



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0030413

(51)⁷ G01C 21/26; G09B 29/10

(13) B

(21) 1-2015-03926

(22) 14/03/2014

(86) PCT/US2014/029282 14/03/2014

(87) WO/2014/144745 18/09/2014

(30) 61/792,762 15/03/2013 US

(45) 25/12/2021 405

(43) 25/12/2015 333A

(73) THE DUN & BRADSTREET CORPORATION (US)

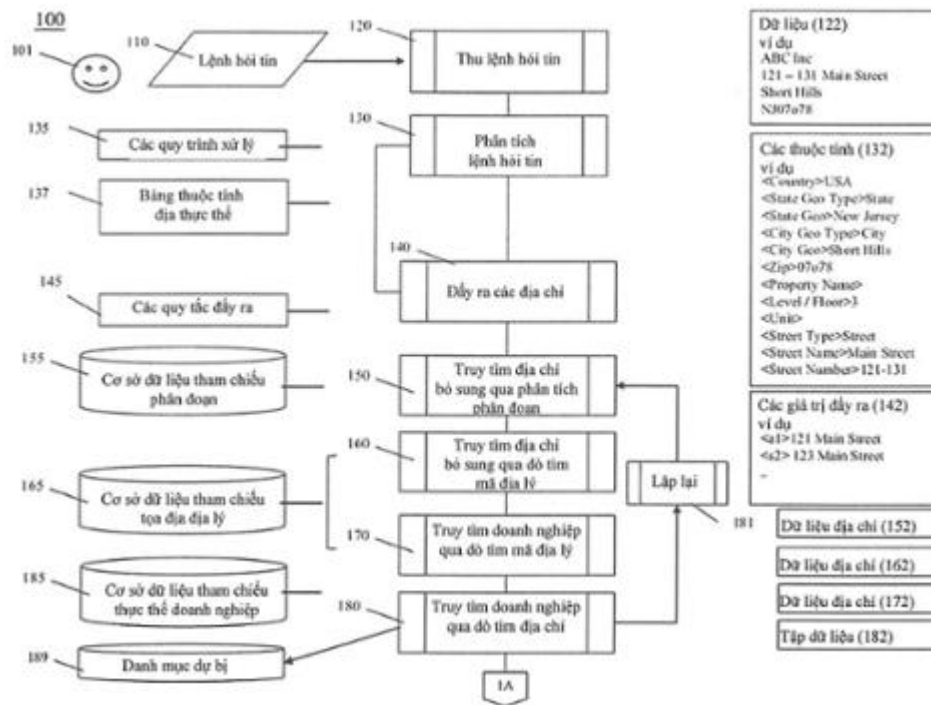
103 JFK Parkway, Short Hills, NJ 07078, USA

(72) KLEIN, Michael (US); SCRIFFIGNANO, Anthony, J. (US); MATTHEWS, Warwick (AU).

(74) Công ty TNHH T&T INVENMARK Sở hữu trí tuệ Quốc tế (T&T INVENMARK CO., LTD.)

(54) PHƯƠNG PHÁP, HỆ THỐNG VÀ THIẾT BỊ LƯU TRỮ DÙNG ĐỂ KHỬ NHẬP NHẰNG KHÔNG TẮT ĐỊNH VÀ SO KHỚP THỰC THỂ THEO ĐỊNH TÍNH Ở DỮ LIỆU ĐỊA ĐIỂM ĐỊA LÝ ĐỐI VỚI CÁC THỰC THỂ DOANH NGHIỆP

(57) Sáng chế đề cập đến phương pháp khử nhập nhằng không tắt định và so khớp thực thể theo định tính ở dữ liệu địa điểm địa lý bao gồm các bước (a) thu dữ liệu mô tả vị trí, (b) ngoại suy, từ dữ liệu nêu trên, địa chỉ có liên kết với vị trí này, (c) nhận dạng một phân đoạn của đường chính có địa chỉ này, (d) xác định một hình đa giác có chu vi bao quanh miền địa lý nằm ở gần phân đoạn này, (e) thu nhận các tọa độ địa lý của một điểm nằm trong hình đa giác, (f) nhận dạng địa chỉ ở các tọa độ địa lý này, và (g) nhận dạng thực thể liên kết với địa chỉ ở các tọa độ địa lý này. Sáng chế còn đề cập đến hệ thống thực hiện phương pháp này, và thiết bị lưu trữ có chứa các lệnh để điều khiển bộ xử lý thực hiện phương pháp này.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến kỹ thuật tạo lập các phương thức, còn được gọi là các tính năng, nhận dạng các doanh nghiệp dựa vào địa chỉ mới để xác định sự liên kết của các doanh nghiệp với một địa điểm vật lý, bằng cách sử dụng các phương pháp vượt quá các tính năng liên quan địa chỉ tiêu chuẩn. Trong trường hợp này, thuật ngữ “phương thức” được dùng để chỉ cách thức mà theo đó các địa chỉ được xác định khác nhau, ví dụ, số và tên đường phố, các tọa độ vĩ độ/kinh độ, hoặc tên tòa nhà, có thể được liên kết với cùng một địa điểm hoặc vị trí vật lý. Tính năng này có thể được sử dụng đáp lại lệnh hỏi tin, được khởi tạo bởi người dùng trực tiếp, hệ thống, ứng dụng, hoặc phương pháp bất kỳ khác, được dành để nhận dạng các thực thể doanh nghiệp bằng cách sử dụng thông tin địa chỉ gồm một trong số các trường dữ liệu, số và tên đường phố chẳng hạn, làm một thuộc tính hỏi tin. Tính năng này có thể được thi hành nhờ sử dụng các giải pháp không tất định khác nhau bao gồm giải pháp khử nhập nhằng không gian địa lý dựa vào khoảng, tên và hình đa giác làm ví dụ. Trong ngữ cảnh này, thuật ngữ “không tất định” có nghĩa là cách thức mà theo đó các địa chỉ sẽ được xác định trong lệnh hỏi tin, kể cả dùng nhiều cách thức sử dụng các thành phần dữ liệu khác nhau được gọi là các dấu hiệu thay thế linh hoạt, và giải pháp mà theo đó chúng sẽ được liên kết với các địa điểm vật lý, là không định trước và không giới hạn. Kết quả xử lý lệnh hỏi tin như vậy bao gồm thông tin phản hồi phản ánh chất lượng được suy ra, phương thức so khớp, và các thuộc tính khác để xác định sự kết nối của các doanh nghiệp với địa chỉ hỏi tin và địa điểm vật lý có liên kết.

Một ví dụ của tính năng này có thể là cho phép liên kết doanh nghiệp với một địa điểm khác với vị trí vật lý và địa chỉ của doanh nghiệp, ví dụ (a) doanh nghiệp trong trung tâm mua sắm, có thể được coi là vị trí vật lý theo chiều ngang, (b) doanh nghiệp trong tòa nhà văn phòng cao tầng, có thể được coi là vị trí vật lý theo chiều dọc, hoặc (c) doanh nghiệp được liên kết với một vùng xác định dựa trên chu vi, ví dụ, khu vực thuế quan riêng.

Tình trạng kỹ thuật của sáng chế

Các giải pháp được mô tả trong đoạn này là các giải pháp có thể được tìm kiếm, nhưng không nhất thiết là các giải pháp đã được nghĩ đến hoặc tìm kiếm trước đây. Do đó, trừ khi có quy định khác, các giải pháp này không thể là giải pháp kỹ thuật đã biết và không được thừa nhận là giải pháp kỹ thuật đã biết do đưa vào đoạn này.

Các tính năng truy nhập và tìm kiếm cơ sở dữ liệu hữu hiệu rất quan trọng để sử dụng có hiệu quả dữ liệu được duy trì trong các cơ sở dữ liệu tham chiếu. Điều cốt yếu đối với mục tiêu này là khả năng cho phép truy tìm có hiệu quả kết quả so khớp, tức là, kết quả so khớp dữ liệu tham chiếu với lệnh hỏi tin có chứa dấu hiệu địa điểm, để nhận dạng và lựa chọn các kết quả so khớp một cách hữu hiệu và có hiệu quả, và để cung cấp thông tin phản hồi có thể dùng để đưa ra quyết định về việc sử dụng các kết quả so khớp.

Cụ thể để nhận dạng một thực thể doanh nghiệp ở một địa điểm địa lý, công nghệ hiện thời xem xét một số trường dữ liệu cụ thể và hữu hạn, ví dụ “địa chỉ đường phố” vật lý đã được đăng ký, công bố hoặc liên kết theo cách khác của thực thể doanh nghiệp. Công nghệ hiện thời thường được xác nhận dựa trên các giả thiết là a) các địa chỉ thực thể doanh nghiệp, tức là, thông tin địa điểm, sẽ được biểu diễn phù hợp, và b) các hệ phương pháp truy tìm dạng biểu diễn cụ thể của một địa điểm, như một địa chỉ đường phố cho tòa nhà nhiều người thuê, sẽ cần được hạn chế ở việc truy tìm các dự bị dùng chung dạng biểu diễn địa điểm địa lý nêu trên. Ví dụ, công nghệ hiện thời này thường dựa trên các phép so sánh oristic “số đường + tên đường + thành phố”, để đưa ra ý kiến về độ chính xác dựa vào (a) sự tương tự của các ký tự giữa lệnh hỏi tin và dự bị, (b) số ký tự được so khớp hoặc thông tin tương quan cơ bản khác sau khi tính đến các sửa đổi về chính tả định trước, như nhiều cách đánh vần một từ cụ thể có sử dụng dấu gạch nối, viết chữ hoa, ngắt từ, đánh dấu chấm câu, các chữ viết tắt đã biết và các từ đồng nghĩa, hoặc (c) các giải pháp khác được dùng để so sánh các thành phần dữ liệu. Vì vậy, công nghệ hiện thời này không thể mang lại kết quả thỏa đáng khi địa điểm có thể được mô tả theo nhiều cách khác nhau, khi một vị trí có nhiều địa chỉ chẳng hạn.

Ví dụ khác, công nghệ hiện thời có thể sử dụng vĩ độ và kinh độ để liên kết các địa chỉ khác nhau với cùng một địa điểm vật lý. Điều này không thể mang lại kết quả thỏa đáng khi địa điểm trải rộng trong một vùng địa lý rộng, như trung tâm hoặc khu

phổ mua sắm, và việc sử dụng vĩ độ/kinh độ làm thông tin nhận dạng có thể liên kết sai các thực thể không liên kết với trung tâm mua sắm nhưng nằm trong cùng một vùng địa lý.

Bản chất kỹ thuật của sáng chế

Do đó, cần phải cải tiến các hệ thống và các phương pháp tìm kiếm và so khớp hiện thời, bao gồm (1) quy trình theo đó lệnh hỏi tin được xử lý để hệ thống hóa và tổng hợp dấu hiệu hỏi tin để tăng đến mức tối đa cơ hội nhận dạng các dự bị thỏa đáng, (2) quy trình theo đó dữ liệu tham chiếu như được duy trì trong cơ sở dữ liệu hoặc chuỗi các cơ sở dữ liệu được truy nhập, đánh giá và sử dụng để nhận dạng so khớp dự bị cho lệnh hỏi tin, (3) quy trình theo đó các dự bị được đánh giá, lựa chọn và ưu tiên, và (4) quy trình theo đó kết quả so khớp và dữ liệu về thực nghiệm so khớp được cấp cho người hỏi tin hoặc hệ thống hỏi tin, bao gồm thông tin chỉ báo độ tin cậy mô tả mức độ tương đối của kết quả so khớp với lệnh hỏi tin, và các thuộc tính để chỉ báo phản hồi về dữ liệu được dùng để truyền kết quả so khớp.

