

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2024年11月7日(07.11.2024)



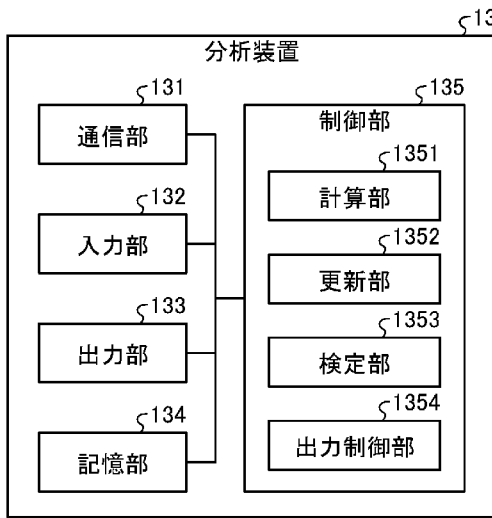
(10) 国際公開番号

WO 2024/228295 A1

- (51) 国際特許分類:
G09C 1/00 (2006.01)
- (21) 国際出願番号: PCT/JP2024/008140
- (22) 国際出願日: 2024年3月4日(04.03.2024)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2023-075875 2023年5月1日(01.05.2023) JP
- (71) 出願人: エヌ・ティ・ティ・コミュニケーションズ株式会社(NTT COMMUNICATIONS CORPORATION) [JP/JP]; 〒1008019 東京都千代田区大手町二丁目3番1号 Tokyo (JP).
- (72) 発明者: 田中 哲士 (TANAKA, Satoshi); 〒1008019 東京都千代田区大手町二丁目3番1号 エヌ・ティ・ティ・コミュニケーションズ株式会社内 Tokyo (JP). 櫻井 陽一 (SAKURAI, Yoichi); 〒1008019 東京都千代田区大手町二丁目3番1号 エヌ・ティ・ティ・コミュニケーションズ株式会社内 Tokyo (JP). 澤田 匡史 (SAWADA, Masashi); 〒1008019 東京都千代田区大手町二丁目3番1号 エヌ・ティ・ティ・コミュニケーションズ株式会社内 Tokyo (JP).
- (74) 代理人: 弁理士法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関3丁目8番1号 虎ノ門ダイビルイースト Tokyo (JP).

(54) Title: ANALYSIS DEVICE, ANALYSIS METHOD, AND ANALYSIS PROGRAM

(54) 発明の名称: 分析装置、分析方法及び分析プログラム



- 13 Analysis device
131 Communication unit
132 Input unit
133 Output unit
134 Storage unit
135 Control unit
1351 Calculation unit
1352 Updating unit
1353 Verification unit
1354 Output control unit

(57) Abstract: An analysis device according to an embodiment of the present invention includes a verification unit and an output control unit. The verification unit performs, by secret computation, verification of a Cox proportional hazard regression analysis, using at least one of a partial regression coefficient, a Hessian matrix, and a score vector obtained by the Cox proportional hazard regression analysis by secret computation. The output control unit outputs the results of the verification by the verification unit.

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

(57) 要約: 実施形態の分析装置は、検定部及び出力制御部を有する。検定部は、秘密計算によるC o x 比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、C o x 比例ハザード回帰分析の検定を秘密計算により行う。出力制御部は、検定部による検定の結果を出力する。

明 細 書

発明の名称：分析装置、分析方法及び分析プログラム

技術分野

[0001] 本発明は、分析装置、分析方法及び分析プログラムに関する。

背景技術

[0002] 従来、データを秘匿したまま統計的な演算を行い、演算の結果得られた統計量をユーザに提供する秘密計算システムが知られている。例えば、秘密計算システムは、重要な個人情報を取り扱う医療分野等におけるデータの分析に利用される場合がある。

[0003] また、暗号化された状態のデータに対し統計処理を行う秘密計算システムが知られている。例えば、暗号化された状態のデータを使って、ロジスティック回帰分析のパラメータを求める技術が知られている（例えば、特許文献2を参照）。

[0004] また、ロジスティック回帰分析に類似する手法であり、生存時間解析に用いられるCox比例ハザード回帰分析が知られている（例えば、特許文献3を参照）。

先行技術文献

特許文献

[0005] 特許文献1：国際公開第2019／124260号

特許文献2：特開2020－042128号公報

特許文献3：特開2008－299370号公報

非特許文献

[0006] 非特許文献1：日本電信電話株式会社，秘密計算のシステムとその原理，[online]，[令和5年1月6日検索]、インターネット<URL：https://www.rd.ntt/sil/project/sc/secure_computation.html>

発明の概要

発明が解決しようとする課題

[0007] しかしながら、従来の技術では、秘密計算によって行われたC o x 比例ハザード回帰分析の検定を行うことが難しい場合があるという問題がある。

課題を解決するための手段

[0008] 上述した課題を解決し、目的を達成するために、本発明の分析装置は、秘密計算によるC o x 比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、前記C o x 比例ハザード回帰分析の検定を秘密計算により行う検定部と、前記検定部による検定の結果を出力する出力制御部と、を有することを特徴とする。

発明の効果

[0009] 本発明によれば、秘密計算によって行われたC o x 比例ハザード回帰分析の検定を行うことができる。

図面の簡単な説明

[0010] [図1]図1は、実施形態に係る分析システムの構成例を示す図である。

[図2]図2は、実施形態に係る分析装置の構成例を示す図である。

[図3]図3は、学習データの例を示す図である。

[図4]図4は、記号の定義を示す図である。

[図5]図5は、W a l d 検定の流れを示すフローチャートである。

[図6]図6は、S c o r e 検定の流れを示すフローチャートである。

[図7]図7は、尤度比検定の流れを示すフローチャートである。

[図8]図8は、対数部分尤度を計算する処理の流れを示すフローチャートである。

[図9]図9は、分析プログラムを実行するコンピュータの一例を示す図である。

発明を実施するための形態

[0011] 以下に、本願に係る分析装置、分析方法及び分析プログラムの実施形態を図面に基づいて詳細に説明する。なお、本発明は、以下に説明する実施形態により限定されるものではない。

