



(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)



1-0033572

(51)⁷ G06F 17/30; G06Q 30/02

(13) B

(21) 1-2017-00815

(22) 07/08/2015

(86) PCT/JP2015/072564 07/08/2015

(87) WO 2016/021726 11/02/2016

(30) 2014-162623 08/08/2014 JP; 2015-000945 06/01/2015 JP

(45) 25/10/2022 415

(43) 25/08/2017 353A

(73) Hakuhodo DY Holdings Inc. (JP)

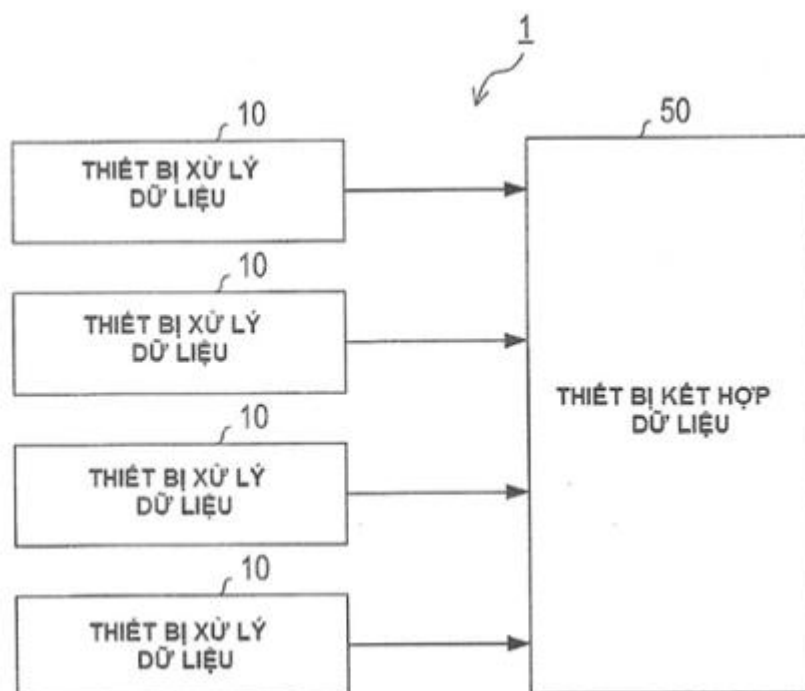
3-1, Akasaka 5-chome, Minato-ku, Tokyo 107-6322 Japan

(72) DOMOTO, Ryo (JP); TOKUHISA, Shinya (JP).

(74) Công ty TNHH Trà và cộng sự (TRA & ASSOCIATES CO.,LTD)

(54) HỆ THỐNG XỬ LÝ THÔNG TIN

(57) Sáng chế đề cập đến hệ thống xử lý thông tin gồm bộ phận lưu trữ (55) và bộ phận kết hợp (51). Bộ phận kết hợp được cấu hình để lưu cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai (155) được lưu bởi bộ phận lưu trữ. Cơ sở dữ liệu thứ nhất được cung cấp với dữ liệu đặc trưng của mỗi phần tử ảo, dữ liệu đặc trưng được tạo ra bằng cách tích hợp dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng dựa vào dữ liệu đặc trưng của mỗi phần tử của nhóm thứ nhất. Cơ sở dữ liệu thứ hai được cung cấp với dữ liệu đặc trưng của mỗi phần tử của nhóm thứ hai. Mỗi dữ liệu đặc trưng được cung cấp trong cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai, gồm dữ liệu tham chiếu đại diện cho loại đặc trưng chung. Bộ phận kết hợp kết hợp cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai.



Lĩnh vực kỹ thuật đề cập

Sáng chế đề cập đến hệ thống xử lý thông tin để quản lý cơ sở dữ liệu.

Tình trạng kỹ thuật của sáng chế

Việc phân tích hành vi mua sắm của khách hàng thường được thực hiện dựa vào dữ liệu bán hàng. Để hỗ trợ các hoạt động thương mại, việc phân tích cũng được thực hiện đối với hành vi giao tiếp của khách hàng với phương tiện truyền thông đại chúng và các nội dung mạng.

Bằng cách đưa ra các bản câu hỏi và các câu hỏi trực diện, các thông tin khác như hành vi mua sắm của khách hàng, hành vi giao tiếp của khách hàng với phương tiện truyền thông/ các nội dung mạng, phong cách sống và các thông tin tương tự cũng được thu thập.

Trong những năm gần đây, mỗi công ty có một cơ sở dữ liệu lớn với dữ liệu như vậy về khách hàng. Tuy nhiên, mỗi công ty đều không sẵn lòng cung cấp dữ liệu về khách hàng ra ngoài công ty, chủ yếu để bảo vệ thông tin cá nhân. Dữ liệu này, khi được cung cấp ra bên ngoài công ty nắm giữ dữ liệu, được cung cấp bằng cách mã hóa, với việc loại bỏ thông tin quan trọng dẫn đến việc nhận dạng khách hàng, hoặc các thay đổi có chủ tâm để bao gồm nhiều lỗi (nhiều) (xem Tài liệu sáng chế 1).

Tài liệu sáng chế 1: Công bố Đơn sáng chế không thẩm định Nhật Bản số: 2014-109647

Bản chất kỹ thuật của sáng chế

Như đã mô tả trên đây, việc cung cấp dữ liệu về các khách hàng từ các công ty nắm giữ dữ liệu bị giới hạn bởi quan điểm bảo vệ thông tin cá nhân. Các công ty nắm giữ dữ liệu có xu hướng không sẵn lòng cung cấp không chỉ dữ liệu về các khách hàng mà còn đối với cả dữ liệu hạt, do giá trị thông tin cao của nó và quan ngại về các bất lợi gây ra bởi việc cung cấp. Do đó, theo lĩnh vực kỹ thuật đã biết, khó sử dụng hiệu quả các dữ liệu khác nhau có sẵn trong xã hội.

Vì vậy theo một khía cạnh của sáng chế, người ta mong muốn có thể tạo ra một công nghệ mới liên quan đến các cơ sở dữ liệu mà có thể sử dụng một cách hiệu quả các dữ liệu khác nhau có sẵn trong xã hội.

Hệ thống xử lý thông tin theo một khía cạnh của sáng chế bao gồm một bộ phận lưu trữ cấu hình để lưu trữ cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai, và một bộ phận kết hợp được cấu hình để kết hợp cơ sở dữ liệu thứ nhất với cơ sở dữ liệu thứ hai. Theo một khía cạnh của sáng chế, hệ thống xử lý thông tin có thể bao gồm thêm một bộ phận phát sinh được cấu hình để tạo ra cơ sở dữ liệu thứ nhất.

Cơ sở dữ liệu thứ nhất là một cơ sở dữ liệu bao gồm dữ liệu đặc trưng của mỗi phần tử ảo liên quan đến các phần tử của nhóm thứ nhất. Dữ liệu đặc trưng của mỗi phần tử ảo có thể được tạo ra bằng cách hợp nhất dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng, dựa vào dữ liệu đặc trưng của phần tử ở nhóm thứ nhất. Dữ liệu đặc trưng của một phần tử đại diện cho các đặc trưng liên quan đến phần tử. Ví dụ, bộ phận phát sinh có thể được cấu hình để xử lý thông kê dữ liệu đặc trưng của phần lớn phần tử đối với phần lớn phần tử giống hoặc tương tự về đặc trưng trong nhóm thứ nhất; theo cách đó hợp nhất dữ liệu đặc trưng của phần lớn phần tử đối với mỗi phần tử giống hoặc tương tự về đặc trưng để tạo ra dữ liệu đặc trưng của mỗi phần tử ảo.

Cơ sở dữ liệu thứ hai bao gồm dữ liệu đặc trưng đối với mỗi phần tử của nhóm thứ hai. Nhóm thứ hai là giống hoặc khác với nhóm thứ nhất. Mỗi dữ liệu đặc trưng của cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai được cấu hình để bao gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp. Dữ liệu tham chiếu đại diện cho một loại đặc trưng chung giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai.

Bộ phận kết hợp kết hợp cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng được đại diện bởi dữ liệu tham chiếu giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai.

Theo hệ thống xử lý thông tin trong một khía cạnh của sáng chế này, mỗi dữ liệu đặc trưng được tích hợp để tạo ra một cơ sở dữ liệu gồm dữ liệu đặc trưng cho mỗi phần tử ảo. Dữ liệu đặc trưng của một phần tử ảo trong đó các đặc trưng của phần tử ảo được trộn. Ví dụ, trong trường hợp ở đó nhóm thứ nhất là một nhóm khách hàng, và

dữ liệu đặc trưng trước khi tích hợp là dữ liệu đặc trưng của một khách hàng đơn đại diện cho các đặc trưng của một khách hàng, dữ liệu đặc trưng của một phần tử ảo tương ứng với dữ liệu trong đó các đặc trưng của phần lớn khách hàng được trộn.

Bằng cách biến đổi dữ liệu đặc trưng của mỗi phần tử thành dữ liệu đặc trưng của mỗi phần tử ảo, hạt thông tin của dữ liệu đặc trưng có thể được tăng. Việc bảo vệ thông tin đối với thông tin chi tiết có thể đạt được. Do đó, theo một khía cạnh của sáng chế, có thể giảm các rào cản đối với việc cung cấp dữ liệu bởi các nhà nắm giữ dữ liệu do việc bảo vệ thông tin. Do đó, theo một khía cạnh của sáng chế này, hệ thống có thể được cung cấp mà có thể sử dụng một cách hiệu quả các dữ liệu khác nhau có sẵn trong xã hội.

Phần lớn phần tử giống hoặc tương tự về đặc trưng đã mô tả trên đây có thể là sự pha trộn cả phần tử giống về đặc trưng và bộ phận tương tự về đặc trưng. Thuật ngữ “sự tương tự” trong phần mô tả này có thể được hiểu như một thuật ngữ rộng bao gồm khái niệm “giống” như một khái niệm phụ.

Các kỹ thuật phân cụm đã biết đến rộng rãi (các kỹ thuật phân loại) có thể được sử dụng lúc hợp nhất dữ liệu đặc trưng của phần lớn phần tử giống nhau hoặc tương tự về đặc trưng. Các kỹ thuật phân cụm khác nhau cũng được biết đến trong lĩnh vực xử lý ngôn ngữ, nhận dạng mẫu, trí tuệ nhân tạo và tương tự. Các kỹ thuật phân cụm nổi tiếng bao gồm kỹ thuật phân cụm bằng phần cứng như k-means (phương pháp k-means) và BIRCH (phân cụm và rút gọn lặp cân bằng sử dụng hệ thứ bậc), và các kỹ thuật phân cụm bằng phần mềm như LSA (phân tích ngữ nghĩa tiềm ẩn), pLSA (phân tích ngữ nghĩa tiềm ẩn theo xác suất), và LDA (phương pháp phân phối Dirichlet tiềm ẩn).

Theo một khía cạnh của sáng chế này, bộ phận phát sinh có thể được cấu hình để sử dụng các kỹ thuật phân cụm nổi tiếng này hoặc các kỹ thuật phân cụm khác để phân cụm một bộ dữ liệu đặc trưng liên quan đến nhóm thứ nhất thành từng dữ liệu đặc trưng giống hoặc tương tự về đặc trưng. Bộ phận phát sinh có thể được cấu hình để tích hợp từng cụm dữ liệu của phần lớn dữ liệu đặc trưng tương ứng với phân cụm.

Theo một khía cạnh của sáng chế này, bộ phận phát sinh có thể được cấu hình để xử lý thống kê từng cụm của phần lớn dữ liệu đặc trưng tương ứng với phân cụm. Đặc biệt là, bộ phận phát sinh có thể được cấu hình để tính cho mỗi cụm một thống kê

đối với mỗi tham số của phần lớn dữ liệu đặc trưng tương ứng với cụm. Các ví dụ về thống kê gồm một giá trị trung bình, một giá trị lớn nhất, một giá trị nhỏ nhất, và một giá trị đại diện sự phân bố (tỷ lệ hoặc tương tự). Giá trị tham số của dữ liệu đặc trưng và thống kê của nó có thể là một lượng vô hướng hoặc một lượng vectơ. Dữ liệu đặc trưng của một phần tử ảo có thể được tạo ra cho mỗi cụm. Theo một khía cạnh của sáng chế, dữ liệu đặc trưng của một bộ phận ảo có thể được tạo ra để bao gồm con số thống kê như một giá trị tham số đại diện cho dữ liệu đặc trưng liên quan đến bộ phận ảo.