Để đáp ứng các nhu cầu này, sáng chế đề xuất phương pháp bao gồm các bước (a) thu dữ liệu mô tả vị trí, (b) ngoại suy, từ dữ liệu này, địa chỉ liên kết với vị trí, (c) nhận dạng một phân đoạn của con đường chính có địa chỉ này, (d) xác định một hình đa giác có chu vi bao quanh miền địa lý nằm ở gần phân đoạn này, (e) thu nhận các tọa độ địa lý của một điểm nằm trong hình đa giác, (f) nhận dạng địa chỉ tại các tọa độ địa lý này, và (g) nhận dạng thực thể có liên kết với địa chỉ ở các tọa độ địa lý này. Sáng chế còn đề xuất hệ thống để thực hiện phương pháp này, và thiết bị lưu trữ chứa các lệnh để điều khiển bộ xử lý thực hiện phương pháp này.

Sáng chế xuất phát từ giải pháp kỹ thuật đã biết trong đó mục đích của quy trình được đề xuất ở đây là nhận dạng các thực thể doanh nghiệp có địa điểm hỏi tin là một thuộc tính, nhận dạng này dựa vào các giá trị như dấu hiệu thay thế khử nhập nhằm không gian địa lý dựa vào khoảng, tên và hình đa giác và linh hoạt, trong đó dữ liệu hỏi tin gọi đến là dữ liệu mô tả thực thể địa chỉ giả định (“địa điểm”) mà một hoặc nhiều thực thể doanh nghiệp đặt tại đó.

Sáng chế đề cập đến hệ thống và phương pháp tự động để phân biệt giữa những lần so khớp của máy, tức là, so khớp dữ liệu tham chiếu với lệnh hỏi tin, mà không cần sự can thiệp của con người, nhờ đó cho phép có tính nhất quán và khả năng mở rộng, và

để cho con người tập trung vào những tình huống cần nghiên cứu sâu hoặc nghiên cứu bổ sung để thực hiện việc quản lý dữ liệu phù hợp. Thuật ngữ “khả năng mở rộng” có nghĩa là giải pháp này không bị hạn chế bởi công nghệ cụ thể đối với giải pháp kỹ thuật.

Các kỹ thuật được mô tả ở đây có các tính năng không được giải quyết theo giải pháp kỹ thuật đã biết. Cụ thể, các kỹ thuật được mô tả ở đây cho phép linh hoạt trong việc truy nhập, lưu giữ và sử dụng các phần tử dữ liệu dự báo có giá trị cao hoặc dấu hiệu dẫn xuất khác được liên kết với một địa điểm địa lý, đã được phê chuẩn trước đó và gộp chung vào các cơ sở dữ liệu của các đặc tính, dấu hiệu hoặc thẻ bài riêng của địa điểm và các địa điểm, cần được sử dụng trong quy trình so khớp. Các kỹ thuật được mô tả ở đây cho phép sự linh hoạt này được xác định khi các phần tử dữ liệu được tạo sẵn cho cơ sở dữ liệu liên quan đến dữ liệu tương đối của dữ liệu đó, và cách thức mà dữ liệu sẽ được sử dụng thông qua thực nghiệm so khớp khi nhận dạng sự so khớp và các thuộc tính đi kèm có thể được người dùng trực tiếp sử dụng để đưa ra quyết định về việc sử dụng dữ liệu và dấu hiệu dẫn xuất đó.

Mô tả vắn tắt các hình vẽ

Fig.1A và Fig.1B cùng là lưu đồ của phương pháp truy tìm các thực thể được liên kết với một địa điểm địa lý.

Fig.2 là sơ đồ khối của hệ thống sử dụng các phương pháp được mô tả ở đây.

Mô tả chi tiết sáng chế

Trong một vài đoạn dưới đây là các xác định của một số thuật ngữ được sử dụng theo sáng chế.

Địa điểm là vị trí vật lý, có thể có nhiều địa chỉ được mô tả theo nhiều cách khác nhau, và có thể có các địa chỉ liên kết với một thực thể như doanh nghiệp hoặc nhà ở, các địa chỉ mà tại đó không có tòa nhà thực hoặc doanh nghiệp hoạt động, hoặc tổ hợp của các vị trí vật lý.

Khoảng là một cách liên kết các địa chỉ dựa vào khoảng cách địa lý giữa các địa chỉ vật lý, như “123 – 131 Main Street”.

Tên là một cách nhận dạng địa điểm doanh nghiệp ngoài địa chỉ chuẩn như theo tên tòa nhà, ví dụ “1 New York Plaza” hoặc “trung tâm mua sắm ở Short Hills”.

Khử nhập nhằng không gian địa lý dựa vào hình đa giác là một cách xác định

mức độ gần theo địa lý giữa nhiều địa chỉ bằng cách sử dụng các công thức toán học và các thuật toán. Đây là khái niệm đằng sau các hệ thống truyền thông địa lý (GIS - Geographic Information System) như hệ định vị toàn cầu (GPS - Global Positioning System).

Vùng dựa trên chu vi, cũng có thể được gọi là phân đoạn hóa, là một giải pháp để liên kết các địa chỉ khác nhau với một ký hiệu nhận dạng địa điểm chung, ví dụ tổ hợp nhà hoặc khu kinh doanh với nhiều doanh nghiệp, tòa nhà, khu vực thuế quan, khu vực mở cửa, hoặc các đơn vị tương tự khác.

Phương thức nhận dạng là một cách xác định mức độ tương tự giữa hai giá trị khác nhau, như giữa địa chỉ hỏi tin dựa vào số đường và tên đường, và địa chỉ có vĩ độ và kinh độ cụ thể. Phương thức này có chính các giá trị dữ liệu cũng như siêu dữ liệu về các giá trị này. Ví dụ về siêu dữ liệu, là dữ liệu mô tả về dữ liệu, trong trường hợp này có thể chỉ báo rằng tên và con số trong lệnh hỏi tin là địa chỉ đường phố.

Phương thức so khớp được dùng để chỉ các quy trình khác nhau có thể được thi hành để sử dụng dữ liệu từ lệnh hỏi tin để nhận dạng dữ liệu từ cơ sở dữ liệu hoặc nguồn khác, sử dụng các quy trình quen thuộc với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này; phương thức này có thể bao gồm “so khớp”, “tìm kiếm”, “dò tìm”, hoặc các quy trình khác để liên kết lệnh hỏi tin với cơ sở dữ liệu hoặc nguồn khác.

Dấu hiệu là thuật ngữ chung được dùng để chỉ dữ liệu. Dấu hiệu thay thế là dữ liệu khác có thể được sử dụng để nhận dạng đặc tính, như “tỉnh thành” có thể được dùng làm giá trị thay thế cho “tiểu bang”. Dấu hiệu dẫn xuất là thông tin có thể được tính hoặc suy ra từ dữ liệu khác, ví dụ, vĩ độ và kinh độ có thể được suy ra từ địa chỉ đường, thành phố và tiểu bang.

Lĩnh hoạt có nghĩa là loại của dấu hiệu dùng để nhận dạng các đặc tính của thực thể hoặc lệnh hỏi tin đối với các thực thể có thể là khác nhau với các lệnh hỏi tin hoặc các thực thể khác nhau.

Phản hồi là thông tin về kết quả của thực nghiệm so khớp, như chất lượng suy ra và kết quả của thực nghiệm đó. Thông tin phản hồi này có thể được dùng để cho phép người dùng trực tiếp đưa ra quyết định về việc sử dụng kết quả của thực nghiệm.

Các cơ sở dữ liệu tham chiếu là các kho chứa dữ liệu đáp ứng được các chuẩn dựa trên chất lượng liên quan đến dữ liệu, ví dụ toàn vẹn dữ liệu, đầy đủ, chính xác và

kip thời. Thuật ngữ “cơ sở dữ liệu tham chiếu” có thể liên hệ đến một hoặc nhiều kho chứa dữ liệu như vậy.

Chiến lược là giải pháp để tiếp nhận lệnh hỏi tin, sử dụng dữ liệu hỏi tin để nhận dạng các dự bị từ cơ sở dữ liệu tham chiếu, và cung cấp từ cơ sở dữ liệu tham chiếu này các dự bị có thể chấp nhận được với người dùng trực tiếp.

Khái niệm “doanh nghiệp” như được sử dụng ở đây là ví dụ về một loại thực thể ở một địa điểm. Tuy nhiên, sáng chế không chỉ giới hạn ở các loại thực thể mà người dùng trực tiếp có thể quan tâm; các ví dụ khác bao gồm khu nhà ở, khu vực thuế quan, và khu vực mở cửa không phát triển.

Hệ thống được mô tả ở đây (a) nhận lệnh hỏi tin để bắt đầu tìm kiếm thực thể có liên kết với một địa điểm địa lý cụ thể, (b) sử dụng lệnh hỏi tin để xác định chiến lược thực hiện cuộc tìm kiếm, (c) nhận dạng các thực thể dự bị có liên kết với địa điểm địa lý, và (d) xuất ra các thực thể này và thông tin liên quan đến quy trình bao gồm mức độ và chất lượng của các dự bị vì chúng liên quan đến lệnh hỏi tin.

Các ví dụ được mô tả ở đây giả định lệnh hỏi tin có chứa cả dữ liệu tên và địa chỉ thực thể. Đây là một ví dụ; tính năng này có thể được sử dụng để xử lý lệnh hỏi tin chỉ có một địa chỉ để nhận dạng các địa chỉ khác có liên kết với địa chỉ hỏi tin, hoặc các thực thể ở các địa chỉ hỏi tin và ở các địa chỉ khác có liên kết với địa chỉ hỏi tin. Do vậy, lệnh hỏi tin có thể có tên và địa chỉ để tìm một doanh nghiệp cụ thể, hoặc chỉ có địa chỉ để tìm nhiều doanh nghiệp ở địa chỉ đó.

Fig.1A và Fig.1B cùng là lưu đồ của phương pháp 100 để truy tìm các thực thể có liên kết với một địa điểm địa lý. Fig.1A được ghép nối với Fig.1B bằng bộ ghép nối có nhãn là “1A”.

Tóm lại, phương pháp 100 nhận lệnh hỏi tin 110 từ người dùng 101, và thi hành các quy trình 120, 130 và 140 để trích ra, tạo lập, nhận dạng và định tính các địa điểm và các thực thể có liên quan sẽ được đệ trình để truy tìm dự bị bằng cách sử dụng các quy trình 150, 160, 170 và 180. Tập hợp các quy trình này có thể truy tìm nhiều bản ghi so khớp, các bản ghi so khớp này sẽ được khử lặp và lọc trong quy trình 181, và tiếp đó được tính điểm định tính, cả khách quan và chủ quan, trong các quy trình 193 và 194. Đầu ra, tức là, kết quả 196, được tạo ra và gửi trả về người dùng 101 ở quy trình 195.

