

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4185399号
(P4185399)

(45) 発行日 平成20年11月26日(2008.11.26)

(24) 登録日 平成20年9月12日(2008.9.12)

(51) Int.Cl. F 1
G 0 6 Q 10/00 (2006.01) G 0 6 F 17/60 1 7 2

請求項の数 4 (全 14 頁)

<p>(21) 出願番号 特願2003-145473 (P2003-145473) (22) 出願日 平成15年5月22日(2003.5.22) (65) 公開番号 特開2004-348489 (P2004-348489A) (43) 公開日 平成16年12月9日(2004.12.9) 審査請求日 平成17年8月5日(2005.8.5)</p>	<p>(73) 特許権者 000004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号 (74) 代理人 100077481 弁理士 谷 義一 (74) 復代理人 100144244 弁理士 成毛 敏明 (74) 代理人 100088915 弁理士 阿部 和夫 (72) 発明者 岩瀬 成人 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内 審査官 田中 伸次</p>
---	--

最終頁に続く

(54) 【発明の名称】 顧客データ管理装置、顧客データ管理方法および顧客データ管理用プログラムならびに顧客データ管理用プログラムを格納した記録媒体

(57) 【特許請求の範囲】

【請求項1】

顧客の名義および住所を含む顧客情報を格納する入力データベースおよび登録済みの顧客情報を格納する名寄せデータベースの双方に接続され、

前記入力データベースに格納された顧客情報中の住所情報を解析して、都道府県から番地・号までを住所コードに変換し、住所情報中にビル名および部屋番号がある場合には、前記住所コードとともに該ビル名および該部屋番号を付加した顧客照合用住所データを出力する住所クレンジング手段と、

顧客情報中の名義情報を解析し、名義の種別が個人の場合には、名義を姓と名とに分割し、異体字を1つの文字に統一し、濁音を清音化して顧客照合用名義データを生成し、名義の種別が法人の場合には、名義の小文字を大文字に変換して解析し、変換後の名義と、該変換後の名義中の主要名義および固有名義とを抽出して顧客照合用名義データを生成し、前記名義の種別とともに前記顧客照合用名義データを出力する名義クレンジング手段と、

前記住所クレンジング手段および前記名義クレンジング手段から出力される顧客データに基づいて前記名寄せデータベースに格納されている顧客情報を検索して照合データを絞り込むためのフィルタリング手段と、

前記顧客データと前記絞り込まれた照合データとの間で、住所情報および名義それぞれの一一致度を算出し、予め定められている照合判定ルールにより前記それぞれの一一致度から前記顧客データの一一致度を判断するマッチング手段と、

前記一致度に応じて新規顧客の顧客データと判断された場合に当該顧客データを前記名寄せデータベースに新規登録する名寄せデータ更新手段とを有する顧客データ管理装置において、

前記フィルタリング手段は、

検索条件対応テーブルを格納する検索条件対応テーブル格納部と、

前記名義クレンジング手段により出力された前記名義の種別と、前記住所クレンジング手段により出力された前記顧客照合用住所データの精度とをキーとして、前記検索条件対応テーブルを検索し、検索条件を決定して、該検索条件で前記名寄せデータベースから抽出した照合対象を出力するフィルタリング部とを有し、

前記検索条件テーブルは、

名義の種別と、顧客データの住所の精度と、顧客データの検索条件とを対応付けて有しており、

前記名義の種別が個人の場合、前記顧客データの住所の精度として、(A)「字丁目」までの記載または「字丁目」より詳しい記載があるときには、該顧客データの字丁目までの住所コードを検索条件とし、該顧客データの住所の精度として、(B)「町大字」までの記載があるときには、該顧客データの町大字までの住所コードと姓とを検索条件とし、該顧客データの住所の精度が、前記(A)でなく、かつ前記(B)でないときには、該顧客データの姓と名とを検索条件とし、

前記名義の種別が法人の場合、前記顧客データの住所の精度として、(C)「町大字」までの記載または「町大字」より詳しい記載があるときには、該顧客データの町大字までの住所コードを検索条件とし、該顧客データの住所の精度として、(D)「市区」までの記載があるときには、該顧客データの市区までの住所コードと固有名義の先頭1文字とを検索条件とし、該顧客データの住所の精度が、前記(C)でなく、かつ前記(D)でないときには、該顧客データの固有名義を検索条件とすることを定義したテーブルである

ことを特徴とする顧客データ管理装置。

【請求項2】

入力データベースに格納される顧客の名義および住所を含む顧客情報と、名寄せデータベースに格納される登録済みの顧客情報とに基づいて顧客の名寄せを行う顧客データ管理方法であって、

住所クレンジング手段が、前記入力データベースに格納された顧客情報中の住所情報を解析して、都道府県から番地・号までを住所コードに変換し、住所情報中にビル名および部屋番号がある場合には、前記住所コードとともに該ビル名および該部屋番号を付加した顧客照合用住所データを出力する住所クレンジングステップと、

名義クレンジング手段が、顧客情報中の名義情報を解析し、名義の種別が個人の場合には、名義を姓と名とに分割し、異体字を1つの文字に統一し、濁音を清音化して顧客照合用名義データを生成し、名義の種別が法人の場合には、名義の小文字を大文字に変換して解析し、変換後の名義と、該変換後の名義中の主要名義および固有名義とを抽出して顧客照合用名義データを生成し、前記名義の種別とともに前記顧客照合用名義データを出力する名義クレンジングステップと、

フィルタリング手段が、前記住所クレンジングステップおよび前記名義クレンジングステップの結果として出力される顧客データに基づいて前記名寄せデータベースに格納されている顧客情報を検索して照合データを絞り込むためのフィルタリングステップと、