Cơ sở dữ liệu thứ nhất có thể được tạo ra bởi bộ phận phát sinh, có thể được tạo ra thủ công bởi con người, hoặc có thể được tạo ra bằng phương pháp khác (ví dụ, sự kết hợp của xử lý thông tin bằng máy tính và công việc thủ công bởi con người).

Theo một khía cạnh của sáng chế, cơ sở dữ liệu thứ hai cũng có thể được tạo ra theo cách giống như cơ sở dữ liệu thứ nhất. Dữ liệu đặc trưng của mỗi bộ phận của cơ sở dữ liệu thứ hai có thể là dữ liệu đặc trưng của một bộ phận ảo. Dữ liệu đặc trưng của mỗi phần tử ảo có thể được tạo ra bằng cách hợp nhất dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng, dựa vào dữ liệu đặc trưng của mỗi bộ phận của một nhóm giống hoặc khác với nhóm thứ nhất. Theo một khía cạnh của sáng chế này, hệ thống xử lý thông tin có thể bao gồm một bộ phận phát sinh để tạo ra cơ sở dữ liệu thứ hai.

Theo một khía cạnh của sáng chế này, cơ sở dữ liệu thứ hai có thể là một cơ sở dữ liệu bao gồm dữ liệu đặc trưng không được tích hợp. Cơ sở dữ liệu thứ hai có thể gồm dữ liệu nguồn đơn làm dữ liệu đặc trưng. Mỗi dữ liệu nguồn đơn là dữ liệu dựa vào thông tin được thu thập từ một nguồn thông tin đơn. Dữ liệu nguồn đơn là dữ liệu đặc trưng của một phần tử không được thống kê và mô hình hóa, và có thể là dữ liệu đặc trưng đại diện cho các đặc trưng liên quan đến một bộ phận đơn.

Một trong các nhóm thứ nhất và nhóm thứ hai đã mô tả trên đây có thể là một tập hợp bao gồm ít nhất một trong nhiều người, các đồ vật, các dịch vụ, và các địa điểm. Bộ phận có thể là một yếu tố của tập hợp này. Tức là, bộ phận có thể là một người, một đồ vật, một dịch vụ, và một địa điểm bất kỳ. Các đồ vật, các dịch vụ, và các địa điểm có thể là các đồ vật, các dịch vụ, và các địa điểm liên quan đến người.

Theo một khía cạnh của sáng chế, một trong số các nhóm thứ nhất và nhóm thứ hai có thể là một tập hợp về người, và còn lại của nhóm thứ nhất và nhóm thứ hai có thể là một tập hợp gồm ít nhất một trong số các đồ vật, các dịch vụ, và các địa điểm. Mỗi nhóm trong số các nhóm thứ nhất và nhóm thứ hai có thể là một tập hợp gồm ít nhất một trong số các đồ vật, các dịch vụ, và các địa điểm.

Theo một khía cạnh của sáng chế, hệ thống xử lý thông tin có thể bao gồm phần lớn bộ phận phát sinh được cấu hình để tạo ra các cơ sở dữ liệu cho các nhóm giống hoặc khác nhau, và bộ phận kết hợp được cấu hình để kết hợp phần lớn cơ sở dữ liệu được tạo ra bởi phần lớn bộ phận phát sinh. Mỗi bộ phận phát sinh có thể được cấu hình để tạo ra một cơ sở dữ liệu gồm dữ liệu đặc trưng cho mỗi phần tử ảo trong đó dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng được tích hợp dựa vào dữ liệu đặc trưng mỗi phần tử của nhóm tương ứng.

Mỗi dữ liệu đặc trưng của phần lớn cơ sở dữ liệu có thể gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp. Dữ liệu tham đại diện cho một loại đặc trưng chung của cơ sở dữ liệu được kết hợp. Bộ phận kết hợp có thể được cấu hình để kết hợp phần lớn cơ sở dữ liệu để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng được đại diện bởi dữ liệu tham chiếu trong số các cơ sở dữ liệu.

Theo một khía cạnh của sáng chế, hệ thống xử lý thông tin có thể được tạo ra bao gồm: một bộ phận lưu trữ được cấu hình để lưu trữ phần lớn cơ sở dữ liệu có dữ liệu đặc trưng của mỗi phần tử ảo; và một bộ phận kết hợp được cấu hình để kết hợp phần lớn cơ sở dữ liệu được lưu trữ trong bộ phận lưu trữ. Thậm chí với hệ thống xử lý thông tin này, dữ liệu có sẵn trong xã hội có thể được sử dụng một cách hiệu quả.

Theo một khía cạnh của sáng chế này, bộ phận kết hợp có thể được cấu hình để kết hợp phần lớn cơ sở dữ liệu qua một loại cơ sở dữ liệu cụ thể của nhóm giống hoặc nhóm khác với một trong các nhóm tương ứng với phần lớn cơ sở dữ liệu. Loại cơ sở dữ liệu cụ thể có dữ liệu nguồn đơn làm dữ liệu đặc trưng cho mỗi phần tử của nhóm. Dữ liệu nguồn đơn gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp các cơ sở dữ liệu. Dữ liệu tham chiếu đại diện cho một loại đặc trưng chung đối với cơ sở dữ liệu được kết hợp.

Theo một khía cạnh của sáng chế này, bộ phận kết hợp có thể được cấu hình để kết hợp phần lớn cơ sở dữ liệu thông qua loại cơ sở dữ liệu cụ thể, bằng cách kết hợp

loại cơ sở dữ liệu cụ thể với mỗi cơ sở dữ liệu khác để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu.

Dữ liệu nguồn đơn đại diện cho các đặc trưng của phần tử vì việc tích hợp trên đây chưa được thực hiện. Do đó, bằng cách kết hợp các cơ sở dữ liệu qua dữ liệu nguồn đơn, có thể liên kết thích hợp hơn dữ liệu đặc trưng của mỗi phần tử ảo trong số các cơ sở dữ liệu.

Theo một khía cạnh của sáng chế, các chức năng của bộ phận phát sinh và bộ phận kết hợp đã mô tả trên đây có thể được thực hiện bằng phần cứng. Chức năng của mỗi bộ phận này có thể được thực hiện bằng một máy tính bởi một chương trình. Các chức năng của những bộ phận này có thể được thực hiện theo cách phân bố bằng phần lớn máy tính. Một chương trình để làm cho máy tính thực hiện ít nhất một chức năng của những bộ phận này có thể được cung cấp cho máy tính. Chương trình có thể được ghi lại trên một phương tiện ghi không chuyển tiếp đọc được bằng máy tính như một bộ nhớ bán dẫn, một đĩa từ tính, một đĩa quang học và thiết bị tương tự.

Theo một khía cạnh của sáng chế này, thiết bị lưu trữ đọc được bằng máy tính không chuyển tiếp có thể được bố trí để lưu một chương trình máy tính thực hiện ít nhất một chức năng của bộ phận phát sinh và bộ phận kết hợp. Theo một khía cạnh của sáng chế này, có thể tạo ra một hệ thống xử lý thông tin gồm một máy tính (bộ xử lý) và một bộ nhớ, bộ nhớ lưu chương trình đã mô tả trên đây.

Theo một khía cạnh của sáng chế, sáng chế đề xuất một phương pháp kết hợp cơ sở dữ liệu gồm việc tạo ra hoặc có được một cơ sở dữ liệu thứ nhất và kết hợp cơ sở dữ liệu thứ nhất với cơ sở dữ liệu thứ hai. Theo một khía cạnh khác của sáng chế, sáng chế đề xuất một phương pháp kết hợp cơ sở dữ liệu gồm việc tạo ra hoặc có được một cơ sở dữ liệu thứ nhất, tạo ra hoặc có được một cơ sở dữ liệu thứ hai, và kết hợp cơ sở dữ liệu thứ nhất với cơ sở dữ liệu thứ hai.

Cơ sở dữ liệu được mô tả trên đây có thể là một cơ sở dữ liệu về khách hàng. Cơ sở dữ liệu có thể được cấu hình để bao gồm dữ liệu đặc trưng đại diện cho ít nhất một đặc trưng của một người, một đồ vật, một dịch vụ, và một địa điểm tương ứng với mỗi khách hàng. Theo hệ thống xử lý thông tin của một khía cạnh của sáng chế, dữ liệu về khách hàng có thể được cung cấp đến bộ phận kết hợp trong khi thông tin cá

nhân của các khách hàng được bảo vệ. Bộ phận kết hợp có thể tạo ra một cơ sở dữ liệu gồm dữ liệu về các khách hàng đã kết hợp của các nhà cung cấp dữ liệu khác nhau.

Trong trường hợp ở đó mỗi cơ sở dữ liệu bao gồm dữ liệu đặc trưng đại diện cho các đặc trưng của từng khách hàng, dữ liệu tham khảo có thể được cấu hình làm dữ liệu đại diện cho thuộc tính nhân khẩu của khách hàng. Các thuộc tính nhân khẩu có trong dữ liệu về các khách hàng được giữ bởi các công ty khác nhau. Bằng cách kết hợp các cơ sở dữ liệu dựa vào dữ liệu tham khảo đại diện cho thuộc tính nhân khẩu của khách hàng, có thể kết hợp một cách thích hợp dữ liệu đặc trưng của các cơ sở dữ liệu khác nhau. Tất nhiên, dữ liệu tham khảo có thể là dữ liệu đại diện cho một đặc trưng của khách hàng khác với thuộc tính nhân khẩu của khách hàng, ngoài hoặc trong địa điểm của thuộc tính nhân khẩu của khách hàng.

Dữ liệu đặc trưng về các khách hàng có thể gồm dữ liệu đại diện cho ít nhất một đặc trưng về hành vi tiêu dùng của khách hàng, hoạt động giao tiếp với phương tiện truyền thông đại chúng, hành vi giao tiếp với nội dung mạng, phong cách sống, sự am hiểu về giá trị và thương hiệu, như là dữ liệu không tham khảo khác với dữ liệu tham khảo. Bằng cách kết hợp dữ liệu đặc trưng gồm các thông tin khác nhau về khách hàng trong số các cơ sở dữ liệu, có thể tách bằng cách phân tích các loại thông tin thống kê có ý nghĩa khác nhau về khách hàng (người tiêu dùng) từ cơ sở dữ liệu kết hợp.

Dữ liệu không tham khảo đưa ra ở đây làm ví dụ có thể cũng là dữ liệu tham khảo. Dữ liệu tham khảo có thể là bất kỳ dữ liệu nào đại diện cho một loại đặc trưng chung trong số các cơ sở dữ liệu được kết hợp trực tiếp từ phần lớn cơ sở dữ liệu.

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối thể hiện cấu hình của hệ thống xử lý thông tin theo phương án thực hiện thứ nhất.

Fig.2A là sơ đồ khối thể hiện cấu hình của thiết bị xử lý dữ liệu, và Fig.2B là sơ đồ khối thể hiện việc xử lý được thực hiện bởi bộ phận tính toán của thiết bị xử lý dữ liệu.

Fig.3A đến Fig.3D là một sơ đồ thể hiện cấu trúc của dữ liệu đặc trưng được giữ bởi một cơ sở dữ liệu liên quan đến khách hàng.

Fig.4 là sơ đồ thể hiện một ví dụ về cơ sở dữ liệu sau xử lý.

Fig.5A là sơ đồ khối thể hiện cấu hình của thiết bị kết hợp dữ liệu, và Fig.5B là sơ đồ khối thể hiện việc xử lý được thực hiện bởi một bộ phận tính toán của thiết bị kết hợp dữ liệu.

Fig.6A là sơ đồ thể hiện một ví dụ về cách kết hợp các cơ sở dữ liệu sau xử lý.

Fig.6B là sơ đồ thể hiện cấu trúc của các cơ sở dữ liệu sau xử lý để kết hợp.

Fig.6C là sơ đồ thể hiện một ví dụ về cấu trúc của cơ sở dữ liệu kết hợp dựa vào các cơ sở dữ liệu sau xử lý.

Fig.7 là sơ đồ thể hiện một ví dụ phân tích cơ sở dữ liệu kết hợp.

Fig.8A là sơ đồ thể hiện một ví dụ về cách kết hợp các cơ sở dữ liệu sau xử lý theo phương án thực hiện thứ hai

Fig.8B là sơ đồ thể hiện một ví dụ về cấu trúc của cơ sở dữ liệu kết hợp theo phương án thực hiện thứ hai.

Fig.9A là sơ đồ minh họa cấu trúc của các cơ sở dữ liệu liên quan đến khách hàng theo phương án thực hiện thứ ba.

Fig.9B là sơ đồ minh họa cấu trúc của các cơ sở dữ liệu liên quan đến khách hàng theo phương án thực hiện thứ tư.

