

(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 2 月 27 日(27.02.2014)



(10) 国際公開番号
WO 2014/030302 A1

- (51) 国際特許分類:
G06F 21/60 (2013.01) *G06Q 50/22* (2012.01)
G06F 17/30 (2006.01)
- (21) 国際出願番号: PCT/JP2013/004624
- (22) 国際出願日: 2013 年 7 月 31 日(31.07.2013)
- (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 2012-181684 2012 年 8 月 20 日(20.08.2012) 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: 豊田 由起(TOYODA, Yuki); 〒1088001 東
京都港区芝五丁目 7 番 1 号日本電気株式会社内
Tokyo (JP).
- (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 (JP).
- (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,
IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS,
LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH,
PL, PT, QA, RO, RS, RU, RW, 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,
MZ, 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 FOR EXECUTING ANONYMIZATION AND ANONYMIZATION PRO-
CESSING METHOD

(54) 発明の名称: 匿名化を実行する情報処理装置及び匿名化処理方法

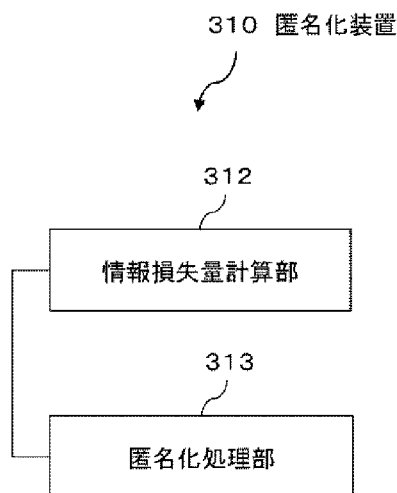


FIG. 1:
310 Anonymization device
312 Information loss amount calculation unit
313 Anonymization processing unit

(57) Abstract: This invention provides an information processing device for anonymizing a dataset so as to conform to a utilization purpose. The information processing device calculates the data loss amount corresponding to each attribute included in a first piece of personal data to be anonymized, decides on the attribute to be manipulated on the basis of priorities corresponding to the respective attributes and the data loss amount, and generates and outputs a second piece of personal data in which the attribute value of the decided attribute of the first piece of personal data has been manipulated.

(57) 要約: 本発明は、データセットを、利用目的に合致するようにして匿名化する情報処理装置を提供する。その情報処理装置は、匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、その属性のそれぞれに対応する優先度とその情報損失量とに基づいて加工の対象とするその属性を決定し、その第1の個人データのその決定した属性の属性値を加工した第2の個人データを生成し、出力する

明 細 書

発明の名称：匿名化を実行する情報処理装置及び匿名化処理方法 技術分野

[0001] 本発明は、個人データを匿名化する情報処理装置、匿名化処理方法及びプログラムに関する。

背景技術

[0002] 診療情報の電子化が進み、各医療機関や健康保険組合等においては診療情報が蓄積されつつある。これらの診療情報を広く活用することは、疫学研究の発展、医療技術及び新薬開発に繋がると考えられている。そこで、その蓄積された診療情報のプライバシーを確保した上で、その診療情報を研究機関等において利用できることが期待されている。

[0003] 情報の利用においてプライバシーを確保するための手法の一つに匿名化がある。この匿名化とは、例えば上述の診療情報のように、他人には知られたくない情報を含むデータに対して、個人の特定を防ぐための加工を施す技術である。以後、上述のような処理対象のデータの一塊をデータセットと呼ぶ。また、データセットを構成する、個人毎に対応するデータの塊を個人データレコードと呼ぶ。また、個人データレコードを構成する、個人の年齢、その個人が罹患した疾病名などの最小単位の情報を、属性と呼ぶ。

[0004] 非特許文献1は、匿名化の代表的な技術の一つであるk-匿名化を開示する。このk-匿名化は、データセットに含まれる個人データレコードのそれぞれを、個人の特定確率が $1/k$ （k-匿名化のk）以下となるように加工し、ある一定の匿名性を保証する技術である。k-匿名化における加工は、例えば、データセットを構成する複数の個人データレコード間で、特定の属性の値が同じになるように、その値を曖昧に（汎化とも呼ばれる）するなどの加工である。

[0005] 汎化によるk-匿名化には、トップダウンアプローチとボトムアップアプローチとがある。トップダウンアプローチは、最も汎化された状態の個人デ

ータレコードに含まれる属性の値を、 k -匿名性が破綻しない範囲で具体化していく手法である。また、ボトムアップアプローチは、加工されていない個人データレコードの元の値を、 k -匿名性が確保されるように汎化していく手法である。

[0006] 非特許文献2は、トップダウンアプローチの代表的な手法の1つを示す。非特許文献2に示される手法は、あるデータセットにおける k -匿名性を満たすために、そのデータセットの個人データレコードを、次に示すように加工することにより匿名化する手法である。

[0007] トップダウンアプローチでは、匿名化対象のデータセットにおいて、全ての個人データレコード間で、匿名化対象の全ての属性の値が属性毎に同じ値に汎化されている状態を初期状態とする。

[0008] 第1のステップは、その匿名化対象の属性の中から、ある属性を1つ選択する。

[0009] 第2のステップは、その全ての個人データレコードのそれぞれに含まれる、第1のステップで選択した属性の値の中央値を求める。

[0010] 第3のステップは、その求めた中央値を境界として、それらの個人データレコードのその属性の値に基づいて、その個人データレコードを2つのグループに分ける。

[0011] 第1のステップから第3のステップまでの処理をくりかえし、各グループの個人データレコードの数が k （ k -匿名化或いは k -匿名性の「 k 」、以後同様）個を満たさなくなった時点で処理を終了とする。このように、最も汎化された初期状態から、ある値を境界として属性の値をグループ化することを分割という。尚、結果として出力されるグループは、それらのグループのそれぞれの個人データレコードの数が k 個を満たさなくなる直前の、グループである。

[0012] また、非特許文献3は、ボトムアップアプローチの代表的な手法の1つを示す。非特許文献3に示される手法は、あるデータセットにおいて、個人データが k -匿名性を満たすように、ある属性の値を元の値から汎化していくこ

とにより匿名化する手法である。

- [0013] 特許文献1は、k-匿名化を取り入れたデータ匿名化装置を開示する。特許文献1のデータ匿名化装置は、データセットを構成する全ての個人データレコードを辺で結んだ完全グラフを生成し、この完全グラフをクラスタに分割し、分割したクラスタ単位で属性を汎化する。こうして、そのデータ匿名化装置は、トップダウンアプローチによるk-匿名化を実現する。

先行技術文献

特許文献

- [0014] 特許文献1：特開2012-022315号公報

非特許文献

- [0015] 非特許文献1：L. Sweeney, "Achieving k-anonymity privacy protection using generalization and suppression", International Journal on Uncertainty, Fuzziness and Knowledge based Systems, 10(5), pp.571-588, 2002.
非特許文献2：K, LeFevre, "Mondrian Multidimensional k-anonymity", ICD E '06 Proceedings of the 22nd International Conference on Data Engineering
非特許文献3：J, Xu, "Utility-Based Anonymization Using Local Recoding", KDD' 06 Proceeding of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining

発明の概要

発明が解決しようとする課題

- [0016] 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、データセットを、利用目的に合致するようにして匿名化することができないという問題点がある。
- [0017] その理由は、上述した特許文献及び非特許文献が開示する匿名化技術では、加工対象の属性が、利用目的を考慮されることなく、利用目的とは関係のない順番で選択され、加工されるからである。

[発明の目的]

本発明の目的は、上述した問題点を解決する情報処理装置、匿名化処理方法、及びプログラムを提供することにある。

課題を解決するための手段

[0018] 本発明の情報処理装置は、匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、出力する情報損失量計算手段と、前記属性のそれぞれに対応する優先度と前記情報損失量とに基づいて加工の対象とする前記属性を決定し、前記第1の個人データの前記決定した属性の属性値を加工した第2の個人データを生成し、出力する匿名化処理手段と、を含む。

[0019] 本発明の匿名化処理方法は、コンピュータが、匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、出力し、前記属性のそれぞれに対応する優先度と前記情報損失量とに基づいて加工の対象とする前記属性を決定し、前記第1の個人データの前記決定した属性の属性値を加工した第2の個人データを生成し、出力する。

[0020] 本発明の不揮発性記録媒体は、匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、出力する処理と、前記属性のそれぞれに対応する優先度と前記情報損失量とに基づいて加工の対象とする前記属性を決定する処理と、前記第1の個人データの前記決定した属性の属性値を加工した第2の個人データを生成し、出力する処理と、をコンピュータに実行させるプログラムを記録する。

発明の効果

[0021] 本発明は、データセットを、利用目的に合致するようにして匿名化することが可能になるという効果がある。

図面の簡単な説明

[0022] [図1]図1は、第1の実施形態に係る匿名化装置の構成を示すブロック図である。

[図2]図2は、第1乃至3の実施形態に係る匿名化装置を含むシステムの構成

を示すブロック図である。

[図3]図3は、第1及び2の実施形態における個人データの一例を示す図である。

[図4]図4は、第1及び2の実施形態における匿名化済個人データの一例を示す図である。

[図5]図5は、第1乃至3の実施形態に係る匿名化装置を実現するコンピュータのハードウェア構成を示すブロック図である。

[図6]図6は、第1の実施形態における匿名化装置の動作を示すフローチャートである。

[図7]図7は、第2の実施形態に係る匿名化装置の構成を示すブロック図である。

[図8]図8は、第2の実施形態における優先度決定情報の一例を示す図である。

[図9A]図9Aは、第2の実施形態における匿名化装置の動作を示すシーケンス図である。

[図9B]図9Bは、第2の実施形態における匿名化装置の動作を示すシーケンス図である。

[図10A]図10Aは、第2の実施形態における匿名化装置の動作を示すシーケンス図である。

[図10B]図10Bは、第2の実施形態における匿名化装置の動作を示すシーケンス図である。

[図11]図11は、第2の実施形態における最大に汎化された状態のデータセット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分割値候補の例を示す図である。

[図13]図13は、第2の実施形態における分割値候補の例を示す図である。
分割値候補の例を示す図である。

[図14]図14は、データセットが分割されるイメージを示す図である。

[図15]図15は、第3の実施形態に係る匿名化装置の構成を示すブロック図

である。

[図16]図 1 6 は、第 3 の実施形態における優先度決定情報の一例を示す図である。

[図17]図 1 7 は、第 3 の実施形態における個人データの一例を示す図である。

[図18A]図 1 8 A は、第 3 の実施形態における匿名化装置の動作を示すシーケンス図である。

[図18B]図 1 8 B は、第 3 の実施形態における匿名化装置の動作を示すシーケンス図である。

[図18C]図 1 8 C は、第 3 の実施形態における匿名化装置の動作を示すシーケンス図である。

[図19]図 1 9 は、第 3 の実施形態における匿名化装置による匿名化処理の途中段階の個人データの例を示す図である。

[図20]図 2 0 は、第 3 の実施形態における匿名化装置による匿名化処理の途中段階の個人データの例を示す図である。

[図21]図 2 1 は、第 3 の実施形態における匿名化装置による匿名化処理の途中段階の個人データの例を示す図である。

発明を実施するための形態

[0023] 本発明を実施するための形態につい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、同様の機能を備える構成要素には同様の符号が与えられている。

[0024] <<<第 1 の実施形態>>>

図 1 は、本発明の第 1 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）3 1 0 の構成を示すブロック図である。

[0025] 図 1 に示すように、本実施形態の匿名化装置 3 1 0 は、情報損失量計算部 3 1 2 及び匿名化処理部 3 1 3 を含む。図 1 に示す構成要素は、ハードウェ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。

[0026] 図2は、本発明の第1の実施形態に係る匿名化装置310を含むシステムの構成を示すブロック図である。

[0027] 図2に示すように、そのシステムは、個人データ記憶装置100と、匿名化済個人データ記憶装置200と、匿名化装置310とから構成されている。

[0028] 個人データ記憶装置100は、匿名化対象の個人データ（第1の個人データ）である、データセット（以後、データセットspと呼ぶ）を記憶する。尚、データセットspは、データレコード（以後、データレコードrpと呼ぶ）の集合である。そして、データレコードrpは、特定の個人に対応する、複数の属性の属性値を含む。

[0029] 個人データは、例えば、医療機関などが保有する診療情報である。この場合、データレコードrpに含まれる属性値は、氏名、生年、診療年月、病名などの属性の属性値である。

[0030] 図3は、個人データ記憶装置100に記憶される個人データであるデータセットsp110の一例を示す図である。データセットsp110は、複数のデータレコードrp111を含む。

[0031] データレコードrp111は、「氏名」、「生年」、「診療年月」及び「病名」の属性のそれぞれの属性値を含む。ここで、「氏名」の属性は、識別子である。「生年」及び「診療年月」の属性は、組み合わせると個人を特定する恐れがある準識別子である。また、「病名」の属性は、他人には知られたくないセンシティブ情報である。尚、これらの準識別子或いはセンシティブ情報とした属性は、一例である。即ち、匿名化装置310において、データセットsp110に含まれる属性の内、いずれの属性を準識別子或いはセンシティブ情報として扱うかは、任意である。

[0032] 図3に示すように、例えば、氏名の属性が「patient A」のデータレコードrp111は、生年の属性として「1949」、診療年月の属性として「201006」、傷病名の属性として「Disease A」の属性値

を含む。図3に示す個人データのデータセットs p 1 1 0は、一例であって、データレコードr p 1 1 1に、準識別子及びセンシティブ情報のそれぞれとして任意の属性の属性値を含んでよい。

[0033] 匿名化済個人データ記憶装置200は、匿名化済み個人データ（第2の個人データ）である匿名化済データセット（以後、匿名化済データセットs aと呼ぶ）を記憶する。尚、匿名化済データセットs aは、データレコードr p 1 1 1が匿名化された匿名化済データレコード（以後、匿名化済データレコードr aと呼ぶ）の集合である。

[0034] 図4は、匿名化済個人データ記憶装置200に記憶される匿名化済個人データである匿名化済データセットs a 2 1 0の一例を示す図である。匿名化済データセットs a 2 1 0は、匿名化装置310によりデータセットs p 1 1 0の属性値が加工（例えば、汎化）されて匿名化された後の、データセットである。即ち、匿名化済データセットs a 2 1 0は、データレコードr p 1 1 1に替えて、データレコードr p 1 1 1が加工された匿名化済データレコードr a 2 1 1を含む。

[0035] 図4に示す匿名化済データセットs a 2 1 0は、「生年」「診療年月」が加工（汎化）され、データセットs p 1 1 0に比べて準識別子の属性値が曖昧化されている。

[0036] 情報損失量計算部312は、データセットs p 1 1 0における、属性のそれぞれに対応する情報損失量（以後、情報損失量I L Aと呼ぶ）を算出し、出力する。

[0037] ここで、情報損失量I L Aは、データセットs p 1 1 0に含まれるいずれかの属性を加工した場合に増加する、属性の情報の抽象性（以後、情報抽象性i aと呼ぶ）の量である。尚、情報抽象性i aは、属性の情報、即ちその属性の属性値、の抽象性である。

[0038] 情報損失量計算部312は、必要に応じて、以下に示す様々な方法を用いて情報損失量I L Aを算出してよい。

[0039] 例えば、第1の方法として、情報損失量計算部312は、ある属性につい

て、その属性の汎化後の属性値の範囲を、同じ属性の汎化前の属性値の範囲で除して1つのデータレコードの情報の損失量（以後、情報損失量 I L R と呼ぶ）を算出する。次に、情報損失量計算部 3 1 2 は、データレコードの個数分だけ情報損失量 I L R を加算し、情報損失量 I L A を算出する。

[0040] 具体的には、情報損失量計算部 3 1 2 は、図 3 に示すデータセット s p 1 1 0 が最大に汎化されとした場合の、匿名化対象の各属性の情報損失量 I L A を以下のように算出する。

[0041] この場合、汎化前と汎化後とのそれぞれの匿名化対象の属性の属性値の範囲は、同一である。従って、情報損失量計算部 3 1 2 は、1つの属性値に対する情報損失量 I L R を、常に「1」と算出する。