Quy trình 100 sử dụng các quy tắc xử lý 135, bảng thuộc tính địa thực thể 137, các quy tắc đẩy ra 145, cơ sở dữ liệu tham chiếu phân đoạn 155, cơ sở dữ liệu tham chiếu tọa độ địa lý 165, cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185, danh mục dự bị 189, các quy tắc tính điểm 190, các quy tắc xếp hạng 191 và các quy tắc quyết định 192.

Bảng thuộc tính địa thực thể 137 là bảng các thuộc tính có thể nhận biết, tức là, các trường dữ liệu, có thể liên kết với dữ liệu để có thể nhận dạng địa điểm địa lý.

Các quy tắc đẩy ra 145 có logic được sử dụng để xác định các địa chỉ khả thi cụ thể dựa vào giá trị đầu vào, ví dụ, ở Mỹ, các địa chỉ ở một bên đường là các số toàn lẻ hoặc toàn chẵn.

Cơ sở dữ liệu tham chiếu phân đoạn 155 chứa thông tin liên quan đến một địa điểm cụ thể, ví dụ, để nhận dạng phân liên kề của đường hoặc đường chính khác giữa các điểm dừng quan trọng như các đường giao nhau.

Cơ sở dữ liệu tham chiếu tọa độ địa lý 165 chứa thông tin không gian địa lý được biểu diễn dưới dạng các tọa độ địa lý, như vĩ độ và kinh độ, của các địa chỉ cụ thể.

Cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 là danh mục các thực thể doanh nghiệp được liên kết với các địa chỉ cụ thể.

Danh mục dự bị 189 là danh mục các doanh nghiệp từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 mà thực sự nằm tại các địa chỉ và các địa điểm được nhận dạng từ cơ sở dữ liệu tham chiếu tọa độ địa lý 165.

Các quy tắc tính điểm 190 là các quy tắc dùng để gán điểm định tính cho mỗi mục trong danh mục dự bị 189 dựa vào sự tương tự của nó so với lệnh hỏi tin 110.

Các quy tắc xếp hạng 191 là các quy tắc để xác định cách thức các mục trong danh mục dự bị 189 được xếp hạng dựa vào điểm số được suy ra khi sử dụng các quy tắc tính điểm 190.

Các quy tắc quyết định 192 là các quy tắc để xác định những mục nào trong danh mục dự bị 189 cần được trình diễn cho người dùng dựa vào bậc xếp hạng như được suy ra từ các quy tắc xếp hạng 191.

Lệnh hỏi tin 110 là yêu cầu bắt đầu tìm kiếm thông tin về một địa điểm cụ thể. Việc tìm kiếm sẽ dựa vào dấu hiệu có trong lệnh hỏi tin 110. Lệnh hỏi tin 110 chứa các phần tử dữ liệu, các phần tử dữ liệu này chứa thông tin cụ thể về địa điểm dưới dạng các

trường dữ liệu, các trường dữ liệu này là toàn bộ, hoặc một tập hợp con, các trường dữ liệu có thể được yêu cầu từ người dùng trực tiếp qua hệ thống hoặc ứng dụng giao diện với quy trình 100, và còn có thể chứa các dấu hiệu bổ sung và gần như không giới hạn về địa điểm. Lệnh hỏi tin 110 có thể được cấp cho phương pháp 100 từ người dùng 101, tức là, người dùng trực tiếp hoặc quy trình tự động, ví dụ như được suy ra từ các tệp tin được đệ trình bằng cách sử dụng các tính năng máy xử lý theo bố, hoặc từ lệnh hỏi tin riêng lẻ được xử lý bằng cách sử dụng các màn hình nhập dữ liệu trực tuyến. Lệnh hỏi tin 110 chứa dấu hiệu hỏi tin có thể bao gồm một trong số các phần tử dữ liệu, ví dụ, dấu hiệu hỏi tin có thể được biểu diễn dưới dạng một trường ghép nối của số đường, tên đường và thành phố, hoặc các trường dữ liệu tách biệt của (1) số đường, (2) tên đường, và (3) thành phố.

Theo một ví dụ, lệnh hỏi tin 110 có thể có các trường dữ liệu tách riêng với thông tin trong Bảng 1.

Bảng 1

Các dấu hiệu hỏi tin	ABC Inc 121 – 131 Main Street Short Hills NJ07078
----------------------	---

Phương pháp 100 bắt đầu với quy trình 120.

Quy trình 120 nhận lệnh hỏi tin 110 chứa dữ liệu, tức là, dấu hiệu hỏi tin, mô tả vị trí. Quy trình 120 tạo cấu trúc và diễn dịch dấu hiệu hỏi tin thành một khuôn chung, tức là, dữ liệu 122.

Theo một ví dụ, dữ liệu 122 có thể chứa thông tin trong bảng 2.

Bảng 2

Trường	Các dấu hiệu hỏi tin
Tên	ABC Inc
Địa chỉ đường phố	121 – 131 Main Street
Thành phố	Short Hills
Tiểu bang	NJ
Mã bưu điện	07078

Từ quy trình 120, phương pháp 100 chuyển sang quy trình 130.

Quy trình 130 phân tích dữ liệu 122 để nhận dạng các trường dữ liệu cụ thể có liên kết với các thuộc tính trong bảng thuộc tính địa thực thể 137 để gia tăng cơ hội nhận dạng so khớp từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 bằng cách sử dụng các trường dữ liệu này riêng lẻ hoặc phối hợp. Theo một ví dụ, quy trình này có thể tách số đường (121 – 131) ra khỏi tên đường (Main Street). Với việc này, quy trình 130 trích ra, từ dữ liệu 122, các thuộc tính liên quan đến việc tìm kiếm các dự bị tương tự với lệnh hỏi tin, nhờ vậy tạo ra các giá trị cho các thuộc tính 132.

Như nêu trên, bảng thuộc tính địa thực thể 137 là bảng gồm các thuộc tính có thể nhận biết, tức là, các trường dữ liệu, có thể được liên kết với dữ liệu để có thể nhận dạng một địa điểm địa lý. Bảng thuộc tính địa thực thể 137 chứa dữ liệu để gán trọng số dự báo (K) cho mỗi thuộc tính (x) nhằm hỗ trợ cho việc tính toán $K(x)$, trong đó $K(x)$ là trọng số dự báo của phần tử x, cũng như thông tin khác liên quan đến mỗi thuộc tính.

Theo một ví dụ, bảng thuộc tính địa thực thể 137 có thể chứa thông tin trong bảng 3.

Bảng 3

Trường (x)	Trọng số dự báo (K)
Tên	1
Tên đường	.7
Số đường	.5
Tên tòa nhà	.8
Số dãy phòng	.3
Số tầng	.4
Thành phố	.6
Tiểu bang	.9
Mã bưu điện	.5
Quốc gia	1

Theo một ví dụ, các thuộc tính 132 có thể chứa thông tin trong bảng 4.

Bảng 4

Trường	Giá trị
Tên doanh nghiệp	ABC Inc
Tên đường	Main Street
Số đường	121 – 131
Tên tòa nhà	<null>
Số dãy	<null>
Số tầng	<null>
Thành phố	Short Hills
Tiểu bang	NJ
Mã bưu điện	07078
Quốc gia	USA

Quy trình 130 sử dụng các quy tắc xử lý 135 để phân tích các thuộc tính 132, chúng còn có thể cung cấp đầu vào để đề ra các quy tắc doanh nghiệp và siêu dữ liệu mới sẽ được lưu trữ trong các quy tắc xử lý 135, trong đó việc thi hành quy trình 130 sau này có thể sử dụng các quy tắc xử lý mới.

Ví dụ, quy tắc siêu dữ liệu có thể chỉ ra rằng đối với một địa chỉ ở Mỹ, trường 2 ký tự trong các chữ viết hoa là chữ viết tắt tên tiểu bang, hoặc trường chữ cái tiếp sau chữ số là tên đường.

Các quy tắc xử lý 135 bao gồm các quy tắc thực thể và siêu dữ liệu địa chỉ tự động và có thể lặp lại dựa vào các quy trình tiêu chuẩn hóa và chuẩn hóa có logic khử nhập nhằng ngữ nghĩa và con số để diễn dịch dấu hiệu hỏi tin cho mỗi thành phần của địa chỉ. Một ví dụ về các quy tắc xử lý 135 để tiêu chuẩn hóa và chuẩn hóa có thể chứa thông tin trong bảng 5.

Bảng 5

Quy tắc xử lý	Tác động
Tiểu bang	NJ = New Jersey
Địa chỉ	Đối với các thực thể ở Mỹ, số đường đứng trước tên đường
Các mối quan hệ	Tên thành phố có thể dư thừa trong nhiều tiểu bang

thành phần địa chỉ	Mã bưu điện là duy nhất
--------------------	-------------------------

Một ví dụ về quy tắc mới đối với các quy tắc xử lý 135 có thể chứa thông tin trong bảng 6.

Bảng 6

Quy tắc xử lý mới	Tác động
Các doanh nghiệp ở số 125 Main Street Short Hills NJ	DEF Corp là hoạt động kinh doanh mới của ABC Inc
Các doanh nghiệp ở số 125 Main Street Short Hills NJ	ABC Inc đã giảm các hoạt động kinh doanh của mình
Các doanh nghiệp ở số 125 Main Street Short Hills NJ	ABC Inc đã cho thuê không gian chưa sử dụng cho một doanh nghiệp khác

Các quy tắc xử lý 135 còn có thể chứa thông tin để làm sạch, phân tích cú pháp và tiêu chuẩn hóa tất cả các thành phần của các thuộc tính 132.

Làm sạch bao gồm việc loại bỏ các giá trị ngoại lai ra khỏi dữ liệu 122 như dấu chấm hoặc các dạng khác của các ký tự không có giá trị, như dấu chấm trong các chữ viết tắt tên đường hoặc các dấu phẩy và các ký tự đẩy dòng hoặc ký tự xuống dòng được dùng làm dấu tách.

Một ví dụ về quy tắc làm sạch cho các quy tắc xử lý 135 có thể chứa thông tin trong bảng 7.

Bảng 7

Giá trị thuộc tính 132	Giá trị đã được làm sạch
07o78	07078 (thay thế 'o' bằng '0')

Phân tích cú pháp bao gồm việc phân chia dữ liệu 122 để gia tăng cơ hội nhận dạng các thực thể địa chỉ so khớp và/hoặc có liên kết cho lệnh hỏi tin 110. Việc này có thể bao gồm việc phân tích cú pháp dấu hiệu hỏi tin thành các phần tử dữ liệu tách riêng, ví dụ, chia tách địa chỉ thành các phần tử có ý nghĩa cấu thành nó.

Một ví dụ về quy tắc phân tích cú pháp cho các quy tắc xử lý 135 có thể chứa thông tin trong bảng 8, thể hiện giá trị ghép nối của mã bưu điện và tiểu bang được chia

tách thành hai giá trị tách riêng như thế nào.

Bảng 8

Giá trị thuộc tính 132	Giá trị đã được phân tích cú pháp
NJ07o78	NJ 07078

Tiêu chuẩn hóa bao gồm việc liên kết các giá trị thay thế với dữ liệu 122 để gia tăng cơ hội nhận dạng sự so khớp.