Fig.10A là sơ đồ thể hiện cấu trúc của các cơ sở dữ liệu liên quan đến khách hàng theo phương án thực hiện thứ năm.

FIG.10B là sơ đồ thể hiện cấu trúc của các cơ sở dữ liệu liên quan đến khách hàng theo phương án thực hiện thứ sáu.

Fig.11A là sơ đồ thể hiện cấu trúc của các cơ sở dữ liệu liên quan đến khách hàng theo phương án thực hiện thứ bảy.

Fig.11B là sơ đồ thể hiện cấu trúc của các cơ sở dữ liệu liên quan đến khách hàng theo phương án thực hiện thứ tám.

Mô tả chi tiết sáng chế

Dưới đây, các phương án thực hiện của sáng chế sẽ được mô tả cùng với việc tham khảo đến các hình vẽ.

[Phương án thực hiện thứ nhất]

Hệ thống xử lý thông tin 1 theo phương án thực hiện này trên Fig.1 gồm phần lớn thiết bị xử lý dữ liệu 10 và một thiết bị kết hợp dữ liệu 50. Mỗi thiết bị xử lý dữ liệu 10 xử lý một cơ sở dữ liệu 153 về khách hàng được giữ bởi một công ty và cung cấp một cơ sở dữ liệu sau xử lý 155 cho thiết bị kết hợp dữ liệu 50. Thiết bị xử lý dữ liệu 10 được bố trí ở phía công ty nơi cơ sở dữ liệu 153 về khách hàng được giữ.

Dưới đây, cơ sở dữ liệu 153 về các khách hàng được thể hiện một cơ sở dữ liệu liên quan đến khách hàng 153. Một công ty xử lý và cung cấp cơ sở dữ liệu liên quan đến khách hàng 153 được thể hiện một nhà cung cấp dữ liệu. Cơ sở dữ liệu liên quan đến khách hàng 153 bao gồm một nhóm dữ liệu đặc trưng về khách hàng.

Một ví dụ về dữ liệu đặc trưng có thể là dữ liệu đại diện cho ít nhất một đặc trưng về hành vi tiêu dùng của khách hàng, hành vi giao tiếp với phương tiện truyền thông đại chúng, hành vi giao tiếp với các nội dung mạng, phong cách sống, khả năng nhận thức về giá trị và thương hiệu, cùng với ID của khách hàng hoặc thuộc tính nhân khẩu của khách hàng.

ID của khách hàng có thể là một mã định danh mà dữ liệu đại diện cho thuộc tính nhân khẩu học của khách hàng có thể được rút từ một cơ sở dữ liệu lưu thông tin là khách hàng. Hành vi tiêu dùng của khách hàng bao gồm hành vi mua sắm của khách hàng.

Ví dụ, dữ liệu đặc trưng có thể được cấu hình để có dữ liệu mua sắm của khách hàng, cùng với thuộc tính nhân khẩu học của khách hàng. Dữ liệu mua sắm có thể là dữ liệu đại diện cho các hàng hóa đã mua và/ hoặc giá mua. Trong trường hợp này, cơ sở dữ liệu liên quan đến khách hàng 153 có thể có dữ liệu đặc trưng cho mỗi khách hàng và cho mỗi thanh toán.

Thiết bị xử lý dữ liệu 10 xử lý phần lớn dữ liệu đặc trưng tương tự về đặc trưng thành một dữ liệu đặc trưng, dựa vào nhóm dữ liệu đặc trưng của cơ sở dữ liệu liên quan đến khách hàng 153. Thuật ngữ “tương tự” được dùng trong tài liệu này có thể được hiểu như một thuật ngữ rộng bao gồm “giống” như một khái niệm phụ. Thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 như một cơ sở dữ liệu bao gồm dữ liệu đặc trưng đã xử lý làm dữ liệu đặc trưng của một khách hàng ảo. Trong bản mô tả này, cơ sở dữ liệu liên quan đến khách hàng sau khi xử lý cũng được thể hiện như một cơ sở dữ liệu sau xử lý 155. Cơ sở dữ liệu sau xử lý 155, ví dụ, bao gồm dữ liệu đặc

trung của mỗi khách hàng ảo tương ứng với mỗi khách hàng.

Cơ sở dữ liệu sau xử lý 155 được cung cấp cho thiết bị kết hợp dữ liệu 50, ví dụ, bằng sự truyền thông. Để ngăn rò rỉ cơ sở dữ liệu, các khóa mật mã và các đường thuê bao có thể được sử dụng để truyền thông. Theo một ví dụ khác, cơ sở dữ liệu sau xử lý 155 có thể được lưu trong một phương tiện ghi đọc được bằng máy tính xách tay và được cung cấp cho thiết bị kết hợp dữ liệu 50 nhờ việc chuyển của phương tiện ghi. Các ví dụ về phương tiện ghi gồm bộ nhớ bán dẫn, đĩa quang học như đĩa CD-ROM và đĩa DVD, và đĩa từ tính như đĩa cứng.

Thiết bị kết hợp dữ liệu 50 kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 cung cấp từ các thiết bị xử lý dữ liệu 10 tương ứng để tạo ra một cơ sở dữ liệu kết hợp 557. Một ví dụ về nhà cung cấp dữ liệu bao gồm, như mô tả trên đây, một công ty có cơ sở dữ liệu liên quan đến khách hàng 153 bao gồm dữ liệu mua sắm của khách hàng làm dữ liệu đặc trưng.

Các ví dụ về các công ty nắm giữ dữ liệu mua sắm bao gồm các nhà bán lẻ (các thương nhân bán hàng hóa) điều hành các cửa hàng tiện lợi, các siêu thị và hoặc cửa hàng bách hóa, các công ty điều hành các máy bán hàng tự động, các công ty cung cấp dịch vụ thanh toán bằng tiền điện tử, và các công ty cung cấp dịch vụ thanh toán bằng thẻ tín dụng.

Nhà bán lẻ có cơ sở dữ liệu trong đó dữ liệu mua sắm đối với mỗi lần thanh toán có được thông qua hệ thống POS v.v, được tích lũy. Như một hệ thống POS, hệ thống ID-POS cho phép nhận dạng các khách hàng được biết đến.

Đối với các nhà bán lẻ, thực tế nhận một thẻ điểm từ một khách hàng khi thanh toán và tích lũy dữ liệu mua sắm kết hợp với một ID khách hàng xác định từ thẻ điểm, để xây dựng một cơ sở dữ liệu gồm dữ liệu mua sắm cho mỗi lần thanh toán cho phép xác định các khách hàng. Cũng là thực tế để nhập giới tính và tuổi của các khách hàng được xác định bởi một người bán hàng từ sự xuất hiện khi thanh toán, nhờ đó lưu dữ liệu mua sắm cho mỗi lần thanh toán theo cách kết hợp các thuộc tính nhân khẩu học với dữ liệu mua sắm của các khách hàng.

Thiết bị kết hợp dữ liệu 50 có thể xây dựng một cơ sở dữ liệu lớn mà một công ty riêng lẻ không thể xây dựng, bằng cách kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 gồm dữ liệu mua sắm của các nhà cung cấp dữ liệu khác nhau. Ví dụ, có thể tạo ra một

cơ sở dữ liệu kết hợp 557 cho phép phân tích hành vi mua sắm của các khách hàng (người tiêu dùng) đối với một phạm vi hàng hóa rộng.

Như đã thể hiện trên Fig.2A, mỗi thiết bị xử lý dữ liệu 10 bao gồm một bộ phận tính toán 11, một bộ phận lưu trữ 15, một bộ phận nhập 17, và một bộ phận xuất 19. Bộ phận tính toán 11 gồm một CPU 13 để thực hiện việc xử lý theo các chương trình 151 khác nhau. Hơn nữa, bộ phận tính toán 11 gồm một RAM 14 được sử dụng như một bộ nhớ làm việc khi thực hiện việc xử lý bởi CPU13. Như đã thể hiện trên Fig.2B, bộ phận tính toán 11 hoạt động một bộ phận phân cụm 110 và một bộ phận xử lý 115, bởi vì việc thực hiện xử lý bởi CPU 13.

Bộ phận lưu trữ 15 chứa các chương trình 151 khác nhau bao gồm các chương trình để làm cho bộ phận tính toán 11 thực hiện chức năng như một bộ phận phân cụm 110 và một bộ phận xử lý 115, và cơ sở dữ liệu liên quan đến khách hàng 153 trước khi xử lý. Hơn nữa, bộ phận lưu trữ 15 được cấu hình để lưu các cơ sở dữ liệu sau xử lý 155 đã tạo ra bởi vì việc xử lý cơ sở dữ liệu liên quan đến khách hàng 153.

Dữ liệu tạo ra cơ sở dữ liệu liên quan đến khách hàng 153 được nhập vào từ bộ phận nhập 17. Bộ phận nhập 17 có thể là một giao diện thông qua đó dữ liệu tạo ra cơ sở dữ liệu liên quan đến khách hàng 153 có thể được lấy ra từ phương tiện ghi, hoặc một giao diện thông qua đó dữ liệu có thể được lấy ra từ các hệ thống nội bộ bằng phương tiện truyền thông.

Fig.3A đến Fig.3D thể hiện các ví dụ về dữ liệu đặc trưng (các bản ghi) được cung cấp trong các cơ sở dữ liệu liên quan đến khách hàng 153. Theo ví dụ được thể hiện trên Fig.3A, dữ liệu đặc trưng có các tham số đại diện cho giới tính và tuổi của các khách hàng như các thuộc tính nhân khẩu của các khách hàng. Hơn nữa, dữ liệu đặc trưng có các tham số chỉ ra địa điểm mua hàng, mã JAN (Số sản phẩm Nhật Bản), số lượng mua, và đơn giá hàng hóa làm các tham số liên quan đến các hàng hóa khách hàng đã mua. Mã JAN là các mã hàng được sử dụng rộng rãi tại Nhật Bản. Các cơ sở dữ liệu liên quan đến khách hàng 153 bao gồm thông tin chi tiết về các hàng hóa đã mua như thể hiện trên Fig.3A được giữ bởi các nhà bán lẻ như các cửa hàng tiện lợi và các siêu thị.

Theo ví dụ khác thể hiện trên Fig.3B, dữ liệu đặc trưng có các tham số đại diện cho giới tính và tuổi của các khách hàng làm các thuộc tính nhân khẩu của các khách

hàng. Ngoài ra, dữ liệu đặc trưng có các tham số chỉ ra địa điểm mua và giá mua hàng hóa làm các tham số liên quan đến các hàng hóa khách hàng đã mua. Các cơ sở dữ liệu liên quan đến khách hàng 153 bao gồm dữ liệu đặc trưng như thể hiện trên Fig.3B, ví dụ, được giữ bởi các công ty thẻ tín dụng. Các cơ sở dữ liệu liên quan đến khách hàng 153 giữ bởi các công ty thẻ tín dụng là tốt hơn các cơ sở dữ liệu liên quan đến khách hàng 153 được giữ bởi các nhà bán lẻ ở chỗ họ có dữ liệu mua sắm không lờ khi so sánh với các cơ sở dữ liệu được giữ bởi các nhà bán lẻ, nhưng bù cho các cơ sở dữ liệu liên quan đến khách hàng 153 được giữ bởi các nhà bán lẻ trong đó các hàng hóa đã mua cụ thể không nhận dạng được.

Theo một ví dụ khác thể hiện trên Fig.3C, dữ liệu đặc trưng có các tham số đại diện cho giới tính, tuổi và nghề nghiệp của các khách hàng như các thuộc tính nhân khẩu của các khách hàng. Ngoài ra, dữ liệu đặc trưng có các tham số chỉ ra cách sử dụng phần lớn phương tiện truyền thông đại chúng (báo/ tạp chí/ máy thu thanh/ Tivi/ phim ảnh v.v) làm các tham số liên quan đến trạng thái giao tiếp với phương tiện truyền thông đại chúng của các khách hàng. Ví dụ, một trong số các tham số đại diện cho cách sử dụng chỉ ra một giá trị bằng '1' khi một khách hàng sử dụng (đăng ký/ xem, v.v) phương tiện truyền thông đại chúng tương ứng, và chỉ ra một giá trị bằng '0' khi không sử dụng. Loại thông tin này được thu thập, ví dụ, bằng bản câu hỏi hoặc theo cách tương tác, và được lưu trong các cơ sở dữ liệu liên quan đến khách hàng 153 làm dữ liệu đặc trưng. Loại cơ sở dữ liệu liên quan đến khách hàng 153 này thường được giữ bởi các công ty liên quan đến truyền thông đại chúng và không được giữ bởi các nhà bán lẻ.