[0042] 次に、情報損失量計算部 3 1 2 は、データレコード r p 1 1 1 の個数（20 個）分の情報損失量 I L R を加算し、情報損失量 I L A として「20」を算出する。こうして、情報損失量計算部 3 1 2 は、データセット s p 1 1 0 のいずれの属性についても、情報損失量 I L A として「20」を算出する。

[0043] また、情報損失量計算部 3 1 2 は、図 3 に示すデータセット s p 1 1 0 の「生年」の属性の属性値が「1956」以下と「1961」以上とのデータレコード r p 1 1 1 のそれぞれに分割されて、汎化されとした場合の、匿名化対象の各属性の情報損失量 I L A を以下のように算出する。

[0044] この場合、汎化前の「生年」の属性の属性値の範囲は、最小値が「1943」であり、最大値が「1977」である。また、汎化後（汎化される「生年」の属性の属性値が「1956」以下のデータレコード r p 1 1 1）の「生年」の属性の属性値の範囲は、最小値が「1943」であり、最大値が「1956」である。

[0045] 従って、情報損失量計算部 3 1 2 は、「生年」の属性の属性値が「1956」以下の「生年」の属性の情報損失量 $I L R - b i r t h - u 1 9 5 6$ を、以下のように算出する。

[0046] 「 $(1956 - 1943) \div (1977 - 1943) = 0.382$ 」

次に、情報損失量計算部 3 1 2 は、「生年」の属性の属性値が「1956

」以下のデータレコード $rp111$ の個数（9個）分の情報損失量 $ILR-birth-ul1956$ を加算し、「生年」の属性の属性値が「1956」以下のデータレコード $rp111$ の情報損失量 $ILA-birth-ul1956$ として「3.438」を算出する。

[0047] また、汎化後（汎化される「生年」の属性の属性値が「1961」以上のデータレコード $rp111$ ）の「生年」の属性の属性値の範囲は、最小値が「1961」であり、最大値が「1977」である。従って、情報損失量計算部312は、「生年」の属性の属性値が「1961」以上の「生年」の属性の情報損失量 $ILR-birth-ov1961$ を、「 $(1977-1961) \div (1977-1943) = 0.471$ 」と算出する。

[0048] 次に、情報損失量計算部312は、「生年」の属性の属性値が「1961」以上のデータレコード $rp111$ の個数（11個）分の情報損失量 $ILR-birth-ov1961$ を加算し、「生年」の属性の属性値が「1961」以上のデータレコード $rp111$ の情報損失量 $ILA-birth-ov1961$ として「5.181」を算出する。

[0049] 次に、情報損失量計算部312は、情報損失量 $ILA-birth-ul1956$ と情報損失量 $ILA-birth-ov1961$ とを加算して、属性が「生年」の情報損失量 $ILA-birth$ として「8.619」を算出する。

[0050] 同様に、汎化前の「診療年月」の属性の属性値の範囲は、最小値が「200512」であり、最大値が「201107」である。また、汎化後（汎化される「生年」の属性の属性値が「1956」以下のデータレコード $rp111$ ）の「診療年月」の属性の属性値の範囲は、最小値が「200512」であり、最大値が「201107」である。従って、情報損失量計算部312は、「生年」の属性の属性値が「1956」以下のデータレコード $rp111$ の、「診療年月」の属性の情報損失量 $ILR-mc-ul1956$ を、「1」と算出する。

[0051] 次に、情報損失量計算部312は、「生年」の属性の属性値が「1956

」以下のデータレコード $r p 1 1 1$ の個数（9 個）分だけ情報損失量 $I L R - m c - u 1 9 5 6$ を加算し、「生年」の属性の属性値が「1 9 5 6」以下のデータレコード $r p 1 1 1$ の「診療年月」の属性の情報損失量 $I L A - m c - u 1 9 5 6$ として「9」を算出する。

[0052] また、汎化後（汎化される「生年」の属性の属性値が「1 9 6 1」以上のデータレコード $r p 1 1 1$ ）の「診療年月」の属性の属性値の範囲は、最小値が「2 0 0 6 1 2」であり、最大値が「2 0 1 1 0 7」である。従って、情報損失量計算部 3 1 2 は、「生年」の属性の属性値が「1 9 6 1」以上のデータレコード $r p 1 1 1$ の、「診療年月」の属性の情報損失量 $I L R - m c - o v 1 9 6 1$ を、「0. 8 3 2」と算出する。

[0053] 次に、情報損失量計算部 3 1 2 は、「生年」の属性の属性値が「1 9 6 1」以上のデータレコード $r p 1 1 1$ の個数（1 1 個）分だけ情報損失量 $I L R - m c - o v 1 9 6 1$ を加算し、「生年」の属性の属性値が「1 9 6 1」以上のデータレコード $r p 1 1 1$ の「診療年月」の属性の情報損失量 $I L A - m c - o v 1 9 6 1$ として「9. 1 5 2」を算出する。

[0054] 次に、情報損失量計算部 3 1 2 は、情報損失量 $I L A - m c - u 1 9 5 6$ と情報損失量 $I L A - m c - o v 1 9 6 1$ とを加算し、属性が「診療年月」の情報損失量 $I L A - m c$ として「1 8. 1 5 2」を算出する。

[0055] 以上が、第 1 の方法の説明である。

[0056] また、第 2 の方法として、情報損失量計算部 3 1 2 は、以下のようにして情報損失量 $I L A$ を算出してもよい。まず、情報損失量計算部 3 1 2 は、汎化後と汎化前とのそれぞれの、その属性の属性値の種類の数比を、1 つのデータレコードの情報損失量 $I L R$ として算出する。次に、情報損失量計算部 3 1 2 は、データレコードの個数分だけ情報損失量 $I L R$ を加算し、情報損失量 $I L A$ を算出する。

[0057] 匿名化処理部 3 1 3 は、図示しない手段（例えば、匿名化処理部 3 1 3 内の図示しない記憶手段）に記憶されている優先度決定情報に基づいて、属性のそれぞれの優先度（以後、優先度 p と呼ぶ）を決定する。また、匿名化処

理部 3 1 3 は、その優先度 p と情報損失量計算部 3 1 2 が算出した情報損失量 ILA とに基づいて、加工する属性を決定する。即ち、匿名化処理部 3 1 3 は、優先度 p を用いることで利用目的を考慮し、かつ情報損失量 ILA を用いることで匿名化済データセット $sa210$ 全体の情報の損失を低減するように、加工する属性を決定する。

[0058] ここで、優先度決定情報は、優先度 p を決定する情報である。優先度 p は、データセット $sp110$ （データレコード $rp111$ ）に含まれる属性のそれぞれが持っている情報抽象性 ia を増加させないようにする（情報の損失を優先的に防ぐ）度合いを示す情報である。即ち、優先度 p は、匿名化済データセット $sa210$ におけるデータセット $sp110$ に対する情報抽象性 ia の増加を、複数の属性のいずれについて、より少なくするように匿名化するかを優先度を示す。

[0059] 例えば、匿名化処理部 3 1 3 は、属性毎に優先度 p と情報損失量 ILA とを演算（例えば、乗算）した評価値を算出する。尚、匿名化処理部 3 1 3 は、図示しない手段から、特定の優先度 p と特定の情報損失量 ILA との組み合わせに対応する評価値を取得するようにしてもよい。尚、評価値を算出する演算は、情報損失量 ILA が一定ならば優先度 p が高いほど、評価値を大きく算出する演算である。また、評価値を算出する演算は、優先度 p が一定ならば情報損失量 ILA が大きいほど評価値を大きく算出する演算である。これは、特定の優先度 p と特定の情報損失量 ILA との組み合わせに対応する評価値を取得する場合も同様である。

[0060] 続けて、匿名化処理部 3 1 3 は、例えば、評価値が大きい属性ほど汎化されないように、評価値が小さい属性ほど汎化されるように、汎化する属性を決定する。

[0061] 尚、匿名化処理部 3 1 3 は、評価値が小さい属性ほど汎化されないように、評価値が大きい属性ほど汎化されるように、汎化する属性を決定するようにしてもよい。この場合、評価値を算出する演算は、情報損失量 ILA が一定ならば優先度 p が高いほど、及び優先度 p が一定ならば情報損失量 ILA

が大きいほど評価値を小さく算出する演算である。このような演算であることは、特定の優先度 p と特定の情報損失量 $IL A$ との組み合わせに対応する評価値を取得する場合の演算についても同様である。

[0062] 次に、匿名化処理部 313 は、データセット $s p 110$ のその決定した属性を加工した匿名化済データセット $s a 210$ を生成し、出力する。尚、匿名化処理部 313 は、データセット $s p 110$ に対する匿名化済データセット $s a 210$ の差分の情報を生成し、出力するようにしてもよい。

[0063] 尚、匿名化処理部 313 は、加工済みデータセットの匿名性を評価するようにしてもよい。ここで、加工済みデータセットは、それらの属性を加工した場合の、データセットの一部分及び全体のいずれかの任意のものである。続けて、匿名化処理部 313 は、その匿名性を評価した結果が所定の内容である場合に、その加工済みデータセットを、匿名化済みのデータセット一部分及び全体のいずれかの任意のものとして、匿名化済個人データ記憶装置 200 に記録するようにしてもよい。

[0064] 以上が、匿名化装置 310 のコンピュータ装置の機能単位に分割した各構成要素についての説明である。

[0065] 次に、匿名化装置 310 のハードウェア単位の構成要素について説明する。

[0066] 図 5 は、本実施形態における匿名化装置 310 を実現するコンピュータ 700 のハードウェア構成を示す図である。

[0067] 図 5 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。

[0068] CPU 701 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、CPU 701 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 6 に示すフローチャ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。

[0069] そして、CPU 701 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 に示す情報損失量計算部 312 及び匿名化処理部 313 として各種の処理を実行する。

[0070] 尚、CPU 701 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。

[0071] 記憶部 702 は、プログラムやデータを記憶する。記憶部 702 は、個人データ記憶装置 100 及び匿名化済個人データ記憶装置 200 を含んでもよい。

[0072] 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 は、プログラムをコンピュータ読み取り可能に記録する。また、記憶装置 703 は、データをコンピュータ読み取り可能に記録してもよい。記憶装置 703 は、個人データ記憶装置 100 及び匿名化済個人データ記憶装置 200 を含んでもよい。

[0073] 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、例えばタッチパネル、加速度計、ジャイロセンサ、カメラなどでもよい。

[0074] 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。

[0075] 通信部 706 は、個人データ記憶装置 100 や匿名化済個人データ記憶装置 200 及び他の図示しない外部装置とのインタフェースを実現する。通信部 706 は、匿名化処理部 313 の一部として含まれる。

[0076] 以上説明したように、図1に示す匿名化装置310の機能単位のブロックは、図5に示すハードウェ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。

[0077] 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（ソフトウェア）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。

[0078] 以上が、本実施形態における匿名化装置310を実現するコンピュータ700の、ハードウェア単位の各構成要素についての説明である。

[0079] 次に、図1～図6を参照して本実施形態の動作について詳細に説明する。

[0080] 図6は、本実施形態における匿名化装置310の動作を示すフローチャートである。

[0081] 情報損失量計算部312は、データセットsp110の匿名化対象の属性のそれぞれについて、情報損失量ILAを算出する（ステップS601）。

[0082] 次に、匿名化処理部313は、優先度pを決定する情報に基づいて、属性のそれぞれの優先度pを決定する（ステップS602）。

[0083] 次に、匿名化処理部313は、情報損失量ILAと優先度pとに基づいて、加工する属性を決定する（ステップS603）。

[0084] 次に、匿名化処理部313は、データレコードrp111の、決定した属性を加工する（ステップS604）。

[0085] 次に、匿名化処理部313は、属性を加工したデータレコードrp111

を出力する（ステップS605）。

[0086] 上述した本実施形態における第1の効果は、利用目的に合致するように制御して、データセットを匿名化することが可能になる点である。

[0087] 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、情報損失量計算部312が属性のそれぞれに対応する情報損失量ILAを算出し、出力する。第2に、匿名化处理部313が優先度pと情報損失量ILAとに基づいて加工する属性を決定し、決定した属性を加工する。

[0088] 上述した本実施形態における第2の効果は、第1の効果に加えて匿名化されたデータセットにおける情報の損失を低減することが可能になる点である。即ち、この第2の効果は、利用目的に合致するように制御してデータセットを匿名化することと、匿名化されたデータセットにおける情報の損失を低減することとが両立して可能になる点である。これは、利用目的に合致することだけを考慮して匿名化がなされることにより、加工を抑制した属性以外の属性が過度に汎化され、データ全体として大きく情報が損失することを防ぐことが可能になるということである。

[0089] その理由は、第1の効果と同様である。即ち、匿名化处理部313が優先度pと情報損失量ILAとの両方に基づいて加工する属性を決定するようにしたからである。

<<<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。

[0090] 図7は、本実施形態に係る匿名化装置320の構成を示すブロック図である。本実施形態の匿名化装置320は、トップダウンアプローチにより匿名化を行う。

[0091] 図7に示すように、匿名化装置320は、優先度決定情報記憶部321と情報損失量計算部322と匿名化处理部323とを含む。

[0092] 尚、匿名化装置320は、匿名化装置310に替えて、図2に示すシステ

ムに含まれてもよい。

[0093] 優先度決定情報記憶部 3 2 1 は、優先度 p を決定する情報を記憶する。尚、優先度 p を決定する情報は、システムの利用者により予め設定されている。また、優先度 p を決定する情報は、分割属性決定部 3 2 3 3 が図 5 に示す通信部 7 0 6 により、予め外部のシステムから受信するようにしてもよい。

[0094] 図 8 は、優先度決定情報記憶部 3 2 1 に記憶される優先度決定情報 3 2 1 0 の一例を示す図である。図 8 に示すように、優先度決定情報 3 2 1 0 は、インデックスと重み（優先度とも呼ばれる）との組を含む。ここで、そのインデックスは、重みを一意に決定する値である。その重みは、そのインデックスのそれぞれに対応し、属性の重要さを示す数字である。図 8 において、例えば、そのインデックスの「5」に対応する重みは「16」である。

[0095] 尚、図 8 の例に係わらず、そのインデックスは 5 種類に限らず、2 種類以上の任意の数であってよい。また、インデックスは数字に限らず、アルファベット等の表記であってよいし、属性の名前（以後、属性名とも呼ぶ）であってもよい。

[0096] また、その重みは、後述する評価値の算出に使用可能な、任意の数値であってよい。

[0097] また、優先度決定情報記憶部 3 2 1 は、そのインデックスの入力に対して重みを計算する計算式（例えば、「 $\text{重み} = 2 \times (\text{インデックス} - 1)$ 」）を、優先度決定情報として記憶してもよい。

[0098] 情報損失量計算部 3 2 2 は、データセット $s p 1 1 0$ における各属性の情報損失量 $I L A$ を計算し、出力する。

[0099] 匿名化処理部 3 2 3 は、分割属性決定部 3 2 3 3 と分割値決定部 3 2 3 4 と匿名性評価部 3 2 3 5 と汎化実行部 3 2 3 6 とを含む。

[0100] 分割属性決定部 3 2 3 3 は、優先度決定情報記憶部 3 2 1 に記憶された優先度決定情報 3 2 1 0 を利用し、例えば、図 5 に示す入力部 7 0 4 から入力された各属性のインデックスに基づいて、各属性の重みを生成する。