Một ví dụ của quy tắc tiêu chuẩn hóa cho các quy tắc xử lý 135 có thể chứa thông tin trong bảng 9.

Bảng 9

Giá trị thuộc tính 132	Giá trị tiêu chuẩn hóa
NJ	New Jersey

Phương pháp 100 chuyển từ quy trình 130 đến quy trình 140.

Quy trình 140 sử dụng các quy tắc đẩy ra 145 để xử lý các thuộc tính 132, dưới dạng các tổ hợp, các tập hợp và/hoặc các phần tử riêng lẻ, để suy ra nhiều thực thể địa chỉ tiềm năng từ các thuộc tính hỏi tin gốc 132, nhờ vậy đưa ra các giá trị đẩy ra 142. Như vậy, quy trình 140 ngoại suy, từ các thuộc tính 132, một hoặc nhiều địa chỉ liên kết với vị trí được nêu trong lệnh hỏi tin 110. Ngoại suy có nghĩa là hình thành ý kiến hoặc đưa ra ước tính về một điều gì đó từ các sự kiện đã biết. Thuật ngữ “ngoại suy” và thuật ngữ “đẩy ra” được sử dụng thay thế lẫn nhau ở đây. Ví dụ về địa chỉ ngoại suy là việc tạo ra địa chỉ “123 Main Street” từ “121-131 Main Street”. Theo ví dụ này, việc ngoại suy còn nhiều hơn việc trích dữ liệu từ lệnh hỏi tin 110 và, ví dụ, còn có thể bao gồm việc ngoại suy “140 Main street”, nhận được từ việc phân tích các thuộc tính 132 ở quy trình 140, mặc dù “140 Main Street” không được nêu rõ trong lệnh hỏi tin 110 (ví dụ nêu rõ là “131 Main Street”) cũng không có hàm ý nhắc đến (ví dụ hàm ý vốn có là “123 Main Street”).

Quy trình 140 sử dụng, ví dụ, việc tính dựa trên thuật toán oristic đối với các địa chỉ vốn có trong địa chỉ đường phố được biểu diễn dưới dạng khoảng, và cung cấp hướng dẫn về những địa chỉ có thể là hợp lệ đối với trường hợp, ví dụ, chỉ các số lẻ trên cùng một bên của các đường phố nhất định. Ví dụ, vì đã biết rằng ở Mỹ, các địa chỉ ở

một bên đường là toàn số lẻ hoặc toàn số chẵn, và hai chữ số được phân cách bởi dấu gạch ngang chỉ báo điểm đầu và điểm cuối, việc tính toán dựa trên thuật toán oristic xác định là có các địa chỉ tiềm năng khác giữa địa chỉ đầu và địa chỉ cuối, và các địa chỉ đó là toàn số lẻ hoặc toàn số chẵn dựa vào địa chỉ đầu và địa chỉ cuối là số lẻ hay số chẵn. Các giá trị đẩy ra 142 là dấu hiệu hỏi tin bổ sung có thể được sử dụng ở quy trình 170 để nhận dạng các dự bị từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185.

Một ví dụ về quy tắc đẩy ra đại diện cho các quy tắc đẩy ra 145 có thể chứa thông tin trong bảng 10.

Bảng 10

Giá trị thuộc tính 132	Quy tắc đẩy ra	Các giá trị tiềm năng
121 – 131 Main Street	Việc lập địa chỉ tiêu chuẩn ở Mỹ có sử dụng các số lẻ hoặc chẵn ở cùng một bên đường	121 Main Street 123 Main Street 125 Main Street 127 Main Street 129 Main Street 131 Main Street

Phương pháp 100 chuyển từ quy trình 140 đến quy trình 150.

Quy trình 150 thu các giá trị đẩy ra 142, và nhận dạng phân đoạn của đường chính có địa chỉ này. Tức là, quy trình 150, có tham chiếu đến cơ sở dữ liệu tham chiếu phân đoạn 155, tính phân đoạn địa chỉ (tức là, phần liên kề của đường hoặc đường chính khác giữa các điểm dừng quan trọng như các đường giao nhau, chính nó thường sẽ là điểm cuối của các phân đoạn khác) có liên kết với mỗi địa chỉ có trong các giá trị đẩy ra 142, và sau đó truy tìm từ cơ sở dữ liệu tham chiếu phân đoạn 155 địa chỉ hoặc các địa chỉ có liên kết với mỗi phân đoạn trong số các phân đoạn này, và xuất chúng ra dưới dạng dữ liệu địa chỉ 152 có cả dữ liệu địa chỉ đường phố và dữ liệu vị trí không gian địa lý như kinh độ và vĩ độ (dữ liệu không gian địa lý này còn được gọi là “các mã địa lý” hoặc “dữ liệu mã địa lý”). Cơ sở dữ liệu tham chiếu phân đoạn 155 chứa dữ liệu liên kết các địa chỉ theo vùng địa lý, còn được gọi là “các chu vi” và dữ liệu dựa trên chu vi thu thập được biết dưới dạng “dữ liệu tham chiếu theo chu vi”. Với việc này, quy trình 150 xác định một hình đa giác có chu vi bao quanh miền địa lý nằm ở gần phân đoạn.

Một ví dụ về dữ liệu tham chiếu phân đoạn trong cơ sở dữ liệu tham chiếu phân đoạn 155 có thể chứa thông tin trong bảng 11.

Bảng 11

ID phân đoạn	ID vùng địa lý	Tên đường	Từ #	Từ vĩ độ /kinh độ	Đến #	Đến vĩ độ/kinh độ
53AA9BCBB7BF	FF2E	Route 10	120	D17DD2F7	160	F66C7FA6
60C883DB7C6C	3B57	US 1	12	60ED05CC	1200	551D0DC2
42A021CA8914	5CAE	Smith Avenue	1	77EABD92	43	97DC4E06
42365F645E94	5A94	Main Street	10	1337CDF4	460	A5302FD7
CCEC8F47EDCE	75CE	NULL	0	CE958552	0	F3E18C19

Bảng 11 thể hiện một ví dụ về các địa chỉ hiện có dọc theo các phân đoạn cụ thể (được thể hiện trong cột thứ nhất) dưới dạng tập hợp của các thuộc tính bao gồm khoảng số đường (các cột 4 và 6), tên đường (cột 3) và các tọa độ Đề các của các điểm cuối phân đoạn (được thể hiện dưới dạng mã hóa trong các cột 5 và 7).

Một ví dụ về dữ liệu tham chiếu chu vi trong cơ sở dữ liệu tham chiếu phân đoạn 155 có thể chứa thông tin trong bảng 12.

Bảng 12

ID hình dạng	Loại	Mã băm tọa độ
0A4E98F3	Điểm	37CE17DF782339B2
3D139E4D	Hình đa giác	D956C0BCCD601FEA
B12AE4CE	Chuỗi đoạn thẳng	E82354DBD99B91DF
DBF8F7DD	Hình đa giác	DA0C535E76DC4F38

Bảng 12 thể hiện một ví dụ biểu diễn dữ liệu tham chiếu chu vi, với ký hiệu nhận dạng riêng cho mỗi chu vi (cột 1), phân mô tả loại chu vi (cột 2) và phiên bản mã hóa của các tọa độ Đề các của chu vi (cột 3).

Xem xét lại một số thao tác được thực hiện bởi các quy trình 140 và 150 để xác định hình đa giác. Quy trình 140, trong quá trình ngoại suy, thu nhận thuộc tính thứ nhất có liên kết với địa chỉ (ví dụ, 123 Cecil Avenue, là địa chỉ ở một đầu của phân đoạn). Quy trình 150, khi nhận dạng phân đoạn, thu nhận thuộc tính thứ hai có liên kết với phân đoạn (ví dụ, các địa chỉ trên phân đoạn có các số tầng có liên kết với chúng), và

tiếp đó, xác định hình đa giác, xác định kích cỡ và hình dạng của hình đa giác dựa vào các thuộc tính thứ nhất và thứ hai. Ví dụ về việc xác định hình đa giác có thể là để cho quy trình 150 sử dụng thuộc tính thứ nhất và thuộc tính thứ hai để tạo dựng một cạnh của hình đa giác chạy theo toàn bộ chiều dài của phân đoạn.

Như nêu trên, tòa nhà văn phòng cao tầng có thể được xem là vị trí theo chiều dọc. Trong trường hợp này, quy trình 150, thay vì chỉ xác định hình đa giác, có thể xác định một khối đa diện dựa vào các thuộc tính thứ nhất và thứ hai, trong đó hình đa giác là một mặt cắt ngang của khối đa diện. Ví dụ về việc xác định khối đa diện có thể là để cho quy trình 150 sử dụng thuộc tính thứ nhất và thuộc tính thứ hai để tạo dựng khối đa diện có chiều cao được tính dựa vào giá trị lớn nhất của các số tầng có liên kết với phân đoạn này.

Phương pháp 100 chuyển từ quy trình 150 đến quy trình 160.

Quy trình 160 thu dữ liệu địa chỉ 152, và nhận các tọa độ địa lý của một điểm nằm trong hình đa giác, và nhận dạng địa chỉ ở các tọa độ địa lý này. Tức là, quy trình 160, có tham chiếu cơ sở dữ liệu tham chiếu tọa độ địa lý 165, truy tìm các tọa độ địa lý, (ví dụ, vĩ độ và kinh độ), cho mỗi địa chỉ trong số các địa chỉ này, và tiếp đó truy tìm từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 địa chỉ hoặc các địa chỉ có liên kết với mỗi tọa độ này, nhờ vậy tạo ra dữ liệu địa chỉ 162. Theo cách này, quy trình 160 nhận dạng các giá trị hiện có khác cho giá trị đẩy ra 142.

Ví dụ về các địa chỉ khác trong dữ liệu địa chỉ 162 được nhận dạng bởi quy trình 160 từ dữ liệu địa chỉ 152 và cơ sở dữ liệu tham chiếu tọa độ địa lý 165 có thể chứa thông tin trong bảng 13.

Bảng 13

Địa chỉ đẩy ra (từ các giá trị đẩy ra 142)	Giá trị vĩ độ/kinh độ	Các giá trị địa chỉ khác (đến Dữ liệu địa chỉ 162]
121 Main Street	-67.534,61.968	525A Elm Ave 525B Elm Ave
123 Main Street	-44.537,-52.324	
125 Main Street	50.828,60.955	
127 Main Street	28.656,-63.131	1 ABC Plaza
129 Main Street	34.928,-19.079	

131 Main Street	-67.534,61.968	18 Corner Street 20 Corner Street
-----------------	----------------	--------------------------------------

Phương pháp 100 chuyển từ quy trình 160 đến quy trình 170.

Quy trình 170 thu dữ liệu địa chỉ 162 dưới dạng đầu vào, và nhận dạng thực thể được liên kết với địa chỉ ở các tọa độ địa lý. Cụ thể hơn, quy trình 170 thu tập hợp mã địa lý (dữ liệu liên quan đến các vị trí không gian địa lý như vĩ độ và kinh độ) là các thuộc tính của các địa chỉ trong dữ liệu địa chỉ 162, và thực hiện việc truy tìm các doanh nghiệp có liên kết với các địa chỉ đó nhờ sử dụng các phương thức khác nhau của các quy trình so khớp từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185, nhờ vậy tạo ra dữ liệu địa chỉ 172. Cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 có thể có hoặc không có doanh nghiệp liên kết với mỗi địa chỉ nhận được từ quy trình 160.

