(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
2014 年 11 月 13 日 (13.11.2014)

W 画P O | P C T

(10) 国際公開番号
WO 2014/181541 A1

(51) 国際特許分類：
G06F 21/60 (2013.01)

(21) 国際出願番号： PCT/JP2014/002439

(22) 国際出願日： 2014 年 5 月 8 日 (08.05.2014)

(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 2013-099432 2013 年 5 月 9 日 (09.05.2013) JP

(71) 出願人：日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒 1088001 東京都港区芝五丁目 7 番 1 号 Tokyo (JP).

(72) 発明者：高橋 翼 (TAKAHASHI, Tsubasa); 〒 1088001 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内 Tokyo (7P).

(74) 代理人：下坂 直樹 (SHIMOSAKA, Naoki); 〒 1088001 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内 Tokyo (7P).

(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, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, ML, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, ML, NA, RW, SD, SL, 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, 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：
– 国際調査報告 (条約第 21 条(3))

(54) Title: INFORMATION PROCESSING DEVICE THAT VERIFIES ANONYMITY AND METHOD FOR VERIFYING ANONYMITY

(54) 発明の名称：匿名性を検証する情報処理装置及び匿名性検証方法

[図8]

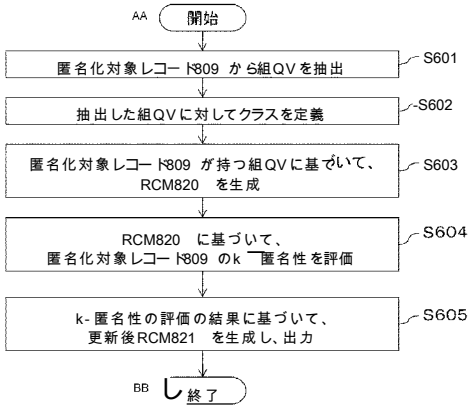


Fig. 8:
S601 Extract set (QV) from record (809) that is subject of anonymization
S602 Define class for extracted set (QV)
S603 Generate RCM (820) on basis of set (QV) had by record (809) that is subject of anonymization
S604 Evaluate k-anonymity of record (809) that is subject of anonymization on basis of RCM (820)
S605 Generate and output post-update RCM (821) on basis of k-anonymity evaluation results
AA Start
BB End

(51) Abstract: The present invention provides an information processing device that enables a reduction in the processing cost of verifying anonymity during anonymization when multi-dimensional data is the subject of anonymization. The information processing device is provided with: a means for generating information indicating the correspondence between a record contained in a data set and a class specifying a unique combination of quasi-identifier attribute values; a means for verifying the anonymity of each record on the basis of the class thereof indicated in the information; and a means for, on the basis of the results of verifying the anonymity, updating and outputting the information in a manner such that the records that do not satisfy the anonymity can be identified.

(57) 要約：本発明は、多次元データを匿名化の対象とする場合に、その匿名化における匿名性検証の計算コストの低減を可能にする情報処理装置を提供する。その情報処理装置は、データセットに含まれるレコードと、準識別子属性値の一意的な組み合わせを特定するクラスとの対応を示す情報を生成する手段と、その情報に示されるそのクラスに基づいて、そのレコード毎の匿名性を検証する手段と、その匿名性を検証した結果に基づいて、その匿名性を満たさないそのレコードを識別可能なように、その情報を更新し、出力する手段と、を備える。

WO 2014/181541 A1

明　細　書

発明の名称　：匿名性を検証する情報処理装置及び匿名性検証方法

技術分野

[0001] 本発明は匿名化の技術に関し、例えば個人情報などの、オリジナルな情報内容のままで公開及び利用されることが好ましくない情報について匿名性を検証する技術に関する。

背景技術

[0002] 近年、購買情報や診療情報等のようなプライバシ情報が、販売サービスや診療サービスの提供者 (サービス事業者) によって蓄積されている。

[0003] そのプライバシ情報は、複数の属性からなり、準識別子と呼ばれる属性を含む。準識別子は、生年や性別などといった個人を特徴付け、それらの組み合わせから個人を特定する可能性を持つ属性である。

[0004] このようなプライバシ情報は、プライバシ侵害の懸念から積極的な二次活用がなされていない。例えば、二次活用とは、プライバシ情報を生成、蓄積しているサービス事業者が、第三者にそのプライバシ情報を提供し、その第三者が自身の提供するサービスを強化するために、そのプライバシ情報を利用することを指す。また、二次活用とは、プライバシ情報を生成、蓄積しているサービス事業者が、第三者に対してそのプライバシ情報の分析などのァゥトソーシングを依頼することなどを指す。

[0005] プライバシ侵害の懸念なしに二次活用ができれば、プライバシ情報を利用した研究の促進、分析・研究結果を用いたサービスの強化を行うことができる。そして、プライバシ情報を保有するサービス事業者以外の第三者も、プライバシ情報の持つ高い有益性を享受することができる。

[0006] 例えば、第三者と想定し得るものとして製薬会社がある。その製薬会社にとっては診療情報を入手することは、困難である。その製薬会社は、その診療情報を入手することができれば、薬品がどのように利用されているのかを知ることができる。更には、その製薬会社は、その診療情報から、薬品の共

起関係や相関関係などを分析することもできる。

[0007] 例えば、プライバシ情報のデータセットは、サービス利用者（個人）を一意に識別するユーザ識別子と、一つ以上の準識別子とを一つのレコードとして含む。そして、サービス提供者は、サービス利用者がサービスを享受する度に、係るレコードを蓄積する。

[0008] ユーザ識別子が付与されたままのプライバシ情報が第三者に提供される場合、その第三者は、そのユーザ識別子を用いることによってサービス利用者を特定することが可能である。そのため、プライバシ侵害の問題が発生し得る。

[0009] また、複数のレコードによって構成されるデータセット（例えば、履歴情報など）の中から、各レコードに付与されている準識別子の組み合わせに基づいて、ある個人を特定できてしまう場合がある。即ち、例えユーザ識別子を取り除いた履歴情報であっても、準識別子の組み合わせに基づいてある個人を特定可能な場合があり、プライバシ侵害が発生し得る。

[001 0] このような特性を有するプライバシ情報のデータセットを、本来の有用性を保ちながら、プライバシを保護した形態に変換する手法として、匿名化（匿名化技術：Ａｎｏｎｙｍｉｚａｔｉｏｎ）が知られている。

[001 1] 非特許文献１は、最もよく知られた匿名性指標である"ｋ－匿名性"を提案する。匿名化対象のデータセットに、係るｋ＿匿名性を充足させる手法は、"ｋ＿匿名化"と呼ばれる。このｋ＿匿名化では、同じ値の準識別子を有するレコードがその匿名化対象のデータセットの中に少なくともｋ個以上存在するように、その準識別子を変換する処理が行われる。この変換処理としては、一般化、切り落とし等の方式が知られている。係る一般化において、元の詳細な情報は、抽象化された情報に変換される。

[001 2] 例えば、特許文献１は、プライバシ情報評価サーバを開示する。第１に、そのプライバシ情報評価サーバは、ユーザ端末から受信したプライバシ情報を加工する。第２に、そのプライバシ情報評価サーバは、その加工したプライバシ情報がｋ-匿名性を満たすか否かを判定する。第３に、そのプライバシ

情報評価サーバは、その判定結果に基づいて、ユーザの識別情報を除いたその加工したプライバシ情報を出力する。

[001 3] また、このようなｋ－匿名化技術を利用する他の関連技術が、非特許文献２に記載されている。非特許文献２では、類似する属性値を持つレコードの集合 (以降、クラスタと呼ぶ) を逐次生成し、クラスタに含まれるレコードにおいて、一般化や切り落としにより共通の属性値を生成することで多次元データをｋ－匿名化する手法が提案されている。

[0014] そのプライバシ情報評価サーバのｋ-匿名性判定部は、ｋ-匿名性判定のフィードバックに基づいて、ボトムアップ処理或いはトップダウン処理により、そのプライバシ情報を一般化する。

[001 5] また、特許文献２は、公開情報のプライバシ保護装置を開示する。第１に、そのプライバシ保護装置は、入力されたデータの各準識別子を加工し、一般化を行う。第２に、そのプライバシ保護装置は、その一般化された準識別子の全てから構成される表が、所定のｋ-匿名性を満たすことを判定する。第３に、そのプライバシ保護装置は、その判定した結果に基づいて、最適なデータセットを出力する。

先行技術文献

特許文献

[001 6] 特許文献１ :特開２０１１_１８０８３９

特許文献２ :特開２０１２-００３４４０

非特許文献

[001 7] 非特許文献１ :Ｌ．Ｓｗｅｅｎｅｙ、"ｋ—ａｎｏｎｙｍｉｔｙ：ａ ｍｏｄｅｌ ｆｏｒ ｐｒｏｔｅｃｔｉｎｇ ｐｒｉｖａｃｙ”、 Ｉｎｔｅｒｎａｔｉｏｎａｌ Ｊｏｕｒｎａｌ ｏｎ Ｕｎｃｅｒｔａｉｎｔｙ，Ｆｕｚｚｉｎｅｓｓ ａｎｄ Ｋｎｏｗｌｅｄｇｅ—ｂａｓｅｄ Ｓｙｓｔｅｍｓ，１０（５），ｐｐ．５５５-５７０，２００２．

非特許文献２ :Ｋ．ＬｅＦｅｖｒｅ，Ｄａｖｉｄ Ｊ．ＤｅＷｉｔｔ ａｎｄ Ｒａｇｈｕ Ｒａｍａｋｒｉｓｈｎａｎ，"Ｍｏｎｄｒｉａｎ

Ｍｕｌｔｉｄｉｍｅｎｓｉｏｎａｌ ｋ—Ａｎｏｎｙｍｉｔｙ", ＩＣＤＥ２００６．Ｐｒｏｃｅｅｄｉｎｇｓ ｏｆ ｔｈｅ ２２ｎｄ Ｉｎｔｅｒｎａｔｉｏｎａｌ Ｃｏｎｆｅｒｅｎｃｅ ｏｎ Ｄａｔａ Ｅｎｇｉｎｅｅｒｉｎｇ Ｐａｇｅ ２５，２００６．

発明の概要

発明が解決しようとする課題

[001 8] しかしながら、上述した特許文献及び非特許文献に記載された技術においては、多次元データを匿名化の対象とした場合、その次元数に比例して、匿名性検証の計算コストが増加するという問題点がある。

[001 9] その理由は、上述した特許文献及び非特許文献に記載された技術のいずれにおいても、多次元データを対象とする匿名化における、匿名化検証の処理特性を考慮していないからである。以下に、この理由を詳細に説明する。

[0020] ｋ＿匿名化の際には、データセットがｋ＿匿名性を満たすかどうかを判定する匿名性検証を多数実施する。特に、有効性 (Ｕｔｉｌｉｔｙ) の高いｋ－匿名化データを生成するためには、各属性 (準識別子) を有効性が高くなるように段階的に加工 (再符号化) する必要がある。そして、各属性を加工される度にその匿名性検証を実施する必要がある。

[0021] ここで、加工する属性が複数であることは、即ち多次元データを匿名化の対象としていることである。

[0022] その匿名性検証では、まず検証の対象であるデータセットについて、全ての準識別子に対して同一の値の組を持つ、レコード数またはユニークなユーザ識別子の種類数が、数えられる。ここで、そのデータセットは、そのユーザ識別子を含むそのレコードを含む。

[0023] ここで、全ての準識別子に対して同一の値の組を持つそのレコードの集合を等価クラス (Ｅｑｕｉｖａｌｅｎｔ Ｃｌａｓｓ) またはクラスと呼ぶ。

[0024] あるクラスのレコード数 (またはユニークなユーザ識別子の種類数) が ｋ 以上である場合、そのクラスに含まれるレコード集合はｋ＿匿名性を充足していると判断される。また、あるクラスのレコード数がｋ未満である場合、

そのクラスに含まれるレコード集合はｋ—匿名性を充足しないと判断される。この匿名性検証を、あるデータセットに含まれる全てのレコードに対して実施し、全てのそのレコードがｋ—匿名性を満たせば、そのデータセットはｋ—匿名性を満たすと判断される。

