



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0709707-7 A2**

(22) Data de Depósito: 11/05/2007
(43) Data da Publicação: 26/07/2011
(RPI 2116)



* B R P I 0 7 0 9 7 0 7 A 2 *

(51) **Int.Cl.:**
G06F 7/00 2006.01
G06F 17/00 2006.01

(54) **Título: ÍNDICES DE LOCALIDADE E MÉTODO PARA INDEXAR LOCALIDADES**

(30) Prioridade Unionista: 12/05/2006 US 11/433,104

(73) Titular(es): Tele Atlas North America, Inc

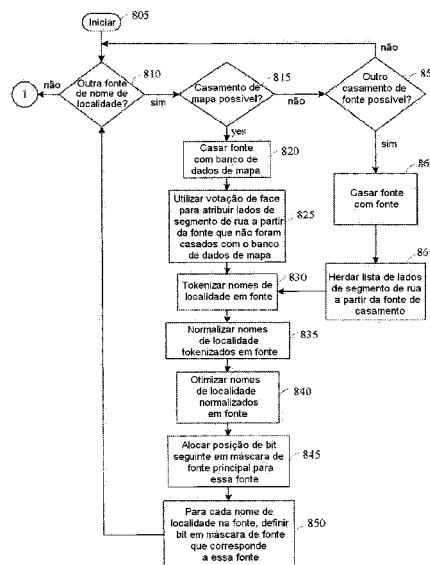
(72) Inventor(es): Michael Geilich

(74) Procurador(es): Nellie Anne Daniel-shores

(86) Pedido Internacional: PCT US2007068805 de 11/05/2007

(87) Publicação Internacional: WO 2007/134249 de 22/11/2007

(57) **Resumo:** ÍNDICES DE LOCALIDADE E MÉTODO PARA INDEXAR LOCALIDADES Índices de localidade são apresentados para uso com bancos de dados e mapas eletrônicos. Cada aspecto geográfico em um banco de dados geográfico é associado a nomes de localidade a partir de várias fontes de nome de localidade. Tokenização, normalização, otimização e casamento sensíveis ao contexto, de nomes de localidade, eliminam nomes de localidade duplicatas e variantes, enquanto preservam nomes significativamente diferentes. Uma tabela de nomes de localidade inclui a representação analisada de cada nome de localidade e outras informações associadas, e um token primário para indexação é identificado. Uma máscara de fonte principal é criada por alocar um bit para cada fonte de nome de localidade utilizada no método. Uma máscara de fonte separada é armazenada para cada aspecto geográfico associado a uma localidade, um conjunto de bits para cada fonte na qual a localidade pode ser encontrada. Nomes de localidade associados a cada aspecto geográfico são indexados em uma tabela de aspectos geográficos em ordem de prevalência para uso em uma dada aplicação.



"ÍNDICES DE LOCALIDADE E MÉTODO PARA INDEXAR LOCALIDADES"

Reivindicação de prioridade

O pedido de patente US número 11/433.104, intitulado LOCALITY INDEXES AND METHOD FOR INDEXING LOCALITIES, de Michael Gelich, depositado em 12 de maio de 5 2006 (número do dossiê do procurador TELA-07767USO).

CAMPO DA INVENÇÃO

A presente invenção refere-se a índices de localidades para bancos de dados geográficos, e mais particularmente a estruturas de dados em bancos de dados geográficos utilizados para indexar nomes de localidade e características geográficas associadas 10 contidas nas localidades.

ANTECEDENTES DA INVENÇÃO

Nos últimos anos, os consumidores foram dotados de uma variedade de dispositivos e sistemas para permitir aos mesmos localizar endereços de rua específicos em um mapa digital. Esses dispositivos e sistemas têm a forma de sistemas de navegação no 15 veículo que permitem aos motoristas navegar por ruas e estradas, dispositivos portáteis como assistentes pessoais digitais ("PDAs"), dispositivos de navegação pessoais e telefones celulares que podem fazer o mesmo, e aplicações de Internet nas quais usuários podem gerar mapas mostrando localizações desejadas. O aspecto comum em todos esses e outros tipos de dispositivos e sistemas é um banco de dados geográficos de características 20 geográficas e software para acessar e manipular o banco de dados geográficos em resposta a entradas de usuário. Essencialmente, em todos esses dispositivos e sistemas um usuário pode entrar uma localização alvo e o resultado retornado será a posição da localização alvo.

Tipicamente, os usuários entrarão um endereço, o nome de uma firma, como um restaurante, um centro de cidade, ou um marco de destino, como a Ponte Golden Gate, e 25 então receberão a localização do lugar solicitado, ou característica. A localização pode ser mostrada em uma exibição de mapa, ou pode ser utilizada para calcular e exibir orientações de direção até o local ou utilizada de outras maneiras.

Tipicamente, aplicações utilizam métodos de busca top-down que buscam a localidade na qual uma característica geográfica desejada está localizada, a seguir buscam 30 a característica geográfica naquela localidade. Os exemplos de características geográficas que podem ser encontradas em uma localidade são endereços, marcos e locais comerciais. As aplicações utilizam também métodos de busca bottom-up que buscam todas as características geográficas que casam com certos critérios, a seguir escolhem a característica geográfica desejada a partir da lista de localidades na qual as características 35 geográficas casadas estão localizadas.

Atualmente, bancos de dados geográficos não são abastecidos com índices de localidade ou têm índices de localidade que são de funcionalidade limitada ao buscar

características geográficas em localidades. Um índice de localidade pode ser utilizado para selecionar um nome de localidade e informações associadas para exibição para um usuário. Uma localidade é, por exemplo, uma cidade ou cidade pequena em um estado (EUA), província (Canadá), condado ou outra característica geográfica principal. Para bancos de dados geográficos atualmente tendo índices de localidade, os índices são basicamente listas de nomes de localidade, ordenados por fonte de nomes, com duplicação de nomes entre fontes. Nomes de localidades podem ser encontrados em muitas fontes de nomes de localidade, como fontes administrativa, postal e coloquial. O termo “nome de localidade” nesse pedido é utilizado para se referir a qualquer dado que pode ser utilizado como uma descrição de localidade. Além das fontes listadas acima, os próprios códigos postais podem ser utilizados como nomes de localidade. Também números de centrais de telefone indicam localidade em alguns países e podem ser utilizados como nomes de localidade. Na Alemanha, prefixos de placa de licença de carro indicam localidade e podem ser utilizados como nomes de localidade. O que se segue é uma discussão de estado da técnica de banco de dados geográficos independente de se um banco de dados geográfico é fornecido ou não com um índice de localidade.

Atualmente, um banco de dados geográfico povoado com informações de localidade a partir de várias fontes de nome de localidade conterá entradas duplicatas para uma localidade se o nome de localidade aparecer em múltiplas fontes de nome de localidade. Os fabricantes de sistema ou dispositivo ou desenvolvedores de aplicações não fundem as localidades duplicatas em um único conjunto de nomes ou fazem uma fusão incompleta devido a diferenças na representação das duplicatas através de fontes de localidade, como soletração, pontuação, abreviatura ou outras diferenças entre as duplicatas. Desse modo, quando um usuário consulta então uma aplicação de banco de dados geográficos para uma localidade, o sistema ou dispositivo do usuário pode listar o mesmo nome de localidade múltiplas vezes se o nome de localidade aparecer em múltiplas fontes de nome de localidade. Isso é confuso para o usuário que deve escolher entre nomes idênticos ou quase idênticos exibidos para a tela do dispositivo ou sistema do usuário. Existe um problema adicional na lista de nomes de localidade se o usuário for incapaz de diferenciar entre localidades duplicatas efetivas e localidades separadas tendo nomes iguais ou levemente variantes. O problema de nomes de localidade duplicatas a partir de múltiplas fontes de nomes de localidade é exacerbado em alguns dispositivos de navegação que têm memória limitada. Por exemplo, alguns dispositivos podem conter somente dois nomes de localidade por característica geográfica. Para uma característica geográfica associada a mais de dois nomes de localidade, qualquer seleção de dois dos nomes de localidade para usar no dispositivo pode ser inferiores à ótima porque localidades que são duplicatas porém separadas e localidades tendo nomes de localidade mais prevalentes podem estar ausentes

na seleção. Uma localidade separada duplicata ausente pode levar um usuário a pegar uma localidade incorreta devido a sua exclusividade aparente em uma lista. Para bancos de dados geográficos tendo índices de localidade, a falha em fundir localidades duplicatas também cria índices de localidade que são de difícil controle em tamanho, especialmente para dispositivos de navegação de memória limitada.

Atualmente, para localidades tendo nomes iguais ou levemente variantes que compartilham as mesmas características geográficas exatas, entradas de nome duplicatas não são eliminadas a partir de índices de localidade da técnica anterior. Para localidades tendo nomes iguais ou levemente variantes que compartilham pelo menos uma característica geográfica, as entradas de nome não são fundidas em uma única entrada em índices de localidade da técnica anterior. Um banco de dados geográfico povoado com informações de localidade a partir de várias fontes de nome de localidade pode conter nomes levemente variantes para uma localidade se pelo menos duas das fontes diferentes tiverem nomes levemente variantes para a localidade. Por exemplo, Ho-Ho-Kus, Nova Jérsei, é conhecida por nomes levemente diferentes em diferentes fontes, como Ho-Ho-Kus, Ho Ho Kus ou Ho-Ho-Kus (Hohokus). Para índices de localidade da técnica anterior, a falha em eliminar entradas de banco de dados geográficos tendo nomes de localidade levemente variantes cria índices de localidade que são de difícil controle em tamanho, especialmente para dispositivos de navegação de memória limitada, e confusão para usuários que tentam distinguir entre esses nomes de localidade levemente diferentes. Para localidades de nome duplicata ainda assim separadas, a técnica anterior distingue atualmente entre as localidades por exibir informações adicionais, como o condado no qual a localidade está localizada. Para essas localidades, cidades próximas, bem conhecidas ou prevalentes exibidas como informações adicionais com as localidades seriam mais úteis para um usuário porque nomes de cidade e localizações são mais prováveis de serem reconhecíveis para o usuário do que nomes de condado nos EUA.

A Figura 1 ilustra um diagrama mostrando um exemplo de definições de localidade que não são tratadas de forma consistente em uso comum. Os exemplos de definições de localidade são “local postal” e “subdivisão de condado”. Na Figura 1, em uso comum, Allston é considerado como sendo parte de Boston. Allston é mostrado contido na Subdivisão de condado: Boston. Ao contrário, Manhattan é considerado como sendo parte da Cidade de Nova York, porém Manhattan é uma Subdivisão de Condado e Cidade de Nova York é um Local postal bem como um Local Incorporado. Na Figura 1, Subdivisão de condado: Manhattan é mostrado contido no Local postal: Cidade de Nova York. Tais contradições ilustram a diferença entre definições de uso comum e localidade formal.

Além disso, em outro exemplo de definições de localidade que não são tratadas consistentemente em uso comum, certas características geográficas no estado de Nova

York estão contidas nas localidades parcialmente em sobreposição conhecidas em uso comum como SoHo, Manhattan e Cidade de Nova York. Como mencionado acima, a cidade de Nova York pode ser encontrada em uma fonte de nome de localidade de Local Postal, e Manhattan pode ser encontrado em uma fonte de nome de localidade de Local incorporado.

5 SoHo, por outro lado, não pode ser encontrado em uma fonte de nome de localidade e é conhecido coloquialmente. SoHo estará ausente em um índice de localidade com base somente em definições de localidade formal.

Além disso, índices de localidade de banco de dados geográficos atuais não são ordenados por prioridade, ou sua importância para uso comum. Além disso, para cada
10 característica geográfica em um banco de dados geográfico, localidades associadas a uma característica geográfica não são priorizadas para a característica geográfica. Para um dispositivo de memória limitada que pode armazenar somente alguns nomes de localidade para cada característica geográfica, sem priorização de localidades, um desenvolvedor de aplicações deve escolher alguns nomes de localidade para uma característica geográfica associada a mais de algumas localidades. Preferivelmente, as localidades de prioridade
15 mais elevada associadas a uma característica geográfica, ou aquelas localidades que são as mais conhecidas ou mais prevalentes em uso comum, seriam exibidas para o dispositivo de um usuário. Ao apresentar uma lista de localidades para um usuário, os nomes de prioridade mais elevada associados a características geográficas devem ser utilizados uma vez que serão os mais reconhecíveis.
20

Além disso, o componente de nome mais importante, ou token primário, de um nome de localidade, como "Hadley" no nome "South Hadley", não é identificado em alguns índices de localidade de banco de dados geográfico atuais. Quando algumas aplicações de navegação atualmente comercialmente disponíveis buscam a cidade Hadley em
25 Massachusetts, Hadley é recuperado, porém South Hadley não é recuperado. Para encontrar South Hadley, o usuário tem de iniciar com "S" e separar através de muitas opções que começam com "South".

Um índice de localidade de banco de dados geográfico é necessário de tal modo que nomes de localidade duplicatas e localidades conhecidas por nomes levemente
30 variantes são fundidos, se e somente se representarem a mesma localidade, para eliminar confusão para um usuário que deve de outro modo escolher entre uma lista de nomes idênticos ou levemente variantes, especialmente para dispositivo de memória limitada. Um tal índice de localidade também é necessário para reduzir o tamanho do índice de outro modo de difícil controle. Embora faça a fusão de localidades com nomes variantes e duplicatas, há também necessidade de preservar nomes de localidade significativamente
35 diferentes. Um índice de localidade é necessário de tal modo que nomes de localidade duplicatas que representam localidades separadas são distinguidos. De outro modo, o

usuário não tem meio para diferenciar dois lugares diferentes com o mesmo nome. Além disso, um índice de localidade flexível é necessário de tal modo que definições de localidade formais não tratadas consistentemente em uso comum são consideradas, e de tal modo que o índice não se baseia nessas definições de localidade formais. Um índice de localidade é necessário que é ordenado por prioridade de localidade para cada característica geográfica associada a múltiplas localidades. A ordenação por prioridade permite que nomes mais importantes sejam escolhidos para serem incluídos em aplicações de memória limitada e identifica o melhor nome para apresentar ao usuário. Finalmente, um índice de localidade é necessário de tal modo que o componente de nome mais importante para uma localidade faz parte do índice para assegurar que uma busca para o componente de nome retornará uma lista expandida de todas as localidades relevantes.

SUMÁRIO DA INVENÇÃO

Descrito em termos gerais, um índice de localidade é provido para uso com mapas eletrônicos e bancos de dados eletrônicos, bem como um método e sistema para criar o índice.

Nomes de localidade a partir de várias fontes de nomes de localidade são associados a características geográficas para cada característica geográfica em um banco de dados geográfico. Tokenização sensível a contexto, normalização, otimização e casamento de nomes de localidade permitem eliminação e fusão de nomes de localidade duplicatas e variantes, enquanto preserva nomes significativamente diferentes. Nomes de localidade duplicatas são eliminados, se e somente se representarem a mesma localidade, para reduzir confusão para um usuário que deve de outro modo escolher entre uma lista de nomes idênticos ou similares. Entradas de banco de dados geográfico para localidades conhecidas por nomes levemente variantes são fundidas em uma única entrada se as localidades compartilharem pelo menos uma característica geográfica em comum. Localidades separadas tendo nomes de localidade duplicatas ou levemente variantes são distinguidas por adornar as mesmas com o nome de uma localidade próxima se e somente se representarem localidades diferentes, novamente para reduzir confusão para um usuário que deve de outro modo escolher entre uma lista de nomes idênticos, ou nomes que são distinguidos em modos que são menos significativos para o usuário, por exemplo, por adornar com nomes de condado cujas localizações não são genericamente conhecidas para os usuários.

Uma tabela de nome de localidade é criada e inclui o nome completo da localidade, o token primário da localidade para indexar e outras informações associadas, como um adorno, informações de centro de cidade e tamanho da localidade. Uma máscara de fonte principal é criada por alocar um bit para cada fonte de nome de localidade utilizada no método. Para cada característica geográfica em uma tabela de prioridade de localidade de

característica, uma máscara de fonte separada é armazenada para cada localidade associada à característica geográfica, um conjunto de bits para cada fonte na qual a localidade pode ser encontrada. Nessa tabela estão links com a tabela de nome de localidade e uma prioridade para cada localidade associada a uma característica geográfica.

5 A tabela de localidade de característica também inclui links com a tabela de encontrar característica, que inclui informações de características geográficas associadas para cada característica geográfica.

Os nomes de localidade para cada característica geográfica são indexados em ordem de prioridade. Na modalidade preferida, a localidade de prioridade mais elevada associada a uma característica geográfica é aquela encontrada em uma fonte de nome postal preferido, a seguir a prioridade das localidades restantes é determinada pelo número de conjunto de bits em cada máscara de fonte de localidade. Em um tal índice, uma primeira localidade tem uma prioridade mais elevada do que a segunda localidade se a primeira localidade for mais conhecida ou prevalente em uso comum.

15 A ordenação por prioridade permite que nomes mais importantes sejam escolhidos para serem incluídos em aplicações de memória limitada e identifica o melhor nome para apresentar para o usuário em uma busca bottom-up. O tamanho de difícil controle do índice de localidade que teria contido nomes de localidade duplicatas e levemente variantes é desse modo reduzido. Além disso, o índice de localidade leva em consideração definições de localidade que não são tratadas consistentemente em uso comum porque o índice não se baseia nessas definições de localidade formais. Finalmente, o componente de nome mais importante para uma localidade a partir da etapa de tokenização é parte do índice para assegurar que uma busca para o componente de nome retornará uma lista expandida de todas as localidades relevantes.

25 BREVE DESCRIÇÃO DOS DESENHOS

A Figura 1 ilustra um diagrama mostrando um exemplo de definições de localidade que não são tratadas consistentemente em uso comum.

A Figura 2 ilustra um diagrama mostrando uma hierarquia de áreas administrativas dos Estados Unidos.

30 A Figura 3 ilustra um exemplo da necessidade de diferenciar entre endereços com o mesmo nome, como "Adams Street" que são localizados em quatro localidades diferentes em uma localidade, como "Boston, Massachusetts."

A Figura 4 ilustra um exemplo de localidades oficiais e vizinhanças com o mesmo nome como "Brentwood, Califórnia" que podem ser distinguidas através do uso de múltiplos tipos de fontes de nome de localidade.

35 A Figura 5 ilustra um exemplo de pequenas aldeias que podem ser listadas em fontes oficiais porém que não têm limites claramente delineados, como "Quechee, Vermont"

que são necessários para inclusão em um índice de localidade abrangente.

A Figura 6 ilustra um exemplo de vizinhanças, que são nomes de localidade não oficiais, como “Greenwich Village” na Cidade de Nova York, que são necessários para inclusão em um índice de localidade abrangente.

5 A Figura 7 ilustra um exemplo de aldeias localizadas em uma circunscrição, como “Forest Hills” na circunscrição de Queens na Cidade de Nova York, que são necessárias para inclusão em um índice de localidade abrangente.

10 As Figuras 8A e 8B mostram uma modalidade de um fluxograma de processo para ligar localidades a características geográficas em um banco de dados geográfico, tokenização, normalização, otimização e casamento de nomes de localidade e criação de um índice de localidades ordenadas por prioridade.

A Figura 9 ilustra um exemplo de votação de face utilizado para determinar um nome de localidade para uma rua associada a um nome de localidade desconhecido.

15 A Figura 10 mostra dois exemplos de máscaras de fonte de nome de localidade para os Estados Unidos e para o Canadá.

A Figura 11 mostra uma modalidade de um algoritmo para reduzir o conjunto de nomes de localidade através de casamento de nomes de localidade.

A Figura 12 mostra uma modalidade de um algoritmo para determinar a prioridade de nomes de localidade para uma dada característica geográfica.

20 A Figura 13 mostra uma modalidade de arquivos de índice de localidade incluindo uma tabela de Prioridade de Localidade de característica, uma tabela de Nome de localidade e uma tabela de Encontrar característica.

A Figura 14 ilustra um exemplo para o qual uma aplicação de navegação pode acomodar inconsistência quando uma cidade próxima é especificada de forma errônea.

25 A Figura 15 mostra um diagrama de blocos de um sistema exemplar que pode ser utilizado com modalidades.

DESCRIÇÃO DETALHADA

30 Para criar um melhor índice de localidade, uma lista completa de nomes de localidade deve primeiramente ser criada por coletar nomes a partir de uma variedade de fontes de nomes de localidade, fontes de nomes de localidade administrativa, postal e coloquial, entre outras. A utilização de nomes de localidade a partir de qualquer número e tipo de fontes permite um esquema universal para dados internacionais. Sem essa característica somente um número fixo de fontes pode ser utilizado, como fontes de nome postal ou administrativo, nomes importantes potencialmente ausentes e limitando os tipos de fontes que podem ser utilizados em países diferentes.