[0101] 次に、分割属性決定部 3 2 3 3 は、その生成した重みと情報損失量 $I L A$

とに基づいて、分割軸の属性（加工の対象とする属性とも呼ばれる、以後、属性と呼ぶ）を決定する。

- [0102] 分割属性は、データセット（例えば、データセット s p 1 1 0）を分割する場合に、その分割属性の属性値を基準として分割する属性のことである。ここで、データセットを分割することは、そのデータセットに含まれるデータレコードをグループ分けすることである。即ち、分割属性決定部 3 2 3 3 は、データセット（例えば、データセット s p 1 1 0）を分割する場合に、その分割属性の属性値の範囲を基準としてその分割を行う。その範囲は、例えば、ある値より大きい値と小さい値とである。或いは、その範囲は、地理的な領域や、物事の種類や、事象との関連性やであってよい。
- [0103] 分割値決定部 3 2 3 4 は、必要な匿名性を満たすように、その分割属性の分割値を決定する。その分割値は、例えば、属性値が数値で示される場合、属性値が取り得る範囲内の数値である。或いは、その範囲は、属性値が地理的な領域である場合、その領域を示す識別情報（例えば、県名）の集合であってよい。また、その範囲は、属性値が物事（例えば、趣味）の種類である場合、その種類を分類する識別情報（例えば、屋外で行うもの）であってよい。また、その範囲は、属性値が事象との関連性である場合、関連性の有無であってよい。
- [0104] 匿名性評価部 3 2 3 5 は、あるデータセットが分割された場合にその分割されたデータセットのそれぞれが、必要な匿名性を満たすか否かを判定する。具体的には、匿名性評価部 3 2 3 5 は、例えば、あるデータセットが、2つのグループに分割された場合、その2つのグループのそれぞれが、少なくとも k 個のデータレコード r p 1 1 1 を含むように、そのデータセットを分割できるか否かを判定する。ここで、その k 個の「k」は、k-匿名性或いは k-匿名化の「k」である。以後の k 個についても、同様である。
- [0105] 汎化実行部 3 2 3 6 は、決定された分割値に基づいて、決定された属性の属性値を汎化（加工）し、出力する。
- [0106] 以上説明した匿名化装置 3 2 0 は、図 1 に示す匿名化装置 3 1 0 と同様に

、図5に示すコンピュータ700によって実現してもよい。

[0107] 次に本実施形態の動作について、図面を参照して詳細に説明する。

[0108] 図9A、図9B、図10A及び図10Bは、本実施形態の動作を示すシーケンス図である。

[0109] 図9Aにおいて、分割属性決定部3233は、例えば図5に示す入力部704からのシステムの利用者による分割属性決定要求の入力を、受け付ける（ステップS801）。

[0110] ここで、分割属性決定要求は、例えば、k-匿名性のkの値「5」と、属性名及び対応するインデックス「生年：4、診療年月：1」とを含む。

[0111] 尚、匿名化済データセットを利用するその利用者は、汎化(加工)の度合を抑制したい属性ほど大きいインデックスの値を指定する。

[0112] 次に、分割属性決定部3233は、例えば図5に示す記憶部702に、その受け付けた分割属性決定要求に含まれるkの値「5」と属性名及び対応するインデックス「生年：4、診療年月：1」とを記憶する（ステップS802）。

[0113] 次に、分割属性決定部3233は、優先度決定情報3210を利用し、属性名及び対応するインデックス「生年：4、診療年月：1」に基づいて重みを生成する（ステップS803）。

[0114] ここでは、分割属性決定部3233は、「生年」の属性に対応するその重みを「8」、「診療年月」の属性に対応するその重みを「1」と算出する。

[0115] 次に、分割属性決定部3233は情報損失量計算部322へ情報損失量ILAの計算要求を送信する（ステップS804）

次に、その情報損失量ILAの計算要求を受信した情報損失量計算部322は、個人データ記憶装置100へデータセットsp110の取得要求（以後、個人データ取得要求とも呼ぶ）を送信する（ステップS805）。

[0116] 次に、データセットsp110を受信した情報損失量計算部322は、情報損失量ILAを計算し、計算した情報損失量ILAを分割属性決定部3233へ送信する（ステップS806）。

[0117] ここで、情報損失量計算部322の情報損失量 IL_A の計算の動作について、詳細に説明する。

[0118] 情報損失量計算部322は、例えば以下に示す式1を用いて、1つのデータレコード r_{p111} の情報損失量 IL_R を算出する。

[0119] [数1]

$$IL_1 = \frac{pt_{a-max} - pt_{a-min}}{pt_{b-max} - pt_{b-min}} \quad \cdot \cdot \cdot \quad (式1)$$

[0120] ここで、 pt_{a-max} は、汎化後の属性値の最大値である。また、 pt_{a-min} は、汎化後の属性値の最小値である。また、 pt_{b-max} は、汎化前の属性値の最大値である。また、 pt_{b-min} は、汎化前の属性値の最小値である。

[0121] 本実施形態はトップダウンアプローチを用いた匿名化の実施形態であるため、データセット s_{p110} の匿名化対象の属性の属性値は、全て同一の値になるように汎化されるものとする。

[0122] 図11は、図3に示すデータセット s_{p110} の匿名化の対象である属性のそれぞれの属性値が同一の値に汎化された場合のデータセット s_{t120} を示す図である。即ち、図11に示すデータセット s_{t120} は、データセット s_{p110} が最大に汎化された状態のデータセットである。

[0123] この場合、 pt_{a-max} は、例えば、図11に示すデータセット s_{t120} において、属性名が「生年」の属性の属性値の最大値である「1977」（「1943～1977」の「1977」）である。また、 pt_{a-min} は、例えば、データセット s_{t120} において、属性名が「生年」の属性の、属性値の最小値である「1943」（「1943～1977」の「1943」）である。また、 pt_{b-max} は、例えば、図3に示すデータセット s_{p110} において、属性名が「生年」の属性の、属性値の最大値である「1977」である。また、 pt_{b-min} は、例えば、データセット s_{p110} において、属性名が「生年」の属性の、属性値の最小値である「19

43」である。

[0124] 従って、属性名が「生年」の属性の、1つのデータレコードrp111の情報損失量ILR-birthは、以下のように「1」が算出される。

[0125] 「情報損失量ILR-birth」= (1977-1973) ÷ (1977-1943) = 1

また、データセットsp110に含まれるデータレコードrp111の数は、20である。従って、属性名が「生年」の属性の情報損失量ILA-birthは、以下のように「1」が算出される。

[0126] (「情報損失量ILR-birth」) × (データレコードrp111の数) = 1 × 20 = 20

同様に、診療年月の全体の情報損失量ILA-mcは、「20」が算出される。

[0127] 尚、情報損失量計算部322は、汎化後と汎化前とのそれぞれのその属性の属性値の種類の数の比を、1つのデータレコードrp111の情報損失量ILRとして、算出するようにしてもよい。

[0128] 以上が、情報損失量計算部322の情報損失量ILAの計算の動作についての詳細な説明である。

[0129] 図9Aの説明に戻る。尚、上述の説明において、ステップS803と、ステップS804、ステップS805及びステップS806との処理順序は、任意の順序であってよい。即ち、その順序は、逆であってもよいし、同時であってもよい。

[0130] 次に、分割属性決定部3233は、分割属性を決定する（ステップS807）。

[0131] ここで、分割属性決定部3233による、分割属性の決定の動作について、詳細に説明する。

[0132] 分割属性決定部3233は、重みと情報損失量ILAとを含んだ評価式を用いて評価値を算出し、分割属性を決定する。以下に示す式2は、評価式の一例である。

[0133] [数 2]

評価値 = 重み × 情報損失量 I L A . . . (式 2)

例えば、データセット s p 1 1 0 の属性名が「生年」の属性の評価値は、重みが「8」、情報損失量 I L A - b i r t h が「20」なので、「160」である。同様に、属性名が「診療年月」の属性の評価値は、重みが「1」、情報損失量 I L A - m c が「20」なので、「20」である。

[0134] 次に、分割属性決定部 3 2 3 3 は、算出した評価値が最大の属性を分割属性として決定する。例えばデータセット s p 1 1 0 の場合、分割属性決定部 3 2 3 3 は、属性名が「生年」の属性の評価値が、属性名が「診療年月」の属性の評価値よりも大きいので、属性名が「生年」の属性をその分割属性として決定する。

[0135] 尚、評価値を算出する式は、式 2 に限らず、優先度 p（例えば、式 2 の「重み」のように、大きいほど優先度が高いことを示す値）が高いほど及び情報損失量 I L A が大きいほど演算結果が大きくなるような、任意の評価式でよい。

[0136] 以上が、分割属性決定部 3 2 3 3 による分割属性の決定の動作についての説明である。

[0137] 次に、図 9 B において、分割属性決定部 3 2 3 3 は、分割値決定部 3 2 3 4 へ分割値決定要求を送信する（ステップ S 8 0 8）。その分割値決定要求は、分割属性決定部 3 2 3 3 により決定された分割属性の属性名の「生年」を含む。

[0138] その分割値決定要求を受信した分割値決定部 3 2 3 4 は、個人データ記憶装置 1 0 0 へ個人データ取得要求を送信する。（ステップ S 8 0 9）

データセット s p 1 1 0 を受信した分割値決定部 3 2 3 4 は、分割値を決定する（ステップ S 8 1 0）。

[0139] ここで、分割値決定部 3 2 3 4 による分割値の決定の動作について、詳細に説明する。

[0140] その分割値は、指定された属性を分割軸としてデータセットを分割する時

の、閾値である。例えば、分割値「生年：１９５６」は、「生年」の属性が「１９５６」以下のデータレコード $r p 1 1 1$ と「１９５６」を超えるデータレコード $r p 1 1 1$ とに、データセット $s p 1 1 0$ を分割することを示す。

[0141] 図１２は、データセット $s p 1 1 0$ の分割値候補 $1 1 0 1 \sim 1 1 1 1$ の例を示す図である。

[0142] まず、分割値決定部 3 2 3 4 は、図１２に示すように、データセット $s p 1 1 0$ のデータレコード $r p 1 1 1$ を、分割属性決定部 3 2 3 3 が決定した属性を、属性値が小さい順番に並べる。

[0143] 次に、分割値決定部 3 2 3 4 は、分割値候補 $1 1 0 1 \sim 1 1 1 1$ を抽出する。分割値決定部 3 2 3 4 が抽出する分割値候補 $1 1 0 1 \sim 1 1 1 1$ は、分割されたデータセット $s p 1 1 0$ の前半部分（第３の個人データとも呼ばれる）と後半部分（第４の個人データとも呼ばれる）とのそれぞれのデータレコード $r p 1 1 1$ の数が k 個以上になる分割値の候補である。例えば、「生年」の属性において属性値の「１９５１」を分割値とすると、その前半部分は、その属性値が １ ９ ５ １ 以下の ５ 個のデータレコード $r p 1 1 1$ を含む。また、その後半部分は、１ ９ ５ ２ 以上の １ ５ 個のデータレコード $r p 1 1 1$ を含む。この場合、その前半部分とその後半部分とのそれぞれは、いずれも ５ 個以上である。

[0144] 図１２に示すデータセット $s p 1 1 0$ において、分割値決定部 3 2 3 4 は、分割値候補 $1 1 0 1 \sim 1 1 1 1$ を抽出する。

[0145] 次に、分割値決定部 3 2 3 4 は、各分割値候補 $1 1 0 1 \sim 1 1 1 1$ に対応する情報損失量 $I L A$ を計算する。例えば、分割値決定部 3 2 3 4 は、式 1 を用いて情報損失量 $I L A$ を算出する。尚、分割値決定部 3 2 3 4 は、式 1 に限らず、他の算出式を用いて情報損失量 $I L A$ を算出してもよい。

[0146] 具体的には、例えば分割値候補 $1 1 0 5$ でデータセット $s p 1 1 0$ を分割する場合、分割値決定部 3 2 3 4 は、以下の様に情報損失量 $I L A$ を計算する。

- [0147] 分割値決定部3234は、図12に示すように、分割属性の「生年」の属性値により、昇順でデータセットsp110をソートする。分割値候補1105の分割値でデータセットsp110を分割した場合、その分割された前半部分のデータレコードrp111の一つの情報損失量ILRは、 $(1956 - 1943) \div (1977 - 1943) = 0.382$ である。
- [0148] 従って、その前半部分の情報損失量ILRの合計は、データレコードrp111の数が9個であるので、 $0.382 \times 9 = 3.438$ である。
- [0149] また、その分割された後半部分のデータレコードrp111の1つの情報損失量ILRは、 $(1977 - 1961) \div (1977 - 1943) = 0.471$ である。
- [0150] 従って、その後半部分の情報損失量ILRの合計は、データレコードrp111の数が11個であるので、 $0.471 \times 11 = 5.181$ である。
- [0151] 従って、分割値候補1105で分割された場合の合計の情報損失量ILAは、 $3.438 + 5.181 = 8.619$ である。
- [0152] 同様にして算出される、分割値候補1101～1104で分割された場合の、「生年」の属性のそれぞれの情報損失量ILAは、「11.76」、「12.47」、「10.67」及び「10.23」である。また、同様にして算出される、分割値候補1106～1111で分割された場合の、「生年」の属性のそれぞれの情報損失量ILAは、「10.00」、「10.05」、「9.88」、「10.14」、「10.70」及び「10.73」である。
- [0153] 分割値候補1101～1111のそれぞれに対する情報損失量ILAを計算した分割値決定部3234は、その情報損失量ILAが最小である分割値候補1105の「生年：1956」を分割値として決定する。
- [0154] 以上が、分割値決定部3234による分割値の決定の動作についての説明である。
- [0155] 図9Bの説明に戻る。次に、その分割値を決定した分割値決定部3234は、その決定した分割値「生年：1956」を匿名性評価部3235へ送信

する（ステップS 8 1 1）。換言すると、分割値決定部3 2 3 4は、匿名性評価部3 2 3 5に、匿名性の評価を要求する。

[0156] その分割値「生年：1 9 5 6」を受信した匿名性評価部3 2 3 5は、匿名性評価を行う（ステップS 8 1 2）。

[0157] ここで、匿名性評価部3 2 3 5による匿名性評価の動作について、詳細に説明する。

[0158] 匿名性評価とは、匿名性の指標を満たすか否かを評価することである。匿名性評価部3 2 3 5は、データセットs p 1 1 0の分割されたその前半部分（第3の個人データ）とその後半部分（第4の個人データ）とについて更に分割を行った場合に、その更に分割された部分が匿名性の指標を満たすか否かを評価する。即ち、その前半部分及びその後半部分のそれぞれについて、データレコードr p 1 1 1の数が2 k 個以上か否かを評価する。

[0159] 匿名性評価部3 2 3 5は、受信した分割値で分割されたその前半部分及びその後半部分のそれぞれのデータレコードr p 1 1 1の数を計数する。例えば、その分割値「生年：1 9 5 6」で分割した場合、匿名性評価部3 2 3 5は、その前半部分のデータレコードr p 1 1 1の数を9 個、その後半部分のデータレコードr p 1 1 1の数を1 1 個と計数する。

[0160] 以上が、匿名性評価部3 2 3 5による匿名性評価の動作についての説明である。

[0161] 図9 Bの説明に戻る。次に、匿名化処理部3 2 3 は、匿名性の指標を満たさないと匿名性評価部3 2 3 5が評価した部分（例えば、その分割値「生年：1 9 5 6」で分割された前半部分）について、ステップS 8 1 3からステップS 8 1 5の処理を実行する。また、匿名化処理部3 2 3 は、匿名性の指標を満たすと匿名性評価部3 2 3 5が評価した部分（例えば、その分割値「生年：1 9 5 6」で分割された後半部分）について、ステップS 8 2 1以降の処理を実行する。

[0162] その分割値「生年：1 9 5 6」で分割された前半部分のデータレコードr p 1 1 1の数は2 k 個未満であった。そこで、匿名性評価部3 2 3 5は、「

生年：１９４３～１９５６」を含む汎化実行要求を汎化実行部３２３６へ送信する（ステップＳ８１３）

その汎化実行要求を受信した汎化実行部３２３６は、「生年」の属性の属性値が「１９４３」～「１９５６」のデータレコードｒｐ１１１を汎化する（ステップＳ８１４）。

[0163] 具体的には、汎化実行部３２３６は、「生年」の属性の属性値が「１９４３」～「１９５６」に該当するデータレコードｒｐ１１１の、「生年」の属性の属性値を「１９４３～１９５６」に、「診療年月」の属性の属性値を「２００５１２～２０１１０７」に書き換える。