[0025] その匿名性検証では、その対象とするデータセットで出現する組Ｑについて、そこに含まれるレコード数またはユニークなユーザ識別子の種類数を数える必要がある。ここで組Ｑは、全ての準識別子の全ての属性値の、組み合わせである。このとき、その準識別子の数が増加すると、組Ｑの組み合わせ数が増大する。

[0026] 例えば、ｄ（ｄ≧１）個の準識別子ｑｉ（ｉ＝１、・・・、ｄ）があり、各準識別子ｑｉの属性値の種類数をｌｑｉｌとする。この場合、その匿名性検証で考慮すべき、その組み合わせ数の最大値は、ｌｑ１｜Ｘ｜ｑ２｜Ｘ∽Ｘｌｑｄｌである。各準識別子が再符号化されることで、その組み合わせ数は増減するが、基本的にはその準識別子の数に比例して、その組み合わせ数は増大する。

[0027] また、その匿名性検証において、組Ｑを抽出する処理では、その準識別子の数であるｄ回のデータ参照が必要になる。そのレコード数をＮ（Ｎ≧１）とすると、ｄＮ回の参照が必要となる。更に、その各準識別子が分散して配置されている場合、１回の参照コストも大きくなり、トータルの参照コストが更に大きくなる。

[0028] 本発明の目的は、上述した問題点を解決できる情報処理装置、匿名性検証方法、及びそのためのプログラム或いはそのプログラムを記録したコンピュータ読み取り可能な非一時的記録媒体を提供することにある。

課題を解決するための手段

[0029] 本発明の一様態における情報処理装置は、データセットに含まれるレコードと、前記レコードに含まれる準識別子の値である準識別子属性値の、一意的な組み合わせを特定するクラスとの対応を示すレコードークラス対応情報（Ｒｅｃｏｒｄ-Ｃｌａｓｓ　Ｍａｐ、ＲＣＭ）を生成する対応情報生成手

段と、前記レコード—クラス対応情報に示される前記クラスに基づいて、前記レコード毎の匿名性を検証する匿名性検証手段と、前記匿名性を検証した結果に基づいて、前記匿名性を満たさない前記レコードを識別可能なように、前記レコード–クラス対応情報を更新し、出力する対応情報更新手段と、を含む。

[0030] 本発明の一様態における匿名化方法は、情報処理装置が、データセットに含まれるレコードと、前記レコードに含まれる準識別子の値である準識別子属性値の、一意的な組み合わせを特定するクラスとの対応を示すレコードークラス対応情報を生成し、前記レコードークラス対応情報に示される前記クラスに基づいて、前記レコード毎の匿名性を検証し、前記匿名性を検証した結果に基づいて、前記匿名性を満たさない前記レコードを識別可能なように、前記レコード—クラス対応情報を更新し、出力する。

[0031] 本発明の一様態におけるコンピュータ読み取り可能な非一時的記録媒体は、データセットに含まれるレコードと、前記レコードに含まれる準識別子の値である準識別子属性値の、一意的な組み合わせを特定するクラスとの対応を示すレコードークラス対応情報を生成し、前記レコードークラス対応情報に示される前記クラスに基づいて、前記レコード毎の匿名性を検証し、前記匿名性を検証した結果に基づいて、前記匿名性を満たさない前記レコードを識別可能なように、前記レコードークラス対応情報を更新し、出力する処理をコンピュータに実行させるプログラムを記憶する。

発明の効果

[0032] 本発明は、多次元データを匿名化の対象とする場合に、その匿名化における匿名性検証の計算コストの低減を可能にするという効果がある。

図面の簡単な説明

[0033] [図1] 図1は本発明の第1の実施形態に係る匿名性検証装置の構成を示すブロック図である。

[図2] 図2は第1の実施形態における匿名化対象データセットの一例を示す図である。

[図3] 図3は第1の実施形態におけるクラス定義の一例を示す図である。

[図4] 図4は第1の実施形態におけるレコード—クラスマップの一例を示す図である。

[図5] 図5は第1の実施形態における匿名性検証結果の一例を示す図である。

[図6] 図6は第1の実施形態における更新後レコード—クラスマツプの一例を示す図である。

[図7] 図7は、第1の実施形態に係る匿名性検証装置を実現するコンピュータのハードウエア構成を示すプロック図である。

[図8] 図8は第1の実施形態における匿名性検証装置の動作を示すフローチャート である。

[図9] 図9は本発明の第2の実施形態に係る匿名性検証装置の構成を示すプロック図である。

[図10] 図10は第2の実施形態におけるレコード—クラスマップの一例を示す図である。

[図11] 図11は第2の実施形態におけるカレントステートテーブルの一例を示す図である。

[図12] 図12は第2の実施形態における中間データの一例を示す図である。

[図13] 図13は第2の実施形態における属性の一般化階層の一例を示す図である。

[図14] 図14は第2の実施形態における匿名性検証結果の一例を示す図である。

[図15] 図15は第2の実施形態における更新後カレントステートテーブルの一例を示す図である。

[図16] 図16は第2の実施形態における更新後レコード—クラスマップの一例を示す図である。

[図17] 図17は第2の実施形態における匿名性検証装置の動作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

[0034] 以下、本発明の実施形態に関して、図面を参照して詳細に説明する。尚、各図面及び明細書記載の各実施形態において、同様の構成要素には同様の符号を付与し、適宜説明を省略する。

また、以下の実施形態及び各図において、本発明の本質に関わらない構成については一般的な技術を採用することとし、本実施形態における詳細な説明及び図示は省略する。

[0035] <<< 第１の実施形態>>>

図１は、第１の実施形態に係る匿名性検証装置（情報処理装置とも呼ばれる）１００の構成を示すブロック図である。図１に示すように、匿名性検証装置１００は、ＲＣＭ生成部（対応情報生成手段とも呼ばれる）１０２と、匿名性検証部１０６と、ＲＣＭ更新部（対応情報更新手段とも呼ばれる）１０８とを含む。尚、図１に示す構成要素は、ハードウエア単位の構成要素でも、コンピュータ装置の機能単位に分割した構成要素でもよい。ここでは、図１に示す構成要素は、コンピュータ装置の機能単位に分割した構成要素として説明する。

[0036] まず、本実施形態に係る匿名性検証装置１００の動作の概要を説明する。

[0037] 匿名性検証装置１００は、匿名化対象データセット（単に、データセットとも呼ばれる）に含まれる全ての匿名化対象レコード（単に、レコードとも呼ばれる）８０９のそれぞれから、組Ｑを抽出する。ここで、組Ｑは、「準識別子の値（準識別子属性値とも呼ばれる）の、全種類の準識別子についての、一意的な組み合わせ」を示す。尚、準識別子の値は、準識別子の具体的内容を表すデータであり、数値に限定されない（例えば、文字などの）情報である。

[0038] 次に、匿名性検証装置１００は、その抽出した各組Ｑを一意的に特定するクラスを定義し、それらの匿名化対象レコード８０９のそれぞれをそれらのクラスに対応付けるレコードークラスマップ（ＲＣＭ、レコードークラス対応情報とも呼ばれる）を生成する。

[0039] そして、匿名性検証装置１００は、そのレコード―クラスマップに基づい

て、各匿名化レコードr'のｋ—匿名性を検証する。ここで、匿名化レコードr'とは、匿名化対象レコード８０９に対応する、匿名化対象レコード８０９が匿名化された或いは匿名化されるレコードである。

[0040] そして、匿名性検証装置１００は、それらのｋ＿匿名性を検証した結果に基づいて、そのｋ＿匿名性を満たさない匿名化レコードド を識別可能なように、そのレコードークラスマップを更新する。

[0041] 次に、本実施形態における匿名性検証装置１００が備える各構成要素について説明する。

＝＝＝ＲＣＭ生成部１０２＝＝＝

ＲＣＭ生成部１０２は、匿名化対象データセット８００に含まれる全ての匿名化対象レコード８０９から、組Ｑを抽出する。

[0042] 図２は、本実施形態における匿名化対象データセット８００の一例を示す図である。

[0043] 図２に示すように、匿名化対象データセット８００は、複数の匿名化対象レコード８０９を含む。匿名化対象レコード８０９は、匿名化対象レコード８０９自身を識別するＲＩＤ（Ｒｅｃｏｒｄ Ｉｄｅｎｔｉｆｉｅｒ、レコード識別子、レコードＩＤ）と、１つ以上の準識別子とを含む。その準識別子は、例えば、 性別」、 生年」、 診療月」及び 傷病名」である。尚、匿名化対象データセット８００は、図２に示す例に係わらず、任意の属性を含んでよい。また、その準識別子は、匿名化対象データセット８００に含まれる属性の内、任意の属性であってよい。

[0044] 匿名化対象データセット８００は、オリジナルな情報内容のままで公開や利用されることが好ましくない、個人情報等の情報である。

[0045] 例えば、ＲＣＭ生成部１０２は、図２に示す匿名化対象データセット８００から、以下の６種類の組Ｑを抽出する。

[0046] {"男性"、"１９３０〜１９３９"、"４〜６"、"Ａ、Ｂ、Ｃ"}、{"男性„、"１９４０〜１９４９"、"４〜６"、"Ｘ、Ｙ、Ｚ"}、{"男性"、"１９５０〜１９５９"、"４〜６"、"Ｘ、Ｙ、Ｚ"}、{"

女性"、"１９３０～１９３９"、"４～６"、"Ａ、Ｂ、Ｃ"}、{"女性"、"１９４０～１９４９"、"４～６"、"Ｘ、Ｙ、Ｚ"}、{"女性"、"１９５０～１９５９"、"４～６"、"Ｘ、Ｙ、Ｚ"}。

[0047] 上述したように、これらの組Ｑは、準識別子の値の、全種類の準識別子についての、一意的な組み合わせである。尚、その組Ｑは、匿名性検証部１０６がｋ－匿名性検証に用いる任意の種類の準識別子の値の、一意的な組み合わせであってもよい。

[0048] 次に、ＲＣＭ生成部１０２は、その抽出した組Ｑのそれぞれに対してクラスを定義する。

[0049] 図３は、本実施形態における、クラス定義表８１０の一例を示す図である。図３に示すようにクラス定義表８１０は、その抽出された組ＱとＣＩＤ（Ｃｌａｓｓ　Ｉｄｅｎｔｉｆｉｅｒ、クラス識別子、クラスＩＤ）との対応を示す。そのクラスＩＤは、その組Ｑのそれぞれに対して定義されたＩＤ情報であり、そのクラスを一意に識別する。

[0050] ＲＣＭ生成部１０２は、各クラスのＣＩＤに、重複がなく一意に識別できる任意の値を付与する。例えば、ＲＣＭ生成部１０２は、その組Ｑの抽出順に対して、１から始まる連番を付与する。

[0051] 次に、ＲＣＭ生成部１０２は、匿名化対象レコード８０９のそれぞれに対応する組Ｑに基づいて、レコード－クラスマップを生成する。レコード－クラスマップは、匿名化対象レコード８０９とその定義したクラスとを対応付ける。同一のクラスに属する匿名化対象レコード８０９の集合は、同一の組Ｑ（即ち、同一の準識別子の値の一意的な組み合わせ）に対応する匿名化対象レコード８０９の集合である。

[0052] 例えば、ＣＩＤが「1」である組Ｑの「{"男性"、"１９３０～１９３９"、"４～６"、"Ａ、Ｂ、Ｃ"}」を持つ匿名化対象レコード８０９の集合は、ＲＩＤが「1」及び「2」の匿名化対象レコード８０９である。よって、ＲＩＤが「1」及び「2」の匿名化対象レコード８０９は、ＣＩＤが「1」のクラスに属する。ＲＣＭ生成部１０２は、ＲＩＤが「3」～「１０

」の匿名化対象レコード809についても、同様にその組Qに基づいて、その各クラスのいずれかに対応付ける。

[0053] 図4は、本実施形態におけるRCM820の一例を示す図である。図4に示すように、RCM820は、「1Dと○1Dとの対応、即ち匿名化対象レコード809とクラスとの対応を示す。