Embora a linguagem utilizada nessa descrição seja específica para os Estados Unidos, em modalidades, os mesmos princípios podem ser aplicados internacionalmente

somente com ajustes nominais. Os exemplos de equivalentes de fonte de nome de localidade estrangeira incluem a Ordnance Survey and Royal Mail no Reino Unido, e Stats Can and Canada Post no Canadá.

Em modalidade, para um dado conjunto de fontes de nomes de localidade, uma lista de nomes de localidades é tirada de cada fonte de nomes de localidade. Em modalidades, as fontes são aquelas contendo localidades em um ou mais estados selecionados, territórios, províncias, ou distritos, por exemplo. Na modalidade preferida, as fontes são aquelas contendo localidades nos Estados Unidos. Nos Estados Unidos, por exemplo, fontes de nomes de localidade incluem, porém não são limitadas a:

1. Federal Information Processing Standards 55 (FIPS55). Esse componente do banco de dados United States Geological Survey (USGS) TIGER está no domínio público (<http://geonames.usgs.gov/fips55.html>). FIPS55 é uma fonte padrão que descreve estrutura de localidade para localidades administrativas como definido pelo governo, por exemplo, códigos para lugares povoados nomeados, divisões de condado primário, e outras localizações dos Estados Unidos, Porto Rico e as áreas afastadas.

2. Arquivo de Estado/cidade do United States Postal Service (USPS). Esse arquivo é um componente do produto USPS ZIP+4. Esses nomes de cidade e estado são encontrados na gama de endereços ou nível de código ZIP. Códigos ZIP de cinco dígitos e extensões de quatro dígitos (ZIP+4) são tratados como nomes de localidade em um índice e apontam para o conjunto apropriado de nomes no Arquivo de estado/cidade USPS. Embora seja genericamente somente um nome de localidade postal preferido para cada local, o serviço postal também inclui qualquer número de nomes de localidade postal permissível e não permissível para o mesmo local. Um nome de localidade postal “preferido” é o nome que USPS recomenda para uso ao endereçar correspondência. Um nome de localidade postal “permissível” é um nome suposto que o USPS aprovou e permite para entrega de correspondência. Um nome de localidade postal “não permissível” é um que USPS não permite para entrega de correspondência. Em modalidades, o índice de localidade incluirá todos os nomes de localidade postal preferidos e permissíveis para cada característica geográfica.

3. Geographic Names Information System (GNIS) fornecido pelo United States Geological Survey (USGS). Esse é um banco de dados de domínio público de nomes de localidade nos Estados Unidos, incluindo os cinquenta estados e os territórios. GNIS lista nomes de cidades, seus pontos centrais, suas populações e informações similares.

4. Pontos de interesses (POIs) para centros de cidades.

5. POIs para Agências de Correios USPS.

6. Sistema de Codificação e referência geográfico topologicamente integrado (TIGER) do United States Census Bureau Registro tipo C para entidade “P” (lugares

incorporado em TIGER).

7. TIGER Registro tipo C para entidade "M" (Subdivisões de condado em TIGER).

5 Nomes de localidade que são inteiramente contidos em um estado podem ser associados ao estado para fins de indexação. Localidades que não são inteiramente contidas em um estado, como certos códigos postais nos Estados Unidos, podem ser indexados múltiplos de acordo com seus estados de contenção. A Figura 2 ilustra um diagrama que mostra uma hierarquia de áreas administrativas dos Estados Unidos. Essas áreas administrativas são inteiramente contidas nos grupos mostrados centralmente no diagrama como Nação, Regiões, Divisões, Estados e Contados. Esse diagrama mostra que 10 subdivisões de condado estão contidas em contados. Lugares administrativos, mostrados como "Lugares" na Figura 2, são inteiramente contidos em um estado. Lugares administrativos podem atravessar fronteiras de condado e subdivisão de condado. Áreas metropolitanas, áreas urbanas e mesmo códigos ZIP podem atravessar fronteiras de estado, e desse modo são somente inteiramente contidas na Nação, como mostrado na Figura 2.

15 A Figura 1 ilustra um diagrama de exemplo mostrando que localidades nos Estados Unidos não podem ser automaticamente modeladas de forma útil para aplicações de navegação utilizando somente um conjunto fixo de regras para manipular nomes a partir de múltiplas fontes de localidade. Locais postais e subdivisões de condado são encontrados em fontes oficiais. Na Figura 1, em Massachusetts, o Local Postal de Allston é inteiramente 20 contido na Subdivisão de condado de Boston. Em Nova York, entretanto, a Subdivisão de condado de Manhattan é inteiramente contida no Lugar Postal da Cidade de Nova York. Desse modo, uma fonte de nome de localidade de Subdivisão de Condado não pode ser necessariamente utilizada para determinar Locais Postais em uma subdivisão de condado específica. Similarmente, uma fonte de nome de localidade de Local Postal não pode ser 25 necessariamente utilizada para determinar uma Subdivisão de condado em um local postal específico. O uso comum de nomes de localidade a partir de fontes diferentes varia com a geografia. Essa variação deve ser considerada ao indexar nomes de localidade a partir de múltiplas fontes.

30 Em modalidades, o seguinte exemplo de caso de uso, como utilizado por um usuário de uma aplicação de software ou dispositivo que acessa o banco de dados geográfico, ilustra os benefícios de utilizar nomes de localidade a partir de múltiplas fontes para construir um índice. Se somente uma fonte de nomes for utilizada, nomes importantes são omitidos. Nomes postais, nomes administrativos, e mesmo nomes coloquiais são todos importantes.

35 Sem fontes de nome postal em Índice:

Entrar estado -> Vermont

Entrar cidade -> Quechee

Cidade não encontrada: Quechee

Com fontes de nome postal em Índice:

Entrar estado -> Vermont

Entrar cidade -> Quechee

5

Encontrado ->

Quechee

Sem fontes de nome administrativo no Índice:

Entrar estado -> Nova York

Entrar cidade -> Manhattan

10

Cidade não encontrada: "Manhattan"

Com fontes de nome administrativo no Índice:

Entrar estado -> Nova York

Entrar cidade -> Manhattan

Encontrado: "Manhattan"

15

Em modalidades, os seguintes quatro exemplos de caso de uso mostram que outro benefício de compilar nomes de localidade a partir de múltiplas fontes de nome de localidade é diferenciar entre endereços de rua ambíguos em uma localidade. Uma cidade nos Estados Unidos pode ter endereços de rua duplicatas localizados em diferentes partes da cidade. Isso é especialmente verdadeiro em cidades grandes, como Boston, Massachusetts. Como mencionado acima, Boston pode ser encontrado como uma Subdivisão de condado na fonte de nome de localidade Administrativa FIPS55. Em modalidades, o primeiro desses quatro exemplos de caso de uso mostra um caso não problemático típico quando um endereço de rua específico é exclusivo em uma cidade, não há problema para fins de navegação, mesmo se a cidade for grande. Um exemplo disso é Newbury Street em Boston. Esse nome de rua tem dez quarteirões de comprimento e não é duplicado em nenhuma outra parte em Boston:

20

25

Com fontes de nome administrativo no índice:

Entrar estado -> Massachusetts

Entrar cidade -> Boston

30

Entrar rua -> Newbury Street // exclusivo independente de número de casa

Nesse ponto, o destino precisa esperar mais entrada a partir do usuário, como um número específico da rua, a interseção mais próxima ou o quarteirão mais próximo. Quando a entrada é fornecida, um destino é localizado em um mapa para o usuário:

Entrar número de rua -> 173

35

Encontrado: "173 Newbury Street, Boston, Massachusetts"

Em modalidades, o segundo desses quatro exemplos de caso de uso ocorre quando o nome de rua é duplicado em uma cidade, porém o número da casa serve para

tornar o destino exclusivo. Uma rua longa que se estende através de várias cidades menores em uma cidade grande é um tal exemplo. Por exemplo, Commonwealth Avenue se estende através de Boston, bem como de cidades menores de Allston e Chestnut Hill em Boston. Como mencionado acima, Boston é uma Subdivisão de condado em fonte de nome de localidade administrativa. Allston e Chestnut Hill são cidades pequenas que podem ser encontradas em fontes de nome de localidade Postal sob códigos postais 02134 e 02467, respectivamente.

Sem fontes de nome administrativo no Índice:

Entrar estado -> Massachusetts
 Entrar cidade -> Boston
 Entrar rua -> Commonwealth Avenue
 Entrar número de rua -> 2000
 Número de rua não encontrado: "2000"

Como Boston não é um nome postal legítimo para o código postal 02467, de acordo com o U.S. Postal Service, "2000 Commonwealth Ave., Chestnut Hill, Massachusetts 02467" não é encontrado no exemplo acima para Boston embora Chestnut Hill seja uma cidade pequena dentro de Boston.

Com fontes de nome administrativo e postal no Índice:

Entrar estado -> Massachusetts
 Entrar cidade -> Boston
 Entrar rua -> Commonwealth Avenue

Nesse ponto, verifica-se que Commonwealth Avenue se estende através de Boston, Allston e Chestnut Hill. O destino precisa esperar mais entrada a partir do usuário, como um número específico de rua, a interseção mais próxima ou o quarteirão mais próximo. Quando a entrada é fornecida, um destino é localizado em um mapa para o usuário:

Entrar número de rua -> 2000
 Encontrado: "2000 Commonwealth Avenue, Chestnut Hill, Massachusetts"

Em modalidades, o terceiro desses quatro exemplos de caso de uso, como ilustrado na Figura 3 é similar ao segundo exemplo de caso de uso, exceto que quatro Adams Streets diferentes podem ser encontradas em quatro localidades diferentes em Boston. A Figura 3 ilustra a necessidade de diferenciar entre endereços com o mesmo nome, como "Adams Street", que são localizados em quatro localidades diferentes em uma localidade, como Boston, Massachusetts:

Sem fontes de nome postal no Índice:

Entrar estado -> Massachusetts
 Entrar cidade -> Boston
 Entrar rua -> Adams Street

Por favor escolher entre ->

Adams St. Boston // a aplicação encontra quatro

Adams St. Boston // Adams streets separadas na cidade

Adams St. Boston // de Boston e o usuário é incapaz

5 Adams St. Boston // de diferenciar entre essas quatro opções

Com fontes de nome postal no Índice:

Entrar estado -> Massachusetts

Entrar cidade -> Boston

Entrar rua -> Adams Street

10 Por favor escolher entre ->

Adams St., Charlestown

Adams St., Hyde Park

Adams St., Roxbury

Adams St., Dorchester

15 Entrar número da rua -> // o usuário continua entrando o número da rua

Nesse exemplo de caso de uso, a aplicação processa cada entrada de usuário antes de solicitar mais informações a partir do usuário. Em outras modalidades, para “Com fontes de nome postal no Índice”, o usuário entra a cidade de Boston, a rua de Adams Street, e um número de rua antes que a aplicação processe essas três entradas. Considerando que o número de rua não é duplicado nas cidades pequenas de Charlestown, Hyde Park, Roxbury e Dorchester, o nome de rua e número serão encontrados para uma dessas quatro cidades e indicados em um mapa para exibição para o usuário.

25 Em modalidades, o quarto desses quatro exemplos de caso de uso mostra que mesmo números de rua, por exemplo “2 Adams St.”, são duplicados em ruas separadas com o mesmo nome em uma cidade. Nesse caso, a única resposta adequada é apresentar ao usuário uma lista de cidades menores nas quais as duplicatas são localizadas, para derivar um destino único. Desse modo, utilizando o exemplo a partir do terceiro exemplo de caso de uso acima:

Com fontes de nomes administrativo e postal no Índice:

30 Entrar estado -> Massachusetts

Entrar cidade -> Boston

Entrar rua -> Adams Street

Entrar número de rua -> 2

Por favor escolher entre ->

35 2 Adams St., Charlestown

2 Adams St., Hyde Park

2 Adams St., Roxbury

2 Adams St., Dorchester

Em modalidades, e outro exemplo de caso de uso como ilustrado na Figura 4, localidades oficiais e vizinhanças com o mesmo nome como “Brentwood, Califórnia” podem ser distinguidas através do uso de múltiplos tipos de fontes de nome de localidade.

5 Brentwood, Califórnia é tanto um local administrativo oficial próximo a San Francisco, como também uma vizinhança famosa porém não oficial de Los Angeles que é um nome postal permissível porém não preferido. A Figura 4 mostra as duas localidades Brentwood na Califórnia. Os dois locais contêm endereços que são prevalentes para fins de navegação e uma boa aplicação de navegação distinguirá entre os mesmos para o usuário:

10 Entrar estado -> Califórnia
 Entrar cidade -> Brentwood
 Por favor escolher entre ->
 Brentwood (cidade próximo a San Francisco)
 Brentwood (vizinhança de Los Angeles)

15 Utilizando esse mesmo exemplo de caso de uso, em outras modalidades, se o usuário entrar o estado, cidade e nome de rua antes da aplicação processar as entradas de usuário, a aplicação pode determinar o Brentwood correto. Por exemplo:

Entrar estado -> Califórnia
 Entrar cidade -> Brentwood
 20 Entrar nome de rua -> Concord Avenue
 Entrar número de rua -> 767
 Encontrado: “767 Concord Avenue, Brentwood (cidade próximo a San Francisco), Califórnia”

Em modalidades, em um exemplo de caso de uso adicional como ilustrado na
 25 Figura 5, pequenas aldeias que podem ser listadas em fontes oficiais porém que não têm fronteiras claramente delineadas, como “Quechee, Vermont” são necessárias para inclusão em um índice de localidade abrangente. A aldeia de Quechee, Vermont é um destino turístico de cidade pequena popular. Simon Pierce Glassblowing pode ser encontrado nas Páginas amarelas como 1760 Quechee Main Street, Quechee, Vermont 05059. Quechee,
 30 entretanto, não é uma localidade administrativa, nem o Serviço Postal dos Estados Unidos reconhece esse endereço. O código postal 05059 é um código ZIP de “Caixa de correios apenas” que contém muito poucos endereços de rua. Desse modo, Quechee Main Street não é uma rua reconhecida em Quechee. A área em volta do centro de Quechee é conhecida como White River Junction e Hartford. A Figura 5 ilustra um mapa futuro de
 35 Quechee com um limite de aldeia delineado possível. Uma boa aplicação de navegação necessita reconhecer endereços como são publicados em catálogos de Páginas Amarelas, quer sejam endereços postais legítimos ou lugares incorporados ou não:

Entrar estado -> Vermont
 Entrar cidade -> Quechee
 Entrar rua -> Quechee Main street
 Entrar número -> 1760

5 Encontrado: "1760 Quechee Main Street, White River Junction, Vermont"

Infelizmente, o nome de localidade Quechee não pode ser ligado ao endereço de rua porque o limite de Quechee não é conhecido. Em vez disso, White River Junction é a localidade designada para o endereço de rua. Essa opção está de acordo com os endereços Postais. Uma aplicação de navegação pode determinar que encontrou a localização desejada através do uso do índice de localidade, criado como discutido abaixo. Embora Quechee não seja a localidade para "1760 Quechee Main Street", o índice de localidade pode expandir a localidade Quechee para localizar a rua em White River Junction, Vermont. Uma aplicação de navegação pode pedir a confirmação do usuário quando a localidade casada difere da entrada de usuário. Embora somente uma rua tenha sido encontrada, poderia ser somente um casamento possível, que o usuário da aplicação de navegação pode aceitar ou recusar. Aperfeiçoamentos em mapas poderiam tornar a resposta correta possível no futuro com a adição da fronteira de Quechee. Nesse caso, o nome da localidade na qual "1760 Quechee Main Street" está localizado será na realidade Quechee.

Em modalidades, em um exemplo de caso de uso adicional como ilustrado na Figura 6, vizinhanças, que são nomes de localidade não oficiais, como "Greenwich Village" na Cidade de Nova York, são necessários para inclusão em um índice de localidade abrangente. Há vários nomes de localidade nos Estados Unidos que são importantes para navegação, ainda não publicados em qualquer fonte postal ou administrativa. Uma classe de tais nomes são vizinhanças famosas. Os exemplos incluem Greenwich Village e SoHO na Cidade de Nova York e Haight-Ashbury em San Francisco. Esses lugares são grandes o bastante para conter segmentos de rua, endereços, firmas e outros pontos de interesse. Boas aplicações de navegação incluirão a capacidade de localizar lugares famosos e os endereços de rua nos mesmos, quer os mesmos sejam nomes postais ou administrativos oficiais ou não.

30 Sem nomes de várias fontes:

Entrar estado -> Nova York
 Entrar cidade -> Greenwich Village
 Cidade não encontrada : "Greenwich village"

Com nomes de várias fontes:

35 Entrar estado -> Nova York
 Entrar cidade -> Greenwich Village //nem nome postal nem administrativo
 Entrar rua -> // o usuário continua por entrar o nome da rua

Nesse exemplo de caso de uso, o uso de nomes de várias fontes, um mapa aperfeiçoado poderia incluir o limite de Greenwich Village. A Figura 6 mostra que o Greenwich Village pode ser definido como a área de Manhattan limitada por Spring e 14th Streets, entre Greenwich St. E Broadway. Utilizando um mapa com essas informações, o diálogo continuaria:

Entrar rua -> Carmine Street

Entrar número de rua -> 13

Encontrado: "13 Carmine Street, Greenwich Village, Nova York"

Em modalidades, em um exemplo de caso de uso adicional como ilustrado na Figura 7, aldeias localizadas em uma circunscrição, como "Forest Hills" na circunscrição de Queens na cidade de Nova York, são necessárias para inclusão em um índice de localidade abrangente. Nomes de localidade a partir de diferentes fontes podem ser utilizados para determinar quais das circunscrições da cidade de Nova York um nome de rua pode ser localizado. A cidade de Nova York é composta de cinco circunscrições. Todos exceto um deles, Queens, é independente como um nome de localidade. Em Queens, entretanto, dez das localidades contidas são definidas. Ao procurar um endereço em Queens, o usuário não necessita saber a localidade em Queens na qual o endereço está localizado. O índice de localidade, discutido abaixo, pode determinar qual aldeia contém o endereço, se o endereço é exclusivamente contido somente em uma aldeia:

Entrar estado -> Nova York

Entrar cidade -> Queens

Entrar rua -> 70th Rd.

Entrar número de rua -> 10700

Encontrado: "10700 70th Road, Forest Hills, Nova York"

Para esse exemplo de caso de uso, o índice de localidade pode também manejar solicitações para os nomes de aldeias localizadas em Queens:

Entrar estado -> Nova York

Entrar cidade -> Forest Hills

Entrar rua -> 70th Rd

Entrar número de rua -> 10700

Encontrado: "10700 70th Road, Forest Hills, Nova York"

As Figuras 8A e 8B mostram uma modalidade de um fluxograma de processo para ligar localidades com características geográficas em um banco de dados geográfico, tokenização, normalização, otimização e casamento de nomes de localidade e criação de um índice de localidades ordenadas por prioridade. Em modalidades, os exemplos de características geográficas que podem ser encontradas em uma localidade incluem porém não são limitadas a ruas, segmentos de rua, margens de segmento de rua, faces de bloco,

marcos, parques estaduais, rodovias, linhas de ferry, rotas de ônibus, centros de distribuição, locais comerciais e locais residenciais. Um segmento de rua é uma porção de uma rua, uma gama de endereços ou um endereço único. Uma margem de segmento de rua é um lado de rua de um segmento de rua. Uma face de bloco é uma de quatro faces que
5 constituem um bloco de cidade.

Para um dado conjunto de fontes de nomes de localidade a partir de cima e para um dado banco de dados geográfico de propriedade, o processo começa na etapa 805. Se outro nome de localidade existir para processar na etapa 810, na etapa 815, o processo determina se o casamento de mapa é possível se a fonte contém características geográficas
10 que casam com aquelas no banco de dados geográficos. Se na etapa 815, o casamento de mapa para a fonte for considerado como possível, na etapa 820, o casamento de mapa associa diretamente nomes de localidade a partir da fonte de nome de localidade com características geográficas no banco de dados geográfico. A associação direta pode ser executada automaticamente através de fusão, ou casamento de atributos ou manualmente
15 por inspeção. A associação direta é tipicamente utilizada para fontes de nome de localidade que compartilham atributos com o banco de dados geográficos. Na modalidade preferida, fusão pode ser utilizado quando a fonte de nomes de localidade tem informações espaciais ligadas a ela indicando sua localização e extensão na terra. A associação direta é feita por sobreposição de localidades a partir da fonte de nomes de localidade espacialmente no
20 banco de dados geográfico, atribuição de uma localidade a quaisquer características de banco de dados geográfico que ocorrem no limite daquela localidade. O casamento de atributos é executado por casamento de atributos comuns entre uma fonte e o banco de dados geográfico, que permite então que uma associação direta seja feita. Atributos que podem ser casados são aqueles que podem ser representados por strings ou números. A
25 associação indireta é tipicamente utilizada para as outras fontes.