Đầu ra của quy trình 170 là thuộc tính của dữ liệu thực thể doanh nghiệp đến dữ liệu địa chỉ 162 để tạo lập dữ liệu địa chỉ 172.

Ví dụ về tập hợp con dữ liệu địa chỉ được truy tìm từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 sử dụng các địa chỉ thu được từ quy trình 160 và được bổ sung vào dữ liệu địa chỉ 172 có thể chứa thông tin trong bảng 14. “HQ (Headquarter)” là chữ viết tắt của “trụ sở chính”.

Bảng 14

Mã địa lý từ quy trình 160	Địa chỉ tìm được từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185	Tên doanh nghiệp tìm được từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 (bổ sung vào dữ liệu địa chỉ 172)
87.284,-25.211	121 Main Street	ABC Automobile Sales
-66.789,-43.111	525A Elm Street	Joe’s Pizza
54.39,5.426	525B Elm Street	<không có doanh nghiệp>
68.649,54.966	123 Main Street	<không có doanh nghiệp>
70.807,-43.812	125 Main Street	ABC Corporate HQ
-67.399,-23.047	127 Main Street	<không có doanh nghiệp>
-81.73,49.476	1 ABC Plaza	ABC Finance Division
16.153,-67.527	129 Main Street	<không có doanh nghiệp>
21.82,5.124	131 Main Street	ABC Automobile Repair
13.322,-71.755	18 Corner Street	Mary’s Takeout Kitchen
86.351,4.892	20 Corner Street	<không có doanh nghiệp>

Phương pháp 100 chuyển từ quy trình 170 đến quy trình 180.

Quy trình 180 thu dữ liệu địa chỉ 172, và sử dụng các địa chỉ trong dữ liệu địa chỉ 172 làm đầu vào, sử dụng các hệ phương pháp khác nhau (như các phương pháp dò tìm dựa trên văn bản) để phân biệt các dự bị (tức là, các thực thể tương tự như các thực thể có liên kết với các địa chỉ trong dữ liệu địa chỉ 172) trong cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185, nhờ vậy tạo ra tập dữ liệu 182

Bảng 15 thể hiện một ví dụ của tập dữ liệu 182 được tạo bởi quy trình 180 từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185.

Bảng 15

Địa chỉ từ quy trình 170 (dữ liệu địa chỉ 172)	Các biến thể được tạo bởi quy trình 180	Các tên doanh nghiệp tìm được từ cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 (được bổ sung vào tập dữ liệu 182)
121 Main Street, Short Hills	12* Main, 07078 1** Main, Short Hills nnn MN, XRTHLS	ABC Automobile Sales Joe's Pizza ABC Corporate HQ ABC Automobile Repair
525A Elm Street, Short Hills	52* Elm, 07078 5** Elm, Short Hills Nnnn ELM, XRTHLS	Joe's Pizza
18 Corner Street, Short Hills	1* Corner, 07078 1* Cnr, 07078 nn KRNR, XRTHLS	Mary's Takeout Kitchen

Quy trình 181 là quy trình đệ quy để tạo nấc các quy trình từ 150 đến 180 theo kiểu đệ quy cho đến khi đáp ứng được một hoặc nhiều tiêu chuẩn trong tập hợp các tiêu chuẩn đã xác định, ví dụ, việc lặp truy tìm giảm xuống dưới các mức định tính và định lượng nhất định. Kết quả là danh mục dự bị 189.

Bảng 16 cung cấp một ví dụ của các chuẩn định lượng và định tính có thể được sử dụng trong quy trình 181.

Bảng 16

Loại tiêu chuẩn	Xác định tiêu chuẩn
Tiêu chuẩn định lượng	Tối đa là 50 bản ghi dự bị Thời gian đáp tối đa là 0,1 giây Tối đa là 25 bản ghi
Tiêu chuẩn định tính	Dự bị phải là các trung tâm chính (tức là, văn phòng chi nhánh không được chấp nhận)

Phương pháp 100 chuyển từ quy trình 180 đến quy trình 193, được thể hiện trên Fig.1B. Danh mục dự bị 189 được thể hiện trên cả Fig.1A và Fig.1B.

Quy trình 193 nhận danh mục dự bị 189, và sử dụng các quy tắc tính điểm 190 để tạo lập danh mục dự bị được phân biệt theo định tính, nhờ vậy tạo ra danh mục 193A. Các ví dụ về quy tắc tính điểm 190 có thể bao gồm các quy tắc để phân biệt bằng cách sử dụng điểm số dựa vào, ngoài những điều khác, (a) bản chất của các phương thức truy

tìm được sử dụng ở các quy trình từ 150 đến 180 để gộp mục nhập địa chỉ hoặc thực thể doanh nghiệp cụ thể vào trong danh mục dự bị 189 (tức là, phương pháp hoặc tổ hợp các phương pháp đã được dùng cho việc phân biệt các dự bị riêng lẻ trong danh mục dự bị 189) và (b) mức độ ngoại suy (ví dụ, một địa chỉ cụ thể được truy tìm bằng hoạt động trực tiếp của phương thức truy tìm theo lệnh hỏi tin 110 hoặc theo địa chỉ đã tự truy tìm được trong lần lặp trước của quy trình đệ quy 181) hoặc trừu tượng hóa (ví dụ, khoảng cách tính theo mét có địa chỉ tìm được từ tập hợp tọa độ danh định có liên kết với lệnh hỏi tin 110) được sử dụng trong các phương thức truy tìm.

Quy trình 193 sử dụng giải pháp gán trọng số tương đối dựa vào bảng thuộc tính địa thực thể 137 và các quy tắc tính điểm 190, để tạo ra tập hợp các điểm số phụ có trọng số cho mỗi mục nhập trong danh mục dự bị 189, các điểm số phụ này được sử dụng cho việc xếp hạng theo định tính danh mục dự bị 189. Việc gán trọng số này xem xét nhiều đặc tính liên quan đến siêu dữ liệu và các giá trị dữ liệu thực nên việc gán trọng số cho một dự bị có thể xem xét các chuẩn như dãy lặp đã tạo ra dự bị, phương thức hoặc các phương thức truy tìm dùng cho việc phân biệt dự bị (và có thể có ma trận để gán điểm số cho các tổ hợp của các phương thức truy tìm, số lần lặp, v.v.), và các đánh giá định tính khác về sự tương tự giữa mỗi mục nhập trong danh mục dự bị 189 và lệnh hỏi tin nguồn 110 (ví dụ, số đo của khoảng cách hiệu chỉnh giữa tên đường của lệnh hỏi tin 110 và mục nhập trong danh mục dự bị 189). Ví dụ, dự bị C_1 tìm được bằng cách sử dụng cả quy trình phân tích phân đoạn hóa (quy trình 150) và quy trình truy tìm dựa vào tên (các quy trình 170 và 180) có thể nhận điểm số 100 cho điểm số phụ S_1 , còn dự bị C_2 tìm được bằng cách sử dụng quy trình truy tìm dựa vào tọa độ địa lý (quy trình 160) chỉ có thể nhận điểm số 90 cho điểm số phụ S_1 .

Bảng 17 là ví dụ về các điểm số phụ có trọng số được sử dụng theo quy tắc tính điểm 190 và quy trình 193.

Bảng 17

Dự bị	Điểm số giao điểm dò tìm địa lý	Điểm số mật độ	Điểm số chứng thực dò tìm địa lý	Điểm số khởi đầu/đẩy ra
C001	59	18	34	57
C002	12	52	42	76
C003	77	78	2	15
C004	86	44	76	79
C005	91	75	71	96

Với bảng trên đây, cột 1 thể hiện dự bị trong danh mục dự bị 189 và các cột còn lại thể hiện điểm số dựa trên sự tương tác của các phương thức truy tìm trong việc phân biệt dự bị (cột 2), sự đánh giá về mật độ xây dựng của địa chỉ dự bị (cột 3), mức độ chứng thực địa chỉ theo nhiều hệ phương pháp định vị không gian địa lý (cột 4), và điểm số thể hiện mức độ ngoại suy hoặc trừu tượng hóa gắn với các hệ phương pháp truy tìm cụ thể được sử dụng để phân biệt dự bị (cột 5).

Tập hợp các điểm số phụ được tạo ra trong quy trình 193 (và được nêu làm ví dụ trong Bảng 17) cho mỗi dự bị sẽ được sắp xếp và tính điểm bằng cách sử dụng các quy tắc xếp hạng 191.

Bảng 18 là ví dụ về kết quả cuối cùng của việc thi hành quy tắc xếp hạng 191.

Bảng 18

Dự bị	Xếp hạng có trọng số
C005	1
C004	2
C002	3
C001	4
C003	5

Phương pháp 100 chuyển từ quy trình 193 đến quy trình 194.

Quy trình 194 nhận danh mục dự bị có tính điểm, tức là, danh mục 193A, từ quy trình 193, và lọc danh mục 193A bằng cách loại bỏ các dự bị có xếp hạng, điểm số,

điểm số phụ hoặc tổ hợp của chúng không đáp ứng các chuẩn bao hàm và/hoặc loại trừ được xác định trong các quy tắc quyết định 192, nhờ vậy tạo ra danh mục 194A.

Bảng 19 là ví dụ về trọng số tương đối của các hệ số tính điểm được dùng trong quy trình 194.

Bảng 19

Hệ số/điểm số	Trọng số
Điểm số giao điểm dò tìm địa lý	0,5
Điểm số mật độ	0,25
Điểm số chứng thực dò tìm địa lý	0,6
Điểm số khởi đầu/đẩy ra	0,33

Bảng 20 là ví dụ về các quy tắc bao hàm/loại trừ trong quy trình 194.

Bảng 20

Quy tắc	Tác động
Mã địa lý không được tạo ra	Loại trừ
Điểm số mật độ < 20 và Điểm số khởi đầu/đẩy ra < 50	Loại trừ

Bảng 21 là ví dụ về các kết quả của quy trình 194, tức là danh mục 194A.

Bảng 21

Dự bị	Xếp hạng có trọng số	Kết quả của các quy tắc
C001	4	Đạt
C002	3	Không đạt
C003	5	Không đạt
C004	2	Đạt
C005	1	Đạt

Phương pháp 100 chuyển từ quy trình 194 đến quy trình 195.

Quy trình 195 nhận danh mục 194A, tức là, danh mục dự bị định tính $C_1...C_T$, trong đó C_T là dự bị có chất lượng thấp nhất được coi là chấp nhận được để đưa vào kết quả sẽ được gửi trả về theo phương pháp 100.

Bảng 22 là ví dụ về danh mục dự bị 189 nhận được, với các dự bị C002 và C003 được loại trừ vì chúng không đạt được các quy tắc, mặc dù chúng đều được tính điểm ở mức đáp ứng được chuẩn chấp nhận định tính.

