　ヴァニシュリー・ラオ
アメリカ合衆国 カリフォルニア州 ９４０４３ マウンテン・ビュー サイプレス・ポイント・ドライブ ５０５ ユニット ２６６
審査官　吉田 歩

(56)参考文献　国際公開第２０１６／１２０９７５（WO，A１）
米国特許出願公開第２０１４／０２８１５１２（ＵＳ，Ａ１）

(58)調査した分野　(Int.Cl.，ＤＢ名)
Ｇ０９Ｃ　１／００
Ｇ０６Ｆ　１６／００