- [0012] まず、図 1 を用いて、分析システムの構成を説明する。分析システムは、秘密計算を利用してデータの分析を行うためのシステムである。
- [0013] 図 1 に示すように、分析システム 1 は、秘密計算システム 10 を含む。また、秘密計算システム 10 は、ネットワーク N を介して提供装置 20 及び提供装置 30 と接続される。例えば、ネットワーク N はインターネットである。また、秘密計算システム 10 は、端末装置 40 と接続される。
- [0014] 提供装置 20 及び提供装置 30 は、データ提供者側の装置である。提供装置 20 及び提供装置 30 は、秘密計算システム 10 にデータを提供（登録）する。
- [0015] 提供装置 20 及び提供装置 30 によって提供されるデータは、秘匿化されることが望ましい情報（例えば、個人の氏名、住所等の個人情報）を含む。例えば、提供装置 20 及び提供装置 30 は、医療機関で利用される診療データ又は健康診断データを提供する。ただし、提供装置 20 及び提供装置 30 によって提供されるデータは、医療機関で利用されるデータに限られない。
- [0016] 秘密計算システム 10 は、データ蓄積部 11 及びデータ処理部 12 を有する。データ蓄積部 11 は、秘密分散によりデータを蓄積する複数の蓄積装置（蓄積装置 111、蓄積装置 112、蓄積装置 113）を含む。また、データ処理部 12 は、秘密計算によりデータを処理する複数の計算装置（計算装置 121、計算装置 122、計算装置 123）を含む。なお、蓄積装置の数及び計算装置の数は、図 1 に示す例に限られない。
- [0017] 秘密計算システム 10 は、非特許文献 1（掲載 URL : https://www.rd.ntt/sil/project/sc/secure_computation.html）に記載された方法に従って、秘密分散及び秘密計算を実行することができる。
- [0018] まず、秘密計算システム 10 に提供されたデータは、複数のシェアに分割される（断片化）。そして、複数のシェアのそれぞれは、データ蓄積部 11 に含まれる複数の蓄積装置に分散して蓄積される。図 1 の例では、提供されたデータが 3 つのシェアに分割される。そして、蓄積装置 111、蓄積装置 112、蓄積装置 113 が、それぞれ 1 つずつシェアを蓄積する。

- [0019] データ処理部 1 2 は、データ蓄積部 1 1 に蓄積されたシェアに対し、秘密計算を実行する。データ処理部 1 2 は、複数の計算装置を使ったマルチパーティ計算により秘密計算を実行する。図 1 の例では、データ処理部 1 2 は、計算装置 1 2 1、計算装置 1 2 2、計算装置 1 2 3 により秘密計算を実行する。
- [0020] データ処理部 1 2 は、シェアを復元することなく各種の統計演算を行うことができる。例えば、データ処理部 1 2 は、ソート、結合等のテーブルの操作、レコード数の集計、総和、平均、最大値、最小値、標本分散等の統計量の計算、t 検定等の統計的検定を行うことができる。さらに、データ処理部 1 2 は、回帰分析及び主成分分析といった統計的分析を行うことができる。
- [0021] 分析装置 1 3 は、データ処理部 1 2 を利用してデータの分析を行う。分析装置 1 3 は、データ処理部 1 2 によって実行された秘密計算の結果に基づき、分析結果をデータ利用者側の端末装置 4 0 に提供する。利用者は、端末装置 4 0 を介してデータの分析結果を得ることができる。
- [0022] 例えば、秘密計算システム 1 0 には、個人ごとの属性及び身体に関するデータが提供される場合がある。属性及び身体に関するデータは秘匿化されることが望ましい個人情報である。属性及び身体に関するデータには、例えば年齢、性別、身長、体重等が含まれる。データ蓄積部 1 1 は、提供されたデータを断片化したシェアを各蓄積装置に格納する。
- [0023] なお、分割された個々のシェアは、単独では意味のないデータである。そのため、1 つのシェアから元のデータを復元することはできない。一方、複数のシェアを揃えることで元のデータを復元することが可能になる。
- [0024] データの利用者は、登録されたデータそのものを閲覧することはできないが、分析装置 1 3 及び端末装置 4 0 を介して、データの分析結果を閲覧することができる。例えば、データに個人の性別及び体重が含まれている場合、利用者は、各個人の性別及び体重を閲覧することはできないが、データの分析結果である「男性の平均体重」を閲覧することができる。
- [0025] 一例として、データ蓄積部 1 1 は、Shamir の閾値秘密分散法という手法を

使って秘密分散を行うことができる。このとき、データ蓄積部 11 は、元のデータを切片とする多項式を通る 3 つの座標をシェアとして各サーバに保管する。また、多項式の傾きはランダムに決定されるため、元のデータが同じであってもシェアが毎回同じであるとは限らない。なお、元のデータは、数値であってもよいし、数値に変換済みのデータであってもよい。

[0026] 秘密計算システム 10 は、複数のシェアから元のデータを復元することができる。多項式が 1 次式であれば、秘密計算システム 10 は、2 つの座標（シェアに相当）を結ぶ直線と軸との交点から切片（元のデータに相当）を求めることができる。一方で、1 つの座標からは直線が定まらないため、元のデータを復元することはできない。

[0027] また、前述の通り、データ処理部 12 は、シェアを復元することなく元のデータに対し秘密計算を実行することができる。例えば、座標で表されたシェア同士を加算した結果は、各シェアの元のデータ同士を加算した結果のシェアに相当する。

[0028] 分析装置 13 は、端末装置 40 からの要求に応じて、データ処理部 12 に秘密計算による処理を実行させる。なお、データ処理部 12 又は端末装置 40 が、分析装置 13 と同等の機能を実現してもよい。例えば、分析システム 1 は、分析装置 13 を有さない構成であってもよい。その場合、端末装置 40 がデータ処理部 12 と接続され、分析装置 13 と同等の処理を実行する。さらに、シェアに基づく統計演算は、データ処理部 12 ではなく端末装置 40 によって実行されてもよい。

[0029] 第 1 の実施形態では、分析装置 13 が秘密計算により Cox 比例ハザード回帰分析、及び Cox 比例ハザード回帰分析の検定を行う場合の例を説明する。

[0030] 図 2 を用いて、分析装置 13 の構成を説明する。図 2 は、実施形態に係る分析装置の構成例を示す図である。

[0031] 分析装置 13 の各部について説明する。図 2 に示すように、分析装置 13 は、通信部 131、入力部 132、出力部 133、記憶部 134 及び制御部

135を有する。

[0032] 通信部131は、他の装置の間でデータの通信を行う。例えば、通信部131はNIC (Network Interface Card) である。通信部131は他の装置との間でデータの送受信を行うことができる。

[0033] 入力部132は、データの入力を受け付けるためのインタフェースである。入力部132は、例えばマウス及びキーボード等の入力装置と接続される。

[0034] 出力部133は、データを出力するためのインタフェースである。出力部133は、例えばディスプレイ及びスピーカ等の出力装置と接続される。

[0035] 記憶部134は、HDD (Hard Disk Drive)、SSD (Solid State Drive)、光ディスク等の記憶装置である。なお、記憶部134は、RAM (Random Access Memory)、フラッシュメモリ、NVRAM (Non Volatile Static Random Access Memory) 等のデータを書き換え可能な半導体メモリであってもよい。記憶部134は、分析装置13で実行されるOS (Operating System) 及び各種プログラムを記憶する。

[0036] 制御部135は、分析装置13全体を制御する。制御部135は、例えば、CPU (Central Processing Unit)、MPU (Micro Processing Unit)、GPU (Graphics Processing Unit) 等の電子回路や、ASIC (Application Specific Integrated Circuit)、FPGA (Field Programmable Gate Array) 等の集積回路である。また、制御部135は、各種の処理手順を規定したプログラムや制御データを格納するための内部メモリを有し、内部メモリを用いて各処理を実行する。