Em modalidades, na etapa 820 quando as fontes de nomes de localidade compartilham atributos com o banco de dados geográficos, uma associação direta com as características geográficas no banco de dados geográficos é feita por casamento de atributos na fonte contra os mesmos atributos no mapa ou banco de dados geográficos. Por
30 exemplo, o casamento de faixa pode ser utilizado para casar atributos de endereços entre uma fonte de localidade e o banco de dados geográficos. O casamento de faixa pode ser feito utilizando qualquer fonte que tenha nomes de localidade associadas a detalhe de rua, incluindo TIGER, e o diretório de Nomes de Lugar de Cidade USPS. Os códigos de Subdivisão de condado (entidade "M") e Lugar incorporado (entidade "P") são diretamente propagados a partir das características geográficas TIGER casadas sobre as características
35 geográfica no mapa ou banco de dados de interesse. O casamento de faixa toma um nome de rua, faixa de números de casa, e localidade a partir de TIGER e tenta casar esses itens

com um segmento de rua correspondente no banco de dados geográficos de propriedade, de interesse. Em TIGER, cada lado de um bloco de rua não somente tem faixa de endereço, tem tags que representam o tipo de entidade P (nome de lugar incorporado) naquela localização, o tipo de entidade M (nome de subdivisão de condado) naquela localização, um
5 código de estado, um código de bloco, um código de tratado, bem como Minor Civil Division (MCD). As faixas que casam tornam possível transferir informações a partir de TIGER sobre o banco de dados geográficos. Um casamento de faixa pode ser um casamento exato de segmentos de rua, segmentos de rua que tocam ou são exatamente alinhados ou segmentos de rua que parcialmente sobrepõem.

10 Na etapa 820, onde o Arquivo Estado/cidade de USPS é a fonte de nomes de localidade, as faixas de endereços de entrega do catálogo ZIP+4 do USPS da fonte são geocodificados contra o mapa ou banco de dados. Em modalidades, códigos ZIP a partir dessa fonte são tratados como os próprios nomes de localidades. Códigos ZIP a partir dessa fonte também apontam para o conjunto apropriado de nomes de localidade no
15 arquivo Estado/cidade. Para cada casamento bem sucedido, o código ZIP de cinco dígitos e um código plus4 de quatro dígitos a partir do ZIP+4 é tratado como um nome de localidade e são propagados sobre a característica geográfica correspondente.

Na etapa 825, para características geográficas em um banco de dados geográfico que não foram casadas com a fonte de nomes de localidade, a votação de face é utilizada
20 para casar as características geográficas com outras características no banco de dados geográfico, desse modo herdando atribuições de localidade a partir das características casadas. A Figura 9 ilustra um exemplo de votação de face utilizada para determinar um nome para uma face de bloco de cidade no banco de dados geográfico associado a um nome de localidade desconhecido. Em modalidades, furos ou características geográficas
25 não casadas na cobertura para as fontes de nomes TIGER são eliminados por um processo de “votação de face”. Para um bloco de cidade que tem uma face de bloco associada a um nome de cidade desconhecida, a votação de face determina um nome de cidade para a face de bloco com base nos nomes de cidades que correspondem a faces de blocos que circundam a mesma, ou faces de bloco que conectam a face de bloco dada a si própria. A
30 Figura 9 ilustra a votação de face para um bloco de cidade, de tal modo que para uma face de bloco dada, as faces de bloco utilizadas em votação de face são duas faces de blocos adjacentes à mesma e uma face de bloco oposta à mesma. As faces de bloco da Figura 9 também podem ser visualizadas como características geográficas que são cada um lado de um segmento de rua. As faces de bloco adjacentes e opostas são examinadas em
35 modalidades, a localidade dominante na qual a face não atribuída está localizada é determinada por um voto da maioria das outras faces adjacentes e opostas. Esse processo propaga códigos de Subdivisão de Condado e Lugar Incorporado e seus nomes associados

sobre quaisquer características geográficas não codificadas a partir das características geográficas codificadas adjacentes e opostas, que em modalidades são faces de bloco.

Por exemplo, na Figura 9, o lado norte de um segmento de rua de bloco da Rua do Centro é associado a um nome de cidade desconhecida porque é uma característica geográfica que não estava associada a nenhuma localidade na fonte de nome de localidade. As outras faces de bloco, ou o lado Leste da Primeira Rua um segmento de rua de bloco, o lado Sul da Rua Principal um segmento de rua de bloco e o lado Oeste da Segunda rua um segmento de rua de bloco, entretanto, foram encontradas como associadas a "Boston". Como três desses três segmentos de rua para o bloco foram associados à Boston, o voto de face é três de três, e a Rua do Centro também estará associada à Boston. Se dois desses três segmentos de rua forem associados a uma cidade específica, o voto de face é dois de três, e a Rua do Centro também será associada à cidade específica. No caso de um empate, onde os três segmentos de rua são individualmente associados a uma cidade diferente, então o voto de face é um de três. Uma vez que não há maioria de votos nesse caso, a Rua do Centro será associada à cidade de uma das ruas adjacentes mais próxima a ela, que nesse caso é a Primeira rua ou a segunda rua.

Em modalidades, a votação de face pode ser utilizada para outras características geográficas além de faces de bloco de cidade, como lados de segmento de rua ou margens de estrada. Em modalidades, a votação de face pode ser utilizada para dois ou mais outros lados de segmento de rua além do segmento de rua associado a um nome de cidade desconhecida. Em modalidades, a votação de face também pode ser utilizada onde duas ou mais das faces de bloco são associadas a nomes de cidade desconhecida. Nesse caso, a maioria de votos é tirada das faces de bloco restantes, e a maioria de votos ou um empate é encontrado e tratado como discutido acima. Em modalidades, a votação de face pode ser utilizada para associar as faces de bloco com outros nomes de localidade além de cidades ou cidades pequenas. Por exemplo, nomes de localidade no Arquivo de Estado/cidade de USPS são o código ZIP de cinco dígitos e um código de construção de quatro dígitos a partir do arquivo ZIP+4.

Outras modalidades de votação de face incluem um voto ponderado ou um voto de comprimento linear em vez de maioria de votos. Em modalidades utilizando um voto ponderado, certas faces de bloco adjacentes a uma face de bloco não associada a uma localidade têm preferência, ou têm peso maior no processo de votação. Um voto ponderado pode ter qualquer componente de ponderação que mede a confiança das atribuições de face de bloco adjacentes. Por exemplo, preferência poderia ser dada a faces de bloco que corresponde a ruas principais ou que são localizadas em regiões maiores. O comprimento das faces de bloco é outra ponderação. Em modalidades, utilizando um voto de comprimento linear, para uma dada face de bloco não associada a uma localidade, para

cada localidade conhecida associada a faces de bloco adjacentes à face de bloco dada, o comprimento total das faces de bloco é tomado para determinar qual localidade associada a faces de bloco adjacentes tem faces de bloco do comprimento linear total mais longo. Essa localidade resultante é então atribuída à face de bloco dada não associada a uma
5 localidade.

Na Figura 8A, se na etapa 815 o casamento de mapa não for possível porque a fonte não compartilha nenhum atributo com o banco de dados geográficos, na etapa 855, o casamento de nome de fonte cruzada é empregado em modalidades. Cruzamento de fontes é associação indireta de nomes de localidade na fonte, ou primeira fonte, com aqueles de
10 outra fonte já associada diretamente a características geográficas no banco de dados geográfico. Na etapa 855, se o casamento de nome de fonte cruzada for possível porque uma segunda fonte já diretamente associada a características geográficas no banco de dados geográficos for encontrada com nomes de localidade em casamento com uma primeira fonte, na etapa 860 a primeira fonte é casada com a segunda fonte. Na etapa 865,
15 cada nome de localidade na primeira fonte herda as associações com características geográficas a partir da segunda fonte, e é desse modo associado indiretamente com a característica geográfica específica. Em modalidades, os exemplos, de características geográficas herdadas são lados de segmento de rua, faces de bloco e linhas de ferry. Em modalidades, os dados FIPS55 são uma fonte de nomes úteis para casamento de nome de
20 cruzamento de fonte. Por exemplo, as localidades GNIS para fonte de Lugares Povoados é casada contra os nomes de localidade na fonte FIPS55 em um estado e condado. Onde os casamentos são feitos, os nomes GNIS herdam as associações com lados de segmento de rua a partir de seus nomes FIPS55 em casamento. A partir da etapa 865, o processo se move para a etapa 830, como discutido abaixo. Se na etapa 855 o casamento de
25 cruzamento de fonte não for possível para a fonte, a fonte não é utilizável no processo, e o processo retorna para selecionar outra fonte de localidade na etapa 810.

Os nomes de localidade tirados das várias fontes de nomes de localidade são tokenizados, normalizados, otimizados e/ou casados, fundidos ou adornados para eliminar nomes de localidade duplicatas e variantes, em modalidades. Na modalidade preferida,
30 todas as etapas de tokenização, normalização, otimização, casamento e fusão ou adorno são executadas. Esse processo reduz o número de nomes de localidade para cada localidade que tem dois ou mais nomes similares, enquanto também preserva nomes de localidade que são significativamente diferentes. Essas etapas acomodam diferenças em codificação de nome entre as várias fontes. Um exemplo de nomes de localidade similares a
35 partir de várias fontes é a cidade de HO-Ho-Kus, Nova Jérsei, que aparece como a seguir em várias fontes de nomes de localidade:

TIGER registro tipo C: Ho-Ho-Kus Twnshp

Estado/cidade de USPS: HO HO KUS Township

Centro de assentamento POI: HO-HO-KUS

FIPS55-3: Ho-Ho-Kus (Hohokus)

GNIS: Ho-Ho-Kus

5 A partir das etapas 825 e 865 na Figura 8A, o processo se move para a etapa 830. Na etapa 830, a primeira parte do processo de casamento de nomes, tokenização, ou análise pode dividir um nome de localidade em tantos quanto aproximadamente dez tokens ou componentes, em modalidades. Muitas técnicas podem ser utilizadas para tokenizar nomes de localidade. A finalidade dessa etapa é dividir o componente ou porção significativa do nome de localidade, ou o “corpo” do nome, para fins de indexação. Os outros componentes, como prefixos ou sufixos serão individualmente componentes separados. Nomes de localidade são então representados por tokens em um índice, desse modo permitindo que o desenvolvedor de aplicações indexe a porção significativa do nome. Por exemplo, tanto Amherst como South Amherst será então indexado sob “A” se desejado. A
10 eliminação de duplicatas em modalidades permitirá que os usuários finais tenham acesso a mais nomes em aplicações de memória limitada e evitará confusão para o usuário de ver o mesmo nome apresentado múltiplas vezes.

 A tokenização de nomes de localidade a partir das duas primeiras fontes de nome de localidade listadas acima para o exemplo de Ho-Ho-Kus, Nova Jérsei produz os
20 seguintes tokens de sufixo e corpo:

 Corpo: Ho-Ho-Kus, sufixo: Twnshp

 Corpo: HO HO KUS, sufixo: Township

 A tokenização é útil para isolar aqueles componentes que definem um nome exclusivo e por associação, aqueles tokens que podem ser ignorados no processo de
25 casamento. A maioria dos usuários finais desejará que “Rutland” case com “Rutland Township”, isto é, que o termo “Township” seja tratado como insignificante. Ao mesmo tempo, a maioria dos usuários finais desejará que “Boston” não case com “South Boston”, isto é, o termo “South” seja tratado como significante. Outro motivo para tokenização é oferecer a um desenvolvedor de aplicações de software, flexibilidade em apresentar nomes
30 de localidade para o usuário final porque a porção significativa do nome será indexada. Por exemplo, por tokenização “Hollywood” e “West Hollywood”, ambos serão apresentados como opções de seleção a um usuário final que entra uma busca de mapa para “Hollywood.” Isso ocorre porque o token de “Corpo” para os dois será “Hollywood”, visto que West Hollywood será tokenizado como Corpo: Hollywood, Prefixo: West, e Hollywood será
35 tokenizado como Corpo: Hollywood.

 Em outra modalidade, a tokenização ajuda a determinar a expansão correta de abreviaturas sensíveis a contexto. Por exemplo, um token de prefixo de localidade “St.” mais

provavelmente se refere a “Saint”, ao passo que um token de sufixo de localidade “St.” mais provavelmente se refere a “Estado”.

O que se segue são outros tipos de tokens e exemplos desses tokens:

- 5 Pré-direção – direção dianteira (“North” Adams)
- Pré-tipo – tipo dianteiro (“Lake” Isabella)
- Prefixo – dianteira, porém não uma direção ou tipo (“Old” Orchard Beach)).
- PreNome – palavras que não são tipo antes do corpo (Lake “of the” woods)
- Corpo – peça principal utilizada para fins de índice (Lake “Isabella”)
- Tipo posterior – tipo traseiro (Imperial “Beach”)
- 10 Direção posterior – token de direção traseira (Leisure Village “West”)
- Sufixo – traseira, porém não uma direção ou tipo (Manchester “By the sea”)
- Divisão – identificador numérico especificando divisões da localidade (Meredosia “1”)

- 15 Adorno – informação suplementar parentética, como nome de condado para esclarecer a localização de um nome de localidade (Middletown “(Bethlehem)”).

Na etapa 835 da Figura 8A, a normalização de tokens a partir da etapa de tokenização envolve genericamente um ou mais dos seguintes processos: expandir abreviaturas, reduzir ou remover pontuação, utilizar caso compatível (maiúscula ou minúscula) e remover espaços embutidos, em modalidades. Em modalidades, as abreviaturas padrão para direcionais e para tipos são expandidas. Por exemplo, a abreviatura direcional “N” é expandida para “North”. Para abreviaturas de tipo, por exemplo, “Mt.” É expandida para “Mount” e “AFB” é expandida para “Air Force Base.” Dado que nomes que aparecem em fontes diferentes podem ser representados de forma diferente, a normalização adequada de abreviaturas é crítica para o processo de casamento. Em modalidades, espaços incorporados e pontuação são removidos. Em modalidades, escrita em letra maiúscula pode ser normalizada utilizando letra maiúscula ou letra minúscula consistente para os tokens de nome de localidade. A escrita em letra maiúscula também pode ser normalizada por escrever em letra maiúscula somente a primeira letra de cada token, em modalidades. Além disso, diferenças de escrita em letra maiúscula podem ser acomodadas no processo de casamento em vez de no processo de normalização, em modalidades. Na modalidade preferida, a escrita em letra maiúscula é normalizada para letra maiúscula consistente. Utilizando o exemplo de Ho-Ho-Kus, Nova Jérsei, a normalização dos tokens produz os seguintes resultados:

- 35 Corpo: HOHOKUS, sufixo: TOWNSHIP
- Corpo: HOHOKUS, sufixo: TOWNSHIP

O seguinte exemplo de caso de uso ilustra os benefícios das características de tokenização e normalização que podem ser armazenadas no índice de localidade, cuja

criação é discutida abaixo. Sem essas características no índice, várias abreviaturas aparecem como nomes de cidades diferentes. Com essas características no índice, as abreviaturas são colocadas em uma forma comum, permitindo que o desenvolvedor de aplicações dobre a lista em uma única entrada não ambígua. Embora a escrita em letra maiúscula de tokens seja normalizada em letra maiúscula consistente para facilitar casamento, tokens são tipicamente apresentados ao usuário somente com a primeira letra de cada token maiúscula.

Sem nomes de localidade tokenizados e normalizados no Índice:

Entrar cidade -> Randolph
 10 Por favor escolher entre ->
 Randolph Hghts
 Randolph Heights
 Randolf Hts

Com nomes de localidade tokenizados e normalizados no Índice:

15 Entrar cidade -> Randolph
 Você escolheu: Randolph Heights

O seguinte exemplo de caso de uso ilustra os benefícios de tokenização e normalização de tokens direcionais em nomes de localidade. Por identificar tokens direcionais, nomes de localidade podem ser indexados por seu corpo, em vez de por direção. Após direcionais serem normalizados, um desenvolvedor de aplicações necessita somente checar em relação a tokens normalizados porém não quaisquer abreviaturas desses tokens.

Sem nomes de localidade tokenizados e normalizados no Índice:

Entrar cidade -> Boston
 25 Encontrado: Boston
 Entrar cidade -> South B
 Por favor escolher entre ->
 South Bank
 South Barrister
 30 South Barnstable
 South Boston
 Entrar cidade -> S. Boston
 Cidade não encontrada: "S. Boston"
 Entrar cidade -> South Boston
 35 Encontrado: "South Boston"

Com nomes de localidade tokenizados e normalizados no índice:

Entrar cidade -> Boston

Por favor escolher entre ->

Boston

South Boston

Na etapa 840 da Figura 8A, a otimização para dois ou mais nomes de localidade
5 similares a partir da etapa de normalização associa genericamente cada nome de localidade
similar com características geográficas contidas na localidade, em modalidades. Os
exemplos de características geográficas incluem ruas, segmentos de rua, marcos, parques
estaduais, rodovias, locais comerciais e locais residenciais. No exemplo de Ho-Ho-Ku, Nova
Jérsei, a otimização encontrará as mesmas características geográficas para HoHokus e
10 para HOHOKUS.

Na etapa 845 da Figura 8A, em uma máscara de fonte principal, o bit seguinte na
máscara de fonte é alocado à fonte. Em modalidades, a máscara é exclusiva em um país.
Em outras modalidades, a máscara poderia ser exclusiva para qualquer área geográfica,
como um estado ou continente. A Figura 10 mostra dois exemplos de máscaras de fonte de
15 nome de localidade para os Estados Unidos e para o Canadá. Em modalidades, cada
posição de bit na máscara de fonte representa uma única fonte de nome de localidade. A
máscara pode conter uma ou mais fontes de nome de localidade administrativa, postal ou
outra. A máscara é exclusiva em um país e não envolve prioridade de fontes de nome de
localidade. Para cada valor de bit na coluna "Valor de bit decimal", uma fonte de nome de
20 localidade na coluna "Fonte de nome de localidade" é alocada ao valor de bit. Para fins de
indexação, a máscara de fonte de localidade permite a flexibilidade para definir tipos
diferentes de nomes de localidade para se adequar melhor à aplicação final. Em
modalidades, as fontes na máscara indicada como "Trump" podem ser utilizadas para dar
prioridade máxima a nomes de localidade que são encontrados nessas fontes para fins de
25 indexação. Para cada nome de localidade na fonte, uma máscara de fonte individual
também é criada, mostrando as fontes nas quais o nome de localidade aparece.

Na etapa 850, a posição de bit seguinte na máscara de fonte para cada nome de
localidade na fonte é definida para essa fonte. Nomes que aparecem em múltiplas fontes
terão bits definidos na máscara para cada fonte na qual aparecem. Por exemplo, o nome
30 "Boston" é simultaneamente um nome de subdivisão de condado, um lugar administrativo e
o nome postal preferido para diversos códigos ZIP. Nomes que não aparecem em múltiplas
fontes terão somente um único conjunto de bits em sua máscara correspondendo a sua
fonte. O processo retorna para a etapa 810 para processar a próxima fonte de nomes de
localidade se existir uma.

35 Se na etapa 810 da Figura 8A não houver fontes de localidade restantes para
processar, o processo se move para a etapa 868 na Figura 8B. Na etapa 868, os nomes
otimizados a partir de todas as fontes utilizáveis são casados. As fontes utilizáveis são

aquelas para as quais o casamento de mapa foi possível na etapa 815 e aquelas fontes para as quais outro casamento de fonte foi possível na etapa 855 na Figura 8A. O casamento concatena os tokens normalizados em nomes completos e compara os mesmos para determinar se podem ser considerados um casamento, em modalidades. Em modalidades, a normalização de caso de nome de localidade ou diferenças de escrita em letra maiúscula poderia ser executada nessa etapa de casamento de nome em vez da etapa de normalização acima. Em modalidades, a lógica de casamento insensível o caso poderia ser utilizada nessa etapa de casamento. Para cada estado nos Estados Unidos, todos os nomes de localidades a partir das fontes designadas são casados em modalidades.