Theo một ví dụ khác thể hiện tại Fig.3D, dữ liệu đặc trưng có các tham số đại diện cho giới tính, tuổi, và nơi cư trú của các khách hàng làm các thuộc tính nhân khẩu của các khách hàng. Ngoài ra, dữ liệu đặc trưng bao gồm các tham số đại diện cho cách sử dụng các trang Internet làm các tham số liên quan đến trạng thái giao tiếp với các nội dung mạng của các khách hàng. Ví dụ, mỗi tham số đại diện cho cách sử dụng chỉ ra một giá trị bằng '1' khi các khách hàng đã sử dụng (truy cập) trang tương ứng, và chỉ ra một giá trị bằng '0' khi không sử dụng. Loại thông tin này, ví dụ, được lưu trong một máy chủ nhất định thông qua một chương trình tạo nhật ký truy cập được cài đặt trong các thiết bị thông tin đầu cuối của các khách hàng với sự chấp thuận

trước của các khách hàng. Các cơ sở dữ liệu liên quan đến khách hàng 153 như vậy thường được giữ bởi các công ty IT mà không phải là các nhà bán lẻ.

Bộ phận xuất 19 được cấu hình để truyền đi cơ sở dữ liệu sau xử lý 155 được tạo ra dựa vào cơ sở dữ liệu liên quan đến khách hàng đã mô tả trên đây từ thiết bị xử lý dữ liệu 10. Ví dụ, Bộ phận xuất 19 được cấu hình để có giao diện truyền thông qua đó cơ sở dữ liệu sau xử lý 155 có thể được cung cấp cho thiết bị kết hợp dữ liệu 50 bằng truyền thông, hoặc một giao diện qua đó cơ sở dữ liệu sau xử lý 155 có thể đưa ra phương tiện ghi.

Bộ phận cụm 110 đọc một nhóm dữ liệu đặc trưng từ cơ sở dữ liệu liên quan đến khách hàng 153 trước khi xử lý lưu trong bộ phận lưu trữ 15, và phân cụm nhóm dữ liệu đặc trưng thành phần lớn dữ liệu đặc trưng tương tự nhau về đặc trưng. Với thuật phân cụm, các phương pháp đã biết như được mô tả trên đây có thể được sử dụng.

Một ví dụ về phương pháp phân cụm đơn giản có thể là một kỹ thuật sắp xếp dữ liệu đặc trưng trên không gian đặc trưng tương ứng làm các vectơ đặc trưng, và nhận dạng và phân cụm dữ liệu đặc trưng tương tự về đặc trưng dựa vào sự tương tự cosin giữa các vectơ đặc trưng.

Cũng có thể có một ma trận rút gọn chiều vectơ đặc trưng trên mỗi vectơ đặc trưng, sao cho một nhóm các vectơ đặc trưng tương tự về đặc trưng của khách hàng được phân tách tốt ra khỏi các vectơ đặc trưng khác trên không gian đặc trưng. Bằng cách điều chỉnh ma trận rút gọn chiều, có thể đạt được việc phân cụm phù hợp hơn.

Kỹ thuật phân cụm, giống như kỹ thuật pLSA (phân tích ngữ nghĩa tiềm ẩn theo xác suất), cũng được biết đến trong đó một dữ liệu đặc trưng được phân loại thành nhiều phân cụm theo phương pháp xác suất. Ví dụ, 80% của một dữ liệu đặc trưng có thể được phân loại thành phân cụm thứ nhất, và 20% còn lại được phân loại thành phân cụm thứ hai.

Cách xác định sự tương tự về các đặc trưng và cách để phân cụm nhóm dữ liệu đặc trưng có thể được xác định bởi người vận hành thiết bị kết hợp dữ liệu 50, xét đến mục đích phân tích dữ liệu dựa vào cơ sở dữ liệu kết hợp 557. Người vận hành thiết bị kết hợp dữ liệu 50 có thể cung cấp một chương trình điều hành bộ phận tính toán 11 như bộ phận phân cụm 110 và một bộ phận xử lý 115 đến các nhà cung cấp dữ liệu, và

có được từ nhà cung cấp dữ liệu các cơ sở dữ liệu sau xử lý 155 phù hợp với mục đích phân tích dữ liệu.

Theo một khía cạnh, hệ thống xử lý thông tin 1 có thể tạo ra các cơ sở dữ liệu sau xử lý 155 trong đó thông tin cá nhân được bảo vệ nhờ việc xử lý thống kê các cơ sở dữ liệu liên quan đến khách hàng 153 được giữ bởi các nhà cung cấp dữ liệu. Trong trường hợp này, tốt hơn là việc phân cụm được thực hiện sao cho mỗi phân cụm chứa dữ liệu đặc trưng nhiều hơn lượng thông tin cá nhân có thể bảo vệ bởi vì kết quả của việc xử lý thống kê (sự tích hợp). Nếu một phân cụm được tạo ra mà không có dữ liệu đặc trưng nhiều hơn lượng thông tin cá nhân có thể bảo vệ, có thể được nhận thức được để bỏ dữ liệu đặc trưng thuộc phân cụm đó để dữ liệu không được xử lý trong bộ phận xử lý 115.

Bộ phận xử lý 115 thực hiện xử lý thống kê để tích hợp các dữ liệu đặc trưng thuộc cùng một phân cụm cho phần lớn dữ liệu đặc trưng được phân cụm bởi bộ phận phân cụm 110, nói cách khác, cho mỗi phân cụm. Cụ thể là, bộ phận xử lý 115 tính toán thống kê cho mỗi tham số của phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm cho mỗi cụm, và tạo ra dữ liệu đặc trưng bao gồm các thống kê này các giá trị tham số, làm dữ liệu đặc trưng của một khách hàng ảo tương ứng với phân cụm. Bộ phận xử lý 115 tạo ra dữ liệu đặc trưng của một khách hàng ảo cho mỗi phân cụm, nhờ đó tạo ra các cơ sở dữ liệu sau xử lý 155 có dữ liệu đặc trưng cho tất cả các phân cụm, và ghi các cơ sở dữ liệu 155 trong bộ phận lưu trữ 15.

Theo ví dụ thể hiện trên Fig.4, bộ phận xử lý 115 tính toán tỷ lệ giới tính được chỉ ra bởi phần lớn dữ liệu đặc trưng trong một phân cụm như một thống kê về một tham số giới tính, và ghi tỷ lệ dữ liệu đặc trưng của một khách hàng ảo tương ứng với phân cụm như một giá trị tham số giới tính. Bộ phận xử lý 115 tính độ tuổi trung bình đại diện cho các dữ liệu đặc trưng trong phân cụm như một thống kê về tham số tuổi, và ghi giá trị trung bình như một giá trị tham số về tuổi trong dữ liệu đặc trưng của khách hàng ảo tương ứng với phân cụm.

Theo ví dụ thể hiện tại phần trên ở Fig.4, mỗi dữ liệu đặc trưng trước khi xử lý có một tham số mua sắm đại diện cho toàn bộ lượng tiền mà một khách hàng đã thanh toán cho hàng hóa có mã JAN tương ứng tại địa điểm mua sắm tương ứng, đối với mỗi sự kết hợp của mã JAN và địa điểm mua sắm. Trong trường hợp này, bộ phận xử lý

115 có thể tính một giá trị trung bình của tổng giá trị đại diện bởi phần lớn dữ liệu đặc trưng trong một phân cụm đối với mỗi sự kết hợp của mã JAN và địa điểm mua sắm như một thống kê liên quan đến tham số mua sắm, và ghi giá trị trung bình trong dữ liệu đặc trưng của một khách hàng ảo tương ứng với phân cụm.

Mỗi dữ liệu đặc trưng sau khi xử lý được thể hiện tại phần dưới ở Fig.4 đại diện cho giá mua sắm (giá trị trung bình) các hàng hóa được thanh toán bởi một khách hàng ảo đối với mỗi sự kết hợp của mã JAN và địa điểm mua sắm, cùng với các thuộc tính nhân khẩu (giới tính và tuổi) của khách hàng ảo.

Theo một ví dụ khác, bộ phận xử lý 115 có thể tính một giá trị lớn nhất và một giá trị nhỏ nhất của tổng giá trị như một giá trị đại diện cho sự phân bố giá mua sắm, thay vì giá trị trung bình của tổng giá trị. Đó là, bộ phận xử lý 115 đối với mỗi sự kết hợp của mã JAN và địa điểm mua sắm, có thể ghi giá trị lớn nhất và giá trị nhỏ nhất của tổng giá trị, làm các giá trị thống kê liên quan đến tham số mua sắm, trong dữ liệu đặc trưng của khách hàng ảo.

Theo Fig.4, dữ liệu đặc trưng trước và sau khi xử lý về cơ bản là giống nhau, ngoại trừ dữ liệu đặc trưng sau khi xử lý đại diện cho các số liệu thống kê, và có một tham số chỉ ra số lượng dữ liệu đặc trưng trước khi xử lý (số lượng mẫu) được bổ sung vào dữ liệu đặc trưng sau khi xử lý.

Tuy nhiên, bộ phận xử lý 115 có thể cũng tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi khách hàng ảo (phân cụm) như thể hiện trong phần dưới của Fig.4 từ các cơ sở dữ liệu liên quan đến khách hàng 153 có dữ liệu đặc trưng như thể hiện trên Fig.3A. Đó là, các loại tham số của dữ liệu đặc trưng không cần trùng khớp trước và sau khi xử lý. Một hoặc phần lớn tham số trước khi xử lý có thể được thay thế trong khi xử lý bằng các loại tham số khác mà có thể được tính từ đó.

Một ví dụ được mô tả trên đây trong đó tỷ lệ (phần trăm), giá trị trung bình, hoặc tập hợp các giá trị lớn nhất và nhỏ nhất đã được tính toán thống kê. Ngoài ra, số trung bình, hoặc tập hợp trung bình và số chênh lệch có thể được tính như một thống kê. Loại số liệu thống kê để tính có thể được xác định theo mục đích phân tích dữ liệu, sao cho thông tin quan trọng đối với việc phân tích không bị mất.

Trong thiết bị xử lý dữ liệu 10, một nhóm dữ liệu đặc trưng thuộc các phân cụm được xử lý (được tích hợp) thành dữ liệu đặc trưng đại diện cho các số thống kê cho

mỗi phân cụm theo cách này, và cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng sau khi xử lý được đưa ra đến thiết bị kết hợp dữ liệu 50.

Bây giờ, một cấu hình của thiết bị kết hợp dữ liệu 50 sẽ được mô tả với việc tham chiếu đến Fig.5A và Fig.5B. Thiết bị kết hợp dữ liệu 50 bao gồm, như thể hiện trên Fig.5A, một bộ phận tính toán 51, một bộ phận lưu trữ 55, một bộ phận nhập 57, và một bộ phận xuất 59. Bộ phận tính toán 51 gồm một CPU 53 thực hiện việc xử lý theo các chương trình 551 khác nhau. Bộ phận tính toán 51 bao gồm thêm một RAM 54 được dùng làm bộ nhớ làm việc khi thực hiện việc xử lý bởi CPU 53. Bộ phận tính toán 51 có chức năng như một bộ xử lý tổng hợp dữ liệu 510 như thể hiện trên Fig.5B bằng cách thực hiện việc xử lý trong CPU 53.

Bộ phận lưu trữ 55 lưu các chương trình 551 551 gồm một chương trình để điều khiển bộ phận tính toán 51 làm bộ xử lý tổng hợp dữ liệu 510. Bộ phận lưu trữ 55 cũng được cấu hình để lưu các cơ sở dữ liệu sau xử lý được cung cấp từ các thiết bị xử lý dữ liệu 10, và cơ sở dữ liệu kết hợp 557 được tạo ra bằng cách kết hợp các cơ sở dữ liệu sau xử lý 155. Các cơ sở dữ liệu sau xử lý 155 cung cấp từ các thiết bị xử lý dữ liệu 10 được đưa vào các thiết bị kết hợp dữ liệu 50 thông qua bộ phận nhập 57. Bộ phận nhập 57 có thể là một giao diện qua đó các cơ sở dữ liệu sau xử lý 155 có thể được đưa vào vào từ phương tiện ghi, hoặc có thể là một giao diện qua đó các cơ sở dữ liệu sau xử lý 155 được chuyển từ các thiết bị xử lý dữ liệu 10 có thể được đưa vào bằng truyền thông. Bộ phận xuất 59 được sử dụng để truyền, ví dụ, cơ sở dữ liệu kết hợp 557 và/hoặc các kết quả phân tích dữ liệu dựa vào cơ sở dữ liệu kết hợp 557 ra bên ngoài thiết bị kết hợp dữ liệu.