[0054] ===匿名性検証部106===

匿名性検証部106は、RCM820に基づいて、匿名化レコードr'のk−匿名性を評価する。

[0055] 具体的には、匿名性検証部106は、RCM820中のCID毎の頻度（RIDの数）を計数し、その頻度に基づいて匿名化レコードr'のk—匿名性を評価する。即ち、その頻度がk以上のクラスは、そのk_匿名性を満たす。従って、そのk—匿名性を満たすそのクラスに属する匿名化レコードr，も、そのk_匿名性を満たす。

[0056] 図5は、匿名性検証部106によるk-匿名性の評価の結果である、匿名性検証結果830の一例を示す図である。

[0057] 図5に示す匿名性検証結果830は、図4のRCM820に対してk=2のk_匿名性を検証した結果を示す。各行は、CIDと、そのCIDに対応するRIDの頻度（クラス毎のRID数=匿名化対象レコード809の数）と、その頻度からそのk_匿名性を充足するか否かを評価した検証結果とを含む。検証結果は、「○K」の場合にそのk-匿名性を充足することを示し、「NG」の場合にそのk-匿名性を充足しないことを示す。

[0058] 図5は、CIDの「3」及び「6」に対応するクラスがそのk—匿名性を満たさないことを示す。従って、それらのクラスに属する、RIDの「5」及び「10」に対応する匿名化レコードr'がそのk_匿名性を満たさないことを示す。

[0059] このとき、匿名性検証部106は、CIDだけを参照してk_匿名性を検証することができ、N（レコード数）回の属性値の参照で匿名性の検証を実現できる。

[0060] 即ち、レコード数をＮ（Ｎ≧１）、準識別子数をｄ（ｄ≧１）とした場合、ＲＣＭ８２０の生成には、ｄＮ回の匿名化対象データセット８００の参照が必要になる。しかし、以降の処理において匿名性検証を実施する際には、匿名性検証部１０６は、Ｎ回の参照、または検証したい匿名化レコードｒ'の属するそのクラスのそのレコード数だけの参照回数でｋ—匿名性を検証できる。

[0061] ＝＝＝ＲＣＭ更新部１０８＝＝＝

ＲＣＭ更新部１０８は、匿名性検証結果８３０に基づいて、更新後ＲＣＭを生成し、出力する。ＲＣＭ更新部１０８は、ｋ＿匿名性を満たさない匿名化レコードｒ'が明らかになるように、ＲＣＭ８２０を更新して更新後ＲＣＭを生成する。

[0062] 具体的には、ＲＣＭ更新部１０８は、ＲＣＭ８２０からｋ＿匿名性を満たさない匿名化レコードｒ'に対応する行を削除する。また、ＲＣＭ更新部１０８は、ｋ—匿名性を満たさないその匿名化レコードｒ'に対応するクラス１Ｄを、ｋ—匿名性を満たす匿名化レコードｒ'に対応するクラスＩＤと区別可能に管理するようにしてもよい。例えば、ＲＣＭ更新部１０８は、クラスＩＤにｋ＿匿名性を満たさないことを示す情報（例えば、所定の値）を付加するようにしてもよい。

[0063] 図６は、更新後ＲＣＭ８２１の一例を示す図である。図６に示すように、更新後ＲＣＭ８２１は、図４に示すＲＣＭ８２０から、ＲＩＤが「５」及び「６」の行が削除されている。即ち、更新後ＲＣＭ８２１は、ＲＣＭ８２０からｋ＿匿名性を満たさない匿名化レコードｒ'に対応する行（一般的に、情報とも呼ばれる）を削除したものである。

[0064] 以上が、匿名性検証装置１００の機能単位の各構成要素についての説明である。

[0065] 次に、匿名性検証装置１００のハードウエア単位の構成要素について説明する。

[0066] 本実施形態において、匿名性検証装置１００は、コンピュータ等の情報処

理装置によって実現することができる。匿名性検証装置１００及び後述する第２の実施形態における匿名化装置２００における各構成要素（機能プロック）は、情報処理装置が備えるハー ドウエア資源において、プログラムが実行されることによって実現される。プログラムは、コンピュータ・プロダラム或いはソフ トウェア・プログラムとも呼ばれる。

[0067] 例えば、匿名性検証装置１００は、コンピュータのＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）、主記憶装置、補助記憶装置等のハ- ドウエアと、記憶装置等から主記憶装置にロー ドされたプログラムとが協働することによって実現される。

[0068] 但し、プログラムの実装形態は、図１に示したプロック構成（ＲＣＭ生成部１０２、匿名性検証部１０６、ＲＣＭ更新部１０８）には限定されず、当業者が採用し得る様々な実装形態を適用可能である（以下の各実施形態においても同様）。尚、匿名性検証装置１００及び後述する各実施形態に係る匿名化装置は、専用の装置によって実現してもよい。

[0069] 図７は、本実施形態における匿名性検証装置１００を実現するコンビュータ７００のハー ドウエア構成を示す図である。

[0070] 図７に示すように、コンピュータ７００は、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）７０１、記憶部７０２、記憶装置７０３、入力部７０４、出力部７０５及び通信部７０６を含む。更に、コンピュータ７００は、外部から供給される記録媒体（または記憶媒体）７０７を含む。記録媒体７０７は、情報を非一時的に記憶する不揮発性記録媒体であってもよい。

[0071] ＣＰＵ７０１は、オペレーティングシステム（不図示）を動作させて、コンピュータ７００の、全体の動作を制御する。また、ＣＰＵ７０１は、例えば記憶装置７０３に装着された記録媒体７０７から、プログラムやデータを読み込み、読み込んだプログラムやデータを記憶部７０２に書き込む。ここで、そのプログラムは、例えば、後述の図８に示すフローチヤー トの動作をコンピュータ７００に実行させるプログラムである。

[0072] そして、ＣＰＵ７０１は、読み込んだプログラムに従って、また読み込んだデータに基づいて、図１に示すＲＣＭ生成部１０２、匿名性検証部１０６及びＲＣＭ更新部１０８として各種の処理を実行する。

[0073] 尚、ＣＰＵ７０１は、通信網 (不図示)に接続されている外部コンピュータ (不図示)から、記憶部７０２にプログラムやデータをダウンロードするようにしてもよい。

[0074] 記憶部７０２は、プログラムやデータを記憶する。記憶部７０２は、匿名化対象データセット８００、クラス定義表８１０、ＲＣＭ８２０、匿名性検証結果８３０及び更新後ＲＣＭ８２１を記憶してよい。

[0075] 記憶装置７０３は、例えば、光ディスク、フレキシブルディスク、磁気光ディスク、外付けハードディスク及び半導体メモリであって、記録媒体７０７を含む。記憶装置７０３ (記録媒体７０７) は、プログラムをコンピュータ読み取り可能に記憶する。また、記憶装置７０３は、データを記憶してもよい。記憶装置７０３は、匿名化対象データセット８００、クラス定義表８１０、ＲＣＭ８２０、匿名性検証結果８３０及び更新後ＲＣＭ８２１を記憶してよい。

[0076] 入力部７０４は、例えばマウスやキーボード、内蔵のキーボタンなどで実現され、入力操作に用いられる。入力部７０４は、マウスやキーボード、内蔵のキーボタンに限らず、例えばタツチパネルなどでもよい。入力部７０４は、例えば１０３の一部として含まれる。この場合、ＲＣＭ生成部１０２は、処理開始の指示や、匿名化対象データセット８００の指定を、入力部７０４を介して受け取るようにしてよい。

[0077] 出力部７０５は、例えばディスプレイで実現され、出力を確認するために用いられる。出力部７０５は、例えばＲＣＭ生成部１０２、匿名性検証部１０６及びＲＣＭ更新部１０８の一部として含まれる。この場合、各構成要素は、処理の結果や異常の発生などを、出力部７０５を介して出力するようにしてよい。

[0078] 通信部７０６は、外部とのインタフェースを実現する。通信部７０６は、

例えばＲＣＭ生成部１０２の一部として含まれる。この場合、ＲＣＭ生成部１０２は、通信部７０６を介して、外部装置から匿名化対象データセット８００を取得するようにしてよい。

[0079] 以上説明したように、図１に示す匿名性検証装置１００の機能単位のブロックは、図７に示すハードウエア構成のコンピュータ７００によって実現される。但し、コンピュータ７００が備える各部の実現手段は、上記に限定されない。すなわち、コンピュータ７００は、物理的に結合した１つの装置により実現されてもよいし、物理的に分離した２つ以上の装置を有線または無線で接続し、これら複数の装置により実現されてもよい。

[0080] 尚、上述のプログラムのコードを記録した記録媒体７０７が、コンピュータ７００に供給され、ＣＰＵ７０１は、記録媒体７０７に格納されたプログラムのコードを読み出して実行するようにしてもよい。或いは、ＣＰＵ７０１は、記録媒体７０７に格納されたプログラムのコードを、記憶部７０２、記憶装置７０３またはその両方に格納するようにしてもよい。すなわち、本実施形態は、コンピュータ７００（ＣＰＵ７０１）が実行するプログラム（ソフトウエア）を、一時的にまたは非一時的に、記憶する記録媒体７０７の実施形態を含む。

[0081] 以上が、本実施形態における匿名性検証装置１００を実現するコンピュータ７００の、ハードウエア単位の各構成要素についての説明である。

[0082] 次に本実施形態の動作について、図１～図８を参照して詳細に説明する。

[0083] 図８は、本実施形態の動作を示すフローチャートである。尚、このフローチャートによる処理は、前述したＣＰＵ７０１によるプログラム制御に基づいて、実行されても良い。また、処理のステップ名については、Ｓ６０１のように、記号で記載する。

[0084] ＲＣＭ生成部１０２は、匿名化対象データセット８００に含まれる全ての匿名化対象レコード８０９から、組Ｑを抽出する（ステップＳ６０１）。

[0085] 例えば、匿名化対象データセット８００は、図７に示す記憶部７０２或いは記憶装置７０３に予め記憶されていてよい。また、ＲＣＭ生成部１０２は

、図７に示す入力部７０４を介して操作者が入力した、匿名化対象データセット８００を取得するようにしてもよい。また、ＲＣＭ生成部１０２は、図７に示す通信部７０６を介して図示しない機器から、匿名化対象データセット８００を受信するようにしてもよい。また、ＲＣＭ生成部１０２は、図７に示す記憶装置７０３を介して、記録媒体７０７に記録された匿名化対象データセット８００を取得するようにしてもよい。

[0086] 次に、ＲＣＭ生成部１０２は、その抽出した組Ｑのそれぞれに対してクラスを定義する（ステップＳ６０２）。

[0087] 次に、ＲＣＭ生成部１０２は、匿名化対象レコード８０９のそれぞれが持つ組Ｑに基づいて、匿名化対象レコード８０９とそれらのクラスとを対応付けたＲＣＭ８２０を生成する（ステップＳ６０３）。

[0088] 次に、匿名性検証部１０６は、ＲＣＭ８２０に基づいて、各匿名化レコードｒ'のｋ－匿名性を評価する（ステップＳ６０４）。

[0089] 次に、ＲＣＭ更新部１０８は、そのｋ-匿名性の評価の結果に基づいて、ＲＣＭ８２０を更新して更新後ＲＣＭ８２１を生成し、出力する（ステップＳ６０５）。

[0090] 例えば、ＲＣＭ更新部１０８は、更新後ＲＣＭ８２１を図７に示す出力部７０５を介して出力する。また、ＲＣＭ更新部１０８は、図７に示す通信部７０６を介して、図示しない機器に更新後ＲＣＭ８２１を送信するようにしてもよい。また、ＲＣＭ更新部１０８は、図７に示す記憶装置７０３を介して、記録媒体７０７に更新後ＲＣＭ８２１を記録するようにしてもよい。

[0091] 上述したように、本実施形態の匿名性検証装置１００は、匿名化対象データセット８００のような多次元データを対象とした匿名化における匿名性検証の際に、匿名化レコードｒ'の数（例えば、Ｎ（Ｎ≧１））の回数だけＲＣＭ８２０を参照する。そして、匿名性検証装置１００がＲＣＭ８２０を参照する回数は、準識別子の数（例えばｄ（ｄ≧１））に係わらない。即ち、匿名性検証装置１００は、匿名性検証の際のデータ参照をＮ回に抑えることができる。一方、関連技術では、そのデータ参照がｄＮ回実施される。