[0037] 制御部135は、各種のプログラムが動作することにより各種の処理部として機能する。例えば、制御部135は、計算部1351、更新部1352、検定部1353及び出力制御部1354を有する。

[0038] 計算部1351は、秘密計算によりCox比例ハザード回帰の計算を行う。計算部1351は、Cox比例ハザード回帰モデルに説明変数を入力し、目的変数を出力させる。

[0039] 図4に、以降の説明で使用する記号の定義を示す。図4は、記号の定義を示す図である。

[0040] 図4に示すように、 N はレコード数である。 k は属性（説明変数）の数である。データセット X は、 $N \times k$ 行列の各レコードの属性情報を含む。例えば、 i 番目のレコードは $(x_1, \dots, x_k)$ である。

[0041] 結果は、あるレコードの生存又は死亡を示す情報である。 t は観測時間である。 y は0（生存）又は1（死亡）の2値である。

[0042] 偏回帰係数及びスコアベクトルは列ベクトルで表される。スコアベクトルは、対数部分尤度の1次導関数である。また、ヘッセ行列は、対数部分尤度の2次導関数である。

[0043] また、偏回帰係数、スコアベクトル、ヘッセ行列、対数部分尤度には、下付きの0が付される。対数部分尤度の計算方法については後述する。

[0044] 更新部1352は、計算部1351によって計算された目的変数が、正解の値に近づくように、秘密計算によりCox比例ハザード回帰モデルのパラメータを更新する。

[0045] 計算部1351と更新部1352の処理が1回又は複数回行われることにより、Cox比例ハザード回帰モデルの学習が行われる。

[0046] 図3は、学習データの例を示す図である。図3の「年齢」列、「性別」列、「摂取カロリー」列、「体重減少」列の値は、Cox比例ハザード回帰モデルの説明変数である。また、「生存時間」列の値に基づく予測生存率が、Cox比例ハザード回帰モデルの目的変数であり、ハザード関数 $h(t | x)$ に対応する。

[0047] 計算部1351は、「年齢」列、「性別」列、「摂取カロリー」列、「体重減少」列の値をCox比例ハザード回帰モデルに入力し、出力を得る。更新部1352は、計算部1351によって得られた出力が、「生存時間」列の値に基づく予測生存率に近づくように、Cox比例ハザード回帰モデルのパラメータを更新する。

[0048] 検定部1353は、Cox比例ハザード回帰モデルの検定を行う。秘密計

算では、学習データそのものを外部から取得することは困難である。そのため、検定部 1353 は、取得可能な限られた情報を基に検定のための統計量を計算する。

[0049] ここでは、検定部 1353 は、秘密計算による Cox 比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、Cox 比例ハザード回帰分析の検定を秘密計算により行う。検定部 1353 は、偏回帰係数、ヘッセ行列、及びスコアベクトルといった Cox 比例ハザード回帰分析に関する情報を、秘密計算システム 10 から取得する。

[0050] 検定部 1353 が、Wald 検定、Score 検定及び尤度比検定を行う場合の処理を説明する。

[0051] (Wald 検定)

図 5 は、Wald 検定の流れを示すフローチャートである。図 5 に示すように、検定部 1353 は、秘密計算システム 10 から、学習済みの Cox 比例ハザード回帰モデルの偏回帰係数とヘッセ行列を取得する（ステップ S101）。

[0052] 次に、検定部 1353 は、(1) 式のように Wald 検定の統計量を計算する（ステップ S102）。そして、検定部 1353 は、検定を実行し、出力制御部 1354 を介して検定結果を出力する（ステップ S103）。ただし、初期偏回帰係数を 0 とした場合、検定部 1353 は、(2) 式により統計量を計算する。

[0053] [数1]

$$(\vec{\beta} - \vec{\beta}_0)^T H (\vec{\beta} - \vec{\beta}_0) \quad \dots (1)$$

[0054] [数2]

$$\begin{aligned} & \vec{\beta}^T H \vec{\beta} \\ & (\vec{\beta}_0 = \vec{0} \text{ のとき}) \end{aligned} \quad \dots (2)$$

[0055] このように、検定部 1353 は、偏回帰係数及びヘッセ行列を基に、Wald

l d 検定の統計量を計算し、統計量を基に χ^2 乗検定を行う。

[0056] (Score 検定)

図6は、Score 検定の流れを示すフローチャートである。図6に示すように、検定部1353は、秘密計算システム10から、学習済みのCox比例ハザード回帰モデルの学習開始時のヘッセ行列（図4の初期ヘッセ行列）とスコアベクトル（図4の初期スコアベクトル）を取得する（ステップS201）。

[0057] 次に、検定部1353は、(3)式のように、Score 検定の統計量を計算する（ステップS202）。そして、検定部1353は、検定を実行し、出力制御部1354を介して検定結果を出力する（ステップS203）。

[0058] [数3]

$$\overrightarrow{u_0}^t H_0 \overrightarrow{u_0} \quad \dots (3)$$

[0059] このように、検定部1353は、Cox比例ハザード回帰分析の回帰モデルの学習開始時におけるヘッセ行列及びスコアベクトルを基に、Score 検定の統計量を計算し、統計量を基に χ^2 乗検定を行う。

[0060] (尤度比検定)

図7は、尤度比検定の流れを示すフローチャートである。検定部1353は、秘密計算システム10から、学習済みのCox比例ハザード回帰モデルの偏回帰係数とヘッセ行列を取得する。

[0061] また、図3の学習データの各レコードは各患者に対応し、Cox比例ハザード回帰モデルのハザード関数は、患者のある時間での生存確率を表すものとする。学習データの生存時間から、ある時間において患者が死亡しているか否かが導かれる。

[0062] 学習データのレコードのうち、指定された時間において死亡している患者に対応するレコードを死亡レコードと呼ぶ。また、学習データのレコードのうち、指定された時間において死亡していない患者に対応するレコードを残存レコードと呼ぶ。

[0063] 図7に示すように、まず、検定部1353は、学習開始時の偏回帰係数が

ら初期の対数部分尤度を計算する（ステップS 3 0 1）。また、検定部 1 3 5 3 は、学習済みの偏回帰係数から対数部分尤度を計算する（ステップS 3 0 2）。

[0064] 検定部 1 3 5 3 は、（４）式により対数部分尤度を計算する。また、検定部 1 3 5 3 は、（４）式の β を β_0 に置き換えた式により、初期の対数部分尤度を計算する。対数部分尤度の計算の詳細については後述する。

[0065] [数4]

$$\ell = \sum_{i=1}^N \left\{ y_i \left(\vec{x}_i \cdot \vec{\beta} - \log \left(\sum_{j=i}^N \exp(\vec{x}_i \cdot \vec{\beta}) \right) \right) \right\} \quad \dots (4)$$