[0164] 次に、汎化実行部３２３６は、書き換えたデータレコードｒｐ１１１を匿名化済個人データ記憶装置２００へ記録する（ステップＳ８１５）。換言すると、汎化実行部３２３６は、匿名化済個人データ記憶装置２００へ匿名化済み個人データを登録する。

[0165] その分割値の「生年：１９５６」で分割された後半部分のデータレコードｒｐ１１１の数は２ｋ個以上であった。そこで、匿名化装置３２０は、そのデータセットｓｐ１１０の分割された後半部分（第４の個人データ）を新たなデータセットｓｐ（新たな第１の個人データ）として、ステップＳ８２１以降の処理（２回目の匿名化）を実行する。

[0166] 図１０Ａにおいて、匿名性評価部３２３５は、「生年：１９６１～１９７７」を含む再分割要求を分割属性決定部３２３３へ送信する（ステップＳ８２１）。

[0167] 次に、その再分割要求を受信した分割属性決定部３２３３は、優先度決定情報３２１０を利用し、属性名及び対応するインデックス「生年：４、診療年月：１」に基づいて重みを生成する（ステップＳ８２２）

ここでは、分割属性決定部３２３３は、「生年」の属性に対応する重みを「８」、「診療年月」の属性に対応する重みを「１」と算出する。

[0168] 次に、分割属性決定部３２３３は、情報損失量計算部３２２へ情報損失量ＩＬＡの計算要求を行う（ステップＳ８２３）

次に、その情報損失量 I L A の計算要求を受信した情報損失量計算部 3 2 2 は、個人データ記憶装置 1 0 0 へ「生年」の属性の属性値が「1 9 6 1」～「1 9 7 7」であるデータレコード r p 1 1 1（そのデータセット s p 1 1 0 の後半部分）の取得要求を送信する（ステップ S 8 2 4）。換言すると、情報損失量計算部 3 2 2 は、個人データ記憶装置 1 0 0 に個人データの取得を要求する。

[0169] 次に、そのデータセット s p 1 1 0 の後半部分を受信した情報損失量計算部 3 2 2 は、情報損失量 I L A を計算し、計算した情報損失量 I L A を分割属性決定部 3 2 3 3 へ送信する（ステップ S 8 2 5）。

[0170] ここでは、情報損失量計算部 3 2 2 は、そのデータセット s p 1 1 0 の後半部分について、「生年」の属性の情報損失量 I L A - b i r t h - o v 1 9 6 1 を以下のように算出する。

[0171] $(1977 - 1961) \div (1977 - 1943) \times 11 = 5.181$

また、情報損失量計算部 3 2 2 は、「診療年月」の属性の情報損失量 I L A - m c - o v 1 9 6 1 を以下のように算出する。

[0172] $(201107 - 200512) \div (201107 - 200612) = 9.152$

次に、分割属性決定部 3 2 3 3 は、分割属性を決定する（ステップ S 8 2 6）

例えば、分割属性決定部 3 2 3 3 は、式 2 を用いて、「生年」の属性について、重みが「8」、情報損失量 I L A - b i r t h - o v 1 9 6 1 が「5.181」なので、「生年」属性の評価値として「41.448」を算出する。同様に、分割属性決定部 3 2 3 3 は、「診療年月」の属性について、重みが「1」、情報損失量 I L A - m c - o v 1 9 6 1 が「9.152」なので、「診療年月」属性の評価値として「9.152」を算出する。

[0173] 次に、分割属性決定部 3 2 3 3 は、「生年」の属性の評価値が「診療年月」の属性の評価値よりも大きいので、属性名が「生年」の属性を分割属性として決定する。

- [0174] 次に、図 10B において、分割属性決定部 3233 は、属性名の「生年」を含む分割値決定要求を分割値決定部 3234 へ送信する（ステップ S827）。
- [0175] 分割値決定要求を受信した分割値決定部 3234 は、個人データ記憶装置 100 へ個人データ取得要求を送信する（ステップ S828）。
- [0176] ここでは、分割属性決定部 3233 は、2 回目の匿名化の対象となるデータレコード r p 1 1 1（例えば、そのデータセット s p 1 1 0 の後半部分）を、取得することを要求する。
- [0177] 対象となるデータレコード r p 1 1 1 を受信した分割値決定部 3234 は、分割値を決定する（ステップ S829）。
- [0178] 図 13 は、そのデータセット s p 1 1 0 の分割された後半部分であるデータセット s p 1 3 0（新たな第 1 の個人データ）の分割値候補 1 1 2 1 及び分割値候補 1 1 2 2 の例を示す図である。
- [0179] まず、分割値決定部 3234 は、図 13 に示すように、データセット s p 1 3 0 のデータレコード r p 1 1 1 を、分割属性決定部 3233 が決定した属性の属性値の小さい順番に並べる。
- [0180] 次に、分割値決定部 3234 は、分割値候補を抽出する。図 13 に示すデータセット s p 1 3 0 において、分割値決定部 3234 は、分割値候補 1 1 2 1 及び分割値候補 1 1 2 2 を分割値候補として抽出する。
- [0181] 次に、分割値決定部 3234 は、分割値候補 1 1 2 1 及び分割値候補 1 1 2 2 のそれぞれに対する情報損失量 I L A - b i r t h を計算する。図 13 に示すデータセット s p 1 3 0 の場合、分割値決定部 3234 は、分割値候補 1 1 2 1 及び分割値候補 1 1 2 2 で分割した場合のそれぞれの情報損失量 I L A - b i r t h は、「5. 565」及び「4. 820」である。続けて、分割値決定部 3234 は、その情報損失量 I L A - b i r t h が最小である分割値候補 1 1 2 2 の「生年：1963」を分割値として決定する。
- [0182] 次に、分割値を決定した分割値決定部 3234 は、決定した分割値「生年：1963」を匿名性評価部 3235 へ送信する（ステップ S830）。換

言すると、分割値決定部 3 2 3 4 は、匿名性評価部 3 2 3 5 に、匿名性の評価を要求する。

[0183] 分割値「生年：1 9 6 3」を受信した匿名性評価部 3 2 3 5 は、匿名性評価を行う（ステップ S 8 3 1）。

[0184] 匿名性評価部 3 2 3 5 は、分割値の「生年：1 9 6 3」で分割された、それぞれのデータレコード r p 1 1 1 の数を計数する。図 1 4 は、分割値決定部 3 2 3 4 が分割値を分割値候補 1 2 2 2 「生年：1 9 6 3」に決定した場合の、図 1 3 に示すデータセット s p 1 3 0 が分割されるイメージを示す図である。図 1 4 に示す例では、匿名性評価部 3 2 3 5 は、分割した後のその前半部分のデータセット s p 1 4 0 のデータレコード r p 1 1 1 の数を 6 個、分割した後のその後半部分のデータセット s p 1 5 0 のデータレコード r p 1 1 1 の数を 5 個と計数する。

[0185] データセット s p 1 4 0 及びデータセット s p 1 5 0 のそれぞれのデータレコード r p 1 1 1 の数が 2 k 個未満である。そこで、匿名性評価部 3 2 3 5 は、「生年：1 9 6 1～1 9 6 3」を含む汎化実行要求及び「生年：1 9 6 4～1 9 7 7」を含む汎化実行要求を汎化実行部 3 2 3 6 へ送信する（ステップ S 8 1 3）。

[0186] 汎化の実行要求を受信した汎化実行部 3 2 3 6 は、「生年」の属性の属性値が「1 9 6 1」～「1 9 6 3」のデータレコード r p 1 1 1 の汎化及び「生年」の属性の属性値が「1 9 6 4」～「1 9 7 7」のデータレコード r p 1 1 1 の汎化を実行する（ステップ S 8 1 4）

図 1 4 に示すように、データセット s p 1 4 0 は、「生年」の属性の属性値が「1 9 6 1～1 9 6 3」、「診療年月」の属性の属性値が「2 0 0 6 1 2～2 0 1 1 0 5」に汎化される。また、データセット s p 1 5 0 は、「生年」の属性の属性値が「1 9 6 4～1 9 7 7」、「診療年月」の属性の属性値が「2 0 0 7 0 6～2 0 1 1 0 4」に汎化される。

[0187] 次に、汎化実行部 3 2 3 6 は、汎化したデータレコード r p 1 1 1 を匿名化済個人データ記憶装置 2 0 0 へ記録する（ステップ S 8 1 5）

上述した本実施形態における効果は、第１の実施形態の効果と同様に、利用目的に合致するように制御してデータセットを匿名化することと、匿名化されたデータセットにおける情報の損失を低減することとを両立させることができる点である。

[0188] その理由は、分割属性決定部３２３３が優先度 p と情報損失量 ILA とに基づいて評価値を生成し、生成した評価値に基づいて汎化する属性を決定するようにしたからである。

<<<第３の実施形態>>>

次に、本発明の第３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。

[0189] 図１５は、本実施形態に係る匿名化装置３３０の構成を示すブロック図である。本実施形態の匿名化装置３３０は、ボトムアップアプローチにより匿名化を行う。

[0190] 図１５に示すように、匿名化装置３３０は、優先度決定情報記憶部３２１と情報損失量計算部３３２と匿名化处理部３３３とを含む。

[0191] 尚、匿名化装置３３０は、匿名化装置３１０に替えて、図２に示すシステムに含まれてよい。

[0192] 優先度決定情報記憶部３２１は、優先度 p を決定する情報を記憶する。尚、優先度 p を決定する情報は、システムの利用者により予め設定されている。また、優先度 p を決定する情報は、分割属性決定部３２３３が図５に示す通信部７０６により、予め外部のシステムから受信するようにしてもよい。

[0193] 図１６は、優先度決定情報記憶部３２１に記憶される優先度決定情報３３１０の一例を示す図である。図１６に示すように、本実施形態の優先度決定情報３３１０は、優先順位と属性名と閾値との組を１以上含む。優先順位は、例えば、対応する属性名で特定される属性を、汎化する順番を示す。閾値は、例えば、下位の優先順位の属性の情報損失量 ILA から上位の優先順位の属性の情報損失量 ILA を引いた値が、この閾値を超える場合に、下位の

優先順位の属性を汎化する順番を先にする場合の値を示す。

[0194] 図 16 において、優先順位は、数字が小さいほど上位の優先順位であることを示すものとする。即ち、図 16 において、上位の優先順位の属性は、属性の名称が「年齢」の属性であり、下位の優先順位の属性は、属性の名称が「2011年度の診療月」の属性である。

[0195] 尚、優先度決定情報は、優先順位と属性名との組を含むようにしてよい。この場合、匿名化処理部 333 は、例えば、図示しない内部記憶手段に閾値を保持するようにしてもよい。

[0196] 情報損失量計算部 332 は、汎化による情報損失量 I L A を算出し、出力する。情報損失量計算部 332 は、例えば、個人データの属性に含まれる異なる属性値の数を計数し、これを情報損失量 I L A とする。

[0197] 匿名化処理部 333 は、汎化属性決定部 3333 と汎化実行部 3336 と匿名性評価部 3335 とを含む。

[0198] 汎化属性決定部 3333 は、汎化する属性を決定する。汎化属性決定部 3333 は、例えば、次のようにして、汎化する属性を決定する。第一に、汎化属性決定部 3333 は、優先順位が下位の属性の情報損失量 I L A から優先順位が上位の属性の情報損失量 I L A を減算して、情報損失量差分を算出する。次に、汎化属性決定部 3333 は、情報損失量差分とその優先順位が上位の属性の閾値とを比較する。そして、汎化属性決定部 3333 は、情報損失量差分が閾値以上の場合、下位の優先順位の属性を汎化することを決定する。また、汎化属性決定部 3333 は、情報損失量差分が閾値未満の場合、上位の優先順位の属性を汎化することを決定する。

[0199] 尚、汎化属性決定部 3333 による、汎化する属性の決定は、以下のようしてもよい。まず、汎化属性決定部 3333 において、優先順位が上位の属性の優先度 p はその属性の閾値であり、優先順位が下位の属性の優先度 p は「0」である。汎化属性決定部 3333 は、それらの属性のそれぞれについて、評価値 = 情報損失量 I L A + 優先度 p という評価式を利用して評価値を算出する。次に、汎化属性決定部 3333 は、評価値の大きい属性を、汎化

する属性として決定する。尚、優先順位が上位の属性と優先順位が下位の属性とで評価値が同じである場合、汎化属性決定部 3333 は、例えばその優先順位が上位の属性を、汎化する属性として決定するようにしてよい。

[0200] 汎化実行部 3336 は、汎化属性決定部 3333 が決定したその属性を汎化する。

[0201] 匿名性評価部 3335 は、汎化実行部 3336 が汎化したデータセットが匿名性指標を満たしているか否かを判定する。

[0202] 図 17 は、本実施形態の個人データ記憶装置 100 に記憶されるデータセット sp160 の一例を示す図である。図 17 に示すデータセット sp160 の各データレコード rp161 のそれぞれは、「氏名」、「年齢」、「2011 年度診療月」及び「病名」の属性の属性値を含む。尚、本実施形態では、「年齢」と「2011 年度診療年月」を匿名化対象の属性（準識別子である）とする。

[0203] 以上説明した匿名化装置 330 は、図 1 に示す匿名化装置 310 と同様に、図 5 に示すコンピュータ 700 によって実現してもよい。

[0204] 図 18A、図 18B 及び図 18C は、本実施形態における匿名化装置 330 の動作を示すシーケンス図である。

[0205] 図 18A において、汎化属性決定部 3333 は、例えば図 5 に示す入力部 704 からのシステムの利用者による匿名化実行要求の入力を、受け付ける（ステップ S841）。

[0206] ここで、匿名化実行要求は、例えば、k-匿名化の k の値（例えば、「3」）を含む。

[0207] 次に、その匿名化実行要求を受信した汎化属性決定部 3333 は、優先度決定情報記憶部 321 に優先度決定情報取得要求を送信する（ステップ S842）。

[0208] 次に、その優先度決定情報取得要求の応答として優先度決定情報 3310 を受信した汎化属性決定部 3333 は、情報損失量計算部 332 へ情報損失量計算要求を送信する。（ステップ S843）。

- [0209] 次に、その情報損失量計算要求を受信した情報損失量計算部332は、個人データ記憶装置100に個人データ取得要求を送信する。（ステップS844）。
- [0210] 次に、その個人データ取得要求への応答としてデータセットsp160を受信した情報損失量計算部332は、情報損失量ILAを計算し、その計算した情報損失量ILAを汎化属性決定部3333へ送信する（ステップS845）。
- [0211] ここで、情報損失量計算部332は、情報損失量ILAを属性の値の種類数で計算する。即ち、情報損失量計算部332は、「年齢」の属性の属性値の種類が12種類なので、「年齢」の属性の情報損失量ILAbirthを「12」と計算する。また、情報損失量計算部332は、「2011年度診療月」の属性の属性値の種類が10種類なので、「2011年度診療月」の属性の情報損失量ILAmc2011を「10」と計算する。
- [0212] 次に、情報損失量ILAを受信した汎化属性決定部3333は、汎化する属性を決定する（ステップS846）。
- [0213] 例えば、汎化属性決定部3333は、受信した優先度決定情報3310を利用して、受信した情報損失量ILAに基づいて、その汎化する属性を決定する。
- [0214] 例えば、汎化属性決定部3333は、優先順位が「2」の属性である「2011年度の診療月」の情報損失量ILAから、優先順位が「1」の属性である「年齢」の情報損失量ILAを減算して、情報損失量差分を算出する。即ち、 $10 - 12 = -2$ を算出する。次に、汎化属性決定部3333は、その情報損失量差分と優先順位が「1」の属性である「年齢」の閾値（「3」）とを比較する。この場合、 $-2 < 3$ なので、汎化属性決定部3333は、優先順位が「1」の属性である「年齢」を汎化することを決定する。
- [0215] 尚、汎化属性決定部3333は、実施形態1で説明した方法を用いて、汎化する属性を決定してもよい。
- [0216] 次に、汎化属性決定部3333は、その汎化すると決定した属性の属性名