[0092] 上述した本実施形態における効果は、多次元データを匿名化の対象とする場合に、その匿名化における匿名性検証の計算コストの低減を可能にする点である。

[0093] その理由は、ＲＣＭ生成部１０２がＲＣＭ８２０を生成し、匿名性検証部１０６がＲＣＭ８２０に基づいてｋ—匿名性を検証し、ＲＣＭ更新部１０８が検証結果に基づいてＲＣＭ８２０を更新して更新後ＲＣＭ８２１を生成し、出力するようにしたからである。

[0094] <<< 第２の実施形態>>>

次に、本発明の第２の実施形態について図面を参照して詳細に説明する。以下、本実施形態の説明が不明確にならない範囲で、前述の説明と重複する内容については説明を省略する。

[0095] 本実施形態は、匿名化処理中に実施する匿名性検証を、レコード—クラスマップを用いてデータの参照コストを抑制しつつ、実現する方法を開示する。

[0096] 本実施形態は、段階的にｋ－匿名性を満たしながら情報損失が低い状態を探索する場合を開示する。具体的には、本実施形態は、準識別子を最も一般化した状態から少しずつ詳細化（具体化）することと、その詳細化した場合の匿名性を検証することとを繰り返すことで再符号化を行う場合を開示する。

[0097] 図９は、本発明の第２の実施形態に係る匿名化装置（情報処理装置とも呼ばれる）２００の構成を示す機能プロック図である。

[0098] 図９に示すように匿名化装置２００は、ＲＣＭ生成部２０２とカレントステート生成部２０３と再符号化部２０５と匿名性検証部２０６とカレントステート更新部２０７とＲＣＭ更新部２０８とを含む。

[0099] ＝＝＝ＲＣＭ生成部２０２＝＝＝

ＲＣＭ生成部２０２は、匿名化対象データセット８００に基づいて、レコード—クラスマップ（ＲＣＭ）を生成する。

[0100] 図１０は、ＲＣＭ生成部２０２が生成するＲＣＭ８４０の一例を示す図で

ある。尚、ＲＣＭ８４０において、ＣＩＤは１つである。即ち、図１０に示すＲＣＭ８４０は、全ての匿名化レコードｒ'が同じクラスに属していることを示す。

[01 01] ＝＝＝カレントステート生成部２０３＝＝＝

カレントステート生成部２０３は、ＲＣＭ８４０に基づいて、準識別子毎にカレントステートテーブルを生成する。

[01 02] 図１１は、カレントステートテーブル８５０の一例を示す図である。図１１に示すように、カレントステートテーブル８５０は、ＲＩＤ（レコード識別子）とＣＩＤ（クラス識別子）と再符号化値との組を含む。

[01 03] 図１１に示すカレントステートテーブル８５０は、図２に示す匿名化対象データセット８００及び図１０に示すＲＣＭ８４０に基づいて、生成されたものである。そして、カレントステートテーブル８５０は、トップダウン再符号化の場合に対応するカレントステートレコードの例である。そのトップダウン再符号化は、所定のｋ－匿名性を充足しつつ、その準識別子の属性値の範囲や性質がオリジナルの属性値に近づくように、準識別子の値を加工（具体化）する再符号化である。

[01 04] その再符号化値は、準識別子の属性値を加工し、生成したものである。カレントステートテーブル８５０に含まれる再符号化値は、いずれの準識別子に対応するカレントステートテーブル８５０においても、全て同じ(例えば、「ＡＮＹ」）である。「ＡＮＹ」は、ある準識別子が取り得る全ての値を包含する。即ち、カレントステートテーブル８５０は、最も一般化された準識別子の値を再符号化値として含む。

[01 05] 尚、カレントステートレコードは、以下に示すボトムアップ再符号化の場合に対応するカレントステートレコードであってもよい。そのボトムアツプ再符号化は、その準識別子のオリジナルの属性値を、所定のｋ－匿名性を充足するように、加工 (一般化）する再符号化である。

[01 06] この場合、そのカレントステートレコードに含まれる再符号化値は、そのカレントステートレコードに対応する匿名化対象レコード８０９に含まれる

、オリジナルの属性値である。

[01 07] ＝＝＝再符号化部２０５＝＝＝

再符号化部２０５は、カレントステートテーブル８５０の再符号化値を加工し、中間データを生成する。

[01 08] 図１２は、中間データ８６０の一例を示す図である。図１２に示すように、中間データ８６０は、図１１に示すカレントステートテーブル８５０と同様に、ＲＩＤとＣＩＤと再符号化値との組を含む。ここで、中間データ８６０に含まれる再符号化値は、カレントステートテーブル８５０の再符号化値が加工されたものである。

[01 09] 図１３は、一般化階層（Ｇｅｎｅｒａｌｉｚａｔｉｏｎ Ｈｉｅｒａｒｃｈｙ）８７０の一例を示す図である。図１３に示すように、一般化階層８７０は、生年属性の一般化階層であり、最上位に生年属性が取り得る値を全て包含する値のルートを持ち、下位に向かって、順次具体化された値のノードを持つ。

[01 10] 例えば、再符号化部２０５は、図１３に示す一般化階層８７０に基づいて、図１１に示すカレントステートテーブル８５０に含まれる「生年」の属性（準識別子）を再符号化し、図１２に示す中間データ８６０を生成する。

[01 11] 具体的には、再符号化部２０５は、カレントステートテーブル８５０に含まれる「生年」の準識別子を、匿名化対象データセット８００に含まれる生年の準識別子の値と一般化階層８７０とに基づいて、再符号化する。具体的には、再符号化部２０５は、その「生年」の準識別子を、最も一般化された状態である「ＡＮＹ」から１段階詳細化を行った「１９００～１９４９」、「１９５０～１９９９」、「２０００～」へと再符号化する。

[01 12] 尚、中間データ８６０は、トップダウン再符号化の場合に対応する中間データの例である。ボトムアップ再符号化の場合、再符号化部２０５は、その「生年」の準識別子を、そのオリジナルの属性値から１段階一般化を行った値へと再符号化する。

[01 13] ＝＝＝匿名性検証部２０６＝＝＝

匿名性検証部２０６は、中間データ８６０に基づいて、各匿名化レコードr'のｋ＿匿名性を検証する。

[01 14] 例えば、匿名性検証部２０６は、中間データ８６０における、同一の (CｌＤ、再符号化値）の組を持つその匿名化レコードr'の数を計数する。または、ｋ—匿名性の検証は、同一の (ＣＩＤ、再符号化値）の組を持つその匿名化レコードr'に対応する、個人識別子等の属性の種類数を計数するようにしてもよい。

[01 15] 計数したレコード数がｋ以上である、同一の (ＣＩＤ、再符号化値）の組に対応する匿名化レコードr'は、ｋ＿匿名性を満たす。また、計数したレコード数がｋ未満の、同一の (ＣＩＤ、再符号化値）の組に対応する匿名化レコードに は、ｋ—匿名性を満たさない。ここで、同一の (ＣＩＤ、再符号化値）の組を持つその匿名化レコードr'のグループは、同一のクラスに属すると考えることができる。

[01 16] 図１４は、図１２の中間データ８６０に対してｋ—匿名性 (k＝２）を評価した結果を示す図である。図１４に示す匿名性検証結果８８０は、 (ＣＩＤ、再符号化値）の組と頻度とを含む。ここで、その頻度は、その (ＣＩＤ、再符号化値）の組を持つその匿名化レコードr'の数である。ＣＩＤが「１」及び再符号化値が「１９００〜１９４９」の組に対応するクラスはその頻度が「8」でありそのｋ—匿名性を満たすことを示している。また、図１４は、ＣＩＤが「1」及び再符号化値が「１９５０〜１９９９」の (ＣＩＤ、再符号化値）の組に対応するクラスはその頻度が「2」でありそのｋ＿匿名性を満たさないことを示している。

[01 17] ＝＝＝カレントステート更新部２０７＝＝＝

カレントステート更新部２０７は、後述するＲＣＭ更新部２０８が中間データ８６０を採用すると判定した場合、中間データ８６０に含まれる再符号化値を、カレントステートテーブル８５０に反映する。更に、カレントステート更新部２０７は、カレントステートテーブル８５０のＣＩＤを更新し、更新後カレントステートテーブル８５２を生成する。カレントステート更新

部２０７は、更新する新しいＣＩＤとして、図１４に示す匿名性検証結果８８０に含まれるＣＩＤと再符号化値との組み合わせに対して一意となる値を割り当てる。

[01 18] また、カレントステート更新部２０７は、後述の更新後ＲＣＭ８４２に基づいて、その再符号化した準識別子以外の、他の準識別子に対応するカレントステートテーブルのＣＩＤを更新する。

[01 19] また、カレントステート更新部２０７は、全ての準識別子の再符号化が終了した場合に、その更新した結果に対応する情報を出力する。その情報は、例えば、それら全ての準識別子のそれぞれに対応するＣＩＤを更新されたカレントステートテーブルである。または、その情報は、それらのＣＩＤを更新されたカレントステートテーブルに基づいて匿名化対象データセット８００が更新されて生成された匿名化データセットであってよい。

[01 20] 図１５は、図１１に示すカレントステートテーブル８５０が、再符号化値を反映され、図１４に示す匿名性検証結果８８０に基づいてクラスＩＤを更新された、更新後カレントステートテーブル８５２の一例を示す図である。ここで、その再符号化値は、図１２に示す中間データ８６０に含まれる再符号化値である。

[01 21] ＝＝＝ＲＣＭ更新部２０８＝＝＝

ＲＣＭ更新部２０８は、匿名性検証部２０６の匿名性検証結果８８０に応じて、中間データ８６０を採用するか否かを判定する。例えば、ＲＣＭ更新部２０８は、全匿名化対象レコード８０９のそれぞれに対応する匿名化レコード r' の８０％以上がｋ- 匿名性を満たす匿名性検証結果８８０を得られる場合、中間データ８６０を採用すると判定する。また、ＲＣＭ更新部２０８は、その匿名性検証結果８８０を得られない場合、中間データ８６０を採用しないと判定する。

[01 22] 中間データ８６０を採用すると判定した場合、ＲＣＭ更新部２０８は、ＣＩＤを変更された更新後カレントステートテーブル８５２に基づいて、ＲＣＭ８４０を更新する。具体的には、ＲＣＭ更新部２０８は、更新後カレント

ステートテーブル８５２のＲＩＤとＣＩＤとの組み合わせを抽出し、その抽出した内容に基づいて、ＲＣＭ８４０を更新する。

[01 23] ＲＣＭ更新部２０８は、更に、匿名性検証結果８８０に基づいて、ｋ＿匿名性を満たさない匿名化レコードｒ'に対応するＲＩＤとＣＩＤとの組み合わせを削除する。また、ＲＣＭ更新部２０８は、ｋ—匿名性を満たさない匿名化レコードｒ'に対応するクラスＩＤを、ｋ＿匿名性を満たすその匿名化レコードに対応するクラスＩＤと区別可能なように管理するようにしてもよい。例えば、ＲＣＭ更新部２０８は、ｋ—匿名性を満たさない匿名化レコードｒ'に対応するクラスＩＤにｋ＿匿名性を満たさないことを示す情報 (例えば、所定の値）を付加するようにしてもよい。

[01 24] 図１６は、図１０に示すＲＣＭ８４０が、ＣＩＤを更新され、ＲＩＤとＣＩＤとの組み合わせを削除された、更新後ＲＣＭ８４２の一例を示す図である。ここで、そのＣＩＤの更新は、図１５に示す更新後カレントステートテ—ブル８５２に基づいた更新である。また、ＲＩＤとＣＩＤとの組み合わせの削除は、図１４に示す匿名性検証結果８８０に基づいた削除である。

[01 25] また、中間データ８６０を採用しないと判定した場合、ＲＣＭ更新部２０８は、その中間データ８６０を破棄する。即ち、その場合、ＲＣＭ更新部２０８は、その中間データ８６０について、何の処理も実行しなくてよい。