[0066] 続いて、検定部 1 3 5 3 は、初期の対数部分尤度と対数部分尤度から、（５）式のように尤度比検定の統計量を計算する（ステップS 3 0 3）。そして、検定部 1 3 5 3 は、 χ^2 乗検定を行い、検定結果を出力する（ステップS 3 0 4）。

[0067] [数5]

$$2(\ell - \ell_0) \quad \dots (5)$$

[0068] 図 8 は、対数部分尤度を計算する処理の流れを示すフローチャートである。図 8 に示すように、まず、検定部 1 3 5 3 は、秘密計算ソートを用いて、データセットのレコードを時刻に対して昇順に並べる（ステップS 3 1 1）。

[0069] 次に、検定部 1 3 5 3 は、偏回帰係数とレコードの積和値を秘密計算積和により計算する（ステップS 3 1 2）。また、検定部 1 3 5 3 は、積和値の指数値である死亡リスク（（４）式の右辺の $\exp$ ）を秘密計算写像で計算する（ステップS 3 1 3）。

[0070] ここで、検定部 1 3 5 3 は、 $i = 1, \dots, N$ に対して、秘密計算総和を用いて、レコード番号が $i \leq j$ となる j の死亡リスクの総和（（４）式の右辺の右側の Σ ）を計算する（ステップS 3 1 4）。

[0071] 続いて、検定部 1 3 5 3 は、死亡リスクの総和の対数値（（４）式の右辺

の $\log$) を秘密計算写像で計算する (ステップ S 3 1 5)。

[0072] そして、検定部 1 3 5 3 は、積和値と対数値の差を秘密計算の減算 ((4) 式の右辺の減算) で計算する (ステップ S 3 1 6)。さらに、検定部 1 3 5 3 は、秘密計算の積和を用いて、生存又は死亡の情報 ((4) 式の y_i) と減算の結果から対数部分尤度を計算する (ステップ S 3 1 7)。

[0073] このように、検定部 1 3 5 3 は、検定部 1 3 5 3 は、C o x 比例ハザード回帰分析の回帰モデルの学習に用いられた死亡レコードと偏回帰係数の積和を、第 1 の死亡リスクとして秘密計算により計算し、回帰モデルの学習に用いられた残存レコードと偏回帰係数の積和を、第 2 の死亡リスクとして秘密計算により計算し、第 1 の死亡リスクと第 2 の死亡リスクを基に対数尤度を計算し、対数尤度を基に検定を行う。

[0074] [実施形態の効果]

これまで説明してきたように、分析装置 1 3 は、検定部 1 3 5 3 及び出力制御部 1 3 5 4 を有する。検定部 1 3 5 3 は、秘密計算による C o x 比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、C o x 比例ハザード回帰分析の検定を秘密計算により行う。出力制御部 1 3 5 4 は、検定部 1 3 5 3 による検定の結果を出力する。

[0075] これにより、分析装置 1 3 は、秘密計算によって行われた C o x 比例ハザード回帰分析の検定を行うことができる。

[0076] 例えば、検定部 1 3 5 3 は、偏回帰係数及びヘッセ行列を基に、W a l d 検定の統計量を計算し、統計量を基に検定を行う。

[0077] 例えば、検定部 1 3 5 3 は、C o x 比例ハザード回帰分析の回帰モデルの学習開始時におけるヘッセ行列及びスコアベクトルを基に、S c o r e 検定の統計量を計算し、統計量を基に検定を行う。

[0078] 例えば、検定部 1 3 5 3 は、C o x 比例ハザード回帰分析の回帰モデルの学習に用いられた死亡レコードと偏回帰係数の積和を、第 1 の死亡リスクとして秘密計算により計算し、回帰モデルの学習に用いられた残存レコードと

偏回帰係数の積和を、第2の死亡リスクとして秘密計算により計算し、第1の死亡リスクと第2の死亡リスクを基に対数尤度を計算し、対数尤度を基に検定を行う。

[0079] このように、分析装置13は、Wald検定、Score検定及び尤度比検定を行うことができる。

[0080] [システム構成等]

また、図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示のように構成されていることを要しない。すなわち、各装置の分散及び統合の具体的形態は図示のものに限られず、その全部又は一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的又は物理的に分散又は統合して構成することができる。さらに、各装置にて行われる各処理機能は、その全部又は任意の一部が、CPU (Central Processing Unit) 及び当該CPUにて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。なお、プログラムは、CPUだけでなく、GPU等の他のプロセッサによって実行されてもよい。

[0081] また、本実施形態において説明した各処理のうち、自動的に行われるものとして説明した処理の全部又は一部を手動的に行うこともでき、あるいは、手動的に行われるものとして説明した処理の全部又は一部を公知の方法で自動的に行うこともできる。この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。

[0082] [プログラム]

一実施形態として、分析装置13は、パッケージソフトウェアやオンラインソフトウェアとして上記の分析処理を実行する分析プログラムを所望のコンピュータにインストールさせることによって実装できる。例えば、上記の分析プログラムを情報処理装置に実行させることにより、情報処理装置を分析装置13として機能させることができる。ここで言う情報処理装置には、

デスクトップ型又はノート型のパーソナルコンピュータが含まれる。また、その他にも、情報処理装置にはスマートフォン、携帯電話機やPHS (Personal Handyphone System) 等の移動体通信端末、さらには、PDA (Personal Digital Assistant) 等のスレート端末等がその範疇に含まれる。

[0083] また、分析装置13は、ユーザが使用する端末装置をクライアントとし、当該クライアントに上記の分析処理に関するサービスを提供する分析サーバ装置として実装することもできる。例えば、分析サーバ装置は、秘密計算によるCox比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルを入力とし、Cox比例ハザード回帰分析の検定結果を出力とする分析サービスを提供するサーバ装置として実装される。

[0084] 図9は、分析プログラムを実行するコンピュータの一例を示す図である。コンピュータ1000は、例えば、メモリ1010、CPU1020を有する。また、コンピュータ1000は、ハードディスクドライバインタフェース1030、ディスクドライバインタフェース1040、シリアルポートインタフェース1050、ビデオアダプタ1060、ネットワークインタフェース1070を有する。これらの各部は、バス1080によって接続される。

[0085] メモリ1010は、ROM (Read Only Memory) 1011及びRAM (Random Access Memory) 1012を含む。ROM1011は、例えば、BIOS (Basic Input Output System) 等のブートプログラムを記憶する。ハードディスクドライバインタフェース1030は、ハードディスクドライブ1090に接続される。ディスクドライバインタフェース1040は、ディスクドライブ1100に接続される。例えば磁気ディスクや光ディスク等の着脱可能な記憶媒体が、ディスクドライブ1100に挿入される。シリアルポートインタフェース1050は、例えばマウス1110、キーボード1120に接続される。ビデオアダプタ1060は、例えばディスプレイ1130に接続される。