Bộ xử lý tổng hợp dữ liệu 510 kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 lưu trong bộ phận lưu trữ 55 để tạo ra cơ sở dữ liệu kết hợp 557. Cụ thể là, bộ xử lý tổng hợp dữ liệu 510, đề cập đến các tham số trong dữ liệu đặc trưng đóng vai trò làm các giới hạn, kết hợp các cơ sở dữ liệu sau xử lý 155 được kết hợp, để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện cho các tham số trong số các cơ sở dữ liệu sau xử lý 155 được kết hợp.

Như mô tả trên đây, mỗi dữ liệu đặc trưng cung cấp trong các cơ sở dữ liệu sau xử lý 155, thậm chí nếu khác nhau về các nhà cung cấp dữ liệu, cũng có các tham số đại diện cho các thuộc tính nhân khẩu chung của các khách hàng. Ví dụ, dữ liệu đặc

trung có chung một tham số đại diện cho giới tính và một tham số đại diện cho tuổi thậm chí trong dữ liệu đặc trưng khác nhau về các nhà cung cấp dữ liệu.

Vì vậy, bộ xử lý tổng hợp dữ liệu 510, tham chiếu đến các tham số được giữ bởi dữ liệu đặc trưng chung của các cơ sở dữ liệu sau xử lý 155 được kết hợp (dưới đây, được diễn đạt như các tham số chung), kết hợp các cơ sở dữ liệu sau xử lý 155, để kết hợp dữ liệu đặc trưng tương tự về đặc trưng của khách hàng đại diện cho các tham số chung.

Việc kết hợp các cơ sở dữ liệu sau xử lý 155, như thể hiện trên Fig.6A, có thể được thực hiện bằng cách kết hợp lần lượt các cơ sở dữ liệu sau xử lý 155, như bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 thứ nhất lưu trong bộ phận lưu trữ 55 với cơ sở dữ liệu sau xử lý 155 thứ hai, cơ sở dữ liệu sau xử lý 155 thứ hai với cơ sở dữ liệu sau xử lý 155 thứ ba, và kết hợp cơ sở dữ liệu sau xử lý 155 thứ m với cơ sở dữ liệu sau xử lý 155 thứ (m+1) (m là một số nguyên bằng 3 hoặc lớn hơn). Theo một ví dụ khác, bộ xử lý tổng hợp dữ liệu 510 có thể kết hợp các cơ sở dữ liệu sau xử lý 155 lưu trong bộ phận lưu trữ 55 theo sự bố trí hình sao quanh một cơ sở dữ liệu sau xử lý 155 cụ thể, bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 cụ thể với phần còn lại của các cơ sở dữ liệu xử lý 155. (Xem Fig.8A)

Các kỹ thuật hợp nhất dữ liệu khác cũng được biết đến. Cũng như trong bộ xử lý tổng hợp dữ liệu 510 của phương án thực hiện hiện tại, các cơ sở dữ liệu sau xử lý 155 có thể được kết hợp sử dụng một kỹ thuật hợp nhất dữ liệu đã biết. Theo một phương pháp hợp nhất dữ liệu đơn giản, việc khớp (kết hợp) dữ liệu đặc trưng tương tự giữa hai cơ sở dữ liệu sau xử lý 155 có thể được thực hiện theo cách dưới đây.

Ví dụ, đối với các tham số chung để đánh giá độ tương tự (ví dụ, tuổi và giới tính), khoảng cách (ví dụ khoảng cách cosin) giữa các vector đặc trưng có các tham số chung này làm các yếu tố khi được bố trí trên một khoảng không đặc trưng được tính cho tất cả các sự kết hợp của dữ liệu đặc trưng giữa các cơ sở dữ liệu sau xử lý 155. Bằng cách khớp các vector đặc trưng với một khoảng cách ngắn nhất giữa chúng, các cơ sở dữ liệu sau xử lý 155 có thể được kết hợp để kết hợp dữ liệu đặc trưng tương tự về đặc trưng của khách hàng được đại diện bởi các tham số chung.

Khi đánh giá mức tương tự giữa hai dữ liệu đặc trưng theo một khoảng cách trên không gian đặc trưng, một giải pháp cho vấn đề vận chuyển có thể được sử dụng để

thực hiện khớp dữ liệu đặc trưng giữa các cơ sở dữ liệu sau xử lý 155 sao cho chi phí vận chuyển là “nhỏ nhất có thể”.

Cơ sở dữ liệu kết hợp 557 tạo ra bằng cách khớp như vậy có thể được cấu hình để có dữ liệu đã lưu liên kết dữ liệu đặc trưng giữa hai cơ sở dữ liệu sau xử lý 155.

Fig.6C thể hiện một cấu hình của cơ sở dữ liệu kết hợp 557 trong trường hợp cơ sở dữ liệu sau xử lý 155A thể hiện trên Fig.6B với dữ liệu đặc trưng có các tham số về số lượng mẫu, giới tính, tuổi, mua P1, P2,..., mua Q1, Q2,... được kết hợp với một cơ sở dữ liệu sau xử lý 155 được thể hiện trên Fig.6B với dữ liệu đặc trưng có các tham số về số lượng của các mẫu, giới tính, tuổi, mua R1, R2,..., mua Q1, Q2,...

Theo các cơ sở dữ liệu sau xử lý 155A, 155B, các tham số chung là giới tính, tuổi, và mua Q1, Q2,..., Mua P1, P2,..., mua Q1, Q2,..., và mua R1, R2,... là, ví dụ, các tham số mua đối với các sự kết hợp tương ứng của mã JAN và địa điểm mua. Ví dụ, mỗi tham số mua đại diện cho giá mua hoặc mua hoặc không mua của các khách hàng cho sự kết hợp tương ứng.

Ví dụ, Mua P_i (i là một số nguyên bằng 2 hoặc lớn hơn) có thể là một tham số đại diện cho giá mua hoặc mua hoặc không mua bởi khách hàng tại một địa điểm mua khác với Mua P₁, của hàng hóa có cùng mã JAN như mã JAN trong Mua P₁. Mua Q_j, R_j (j là một số nguyên bằng 1 hoặc lớn hơn) có thể là các tham số đại diện cho giá mua hoặc mua hoặc không mua bởi các khách hàng tại một địa điểm mua tương ứng với một tham số j, hàng hóa có mã JAN khác nhau với mã JAN trong Mua P_j.

Trong cơ sở dữ liệu kết hợp 557 có được bằng cách kết hợp các cơ sở dữ liệu sau xử lý 155A, 155B được cấu hình như vậy, dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155A và dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155B kết hợp với nhau được trình bày kết hợp với các số định danh của chúng. Đó là, trong cơ sở dữ liệu kết hợp 557, số định danh dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155B kết hợp với dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155A được ghi kết hợp số định danh dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155A.

Như được thể hiện trên Fig.6C, có thể có một trường hợp trong đó phần lớn dữ liệu đặc trưng trong cơ sở dữ liệu sau xử lý 155B (B20056, B00234, B01123) được kết hợp với một dữ liệu đặc trưng trong cơ sở dữ liệu sau xử lý 155A (A00001).

Mỗi dữ liệu đặc trưng trong các cơ sở dữ liệu sau xử lý 155 tương ứng với phần lớn dữ liệu đặc trưng trước khi xử lý, và số lượng mẫu trong một phân cụm tương ứng với phần lớn dữ liệu đặc trưng trước khi xử lý có thể khác so với các phân cụm khác. Trong trường hợp này, giả định rằng mỗi dữ liệu đặc trưng có trong lượng tương ứng với số lượng mẫu, dữ liệu đặc trưng tương tự nhất về đặc trưng của khách hàng có thể được kết hợp bởi cùng số lượng mẫu, và dữ liệu đặc trưng của số lượng mẫu còn lại sau đó có thể được kết hợp với hầu hết dữ liệu đặc trưng thứ hai bởi cùng số lượng mẫu.

Trên Fig.6C, một bản ghi thể hiện "A00001, B20056, 5.3" chỉ ra rằng dữ liệu đặc trưng có một số định danh A00001 của cơ sở dữ liệu sau xử lý 155A và dữ liệu đặc trưng có một số định danh B20056 của cơ sở dữ liệu sau xử lý 155B được kết hợp với nhau bởi số lượng mẫu 5.3. Tương tự, bản ghi tiếp theo chỉ ra rằng dữ liệu đặc trưng có số định danh A00001 của cơ sở dữ liệu sau xử lý 155A và dữ liệu đặc trưng có số định danh B00234 của cơ sở dữ liệu sau xử lý 155B được kết hợp với nhau bởi số lượng mẫu 8.2. Hơn nữa, bản ghi tiếp theo chỉ ra rằng dữ liệu đặc trưng có số định danh A00001 của cơ sở dữ liệu sau xử lý 155A và dữ liệu đặc trưng có số định danh B01123 của cơ sở dữ liệu sau xử lý 155B được kết hợp với nhau bởi số lượng mẫu 6.5. Dữ liệu đặc trưng có số định danh A00001 của cơ sở dữ liệu sau xử lý 155A được tạo ra dựa vào một phân cụm có số lượng mẫu 20. Dữ liệu đặc trưng có số định danh A00001 của cơ sở dữ liệu sau xử lý 155A được phân bổ và được kết hợp với dữ liệu đặc trưng có số định danh B20056, số định danh B00234, và số định danh B01123 của cơ sở dữ liệu sau xử lý 155B.

Trong cơ sở dữ liệu kết hợp 557 được cấu hình như vậy, các cơ sở dữ liệu sau xử lý 155 được kết hợp với nhau, và cơ sở dữ liệu kết hợp 557 và các cơ sở dữ liệu sau xử lý 155 đã xây dựng lên một cơ sở dữ liệu lớn. Theo ví dụ thể hiện trên Fig.6B và Fig.6C, cơ sở dữ liệu sau xử lý 155A không chứa thông tin về Mua R1, R2,... và cơ sở dữ liệu sau xử lý 155B không chứa thông tin về Mua P1, P2....được kết hợp để xây dựng lên một cơ sở dữ liệu lớn trong đó thông tin Mua P1, P2,..., Mua Q1, Q2,..., và Mua R1, R2,... được liên kết. Vì vậy, có thể phân tích, ví dụ, loại xu hướng các khách hàng mua Mua P1, P2,..., ... hướng về Mua R1, R2,

FIG. 7 là sơ đồ trong đó một khối cho mỗi sự kết hợp của mã JAN và địa điểm mua được bố trí theo chiều ngang, và một dãy khối được sắp xếp cho mỗi khách hàng

ảo theo chiều dọc. Các khối gạch chéo chỉ ra rằng khách hàng ảo tương ứng đã mua hàng hóa tương ứng với mã JAN tại địa điểm mua tương ứng. Các khối không gạch chéo chỉ ra rằng khách hàng ảo tương ứng không mua hàng hóa tương ứng với mã JAN tại địa điểm mua tương ứng.

Theo Fig.7, các khối gạch đứt được tập trung tại một khu vực thể hiện bởi đường nét đứt. Sự tập trung này thể hiện rằng có các người mua cho sự kết hợp mã JAN và địa điểm mua. Do đó, khi quảng cáo hàng hóa tương ứng được đưa cho các khách hàng tương ứng tới các khối không gạch chéo trong khu vực này, hiệu quả của việc quảng cáo được chứng minh đáng kể. Ví dụ, các khách hàng ảo quây trong đường đậm hầu như không thể hiện hành vi mua sắm trong khu vực được thể hiện trong đường nét đứt, nhưng tương ứng với khách hàng mua dựa vào số liệu thống kê. Do đó, khi quảng cáo hàng hóa tương ứng được đưa đến các khách hàng tương ứng, hiệu quả quảng cáo được mong đợi.

Nếu việc phân tích dữ liệu như thể được thực hiện dựa vào dữ liệu thông thường về các khách hàng, có một khả năng rằng thông tin về sự riêng tư của các khách hàng được xác định chi tiết trong quá trình phân tích. Ngược lại, theo phương án thực hiện này, dữ liệu của mỗi dòng thể hiện trên Fig.6B và Fig.7 là dữ liệu về các khách hàng ảo được tạo ra thống kê. Theo phương án thực hiện này, vì việc phân tích dữ liệu đã đề cập trên đây được thực hiện dựa vào dữ liệu đã tạo ra theo cách thống kê, có thể có được thông tin hữu ích dùng cho các hoạt động thương mại bằng cách phân tích dữ liệu trong khi loại bỏ được các nhược điểm trên đây trong lĩnh vực kỹ thuật đã biết.