[0 126] 次に、本実施形態の動作を、図面を参照して詳細に説明する。

[01 27] ＲＣＭ生成部２０２は、匿名化対象データセット８００に基づいて、ＲＣＭ８４０を生成する (ステップＳ６１１）。ここで、ＲＣＭ生成部２０２は、初期状態として全ての匿名化対象レコード８０９のクラスを同じクラス（例えば、ＣＩＤが「1」）に初期化する。

[01 28] 次に、カレントステート生成部２０３は、ＲＣＭ８４０に基づいて、準識別子毎にカレントステートテーブル８５０を生成する (ステップＳ６１３）。

[01 29] 次に、再符号化部２０５は、再符号化の可否を判定する (ステップＳ６１５）。ここで、再符号化の可否とは、匿名化の対象のいずれかの準識別子に

ついての、ｋ―匿名性を満たす再符号化の可否を示す。

[01 30] 再符号化が可と判定された場合 (ステップＳ６１５でＹＥＳ)、処理はステツプＳ６１７へ進む。再符号化が不可と判定された場合 (ステップＳ６１５でＮ○)、処理はステツプＳ６３１へ進む。

[01 31] 例えば、再符号化部２０５は、先行して実行されたステップＳ６２３ (後述)の処理において、中間データ８６０を採用しないと判定された再符号化に対応する準識別子について、再符号化が不可であると判断する。

[01 32] また、再符号化部２０５は、ステップＳ６１７以降の処理を実行しても、ＲＣＭ更新部２０８における中間データ８６０の採用基準を満足できる可能性がある場合に再符号化が可であると判定するようにしてもよい。ここでその採用基準は、例えば、匿名化対象レコード809のそれぞれに対応する匿名化レコードド の８０％以上がｋ＿匿名性を満たす匿名性検証結果８８０を得られるか否かである。そして、再符号化部２０５は、その採用基準を満足できる可能性がない場合に再符号化が不可であると判定するようにしてもよい。

[01 33] 尚、その採用基準を満足できる可能性がない場合は、例えば、前回のステップＳ６１７以降の処理実行による匿名性検証結果８８０が、全匿名化レコ―ドｒ' のちようど８０％がｋ―匿名性を満たすことを示している場合である。換言すると、その場合は、匿名化結果に含めない匿名化レコードｒ' を更に増やすと、匿名性検証結果８８０が、そのｋ＿匿名性を満たす匿名化レコードｒ' が８０％未満であることを示す場合である。

[01 34] 更に、再符号化部２０５は、その再符号化値が一般化階層８７０の最下層のノード (リーフ)のいずれかの値である場合に再符号化が不可であると判定するようにしてよい。ここで、その再符号化値は、ステップＳ６１５の一回目の処理においては、カレントステートテーブル８５０に含まれる再符号化値である。また、その再符号化値は、ステップＳ６１５の二回目以降の処理においては、中間データ８６０に含まれる再符号化値である。

[01 35] 再符号化部２０５は、上述に係わらず、任意の条件或いは条件の組み合わ

せに基づいて、再符号化の可否を判定してよい。

[01 36] この「ＲＣＭ更新部２０８における中間データ８６０の採用基準」に基づく判定は、ｋ＿匿名性を満たさない匿名化レコードド を匿名化結果には含めないことで、データセット全体としてはｋ—匿名性を満たすようにするという想定に基づくものである。

[01 37] 再符号化が可と判定された場合、再符号化部２０５は、再符号化の対象とする準識別子を選択する (ステップＳ６１７）。再符号化の対象とする準識別子は、再符号化可能な準識別子群から選択する。このとき、再符号化部２０５は、準識別子毎の情報損失等の情報量や、準識別子毎に定義された再符号化方法を用いて、再符号化を実施した場合の情報量の変化度合いや、その見積もり値を算出し、それらを用いて、再符号化の対象とする準識別子を１つ選択する。尚、再符号化部２０５は、選択の評価指標として任意の情報量を用いてよい。

[01 38] ここでは、生年属性が再符号化の対象として選択されたものとして、ステップＳ６１９以降のステップの動作を説明する。

[01 39] 再符号化部２０５は、対象準識別子のカレントステートテーブル８５０に含まれる再符号化値を、その対象準識別子に対応する属性毎の再符号化手法に基づいて加工し、中間データ８６０を生成する (ステップＳ６１９）。

[0140] 次に、匿名性検証部２０６は、ステップＳ６１９で生成した中間データ８６０に基づいて、各匿名化レコードｒ' のｋ—匿名性を検証する (ステップＳ６２１）。

[0141] 次に、匿名性検証部２０６は、その検証した結果に基づいて、中間データ８６０を採用するか否かを判定する (ステップＳ６２３）。例えば、図１２に示す中間データ８６０は、図１４の匿名性検証結果８８０が示すように、１０レコード中８レコード (８０％）がｋ—匿名性を満たす。このため、匿名性検証部２０６は中間データ８６０を採用すると判定する。

[0142] 中間データ８６０を採用する場合 (ステップＳ６２３でＹＥＳ）、処理はステツプＳ６２５へ進む。

[0143] 中間データ８６０を採用しない場合（ステップＳ６２３でＮＯ）、処理はステップＳ６１５へ戻る。この場合、その中間データ８６０は、破棄される。

[0144] 次に、カレントステート更新部２０７は、更新後カレントステートテーブル８５２を生成する（ステップＳ６２５）。

[0145] 次に、ＲＣＭ更新部２０８は、更新後ＲＣＭ８４２を生成する（ステップＳ６２７）。

[0146] 次に、カレントステート更新部２０７は、その更新後ＲＣＭ８４２に基づいて、その再符号化した準識別子以外の、他の準識別子に対応するカレントステートテーブルのＣＩＤを更新する（ステツプＳ６２９）。尚、カレントステート更新部２０７は、この更新の処理を、ステツプＳ６１９（再符号化）の前に、実施するようにしてもよい。その場合、カレントステート更新部２０７は、この更新の処理を再符号化の処理対象である準識別子に対応するカレントステートテーブルに対してのみ実施するようにしてもよい。

[0147] 次に、カレントステート更新部２０７は、準識別子のそれぞれに対応する、それらの全ての更新後カレントステートテーブルを出力する（ステップＳ６３１）。尚、カレントステート更新部２０７は、それらの更新後カレントステートテーブルに基づいて、匿名化対象データセット８００を更新して匿名化データセットを生成し、出力するようにしてもよい。

[0148] 上述したように、本実施形態の匿名化装置２００は、匿名化対象データセット８００のような多次元データを対象とした匿名化における匿名性検証の際に、匿名化レコードr'の数「Ｎ（Ｎ≧１）」の２倍の回数だけカレントステートテーブルを参照する。ここで、そのカレントステートテーブルは、カレントステートテーブル８５０或いは更新後カレントステートテーブル８５２である。そして、匿名化装置２００がそのカレントステートテーブルを参照する回数は準識別子の数「ｄ（ｄ≧１）」に係わらない。即ち、匿名性検証装置１００は、匿名性検証の際のデータ参照を２Ｎ回に抑えることができる。一方、関連技術では、そのデータ参照がｄＮ回実施される。

[0149] 上述した本実施形態における効果は、多次元データを匿名化の対象とする場合に、その匿名化における匿名性検証の計算コストの低減を可能にする点である。

[0150] その理由は、以下のような構成を含むからである。即ち、第１に、ＲＣＭ生成部２０２がＲＣＭ８４０を生成する。第２に、カレントステートテーブル生成部２０３がカレントステートテーブル８５０を生成する。第３に、再符号化部２０５が中間データ８６０を生成する。第４に、匿名性検証部２０６が中間データ８６０に基づいてｋ—匿名性を検証する。第５に、カレントステート更新部２０７がその検証結果に基づいてカレントステートテーブル８５０を更新する。第６に、ＲＣＭ更新部２０８がその検証結果及びその更新されたカレントステートテーブル８５０に基づいて、ＲＣＭ８４０を更新して更新後ＲＣＭ８４２を生成し、出力する。

[0151] 以上の各実施形態で説明した各構成要素は、必ずしも個々に独立した存在である必要はない。例えば、各構成要素は、複数の構成要素が１個のモジュールとして実現されてよい。また、各構成要素は、１つの構成要素が複数のモジュールで実現されてもよい。また、各構成要素は、ある構成要素が他の構成要素の一部であるような構成であってよい。また、各構成要素は、ある構成要素の一部と他の構成要素の一部とが重複するような構成であつてもよい。

[0152] 以上説明した各実施形態における各構成要素及び各構成要素を実現するモジュールは、必要に応じ、可能であれば、ハードウェア的に実現されてよい。また、各構成要素及び各構成要素を実現するモジュールは、コンピュータ及びプログラムで実現されてよい。また、各構成要素及び各構成要素を実現するモジュールは、ハードウェア的なモジュールとコンピュータ及びプログラムとの混在により実現されてもよい。

[0153] そのプログラムは、例えば、磁気ディスクや半導体メモリなど、不揮発性のコンピュータ可読記録媒体に記録されて提供され、コンピュータの立ち上げ時などにコンピュータに読み取られる。この読み取られたプログラムは、

そのコンピュータの動作を制御することにより、そのコンピュータを前述した各実施形態における構成要素として機能させる。

[01 54] また、以上説明した各実施形態では、複数の動作をフローチャートの形式で順番に記載してあるが、その記載の順番は複数の動作を実行する順番を限定するものではない。このため、各実施形態を実施するときには、その複数の動作の順番は内容的に支障がない範囲で変更することができる。

[01 55] 更に、以上説明した各実施形態では、複数の動作は個々に相違するタイミングで実行されることに限定されない。例えば、ある動作の実行中に他の動作が発生したり、ある動作と他の動作との実行タィミングが部分的に乃至全部において重複してもよい。

[01 56] 更に、以上説明した各実施形態では、ある動作が他の動作の契機になるように記載しているが、その記載はある動作と他の動作との全ての関係を限定するものではない。このため、各実施形態を実施するときには、その複数の動作の関係は内容的に支障のない範囲で変更することができる。また各構成要素の各動作の具体的な記載は、各構成要素の各動作を限定するものではなし。このため、各構成要素の具体的な各動作は、各実施形態を実施する上で機能的、性能的、その他の特性に対して支障をきたさない範囲内で変更されて良い。

[01 57] 以上、各実施形態を参照して本発明を説明したが、本願発明は上記実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

[01 58] この出願は、２０１３年５月９日に出願された日本出願特願２０１３＿０９９４３２を基礎とする優先権を主張し、その開示の全てをここに取り込む。

符号の説明

[01 59] １００ 匿名性検証装置

１０２ ＲＣＭ生成部

１０６ 匿名性検証部

１０８　ＲＣＭ更新部

２００　匿名化装置

２０２　ＲＣＭ生成部

２０３　カレントステート生成部

２０５　再符号化部

２０６　匿名性検証部

２０７　カレントステート更新部

２０８　ＲＣＭ更新部

７００　コンピュータ

７０１　ＣＰＵ

７０２　記憶部

７０３　記憶装置

７０４　入力部

７０５　出力部

７０６　通信部

７０７　記録媒体

８００　匿名化対象データセット

８０９　匿名化対象レコード

８１０　クラス定義表

８２０　ＲＣＭ

８２１　更新後ＲＣＭ

８３０　匿名性検証結果

８４０　ＲＣＭ

８４２　更新後ＲＣＭ

８５０　カレントステートテーブル

８５２　更新後カレントステートテーブル

８６０　中間データ

８７０　一般化階層

８８０　匿名性検証結果

請求の範囲

[請求項1] データセットに含まれるレコードと、前記レコードに含まれる準識別子の値である準識別子属性値の一意的な組み合わせを特定するクラスとの対応を示すレコードークラス対応情報を生成する対応情報生成手段と、

前記レコードークラス対応情報に示される前記クラスに基づいて、前記レコード毎の匿名性を検証する匿名性検証手段と、

前記匿名性を検証した結果に基づいて、前記匿名性を満たさない前記レコードを識別可能なように、前記レコードークラス対応情報を更新し、出力する対応情報更新手段と、を含む