マッチング手段が、前記顧客データと前記絞り込まれた照合データとの間で、住所情報および名義それぞれの一致度を算出し、予め定められている照合判定ルールにより前記それぞれの一致度から前記顧客データの一致度を判断するマッチングステップと、

名寄せデータ更新手段が、前記一致度に応じて新規顧客の顧客データと判断された場合に当該顧客データを前記名寄せデータベースに新規登録する名寄せデータ更新ステップとを有し、

前記フィルタリングステップでは、

前記フィルタリング手段が、前記名義クレンジングステップで出力された前記名義の

10

20

30

40

50

種別と、前記住所クレンジングステップで出力された前記顧客照合用住所データの精度とをキーとして、検索条件対応テーブル格納部に格納された検索条件対応テーブルを検索し、検索条件を決定して、該検索条件で前記名寄せデータベースから抽出した照合対象を出力し、

前記検索条件テーブルは、

名義の種別と、顧客データの住所の精度と、顧客データの検索条件とを対応付けて有しており、

前記名義の種別が個人の場合、前記顧客データの住所の精度として、(A)「字丁目」までの記載または「字丁目」より詳しい記載があるときには、該顧客データの字丁目までの住所コードを検索条件とし、該顧客データの住所の精度として、(B)「町大字」までの記載があるときには、該顧客データの町大字までの住所コードと姓とを検索条件とし、該顧客データの住所の精度が、前記(A)でなく、かつ前記(B)でないときには、該顧客データの姓と名とを検索条件とし、

10

前記名義の種別が法人の場合、前記顧客データの住所の精度として、(C)「町大字」までの記載または「町大字」より詳しい記載があるときには、該顧客データの町大字までの住所コードを検索条件とし、該顧客データの住所の精度として、(D)「市区」までの記載があるときには、該顧客データの市区までの住所コードと固有名義の先頭1文字とを検索条件とし、該顧客データの住所の精度が、前記(C)でなく、かつ前記(D)でないときには、該顧客データの固有名義を検索条件とすることを定義したテーブルである

ことを特徴とする顧客データ管理方法。

20

【請求項3】

請求項1に記載の顧客データ管理装置としてコンピュータを機能させるための顧客データ管理用プログラム。

【請求項4】

請求項3に記載の顧客データ管理用プログラムを格納した、コンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、顧客データ管理装置、顧客データ管理方法および顧客データ管理用プログラムならびに顧客データ管理用プログラムを格納した記録媒体に関する。より詳細には、本発明は、大規模顧客データを高速・高精度に名寄せすることを可能とする顧客データ管理装置、顧客データ管理方法および顧客データ管理用プログラムならびに顧客データ管理用プログラムを格納した記録媒体に関する。

30

【0002】

【従来の技術】

従来の名寄せシステムにおいては、大規模な顧客データの照合を行なう場合、例えば特許文献1に示すように顧客情報の各項目ごとに一致不一致を判断し、RDB(Relational Data Base)の完全一致検索機能を利用した照合を行なっている。また、取扱う顧客データがそれほど大規模でなく、かつ、高い精度での照合が必要な場合には、対象である顧客データを一度データ・ファイルにダウンロードしたうえで、文字列一致率や単語一致率などにより文字列の一部一致を点数付けして照合を行っている。

40

【0003】

【特許文献1】

特開昭63-282838号公報

【0004】

【発明が解決しようとする課題】

しかしながら、このような従来の名寄せ方法には、以下のような問題があった。すなわち、RDBを用いて文字列一致率や単語一致率などによる照合には一般に長時間を必要とし、大規模な顧客データについて迅速な照合を行なうことは困難であった。また、顧客情報

50

項目としてある「名義」および「住所」の双方の照合度を考慮して最終的な一致度を求めるようなことはRDBを用いた照合方法では不可能である。

【0005】

一方、顧客データをファイル化することにより高度な照合が可能となるが、データのダウンロードそのものに時間を要することに加え、このような顧客データも最終的にはRDB化されるものであり顧客データのRDB化にも長時間が必要とされる。また、取扱うべき顧客データが大規模な場合には、照合やソートにも時間がかかることとなり迅速な照合を求められている現実的なニーズに応えることは難しい。

【0006】

さらに、顧客データをファイル化した場合には、照合ルールの変更に柔軟に対応することが容易ではない。例えば、「名義」の種別が法人ならば「住所」を番地まで絞り込むこととするルールがある場合に、「住所」が町名まで解析できずその結果番地までの絞り込みが不可能なケースも発生し得る。このような場合、「住所」と「名義」とで絞り込むといったようなルールに変更すると、ファイル化した顧客データを利用することが困難となってしまう。

【0007】

本発明はこのような問題に鑑みてなされたもので、その目的とするところは、大規模顧客データを高速・高精度に名寄せすることを可能とする顧客データ管理装置、顧客データ管理方法および顧客データ管理用プログラムならびに顧客データ管理用プログラムを格納した記録媒体を提供することにある。