Hệ thống xử lý thông tin 1 theo phương án thực hiện này đã được mô tả trên đây. Theo phương án thực hiện này, thiết bị xử lý dữ liệu 10 xử lý các cơ sở dữ liệu liên quan đến khách hàng 153, và cung cấp các cơ sở dữ liệu sau xử lý 155 cho thiết bị kết hợp dữ liệu 50. Sau đó, thiết bị kết hợp dữ liệu 50 kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 tạo ra bởi thiết bị xử lý dữ liệu 10.

Dựa vào một nhóm dữ liệu đặc trưng của các cơ sở dữ liệu liên quan đến khách hàng 153 trước khi xử lý, bộ phận tính toán 11 của thiết bị xử lý dữ liệu 10 phân cụm nhóm dữ liệu đặc trưng thành phần lớn dữ liệu đặc trưng tương tự về đặc trưng. Hơn nữa, bộ phận tính toán 11 xử lý thống kê phần lớn dữ liệu đặc trưng tương ứng với mỗi phân cụm và tích hợp phần lớn dữ liệu đặc trưng cho mỗi phân cụm. Theo đó, dữ liệu

đặc trưng của mỗi khách hàng ảo tương ứng với phần lớn khách hàng tương tự về đặc trưng được tạo ra. Đặc biệt, bộ phận tính toán 11 (bộ phận xử lý 115) tính một số liệu thống kê cho mỗi tham số của phần lớn dữ liệu đặc trưng thuộc một phân cụm đối với mỗi cụm. Dữ liệu đặc trưng của mỗi khách hàng ảo được tạo ra gồm số thống kê làm giá trị tham số. Bộ phận tính toán 11 (bộ phận xử lý 115) tạo ra một cơ sở dữ liệu gồm dữ liệu đặc trưng của mỗi khách hàng ảo trong đó phần lớn dữ liệu đặc trưng được tích hợp như vậy, làm cơ sở dữ liệu sau xử lý 155.

Mỗi dữ liệu đặc trưng trong cơ sở dữ liệu sau xử lý 155 bao gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp các cơ sở dữ liệu sau xử lý 155. Dữ liệu tham khảo đại diện cho một loại đặc trưng chung mà cơ sở dữ liệu sau xử lý 155 được kết hợp. Dữ liệu tham khảo theo phương án thực hiện trên đây là giới tính và tuổi của các khách hàng làm các thuộc tính nhân khẩu của các khách hàng.

Bộ phận tính toán 51 của thiết bị kết hợp dữ liệu 50 kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu trong số cơ sở dữ liệu sau xử lý 155.

Đó là, theo hệ thống xử lý thông tin 1, thiết bị xử lý dữ liệu 10 xử lý thống kê và tích hợp dữ liệu đặc trưng của phần lớn khách hàng, và tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi khách hàng ảo. Ví dụ, nếu dữ liệu đặc trưng trước khi xử lý thống kê là dữ liệu của một khách hàng riêng lẻ đại diện cho các đặc trưng của một khách hàng, hoạt động của thiết bị xử lý dữ liệu 10 chuyển đổi dữ liệu đặc trưng từ dữ liệu một khách hàng riêng lẻ thành dữ liệu khách hàng ảo trong đó các đặc trưng của phần lớn khách hàng đã được trộn.

Vì vậy, các nhà cung cấp dữ liệu có thể chuyển đổi các cơ sở dữ liệu liên quan đến khách hàng 153 thành các cơ sở dữ liệu sau xử lý 155 trong đó thông tin cá nhân của các khách hàng được bảo vệ, sử dụng thiết bị xử lý dữ liệu 10, và cung cấp dữ liệu về các khách hàng cho thiết bị kết hợp dữ liệu 50 trong khi thông tin cá nhân được bảo vệ.

Theo cách đó, người điều khiển thiết bị kết hợp dữ liệu 50 có thể có được các cơ sở dữ liệu liên quan đến khách hàng từ các công ty không sẵn lòng cung cấp các cơ sở dữ liệu liên quan đến khách hàng 153 từ quan điểm bảo vệ thông tin, làm các cơ sở dữ

liệu sau xử lý 155. Thiết bị kết hợp dữ liệu 50 có thể tạo ra một cơ sở dữ liệu kết hợp dữ liệu của các khách hàng khác nhau từ các nguồn dữ liệu khác nhau.

Bởi vậy, hệ thống xử lý thông tin 1 có thể xây dựng, như cơ sở dữ liệu kết hợp 557, một cơ sở dữ liệu tích hợp dữ liệu về các khách hàng khác nhau phân tán trong xã hội. Hệ thống xử lý thông tin 1 có thể sử dụng một cách hiệu quả dữ liệu về các khách hàng khác nhau, và thực hiện phân tích dữ liệu đáng kể, dựa vào cơ sở dữ liệu.

Các tham số chung trong số các cơ sở dữ liệu sau xử lý 155 (dữ liệu tham chiếu khi kết hợp các cơ sở dữ liệu) trong phương án thực hiện mô tả trên đây là giới tính và tuổi của các khách hàng. Tuy nhiên, các tham số chung cũng có thể bao gồm các tham số khác. Ví dụ, các tham số chung có thể bao gồm ít nhất một trong số tham số về giới tính, tuổi, nghề nghiệp, nơi cư trú, thu nhập, trình độ học vấn, và thành phần gia đình của khách hàng, làm các thuộc tính nhân khẩu học của các khách hàng.

[Phương án thực hiện thứ hai]

Hệ thống xử lý thông tin 1 theo phương án thực hiện thứ hai được cấu hình để sử dụng một cơ sở dữ liệu bao gồm một dữ liệu nguồn đơn làm cơ sở dữ liệu trung tâm (hub) H (xem Fig.8A), trong thiết bị kết hợp dữ liệu 50. Hệ thống xử lý thông tin 1 kết hợp từng cơ sở dữ liệu sau xử lý 155 được cung cấp từ thiết bị xử lý dữ liệu 10 với cơ sở dữ liệu hub H, theo cách đó kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 thông qua cơ sở dữ liệu hub H.

Hệ thống xử lý thông tin 1 theo phương án thực hiện này có cùng cấu hình phần cứng như ở phương án thực hiện thứ nhất. Hệ thống xử lý thông tin 1 của phương án thực hiện này khác với phương án thực hiện thứ nhất ở chỗ bộ phận lưu trữ 55 của thiết bị kết hợp dữ liệu 50 có cơ sở dữ liệu hub H, ngoài các cơ sở dữ liệu sau xử lý 155 được cung cấp từ thiết bị xử lý dữ liệu 10. Ngoài khía cạnh này, hệ thống xử lý thông tin 1 theo phương án thực hiện này chỉ khác với hệ thống xử lý thông tin của phương án thực hiện thứ nhất trong các hoạt động xử lý của bộ xử lý tổng hợp dữ liệu 510. Do đó, dưới đây, theo yêu cầu, việc mô tả cấu hình tương tự với phương án thực hiện thứ nhất sẽ được bỏ qua.

Dữ liệu nguồn đơn hình thành nên cơ sở dữ liệu hub H có được, ví dụ, từ một điều tra người tiêu dùng “HABIT (nhãn hiệu đã đăng ký)” được thực hiện bởi Hakuhodo Co., Ltd. HABIT thu các thông tin khác nhau về phong cách sống (ý

thức và thái độ sống), nhận thức thấp về giá trị và thương hiệu từ mỗi cộng tác viên điều tra, cùng với thông tin về các thuộc tính nhân khẩu của các cộng tác viên, bằng cách thực hiện phỏng vấn và các bảng câu hỏi. Dữ liệu nguồn đơn cho mỗi khách hàng (cộng tác viên) được tạo ra bởi HABIT đại diện cho đặc trưng của khách hàng không xuất hiện trong dữ liệu mua sắm và tương tự.

Các ví dụ khác về dữ liệu nguồn đơn có thể được cung cấp trong cơ sở dữ liệu hub H là dữ liệu nguồn đơn bao gồm dữ liệu mua của các khách hàng có được từ hệ thống ID-POS, và dữ liệu nguồn đơn đại diện cho trạng thái tiếp xúc với phương tiện thông tin đại chúng có được thông qua các bảng câu hỏi. Ngoài ra, cơ sở dữ liệu hub H có thể được cung cấp với dữ liệu trong đó phần lớn loại dữ liệu nguồn đơn này được kết hợp cho mỗi khách hàng.

Dữ liệu nguồn đơn được cung cấp trong cơ sở dữ liệu hub H tương ứng với dữ liệu đặc trưng đại diện cho đặc trưng của khách hàng, và có các tham số đại diện cho thuộc tính nhân khẩu học chung của các khách hàng làm các tham số chung cho các đặc trưng của các cơ sở dữ liệu sau xử lý 155.

Bộ xử lý tổng hợp dữ liệu 510, như thể hiện trên Fig.8A, kết hợp mỗi cơ sở dữ liệu sau xử lý 155 được cung cấp cơ sở dữ liệu hub H từ các thiết bị xử lý dữ liệu 10, theo cách đó kết hợp phần lớn cơ sở dữ liệu sau xử lý 155 qua cơ sở dữ liệu hub H. Phương pháp kết hợp đã được mô tả trên đây. Đó là, bộ xử lý tổng hợp dữ liệu 510 kết hợp cơ sở dữ liệu hub H với từng cơ sở dữ liệu sau xử lý 155 để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện cho các tham số chung.

Cơ sở dữ liệu kết hợp 557 đại diện cho, ví dụ, mối quan hệ giữa dữ liệu đặc trưng của cơ sở dữ liệu hub H và dữ liệu đặc trưng của mỗi cơ sở dữ liệu sau xử lý 155 được kết hợp, liên kết với các số định danh như trong phương án thực hiện trên đây. Trong cơ sở dữ liệu kết hợp 557 thể hiện trên Fig.8B, liên kết với một số định danh của dữ liệu đặc trưng ở cơ sở dữ liệu hub H (H00001), các số định danh của dữ liệu đặc trưng ở các cơ sở dữ liệu sau xử lý 155A, 155B (A12345, B00044) kết hợp với dữ liệu đặc trưng của cơ sở dữ liệu hub H được ghi. Số lượng mẫu thể hiện trên Fig.8B như được mô tả trên Fig.6C.

Dữ liệu nguồn đơn đại diện cho các đặc trưng khách hàng. Do đó, nếu cơ sở dữ liệu hub H và từng cơ sở dữ liệu sau xử lý 155 tương ứng được kết hợp dựa vào cơ sở

dữ liệu hub H có dữ liệu nguồn đơn làm dữ liệu đặc trưng đại diện cho đặc trưng của khách hàng, dữ liệu đặc trưng giữa các cơ sở dữ liệu sau xử lý 155 có thể liên kết thích hợp thêm.

Dữ liệu nguồn đơn như HABIT bao gồm các thông tin khác nhau đại diện cho đặc trưng của khách hàng. Khi cơ sở dữ liệu sau xử lý 155 tương ứng được cung cấp cơ sở dữ liệu hub H có dữ liệu nguồn đơn như vậy từ các thiết bị xử lý dữ liệu 10 được kết hợp, dữ liệu tham chiếu dùng để kết hợp có thể được thay đổi. Đó là, tùy thuộc vào các loại tham số đại diện cho các đặc trưng của các khách hàng có trong các cơ sở dữ liệu sau xử lý 155, dữ liệu tham chiếu dùng để kết hợp các cơ sở dữ liệu sau xử lý 155 và cơ sở dữ liệu hub H có thể được thay đổi.

Ví dụ, nếu các cơ sở dữ liệu sau xử lý 155 được cung cấp với dữ liệu đặc trưng có một tham số đại diện cho hành vi giao tiếp của khách hàng với nội dung mạng, các tham số chung giữa cơ sở dữ liệu hub H và các cơ sở dữ liệu sau xử lý 155 gồm tham số đại diện cho hành vi giao tiếp của khách hàng với các nội dung mạng, thêm vào hoặc thay cho các tham số đại diện cho thuộc tính nhân khẩu của các khách hàng.

Trong trường hợp này, cơ sở dữ liệu hub H và các cơ sở dữ liệu sau xử lý 155 có thể được kết hợp để kết hợp dữ liệu đặc trưng tương tự về hành vi giao tiếp của khách hàng với các nội dung mạng đại diện cho tham số chung.