情報処理装置。

[請求項2] 前記レコードークラス対応情報と前記準識別子属性値とに基づいて、前記準識別子の種類毎に、前記レコードと前記クラスと前記準識別子属性値に対応する再符号化値との対応を示すカレントステート情報を生成するカレントステート生成手段と、

前記カレントステート情報に含まれる前記再符号化値を加工する再符号化手段と、を含み、

前記匿名性検証手段は、前記再符号化値に基づいて、前記レコード毎の匿名性を検証し、

前記再符号化値と前記匿名性を検証した結果とに基づいて、前記カレントステート情報を更新し、その更新した結果に対応する情報を出力するカレントステート更新手段を更に含み、

前記対応情報更新手段は、前記更新されたカレントステート情報に基づいて、前記レコード—クラス対応情報を更新する

ことを特徴とする請求項1記載の情報処理装置。

[請求項3] 前記再符号化手段は、最上位に前記順識別子が取り得る値を全て包含する値のルートを持ち、下位に向かって順次具体化された値のノードを持つ一般化階層に基づいて、前記再符号化値を生成する

ことを特徴とする請求項２記載の情報処理装置。

[請求項4] 前記対応情報更新手段は、前記匿名性を満たさない前記レコードに対応する情報を、前記レコードクラス対応情報から削除することで、前記レコードクラス対応情報を更新する

ことを特徴とする請求項１乃至３のいずれか１項に記載の情報処理装置。

[請求項5] 前記対応情報更新手段は、前記レコードクラス対応情報に含まれる前記匿名性を満たさない前記レコードに対応する情報に、前記匿名生成を満たさないことを示す情報を付加することで、前記レコードクラス対応情報を更新する

ことを特徴とする請求項１乃至３のいずれか１項に記載の情報処理装置。

[請求項6] 前記対応情報更新手段は、前記レコードクラス対応情報に含まれる前記匿名性を満たさない前記レコードに対応する情報を、前記匿名生成を満たさないことを示す情報に変更することで、前記レコードクラス対応情報を更新する

ことを特徴とする請求項１乃至３のいずれか１項に記載の情報処理装置。

[請求項7] 情報処理装置が、

データセットに含まれるレコードと、前記レコードに含まれる準識別子の値である準識別子属性値の一意的な組み合わせを特定するクラスとの対応を示すレコード—クラス対応情報を生成し、

前記レコードクラス対応情報に示される前記クラスに基づいて、前記レコード毎の匿名性を検証し、

前記匿名性を検証した結果に基づいて、前記匿名性を満たさない前記レコードを識別可能なように、前記レコードクラス対応情報を更新し、出力する

匿名化検証方法。

[請求項8] 前記情報処理装置が、

前記レコードークラス対応情報と前記準識別子属性値とに基づいて、前記準識別子の種類毎に、前記レコードと前記クラスと前記準識別子属性値に対応する再符号化値との対応を示すカレントステート情報を生成し、

前記カレントステート情報に含まれる前記再符号化値を加工し、

前記匿名性の検証は、前記再符号化値に基づいて、前記レコード毎の匿名性を検証し、

前記再符号化値と前記匿名性を検証した結果とに基づいて、前記カレントステート情報を更新し、その更新した結果に対応する情報を出力し、

前記レコードークラス対応情報の更新は、前記更新されたカレントステート情報に基づいて、前記レコードークラス対応情報を更新することを特徴とする請求項7記載の匿名化検証方法。

[請求項9] データセットに含まれるレコードと、前記レコードに含まれる準識別子の値である準識別子属性値の一意的な組み合わせを特定するクラスとの対応を示すレコード—クラス対応情報を生成し、

前記レコードークラス対応情報に示される前記クラスに基づいて、前記レコード毎の匿名性を検証し、

前記匿名性を検証した結果に基づいて、前記匿名性を満たさない前記レコードを識別可能なように、前記レコードークラス対応情報を更新し、出力する処理をコンピュータに実行させる

プログラムを記録したコンピュータ読み取り可能な非一時的記録媒体。

[請求項10] 前記レコードークラス対応情報と前記準識別子属性値とに基づいて、前記準識別子の種類毎に、前記レコードと前記クラスと前記準識別子属性値に対応する再符号化値との対応を示すカレントステート情報を生成し、

前記カレントステート情報に含まれる前記再符号化値を加工し、

前記匿名性の検証は、前記再符号化値に基づいて、前記レコード毎の匿名性を検証し、

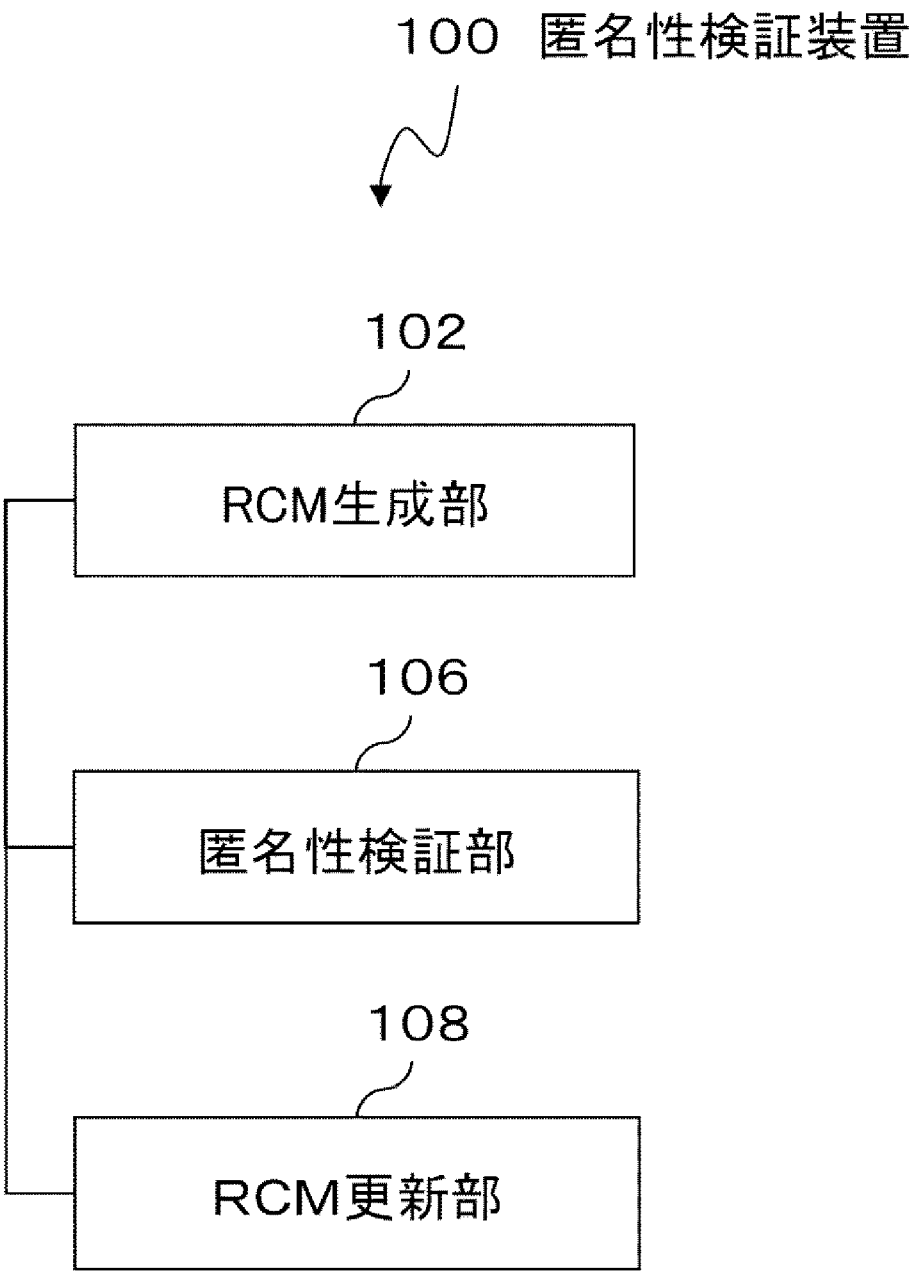
前記再符号化値と前記匿名性を検証した結果とに基づいて、前記カレントステート情報を更新し、その更新した結果に対応する情報を出力し、

前記レコードークラス対応情報の更新は、前記更新されたカレントステート情報に基づいて、前記レコードークラス対応情報を更新する処理をコンピュータに実行させる

請求項９記載のプログラムを記録したコンピュータ読み取り可能な非一時的記録媒体。

WO 2014/181541 PCT/JP2014/002439

[図1]



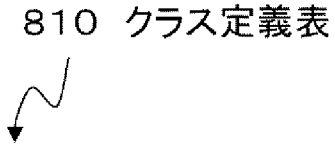
[図2]

800 匿名化対象データセット

RID	性別	生年	診療月	傷病名
1	男性	1930～1939	4～6	A, B, C
2	男性	1930～1939	4～6	A, B, C
3	男性	1940～1949	4～6	X, Y, Z
4	男性	1940～1949	4～6	X, Y, Z
5	男性	1950～1959	4～6	X, Y, Z
6	女性	1930～1939	4～6	A, B, C
7	女性	1930～1939	4～6	A, B, C
8	女性	1940～1949	4～6	X, Y, Z
9	女性	1940～1949	4～6	X, Y, Z
10	女性	1950～1959	4～6	X, Y, Z

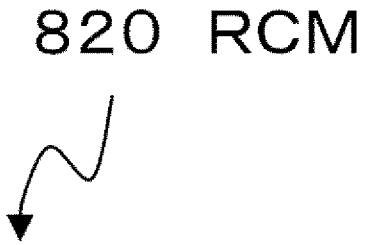
809 匿名化対象レコード

[図3]



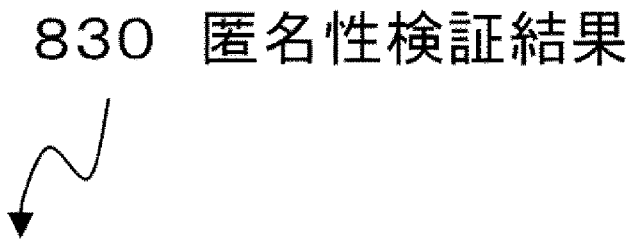
組Q(準識別子属性値の組み合わせ)	CID
“男性”、“1930～1939”、“4～6”、“A, B, C”	1
“男性”、“ 1940～1949 ”、“4～6”、“X, Y, Z”	2
“男性”、“ 1950～1959 ”、“4～6”、“X, Y, Z”	3
“女性”、“ 1930～1939 ”、“4～6”、“A, B, C”	4
“女性”、“ 1940～1949 ”、“4～6”、“X, Y, Z”	5
“女性”、“ 1950～1959 ”、“4～6”、“X, Y, Z”	6

[図4]



RID	CID
1	1
2	1
3	2
4	2
5	3
6	4
7	4
8	5
9	5
10	6

[図5]