（この場合、「年齢」）を含む汎化実行要求を汎化属性実行部に送信する（ステップS 8 4 7）。

[0217] 次に、その汎化実行要求を受信した汎化実行部 3 3 3 6 は、図 1 7 に示すデータセット s p 1 6 0 を図 1 9 に示すデータセット s p 1 6 2 のように汎化する（ステップS 8 4 8）。

[0218] 図 1 9 は、本実施形態の匿名化装置 3 3 0 による匿名化処理の途中段階の（一部が汎化された）データセットの例を示す図である。

[0219] 次に、汎化実行部 3 3 3 6 は、データセット s p 1 6 2 を含む匿名性評価要求を匿名性評価部 3 3 3 5 へ送信する（ステップS 8 4 9）。

[0220] 尚、汎化実行部 3 3 3 6 は、データセット s p 1 6 2 を図 5 に示す記憶部 7 0 2 に格納し、その格納したアドレスを含む匿名性評価要求を匿名性評価部 3 3 3 5 へ送信するようにしてもよい。以下における匿名性評価要求についても同様である。

[0221] 次に、その匿名性評価要求を受信した匿名性評価部 3 3 3 5 は、データセット s p 1 6 2 の匿名性を評価する。図 1 9 のデータセット s p 1 6 2 の場合、匿名性評価部 3 3 3 5 は、「診療月」の属性について k - 匿名性の k の値（「3」）を満たしていないと判定する（ステップS 8 5 0）。

[0222] 次に、図 1 8 B において、匿名性評価部 3 3 3 5 は、汎化属性決定部 3 3 3 3 へ汎化属性決定要求を送信する（ステップS 8 5 1）。

次に、その汎化属性決定要求を受信した汎化属性決定部 3 3 3 3 は、情報損失量計算部 3 3 2 へ情報損失量計算要求を送信する（ステップS 8 5 2）。

[0223] 次に、その情報損失量計算要求を受信した情報損失量計算部 3 3 2 は、情報損失量 I L A を計算し、計算した情報損失量 I L A を汎化属性決定部 3 3 3 3 へ送信する（ステップS 8 5 3）。

[0224] ここで、図 1 9 に示すデータセット s p 1 6 2 の場合、「年齢」の属性の属性値の種類は「2 1 ~ 2 4」、「3 1 ~ 4 0」、「4 1 ~ 5 1」及び「5 2 ~ 5 8」の 4 種類である。また、「2 0 1 1 年度診療月」の属性の属性値

の種類は、１０種類である。従って、情報損失量計算部３３２は、「年齢」及び「２０１１年度診療月」のそれぞれの属性に対応する情報損失量 $ILA - birth$ 及び情報損失量 $ILA - mc 2011$ を、「４」及び「１０」と計算する。

[0225] 次に、その情報損失量 ILA を受信した汎化属性決定部３３３は、汎化する属性の決定をする（ステップＳ８５４）。

[0226] 優先順位が「１」の「年齢」の属性の情報損失量 $ILA - birth$ は「４」、優先順位が「２」の「２０１１年度診療月」の属性の情報損失量 $ILA - mc 2011$ は「１０」なので、情報損失量差分は、以下のとおりである。

[0227] $10 - 4 = 6$

汎化属性決定部３３３は、この情報損失量差分（「６」）と優先順位が「１」の属性である「年齢」の閾値（「３」）とを比較する。この場合、 $6 > 3$ なので、汎化属性決定部３３３は、優先順位が「２」の属性である「２０１１年度診療月」を汎化することを決定する。

[0228] 次に、汎化属性決定部３３３は、その汎化すると決定した属性名（この場合、「２０１１年度診療月」）を含む汎化実行要求を汎化属性実行部に送信する（ステップＳ８５５）。

[0229] 次に、その汎化実行要求を受信した汎化実行部３３３６は、図１９に示すデータセット $sp 162$ を図２０に示すデータセット $sp 163$ のように汎化する（ステップＳ８５６）。

[0230] 図２０は、本実施形態の匿名化装置３３０による匿名化処理の途中段階の（一部が汎化された）データセットの例を示す図である。

[0231] 次に、汎化実行部３３３６は、データセット $sp 163$ を含む匿名性評価要求を匿名性評価部３３３５へ送信する（ステップＳ８５７）

次に、その匿名性評価要求を受信した匿名性評価部３３３５は、データセット $sp 163$ の匿名性を評価する。図２０に示すデータセット $sp 163$ の場合、匿名性評価部３３３５は、「診療月」の属性と「２０１１年度診療

月」の属性とを組み合わせた場合について、 k －匿名性の k の値（「3」）を満たしていないと判定する（ステップS858）。

[0232] 次に、匿名性評価部3335は、汎化属性決定部3333へ汎化属性決定要求を送信する（ステップS859）。

次に、図18Cにおいて、その汎化属性決定要求を受信した汎化属性決定部3333は、情報損失量計算部332へ情報損失量計算要求を送信する（ステップS860）。

[0233] 次に、その情報損失量計算要求を受信した情報損失量計算部332は、情報損失量ILAを計算し、その計算した情報損失量ILAを汎化属性決定部3333へ送信する（ステップS861）

ここで、図20に示すデータセットsp163の場合、「年齢」の属性の属性値の種類は4種類、「2011年度診療月」の属性の属性値の種類は4種類である。従って、情報損失量計算部332は、「年齢」及び「2011年度診療月」のそれぞれの属性に対応する情報損失量ILA-birth及び情報損失量ILA-mc2011を、何れも「4」と計算する。

[0234] 次に、情報損失量ILAを受信した汎化属性決定部3333は、汎化する属性を決定する（ステップS862）。

[0235] 優先順位が「1」の「年齢」の属性の情報損失量ILA-birthは「4」、優先順位が「2」の「2011年度診療月」の属性の情報損失量ILA-mc2011は「4」なので、情報損失量差分は、以下のとおりである。

[0236] $4 - 4 = 0$

汎化属性決定部3333は、この情報損失量差分（「0」）と優先順位が「1」の属性である「年齢」の閾値（「3」）とを比較する。この場合、 $0 < 3$ なので、汎化属性決定部3333は、優先順位が「1」の属性である「年齢」を汎化することを決定する。

[0237] 次に、汎化属性決定部3333は、汎化すると決定した属性の属性名（この場合、「年齢」）を含む汎化実行要求を、汎化実行部3336に送信する

(ステップS 8 6 3)。

[0238] 次に、その汎化実行要求を受信した汎化実行部 3 3 3 6 は、図 2 0 に示すデータセット s p 1 6 3 を図 2 1 に示すデータセット s p 1 6 4 のように汎化する (ステップS 8 6 4)。

[0239] 図 2 1 は、本実施形態の匿名化装置 3 3 0 により匿名化处理されたデータセットの例を示す図である。

[0240] 次に、汎化実行部 3 3 3 6 は、データセット s p 1 6 4 を含む匿名性評価要求を匿名性評価部 3 3 3 5 へ送信する (ステップS 8 6 5)

次に、その匿名性評価要求を受信した匿名性評価部 3 3 3 5 は、データセット s p 1 6 4 の匿名性を評価する。図 2 1 に示すデータセット s p 1 6 4 の場合、匿名性評価部 3 3 3 5 は、データセット s p 1 6 4 が k -匿名性を満たしていると判定する (ステップS 8 6 6)。

[0241] 次に、匿名性評価部 3 3 3 5 は、その匿名性を満たしたデータセット s p 1 6 4 を、匿名化済個人データ記憶装置 2 0 0 へ送信する (ステップS 8 6 7)。

[0242] データセット s p 1 6 4 を受信した匿名化済個人データ記憶部 2 a は、そのデータセット s p 1 6 4 を匿名化済データセット s t 1 2 0 (匿名化済個人データ) として記憶する。(ステップS 8 6 8)

上述した本実施形態における効果は、第 1 の実施形態の効果と同様に、利用目的に合致するように制御してデータセットを匿名化することと、匿名化されたデータセットにおける情報の損失を低減することとを両立して可能にできる点である。

[0243] その理由は、汎化属性決定部 3 3 3 3 が優先順位と閾値と情報損失量 I L A とに基づいて評価値を生成し、生成した評価値に基づいて汎化する属性を決定するようにしたからである。

[0244] 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ってもよい。

[0245] 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。

[0246] 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。

[0247] 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しない範囲で変更することができる。

[0248] 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イミングが部分的に乃至全部において重複していたりしていてもよい。

[0249] 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い。

い。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。

[0250] 以上、各実施形態及び実施例を参照して本発明を説明したが、本発明は上記実施形態及び実施例に限定されるものではない。本発明の構成や詳細には、本発明のスコープ内で当業者が理解しえる様々な変更をすることができる。

[0251] 以上、実施形態を参照して本願発明を説明したが、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。

[0252] この出願は、2012年8月20日に提出された日本出願特願2012-181684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

[0253]

100	個人データ記憶装置
110	データセット s p
111	データレコード r p
130	データセット s p
140	データセット s p
150	データセット s p
160	データセット s p
161	データレコード r p
162	データセット s p
163	データセット s p
164	データセット s p
200	匿名化済個人データ記憶装置
210	匿名化済データセット s a
211	匿名化済データレコード

3 1 0	匿名化装置
3 1 2	情報損失量計算部
3 1 3	匿名化处理部
3 2 0	匿名化装置
3 2 1	優先度決定情報記憶部
3 2 2	情報損失量計算部
3 2 3	匿名化处理部
7 0 0	コンピュータ
7 0 1	C P U
7 0 2	記憶部
7 0 3	記憶装置
7 0 4	入力部
7 0 5	出力部
7 0 6	通信部
7 0 7	記録媒体
1 1 0 5	分割値候補
1 1 2 1	分割値候補
1 1 2 2	分割値候補
3 2 1 0	優先度決定情報
3 2 3 3	分割属性決定部
3 2 3 4	分割値決定部
3 2 3 5	匿名性評価部
3 2 3 6	汎化実行部
1 1 0 1 ~ 1 1 0 4	分割値候補
1 1 0 1 ~ 1 1 1 1	分割値候補
1 1 0 6 ~ 1 1 1 1	分割値候補

請求の範囲

- [請求項1] 匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、出力する情報損失量計算手段と、
前記属性のそれぞれに対応する優先度と前記情報損失量とに基づいて加工の対象とする前記属性を決定し、前記第1の個人データの前記決定した属性の属性値を加工した第2の個人データを生成し、出力する匿名化処理手段と、
を含む情報処理装置。
- [請求項2] 前記優先度は、前記加工した第2の個人データにおける情報の損失を、前記属性のいずれについてより少なくするかを示す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] 前記優先度を決定する情報を記憶する優先度決定情報記憶手段を更に含み、
前記匿名化処理手段は、前記優先度を決定する情報に基づいて前記優先度を決定することを特徴とする請求項1または2記載の情報処理装置。
- [請求項4] 前記匿名化処理手段は、前記情報損失量が一定ならば前記優先度が高いほど、及び前記優先度が一定ならば前記情報損失量大きいほど演算結果が大きくなるような評価式を利用して評価値を算出し、前記算出した評価値が最大の前記属性を前記加工の対象とする属性として決定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] 前記評価式は、情報損失量と優先度とを乗じる演算を含むことを特徴とする請求項4記載の情報処理装置。
- [請求項6] 前記評価式は、情報損失量と優先度とを加算する演算を含むことを特徴とする請求項4記載の情報処理装置。
- [請求項7] 前記匿名化処理手段は、

前記第 1 の個人データが匿名化の対象である前記属性のそれぞれの属性値を同一の値に汎化される場合の、前記情報損失量を算出し、算出した前記情報損失量に基づいて汎化対象である分割属性を決定する分割属性決定手段と、

前記決定した分割属性を軸として前記第 1 の個人データを分割して前記分割属性の属性値を汎化する場合の前記情報損失量が最小であるように、前記分割属性の分割値を決定する分割値決定手段と、

前記決定した分割値で前記第 1 の個人データを分割して生成した第 3 の個人データ及び第 4 の個人データのそれぞれについて、更なる分割が可能か否かを判定する匿名性評価手段と、

前記匿名化評価手段が更なる分割が可能でないと判定した前記第 3 の個人データ及び前記第 4 の個人データの前記分割属性の属性値を汎化し、出力する汎化実行手段と、を含み、

前記分割属性決定手段と前記分割値決定手段とは、前記匿名性評価手段が更なる分割が可能であると判定した前記第 3 の個人データ及び前記第 4 の個人データを新たな第 1 の個人データとして処理する

ことを特徴とする請求項 1 乃至 6 のいずれか 1 項に記載の情報処理装置。

[請求項 8]

前記匿名化処理手段は、

前記第 1 の個人データの匿名化の対象である前記属性のそれぞれに対応する前記情報損失量を算出し、算出した前記情報損失量と前記優先度とに基づいて汎化対象である汎化属性を決定する汎化属性決定手段と、

前記第 1 の個人データに含まれる前記決定した汎化属性の属性値を汎化して、第 5 の個人データを生成する汎化実行手段と、

前記第 5 の個人データについて、所定の匿名性を有しているか否かを判定し、前記第 5 の個人データが所定の匿名性を有していると判定した場合、前記第 5 の個人データを前記第 2 の個人データとして、出

力する匿名性評価手段と、を含み、

前記汎化属性決定手段及び汎化実行手段は、前記匿名性評価手段が前記第5の個人データが所定の匿名性を有していないと判定した場合、前記第5の個人データを新たな第1の個人データとして処理することを特徴とする請求項1乃至6のいずれか1項に記載の情報処理装置。

[請求項9]

コンピュータが、

匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、出力し、

前記属性のそれぞれに対応する優先度と前記情報損失量とに基づいて加工の対象とする前記属性を決定し、

前記第1の個人データの前記決定した属性の属性値を加工した第2の個人データを生成し、出力する、

匿名化処理方法。

[請求項10]

前記コンピュータが、

前記情報損失量が一定ならば前記優先度が高いほど、及び前記優先度が一定ならば前記情報損失量が多いほど演算結果が大きくなるような評価式を利用して評価値を算出し、

前記算出した評価値が最大の前記属性を前記加工の対象とする属性として決定する

ことを特徴とする請求項9記載の匿名化処理方法。

[請求項11]

匿名化対象の第1の個人データに含まれる属性のそれぞれに対応する情報損失量を算出し、出力する処理と、

前記属性のそれぞれに対応する優先度と前記情報損失量とに基づいて加工の対象とする前記属性を決定する処理と、

前記第1の個人データの前記決定した属性の属性値を加工した第2の個人データを生成し、出力する処理と、をコンピュータに実行させる

プログラムを記録した不揮発性記録媒体。

[請求項12]

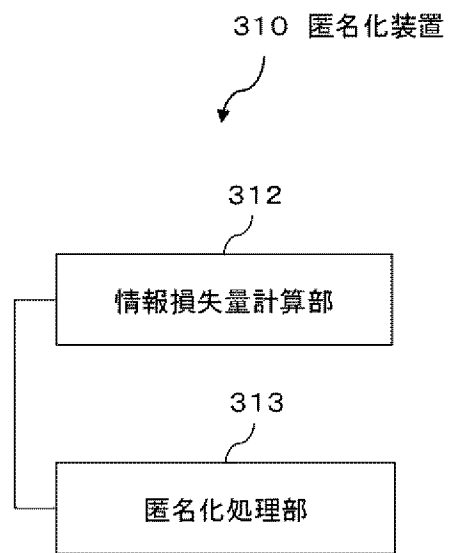
前記属性を決定する処理は、

前記情報損失量が一定ならば前記優先度が高いほど、及び前記優先度が一定ならば前記情報損失量が多いほど演算結果が大きくなるような評価式を利用して評価値を算出する処理と、