[Phương án thực hiện thứ ba]

Hệ thống xử lý thông tin 1 theo phương án thực hiện thứ ba gồm phần lớn thiết bị xử lý dữ liệu 10 và thiết bị kết hợp dữ liệu 50. Các thiết bị xử lý dữ liệu 10 và thiết bị kết hợp dữ liệu 50 có cấu hình phần cứng giống cấu hình phần cứng theo phương án thực hiện thứ nhất và phương án thực hiện thứ hai. Dưới đây, các phần mô tả các cấu hình và các nội dung xử lý trong hệ thống xử lý thông tin 1 theo phương án thực hiện thứ ba mà tương tự với các phần mô tả các cấu hình và các nội dung xử lý theo phương án thực hiện thứ nhất và phương án thực hiện thứ hai sẽ được bỏ qua.

Hệ thống xử lý thông tin 1 theo phương án thực hiện này có các cơ sở dữ liệu liên quan đến khách hàng 1531L, 1531R như thể hiện trên Fig.9A, làm các cơ sở dữ liệu liên quan đến khách hàng 153. Tại khu vực bên trái trên Fig.9A, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1531L có dữ liệu mua sắm của các khách hàng có được từ một hệ thống POS, làm dữ liệu đặc trưng của các khách hàng. Tại khu

vực bên phải trên Fig.9A, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1531R có dữ liệu đặc trưng đại diện cho cách sử dụng các thiết bị được sử dụng bởi người dùng, làm dữ liệu đặc trưng về các khách hàng. Ví dụ về các thiết bị gồm một lò vi sóng.

Theo Fig.9A, cơ sở dữ liệu liên quan đến khách hàng 1531L gồm dữ liệu đặc trưng của mỗi lần thanh toán. Dữ liệu đặc trưng có các tham số đại diện cho giới tính, tuổi, và nơi cư trú của khách hàng, làm các thuộc tính nhân khẩu của các khách hàng. Ngoài ra, dữ liệu đặc trưng có các tham số là địa điểm nơi khách hàng mua các hàng hóa, mã sản phẩm, số lượng mua, và đơn giá. Mã sản phẩm gồm, ví dụ, mã JAN

Theo Fig.9A, cơ sở dữ liệu liên quan đến khách hàng 1531R gồm dữ liệu đặc trưng của mỗi lò vi sóng. Dữ liệu đặc trưng có các tham số đại diện cho giới tính và tuổi của các khách hàng người sở hữu một lò vi sóng, cũng như các tham số đại diện cho địa điểm lắp đặt lò vi sóng, và chức năng và công thức nấu ăn được sử dụng trong lò vi sóng của mỗi múi giờ.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1531L được cấu hình để phân cụm phần lớn dữ liệu đặc trưng tương tự, ví dụ, theo giới tính, tuổi và lịch sử mua trong cơ sở dữ liệu liên quan đến khách hàng 1531L. Thiết bị xử lý dữ liệu 10, tương tự với thiết bị xử lý dữ liệu của phương án thực hiện thứ nhất, được cấu hình để xử lý thống kê phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm cho mỗi phân cụm, theo cách đó tạo ra dữ liệu đặc trưng của một khách hàng ảo trong đó phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được tích hợp cho mỗi cụm.

Bằng cách này, thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng cho mỗi khách hàng ảo tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531L. Dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531L về cơ bản có cùng các tham số như các tham số của dữ liệu đặc trưng của cơ sở dữ liệu liên quan đến khách hàng 151L. Tuy nhiên, dữ liệu đặc trưng của cơ sở dữ liệu sau xử lý 155 thể hiện các số thống kê của phân cụm tương ứng, làm các giá trị tham số. ID của khách hàng có thể được loại bỏ tại thời điểm xử lý.

Một ví dụ khác, thiết bị xử lý dữ liệu 10 có thể được cấu hình để thu thập dữ liệu đặc trưng về bộ phận thanh toán giữ bởi cơ sở dữ liệu liên quan đến khách hàng 1531L

cho mỗi dữ liệu đặc trưng thể hiện ID khách hàng giống nhau, và sau đó chuyển đổi dữ liệu đặc trưng đã thu thành dữ liệu đặc trưng của các bộ phận riêng. Ngoài ra, thiết bị xử lý dữ liệu 10 có thể được cấu hình để phân cụm và xử lý thống kê phần lớn dữ liệu đặc trưng tương tự về lịch sử mua sắm, dựa vào tập hợp dữ liệu đặc trưng của các bộ phận riêng, theo cách đó tạo ra cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531L.

Thiết bị xử lý dữ liệu 10 khác có cơ sở dữ liệu liên quan đến khách hàng 1531R được cấu hình để phân cụm phần lớn dữ liệu đặc trưng tương tự, ví dụ, theo giới tính, tuổi, và chức năng và cách sử dụng các công thức nấu ăn, trong cơ sở dữ liệu liên quan đến khách hàng 1531R. Hơn nữa, thiết bị xử lý dữ liệu 10 được cấu hình để xử lý thống kê phần lớn dữ liệu đặc trưng thuộc cùng phân cụm cho mỗi phân cụm, để tạo ra dữ liệu đặc trưng của một thiết bị ảo trong đó phần lớn dữ liệu đặc trưng thuộc cùng phân cụm được tích hợp cho mỗi cụm. Thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 được cung cấp với dữ liệu đặc trưng của mỗi thiết bị ảo tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531R.

Thiết bị kết hợp dữ liệu 50 được cấu hình để tạo ra cơ sở dữ liệu kết hợp 557 bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531R và cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531L.

Thiết bị kết hợp dữ liệu 50 có thể kết hợp các cơ sở dữ liệu sau xử lý 155 để kết hợp, sử dụng các tham số đại diện cho giới tính và tuổi của các khách hàng và các tham số liên quan đến thực phẩm được giữ bởi các cơ sở dữ liệu sau xử lý 155 làm dữ liệu tham chiếu, để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện bởi dữ liệu đặc trưng trong số các cơ sở dữ liệu sau xử lý 155 được kết hợp.

Trong cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531L, “các tham số liên quan đến thực phẩm” tương ứng với các tham số đại diện cho mã sản phẩm và số lượng mua được giữ bởi dữ liệu đặc trưng. Trong cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1531R, “các tham số liên quan đến thực phẩm” tương ứng với các công thức nấu ăn sử dụng.

Theo cơ sở dữ liệu kết hợp 557 tạo ra như vậy, mối quan hệ tương ứng giữa giới tính, tuổi, thực phẩm đã mua, và công thức nấu ăn có thể được phân tích. Do đó, qua

một lò vi sóng, các công thức nấu ăn đề xuất sử dụng thực phẩm mà một người sử dụng thường mua có thể được giới thiệu cho người sử dụng. Hơn nữa, các phiếu thưởng thực phẩm liên quan đến các công thức nấu ăn sử dụng thường xuyên có thể được bổ sung vào biên lai đưa ra từ một máy tính tiền POS.

Ở trên, như phương án thực hiện thứ ba, một ví dụ về phân cụm và xử lý thống kê cả cơ sở dữ liệu liên quan đến khách hàng 1531L và 1531R cho mỗi dữ liệu đặc trưng tương tự về đặc trưng đã được mô tả. Tuy nhiên, một trong số các cơ sở dữ liệu liên quan đến khách hàng 1531L, 1531L có thể không được xử lý thống kê.

Đó là, cơ sở dữ liệu kết hợp 557 có thể được tạo ra bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 của cơ sở dữ liệu liên quan đến khách hàng 1531L, và cơ sở dữ liệu liên quan đến khách hàng 1531R, hoặc có thể được tạo ra bằng cách kết hợp cơ sở dữ liệu liên quan đến khách hàng 1531L, và cơ sở dữ liệu sau xử lý 155 của cơ sở dữ liệu liên quan đến khách hàng 1531R. Cơ sở dữ liệu kết hợp 557 sau, ví dụ, có thể được cung cấp đến nhà cung cấp dữ liệu của dữ liệu liên quan đến khách hàng 1531L

[Phương án thực hiện thứ tư]

Hệ thống xử lý thông tin 1 theo phương án thực hiện thứ tư, tương tự với phương án thực hiện thứ ba, về cơ bản là hệ thống trong đó chỉ cấu hình của cơ sở dữ liệu liên quan đến khách hàng 153 là khác với cấu hình theo phương án thực hiện thứ nhất và phương án thực hiện thứ hai. Thiết bị xử lý dữ liệu 10 và thiết bị kết hợp dữ liệu 50 cung cấp trong hệ thống xử lý thông tin 1 theo phương án thực hiện này có cùng cấu hình phần cứng tương tự với cấu hình phần cứng theo phương án thực hiện thứ nhất và phương án thực hiện thứ hai.

Trong phần tiếp theo, phần mô tả các cấu hình và các nội dung xử lý trong hệ thống xử lý thông tin 1 của phương án thực hiện thứ tư tương tự với phương án thực hiện thứ nhất và phương án thực hiện thứ hai sẽ được bỏ qua. Các cấu hình phần cứng và các nội dung xử lý của hệ thống xử lý thông tin 1 theo các phương án thực hiện từ phương án thực hiện thứ năm tới phương án thực hiện thứ tám mô tả sau đó về cơ bản tương tự với cấu hình phần cứng và các nội dung xử lý sau của phương án thực hiện thứ nhất và phương án thực hiện thứ hai.

Hệ thống xử lý thông tin 1 theo phương án thực hiện này có một cơ sở dữ liệu liên quan đến khách hàng 1532L, 1532R thể hiện trên Fig.9B làm cơ sở dữ liệu liên

quan đến khách hàng 153. Trong phần bên trái của Fig.9B, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1532L có dữ liệu mua của các khách hàng thu được từ một hệ thống POS, làm dữ liệu đặc trưng về các khách hàng. Trong phần bên phải trên Fig.9B, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1532R có dữ liệu đặc trưng là cách sử dụng các máy bán hàng tự động, làm dữ liệu đặc trưng về các khách hàng.

Trong những năm gần đây, máy bán đồ uống tự động được bố trí màn hình tinh thể lỏng và một máy ảnh đã được biết đến. Máy bán hàng tự động có một chức năng xác định giới tính và tuổi của người sử dụng máy bán hàng tự động, từ hình ảnh đã chụp bởi máy ảnh, và, dựa vào kết quả của việc xác định này, hiển thị một đồ uống đề xuất qua màn hình tinh thể lỏng. Cơ sở dữ liệu kết hợp 557 được mô tả dưới đây giúp xác định đồ uống đề xuất để hiển thị trong máy bán hàng tự động.

Theo Fig.9B, cơ sở dữ liệu liên quan đến khách hàng 1532L được cấu hình theo cách giống như cơ sở dữ liệu liên quan đến khách hàng 1531L theo phương án thực hiện thứ ba. Cơ sở dữ liệu liên quan đến khách hàng 1532L bao gồm dữ liệu đặc trưng đại diện cho lịch sử mua sắm đối với đồ uống.

Theo Fig.9B, cơ sở dữ liệu liên quan đến khách hàng 1532R gồm dữ liệu đặc trưng của mỗi đồ uống được bán trong mỗi máy bán hàng tự động. Mỗi dữ liệu đặc trưng tạo thành cơ sở dữ liệu liên quan đến khách hàng 1532R đại diện cho địa điểm lắp đặt máy và nhà sản xuất máy bán hàng tự động. Ngoài ra, dữ liệu đặc trưng đại diện cho giới tính và tuổi làm thuộc tính nhân khẩu của người mua, cũng như thời tiết và thời gian tại thời điểm bán đồ uống tương ứng. Ngoài ra, dữ liệu đặc trưng đại diện cho mã hàng hóa, số lượng mua, và đơn giá của các đồ uống mà người mua đã mua. Hơn nữa, dữ liệu đặc trưng đại diện cho thông tin hàng tồn kho tại thời điểm bán. Địa điểm lắp máy bán hàng tự động, ví dụ, đại diện cho vĩ độ và kinh độ ở nơi lắp đặt máy bán hàng tự động.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1532L, ví dụ, được cấu hình để phân cụm phần lớn dữ liệu đặc trưng tương tự về giới tính, tuổi và lịch sử mua sắm trong cơ sở dữ liệu liên quan đến khách hàng 1532L, và, như trong phương án thực hiện thứ ba, tạo ra cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1532L.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1532R, ví dụ, được cấu hình để phân cụm phần lớn dữ liệu đặc trưng tương tự về địa điểm lắp đặt, nhà sản xuất và thông tin hàng tồn kho, và tương tự trong cơ sở dữ liệu liên quan đến khách hàng 1532R, và sau đó, đối với mỗi cụm, xử lý thống kê phần lớn dữ liệu đặc trưng thuộc cùng phân cụm, theo cách đó tạo ra dữ liệu đặc trưng của một thiết bị ảo (máy bán hàng tự động) trong đó phần lớn dữ liệu đặc trưng thuộc cùng phân cụm được tích hợp cho mỗi phân cụm.