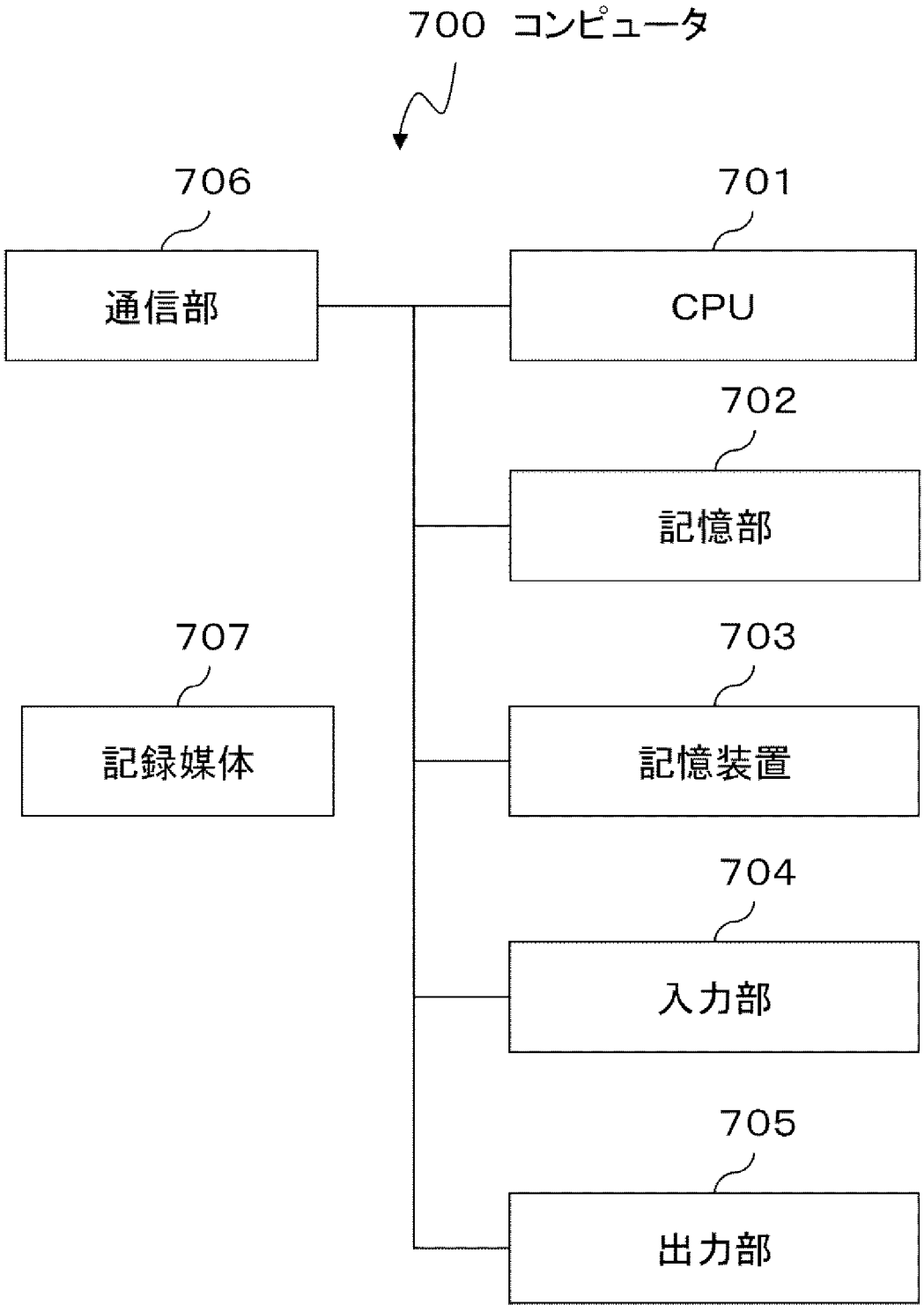
CID	頻度	検証結果
1	2	OK
2	2	OK
3	1	NG
4	2	OK
5	2	OK
6	1	NG

[図6]

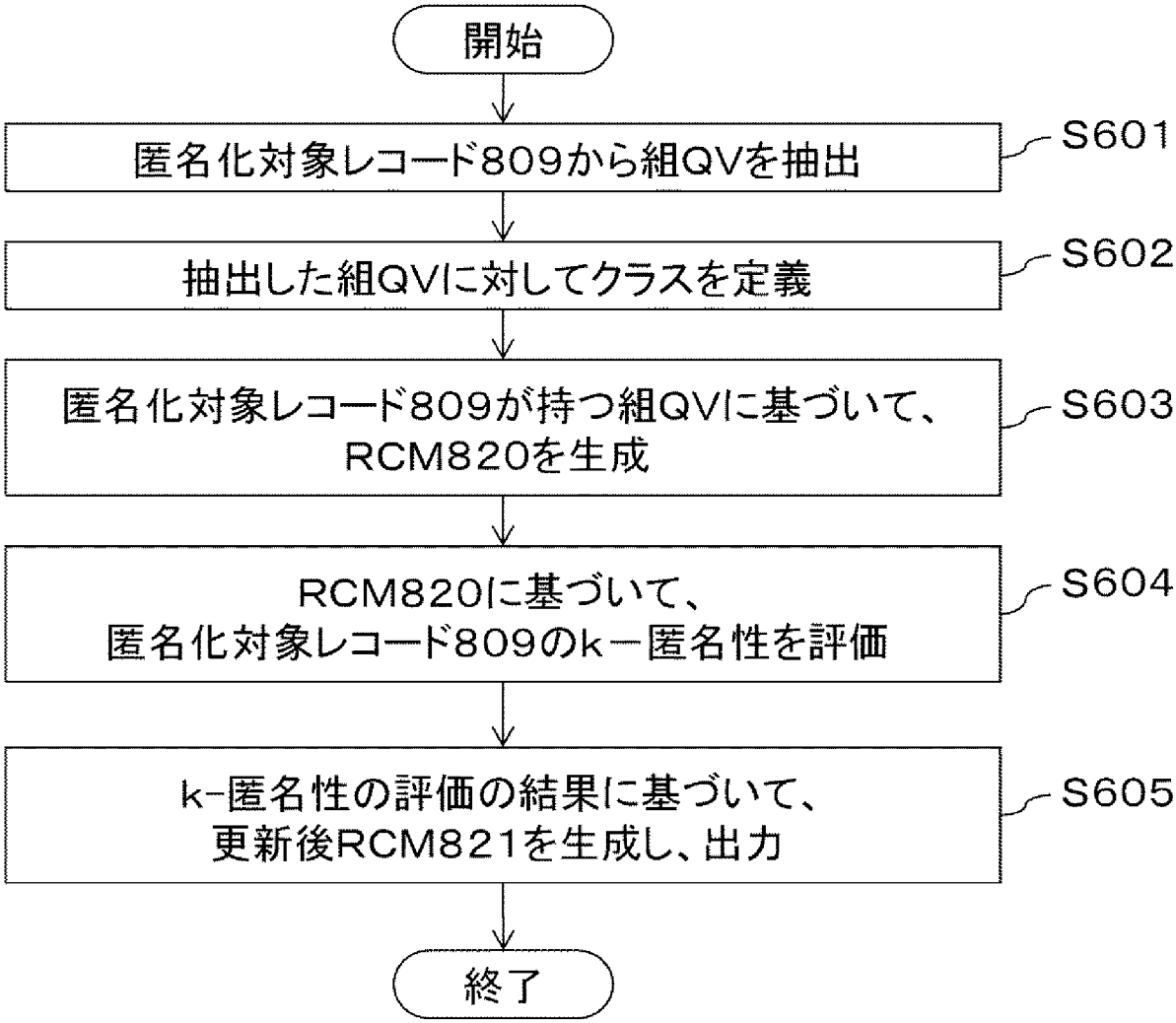
821 更新後RCM

RID	CID
1	1
2	1
3	2
4	2
6	4
7	4
8	5
9	5

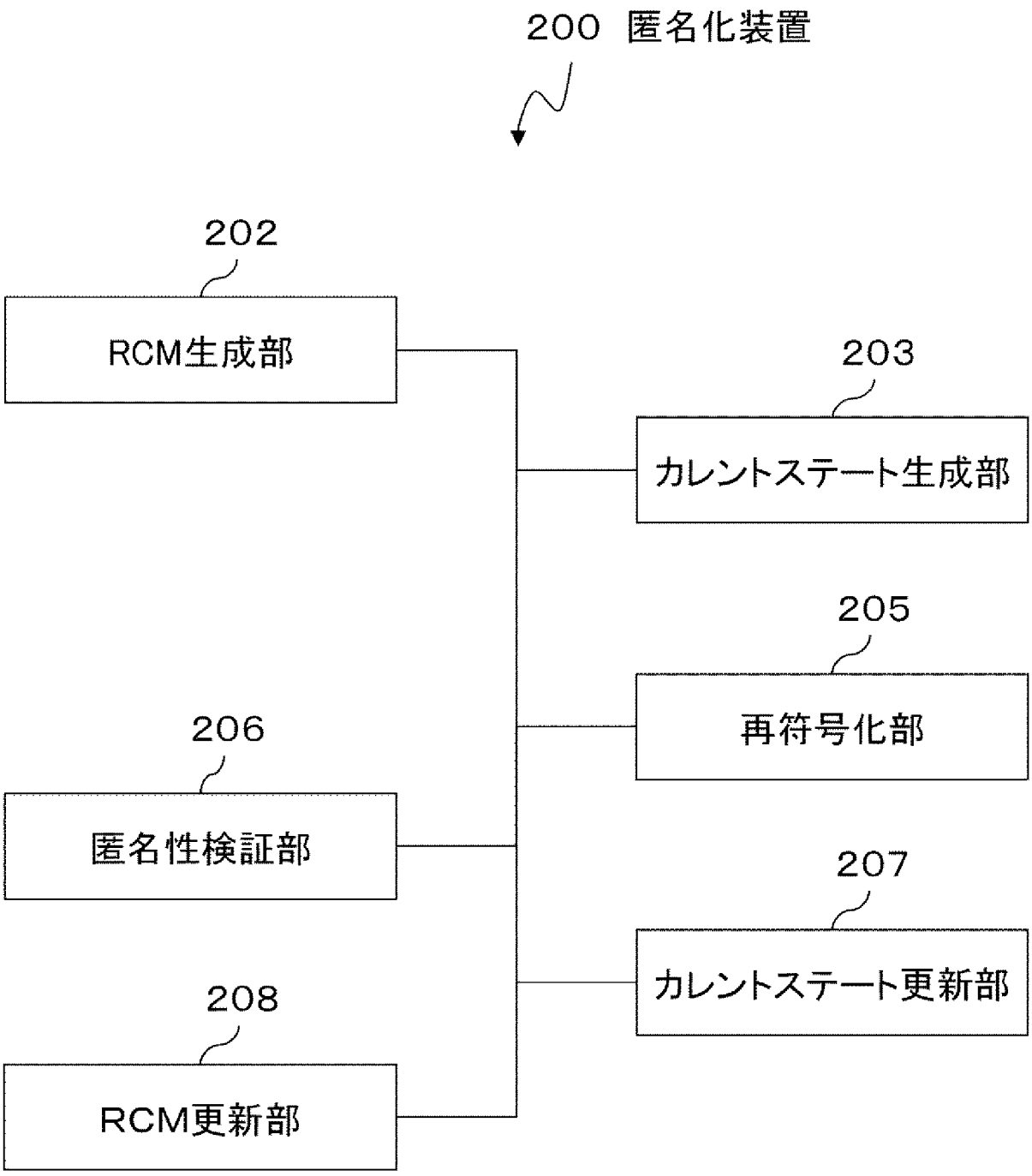
[図7]


700 コンピュータ
706
通信部
701
CPU
702
記憶部
707
記録媒体
703
記憶装置
704
入力部
705
出力部

[図8]


開始
匿名化対象レコード809から組QVを抽出
S601
抽出した組QVに対してクラスを定義
S602
匿名化対象レコード809が持つ組QVに基づいて、
RCM820を生成
S603
RCM820に基づいて、
匿名化対象レコード809のk－匿名性を評価
S604
k-匿名性の評価の結果に基づいて、
更新後RCM821を生成し、出力
S605
終了

[図9]


200 匿名化装置
202
RCM生成部
203
カレントステート生成部
205
再符号化部
206
匿名性検証部
207
カレントステート更新部
208
RCM更新部

[図10]

840 RCM

RID	CID
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1

[図11]