【0008】

【課題を解決するための手段】

本発明は、このような目的を達成するために、請求項1に記載の発明は、顧客の名義および住所を含む顧客情報を格納する入力データベースおよび登録済みの顧客情報を格納する名寄せデータベースの双方に接続され、前記入力データベースに格納された顧客情報中の住所情報を解析して、都道府県から番地・号までを住所コードに変換し、住所情報中にビル名および部屋番号がある場合には、前記住所コードとともに該ビル名および該部屋番号を付加した顧客照合用住所データを出力する住所クレンジング手段と、顧客情報中の名義情報を解析し、名義の種別が個人の場合には、名義を姓と名とに分割し、異体字を1つの文字に統一し、濁音を清音化して顧客照合用名義データを生成し、名義の種別が法人の場合には、名義の小文字を大文字に変換して解析し、変換後の名義と、該変換後の名義中の主要名義および固有名義とを抽出して顧客照合用名義データを生成し、前記名義の種別とともに前記顧客照合用名義データを出力する名義クレンジング手段と、前記住所クレンジング手段および前記名義クレンジング手段から出力される顧客データに基づいて前記名寄せデータベースに格納されている顧客情報を検索して照合データを絞り込むためのフィルタリング手段と、前記顧客データと前記絞り込まれた照合データとの間で、住所情報および名義それぞれの一致度を算出し、予め定められている照合判定ルールにより前記それぞれの一致度から前記顧客データの一致度を判断するマッチング手段と、前記一致度に応じて新規顧客の顧客データと判断された場合に当該顧客データを前記名寄せデータベースに新規登録する名寄せデータ更新手段とを有する顧客データ管理装置において、前記フィルタリング手段は、検索条件対応テーブルを格納する検索条件対応テーブル格納部と、前記名義クレンジング手段により出力された前記名義の種別と、前記住所クレンジング手段により出力された前記顧客照合用住所データの精度とをキーとして、前記検索条件対応テーブルを検索し、検索条件を決定して、該検索条件で前記名寄せデータベースから抽出した照合対象を出力するフィルタリング部とを有し、前記検索条件テーブルは、名義の種別と、顧客データの住所の精度と、顧客データの検索条件とを対応付けて有しており、前記名義の種別が個人の場合、前記顧客データの住所の精度として、(A)「字丁目」までの記載または「字丁目」より詳しい記載があるときには、該顧客データの字丁目までの住所コードを検索条件とし、該顧客データの住所の精度として、(B)「町大字」までの記載があるときには、該顧客データの町大字までの住所コードと姓とを検索条件とし、該顧客

10

20

30

40

50

データの住所の精度が、前記（Ａ）でなく、かつ前記（Ｂ）でないときには、該顧客データの姓と名とを検索条件とし、前記名義の種別が法人の場合、前記顧客データの住所の精度として、（Ｃ）「町大字」までの記載または「町大字」より詳しい記載があるときには、該顧客データの町大字までの住所コードを検索条件とし、該顧客データの住所の精度として、（Ｄ）「市区」までの記載があるときには、該顧客データの市区までの住所コードと固有名義の先頭１文字とを検索条件とし、該顧客データの住所の精度が、前記（Ｃ）でなく、かつ前記（Ｄ）でないときには、該顧客データの固有名義を検索条件とすることを定義したテーブルであることを特徴とする。

【 0 0 0 9 】

請求項２に記載の発明は、入力データベースに格納される顧客の名義および住所を含む顧客情報と、名寄せデータベースに格納される登録済みの顧客情報とに基づいて顧客の名寄せを行う顧客データ管理方法であって、住所クレンジング手段が、前記入力データベースに格納された顧客情報中の住所情報を解析して、都道府県から番地・号までを住所コードに変換し、住所情報中にビル名および部屋番号がある場合には、前記住所コードとともに該ビル名および該部屋番号を付加した顧客照合用住所データを出力する住所クレンジングステップと、名義クレンジング手段が、顧客情報中の名義情報を解析し、名義の種別が個人の場合には、名義を姓と名とに分割し、異体字を１つの文字に統一し、濁音を清音化して顧客照合用名義データを生成し、名義の種別が法人の場合には、名義の小文字を大文字に変換して解析し、変換後の名義と、該変換後の名義中の主要名義および固有名義とを抽出して顧客照合用名義データを生成し、前記名義の種別とともに前記顧客照合用名義データを出力する名義クレンジングステップと、フィルタリング手段が、前記住所クレンジングステップおよび前記名義クレンジングステップの結果として出力される顧客データに基づいて前記名寄せデータベースに格納されている顧客情報を検索して照合データを絞り込むためのフィルタリングステップと、マッチング手段が、前記顧客データと前記絞り込まれた照合データとの間で、住所情報および名義それぞれの一致度を算出し、予め定められている照合判定ルールにより前記それぞれの一致度から前記顧客データの一致度を判断するマッチングステップと、名寄せデータ更新手段が、前記一致度に応じて新規顧客の顧客データと判断された場合に当該顧客データを前記名寄せデータベースに新規登録する名寄せデータ更新ステップとを有し、前記フィルタリングステップでは、前記フィルタリング手段が、前記名義クレンジングステップで出力された前記名義の種別と、前記住所クレンジングステップで出力された前記顧客照合用住所データの精度とをキーとして、検索条件対応テーブル格納部に格納された検索条件対応テーブルを検索し、検索条件を決定して、該検索条件で前記名寄せデータベースから抽出した照合対象を出力し、前記検索条件テーブルは、名義の種別と、顧客データの住所の精度と、顧客データの検索条件とを対応付けて有しており、前記名義の種別が個人の場合、前記顧客データの住所の精度として、（Ａ）「字丁目」までの記載または「字丁目」より詳しい記載があるときには、該顧客データの字丁目までの住所コードを検索条件とし、該顧客データの住所の精度として、（Ｂ）「町大字」までの記載があるときには、該顧客データの町大字までの住所コードと姓とを検索条件とし、該顧客データの住所の精度が、前記（Ａ）でなく、かつ前記（Ｂ）でないときには、該顧客データの姓と名とを検索条件とし、前記名義の種別が法人の場合、前記顧客データの住所の精度として、（Ｃ）「町大字」までの記載または「町大字」より詳しい記載があるときには、該顧客データの町大字までの住所コードを検索条件とし、該顧客データの住所の精度として、（Ｄ）「市区」までの記載があるときには、該顧客データの市区までの住所コードと固有名義の先頭１文字とを検索条件とし、該顧客データの住所の精度が、前記（Ｃ）でなく、かつ前記（Ｄ）でないときには、該顧客データの固有名義を検索条件とすることを定義したテーブルであることを特徴とする。