Muitos algoritmos diferentes são possíveis para casamento de nome. Os exemplos de técnicas de casamento de nome incluem casamento sensível a contexto, casamento fonético e Soundex. O casamento sensível a contexto é casamento string dos nomes ou casamento da soletração de nomes. Esse tipo de casamento é executado com conhecimento de quais tokens estão sendo casados que permitem regras especiais. Por exemplo, no token de corpo, um bom algoritmo de casamento sensível a contexto pode casar “John F. Kennedy” e “John Fitzgerald Kennedy”. Um excelente algoritmo de casamento sensível a contexto pode casar “MLK” e “Martin Luther King.” O casamento fonético, por outro lado, casa os sons de palavras ao contrário da soletração das palavras. Por exemplo, “fish” e “phish” casam foneticamente. Para casamento de nome em vários idiomas, diferentes algoritmos de casamento fonético podem ser utilizados. Soundex é um algoritmo fonético para indexar nomes por seu som quando pronunciado em inglês. O objetivo básico é para nomes com a mesma pronúncia serem codificados na mesma string de modo que o casamento possa ocorrer apesar de diferenças pequenas em soletração. Informações mais detalhadas em relação a algoritmos fonéticos podem ser encontradas no pedido número 11/377.764, depositado em 16 de março de 2006, intitulado “Geographic Feature Name reduction using phonetic algorithms” de Jesse Sheridan.

Em modalidades, para que dois nomes completos casem, as strings devem casar exatamente. Se nomes completos não casarem, em modalidades, um casamento de tokens de corpo é tentado. Tokens de corpo devem casar e tokens de direção e tipo também devem casar para um casamento de token bem sucedido. Desse modo, o casamento dos tokens pode não iniciar com um ou ambos tokens dianteiros, e um token deve ser uma substring dianteira da outra. Desse modo, o casamento de tokens também deve ignorar certos tokens. Em modalidades, variações de soletração pequenas podem ser permitidas entre dois nomes em casamento. Em modalidades, o casamento de nome é implementado relativamente de forma conservativa para evitar casamentos falsos. Desse modo:

“North Boston” não casa com “South Boston”

“South Boston” não casa com “Boston”

“Township of Rutland” não casa com “Rutland Township”

Na etapa 870 da Figura 8B, todos os conjuntos de nomes de localidade casados encontrados na etapa 868 são processados. Cada conjunto de nomes de localidade casados são localidades tendo nomes duplicatas ou levemente variantes. Na etapa 870, se
5 outro conjunto de nomes de localidade casados existir, o processo determina se nomes casados representam geometria de sobreposição na etapa 872. Na etapa 872, nomes casados representam geometria de sobreposição se as localidades se sobrepõem ou mesmo se são somente adjacentes entre si, desde que compartilhem pelo menos uma característica geográfica em comum determinada na etapa de otimização 840.

10 Se na etapa 872 da Figura 8B, os nomes casados representarem geometria de sobreposição, se na etapa 873, a geometria de sobreposição for exata, então na etapa 874, nomes duplicatas exceto um são eliminados a partir das entradas de índice de localidade no banco de dados geográficos. Se todas as características geográficas associadas a um nome de localidade forem iguais àqueles de outro, esses nomes de localidade são duplicatas
15 verdadeiras e todos exceto um são eliminados. Nomes de localidade são eliminados se e somente se os nomes representarem a mesma localidade. Essa etapa elimina localidades duplicatas e reduz o conjunto de nomes de localidade. Para um índice de localidade tendo muitas entradas duplicatas, essa técnica reduzirá muito a quantidade de indexação e espaço exigido pelo índice. No exemplo de Ho-Ho-Kus, Nova Jérsei, os tokens normalizados concatenados juntos para cada nome são ambos “HOHOKUS TOWNSHIP”. Como esses
20 dois nomes de localidade serão determinados como tendo todas as características geográficas em comum a partir da etapa de otimização, esses nomes de localidades são duplicatas verdadeiras e um é eliminado. O processo então retorna para a etapa 870 para determinar se outro conjunto de nomes de localidade casados existe.

25 Se na etapa 873 da Figura 8B a geometria de sobreposição não for exata, ou uma localidade compartilhar pelo menos um porém um número menor do que todas as características geográficas com outra localidade, normalmente uma localidade com um nome levemente diferente, essas localidades são consideradas como sendo a mesma localidade e são fundidas na etapa 875. Por exemplo, “Randolph” e “Randolph Center” em
30 Vermont são duas cidades separadas porém sobrepostas. Como as duas cidades se sobrepõem, compartilham pelo menos uma característica geográfica em comum, são consideradas como sendo a mesma localidade e são fundidas.

Em modalidades, a fusão de nomes de localidade somente ocorre quando as localidades de sobreposição não têm características de não sobreposição que não podem
35 ser distinguidas entre si. Por exemplo, se Randolph e Randolph Center tiverem ambas uma Main Street sem números de rua em sobreposição, as duas cidades podem ser fundidas. Se as duas cidades tiverem um “2 Main Street” por exemplo, entretanto, as cidades não devem

ser fundidas.

O seguinte exemplo de caso de uso ilustra a vantagem de eliminar todos exceto um dos nomes de localidade duplicatas a partir de múltiplas fontes que têm geometria de sobreposição. Sem essa característica, um nome de localidade é listado múltiplo em opções

5

Sem eliminar duplicatas:

Entrar cidade -> Hanover

Por favor escolher entre ->

Hanover (subdivisão de condado)

10

Hanover (lugar administrativo)

Hanover (03755)

Após eliminar duplicatas:

Entrar cidade -> Hanover

Encontrado: "Hanover"

15

O seguinte exemplo de caso de uso também ilustra a vantagem de fundir localidades tendo nomes levemente diferentes. Sem fusão, o usuário pode não saber qual nome levemente diferente é a localidade na qual um destino desejado está localizado. Com a fusão, o usuário não necessita distinguir entre nomes. Por exemplo, as localidades "Randolph", "Randolph Center" e "Randolf Township" se sobrepõem, e desse modo são

20

fundidos em uma área comum, representada pelo nome único "Randolph". Desse modo para uma busca de usuário:

Sem fusão:

Entrar cidade -> Randolph

Entrar rua -> Main Street

25

Por favor escolher entre ->

Main Street, Randolph

Main Street, Randolph Center

Main Street, Randolph Township

Com a fusão:

30

Entrar cidade -> Randolph

Entrar rua -> Main Street

Encontrado: "Main Street, Randolph"

35

Na etapa 876 da Figura 8B, uma união de todas as características a partir dos nomes casados é atribuída ao nome fundido. Por exemplo, em FIPS55, a Subdivisão de Condado de Boston define certa geografia. O Lugar Administrativo de Boston define outra geografia que sobrepõe porém não é necessariamente igual. O lugar postal de Boston define um terceiro conjunto de geografia cobrindo ruas nas quais a correspondência norte-

americana pode ser entregue. A criação de uma união dessas características diferentes forma um conjunto completo de características que são associadas à Boston. A união das características geográficas associadas a cada um desses nomes relacionados à Boston compreende um conjunto das características geográficas incluindo cada uma dessas fontes.

5 Por exemplo, se Adams St. For de interesse para um usuário final, embora Adams St. não faça parte do lugar postal Boston, Adams St. será encontrado para o usuário porque faz parte da Subdivisão do Condado de Boston devido à união de características geográficas a partir do casamento de nomes de localidade de várias fontes de nomes de localidade. Desse modo, uma lista de nomes de localidade exclusivos resulta, com conjunto de bits em
10 uma máscara de fonte correspondendo às fontes nas quais cada nome é encontrado, e uma união de todas as características geográficas às quais cada nome se aplica. O processo então retorna para a etapa 870 para determinar se outro conjunto de nomes de localidade casados existe.

A Figura 11 mostra uma modalidade de um algoritmo para reduzir o conjunto de
15 nomes de localidade através do casamento de nomes de localidade. Para cada nome de localidade A em uma fonte de nomes de localidade, para cada nome B em quaisquer outras fontes que casam com o nome A, atribuir a A quaisquer lados de rua de segmento associados a B ainda não atribuídos a A. Essa é a etapa 876 da Figura 8B acima. Incluir quaisquer bits na máscara de fonte B não incluídos ainda na máscara de fonte A e deletar B.

20 Na etapa 872 da Figura 8B, se os nomes casados não representarem geometria de sobreposição, os nomes casados são adornados para tornar os mesmos distintos na etapa 878. Os nomes casados que não representam geometria de sobreposição são localidades tendo nomes duplicatas ou levemente variantes que são fisicamente separadas. Em modalidades, essas localidades fisicamente separadas são cidades com nomes iguais ou
25 levemente diferentes. Genericamente, tais localidades com nomes duplicatas existem em diferentes condados em um estado. Desse modo, esses nomes duplicatas podem ser distinguidos para um usuário mostrando um adorno, por exemplo o nome de condado no qual a localidade está localizada. Um adorno de localidade é tipicamente mostrado em parênteses ou em aspas próximo ao nome de localidade. Nomes de condado ou outros
30 adornos de borda, entretanto, podem não ser reconhecíveis para os usuários não locais. Em vez disso, os nomes de cidades grandes, facilmente reconhecíveis próximo a cada localidade tendo nomes duplicatas fornecerão melhores informações para o usuário. Desse modo, na etapa 878, um adorno de cidade separado é armazenado no índice de localidade para cada um dos nomes a partir da etapa 872. Informações mais detalhadas em relação à
35 criação desse tipo de adorno podem ser encontradas no pedido número 11/345.877, depositado em 1º de fevereiro de 2006, intitulado "Method for differentiating duplicate or similarly named disjoint localities within a state or other principle geographic unit of interest",

de Michael Geilich. O processo então retorna para a etapa 870 para determinar se outro conjunto de nomes de localidade casados existe.

O seguinte exemplo de caso de uso mostra adornos para localidades separadas tendo nomes duplicatas ou levemente variantes:

5 Adornar com nomes de condado:

Entrar estado -> PA

Entrar cidade -> Bethel

Por favor escolher entre ->

Bethel (Berks)

10 Bethel (Allegheny)

Bethel (Lancaster)

Bethel (Mercer)

Bethel (Sullivan)

Bethel (Wayne)

15 Adornar com nomes de cidades grandes, próximas, facilmente reconhecíveis:

Entrar estado -> PA

Entrar cidade -> Bethel

Por favor escolher entre ->

Bethel (Fredericksburg)

20 Bethel (Pittsburgh)

Bethel (Lancaster)

Bethel (Youngstown)

Bethel (Williamsport)

Bethel (Scranton)

25 Nesse exemplo de caso de uso, a aplicação processa cada entrada de usuário antes de solicitar mais informações a partir do usuário. Em outras modalidades, para “Adornar com nomes de cidades grandes, próximas, facilmente reconhecíveis” se o usuário entrar o estado, cidade e nome de rua antes da aplicação processar essas três entradas de usuário, um destino único pode ser determinado se o endereço de rua for encontrado
30 somente em uma das opções. Por exemplo:

Adornar com nomes de cidades grandes, próximas, facilmente reconhecíveis:

Entrar estado -> PA

Entrar cidade -> Bethel

Entrar nome de rua -> Main Street

35 Encontrado: ‘Main Street, bethel (Fredericksburg)’

Se na etapa 870, outro conjunto de nomes de localidade casados não existir, então na etapa 880 da Figura 8B, o índice é criado. O índice é primeiramente ordenado por

característica geográfica. Para cada característica geográfica, localidades que contêm a característica geográfica são indexados em ordem de prioridade. Nomes de localidade no índice são ordenados por prioridade para permitir que desenvolvedores de aplicações programem seleção dos nomes mais prevalentes para qualquer característica geográfica nas aplicações. Isso provê aos usuários finais os nomes mais prevalentes a partir dos quais selecionar por exemplo, em ambientes de memória limitada. Para um dispositivo de memória limitada que pode armazenar somente alguns nomes de localidade para cada característica geográfica, um desenvolvedor de aplicações pode utilizar o índice de localidade para escolher as localidades com prioridade mais elevada para o usuário para uma característica geográfica associada a mais de algumas localidades. Similarmente, para aplicações de busca bottom-up, a aplicação solicita o endereço, ou característica geográfica, a partir do usuário e apresenta uma lista de localidades a partir da qual o usuário escolhe. Ao apresentar a lista de localidades, os nomes com prioridade mais elevada associados ao endereço podem ser utilizados.

Em modalidades, ordem de prioridade das localidades associadas a uma característica geográfica se baseia na prevalência de cada nome de localidade em uso comum para uma aplicação pretendida. Em modalidades, a priorização com base em uso comum permite que nomes de localidade sejam ordenados de forma diferente para usuários diferentes. No exemplo de localidades de sobreposição como “Cidade de Nova York”, “Manhattan” e “SoHo”, em uso comum, um usuário local conheceria bem a área, mais provavelmente usaria a mais específica das três localidades, ou “SoHo”. Se uma aplicação for destinada a esse usuário local, o nome de localidade de prioridade mais elevada seria mais provavelmente um tendo o número mínimo de fontes no qual o nome de localidade pode ser encontrado. Desse modo, a ordem de prioridade a partir do mais elevado para o mais baixo seria “SoHo”, “Manhattan” a seguir “Cidade de Nova York.”

Utilizando o mesmo exemplo de localidades em sobreposição na cidade de Nova York, em uso comum, um usuário não local não conhece bem a área local, entretanto, mais provavelmente utilizaria a localidade mais conhecida, facilmente reconhecível. Se uma aplicação for destinada a esse usuário não local, o nome de localidade de prioridade mais elevada mais provavelmente seria um tendo o maior número de fontes no qual o nome de localidade pode ser encontrado. Desse modo, a ordem de prioridade a partir do mais elevado para o mais baixo seria “Cidade de Nova York”, “Manhattan”, então “SoHo”.

Em modalidades, algoritmos para determinar ordem de prioridade em uma aplicação podem ser aplicados diferentemente para atender usos comuns diferentes para um usuário. Por exemplo, para um usuário local navegando em uma localidade como uma cidade grande, o usuário poderia querer uma prioridade de nomes de localidade com base em uso comum para um usuário local. Enquanto o mesmo usuário que navega para a

mesma cidade grande de longe, entretanto, o usuário pode querer uma prioridade diferente com base em uso comum para um usuário não local. Após o usuário chegar na cidade grande e cruzar o limite para dentro da cidade, entretanto o usuário pode querer que a prioridade mude de volta para aquela de um usuário local.

5 Muitos esquemas de ordenação de prioridade diferentes são possíveis. Na modalidade preferida, a localidade de prioridade mais elevada associada a uma característica geográfica é aquela encontrada em uma fonte de nome postal preferida, então a prioridade das localidades restantes é determinada pelo número de conjuntos de bits em cada máscara de fonte de localidade. Em modalidades, uma primeira localidade tem uma
10 prioridade mais elevada do que a segunda localidade se a primeira localidade for mais bem conhecida ou prevalente em uso comum. Em modalidades, a prioridade de um nome de localidade é determinada pelo número de fontes no qual o nome pode ser encontrado. O nome de localidade para uma característica geográfica com a prioridade mais elevada é o nome de localidade que pode ser encontrado no número maior de fontes, e desse modo
15 que, tem o conjunto de mais bits em sua máscara de fonte. A ordem de prioridade dos nomes de localidade para uma característica geográfica é da mais elevada para a mais baixa.

Em modalidades, um desenvolvedor de aplicações também pode utilizar a máscara de fonte para cancelar esse esquema de prioridade default por preferir certas fontes de
20 nomes de localidade em relação a outras. Em outras modalidades, a prioridade é definida em termos do tamanho de localidade física maior ou população de localidade maior. Em outras modalidades, prioridade é definida como o número maior de características geográficas, por exemplo, segmentos de rua, em uma localidade. A prioridade pode ser definida também em termos do número maior de características geográficas principais
25 localizadas na localidade, ao contrário do número de características geográficas localizadas na localidade, em outras modalidades. Um exemplo de uma característica geográfica principal é uma rodovia importante. Em modalidades, a prioridade pode ser definida utilizando as máscaras de fonte de localidade para determinar uma preferência de certas fontes de nomes de localidade em relação a outras. Em modalidades, um desenvolvedor de
30 aplicações pode utilizar nomes de localidade a partir de fontes de localidade indicadas como "Trump" na Figura 10 como os nomes de prioridade máxima.

Em modalidades, no caso de empates de prioridade de localidade, uma separação primária é executada utilizando um dos esquemas acima, e onde necessário, por uma separação secundária com base em um dos esquemas acima. Na modalidade preferida,
35 uma separação primária é executada no número de fontes a partir de mais elevada para mais baixa na qual cada localidade pode ser encontrada. Uma separação secundária se baseia, por exemplo, no número de características geográficas, ou segmentos de rua, a

partir da mais elevada para a mais baixa contida em cada localidade.

A Figura 12 mostra uma modalidade de um algoritmo para determinar a prioridade de nomes de localidade para uma dada característica geográfica. Para cada lado de segmento de rua S em um banco de dados geográfico, encontrar todos os nomes de localidade A para os quais S é atribuído. Para cada A, encontrar o nome A com o conjunto de mais bits em sua máscara de fonte. Atribuir A para o nome de prioridade mais elevada seguinte no índice para esse lado de segmento de rua S.

O processo da Figura 8B termina na etapa 890.

A Figura 13 mostra uma modalidade de arquivos de índice de localidade incluindo uma tabela de Prioridade de Localidade de Característica, uma tabela de Nome de Localidade e uma Tabela Encontrar característica. Essas tabelas são finalmente armazenadas em um banco de dados. Em modalidades, na tabela de Prioridade de Localidade de Característica da Figura 13, lista localidades por prioridade para cada característica geográfica. Em modalidades, cada característica geográfica na tabela é associada a um número ID de característica, FF_ID. Os números de ID de característica podem ser seqüenciais porém não têm necessariamente de ser seqüenciais. Os números de ID de característica são também um link para a tabela Encontrar Característica. Em modalidades, cada localidade associada a cada característica geográfica na tabela também é associada a um número de ID de localidade, NAME_ID. Os números de ID de localidade podem ser seqüenciais porém não têm necessariamente de ser seqüenciais. O campo PRIORIDADE indica a prevalência do nome de localidade associado à característica geográfica. Como mencionado acima, muitos esquemas de prioridade existem para priorizar os nomes de localidade associados a cada característica geográfica. PRIORIDADE é um número seqüencial que inicia com "1" como a prioridade mais elevada. A tabela também contém a máscara de fonte de nome de localidade para essa localidade, LOC_MASK, descrito acima.

O formato variável do índice de localidade permite que qualquer número de entradas de tabela seja incluído para cada característica geográfica na tabela de Prioridade de Localidade de Característica. Isso é especialmente importante na América do Norte para nomes postais. Embora haja genericamente somente um nome de localidade postal preferido para cada local, o serviço postal também inclui qualquer número de nomes de localidade postal permissíveis para o mesmo local. O índice de localidade inclui todos os nomes postais preferidos e permissíveis para cada característica geográfica.

Em modalidades, a tabela de Nome de Localidade da Figura 13 é ligada à tabela de Prioridade de Localidade de Característica através dos números de ID de localidade, NAME_ID. A tabela também contém o nome completo da localidade, FULL_NAME, utilizando letras de caso misturado em modalidades. Em modalidades, os nomes de

localidade completos como representados em FIPS55 são utilizados para a codificação final de nomes de localidade completos nessa tabela. Outras fontes para representar nomes de localidade completos podem ser entretanto, utilizados. O campo NAME_KEY da tabela é o componente significativo do nome de localidade para fins de indexar. Em modalidades, 5 NAME_KEY é encontrado a partir de tokenização e normalização do nome de localidade acima. Isso permite que “Hollywood” e “West Hollywood” sejam ambos indexados sob “H”, por exemplo, visto que o token de corpo principal para ambos é “Hollywood”. O campo ADORNO é um pointer para outra entrada na Tabela de Nome de Localidade contendo o nome de localidade de um local ou cidade grande e facilmente reconhecível próximo da 10 localidade. Em modalidades, ADORNMENT é armazenado na tabela somente quando a localidade é uma localidade ambígua em uma subdivisão primária de um condado, como um estado. Em modalidades, o adorno é utilizado para diferenciar localidades duplicatas em uma lista em um sistema ou dispositivo de usuário.