[0086] ハードディスクドライブ1090は、例えば、OS1091、アプリケー

ションプログラム１０９２、プログラムモジュール１０９３、プログラムデータ１０９４を記憶する。すなわち、分析装置１３の各処理を規定するプログラムは、コンピュータにより実行可能なコードが記述されたプログラムモジュール１０９３として実装される。プログラムモジュール１０９３は、例えばハードディスクドライブ１０９０に記憶される。例えば、分析装置１３における機能構成と同様の処理を実行するためのプログラムモジュール１０９３が、ハードディスクドライブ１０９０に記憶される。なお、ハードディスクドライブ１０９０は、ＳＳＤ（Solid State Drive）により代替されてもよい。

[0087] また、上述した実施形態の処理で用いられる設定データは、プログラムデータ１０９４として、例えばメモリ１０１０やハードディスクドライブ１０９０に記憶される。そして、ＣＰＵ１０２０は、メモリ１０１０やハードディスクドライブ１０９０に記憶されたプログラムモジュール１０９３やプログラムデータ１０９４を必要に応じてＲＡＭ１０１２に読み出して、上述した実施形態の処理を実行する。

[0088] なお、プログラムモジュール１０９３やプログラムデータ１０９４は、ハードディスクドライブ１０９０に記憶される場合に限らず、例えば着脱可能な記憶媒体に記憶され、ディスクドライブ１１００等を介してＣＰＵ１０２０によって読み出されてもよい。あるいは、プログラムモジュール１０９３及びプログラムデータ１０９４は、ネットワーク（ＬＡＮ（Local Area Network）、ＷＡＮ（Wide Area Network）等）を介して接続された他のコンピュータに記憶されてもよい。そして、プログラムモジュール１０９３及びプログラムデータ１０９４は、他のコンピュータから、ネットワークインタフェース１０７０を介してＣＰＵ１０２０によって読み出されてもよい。

符号の説明

- [0089] １ 分析システム
 １０ 秘密計算システム
 １１ データ蓄積部

- 1 2 データ処理部
- 1 3 分析装置
 - 1 3 1 通信部
 - 1 3 2 入力部
 - 1 3 3 出力部
 - 1 3 4 記憶部
 - 1 3 5 制御部
 - 1 3 5 1 計算部
 - 1 3 5 2 更新部
 - 1 3 5 3 検定部
 - 1 3 5 4 出力制御部

請求の範囲

- [請求項1] 秘密計算によるC o x比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、前記C o x比例ハザード回帰分析の検定を秘密計算により行う検定部と、
- 前記検定部による検定の結果を出力する出力制御部と、
- を有することを特徴とする分析装置。
- [請求項2] 前記検定部は、前記偏回帰係数及び前記ヘッセ行列を基に、W a l d検定の統計量を計算し、前記統計量を基に検定を行うことを特徴とする請求項1に記載の分析装置。
- [請求項3] 前記検定部は、前記C o x比例ハザード回帰分析の回帰モデルの学習開始時における前記ヘッセ行列及び前記スコアベクトルを基に、S c o r e検定の統計量を計算し、前記統計量を基に検定を行うことを特徴とする請求項1に記載の分析装置。
- [請求項4] 前記検定部は、前記C o x比例ハザード回帰分析の回帰モデルの学習に用いられた死亡レコードと前記偏回帰係数の積和を、第1の死亡リスクとして秘密計算により計算し、前記回帰モデルの学習に用いられた残存レコードと前記偏回帰係数の積和を、第2の死亡リスクとして秘密計算により計算し、前記第1の死亡リスクと前記第2の死亡リスクを基に対数尤度を計算し、前記対数尤度を基に検定を行うことを特徴とする請求項1に記載の分析装置。
- [請求項5] 分析装置によって実行される分析方法であって、
- 秘密計算によるC o x比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、前記C o x比例ハザード回帰分析の検定を秘密計算により行う検定工程と、
- 前記検定工程による検定の結果を出力する出力制御工程と、
- を含むことを特徴とする分析方法。

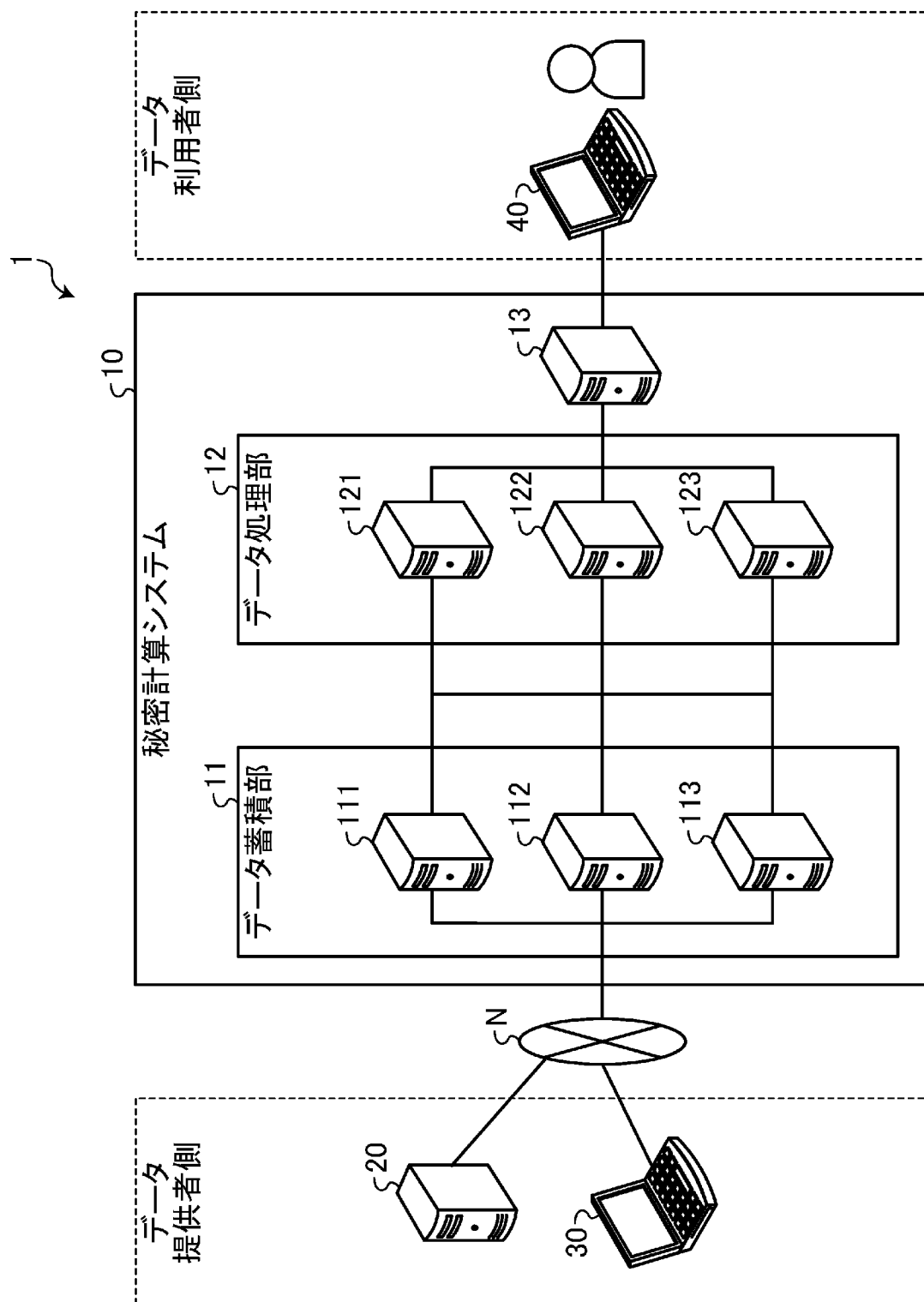
[請求項6] 秘密計算によるC o x比例ハザード回帰分析によって得られた偏回帰係数、ヘッセ行列、及びスコアベクトルの少なくともいずれかを用いて、前記C o x比例ハザード回帰分析の検定を秘密計算により行う検定ステップと、