【 0 0 1 0 】

請求項３に記載の発明は、顧客データ管理用プログラムであって、当該顧客データ管理用プログラムにより、請求項１に記載の顧客データ管理装置としてコンピュータを機能させることを特徴とする。

10

20

30

40

50

【 0 0 1 1 】

請求項 4 に記載の発明は、コンピュータ読み取り可能な記録媒体であって、請求項 3 に記載の顧客データ管理用プログラムを格納していることを特徴とする。

【 0 0 2 1 】

【発明の実施の形態】

以下に、図面を参照して本発明の実施の形態について説明する。

【 0 0 2 2 】

図 1 は本発明の顧客管理システムを説明するための図であり、図 2 は図 1 のシステムで用いる本発明の顧客データ管理方法を説明するための図である。図 1 において、10 は本発明の顧客データ管理方法を実行するコンピュータであり、本発明を実行するためのプログラムが格納されている。11 は入力された顧客データを格納する入力 DB、12 は入力 DB に格納されている顧客データの中から処理されるべきデータを抽出するためのデータ抽出部、13 は「住所」および「名義」以外の顧客データ（例えば、「日付」や「電話番号」など）を変換して正規化するためのデータ変換部である。また、14 は住所照合に必要な各データを作成する住所クレンジング部、15 は名義照合に必要な各データを作成する名義クレンジング部、16 は住所クレンジング部 14 および名義クレンジング部 15 によるクレンジング結果を基にして照合すべき顧客データの検索条件を決定して既に登録されている顧客のデータを格納している名寄せ DB 18 を検索するフィルタリング部である。なお、入力 DB 11 および名寄せ DB 18 は、コンピュータ 10 とインターネットなどの通信網で接続されたものであってもよいことはいうまでもない。

【 0 0 2 3 】

17 は予め定められた種々の検索条件がテーブル化されて格納されている検索条件対応テーブル格納部であり、18 は登録済み顧客の顧客データを格納する名寄せ DB である。また、19 は照合ルール対応テーブル格納部であり、顧客判定ルールや住所照合ルールあるいは名義照合ルールなどの様々な照合ルールがテーブルとして格納されている。20 はマッチング部で、照合ルール対応テーブル格納部 19 に格納されている照合ルールにしたがって、入力 DB 11 に顧客データが登録された顧客と名寄せ DB 18 に顧客データが登録されている顧客とを照合し、新たに入力された顧客データが既に登録済みの顧客か否かを判断する。さらに、21 は名寄せ DB 更新部であり、この名寄せ DB 更新部 21 は、入力 DB 11 に入力されて照合対象とされた顧客データが既に登録済みの顧客のものではないとマッチング部 19 により判断された場合にその顧客データを名寄せ DB 18 に格納して新規に顧客登録する。

【 0 0 2 4 】

この顧客管理システムを用いて顧客データ管理を実行する際の処理フローは以下のとおりである。まず、外部から入力（登録）された顧客データを格納した入力 DB 11 から顧客データが抽出される（S21）。この顧客データとは、例えば、「種別」は個人、「名義」は木村太郎、「住所」は（東京都）足立区千住 1 - 2 - 3 ハイツ北千住 508、「生年月日」は昭和 12 年 1 月 23 日、「電話番号」は 03 - ABC - DEFG、「日付」は平成 15 年 5 月 12 日、などの所定の項目を備えている。必要により、この抽出データのなかの「住所」および「名義」以外の顧客データをデータ変換して正規化を行なう（S22）。

【 0 0 2 5 】

次に、顧客データのうちの「住所」および「名義」についてのデータのクレンジングを実行する（S23 および S24）。住所クレンジング（S23）は住所の照合のために行なうもので、都道府県から番地・号までの住所コード、ならびに、ビル名および部屋番号が出力される。例えば、住所が「（東京都）足立区千住 1 - 2 - 3 ハイツ北千住 508」である場合には、都道府県から番地・号までの「（東京都）足立区千住 1 - 2 - 3」に対応する住所コード「13 / 121 / 040 / 001 / 002 / 003」と、ビル名に対応する「ハイツ北千住」と、部屋番号に対応する「508」とが出力される。

【 0 0 2 6 】

名義クレンジング (S 2 4) は、顧客が個人か法人かの判断を行うと同時に名義照合のための各種データを作成するための処理である。個人の場合には、例えば、姓と名を分割し、異体字や清音化などの揺らぎを正規化した姓および名を作成する。例えば、「木村太郎」は「木村」という姓と「太郎」という名に分割される。また、「澤田一郎」の場合は「澤田」という姓と「一郎」という名に分割されるが、姓の「澤田」は「沢田」にクレンジングする如くである。また、法人の場合には、名義を会社名・支店名・部門名等に分割したうえで、名義の照合用データである揺らぎを削除するクレンジングを行なって、揺らぎ削除名義、主要名義、固有名義を定める。例えば、入力名義が「焼肉屋ジャンジャン」である場合には、名義の文字列中の小文字を大文字に変換することで揺らぎを削除して揺らぎ削除名義を「焼肉屋ジャンジャン」とし、主要名義を「ジャンジャン」、固有名義も「ジャンジャン」とするといった具合である。