Bảng 22

Dự bị	Xếp hạng có trọng số	Tên doanh nghiệp
C001	4	ABC Automobile Sales
C004	2	Joe's Pizza
C005	1	ABC Corporate HQ

Quy trình 195 hợp dịch và gửi trả về kết quả 196, kết quả này bao gồm danh mục dự bị 197 và thông tin phản hồi 198. Thông tin phản hồi 198 là thông tin về mức độ tương tự giữa lệnh hỏi tin 110 và mỗi dự bị trong danh mục dự bị 197.

Thông tin phản hồi 198 được gộp trong kết quả 196 để chỉ báo chất lượng của mỗi dự bị trong danh mục dự bị 197, ví dụ, mức độ tin cậy để dự bị tốt nhất C_n là thực thể doanh nghiệp nằm tại (các) vị trí địa chỉ được quy cho lệnh hỏi tin 110. Thông tin phản hồi 198 còn có thể bao gồm mức độ tương tự tương đối như được biểu diễn bằng sự tương quan tương đối giữa mỗi trường trong dữ liệu 122 và mỗi thành phần của mỗi dự bị trong danh mục dự bị 197, cũng như nguồn của dữ liệu cụ thể đã được sử dụng để điền vào cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185 với các thành phần của mỗi dự bị trong danh mục dự bị 197.

Bảng 23 là ví dụ về thông tin phản hồi 198.

Bảng 23

Tên doanh nghiệp	Địa chỉ	Điểm số tương tự	Phản hồi về tương quan	Phản hồi về địa lý
ABC Automobile Sales	121 Main Street, Short Hills NJ 07078	88	ABZFA	2AC
Joe's Pizza	525A Elm Street, Short Hills NJ 07078	67	BBZZZ	6HA
ABC Corporate HQ	125 Main Street, Short Hills NJ 07078	97	AAABB	3AA

Bảng 23 thể hiện thực thể doanh nghiệp (cột 1), địa chỉ liên kết với thực thể đó (cột 2), điểm số biểu diễn mức độ tương tự của địa chỉ đó với dự bị trong danh mục dự bị 197 (cột 3), chuỗi ký tự với ý nghĩa nhúng trong biểu diễn bản chất của sự tương tự (còn được gọi là sự tương quan) của địa chỉ (cột 4) và chuỗi có ý nghĩa nhúng trong mô tả các thuộc tính của địa chỉ (cột 5).

Do vậy, tóm lại, phương pháp 100 bao gồm các bước 1) nhận lệnh hỏi tin để bắt đầu tìm kiếm các thực thể doanh nghiệp, mỗi thực thể này được liên kết với một địa điểm địa lý cụ thể, 2) làm sạch, phân tích cú pháp và tiêu chuẩn hóa lệnh hỏi tin, 3) đẩy ra các địa chỉ bổ sung dựa vào địa chỉ hỏi tin, 4) xác định các phương pháp tối ưu để tìm kiếm các cơ sở dữ liệu tham chiếu dựa vào một hoặc nhiều dấu hiệu hỏi tin đã được làm sạch, phân tích cú pháp và tiêu chuẩn hóa, 5) truy tìm các dự bị và chọn các thực thể trong cơ sở dữ liệu tham chiếu so khớp với lệnh hỏi tin, và 6) gửi trả về các dự bị từ cơ sở dữ liệu tham chiếu dựa vào mức độ tương tự với lệnh hỏi tin hoặc dữ liệu suy ra được từ lệnh hỏi tin, cùng với thông tin phản hồi về mức độ tương tự.

Quy trình 100 bao gồm các bước để 1) thu dữ liệu đầu vào gồm chuỗi tìm kiếm có nhiều phần tử, 2) chuyển đổi tập hợp con các phần tử này thành tập hợp các thuật ngữ, 3) truy tìm dữ liệu tham chiếu lưu trữ dựa vào các thuật ngữ để nhận dạng các dự bị có khả năng nhất trong việc so khớp với dữ liệu đầu vào, 4) suy luận dự đoán về khả năng nhận dạng tập hợp dự bị so khớp bằng cách sử dụng các dấu hiệu linh hoạt dựa vào lệnh hỏi tin của người dùng trực tiếp chứa cả dữ liệu được dự kiến là một phần của lệnh hỏi tin, dữ liệu thay thế có thể được cấp bởi người dùng trực tiếp, và dữ liệu được suy ra từ dữ liệu hỏi tin, 5) chọn một hoặc nhiều dự bị so khớp dựa vào dự đoán suy luận được,

và 6) cung cấp các kết quả so khớp và thông tin phản hồi để cho phép người dùng trực tiếp quyết định sẽ sử dụng các kết quả này như thế nào.

Phương pháp 100 có chức năng để nhận dạng các thực thể doanh nghiệp được liên kết với các địa điểm địa lý bằng cách sử dụng một tập hợp hữu hạn nhưng không giới nội thời gian của các dấu hiệu có thể được sử dụng bởi các quy trình con trong phương pháp 100 để đưa ra ý kiến về sự tương tự giữa lệnh hỏi tin và các dự bị so khớp. Phương pháp 100 giải quyết một số vấn đề vốn có trong việc nhận dạng và định tính riêng một địa điểm địa lý và sự liên kết của nó với một hoặc nhiều thực thể doanh nghiệp bao gồm 1) sự thiếu chính xác và thay đổi tự nhiên trong việc khái niệm hóa và biểu diễn các dấu hiệu địa chỉ địa lý, 2) sự vắng mặt các ký hiệu nhận dạng không nhập nhằng hoặc các từ khóa ổn định cho các địa điểm địa lý, và 3) phương sai rất rộng ở mức độ chi tiết thường được xem xét khi sử dụng doanh nghiệp để có địa chỉ “đầy đủ”. Bằng cách mở rộng tập hợp linh hoạt và có thể thay đổi của các thuộc tính nhận dạng và so khớp đối với một địa điểm địa lý, các vấn đề này có thể được giải quyết. Tính linh hoạt của kỹ thuật này bao gồm cả siêu dữ liệu và các giá trị dữ liệu thực, và có thể được sử dụng cho cả hai việc 1) điền vào cơ sở dữ liệu với thông tin liên quan đến các địa điểm địa lý và kết nối giữa các địa điểm địa lý và các thực thể doanh nghiệp, cũng như 2) chọn các thực thể doanh nghiệp từ cơ sở dữ liệu dựa vào lệnh hỏi tin và các quy tắc kiểm soát ngưỡng chấp nhận đối với các mục đích nhất định.

Phương pháp 100 có thể giải quyết trường hợp một thực thể duy nhất ở một vị trí, ví dụ, một thực thể trong một tòa nhà, hoặc trường hợp nhiều thực thể ở một vị trí, ví dụ, nhiều người thuê trong trung tâm mua sắm. Đối với trường hợp nhiều thực thể ở một vị trí, mỗi thực thể này vẫn có thể có địa chỉ khác với địa chỉ của các thực thể khác.

Đối với trường hợp một thực thể, phương pháp 100 bao gồm các bước:

trong quy trình 120, thu dữ liệu, tức là, lệnh hỏi tin 110, mô tả vị trí;

trong quy trình 140, ngoại suy, từ dữ liệu này, địa chỉ được liên kết với vị trí;

trong quy trình 150, nhận dạng phân đoạn của con đường chính có địa chỉ này;

trong quy trình 150, xác định hình đa giác có chu vi bao quanh miền địa lý nằm ở gần phân đoạn này;

trong quy trình 160, thu nhận các tọa độ địa lý của một điểm nằm trong hình đa giác này;

trong quy trình 160, nhận dạng địa chỉ tại các tọa độ địa lý này; và
 trong quy trình 170, nhận dạng thực thể có liên kết với địa chỉ ở các tọa độ địa lý này.

Đối với trường hợp nhiều thực thể ở một vị trí, xem xét tình huống có hai thực thể là người thuê trong trung tâm mua sắm, trong đó mỗi thực thể này có địa chỉ khác với địa chỉ của thực thể kia. Trong quy trình 150, hình đa giác có thể có chu vi bao quanh trung tâm mua sắm. Do đó, giả sử rằng các tọa độ địa lý nêu trên là các tọa độ địa lý thứ nhất, điểm nêu trên là điểm thứ nhất, địa chỉ nêu trên tại các tọa độ địa lý là địa chỉ thứ nhất tại các tọa độ địa lý này, và thực thể nêu trên là thực thể thứ nhất. Khi được thi hành, quy trình 160 thu nhận các tọa độ địa lý thứ hai của điểm thứ hai nằm trong hình đa giác, và nhận dạng địa chỉ thứ hai tại các tọa độ địa lý thứ hai này. Tiếp đó, quy trình 170 nhận dạng thực thể thứ hai có liên kết với địa chỉ thứ hai.

Fig.2 là sơ đồ khối của hệ thống 200 sử dụng các phương pháp được mô tả ở đây. Hệ thống 200 có máy tính 205 ghép nối với mạng truyền thông dữ liệu, tức là mạng 230, như mạng Internet.

Máy tính 205 bao gồm giao diện người dùng 210, bộ xử lý 215, và bộ nhớ 220. Mặc dù máy tính 205 được thể hiện ở đây dưới dạng thiết bị độc lập, nhưng không chỉ giới hạn như vậy, mà còn có thể được ghép nối với các thiết bị khác (không được thể hiện) trong hệ thống xử lý phân tán.

Giao diện người dùng 210 bao gồm thiết bị nhập, như bàn phím hoặc hệ thống con nhận dạng tiếng nói, để cho phép người dùng 101 truyền thông thông tin và các lựa chọn hướng dẫn đến bộ xử lý 215. Giao diện người dùng 210 còn bao gồm thiết bị xuất như màn hình hoặc máy in. Cơ cấu điều khiển con trỏ như chuột, bi xoay, hoặc cần điều khiển trò chơi, cho phép người dùng 101 điều khiển con trỏ trên màn hình để truyền thông thông tin và các lựa chọn hướng dẫn khác đến bộ xử lý 215.

Hệ thống 200 còn bao gồm thiết bị người dùng 240 ghép nối truyền thông với máy tính 205 qua mạng 230. Người dùng 101 có thể tương tác với máy tính 205 qua thiết bị người dùng 240, thay cho việc làm như vậy qua giao diện người dùng 210.

Bộ xử lý 215 là thiết bị điện tử được tạo cấu hình của mạch logic để đáp lại và thi hành các lệnh.

Bộ nhớ 220 là vật ghi đọc được bằng máy tính phi nhất thời được mã hóa với chương trình máy tính. Với việc này, bộ nhớ 220 lưu trữ dữ liệu và các lệnh có thể đọc và thi hành bằng bộ xử lý 215 để điều khiển hoạt động của bộ xử lý 215. Bộ nhớ 220 có thể được lắp đặt trong bộ nhớ truy nhập ngẫu nhiên (RAM - Random Access Memory), ổ cứng, bộ nhớ chỉ đọc (ROM - Read Only Memory), hoặc tổ hợp của chúng. Một trong số các thành phần của bộ nhớ 220 là môđun chương trình 225.