 前記検定ステップによる検定の結果を出力する出力制御ステップと、

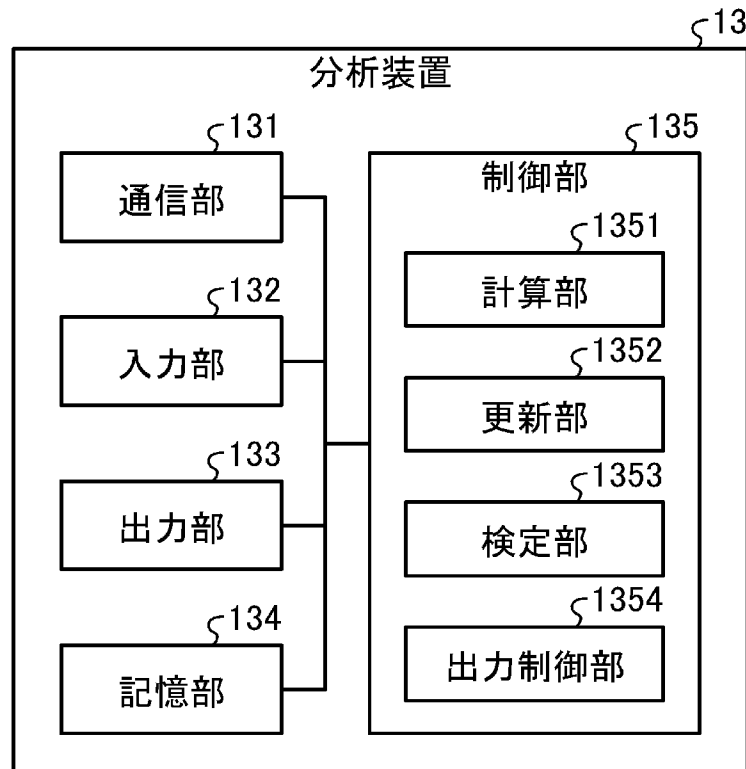
 、

 をコンピュータに実行させることを特徴とする分析プログラム。

[図1]



[図2]



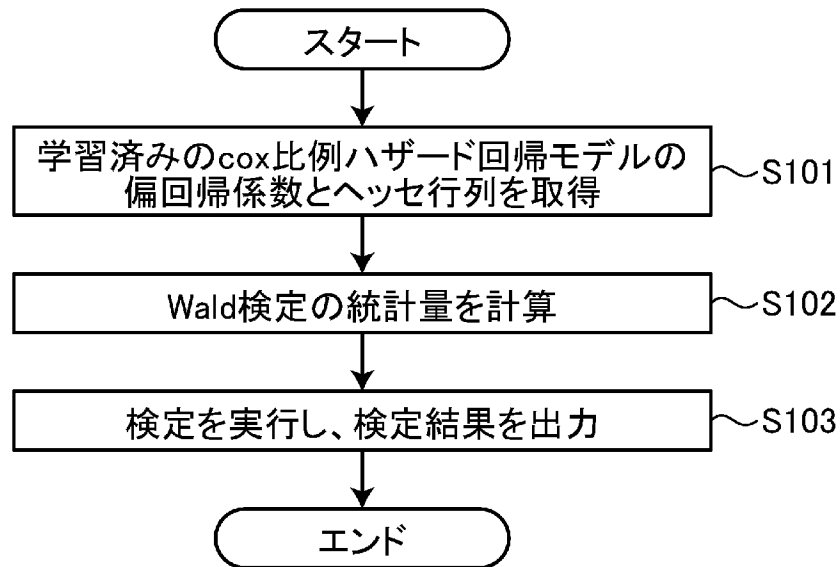
[図3]

No	年齢	性別	摂取カロリー	体重減少	生存時間
1	56	2	3200	55	250
2	32	1	100	-15	400
3	77	2	1450	-10	750
...	...	...	...	...	...

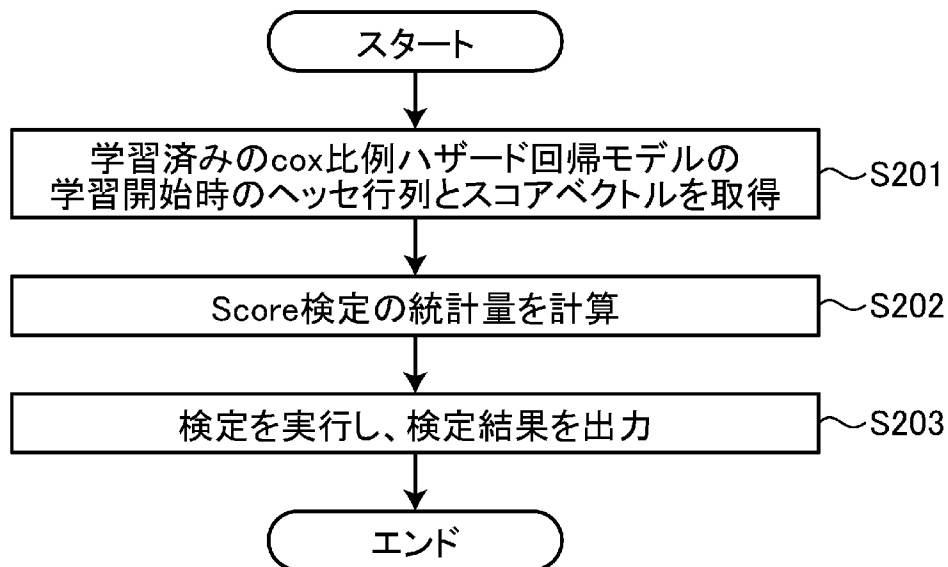
[図4]