10

【 0 0 2 7 】

図 3 は、このような名義クレンジングと住所クレンジングを実行した結果の例を説明するための図で、既に説明した名義のほか、例えば入力名義が「レストラン若菜」のケースでは、名義クレンジングの結果として、「種別」が法人、「会社名」がレストラン若菜、「主要名義」が若菜、「固有名義」が若菜とされ、住所クレンジングの結果として、都道府県から番地・号までの「(東京都)荒川区南千住 1 - 1 0 - 1」に対応する住所コード「1 3 / 1 1 8 / 0 0 7 / 0 0 1 / 0 1 0 / 0 0 0 1」と、ビル名に対応する「小林ビル 1 F」とが出力される。

【 0 0 2 8 】

20

次に、住所クレンジング (S 2 3) 及び名義クレンジング (S 2 4) の結果を基に、検索条件対応テーブル 1 7 を参照して名寄せ D B 1 8 に格納されている顧客データを検索するための検索条件を作成する (S 2 5)。図 4 は、検索条件対応テーブルの項目内容を説明するための一例であり、この例では「名義の種類」、「住所の必要精度」および「検索条件」の項目に分類されている。「名義の種類」としては「法人」か「個人」かが分類され、「住所の必要精度」としては「町大字」、「市区」、「字丁目」、「町大字」などと分類される。例えば法人名義で「住所の必要精度」が「市区」の場合には、顧客データの検索条件は「住所(市区)」と「固有名義の先頭 1 文字」との和とされる。なお、「住所の必要精度」を定義せず、法人名義の固有名義のみ若しくは個人名義の姓および名のみを検索条件とすることも可能である。これらの検索条件の設定に際しては、あらかじめ、検索対象である顧客データの登録状況を調べておき、データベース中の顧客データをフィルタリングして抽出される登録顧客(の顧客データ)が、漏れなく且つ絞り込み充分となる条件にしておく。

30

【 0 0 2 9 】

S 2 6 はフィルタリングの工程であり、住所クレンジング部 1 4 および名義クレンジング部 1 5 によるクレンジング結果を基にして照合すべき顧客データの検索条件を決定して既に登録されている顧客のデータを格納している名寄せ D B 1 8 を検索する。

【 0 0 3 0 】

たとえば、全国の法人名義の顧客データの検索を行なう場合には、その法人の住所が「町大字」まで知られていれば「固有名義」を検索条件に含めなくても十分な絞り込みが可能であるが、「市区」までの住所しか判らないときには固有名義も検索条件に繰り込まない限り十分な絞り込みができない。そのような場合には「住所(市区) + 固有名義の先頭 1 文字」を検索条件とすることとなる。また、通常の名寄せではなく、企業単位で名寄せするような場合には、大企業は全国に支店を有しているために「住所」を検索条件として用いることはできず、その法人の固有名義そのものを用いて全国検索する。同様に、個人名義の顧客データの検索を行う場合には、住所の情報として「字丁目」まであれば十分な絞り込みが可能であるが、「町大字」までの住所しか判らない場合には姓などの条件が必要となり、「住所(町大字) + 姓」を検索条件とすることになる。

40

【 0 0 3 1 】

この検索条件の決定について図 3 に示したクレンジング後の顧客データを例にとって説

50

明すると、名義「木村太郎」の場合は個人名義の顧客データなので、「足立区千住1-2-3」の個人データを検索する。また、名義「レストラン若菜」の場合は法人名義のデータなので、「荒川区南千住」の法人データを検索する。名義「須田総合家具センター」の場合は法人名義のデータであるが、住所が東京都の中央区なのか大阪府の中央区なのか確定されないので、検索条件として住所は使わずに固有名義「須田」のみを検索することとなる。

【0032】

次に、名義及び住所のクレンジング結果から照合ルールを定め(S27)、マッチングを実行する(S28)。

【0033】

図5は一般的な照合ルールの例を説明するための図で、この図中で「確定一致」とあるのは入力DB11に格納されている照合対象とされた顧客データを、既に登録済みの名寄せDB18に格納された顧客データと照合して新規な顧客か既に登録済みの顧客かをチェック(ユーザチェック)する必要がないほどの精度で一致していることを意味しており、この場合には対象顧客データは既に登録済みの顧客のものであると判断されることとなる。「曖昧一致」とあるのは一応は「一致データ」の候補ではあるが安全のためにユーザチェックが必要であると判断されたことを意味している。

【0034】

例えば、「住所照合ルール」として、「号」までが一致している場合には住所一致度を95、「番地」までが一致している場合には住所一致度を90、「丁目」までが一致している場合には一致度を80としたり、「名義照合ルール」として、「文字列一致」の場合には一致した文字の割合に100を乗じて名義一致度とし、「単語一致」の場合には一致した単語の割合に100を乗じて名義一致度とする。そして、このようなルールに基づいて「一致度」を求め、「名義」の一致度が90以上かつ「住所」は「番地」まで一致しておりその一致度が90以上ならば「確定一致」とし、「住所」が「番地」まで一致しており一致度は90以上であっても「名義」の一致度が80以上90未満の場合には「曖昧一致」とする。逆に、「名義」一致度が90以上であっても「住所」が「字丁目」レベルの一致しかなく一致度が80以上90未満の場合にも「曖昧一致」とみなす。

【0035】

最後に、照合対象である入力データが名寄せDB18に格納されている照合データと一致しない場合、すなわち新規の顧客データである場合には、その入力データを新規顧客のデータとして名寄せDBに登録する(S29)。一方、「確定一致」の場合にはその入力データは既に登録済みの顧客のデータであるため削除される。なお、「曖昧一致」と判断された入力データは、一旦は名寄せDB18に登録されるが、その後に入手による詳細なチェックを受け、新規顧客と判断されればそのまま名寄せDB18に格納され、既登録顧客と判断されれば削除される。