Môđun chương trình 225 chứa các lệnh để điều khiển bộ xử lý 215 thi hành các phương pháp được mô tả ở đây. Ví dụ, theo sự điều khiển của môđun chương trình 225, bộ xử lý 215 thi hành các quy trình của phương pháp 100. Thuật ngữ "môđun" như được sử dụng ở đây để chỉ hoạt động chức năng có thể được thể hiện dưới dạng thành phần độc lập hoặc dưới dạng cấu hình tích hợp của nhiều thành phần con. Do vậy, môđun chương trình 225 có thể được thực hiện dưới dạng một môđun hoặc nhiều môđun hoạt động phối hợp với nhau. Hơn nữa, mặc dù môđun chương trình 225 được mô tả ở đây dưới dạng được lắp đặt trong bộ nhớ 220, và do đó được cài đặt trong phần mềm, nhưng nó có thể được thực hiện ở bất kỳ một trong số phần cứng (ví dụ, mạch điện tử), phần sụn, phần mềm, hoặc tổ hợp của chúng.

Bộ xử lý 215 nhận lệnh hỏi tin 110, qua mạng 230 hoặc qua giao diện người dùng 210. Bộ xử lý 215 truy nhập các quy tắc xử lý 135, bảng thuộc tính địa thực thể 137, các quy tắc đẩy ra 145, cơ sở dữ liệu tham chiếu phân đoạn 155, cơ sở dữ liệu tham chiếu tọa độ địa lý 165, và cơ sở dữ liệu tham chiếu thực thể doanh nghiệp 185, các quy tắc tính điểm 190, các quy tắc xếp hạng 191, và các quy tắc quyết định 192, chúng có thể là các thành phần của máy tính 205, ví dụ, được lưu trữ trong bộ nhớ 220, hoặc có thể nằm ngoài máy tính 205 khi đó máy tính 205 truy nhập chúng qua mạng cục bộ (không được thể hiện) hoặc qua mạng 230. Bộ xử lý 215 truy nhập các quy trình 120, 130, 140, 150, 160, 170, 180, 181, 193 và 194, các quy trình này có thể là các thành phần của máy tính 205, ví dụ, được lưu trữ trong bộ nhớ 220, hoặc có thể nằm ngoài máy tính 205 khi đó máy tính 205 truy nhập chúng qua mạng cục bộ hoặc qua mạng 230. Bộ xử lý 215 gửi kết quả 196 trả về người dùng 101 trong quy trình 195.

Dù được thể hiện dưới dạng đã được nạp sẵn vào bộ nhớ 220, nhưng môđun chương trình 225 cũng có thể được tạo cấu hình trên thiết bị lưu trữ 235 để sau đó nạp vào bộ nhớ 220. Thiết bị lưu trữ 235 còn là vật ghi đọc được bằng máy tính phi nhất thời

được mã hóa với chương trình máy tính, và có thể là thiết bị lưu trữ thông thường bất kỳ để lưu trữ mô đun chương trình 225 trong đó. Ví dụ về thiết bị lưu trữ 235 bao gồm đĩa mềm, đĩa compact, băng từ, bộ nhớ chỉ đọc, phương tiện nhớ quang, ổ nhớ tác động nhanh bus nối tiếp đa năng (USB - Universal Serial Bus), đĩa đa năng số, hoặc ổ đĩa zip. Thiết bị lưu trữ 235 cũng có thể là bộ nhớ truy nhập ngẫu nhiên, hoặc loại lưu trữ điện tử khác, nằm trong hệ thống lưu trữ từ xa và được ghép nối với máy tính 205 qua mạng 230.

Các kỹ thuật được mô tả ở đây làm ví dụ và không được hiểu là có ngụ ý hạn chế cụ thể nào đối với sáng chế. Cần phải hiểu rằng, các phương án thay thế, các tổ hợp và các sửa đổi có thể được tìm ra bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Ví dụ, các bước gắn với các quy trình được mô tả ở đây có thể được thực hiện theo thứ tự bất kỳ, trừ khi có quy định khác hoặc được chỉ rõ bởi chính các bước này. Sáng chế dự định bao hàm tất cả các phương án thay thế, sửa đổi và cải biến như vậy trong phạm vi của yêu cầu bảo hộ kèm theo.

Thuật ngữ "bao gồm " hoặc "gồm có" được hiểu là xác định sự có mặt của các dấu hiệu, số nguyên, bước hoặc thành phần đã nêu, nhưng không loại trừ sự có mặt của một hoặc nhiều dấu hiệu, số nguyên, bước hoặc thành phần khác hoặc các tổ hợp của chúng.

YÊU CẦU BẢO HỘ

1. Phương pháp khử nhập nhằng không tất định và so khớp thực thể theo định tính ở dữ liệu địa điểm địa lý bao gồm các bước:

- (a) nhận dữ liệu mô tả vị trí;
- (b) ngoại suy, từ dữ liệu nêu trên, địa chỉ gắn với vị trí này;
- (c) thực hiện quy trình xử lý mà thu được danh sách ứng viên, trong đó quy trình xử lý này bao gồm các bước:
 - nhận dạng đoạn đường lớn bao gồm địa chỉ nêu trên;
 - thu nhận các tọa độ địa lý của điểm tương ứng với địa chỉ nêu trên;
 - xác định hình đa giác có chu vi mà bao gồm vùng địa lý bao gồm các tọa độ địa lý nêu trên;
 - xác định thực thể gắn với địa chỉ nêu trên ở các tọa độ địa lý nêu trên;
 - đưa thực thể nêu trên vào trong danh sách ứng viên nêu trên; và
 - xác định địa chỉ bổ sung được bao gồm đa giác nêu trên;
- (d) gửi địa chỉ bổ sung nêu trên theo cách đệ quy vào trong quy trình nêu trên cho đến khi thỏa mãn tiêu chuẩn; và
- (e) quay lại danh sách ứng viên nêu trên.

2. Phương pháp theo điểm 1,

- trong đó bước ngoại suy nêu trên bao gồm thu thuộc tính thứ nhất gắn với địa chỉ nêu trên,
- trong đó bước nhận dạng đoạn nêu trên bao gồm thu thuộc tính thứ hai gắn với đoạn nêu trên, và
- trong đó bước xác định nêu trên bao gồm xác định kích thước và hình dáng của đa giác nêu trên dựa vào thuộc tính thứ nhất và thứ hai nêu trên.

3. Phương pháp theo điểm 1,

- trong đó bước ngoại suy nêu trên bao gồm thu thuộc tính thứ nhất gắn với địa chỉ nêu trên,
- trong đó bước nhận dạng đoạn nêu trên bao gồm thu thuộc tính thứ hai gắn với đoạn nêu trên,

trong đó phương pháp nêu trên còn bao gồm, trước bước thu nhận nêu trên, xác định kích thước và hình dáng của khối đa diện dựa vào thuộc tính thứ nhất và thứ hai nêu trên, và

trong đó đa giác nêu trên có mặt cắt ngang của khối đa diện.

4. Hệ thống xử lý nhập nhằm không tắt định và so khớp thực thể theo định tính ở dữ liệu địa điểm địa lý bao gồm:

bộ xử lý; và

bộ nhớ ghép nối truyền thông với bộ xử lý nêu trên, và chứa các mã lệnh đọc được bằng bộ xử lý nêu trên khiến cho bộ xử lý này thực hiện các hoạt động sau:

(a) nhận dữ liệu mô tả vị trí;

(b) ngoại suy, từ dữ liệu nêu trên, địa chỉ gắn với vị trí này;

(c) thực hiện quy trình xử lý mà thu được danh sách ứng viên, trong đó quy trình xử lý này bao gồm các bước:

nhận dạng đoạn đường lớn bao gồm địa chỉ nêu trên;

thu các tọa độ địa lý của điểm tương ứng với địa chỉ nêu trên;

xác định hình đa giác có chu vi mà bao gồm vùng địa lý bao gồm các tọa độ địa lý nêu trên;

xác định thực thể gắn với địa chỉ nêu trên ở các tọa độ địa lý nêu trên;

đưa thực thể nêu trên vào trong danh sách ứng viên nêu trên; và

xác định địa chỉ bổ sung được bao gồm đa giác nêu trên;

(d) gửi địa chỉ bổ sung nêu trên theo cách đệ quy vào trong quy trình nêu trên cho đến khi thỏa mãn tiêu chuẩn; và

(e) quay lại danh sách ứng viên nêu trên.

5. Hệ thống theo điểm 4,

trong đó bước ngoại suy nêu trên bao gồm thu thuộc tính thứ nhất gắn với địa chỉ nêu trên,

trong đó bước nhận dạng đoạn nêu trên bao gồm thu thuộc tính thứ hai gắn với đoạn nêu trên, và

trong đó bước xác định nêu trên bao gồm xác định kích thước và hình dáng của đa giác nêu trên dựa vào thuộc tính thứ nhất và thứ hai nêu trên.

6. Hệ thống theo điểm 4,

trong đó bước ngoại suy nêu trên bao gồm thu thuộc tính thứ nhất gắn với địa chỉ nêu trên,

trong đó bước nhận dạng đoạn nêu trên bao gồm thu thuộc tính thứ hai gắn với đoạn nêu trên,

trong đó các hoạt động cũng bao gồm, trước bước thu nhận nêu trên, xác định kích thước và hình dáng của khối đa diện dựa vào thuộc tính thứ nhất và thứ hai nêu trên, và

trong đó đa giác nêu trên có mặt cắt ngang của khối đa diện.

7. Thiết bị lưu trữ chứa các mã lệnh có thể đọc được bằng bộ xử lý khiến cho bộ xử lý này thực hiện các hoạt động sau:

(a) nhận dữ liệu mô tả vị trí;

(b) ngoại suy, từ dữ liệu nêu trên, địa chỉ gắn với vị trí này;

(c) thực hiện quy trình xử lý mà thu được danh sách ứng viên, trong đó quy trình xử lý này bao gồm các bước:

nhận dạng đoạn đường lớn bao gồm địa chỉ nêu trên;

thu các tọa độ địa lý của điểm tương ứng với địa chỉ nêu trên;

xác định hình đa giác có chu vi mà bao gồm vùng địa lý bao gồm các tọa độ địa lý nêu trên;

xác định thực thể gắn với địa chỉ nêu trên ở các tọa độ địa lý nêu trên;

đưa thực thể nêu trên vào trong danh sách ứng viên nêu trên; và

xác định địa chỉ bổ sung được bao gồm đa giác nêu trên;

(d) gửi địa chỉ bổ sung nêu trên theo cách đệ quy vào trong quy trình nêu trên cho đến khi thỏa mãn tiêu chuẩn; và

(e) quay lại danh sách ứng viên nêu trên.

8. Thiết bị lưu trữ theo điểm 7,

trong đó bước ngoại suy nêu trên bao gồm thu thuộc tính thứ nhất gắn với địa chỉ nêu trên,

trong đó bước nhận dạng đoạn nêu trên bao gồm thu thuộc tính thứ hai gắn với đoạn nêu trên, và

trong đó bước xác định nêu trên bao gồm xác định kích thước và hình dáng của đa giác nêu trên dựa vào thuộc tính thứ nhất và thứ hai nêu trên.

9. Thiết bị lưu trữ theo điểm 7,

trong đó bước ngoại suy nêu trên bao gồm thu thuộc tính thứ nhất gắn với địa chỉ nêu trên,

trong đó bước nhận dạng đoạn nêu trên bao gồm thu thuộc tính thứ hai gắn với đoạn nêu trên,

trong đó các hoạt động cũng bao gồm, trước bước thu nhận nêu trên, xác định kích thước và hình dáng của khối đa diện dựa vào thuộc tính thứ nhất và thứ hai nêu trên, và

trong đó đa giác nêu trên có mặt cắt ngang của khối đa diện.