• レコード数	: N
• 属性数	: k
• データセット	: $X = (x_{i,j})_{i=1,\dots,N}^{j=1,\dots,k}$
• 結果	: $(\vec{t} \ \vec{y}) = ((t_1 \ y_1) \ \dots \ (t_N \ y_N))^T$
• 偏帰帰係数	: $\vec{\beta} = (\beta_1 \ \dots \ \beta_k)^T$
• スコアベクトル	: $\vec{u} = (u_1 \ \dots \ u_k)^T$
• ヘッセ行列	: $H = (h_{i,j})_{i=1,\dots,k}^{j=1,\dots,k}$
• 初期偏帰帰係数	: $\vec{\beta}_0$
• 初期スコアベクトル	: $\vec{u}_0$
• 初期ヘッセ行列	: H_0
• 対数部分尤度	: ℓ
• 初期対数部分尤度	: ℓ_0

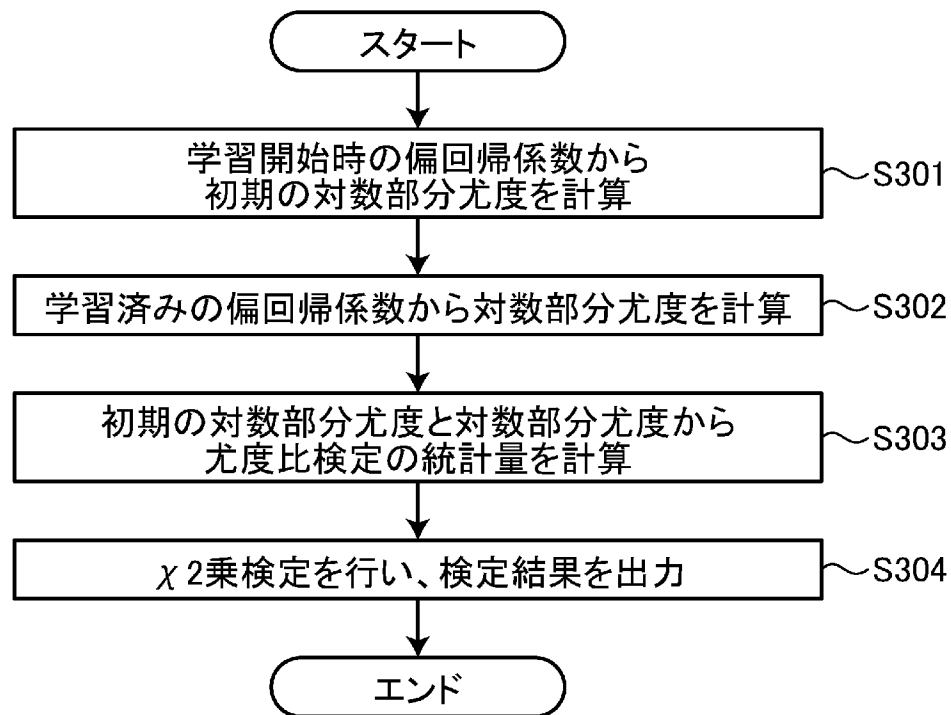
[図5]



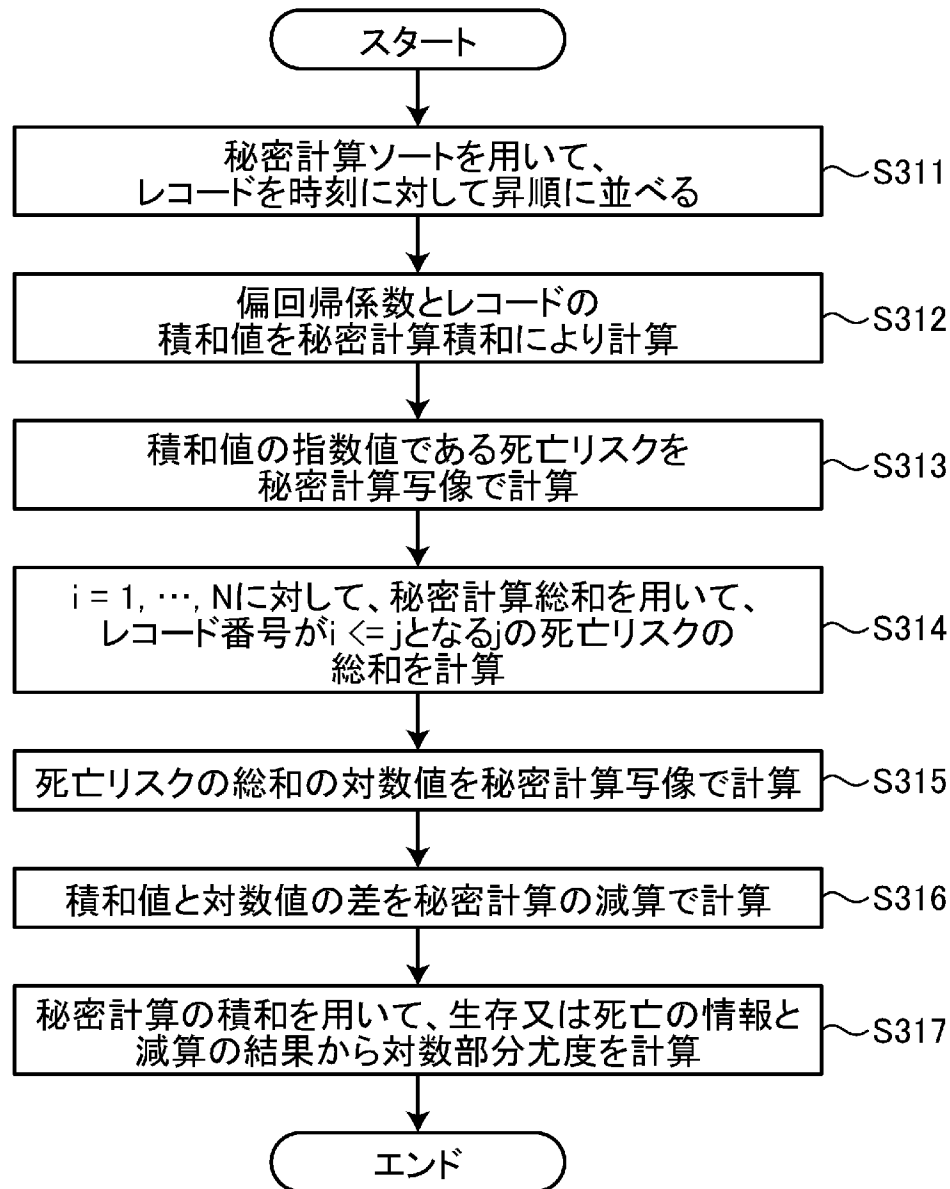
[図6]