【0036】

上述した照合ルールは、必要に応じて変更することも可能である。図6は、個人名義の照合ルールと法人名義の照合ルールとを名寄せ種類別にテーブル化した例である。名義のマッチングを判断するに際しては、名義が法人名ならば主要語や固有名での照合も考えられるが、名義が個人名である場合は姓と名の照合のみである。また、照合すべき住所の範囲も、法人ならば個人よりも件数が少ないので住所の照合度を緩めてもかまわない。入力データ全体が個人または法人の顧客データであれば、本発明の顧客管理システムの起動時に照合ルールを選択すればよいが、個人データと法人データとが混在して入力されてこれらの混在データが照合対象とされる場合には、システムを稼働させた状態で動的に照合ルールを変更する必要が生じる。

【0037】

そこで、図7に示すように、検索条件に対応付けられた照合ルールを検索条件対応テーブルの項目として追加し、名義の種別や住所の精度に応じてどの照合ルールを用いるかを判断してデータ照合を行うこととすることができる。なお、この図の照合ルール項目中の「

10

20

30

40

50

法人寄せ2」および「個人寄せ2」の照合ルールにおいては、フィルタリングの結果として「住所」が「市区」または「丁目」まで一致していることが明らかなので、図6中に示した具体的な照合ルールでは住所に関する照合条件を設ける必要がない。また、「住所」データは「市区」または「丁目」のレベルまでしかないので、照合結果は「曖昧一致」として、後で人手などによる住所データの修正を行うようにしている。

【0038】

なお、本システムのOSをマルチプロセス可能なOSとした場合には、このような照合ルールの変更だけでなく照合作業のさらなる高速化が実現できる。具体例としては図8に示すように、名寄せDB18を個人DB18aと法人DB18bとで構成するようにし、個人DB18aには個人名義の顧客データを格納し、法人DB18bには法人名義の顧客データを格納するようにする。この場合、マッチング部20を個人マッチング部20aと法人マッチング部20bとに分割し、これらの個人マッチング部20aと法人マッチング部20bの各々に、個人寄セルールテーブル20a'と法人寄セルールテーブル20b'を備えてそれぞれの名寄せルールに則ってマッチングさせるように構成し、これらを個人または法人専用とする。そして、住所・名義クレンジング等がなされた入力データが個人名義のものであるか法人名義のものであるかに応じて何れかの処理ルートを選択して照合処理の実行を行なうようにすればよい。

【0039】

また、マッチングの手順を以下のように構成することとしてもよい。すなわち、マッチング部20を、例えば図9に示すように、確定一致マッチング部20cと第1の曖昧一致マッチング部20dおよび第2の曖昧一致マッチング部20eを設け、確定一致マッチング部20cには確定一致ルールテーブル20fを、第1の曖昧一致マッチング部20dおよび第2の曖昧一致マッチング部20eには曖昧一致ルールテーブル20gを備えるように構成する。なお、図9では曖昧マッチング部を2つ備える構成としたが、曖昧一致マッチング部は1つとしてもよく或いは3つ以上としてもよい。

【0040】

マッチング部20をこのように構成するメリットは以下のようなものである。例えば、顧客数がある程度飽和して、入力される顧客データの80%程度は既存の顧客になったような場合には、確定一致とみなせる照合対象である入力データは50%以上あることが期待できる。また、確定一致レベルにある入力データの照合は照合すべきデータ数も少ないため、曖昧一致レベルにある入力データの照合よりも格段に早く終了する。そこで、確定一致のプロセスと曖昧一致のプロセスを分離し、確定一致とされた入力データの処理が早く終了するようにすることにより、全体の処理が高速化される。

【0041】

具体的には図10に示すフローチャートのように、確定一致マッチング部20cにより確定一致ルールに基づいた確定一致マッチングを行ない(S101)、マッチングしたか否かを判断する(S102)。その結果、確定一致したと判定(S102:Yes)された入力データは既に登録されている顧客のデータであるから既存データとして処理される(S103)。一方、確定一致しない(S102:No)と判定された入力データは取り敢えず新規データとして一旦名寄せDB18に格納・登録され(S104)る。確定一致しない状態で名寄せDB18に格納された入力データは曖昧一致マッチング部により、曖昧一致マッチングルールに則ってその一致度が判断される(S105)。曖昧一致マッチングの結果、入力データは曖昧一致であるものとの判断がなされると(S106:Yes)、その入力データは「曖昧一致データ」として取扱われる。一方、曖昧マッチングの結果、入力データは曖昧一致ではないとの判断がなされると(S106:No)、その入力データは「新規データ」とされて名寄せDB18へ登録すべきデータとして取扱われる(S108)。

【0042】

すなわち、この処理においては取扱う入力データは確定一致でないことが判明した時点で名寄せDB18に格納・登録され、1または複数の曖昧一致マッチングプロセスを実行す

10

20

30

40

50

ることが可能となる。曖昧一致マッチングに要する処理時間が確定一致マッチングの処理時間の10倍を要したり、全体の入力データの半分が確定一致する顧客データであるような場合には、極めて迅速に照合作業を実行することが可能となる。

【0043】

【発明の効果】

以上説明したように、本発明によれば、顧客データ管理装置として動作するコンピュータを入力DBおよび名寄せDBの双方に接続し、顧客の名義と住所とを含む顧客情報のうちの住所情報を住所クレンジング部により解析するとともに、名義の種類および顧客照合用の名義データを名義クレンジング部により解析する。そして、住所クレンジング部および名義クレンジング部から出力される顧客データに基づいて名寄せDBに格納されている顧客情報をフィルタリング部により検索して照合データを絞り込む。顧客データと絞り込まれた照合データとはマッチング部により比較されて一致度が判断され、その一致度に応じて新規顧客の顧客データと判断された場合にはその顧客データを名寄せDBに新規登録することとした。このような構成とすると、顧客データのフィルタリングにより照合すべきデータ数を減らすことができ、金融機関等で取扱われる大規模顧客データを高速・高精度に名寄せすることを可能とする顧客管理システムおよびそれに用いられる顧客データ管理装置を提供することが可能となる。