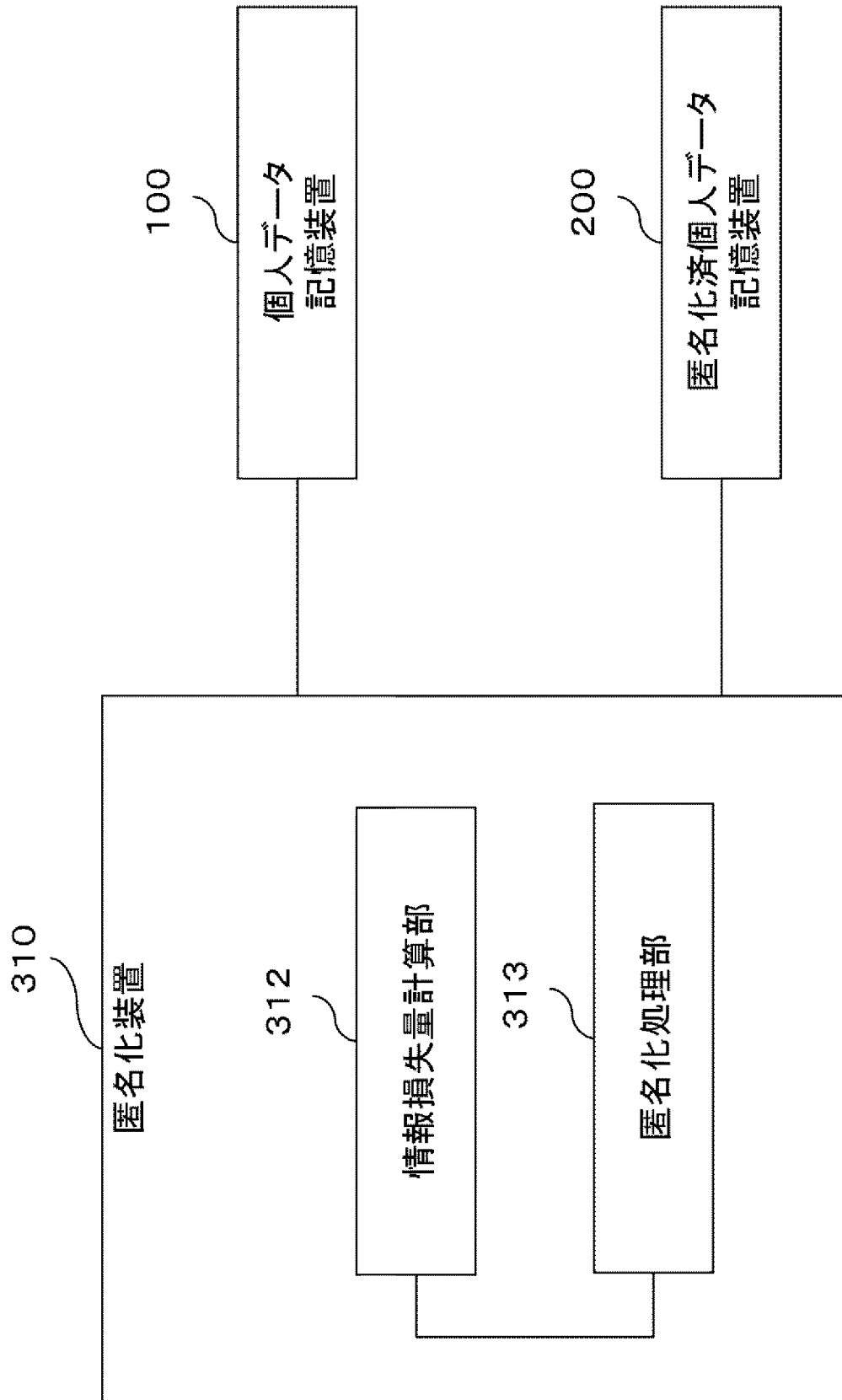
前記算出した評価値が最大の前記属性を前記加工の対象とする属性として決定する処理と、を含む

ことを特徴とする請求項11記載のプログラムを記録した不揮発性記録媒体。

[図1]



[図2]



[図3]

110 データセットsp

氏名	生年	診療月日	傷病名
patientA	1949	201006	DiseaseA
patientB	1962	200612	DiseaseB
patientC	1961	201010	DiseaseC
patientD	1962	201107	DiseaseD
patientE	1951	200905	DiseaseA
patientF	1964	201104	DiseaseB
patientG	1949	201003	DiseaseC
patientH	1954	200912	DiseaseD
patientI	1953	201005	DiseaseA
patientJ	1943	201107	DiseaseB
patientK	1956	200909	DiseaseC
patientL	1972	200706	DiseaseD
patientM	1955	200512	DiseaseE
patientN	1963	201105	DiseaseB
patientO	1961	200701	DiseaseC
patientP	1977	201005	DiseaseE
patientQ	1976	201002	DiseaseG
patientR	1961	201102	DiseaseH
patientS	1951	200710	DiseaseI
patientT	1967	200706	DiseaseK

111 データレコードrp

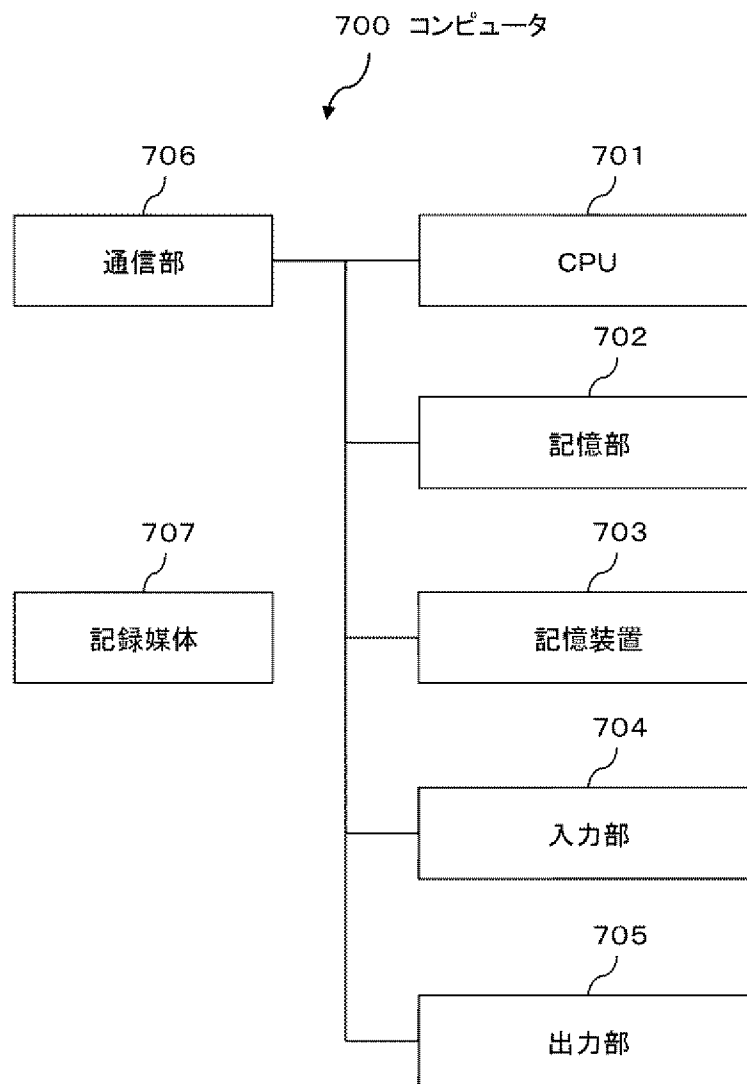
[図4]

210 匿名化済データセットsa

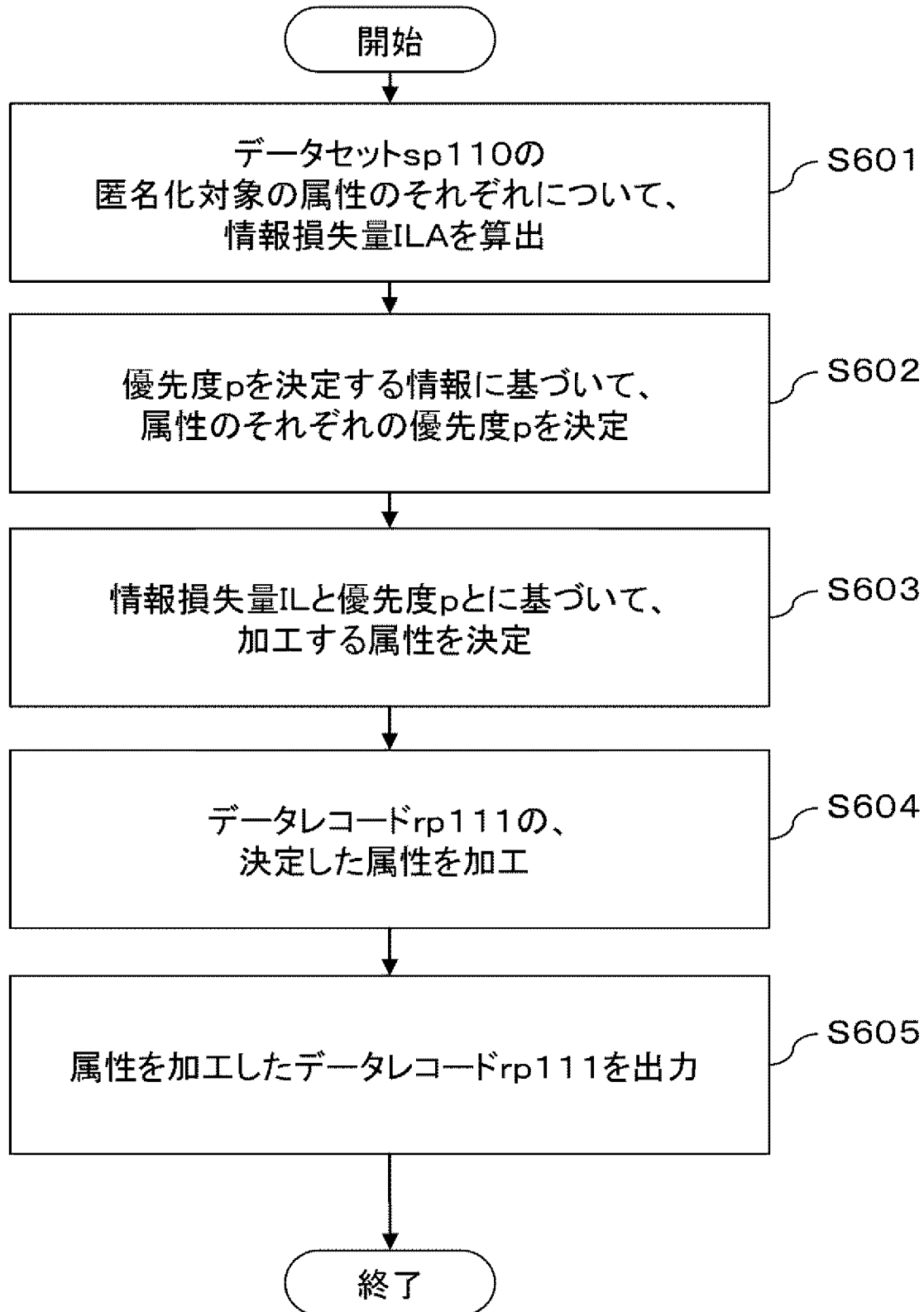
氏名	生年	診療月日	傷病名
patientJ	1943~1956	200512~201107	DiseaseB
patientG	1943~1956	200512~201107	DiseaseC
patientA	1943~1956	200512~201107	DiseaseA
patientS	1943~1956	200512~201107	DiseaseI
patientE	1943~1956	200512~201107	DiseaseA
patientI	1943~1956	200512~201107	DiseaseA
patientH	1943~1956	200512~201107	DiseaseD
patientN	1943~1956	200512~201107	DiseaseE
patientK	1943~1956	200512~201107	DiseaseC
patientO	1961~1963	200612~201107	DiseaseC
patientC	1961~1963	200612~201107	DiseaseC
patientR	1961~1963	200612~201107	DiseaseH
patientD	1961~1963	200612~201107	DiseaseD
patientB	1961~1963	200612~201107	DiseaseB
patientN	1961~1963	200612~201107	DiseaseB
patientF	1964~1977	200706~201104	DiseaseB
patientT	1964~1977	200706~201104	DiseaseK
patientL	1964~1977	200706~201104	DiseaseD
patientQ	1964~1977	200706~201104	DiseaseG
patientP	1964~1977	200706~201104	DiseaseE

211 データレコードra

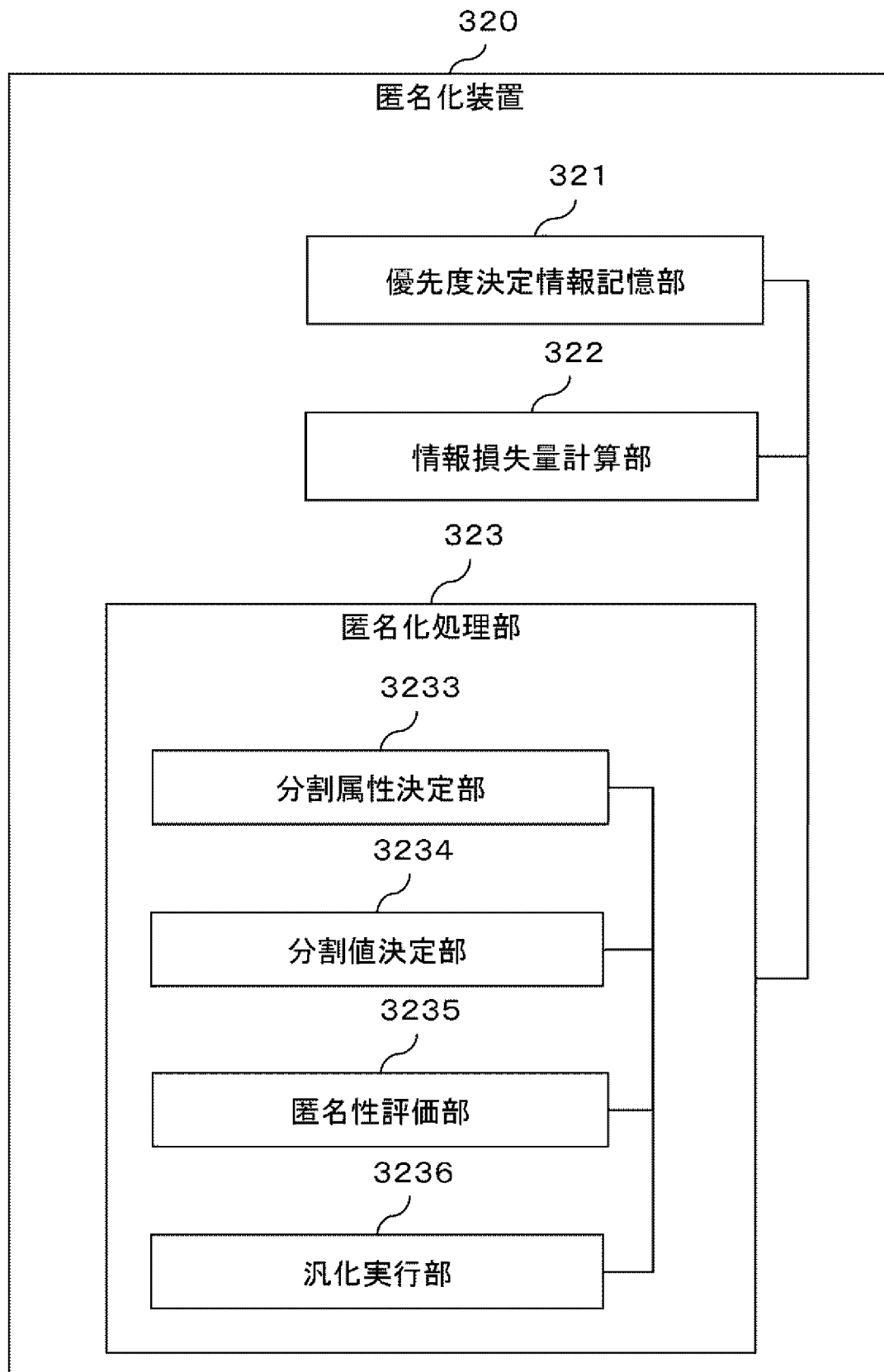
[図5]