O campo NAME_LC é um código de três caracteres para o idioma do nome de 15 localidade. Em modalidades, NAME_LC é definido para cada nome de localidade para indicar o idioma nativo do nome para suportar países de múltiplos idiomas. Em modalidades, NAME_LC pode ser qualquer número de caracteres. LOC_SIZE indica uma contagem do número de características geográficas associadas a esse nome de localidade e pode ser utilizado por desenvolvedores de aplicações para cancelar o esquema de PRIORITY default 20 fornecido na tabela de prioridade de Localidade de Característica. COUNTRY é um código de país e há uma abreviatura de três caracteres do país no qual a localidade está localizada. Em modalidades, COUNTRY pode ser um código de país padrão como ISO 3166-1, que faz parte do padrão ISO 3166 publicado pela primeira vez pelo International Organization for Standardization. Em modalidades, COUNTRY pode ser qualquer número de caracteres. 25 CNETER_ID é um link com características de ponto de centro de cidade encontradas em outro lugar no banco de dados geográficos para essa localidade. Em modalidades, essas características de ponto de centro de cidade são as coordenadas de latitude e longitude de ponto central de localidade, bem como um segmento de rua correspondendo ao centro de cidade. Centros de cidade fornecem um ponto em uma localidade para um usuário quando 30 um endereço de rua específico não é solicitado ou não pode ser encontrado.

Em modalidades, a tabela de Nome de Localidade da Figura 13 pode conter muitos outros tipos úteis de informações sobre localidades. Por exemplo, a inclusão de fonemas na tabela de Nome de Localidade seria útil para aplicações de texto para fala, onde um fonema é um conjunto de sons de fala ou elementos de sinal que são equivalentes de forma 35 cognitiva. Outros exemplos de tipos diferentes de informações que podem ser armazenadas na tabela de Nome de Localidade são uma imagem da prefeitura de uma localidade e o número de telefone do departamento de polícia de uma localidade.

Em modalidades, a tabela Encontrar característica da Figura 13 contém informações sobre cada característica geográfica. FF_ID é um número de ID de característica utilizado para ligar informações de características geográficas à tabela de Prioridade de Localidade de Característica. FEAT_TYPE é o tipo de característica geográfica, como “R” para características de estrada e “F” para características de linha de ferry. FEAT_ID é um link para informações no banco de dados geográfico sobre a característica como nomes de ruas e faixas de endereço. FEAT_ID também provê ligação indireta com outro conteúdo ligado ao banco de dados geográfico como Pontos de Interesse. SIDE é o lado da característica geográfica, por exemplo, uma margem de rua. SIDE inclui “R” para o lado direito, “L” para o lado esquerdo, “B” para os dois lados e “nulo” para “não aplicável.”

Em modalidades, o índice de localidade é fornecido em múltiplos formatos, incluindo formatos internacionais, para permitir integração fácil com bancos de dados geográficos de propriedade. O índice de localidade é fornecido para acomodar dados a partir de qualquer país. Embora o formato seja generalizado, o conteúdo é moldado para incluir fontes de localidade específicas e tipos apropriados em cada país. Uma aplicação de propriedade provê a pronúncia para cada nome de localidade.

Em modalidades, para uso de tabela de índice de localidade, em uma implementação top-down de encontrar um endereço, a localidade é primeiramente resolvida, e então a característica geográfica correta é encontrada na localidade. Uma aplicação de navegação executará primeiramente casamento de nome para encontrar o nome de localidade desejado na tabela Nome de Localidade. Após encontrar a localidade, a tabela Prioridade de Localidade de Característica é buscada utilizando o NAME_ID da localidade escolhida para determinar as características geográficas contidas naquela localidade. Os FF_IDs daquelas características são utilizados como índice na tabela Encontrar Característica para recuperar informações sobre essas características necessárias para encontrar uma característica específica, como nomes de rua e faixas de endereço no caso de segmentos de rua, e então o casamento é executado para selecionar a característica geográfica específica desejada. Por exemplo, [Enter City-> Boston]. “Boston” é casado com os nomes na Tabela Nomes de Localidade, retornando o NAME_ID para “Boston”. [Enter Street -> Adams]. A Tabela de Prioridade de Localidade de Característica é busca para uma lista de FF_IDs cujo NAME_ID é o NAME_ID para “Boston.” A Tabela Encontrar característica é buscada para o FEAT_ID que aponta para “Adams” no banco de dados geográficos. Subseqüentemente, o número de casa desejado pode ser solicitado a partir do usuário e a Tabela Encontrar Característica é buscada para o FEAT_ID que aponta para a faixa de endereços contendo o número de casa solicitado no banco de dados geográficos. A Tabela Encontrar característica pode ser buscada para o FEAT_ID que aponta para o ponto

de latitude e longitude para essa característica no banco de dados geográfico, para exibir para o usuário a localização da característica em um dispositivo ou aplicação de navegação, por exemplo. Para desempenho aperfeiçoado, o índice de localidade será freqüentemente pré-compilado para eliminar muitas dessas referências indiretas.

5 Em modalidades, para uso de índice de localidade, em uma implementação bottom-up de encontrar endereços, uma lista de características geográficas alvo é escolhida primeiramente, então a característica correta é selecionada por resolver a localidade desejada a partir da lista de todas as localidades contendo uma característica por aquele nome. Uma aplicação de navegação executará primeiramente o casamento para encontrar
10 uma lista de características geográficas na tabela Encontrar Característica. Os FF_IDs correspondentes a partir da tabela Encontrar característica são então utilizados como índices para a tabela de Prioridade de Localidade de Característica. As entradas na tabela de prioridade para esses FF_IDs podem ser então varridas para um NAME_ID cujo nome na tabela de Nome de Localidade casa com a localidade desejada. Se o desenvolvedor de
15 aplicações desejar apresentar opções de localidade para o usuário, a aplicação deve considerar os NAME_IDs de localidade em ordem de prioridade, escolhendo os nomes de localidade com prioridade mais elevada que são exclusivos para os FF_IDs em consideração. Esses nomes podem ser então apresentados ao usuário a partir dos quais
20 escolher. Como no caso top-down, o índice de localidade será freqüentemente pré-compilado para eliminar muitas das referências indiretas entre as tabelas.

Em modalidades, o índice de localidade pode ser utilizado para encontrar lugares nomeados como pontos de interesse e marcos. Listas de tais lugares são primeiramente associadas a segmentos de rua a partir do banco de dados geográficos de propriedade. A aplicação casará então o nome do ponto de interesse ou marco desejado para encontrar o
25 segmento de rua. A aplicação utiliza então a implementação de encontrar endereços acima utilizando o segmento de rua em ordem para determinar a localidade correta.

Em modalidades, o índice de localidade pode ser utilizado para encontrar um centro de cidade. Uma aplicação casará em nome a localidade desejada utilizando FULL_NAME e NAME_KEY na tabela de Nome de Localidade para encontrar a entrada correta na tabela.
30 Após a entrada correta ser encontrada, o campo CENTER_ID é utilizado para encontrar a informação de centro de localidade de propriedade correspondente no banco de dados geográfico, como coordenadas de latitude e longitude ou segmento de rua correspondendo ao centro de cidade.

Em modalidades, o índice de localidade pode ser utilizado para tornar desambígua
35 localidade com nomes duplicatas, porém geografia distinta. Uma aplicação casará em nome a localidade desejada utilizando FULL_NAME e NAME_KEY na tabela de Nome de Localidade para encontrar a entrada correta na tabela. Por exemplo, se a localidade for

“Brentwood, Califórnia”, dois casamento serão encontrados como mostrado na Figura 4. O ADORNMENT a partir da tabela de Nome de Localidade será desse modo utilizado para cada localidade Brentwood, por exemplo adornos “Los Angeles” e “San Francisco.” Esses poderiam ser exibidos para um usuário como “Brentwood (Los Angeles)” e “Brentwood (San Francisco)” a partir dos quais o usuário pode escolher.

Em modalidades, o índice de localidade pode ser utilizado para resolver ambigüidade em características de endereço. Por exemplo, para o exemplo de “2 Adams Street” na Figura 3, a aplicação utilizará os nomes de localidades múltiplas ordenadas por PRIORIDADE para cada característica, para distinguir entre os quatro endereços “2 Adams Street” encontrados na localidade de Boston, Massachusetts. A aplicação primeiramente encontrará segmentos de endereço que correspondem a endereços duplicatas no banco de dados geográficos, utilizando o campo FEAT_ID da tabela Encontrar Característica. A aplicação encontrará então os FF_IDs correspondentes na tabela Encontrar Características. Os FF_IDs são então utilizados como índices para a tabela de Prioridade de Localidade de Característica. localidades são recuperadas em ordem a partir da prioridade mais elevada para a mais baixa utilizando PRIORIDADE até que um NAME_ID exclusivo seja encontrado para cada entrada FF_ID. Os NAME_IDs são utilizados como índices para a tabela Nome de Localidade para recuperar um nome de localidade exclusivo, FULL_NAME, para cada endereço duplicata. No exemplo para “2 Adams Street”, nomes de localidade exclusivos serão encontrados em Charlestown, Hyde Park, Roxbury e Dorchester, todas sub-localidades de Boston, Massachusetts.

Em modalidades, o índice de localidade pode ser utilizado para buscar áreas vizinhas para uma característica solicitada em uma aplicação top-down. Em alguns casos uma característica desejada pode não ser encontrada em uma localidade especificada por um usuário e a aplicação de navegação desejará expandir a busca para localidades de contenção maiores ou vizinhas. A aplicação primeiramente casará o nome da localidade desejada na tabela de Nome de Localidade, recuperando o NAME_ID correspondente. Após determinar que não há FF_IDs correspondendo à característica solicita na tabela de Prioridade de Localidade de Característica com esse NAME_ID de localidade, a aplicação encontrará um ou mais FF_IDs na tabela de Prioridade de localidade de característica que contém esse NAME_ID. A cadeia de prioridade pode ser seguida, prioridade mais elevada ou mais baixa, para esses FF_IDs na tabela de Prioridade de Localidade de característica para recuperar outros NAME_IDs que correspondem a esses FF_IDs. A tabela Encontrar Característica pode ser consulta para determinar se o endereço solicitado está dentro de qualquer uma dessas outras localidades relacionadas.

Em modalidades, o exemplo de caso de uso a seguir ilustra a vantagem da característica de priorização do índice de localidade. Sem priorização, não está claro para o

desenvolvedor de aplicações como usar o nome mais reconhecível ao consultar o usuário. Em alguns lugares, nomes postais são os mais comuns. Em outras áreas, nomes administrativos são bem conhecidos. Com a característica de priorização, o nome mais comum pode ser escolhido.

5 Sem priorização:
 Entrar rua -> Broadway
 Por favor escolher entre ->
 Broadway (Charlestown, MA)
 Broadway (Manhattan, NY)

10 Com priorização:
 Entrar rua -> Broadway
 Por favor escolher entre ->
 Broadway (Boston, MA)
 Broadway (New York, NY)

15 Em modalidades, em um exemplo de caso de uso adicional como ilustrado na Figura 14, uma aplicação de navegação pode acomodar inconsistência quando uma cidade próxima é erroneamente especificada. Cidades grandes como Chicago são genericamente circundadas por subúrbios. Os subúrbios são separados, e têm sua própria estrutura administrativa. Em particular, seus nomes de localidade diferem freqüentemente. Um usuário poderia não estar ciente da área suburbana, porém pensando somente na cidade central grande. Um exemplo é encontrado nos subúrbios ao norte de Chicago, como mostrado na Figura 14. Suponha que o usuário deseje localizar "Bryn Mawr Country Club" em Lincolnwood, porém somente conhece a área como Chicago. Se o usuário souber que o endereço da rua é "6600 North Crawford Ave." a entrada poderia prosseguir como a seguir:

25 Entrar estado -> Illinois
 Entrar cidade -> Chicago
 Entrar rua -> North Crawford Avenue

30 A aplicação de navegação observaria uma inconsistência aqui. A aplicação primeiramente buscará todos FF_IDs na tabela de Prioridade de Localidade de Característica onde o NAME_ID aponta para Chicago. A aplicação observará que "North Crawford Avenue" não existe em Chicago. A aplicação buscará todos os FF_IDs na tabela de prioridade de Localidade de Característica onde os FF_IDs aponta para "North Crawford Avenue." A aplicação encontrará "North Crawford Avenue" no subúrbio de Lincolnwood, Chicago. Se a aplicação tivesse encontrado "North Crawford Avenue" em várias localidades,
 35 a aplicação utilizaria o nome de localidade de prioridade mais elevado para esse FF_ID utilizando PRIORIDADE na tabela de Prioridade de Localidade de característica. A aplicação pode observar que "South Crawford Avenue" existe em Chicago. A aplicação então solicita o

número de rua:

Entrar número de rua -> 6600

Encontrado: "6600 North Crawford Avenue, Lincolnwood, Illinois"

5 Nesse exemplo, se o número de rua correto foi encontrado nos dois lugares, a aplicação poderia oferecer para o usuário uma escolha: "6600 South Crawford Avenue, Chicago" ou "6600 North Crawford Avenue, Lincolnwood." Uma vez que o número da rua "6600" não é encontrado em "South Crawford Avenue" em Chicago, essa opção de endereço não é exibida para o usuário. Embora o número da rua "6600" encontrado para "North Crawford Avenue" seja localizado em Lincolnwood e não em Chicago, a aplicação
10 pode assumir que é o endereço que o usuário pretende solicitar.

Em modalidades, em um exemplo de caso de uso adicional, a aplicação pode fornecer manipulação de se uma de entradas de um usuário para a rua ou para a cidade é incompatível e deve ser fixa. O endereço para Chandler Music Hall em seu website é "71-73 Main Street, Randolph, Vermont." Na cidade de Randolph, Main Street é dividido em "North
15 Main Street" e um "South Main Street." "Main Street" também existe na cidade próxima de Randolph Center. Para o usuário final, se a rua for realmente Main Street, então o Hall deve estar em Randolph Center. Se o Hall estiver em Randolph, então é localizado em North Main Street ou em South Main Street. O Hall é na realidade localizado em Randolph, em 71 North Main Street. Se um usuário final estivesse utilizando o endereço de website em uma
20 aplicação top-down, o usuário seria corretamente levado a partir de Randolph para North ou South Main Street, porém a aplicação pediria a ele para uma decisão porque o número de rua 71 existe nas duas ruas. Se o usuário estivesse utilizando o endereço de website em uma aplicação de bottom-up, o usuário seria levado incorretamente de Main Street para Randolph Center. Em modalidades, um modo para uma aplicação de navegação tratar
25 desse tipo de situação é apresentar todas as opções para o usuário:

Entrar estado -> Vermont

Entrar cidade -> Randolph

Entrar rua -> Main Street

Entrar número de rua -> 71

30 Por favor escolher entre ->

71 North Main Street, Randolph

71 South Main Street, Randolph

71 Main Street, Randolph Center

35 Em modalidades, uma ou mais etapas da presente invenção são realizadas automaticamente. A característica automática é implementada utilizando software apropriado. A característica automática cria um aumento substancial em eficiência e velocidade com as quais os índices de localidade são criados.

As modalidades da presente invenção com modificação podem ser aplicadas em dispositivos e aplicações não de navegação. Por exemplo, em uma aplicação espacial de Páginas Amarelas, é desejável encontrar todas as firmas de um certo tipo separadas por distância a partir de um ponto. Em modalidades, a indexação de localidades para esse tipo de aplicação pode utilizar um esquema de prioridade com base em frequência de ocorrência em um catálogo de Páginas amarelas.

A Figura 15 mostra um diagrama de blocos de um sistema exemplar 900 que pode ser utilizado com modalidades da presente invenção. Embora esse diagrama represente componentes como logicamente separados, tal representação é meramente para fins ilustrativos. Será evidente para aqueles versados na técnica que os componentes retratados nessa Figura podem ser combinados ou divididos em componentes separados de software, firmware e/ou hardware. Além disso, será também evidente para aqueles versados na técnica que tais componentes, independente de como são combinados ou divididos, podem executar no mesmo sistema/dispositivo de computação ou podem ser distribuídos entre sistemas/dispositivos de computação diferentes conectados por uma ou mais redes ou outros meios de comunicação apropriados.

Como mostrado na Figura 15, o sistema 900 inclui tipicamente um dispositivo de computação 910 que pode compreender uma ou mais memória 912, um ou mais processadores 914, e um ou mais dispositivos de armazenagem ou repositórios 916 de algum tipo. O sistema 900 pode incluir ainda um dispositivo de exibição 918, incluindo uma interface de usuário gráfico ou GUI 920 operando no mesmo pelo qual o sistema pode exibir mapas, e outras informações para um usuário. O usuário utiliza o dispositivo de computação para solicitar, por exemplo, que uma localidade seja exibida em um mapa ou que orientações de direção sejam exibidas como uma rota em um mapa e/ou como direções de texto. A GUI 920 exibe um exemplo de um par de localidades duplicatas para "Washington, Nova Jérsei" e seus adornos "Easton" e "Hammonton." O usuário selecionará uma das localidades duplicatas a serem exibidas para GUI 920.

Um banco de dados geográficos 930 é mostrado como armazenagem externa para o dispositivo ou sistema de computação 910, porém o banco de dados geográficos 930 em algumas ocorrências pode ser a mesma armazenagem que a armazenagem 916. Em modalidades, entradas de nome de localidade são fundidas para localidades variantes e duplicatas 932 em banco de dados geográficos 930. Em modalidades, o banco de dados geográficos 930 contém uma máscara de fonte principal de fontes de localidade 934. Em modalidades, um índice de localidade incluindo tabelas de Prioridade de Localidade de característica, Nome de localidade e Encontrar característica 936 é armazenado no banco de dados geográficos 930.

Software de criação de banco de dados geográficos de propriedade 930 pode

utilizar fontes e definições de localidade de mundo real 960 para fundir e/ou adornar as entradas de nome de localidade variantes e duplicatas 932, criar a máscara de fonte principal de fontes de localidade 934 e criar o índice de localidade 936. Os exemplos de fontes de localidade de mundo real e definições são descritos acima na discussão para a

5 Figura 2. Informações a partir do banco de dados geográficos 930 são utilizadas por um software de aplicação de dispositivo e conversor de banco de dados geográficos em aplicação 950, que é finalmente utilizado por um usuário do dispositivo de computação 910. O software de aplicação de dispositivo e conversor de banco de dados geográficos em aplicação 950 é mostrado remoto ao dispositivo de computação de usuário 910 porém

10 também pode residir no dispositivo de computação de usuário 910.

Para um exemplo de um software de aplicação de dispositivo e conversor de banco de dados geográficos em aplicação 950, como utilizado por um usuário na Internet, ou em um dispositivo de navegação, o usuário pode selecionar uma localidade a ser exibida em um mapa. Alternativamente, se o usuário solicitar orientações de direção, por exemplo, a

15 localidade pode ser a localidade de partida ou término.

Em modalidades, o tipo de aplicação de software que consulta o usuário pode ser uma aplicação drill-down, top-down ou bottom-up. A abordagem drill down é útil em sistemas de navegação baseados em carro com memória limitada. Em modalidades úteis para dispositivos de memória limitada, o desenvolvedor de aplicações pode incluir no dispositivo somente nomes de localidade que têm classificação elevada em prioridade. Uma aplicação top-down solicita primeiramente ao usuário para entrar uma característica geográfica principal, por exemplo, um estado ou província. A aplicação então solicita que o usuário entre uma localidade, por exemplo uma cidade ou cidade pequena, localizada na característica geográfica principal. A aplicação solicita então que o usuário entre o nome da

20 rua na localidade. Finalmente, a aplicação solicita que o usuário entre o número de rua. Na maioria dos casos, as consultas resultam em especificação de uma característica de banco de dados geográficos não ambígua para uso por uma aplicação, por exemplo, exibir a localidade para o usuário em GUI 920 do dispositivo de exibição 918. Uma aplicação bottom-up solicita primeiramente que o usuário entre um número de casa e nome de rua. A

25 aplicação então exibe todas as localidades nas quais um endereço pode ser encontrado. Finalmente, a aplicação solicita que o usuário escolha ou entre o nome da localidade desejada. A metodologia bottom-up também resulta normalmente em especificação de uma característica de banco de dados geográficos não ambígua que pode ser então utilizada pela aplicação.

35 Em modalidades, o software de aplicação pode utilizar o índice de banco de dados geográficos em uma aplicação drill-down, que permite ao usuário final entrar um nome de localidade parcial ou total, normalmente em um estado dado. Em modalidades, a aplicação

apresenta nomes para o usuário final que casam com a entrada do usuário, e o usuário escolhe a melhor opção. Com o casamento contra os corpos de nome tokenizados, a aplicação pode apresentar tanto “Hollywood” como “West Hollywood” quando quaisquer das primeiras letras de “Hollywood” são entradas pelo usuário final.