10

【図面の簡単な説明】

【図1】本発明の顧客管理システムを説明するための図である。

【図2】本発明の顧客データ管理方法を説明するための図である。

20

【図3】入力名義の名義クレンジングおよび住所クレンジング後の名義例を説明するための図である。

【図4】検索条件対応テーブルの項目内容を説明するための一例を説明するための図である。

【図5】一般的な照合ルールの例を説明するための図である。

【図6】個人名義の照合ルールと法人名義の照合ルールとを名寄せ種類別にテーブル化した例を説明するための図である。

【図7】照合ルールを追加した検索条件対応テーブルを説明するための図である。

【図8】個人顧客と法人顧客の照合を別のプロセスとする場合のシステム構成図である。

【図9】確定一致マッチングと曖昧一致マッチングとを分離したマッチング部の構成を説明するための図である。

30

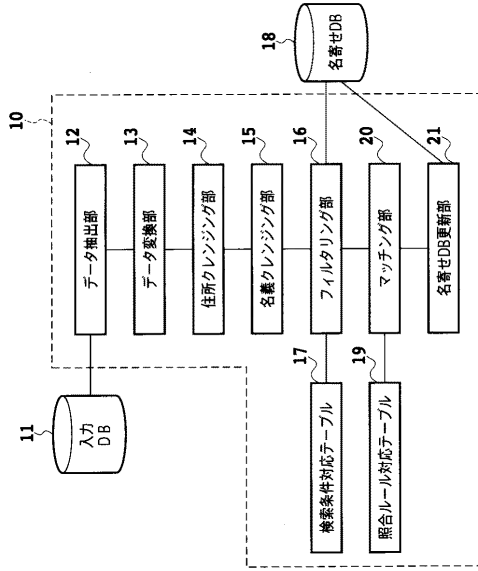
【図10】確定一致マッチングと曖昧一致マッチングとを分離したマッチング工程を説明するための図である。

【符号の説明】

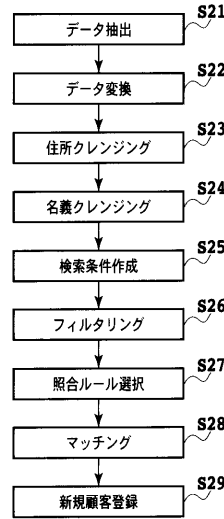
- 10 コンピュータ
- 11 入力DB
- 12 データ抽出部
- 13 データ変換部
- 14 住所クレンジング部
- 15 名義クレンジング部
- 16 フィルタリング部
- 17 検索条件対応テーブル格納部
- 18 名寄せDB
- 19 照合ルール対応テーブル格納部
- 20 マッチング部
- 21 名寄せDB更新部

40

【 図 1 】



【 図 2 】



【 図 3 】

入力名義	入力住所	名義クレンジング結果	住所クレンジング結果
木村太郎	足立区千住1-2-3 ハイイツ北千住508	種別：個人 姓：木村 正规范化姓：木村 名：太郎	13/121/040/001/0 002/0003/508 ビル名：ハイイツ北千住
栗田一郎	荒川区町屋6-43-8	種別：個人 姓：栗田 正规范化姓：栗田 名：一郎	13/118/006/006/00 43/0008
レストラン若菜	荒川区常千住110-1 小林ビルF	種別：法人 会社名：レストラン若菜 住所名義：若菜 住所固有名称：若菜	13/118/007/001/00 10/0001 ビル名：小林ビルF
ABC東日本荒川支店	荒川区荒川6-51-22	種別：法人 会社名：ABC東日本 支店：荒川支店 正式名：東日本ABC株式会社	13/118/001/0006/00 51/0022
須田総合家具センター	中央区日本橋1丁目	種別：法人 住所名義：須田総合家具センター 住所固有名称：須田総合家具	13/102/012/001 27/128/069/001
焼肉屋ジャンジャン	荒川区荒川2-33	種別：法人 会社名：焼肉屋ジャンジャン 住所名義：焼肉屋ジャンジャン 住所固有名称：ジャンジャン	13/118/001/0002/00 33

【 図 4 】

名義の種類	住所の必要精度	検索条件
法人	町大字	住所 (町大字)
	市区	住所 (市区) + 固有名称の先頭1文字
	-	固有名称
個人	字丁目	住所 (字丁目)
	町大字	住所 (町大字) + 姓
	-	姓、名

【 図 5 】

照合判定ルール
 (名義) ≧ 90 & (住所) ≧ 90 → 確定一致
 (名義) ≧ 80 & (住所) ≧ 90 → 曖昧一致
 (名義) ≧ 90 & (住所) ≧ 80 → 曖昧一致