[図6]



[図7]



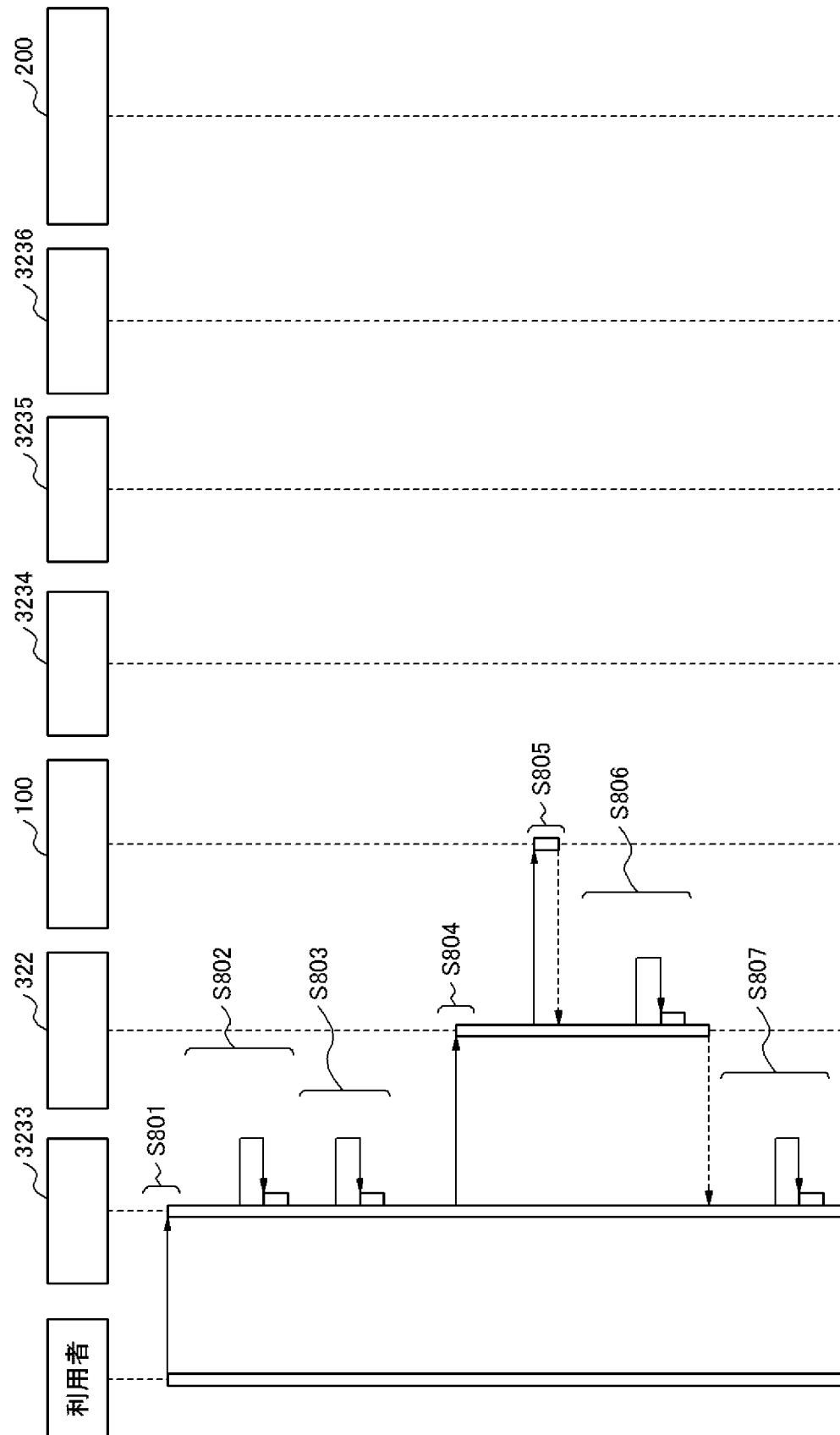
[図8]

3210 優先度決定情報



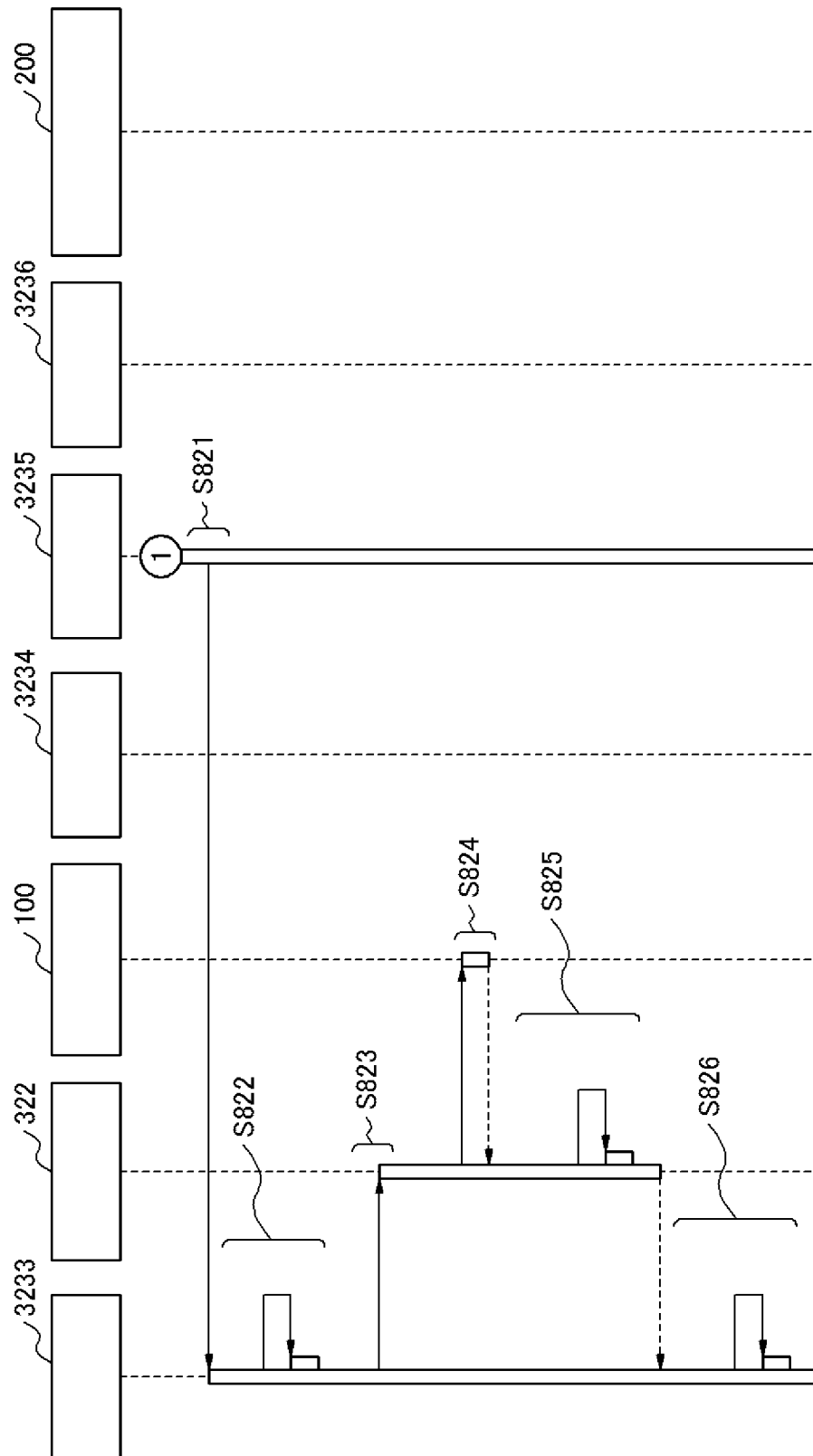
インデックス	重み
1	1
2	2
3	4
4	8
5	16

[図9A]

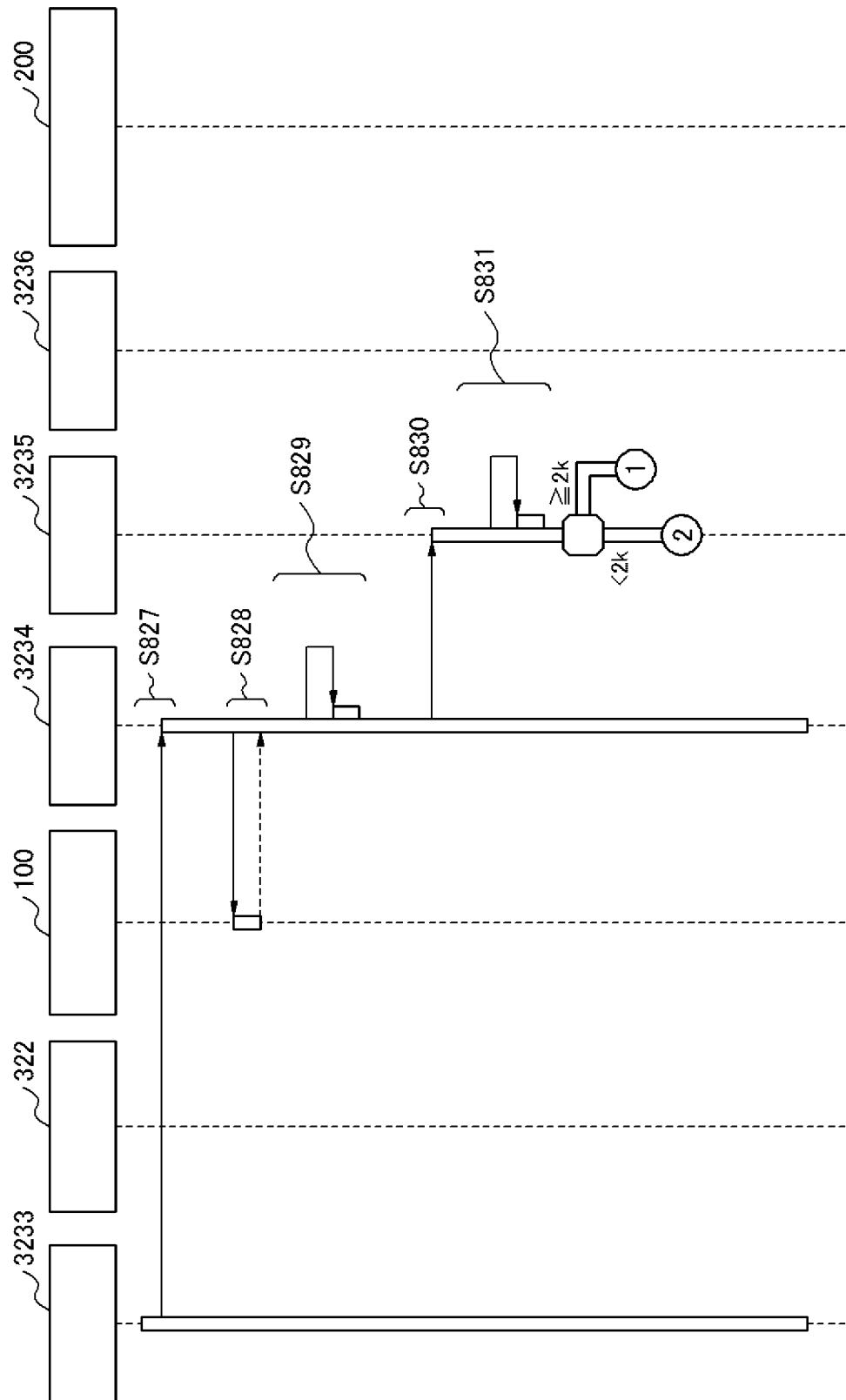


The diagram illustrates a flowchart for determining a value k . It starts at step S808, where a decision is made based on a comparison. If the result is $\geq 2k$, it follows path 1 to step S812, then to step S813, and finally to step S815. If the result is $< 2k$, it follows path 2 to step S814, and then to step S815. The diagram also features a vertical axis with labels 100, 200, 322, 323, 324, 3235, 3236, and 3237.

[図10A]



[図10B]



[図11]

120 データセット

氏名	生年	診療月日	傷病名
patientA	1943~1977	200512~201107	DiseaseA
patientB	1943~1977	200512~201107	DiseaseB
patientC	1943~1977	200512~201107	DiseaseC
patientD	1943~1977	200512~201107	DiseaseD
patientE	1943~1977	200512~201107	DiseaseA
patientF	1943~1977	200512~201107	DiseaseB
patientG	1943~1977	200512~201107	DiseaseC
patientH	1943~1977	200512~201107	DiseaseD
patientI	1943~1977	200512~201107	DiseaseA
patientJ	1943~1977	200512~201107	DiseaseB
patientK	1943~1977	200512~201107	DiseaseC
patientL	1943~1977	200512~201107	DiseaseD
patientM	1943~1977	200512~201107	DiseaseE
patientN	1943~1977	200512~201107	DiseaseB
patientO	1943~1977	200512~201107	DiseaseC
patientP	1943~1977	200512~201107	DiseaseE
patientQ	1943~1977	200512~201107	DiseaseG
patientR	1943~1977	200512~201107	DiseaseH
patientS	1943~1977	200512~201107	DiseaseI
patientT	1943~1977	200512~201107	DiseaseK

121 データレコード

[図12]

110 データセットsp

氏名	生年	診療月日	傷病名
patientJ	1943	201107	DiseaseB
patientG	1949	201003	DiseaseC
patientA	1949	201006	DiseaseA
patientS	1951	200710	DiseaseI
patientE	1951	200905	DiseaseA
patientI	1953	201005	DiseaseA
patientH	1954	200912	DiseaseD
patientN	1955	200512	DiseaseE
patientK	1956	200909	DiseaseC
patientO	1961	200701	DiseaseC
patientC	1961	201010	DiseaseC
patientR	1961	201102	DiseaseH
patientD	1962	201107	DiseaseD
patientB	1962	200612	DiseaseB
patientN	1963	201105	DiseaseB
patientF	1964	201104	DiseaseB
patientT	1967	200706	DiseaseK
patientL	1972	200706	DiseaseD
patientQ	1976	201002	DiseaseG
patientP	1977	201005	DiseaseE

1101 分割値候補
 1102 分割値候補
 1103 分割値候補
 1104 分割値候補
 1105 分割値候補
 1106 分割値候補
 1107 分割値候補
 1108 分割値候補
 1109 分割値候補
 1110 分割値候補
 1111 分割値候補

[図13]

130 データセットsp



氏名	生年	診療月日	傷病名
patientO	1961	200701	DiseaseC
patientC	1961	201010	DiseaseC
patientR	1961	201102	DiseaseH
patientD	1962	201107	DiseaseD
patientB	1962	200612	DiseaseB
patientN	1963	201105	DiseaseB
patientF	1964	201104	DiseaseB
patientT	1967	200706	DiseaseK
patientL	1972	200706	DiseaseD
patientQ	1976	201002	DiseaseG
patientP	1977	201005	DiseaseE

1121 分割値候補

1122 分割値候補

[図14]

140 データセットsp

氏名	生年	診療月日	傷病名
patientO	1961-1963	200612-201105	DiseaseC
patientC	1961-1963	200612-201105	DiseaseC
patientR	1961-1963	200612-201105	DiseaseH
patientD	1961-1963	200612-201105	DiseaseD
patientB	1961-1963	200612-201105	DiseaseB
patientN	1961-1963	200612-201105	DiseaseB