5 Em outras modalidades, a aplicação de software não é uma aplicação drill-down e em vez disso consulta o usuário em relação ao número de rua e rua, localidade e característica geográfica principal em um momento. Na maioria dos casos, a consulta resulta em especificação de uma característica de banco de dados geográficos não ambígua, e o processo retorna a localização para o usuário. Se o usuário entrar um nome
10 de rua de “Main street” e uma localidade de “Springfield”, uma localidade duplicata “Springfield” será encontrada se também tiver uma rua pelo nome de “Main Street.” Se localidades duplicatas existirem para a característica geográfica, então uma lista de localidades e seus adornos pode ser exibida para o usuário para pedir ao usuário para escolher um, como em GUI 920 do dispositivo de exibição 918. Para um exemplo de par de
15 localidades duplicatas para “Washington, Nova Jérsei”, as duas localidades podem ser adornadas com os condados nos quais são encontradas ou com nomes de cidades maiores próximas. “Easton, Nova Jérsei” e “Hammonton, Nova Jérsei” respectivamente são cidades grandes próximas das duas localidades duplicatas. Desse modo, “Washington (Easton), NJ” e “Washington (Hammonton), NJ” são exibidas para a GUI 920 da Figura 15. Nesse
20 exemplo, os adornos são apresentados em parênteses porém podem ser apresentados de outras maneiras, como pelo uso de vírgulas para separar cada localidade duplicata a partir de seu adorno respectivo. O usuário seleciona uma das localidades duplicatas, e a localidade em um mapa ou orientações de direção são então exibidas para o usuário.

A codificação de software apropriado pode ser facilmente preparada por
25 programadores especializados com base nos ensinamentos da presente revelação, como será evidente para aquele versados na técnica de software. As modalidades da presente invenção podem ser implementadas também pela preparação de circuitos integrados de aplicação específica ou por interconexão de uma rede apropriada de circuitos de componentes convencionais, como será prontamente evidente para aqueles versados na
30 técnica.

As modalidades da presente invenção podem incluir um produto de programa de computador que é um meio de armazenagem (meios) tendo instruções armazenadas no mesmo que pode ser utilizado para programar um computador para executar quaisquer dos processos de modalidades da presente invenção. O meio de armazenagem pode incluir,
35 porém não é limitado a, qualquer tipo de disco incluindo discos flexíveis, discos ópticos, DVD, CD-ROMs, microdrive, e discos magneto-ópticos, ROMs, RAMs, EPROMs, EEPROMs, DRAMs, VRAMs, dispositivos de memória flash, cartões ópticos ou magnéticos,

nanossistemas, incluindo lcs de memória molecular, ou qualquer tipo de sistema ou dispositivo apropriado para armazenar instruções e/ou dados.

5 Armazenados em qualquer um dos meios legíveis por computador (meios), as modalidades da presente invenção podem incluir software para controlar tanto o hardware do computador especializado/propósito geral ou microprocessador, e para permitir que o computador ou microprocessador interaja com um usuário humano ou outro mecanismo utilizando os resultados de modalidades da presente invenção. Tal software pode incluir, porém não é limitado a, acionadores de dispositivos, sistemas operacionais, e aplicações de usuário. Finalmente, tais meios legíveis por computador incluem ainda software para
10 executar modalidades da presente invenção, como descrito acima.

Incluído na programação ou software do microprocessador ou computador especializado/propósito final estão módulos de software para implementar os ensinamentos da presente invenção. As modalidades da presente invenção podem ser convenientemente implementadas utilizando um microprocessador ou computador digital especializado ou de
15 propósito geral convencional de acordo com os ensinamentos da presente revelação, como será evidente para aqueles versados na técnica de computador.

A descrição acima da presente invenção foi fornecida para fins de ilustração e descrição. Não se pretende que seja exaustiva ou limite as modalidades da presente invenção a formas precisas reveladas. Muitas modificações e variações serão evidentes
20 para um técnico especializado na arte. As modalidades foram escolhidas e descritas para explicar melhor os princípios da presente invenção e sua aplicação prática, desse modo permitindo que outros versados na técnica entendam a presente invenção para várias modalidades e com várias modificações que são apropriadas para o uso específico considerado. Pretende-se que o escopo da presente invenção seja definido pelas
25 reivindicações a seguir e seus equivalentes.



PI0709707-7

REIVINDICAÇÕES

1. Índice de localidade de banco de dados geográfico, armazenável em um meio de armazenagem, **CARACTERIZADO** por compreender:

5 um ponteiro para pelo menos um aspecto geográfico em um banco de dados geográfico; e

um conjunto de um ou mais nomes de localidade associado à pelo menos um aspecto geográfico, onde um ou mais nomes de localidade são selecionados a partir de uma ou mais fontes de nomes de localidade e são ordenados por prioridade com base em prevalência de um ou mais nomes de localidade em uso comum para uma aplicação pretendida.

10 2. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que os aspectos geográficos compreendem ruas, segmentos de rua, margens de segmento de rua, partes dianteiras de bloco, marcos, parques estaduais, rodovias, centros de remessa, linhas ferroviárias, rotas de ônibus, centros de remessa, locais comerciais e locais residenciais.

15 3. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** por compreender ainda uma máscara de fonte principal criada por alocar um bit para cada uma ou mais fontes de nome de localidade utilizadas no índice.

20 4. Índice, de acordo com a reivindicação 3, **CARACTERIZADO** por compreender ainda uma máscara de fonte de localidade para cada localidade associada a cada aspecto geográfico, onde cada bit na máscara de fonte de localidade é definido se a localidade pode ser encontrada na fonte para a qual um bit correspondente foi alocado na máscara de fonte principal.

5. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a ordem de prioridade pode ser aplicada diferentemente para atender usos comuns diferentes.

25 6. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o uso comum para uma aplicação pretendida compreende o número mínimo de fontes no qual um nome de localidade pode ser encontrado se a aplicação for destinada a um usuário local.

30 7. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o uso comum para uma aplicação pretendida compreende o número máximo de fontes no qual um nome de localidade pode ser encontrado se a aplicação for destinada para um usuário não local.

8. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a prioridade de nomes de localidade para um aspecto geográfico com base em prevalência de cada nome de localidade em uso comum para uma aplicação pretendida compreende uma determinação da localidade com prioridade mais elevada associada a um aspecto geográfico para ser a localidade encontrada em uma fonte de nome postal preferida, então uma determinação de prioridade das localidades restantes associadas ao aspecto geográfico para ser pelo número de bits definidos em cada máscara de fonte de localidade, onde para as

localidades restantes, quanto maior o número de fontes de nome na máscara de fonte para a localidade, mais elevada a prioridade da localidade.

9. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a prioridade de nomes de localidade para um aspecto geográfico com base em prevalência de cada nome de localidade em uso comum para uma aplicação pretendida compreende uma determinação de um número de fontes de nome de localidade na qual a localidade pode ser encontrada a partir da máscara de fonte associada à localidade, em que quanto maior o número de fontes de nome na máscara de fonte para a localidade, mais elevada a prioridade da localidade.

10. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que uma prioridade alternada de nomes de localidades para um aspecto geográfico compreende ser baseado em uma determinação de um entre:

um número de aspectos geográficos em cada localidade, em que quanto maior o número de fontes geográficas na localidade, mais elevada a prioridade da localidade;

um tamanho físico de cada localidade, em que quanto maior o tamanho físico da localidade, mais elevada a prioridade da localidade; e

um tamanho de população de cada localidade, em que quanto maior o tamanho da população da localidade, mais elevada a prioridade da localidade.

11. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que uma prioridade alternada de nomes de localidades para um aspecto geográfico compreende ser baseado em uma determinação de uma preferência de uma certa fonte de nome de localidade em relação a outras utilizando as máscaras de fonte de localidade, em que as localidades tendo um conjunto de bits em suas máscaras de fonte de localidade para certa localidade têm uma prioridade mais elevada do que localidades que não têm.

12. Índice, de acordo com a reivindicação 3, **CARACTERIZADO** pelo fato de que a máscara de fonte principal compreende uma fonte de confiança, onde uma prioridade alternativa de nomes de localidade para um aspecto geográfico compreende ser baseado na fonte de confiança, em que localidades tendo um conjunto de bits em suas máscaras de fonte de localidade para a fonte de confiança têm uma prioridade mais elevada do que localidades que não têm.

13. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que se uma determinação de prioridade de nomes de localidade para um aspecto geográfico resultar em um empate entre localidades, prioridade das localidades de empate compreende ser baseado em uma determinação de um entre:

um número de aspectos geográficos em cada localidade de empate, onde quanto maior o número de fontes geográficas na localidade de empate, mais elevada a prioridade da localidade de empate;

um tamanho físico de cada localidade de empate, onde quanto maior o tamanho físico da localidade de empate, mais elevada a prioridade da localidade de empate;

um tamanho de população de cada localidade de empate, onde quanto maior o tamanho da população da localidade de empate, mais elevada a prioridade da localidade de empate; e

uma preferência de uma certa fonte de nome de localidade em relação a outras utilizando as máscaras de fonte de localidade, em que localidades de empate tendo um conjunto de bits em suas máscaras de fonte de localidade para certa localidade têm uma prioridade mais elevada do que localidades de empate que não têm.

14. Índice, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de que a associação de um nome de localidade a partir de um ou mais nomes de localidade com pelo menos um aspecto geográfico compreende associação direta ou indireta.

15. Índice, de acordo com a reivindicação 14, **CHARACTERIZADO** pelo fato de que a associação direta compreende uma fonte de nome de localidade específica associada a aspectos geográficos em geral, casando quaisquer aspectos geográficos associados ao nome de localidade com pelo menos um aspecto geográfico no banco de dados geográfico utilizando pelo menos um atributo comum entre a fonte de nome de localidade e os aspectos geográficos no banco de dados geográfico.

16. Índice, de acordo com a reivindicação 15, **CHARACTERIZADO** por compreender ainda um voto de facee tirado de aspectos geográficos casados em um mapa adjacente a um aspecto geográfico não casado no banco de dados geográfico para atribuir uma localidade ao aspecto geográfico não casado.

17. Índice, de acordo com a reivindicação 16, **CHARACTERIZADO** pelo fato de que um voto de facee compreende um entre um voto de maioria, um voto ponderado e um voto de comprimento linear.

18. Índice, de acordo com a reivindicação 14, **CHARACTERIZADO** pelo fato de que a associação indireta compreende para uma primeira fonte de nome de localidade que não é associada a aspectos geográficos em geral, nome de localidade de fonte cruzada casando com uma segunda fonte de nome de localidade que é associada a aspectos geográficos é utilizado de tal modo que cada nome de localidade na primeira fonte herda as associações com aspectos geográficos a partir da segunda fonte.

19. Índice, de acordo com a reivindicação 1, **CHARACTERIZADO** por compreender ainda um token principal do nome de localidade, em que o token principal é determinado por um ou mais entre tokenizar, normalizar e otimizar os nomes de localidade, bem como casar o nome de localidade com quaisquer nomes de localidade duplicatas ou similares.

20. Índice, de acordo com a reivindicação 19, **CHARACTERIZADO** pelo fato de que tokenizar compreende dividir os nomes de localidade em tokens, ou componentes.

21. Índice, de acordo com a reivindicação 19, **CHARACTERIZADO** pelo fato de que o token principal compreende o corpo principal ou componente principal apropriado para indexação.

22. Índice, de acordo com a reivindicação 20, **CHARACTERIZADO** pelo fato de que tokens além do token principal compreendem um ou mais entre um token de direção principal, um token de tipo principal, um prenome ou informação de não tipo precedendo o corpo, um prefixo, um tipo traseiro, uma direção traseira, um sufixo, um identificador numérico que especifica divisões da localidade, e um adorno ou nome de cidade facilmente reconhecível, próximo.

23. Índice, de acordo com a reivindicação 19, **CHARACTERIZADO** pelo fato de que normalização compreende uma ou mais entre expandir abreviaturas, reduzir pontuação, remover espaços incorporados e normalizar em letras maiúsculas.

24. Índice, de acordo com a reivindicação 19, **CHARACTERIZADO** pelo fato de que a otimização compreende associar o nome de localidade com aspectos geográficos contidos na localidade.

25. Índice, de acordo com a reivindicação 19, **CHARACTERIZADO** pelo fato de que casar o nome de localidade com quaisquer nomes de localidade duplicatas ou levemente variantes compreende concatenar tokens de nome de localidade e comparar tokens para o nome de localidade com os tokens para quaisquer nomes de localidade similares ou duplicatas para determinar casamentos.

26. Índice, de acordo com a reivindicação 19, **CHARACTERIZADO** pelo fato de que o casamento do nome de localidade com quaisquer nomes de localidade duplicatas ou levemente variantes compreende casar os nomes com base em sua representação fonética ou por outro meio.

27. Índice, de acordo com a reivindicação 26, **CHARACTERIZADO** pelo fato de que o casamento compreende ainda comparar aspectos geográficos a partir da etapa de otimização para o nome de localidade e quaisquer nomes de localidade duplicatas ou levemente variantes para determinar se essas localidades sobrepõem ou estão adjacentes.

28. Índice, de acordo com a reivindicação 27, **CHARACTERIZADO** pelo fato de que se todos os aspectos geográficos casam para o nome de localidade e quaisquer nomes de localidades duplicatas ou levemente variantes esses nomes de localidade representam a mesma localidade, e nomes de localidade duplicatas exceto um nome de localidade, são eliminados a partir do índice.

29. Índice, de acordo com a reivindicação 27, **CHARACTERIZADO** pelo fato de que se um ou mais porém não todos os aspectos geográficos casam para a localidade e quaisquer localidades duplicatas ou similares, esses nomes de localidade são considerados como representando a mesma localidade e são fundidos em um nome de localidade no índice.

30. Índice, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que uma união de todos os aspectos geográficos a partir de localidades que sobrepõem ou são adjacentes é associada ao nome de localidade fundido.

5 31. Índice, de acordo com a reivindicação 27, **CARACTERIZADO** por compreender ainda adornos de cidades famosas, próximas, que são criados e armazenados no índice para localidades separadas resultando se nenhum dos aspectos geográficos casar para a localidade e quaisquer localidades duplicatas ou similares.

10 32. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** por compreender ainda um ou mais de números de identificação de aspecto geográfico, números de identificação de localidade, pontos de latitude e longitude de centro de cidade de localidade, adornos de localidade, nomes completos de localidades e tamanho de localidades.

33. Índice, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o índice é criado automaticamente.

15 34. Método para indexar uma localidade, **CARACTERIZADO** por compreender as etapas de:

receber uma seleção de um ou mais aspectos geográficos a partir de um banco de dados geográfico;

determinar um conjunto de um ou mais nomes de localidade a partir de um conjunto de uma ou mais fontes de nome de localidade;

20 associar os nomes de localidade aos aspectos geográficos do banco de dados geográfico;

priorizar para cada aspecto geográfico os nomes de localidade associados em ordem de prevalência em uso comum para uma aplicação pretendida; e

25 ordenar os nomes de localidade associados a cada aspecto geográfico por prioridade.

35. Sistema que inclui funcionalidade para habilitar um usuário a acessar localidades e aspectos geográficos nas localidades, **CARACTERIZADO** por compreender:

30 um índice de banco de dados geográfico tendo pelo menos um aspecto geográfico em um banco de dados geográfico e um conjunto de um ou mais nomes de localidade associados a pelo menos um aspecto geográfico, em que um ou mais nomes de localidade são selecionados a partir de uma ou mais fontes de nome de localidade e são ordenados por prioridade com base em prevalência do nome de localidade em uso comum para uma aplicação pretendida; e

35 um programa de aplicação que utiliza o índice de banco de dados geográfico em combinação com a exibição de informações de aspecto geográfico e localidade para um usuário e com recebimento de entrada a partir de um usuário.

36. Sistema, de acordo com a reivindicação 36, **CARACTERIZADO** pelo fato de

que a exibição de informações de aspecto geográfico e localidade compreende um ou mais de exibição textual de informações de aspecto geográfico e localidade para um usuário, exibição da localização dos aspectos geográficos em um mapa para o usuário e exibição de informações de roteamento em um mapa para o usuário.

5 37. Sistema, de acordo com a reivindicação 35, **CARACTERIZADO** pelo fato de que o sistema compreende um sistema baseado na Internet.

38. Sistema, de acordo com a reivindicação 35, **CARACTERIZADO** pelo fato de que o sistema compreende um sistema de navegação no veículo.

10 39. Dispositivo portátil que inclui funcionalidade para habilitar um usuário a acessar localidades e aspectos geográficos nas localidades, **CARACTERIZADO** por compreender:

um índice de banco de dados geográfico tendo pelo menos um aspecto geográfico em um banco de dados geográfico e um conjunto de um ou mais nomes de localidade associado a pelo menos um aspecto geográfico, em que um ou mais nomes de localidade são selecionados a partir de uma ou mais fontes de nome de localidade e são ordenados por
15 prioridade com base em prevalência do nome de localidade em uso comum para uma aplicação pretendida; e

um programa de aplicações que utiliza o índice de banco de dados geográfico em combinação com a exibição de informações de aspectos geográficos e localidade para um usuário e com recebimento de entrada a partir de um usuário.

20 40. Dispositivo portátil, de acordo com a reivindicação 39, **CARACTERIZADO** pelo fato de que a exibição de informações de aspecto geográfico e localidade compreende um ou mais de exibição textual de informações de aspecto geográfico e localidade para um usuário, exibição da localização dos aspectos geográficos em um mapa para o usuário e exibição de informações de roteamento em um mapa para o usuário.

25 41. Dispositivo portátil, de acordo com a reivindicação 39, **CARACTERIZADO** pelo fato de que o dispositivo portátil compreende um assistente pessoal digital (PDA).

42. Dispositivo portátil, de acordo com a reivindicação 39, **CARACTERIZADO** pelo fato de que o dispositivo portátil compreende um sistema de navegação pessoal.

30 43. Dispositivo portátil, de acordo com a reivindicação 39, **CARACTERIZADO** pelo fato de que o dispositivo portátil compreende um telefone celular.

44. Programa de aplicações baseado em Sistemas de Informações geográficas (GIS) que inclui funcionalidade para habilitar um usuário a acessar localidades e aspectos geográficos nas localidades, **CARACTERIZADO** por compreender:

35 um índice de banco de dados geográfico tendo pelo menos um aspecto geográfico em um banco de dados geográfico e um conjunto de um ou mais nomes de localidade associado a pelo menos um aspecto geográfico, em que um ou mais nomes de localidade são selecionados a partir de uma ou mais fontes de nome de localidade e são ordenados por

prioridade com base em prevalência do nome de localidade em uso comum para uma aplicação pretendida.

45. Meio legível por máquina, incluindo operações armazenadas no mesmo que, quando processadas por um ou mais processadores, **CARACTERIZADO** pelo fato de fazer com que um sistema execute as etapas de:

5 receber uma seleção de aspectos geográficos a partir de um banco de dados geográfico;

determinar um conjunto de um ou mais nomes de localidade a partir de um conjunto de uma ou mais fontes de nome de localidade;

10 associar os nomes de localidade aos aspectos geográficos a partir do banco de dados geográfico;

priorizar para cada aspecto geográfico os nomes de localidade associados em ordem de prevalência em uso comum para uma aplicação pretendida; e

15 ordenar os nomes de localidade associados a cada aspecto geográfico por prioridade.