住所照合ルール
 同一一致 → (住所) = 95
 番地一致 → (住所) = 90
 丁目一致 → (住所) = 80

名義照合ルール
 文字列一致 → (名義) = 一致した文字の割合 × 100
 単語一致 → (名義) = 一致した単語の割合 × 100

注：上記で(名義)(住所)は評価値を表す

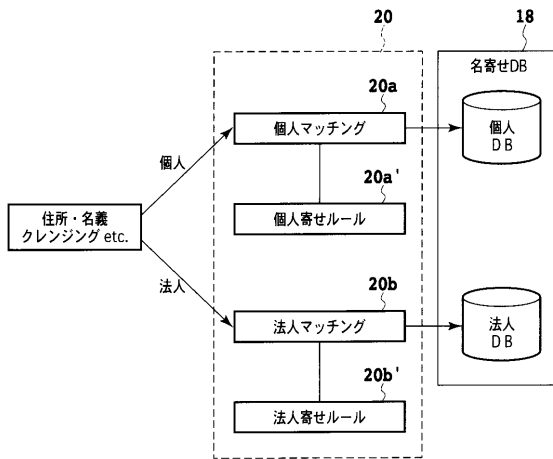
【図 6】

寄せの種類	照合ルール
法人寄せ 1	照合判定ルール (名義) ≧ 90 & (住所) ≧ 90 → 確定一致 (名義) ≧ 80 & (住所) ≧ 90 → 曖昧一致 (名義) ≧ 90 & (住所) ≧ 70 → 曖昧一致 住所照合ルール 号一致 → (住所) = 95 番地一致、号省略 → (住所) = 90 番地一致 → (住所) = 85 丁目一致 → (住所) = 80 町大字一致 → (住所) = 70 名義照合ルール 完全一致 → (名義) = 100 揺らぎ削除一致 → (名義) = 95 主要語一致 → (名義) = 85 固有名一致 → (名義) = 80 文字列一致 → (名義) = 一致した文字の割合 × 100 単語一致 → (名義) = 一致した単語の割合 × 100
法人寄せ 2	照合判定ルール (名義) ≧ 90 → 曖昧一致 名義照合ルール・・・法人寄せ 1 と同じ
法人寄せ 3	照合判定ルール (名義) ≧ 95 → 曖昧一致 名義照合ルール・・・法人寄せ 1 と同じ
個人寄せ 1	照合判定ルール (名義) ≧ 90 & (住所) ≧ 90 → 確定一致 (名義) ≧ 80 & (住所) ≧ 90 → 曖昧一致 (名義) ≧ 90 & (住所) ≧ 80 → 曖昧一致 住所照合ルール 部屋番号一致 → (住所) = 100 号一致 → (住所) = 95 番地一致、号省略 → (住所) = 90 番地一致 → (住所) = 85 丁目一致 → (住所) = 80 名義照合ルール 姓名完全一致 → (名義) = 100 姓名揺らぎ削除一致 → (名義) = 95 姓名読み一致 → (名義) = 90 姓完全一致 → (名義) = 80 姓揺らぎ削除一致 → (名義) = 70
個人寄せ 2	照合判定ルール (名義) ≧ 90 → 曖昧一致 名義照合ルール・・・個人寄せ 1 と同じ
個人寄せ 3	照合判定ルール 電話番号一致 → 曖昧一致

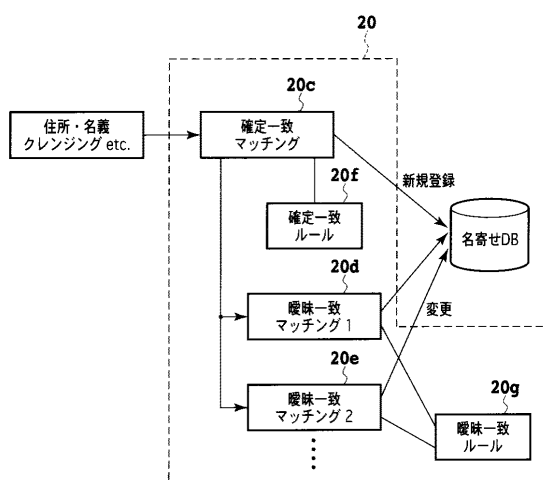
【図 7】

照合ルール	検索条件	住所の精度	名義の種類
法人寄せ 1	住所 (町大字)	町大字	法人
法人寄せ 2	住所 (市区) + 固有名称先頭1文字	市区	法人
法人寄せ 3	固有名称	-	法人
個人寄せ 1	住所 (字丁目)	字丁目	個人
個人寄せ 2	住所 (丁目) + 姓	町大字	個人
個人寄せ 3	姓、名	-	個人

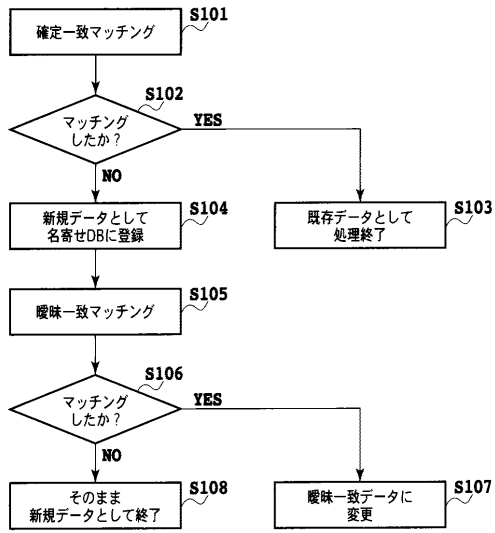
【図 8】



【図 9】



【図10】



フロントページの続き

- (56)参考文献 特開2001-325425(JP,A)
特開平08-297709(JP,A)
データクレンジング・ツール 辞書機能でデータの不統一を解消数百個を超えるDBの統合を支援、ネットワークコンピューティング、日本、株式会社リックテレコム、1999年10月1日、第11巻、第10号、P.86-91
岩瀬成人、外3名、自然言語処理を用いた住所解析方式、第49回(平成6年後期)全国大会講演論文集(3)、社団法人情報処理学会、1994年9月30日、P.3-261~3-262
アグレックス 豊富な日本語辞書機能で実現した名寄せシステムの『トリリアム』 CRMの基本、顧客データを高い精度で一元化する、Computer TELEPHONY、日本、株式会社リックテレコム、2002年4月20日、第5巻、第5号、P.98

(58)調査した分野(Int.Cl., DB名)

G06Q 10/00

JSTPlus(JDreamII)

G-Search