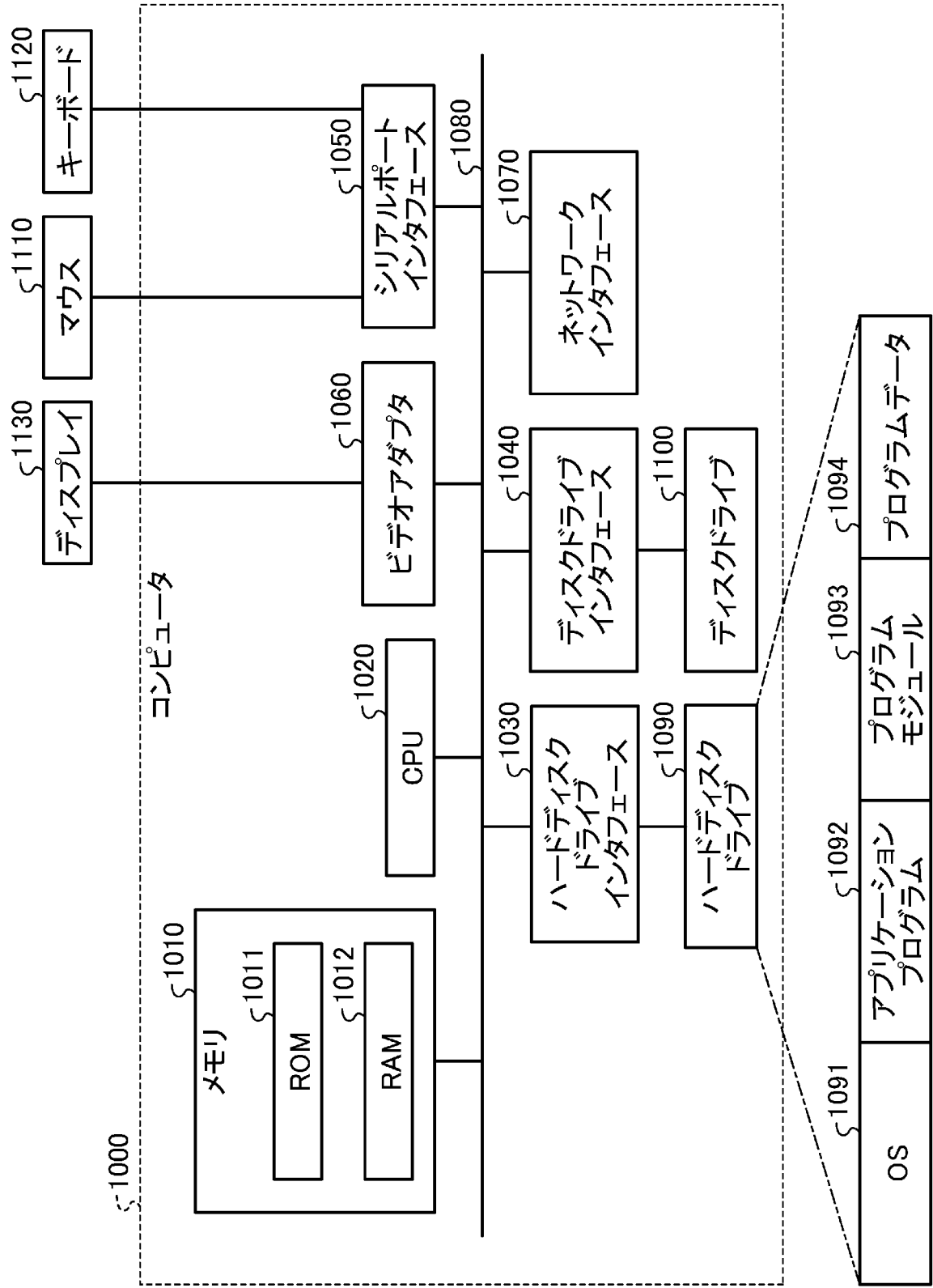
[図7]



[図8]



[図9]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2024/008140

A. CLASSIFICATION OF SUBJECT MATTER**G09C 1/00**(2006.01)i

FI: G09C1/00 650Z

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)

G09C1/00

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2024

Registered utility model specifications of Japan 1996-2024

Published registered utility model applications of Japan 1994-2024

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2022/079904 A1 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 21 April 2022 (2022-04-21)	1-6
A	CN 111506922 A (ALIPAY (HANGZHOU) TECHNOLOGY CO., LTD.) 07 August 2020 (2020-08-07)	1-6
A	川崎 将平 他, プライバシーを保護した尤度比検定, 2014年度 人工知能学会全国大会 (第28回) 論文集, May 2014, 2B5-OS-15b-3, pp. 1-4, (KAWASAKI, Shohei et al. Privacy-preserving likelihood-ratio test.), non-official translation (Proceedings of 28th Annual Conference of the Japanese Society for Artificial Intelligence (2014))	1-6

☐ Further documents are listed in the continuation of Box C.☒ 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

“D” document cited by the applicant in the international application

“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

30 April 2024

Date of mailing of the international search report

14 May 2024

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2024/008140

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2022/079904	A1	21 April 2022	US	2023/0367846	A1	
				EP	4231272	A1	
				CN	116324935	A	
CN	111506922	A	07 August 2020	(Family: none)			

A. 発明の属する分野の分類（国際特許分類（IPC））

G09C 1/00(2006.01)i

FI: G09C1/00 650Z

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

G09C1/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報

1922 - 1996年

日本国公開実用新案公報

1971 - 2024年

日本国実用新案登録公報

1996 - 2024年

日本国登録実用新案公報

1994 - 2024年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2022/079904 A1（日本電信電話株式会社）21.04.2022（2022 - 04 - 21）	1-6
A	CN 111506922 A (ALIPAY (HANGZHOU) TECHNOLOGY CO., LTD.) 07.08.2020（2020 - 08 - 07）	1-6
A	川崎 将平 他, プライバシを保護した尤度比検定, 2014年度 人工知能学会全国大会（第28回）論文集, 2014.05, 2B5-0S-15b-3, pp. 1-4	1-6

☐ C欄の続きにも文献が列挙されている。

☒ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

“A” 特に関連のある文献ではなく、一般的技术水準を示すもの

“D” 国際出願で出願人が先行技術文献として記載した文献

“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

“O” 口頭による開示、使用、展示等に言及する文献

“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献

“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの

“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

“&” 同一パテントファミリー文献

国際調査を完了した日

30.04.2024

国際調査報告の発送日

14.05.2024

名称及びあて先

日本国特許庁(ISA/JP)

〒100-8915

日本国

東京都千代田区霞が関三丁目4番3号

権限のある職員（特許庁審査官）

金沢 史明 5S 4538

電話番号 03-3581-1101 内線 3529

様式 PCT/ISA/210（第2ページ）（2022年7月）

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2024/008140

引用文献			公表日	パテントファミリー文献			公表日
WO	2022/079904	A1	21.04.2022	US	2023/0367846	A1	
				EP	4231272	A1	
				CN	116324935	A	
CN	111506922	A	07.08.2020	(ファミリーなし)			