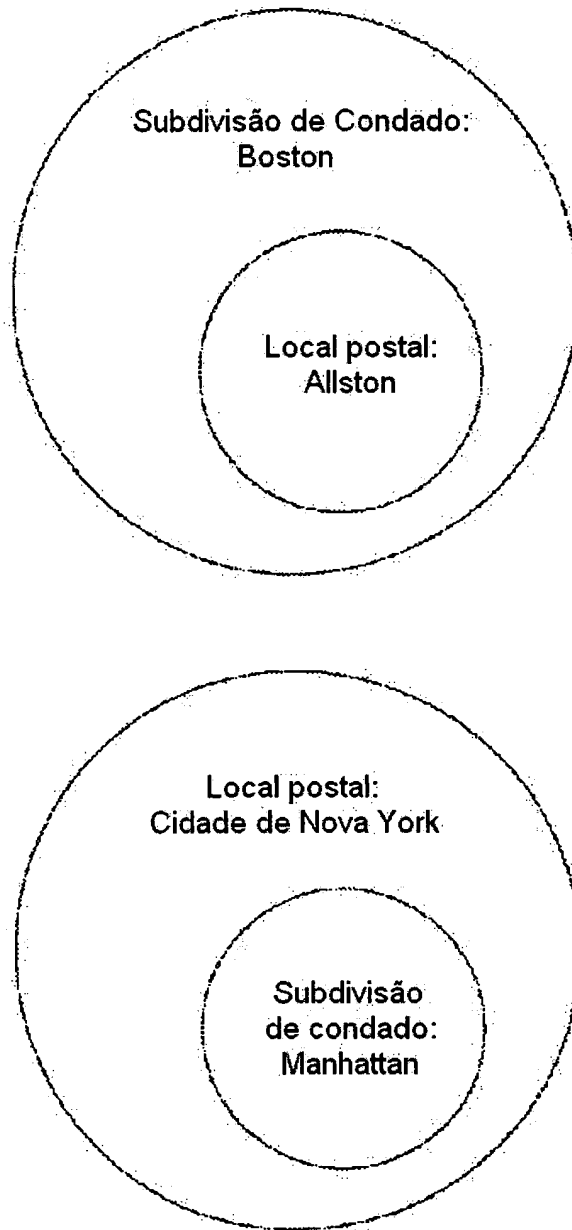


FIG. 1

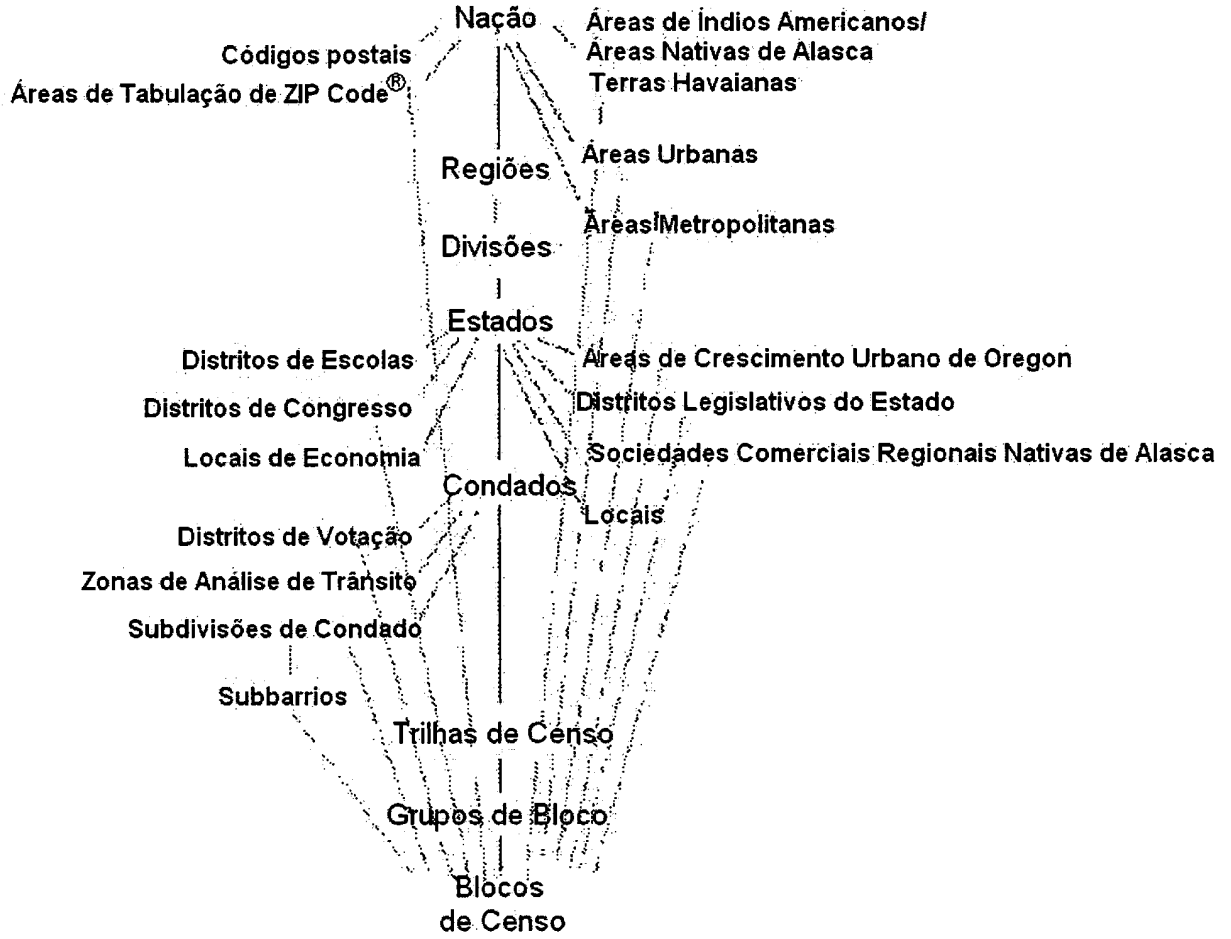


FIG. 2

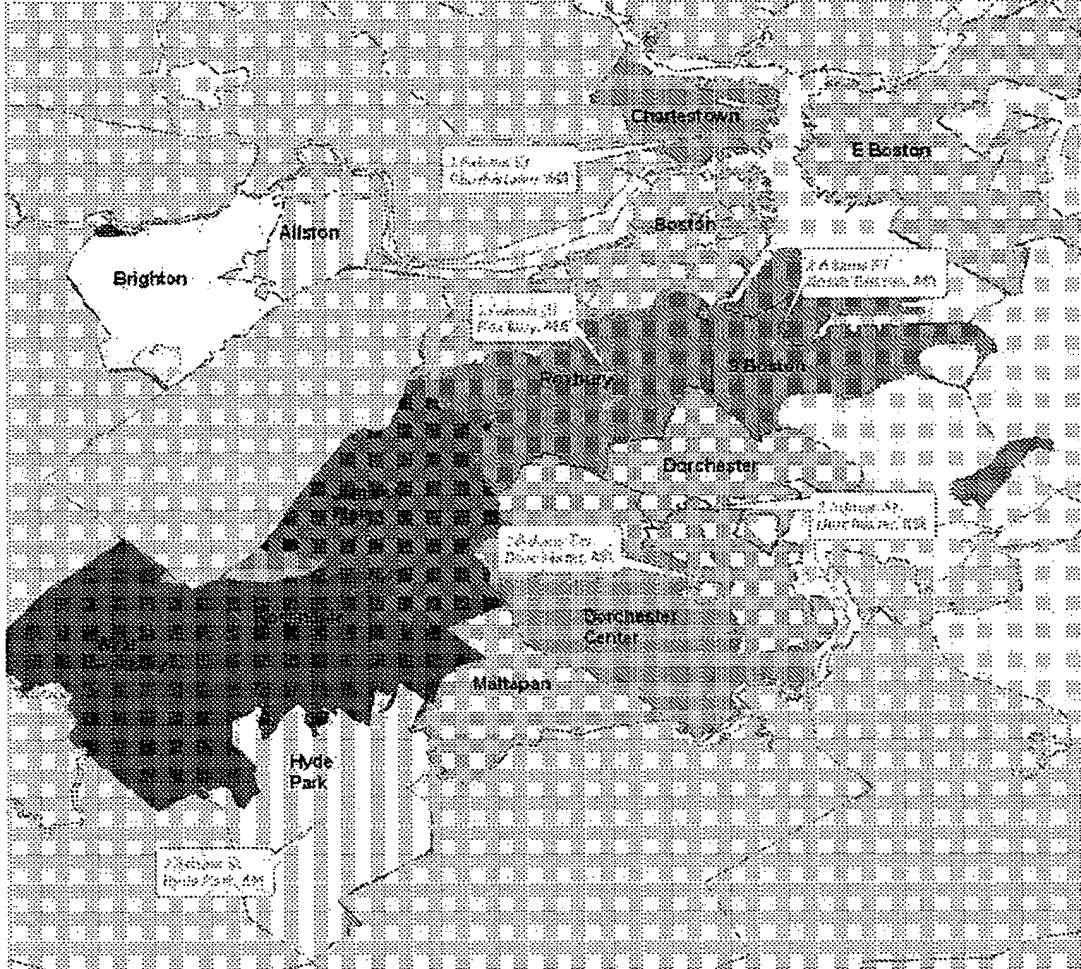


FIG. 3

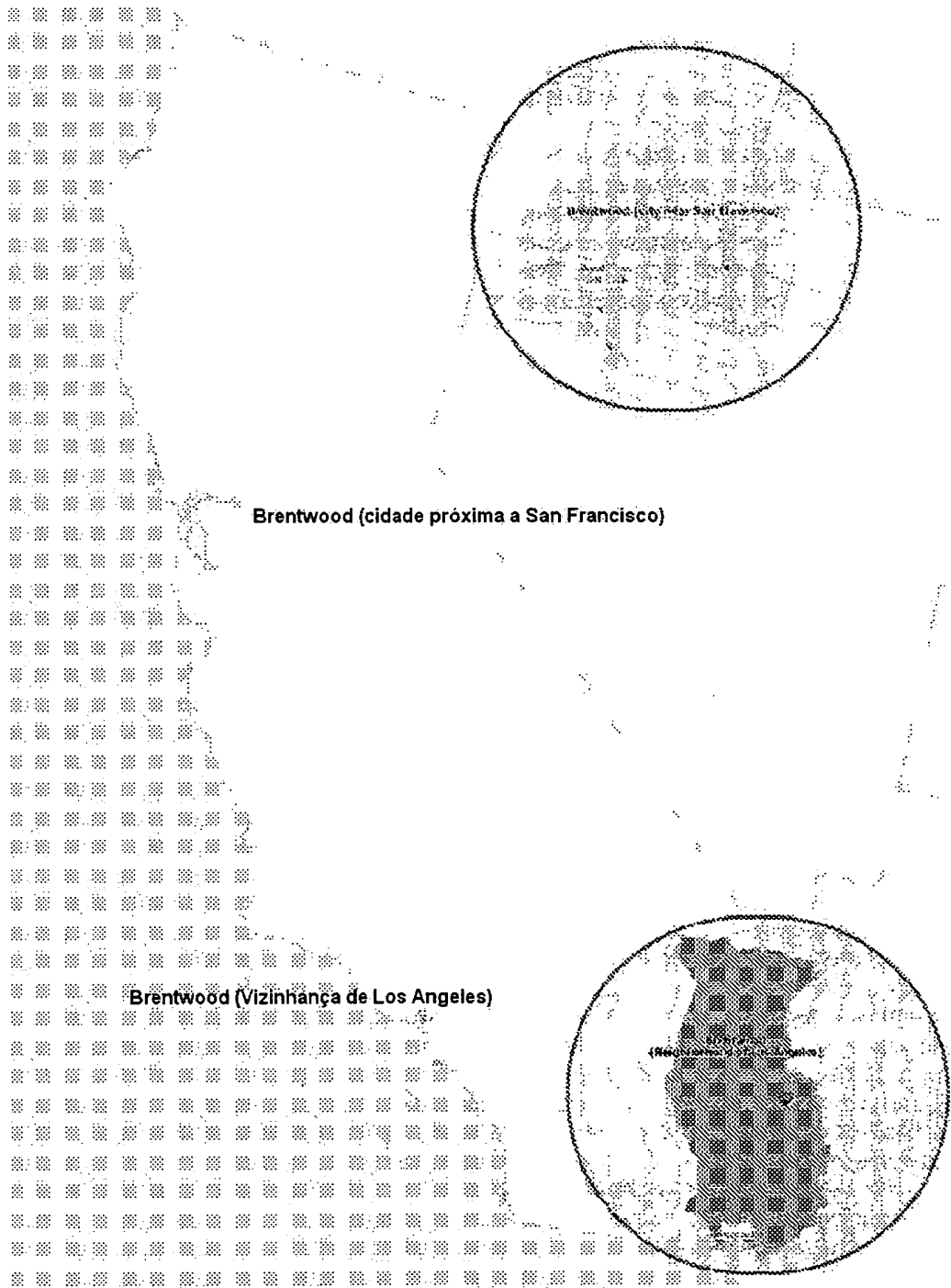


FIG. 4

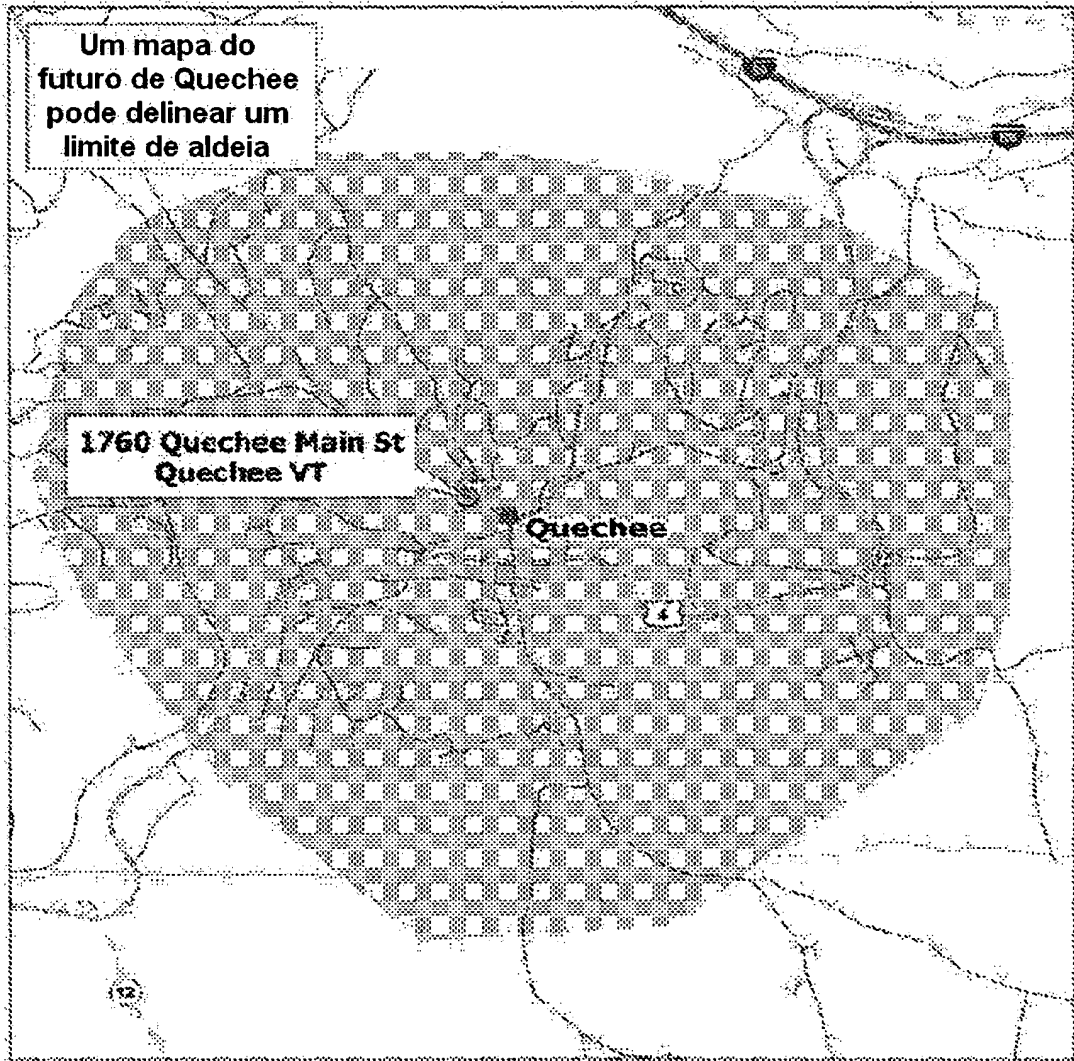


FIG. 5

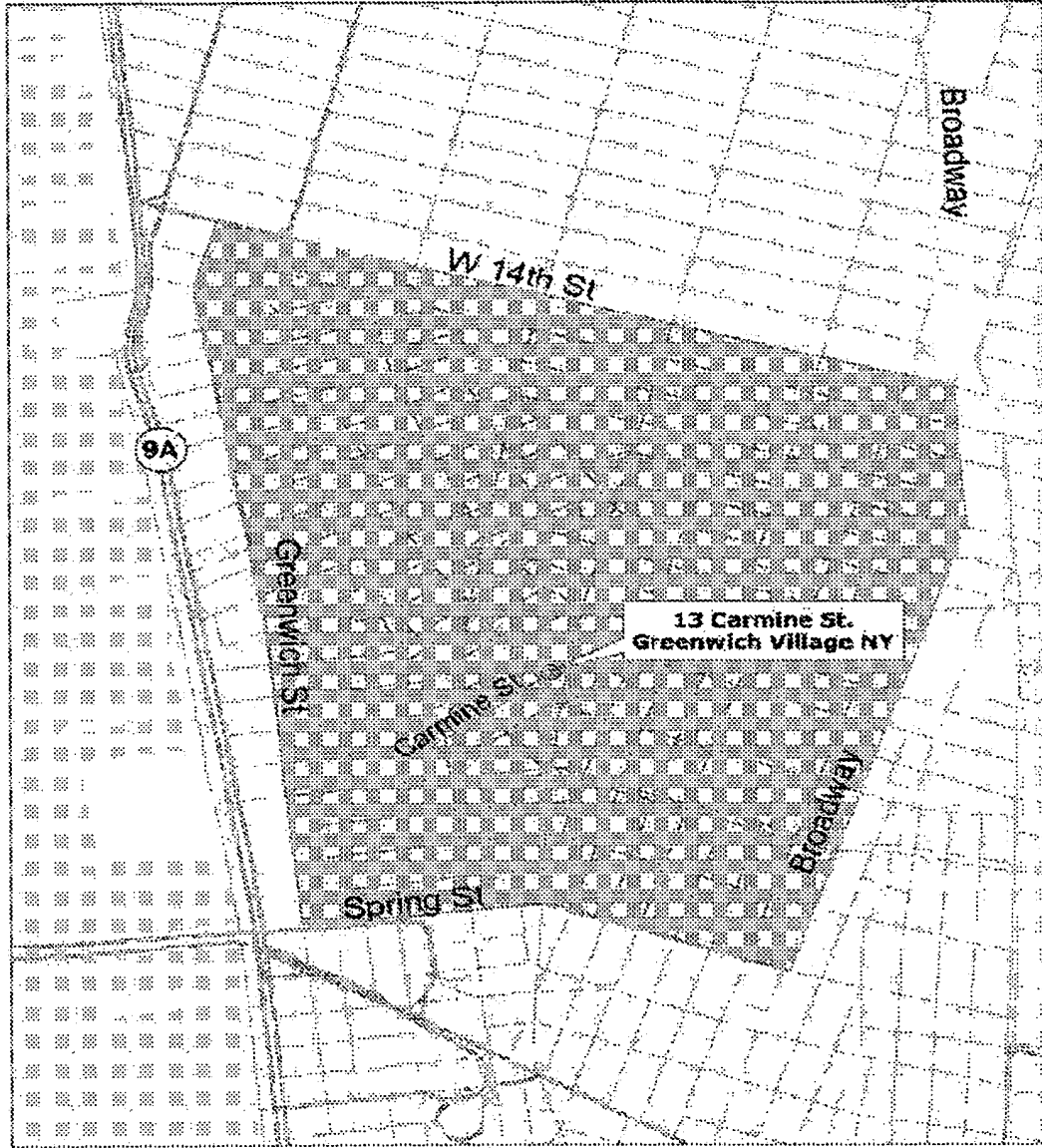


FIG. 6

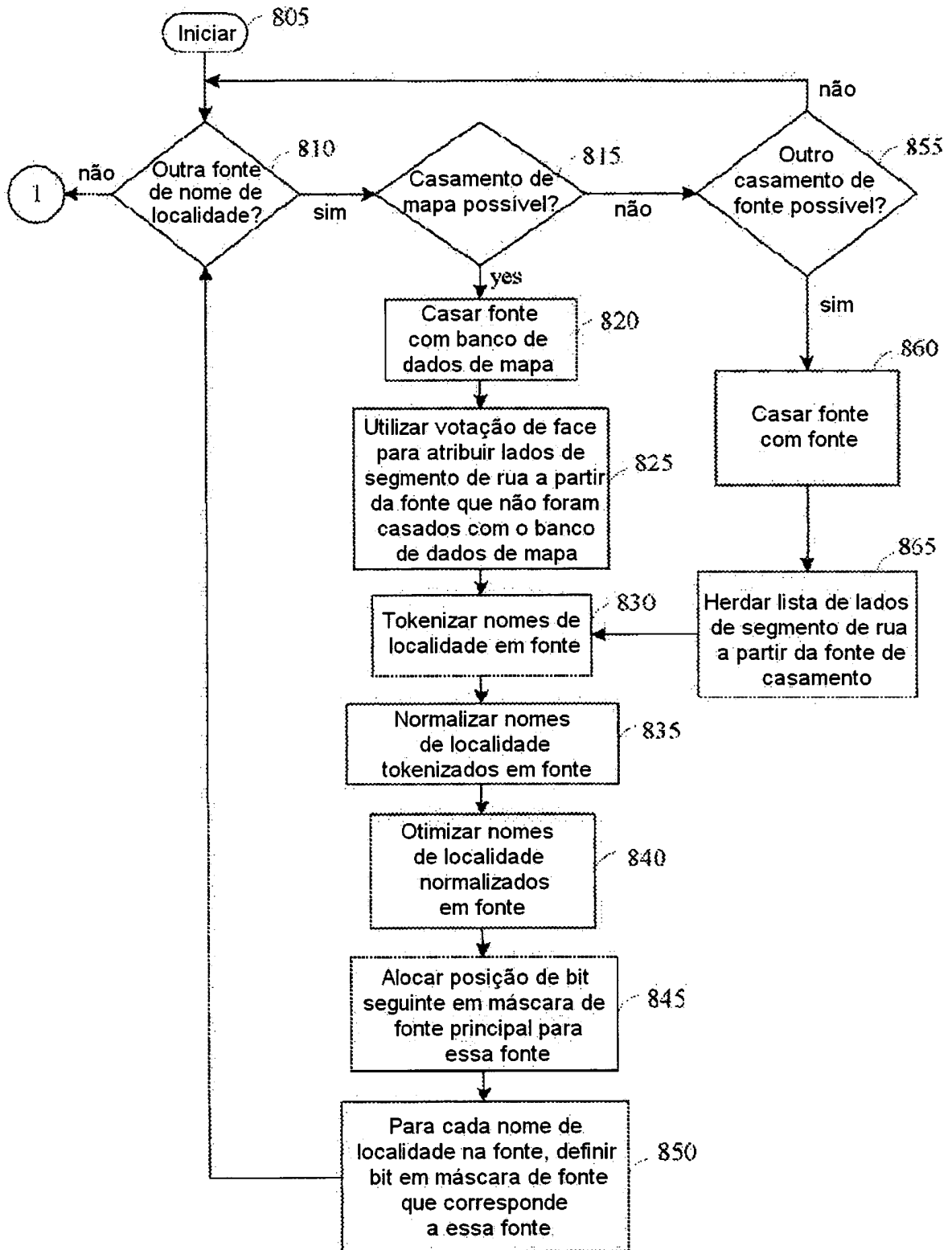


FIG. 8A

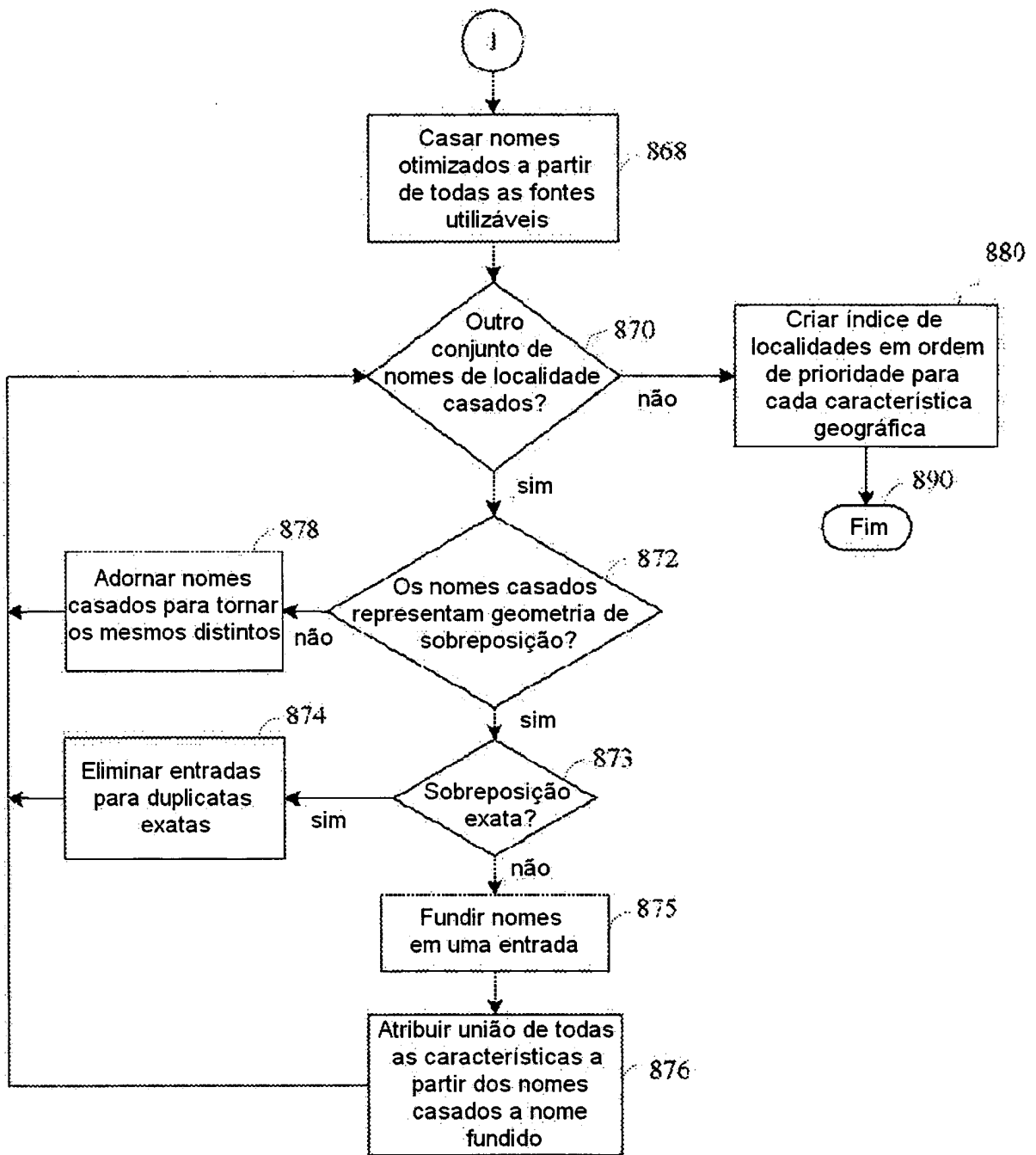


FIG. 8B

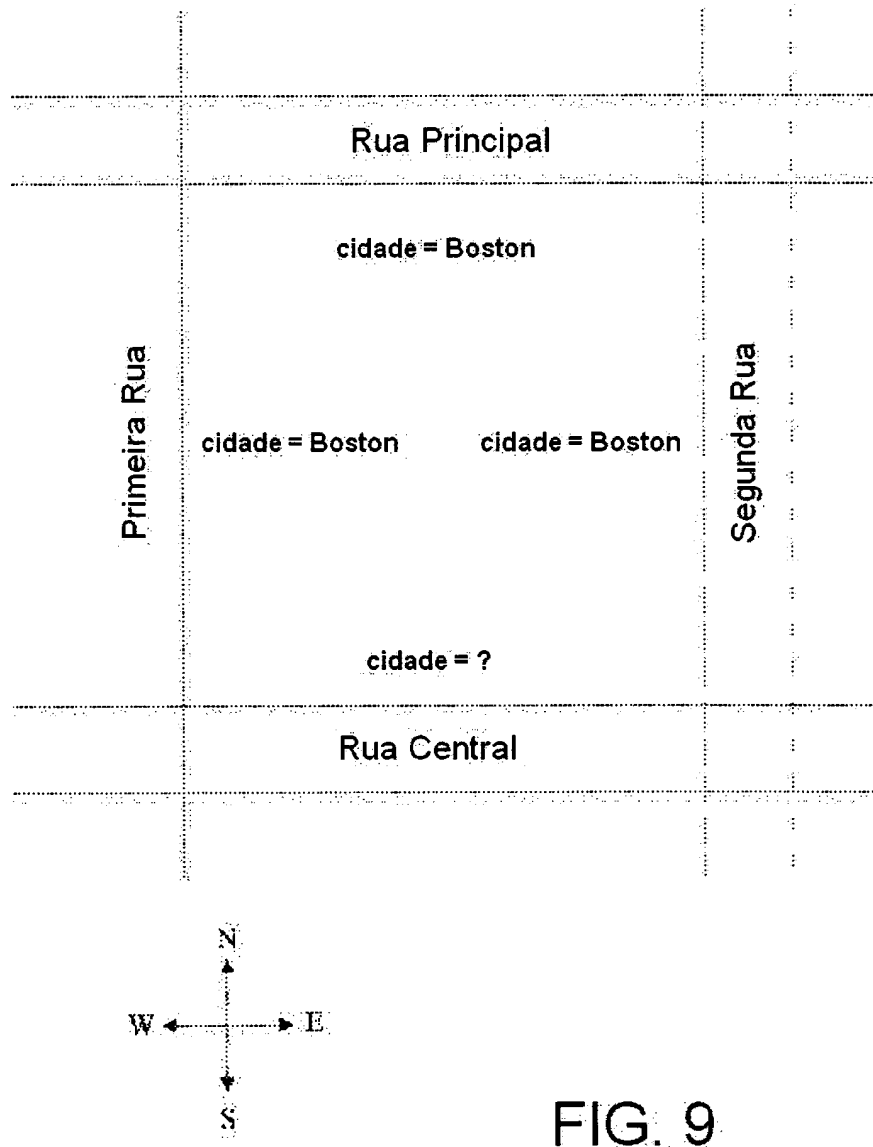


FIG. 9

Valores de Bit Decimal para Fontes de Nome de Localidade dos Estados Unidos
(exclusivo em um estado)

Valor de bit decimal	Fonte de nome de localidade	Trump
1	Nome de subdivisão de condado	
2	Nome do local de propriedade	
4	Nome do local FIPS	
8	USPS ZIP-9 nome preferido	Sim
16	USPS ZIP-5 nome preferido	
32	USPS ZIP-5 nome permissível	
64	Nome de centro de cidade de propriedade	

Valores de bit decimal para fontes de nome de Localidade no Canadá
(exclusivo em uma província)

Valor de bit decimal	Fonte de nome de localidade	Trump
1	Local Tele Atlas	
2	Local FSA	Sim
4	Centro de Cidade Tele Atlas	

FIG. 10

para cada nome de localidade A em uma dada fonte de nome de localidade

```
{
  for each name B in any other sources that matches name A
  {
    assign to A any street segment sides associated with B
    not already assigned to A.

    include any bits in bit mask B not already
    included in bit mask A.

    delete B.
  }
}
```

FIG. 11

```
para cada lado de segmento de rua S em um banco de dados de mapa
{
    find all locality names A to which S is assigned

    for each A
    {
        find the name A with the most bits set in
        its bit mask

        assign A to the next highest priority name
        in Index for this street segment side S
    }
}
```

FIG. 12

Tabela de Prioridade de Localidade de característica listando localidades por prioridade para cada característica geográfica:

Coluna	Descrição
FF_ID	Número de ID de característica geográfica para link para tabela de Encontrar Característica
NAME_ID	Número de ID de localidade para link para tabela de Nome de Localidade
PRIORITY	Prioridade dessa localidade
LOC_MASK	Máscara de Fonte de Nome de Localidade para essa localidade

IDs de localidade de mapeamento de tabela de Nome de localidade para nomes de localidade e informações associadas

Coluna	Descrição