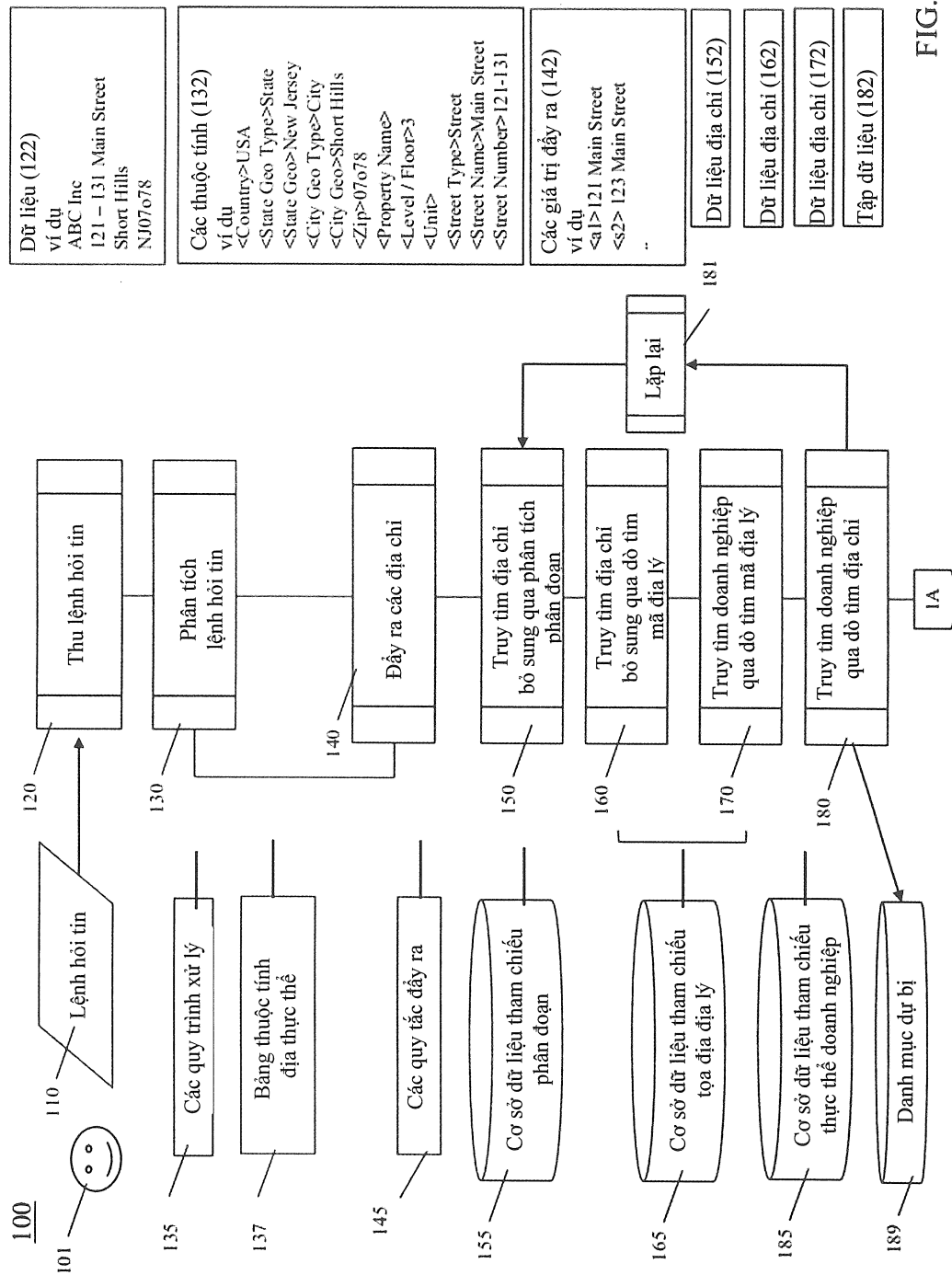


FIG. 1A

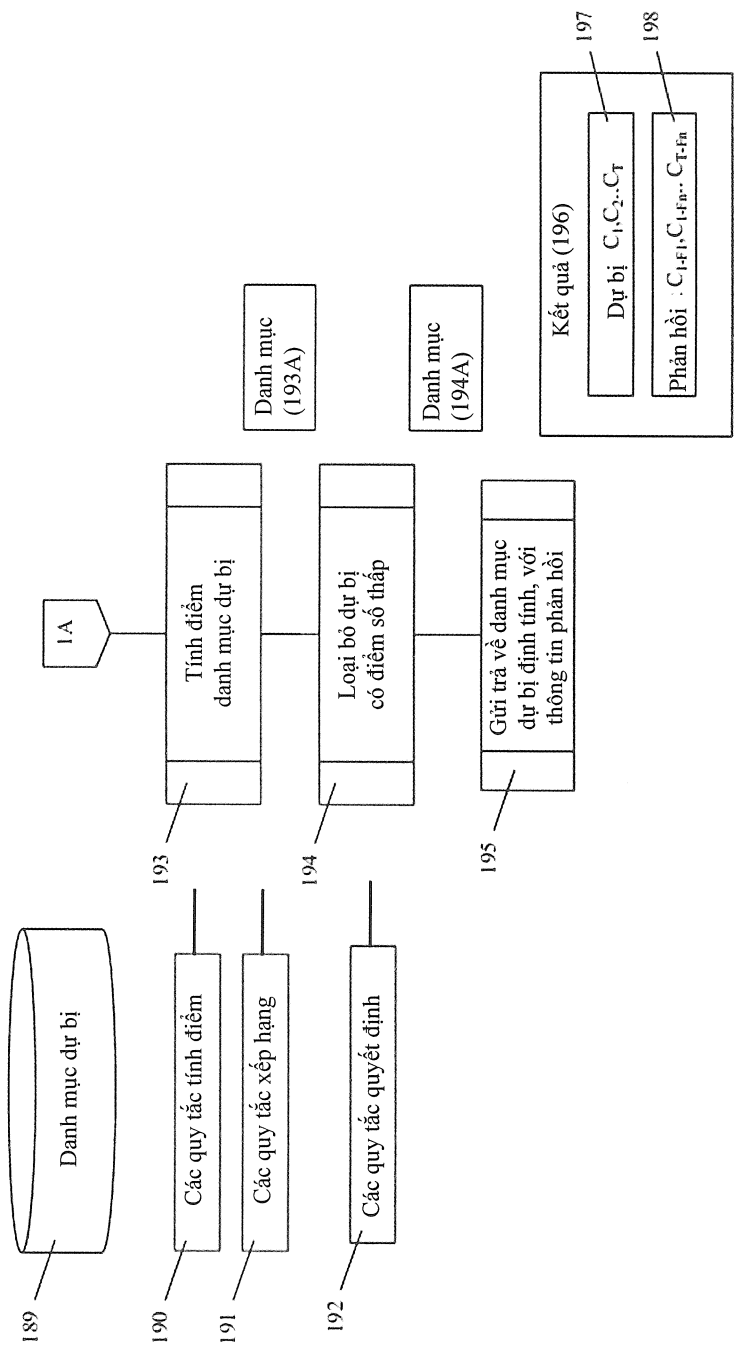


FIG. 1B

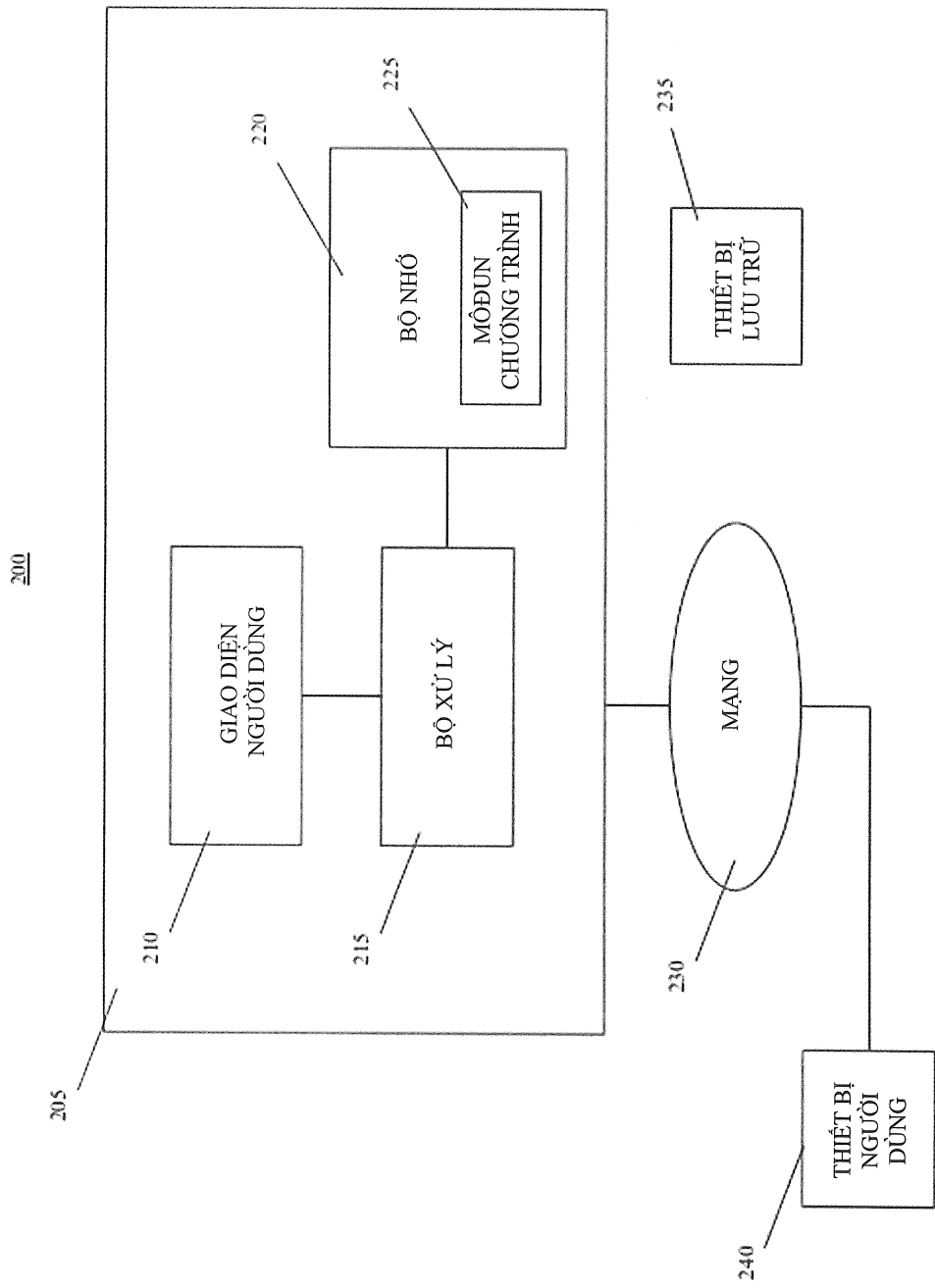


FIG. 2