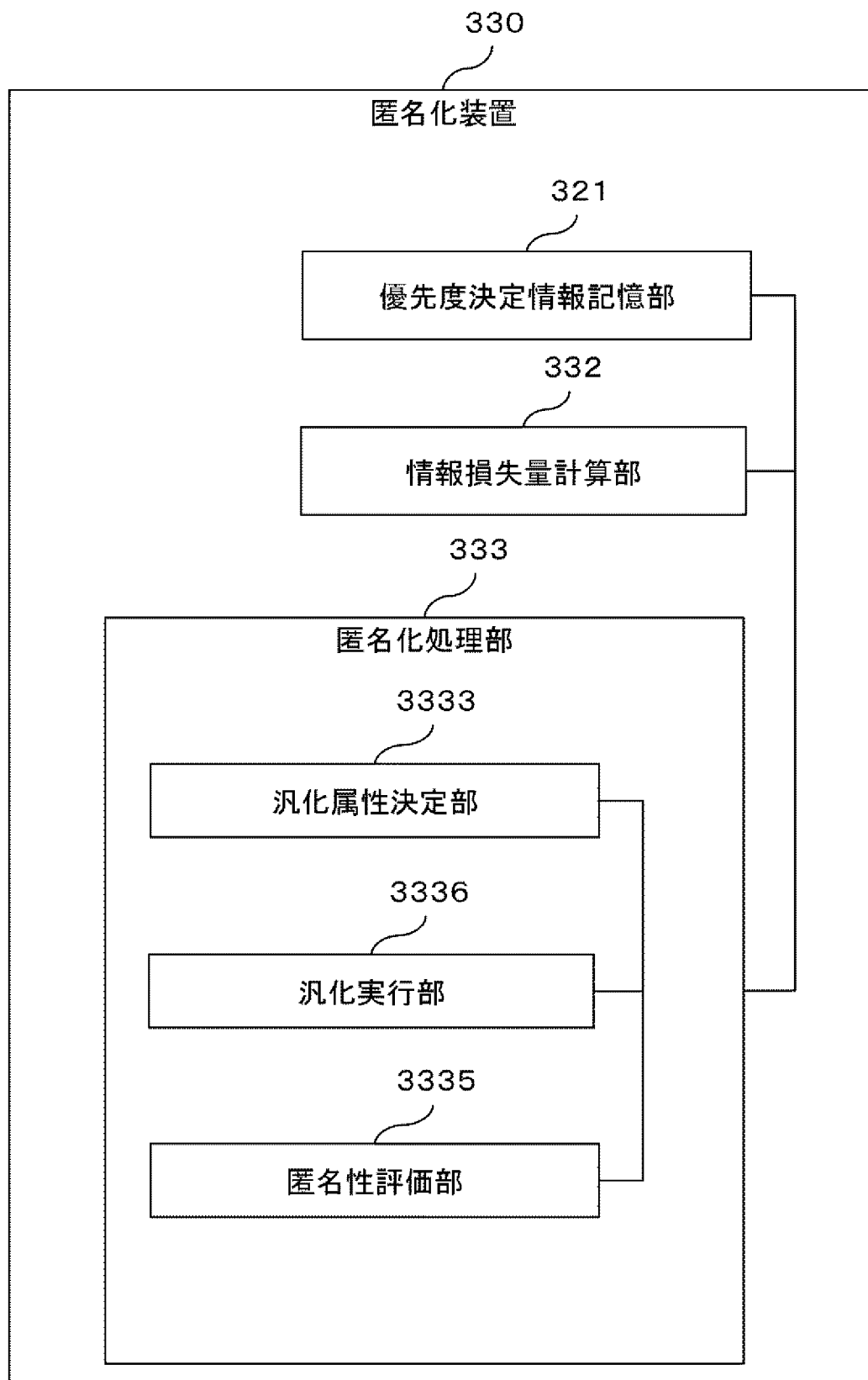
150 データセットsp

氏名	生年	診療月日	傷病名
patientF	1964-1977	200706-201104	DiseaseB
patientT	1964-1977	200706-201104	DiseaseK
patientL	1964-1977	200706-201104	DiseaseD
patientQ	1964-1977	200706-201104	DiseaseG
patientP	1964-1977	200706-201104	DiseaseE

130 データセットsp

氏名	生年	診療月日	傷病名
patientO	1961	200701	DiseaseC
patientC	1961	201010	DiseaseC
patientR	1961	201102	DiseaseH
patientD	1962	201107	DiseaseD
patientB	1962	200612	DiseaseB
patientN	1963	201105	DiseaseB
patientF	1964	201104	DiseaseB
patientT	1967	200706	DiseaseK
patientL	1972	200706	DiseaseD
patientQ	1976	201002	DiseaseG
patientP	1977	201005	DiseaseE

[図15]



[図16]

3310 優先度決定情報



優先順位	属性の名前	閾値
1	年齢	3
2	2011年度の診療月	2

[図17]

160 データセットsp



氏名	年齢	2011年度診療月	病名
患者A	21	4月	病名A
患者B	22	5月	病名B
患者C	24	6月	病名C
患者D	31	7月	病名C
患者E	37	8月	病名B
患者F	40	12月	病名D
患者G	41	12月	病名D
患者H	45	12月	病名B
患者I	51	1月	病名C
患者J	52	2月	病名B
患者K	55	3月	病名C
患者L	58	9月	病名D

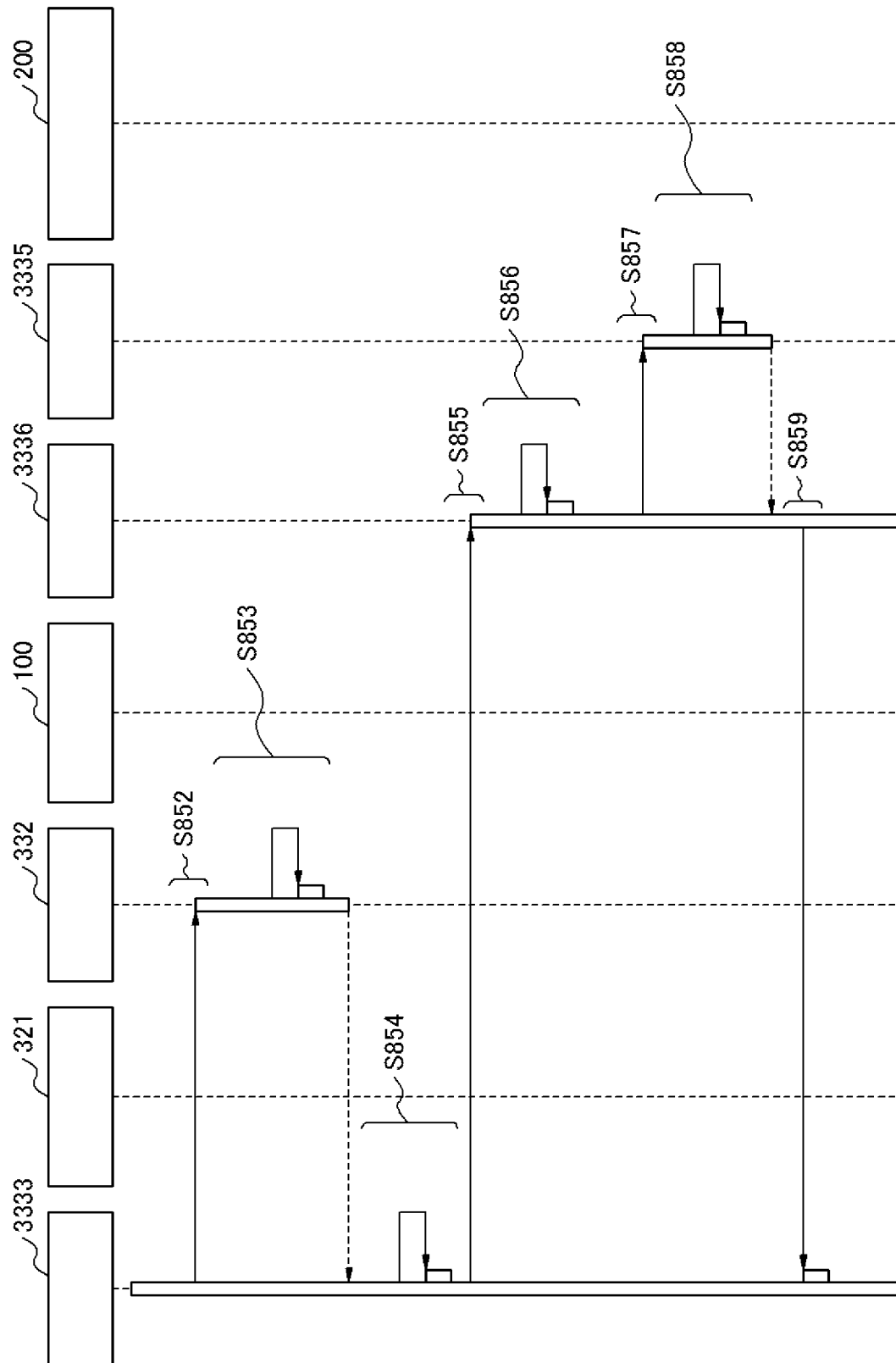
161 データレコードrp



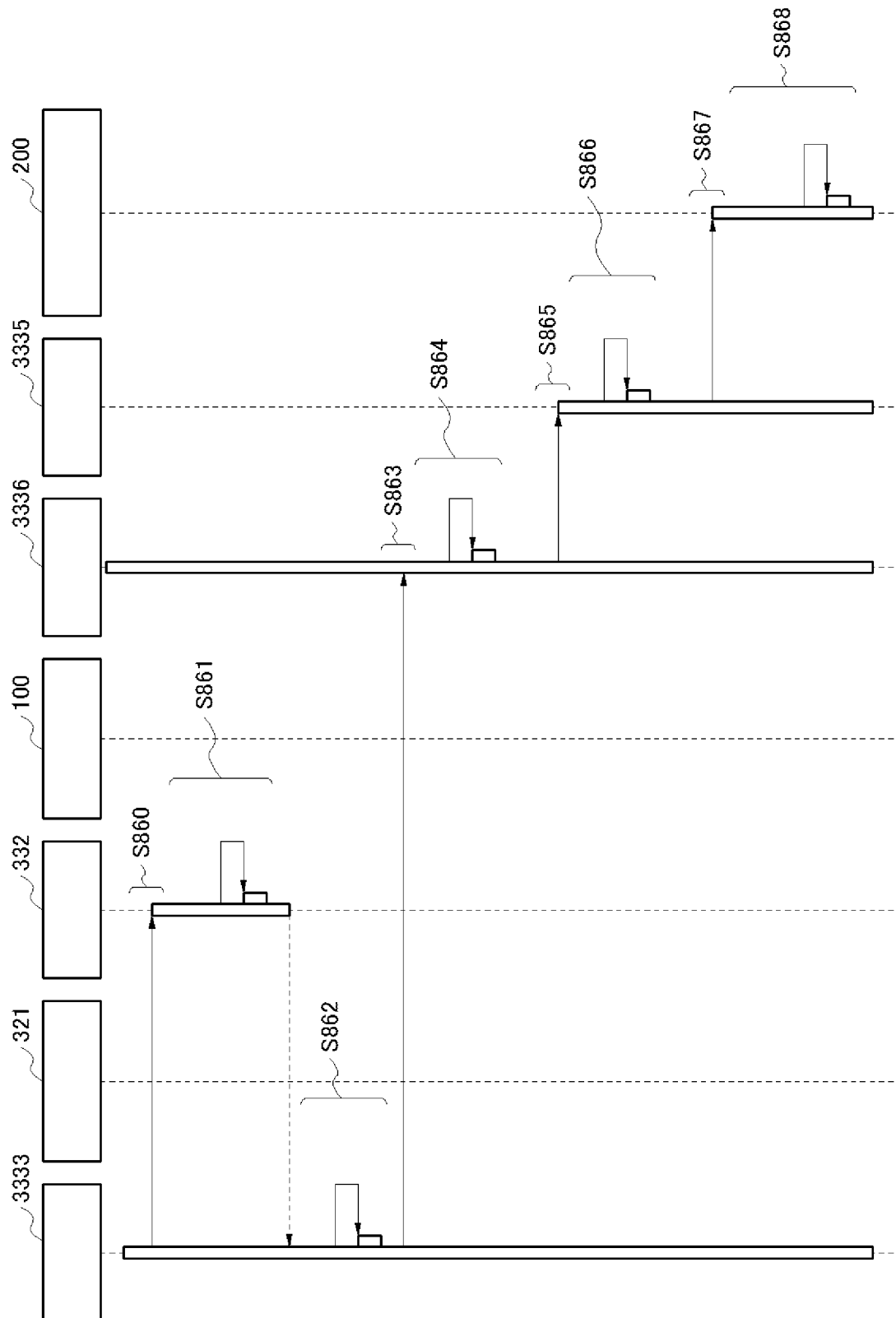
[図18A]



[図18B]



[図18C]



[図19]

162 データセットsp

氏名	年齢	2011年度診療月	病名
患者A	21~24	4月	病名A
患者B	21~24	5月	病名B
患者C	21~24	6月	病名C
患者D	31~40	7月	病名C
患者E	31~40	8月	病名B
患者F	31~40	12月	病名D
患者G	41~51	12月	病名D
患者H	41~51	12月	病名B
患者I	41~51	1月	病名C
患者J	52~58	2月	病名B
患者K	52~58	3月	病名C
患者L	52~58	9月	病名D

161 データレコードrp

[図20]

163 データセットsp



氏名	年齢	2011年度診療月	病名
患者A	21~24	4月~6月	病名A
患者B	21~24	4月~6月	病名B
患者C	21~24	4月~6月	病名C
患者D	31~40	7月~9月	病名C
患者E	31~40	7月~9月	病名B
患者L	52~58	7月~9月	病名D
患者F	31~40	12月	病名D
患者I	41~51	1月~3月	病名C
患者G	41~51	12月	病名D
患者H	41~51	12月	病名B
患者K	52~58	1月~3月	病名C
患者J	52~58	1月~3月	病名B

161 データレコードrp



[図21]

164 データセットsp

氏名	年齢	2011年度診療月	病名
患者A	21～24	4月～6月	病名A
患者B	21～24	4月～6月	病名B
患者C	21～24	4月～6月	病名C
患者D	31～58	7月～9月	病名D
患者E	31～58	7月～9月	病名B
患者L	31～58	7月～9月	病名D
患者F	31～51	12月	病名D
患者G	31～51	12月	病名D
患者H	31～51	12月	病名B
患者J	41～58	1月～3月	病名B
患者I	41～58	1月～3月	病名C
患者K	41～58	1月～3月	病名C

161 データレコードrp

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/004624

A. CLASSIFICATION OF SUBJECT MATTER

G06F21/60(2013.01) i, G06F17/30(2006.01) i, G06Q50/22(2012.01) 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, G06F17/30, G06Q50/22

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	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

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	JP 2012-3440 A (KDDI Corp.), 05 January 2012 (05.01.2012), paragraphs [0032] to [0039], [0048] to [0062] (Family: none)	1-12
A	WO 2011/145401 A1 (Hitachi, Ltd.), 24 November 2011 (24.11.2011), paragraphs [0015] to [0022], [0081] to [0099]; fig. 11, 12 & EP 2573699 A1 & CN 102893553 A	1-12
A	JP 2012-22315 A (NEC (China) Co., Ltd.), 02 February 2012 (02.02.2012), paragraphs [0019] to [0025] & CN 102314565 A	1-12

☐ 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
16 August, 2013 (16.08.13)

Date of mailing of the international search report
27 August, 2013 (27.08.13)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06F21/60(2013.01)i, G06F17/30(2006.01)i, G06Q50/22(2012.01)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（I P C））

Int.Cl. G06F21/60, G06F17/30, G06Q50/22

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 3 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 3 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 3 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	JP 2012-3440 A（K D D I 株式会社） 2012.01.05, 段落【0032】-【0039】，【0048】-【0062】（No Family）	1-12
A	WO 2011/145401 A1（株式会社日立製作所） 2011.11.24, 段落[0015]-[0022], [0081]-[0099], 図 11, 12 & EP 2573699 A1 & CN 102893553 A	1-12

☒ 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

1 6 . 0 8 . 2 0 1 3

国際調査報告の発送日

2 7 . 0 8 . 2 0 1 3

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官（権限のある職員）

金木 陽一

5 S

4 8 7 6

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 4 6

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 2012-22315 A (エヌイーシー (チャイナ) カンパニー, リミテッ ド) 2012.02.02, 段落【0019】～【0025】 & CN 102314565 A	1-12