NAME_ID	Número de ID de nome de localidade para link para tabela de Prioridade de localidade de característica
FULL_NAME	Nome completo de localidade em caso misturado
NAME_KEY	Componente significativo de nome de localidade para indexação
ADORNMENT	Pointer para outra localidade em Tabela de nome de localidade de cidade grande e facilmente reconhecível próximo à localidade
NAME_LC	Código de 3 caracteres para o idioma do nome de localidade
LOC_SIZE	Número de características geográficas associadas à localidade
COUNTRY	Abreviatura de 3 caracteres de país no qual a localidade está localizada
CENTER_ID	Link para característica de ponto de centro de cidade no banco de dados de mapa

Características geográficas de mapeamento de tabela de Encontrar características para tipos e números de ID

Coluna	Descrição
FF_ID	Número de ID de característica geográfica
FEAT_TYPE	R = característica de estrada, F = característica de linha ferry
FEAT_ID	Link para característica geográfica de banco de dados de mapa
SIDE	R = lado direito, L = lado esquerdo, B = os dois lados, nulo – não aplicável (o lado é uma margem de rua, por exemplo)

FIG. 13

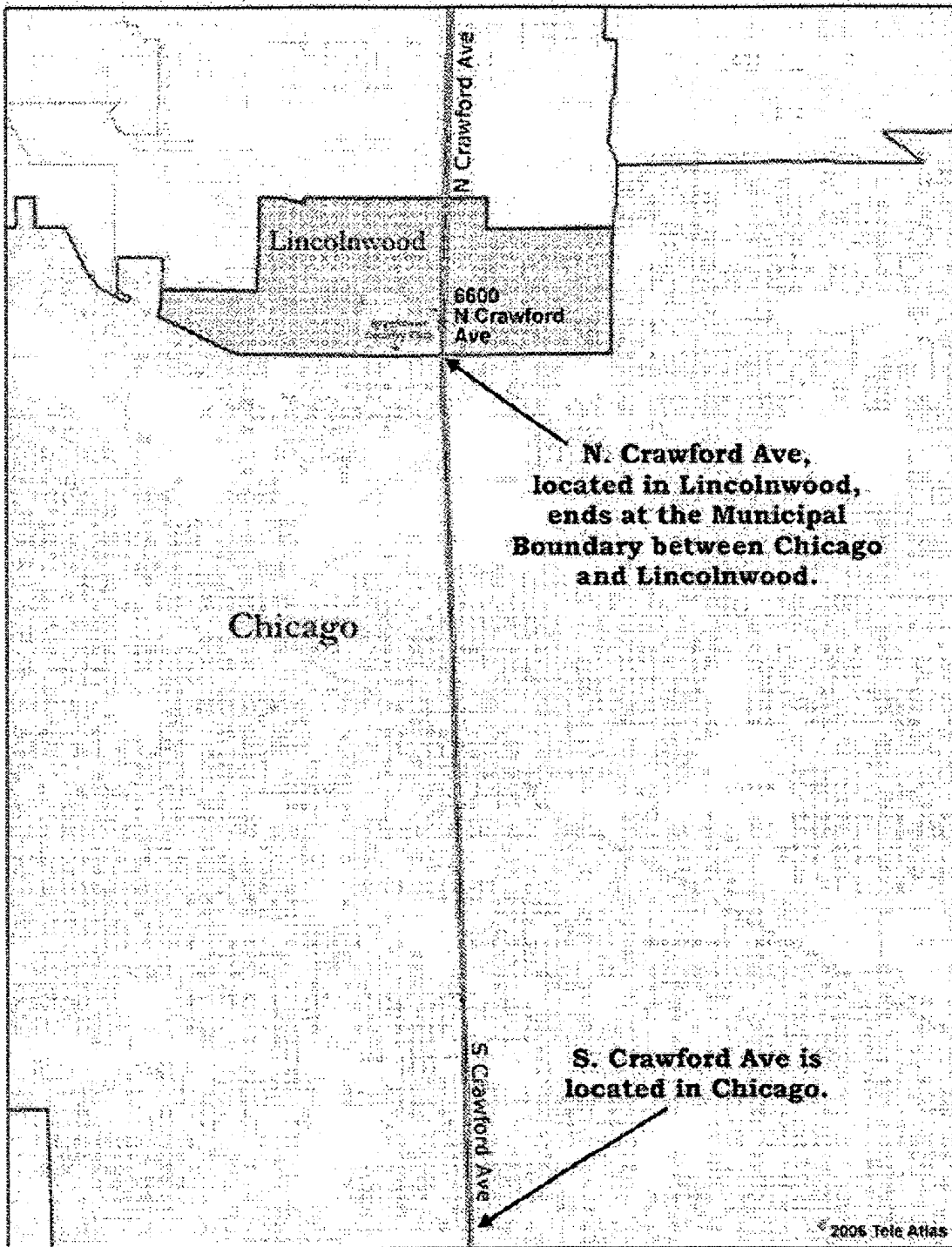


FIG. 14

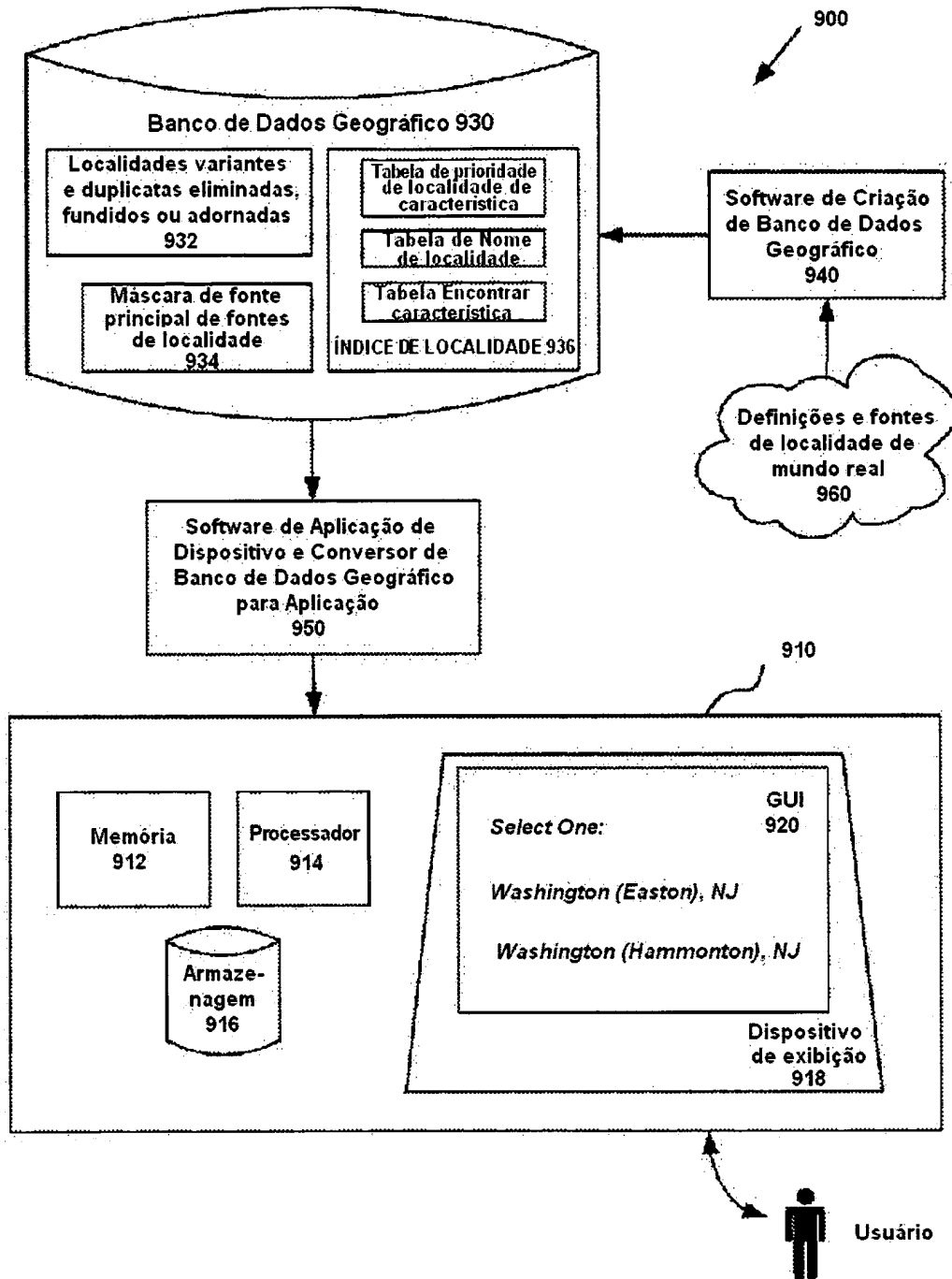


FIG. 15

RESUMO**"ÍNDICES DE LOCALIDADE E MÉTODO PARA INDEXAR LOCALIDADES"**

Índices de localidade são apresentados para uso com bancos de dados e mapas eletrônicos. Cada aspecto geográfico em um banco de dados geográfico é associado a nomes de localidade a partir de várias fontes de nome de localidade. Tokenização, normalização, otimização e casamento sensíveis ao contexto, de nomes de localidade, eliminam nomes de localidade duplicatas e variantes, enquanto preservam nomes significativamente diferentes. Uma tabela de nomes de localidade inclui a representação analisada de cada nome de localidade e outras informações associadas, e um token primário para indexação é identificado. Uma máscara de fonte principal é criada por alocar um bit para cada fonte de nome de localidade utilizada no método. Uma máscara de fonte separada é armazenada para cada aspecto geográfico associado a uma localidade, um conjunto de bits para cada fonte na qual a localidade pode ser encontrada. Nomes de localidade associados a cada aspecto geográfico são indexados em uma tabela de aspectos geográficos em ordem de prevalência para uso em uma dada aplicação.