850 カレントステートテーブル

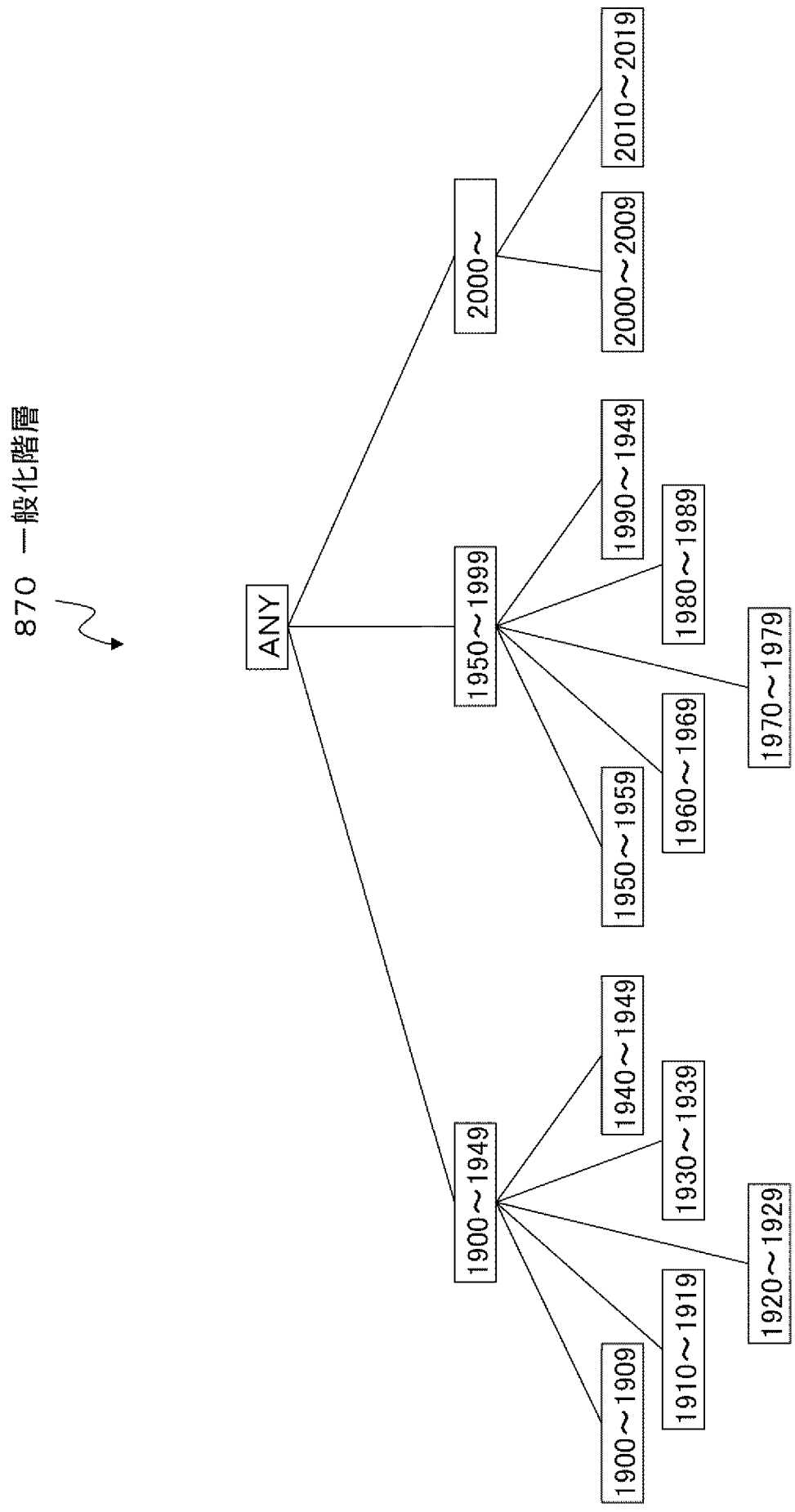
RID	CID	再符号化値
1	1	ANY
2	1	ANY
3	1	ANY
4	1	ANY
5	1	ANY
6	1	ANY
7	1	ANY
8	1	ANY
9	1	ANY
10	1	ANY

[図12]

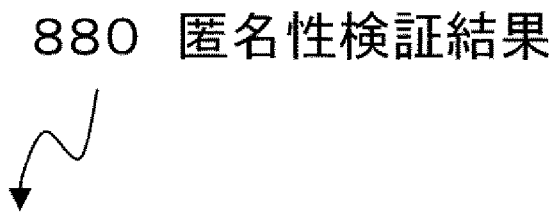
860 中間データ

RID	CID	再符号化値
1	1	1900～1949
2	1	1900～1949
3	1	1900～1949
4	1	1900～1949
5	1	1950～1999
6	1	1900～1949
7	1	1900～1949
8	1	1900～1949
9	1	1900～1949
10	1	1950～1999

[図13]


870 一般化階層
ANY
2000～
2010～2019
2000～2009
1950～1999
1990～1949
1980～1989
1970～1979
1960～1969
1950～1959
1900～1949
1940～1949
1930～1939
1920～1929
1910～1919
1900～1909

[図14]



CID	再符号化値	頻度
1	1900～1949	8
1	1950～1999	2

[図15]

852 更新後カレントステートテーブル

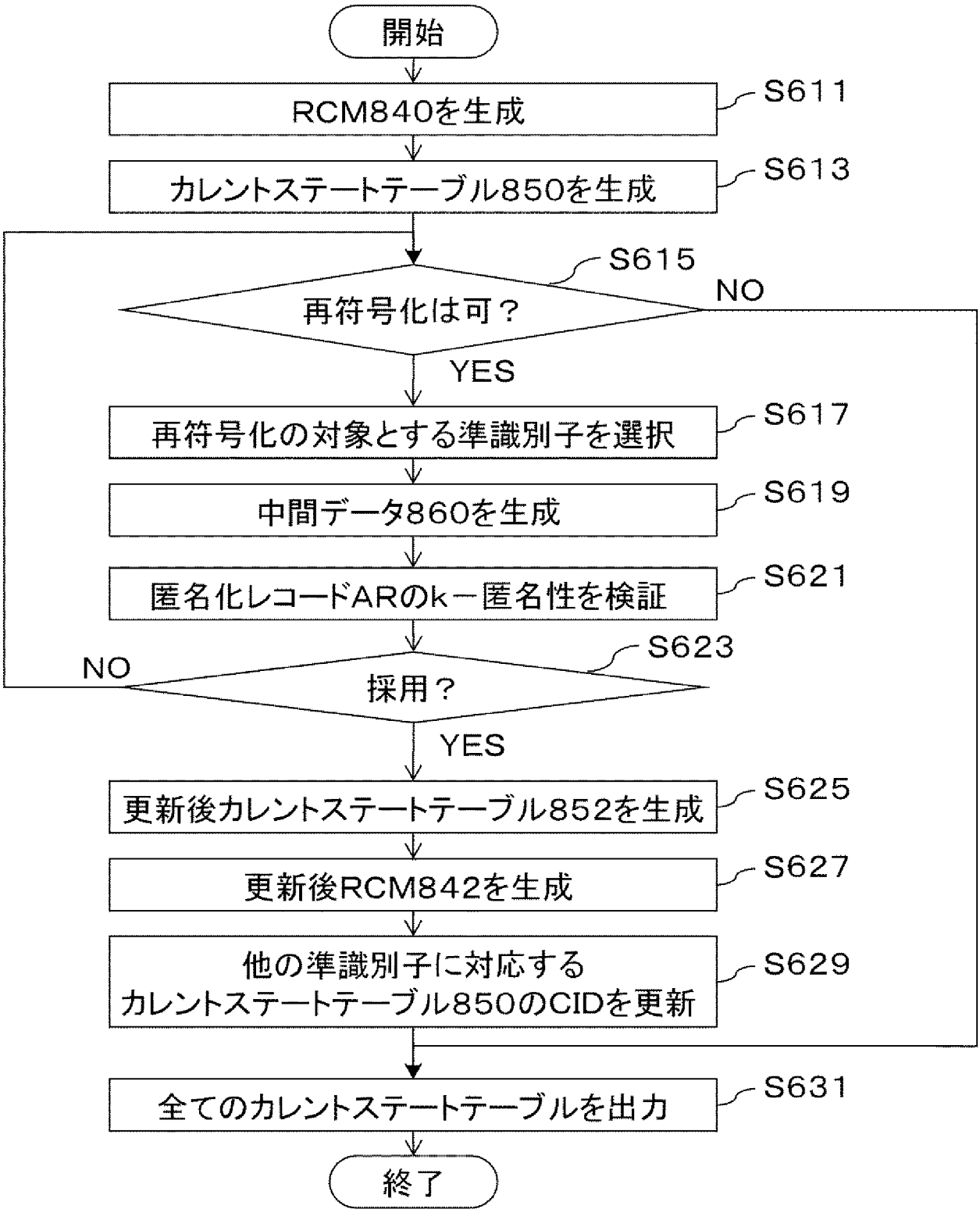
RID	CID	再符号化値
1	2	1900～1949
2	2	1900～1949
3	2	1900～1949
4	2	1900～1949
5	1	1950～1999
6	2	1900～1949
7	2	1900～1949
8	2	1900～1949
9	2	1900～1949
10	1	1950～1999

[図16]

842 更新後RCM

RID	CID
1	2
2	2
3	2
4	2
5	2
6	2
7	2
8	2
9	2

[図17]


開始
RCM840を生成
S611
カレントステートテーブル850を生成
S613
S615
再符号化は可？
NO
YES
再符号化の対象とする準識別子を選択
S617
中間データ860を生成
S619
匿名化レコードARのk－匿名性を検証
S621
S623
NO
採用？
YES
更新後カレントステートテーブル852を生成
S625
更新後RCM842を生成
S627
他の準識別子に対応する
カレントステートテーブル850のCIDを更新
S629
全てのカレントステートテーブルを出力
S631
終了

INTERNATIONAL SEARCH REPORT

International application No.
PCT / JP2 014/ 002439

A. CLASSIFICATION OF SUBJECT MATTER
G06F21/60(2013.01)i

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)
G06F21/60

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2014
Kokai Jitsuyo Shinan Koho 1971-2014 Toroku Jitsuyo Shinan Koho 1994-2014

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	Keiichi HIROTA et al., "Development and evaluation of personal data anonymization platform in information grand voyage project", IEICE Technical Report, vol.110, no.113, 24 June 2010 (24.06.2010), pages 297 to 308	1-10
A	Yu Juan et al., TopDown-KACA: an Efficient Local-recoding Algorithm for k-anonymity, IEEE International Conference on Granular Computing, 2009, GRC '09 [online], 2009.08.19, p.727-732, [retrieval date 2011.12.08], retrieved by utilizing IEEE Xplore	1-10
A	JP 2013-80375 A (Hitachi, Ltd.), 02 May 2013 (02.05.2013), entire text; all drawings (Family: none)	1-10

[X] 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
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 July, 2014 (10.07.14)	05 August, 2014 (05.08.14)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT / JP2 014/ 002439

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2013/031997 A1 (NEC Corp .), 07 March 2013 (07.03.2013), entire text ; all drawings (Family: none)	1-10

国際調査報告

国際出願番号 PCT／JP2014／002439

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G06F2 1/60 (2013. 01) i

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))

Int.Cl. G06F2 1/60

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－
日本国公開実用新案公報 1971－2
日本国実用新案登録公報 1996－
日本国登録実用新案公報 1994－2

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)

年

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	廣田 啓一 他，情報大航海プロジェクトにおける個人情報匿名化基盤の構築と検証，電子情報通信学会技術研究報告 Vol. 110 No. 113, 2010. 06. 24, pp. 297-308	1-10
A	Yu Juan、外3名, TopDown-KACA : an Efficient Local-recoding Algorithm for k-anonymity, IEEE international Conference on Granular Computing, 2009, GRC' 09 [online] , 2009. 08. 19, p. 727-732, [検索日 2011. 12. 08], IEEE Xplore より検索	1-10

☑ 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、一般的技術水準を示すもの
「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矛盾するものではなく、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	国際調査報告の発送日
10. 07. 2014	05. 08. 2014

国際調査機関の名称及びあて先	特許庁審査官 (権限のある職員)	
日本国特許庁 (ISA／JP) 郵便番号100－8915 東京都千代田区霞が関三丁目4番3号	岸野 徹 電話番号 03-3581－1101 内線 3546	5S 3983

様式PCT／ISA／210 (第2ページ) (2009年7月)

Ｃ（続き）．　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	JP 2013-80375 A (株式会社 日立製作所) 2013. 05. 02, 全文、全図 (ファミリーなし)	1-10
A	Wo 2013/031997 A1 (日本電気株式会社) 2013. 03. 07, 全文、全図 (ファミリーなし)	1-10