Trong cách này, thiết bị xử lý dữ liệu 10 được cấu hình để tạo ra cơ sở dữ liệu sau xử lý 155 bao gồm dữ liệu đặc trưng cho mỗi thiết bị ảo, tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1532R.

Thiết bị kết hợp dữ liệu 50 được cấu hình để tạo ra cơ sở dữ liệu kết hợp 557 bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1532L và cơ sở dữ liệu sau xử lý 155 mà tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1532R.

Ví dụ, thiết bị kết hợp dữ liệu 50 sử dụng các tham số đại diện cho giới tính và tuổi của các khách hàng cũng như tham số liên quan đến địa điểm làm dữ liệu tham chiếu kết hợp các cơ sở dữ liệu sau xử lý 155 để kết hợp để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện cho dữ liệu tham chiếu giữa các cơ sở dữ liệu sau xử lý 155 được kết hợp.

“Tham số liên quan đến địa điểm” được sử dụng trong tài liệu này tương ứng với một tham số đại diện cho địa điểm mua giữ bởi dữ liệu đặc trưng trong cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1532L. Trong cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1532R, “tham số liên quan đến địa điểm” tương ứng với một tham số đại diện cho địa điểm của máy bán hàng tự động.

Theo cơ sở dữ liệu kết hợp 557 được tạo ra như vậy, ví dụ, mối quan hệ tương ứng giữa hành vi mua sắm của người tiêu dùng trong các cửa hàng như các cửa hàng tiện lợi hoặc các siêu thị và hành vi mua sắm của người tiêu dùng tại các máy bán hàng tự động có thể được phân tích. Do đó, dựa vào cơ sở dữ liệu kết hợp 557 này, tại một máy bán hàng tự động, một đồ uống được mua tại các cửa hàng trong khu vực bởi những người mua tương tự về tuổi và giới tính của một người sử dụng đứng trước máy

bán hàng tự động có thể được đề xuất đến người sử dụng. Trong một cửa hàng, một phiếu thưởng của đồ uống đã mua tại các máy bán hàng tự động trong khu vực đó bởi những người dùng tương tự về tuổi và giới tính của một người mua có thể được bổ sung vào biên lai đưa ra từ máy tính tiền POS.

Tương tự với phương án thực hiện thứ ba, một cơ sở dữ liệu liên quan đến khách hàng 1532L và cơ sở dữ liệu liên quan đến khách hàng 1532R có thể không được xử lý thống kê. Nói cách khác, cơ sở dữ liệu kết hợp 557 có thể là cơ sở dữ liệu kết hợp cơ sở dữ liệu sau xử lý 155 của cơ sở dữ liệu liên quan đến khách hàng 1532L, và cơ sở dữ liệu liên quan đến khách hàng 1532R, hoặc có thể là cơ sở dữ liệu kết hợp cơ sở dữ liệu liên quan đến khách hàng 1532L, và cơ sở dữ liệu sau xử lý 155 của cơ sở dữ liệu liên quan đến khách hàng 1532R. Ý tưởng kỹ thuật này cũng có thể ứng dụng được với các phương án thực hiện từ thứ năm đến thứ tám được mô tả dưới đây.

[Phương án thực hiện thứ năm]

Dưới đây, phần mô tả các cấu hình và các nội dung xử lý trong hệ thống xử lý thông tin 1 theo phương án thực hiện thứ năm tương tự với cấu hình và nội dung xử lý của phương án thực hiện thứ nhất và phương án thực hiện thứ hai sẽ được bỏ qua. Hệ thống xử lý thông tin 1 theo phương án thực hiện này có các cơ sở dữ liệu liên quan đến khách hàng 1533L, 1533R thể hiện trên Fig.10A, làm cơ sở dữ liệu liên quan đến khách hàng 153.

Tại phần trái trên Fig.10A, thể hiện ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1533L có dữ liệu đặc trưng đại diện cho các đặc trưng của mỗi người sử dụng sở hữu một thiết bị đầu cuối di động điển hình bởi điện thoại thông minh, làm dữ liệu đặc trưng về khách hàng. Trong phần phải trên Fig.10A, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1533R có dữ liệu đặc trưng đại diện cho cách sử dụng của nhà ga, làm dữ liệu đặc trưng về các khách hàng.

Theo Fig.10A, cơ sở dữ liệu liên quan đến khách hàng 1533L có dữ liệu đặc trưng đối với mỗi thiết bị đầu cuối di động. Dữ liệu đặc trưng đại diện cho giới tính, tuổi và nơi cư trú làm các thuộc tính nhân khẩu của một người sử dụng sở hữu một thiết bị đầu cuối di động. Ngoài ra, dữ liệu đặc trưng đại diện cho lịch sử di động của người sử dụng, và thời tiết tại mỗi điểm di động. Hơn nữa, dữ liệu đặc trưng đại diện

cho cách sử dụng phiếu thưởng đã được sử dụng bởi người sử dụng qua một chương trình ứng dụng được cài trong thiết bị đầu cuối di động.

Cơ sở dữ liệu liên quan đến khách hàng 1533R có dữ liệu đặc trưng cho mỗi kết hợp của nhà ga, cửa bán vé, thời tiết, và múi giờ. Dữ liệu đặc trưng tạo thành cơ sở dữ liệu liên quan đến khách hàng 1533R đại diện cho giới tính và tuổi của khách đi qua cửa một cửa bán vé, và múi giờ và thời tiết tại thời điểm đi qua, làm dữ liệu giao thông của nhà ga và cửa bán vé tương ứng. Giới tính và tuổi của khách đi qua có thể có được từ thông tin của khách đi vé tháng, hoặc từ hình ảnh được chụp bởi một máy ảnh.

Ví dụ, thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1533L phân cụm phần lớn dữ liệu đặc trưng tương tự về giới tính, tuổi, tình trạng di chuyển và cách sử dụng phiếu thưởng trong cơ sở dữ liệu liên quan đến khách hàng 1533L. Sau đó, đối với mỗi phân cụm, phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được xử lý thống kê để tạo ra dữ liệu đặc trưng của một thiết bị đầu cuối di động ảo trong đó phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được tích hợp. Bằng cách này, thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi thiết bị đầu cuối di động ảo, tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1533L.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1533R phân cụm phần lớn dữ liệu đặc trưng về đặc trưng, trong cơ sở dữ liệu liên quan đến khách hàng 1533R. Phần lớn dữ liệu đặc trưng được xử lý thống kê cho mỗi phân cụm để tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi sự kết hợp ảo của nhà ga, cửa bán vé, thời tiết, múi giờ.

Thiết bị kết hợp dữ liệu 50 được cấu hình để tạo ra cơ sở dữ liệu kết hợp 557 bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1533L và cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1533R.

Ví dụ, thiết bị kết hợp dữ liệu 50 sử dụng các tham số đại diện cho giới tính và tuổi được giữ chung bởi cơ sở dữ liệu xử lý sau 155, cũng như các tham số liên quan đến địa điểm và thời tiết làm dữ liệu tham chiếu, để kết hợp các cơ sở dữ liệu sau xử lý 155 kết hợp để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện cho dữ liệu tham chiếu giữa các cơ sở dữ liệu hậu xử lý 155 được kết hợp.

Theo cơ sở dữ liệu kết hợp 557 được tạo ra như vậy, ví dụ, mỗi quan hệ tương ứng giữa hành vi tiêu dùng và cách sử dụng phiếu thưởng mua hàng ở gần một nhà ga và các xu hướng của khách đi qua trong nhà ga và cửa bán vé có thể được phân tích. Do đó, dựa vào cơ sở dữ liệu kết hợp 557, các phiếu thưởng mua hàng cho mục đích để thúc đẩy việc sử dụng tại các cửa hàng ở gần nhà ga có thể được gửi chính xác đến các thiết bị đầu cuối di động của người sử dụng đã đi qua cửa bán vé.

[Phương án thực hiện thứ sáu]

Dưới đây, phần mô tả các cấu hình và các nội dung xử lý trong hệ thống xử lý thông tin 1 theo phương án thực hiện thứ sáu tương tự với cấu hình và nội dung xử lý của phương án thực hiện thứ nhất và phương án thực hiện thứ hai sẽ được bỏ qua. Hệ thống xử lý thông tin 1 theo phương án thực hiện này có các cơ sở dữ liệu liên quan đến khách hàng 1534L, 1534R thể hiện trên Fig.10B, làm cơ sở dữ liệu liên quan đến khách hàng 153.

Trong phần trái trên Fig.10B, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1534L gồm dữ liệu đặc trưng về các khách hàng có sở hữu các thẻ ETC. Các thẻ ETC là các thẻ tín dụng cho hệ thống thu phí điện tử (ETC) trong các mạng lưới thu phí đường bộ, phổ biến ở Nhật Bản. Trong phần phải trên Fig.10B, thể hiện một ví dụ thể hiện cơ sở dữ liệu liên quan đến khách hàng 1534R gồm dữ liệu đặc trưng của mỗi sự kết hợp giữa giao lộ và cửa thu phí, làm dữ liệu đặc trưng về khách hàng.

Cơ sở dữ liệu liên quan đến khách hàng 1534L bao gồm dữ liệu đặc trưng cho mỗi thẻ ETC. Dữ liệu đặc trưng đại diện cho một ID của một thiết bị trên xe được kết hợp với thẻ ETC. Ngoài ra, dữ liệu đặc trưng đại diện cho giới tính, tuổi và nơi cư trú của một khách hàng sở hữu thẻ ETC. Hơn nữa, dữ liệu đặc trưng đại diện cho các ID của các cửa thu phí và ngày và thời gian đi qua các cửa thu phí.

Dữ liệu đặc trưng tạo thành cơ sở dữ liệu liên quan đến khách hàng 1534R đại diện cho số thẻ ETC của xe cộ đi qua cửa thu phí tương ứng của giao lộ tương ứng, ID thiết bị trên xe, và ngày và thời gian đi qua cửa. Ngoài ra, dữ liệu đặc trưng gồm thông tin về số xe cộ đi qua và các phương tiện xung quanh.

Ví dụ, thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1534L phân cụm phần lớn dữ liệu đặc trưng tương tự về giới tính, tuổi và trạng thái di chuyển

trong cơ sở dữ liệu liên quan đến khách hàng 1534L. Sau đó, đối với mỗi phân cụm, phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được xử lý thống kê để tạo ra dữ liệu đặc trưng của mỗi khách hàng ảo (thẻ) trong đó phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được tích hợp. Bằng cách này, thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi khách hàng ảo (thẻ), tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1534L.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1534R phân cụm phần lớn dữ liệu đặc trưng tương tự về đặc trưng, trong cơ sở dữ liệu liên quan đến khách hàng 1534R. Phần lớn dữ liệu đặc trưng được xử lý thống kê cho mỗi phân cụm để tạo ra cơ sở dữ liệu sau xử lý 155 bao gồm dữ liệu đặc trưng của mỗi cổng thu phí ảo.

Thiết bị kết hợp dữ liệu 50 được cấu hình để tạo ra cơ sở dữ liệu kết hợp 557 bằng cách kết

hợp cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1534L và cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1534R.

Dựa vào cơ sở dữ liệu kết hợp 557 được tạo ra như vậy, các vé thưởng mua hàng cho mục đích thúc đẩy việc sử dụng các phương tiện gần các cửa thu phí có thể được phân bổ thích hợp đến các khách hàng thông qua các thiết bị gắn trên xe như các hệ thống dẫn đường xe hoặc các công ty thẻ tín dụng.

[Phương án thực hiện thứ bảy]

Dưới đây, phần mô tả các cấu hình và các nội dung xử lý trong hệ thống xử lý thông tin 1 theo phương án thực hiện thứ bảy tương tự với cấu hình và nội dung xử lý của phương án thực hiện thứ nhất và phương án thực hiện thứ hai sẽ được bỏ qua. Hệ thống xử lý thông tin 1 theo phương án thực hiện này có cơ sở dữ liệu liên quan đến khách hàng 1535L, 1535R thể hiện trên Fig.11A, làm cơ sở dữ liệu liên quan đến khách hàng 153.

Trong phần trái của Fig.11A, thể hiện một ví dụ cơ sở dữ liệu liên quan đến khách hàng 1535L gồm dữ liệu đặc trưng của mỗi khách đi qua phía trước của biển chỉ dẫn kỹ thuật số cho mỗi biển chỉ dẫn kỹ thuật số, làm dữ liệu đặc trưng của các khách

hàng. Trong phần phải trên Fig.11A, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1535R gồm dữ liệu đặc trưng của mỗi khách qua các cửa bán vé cho mỗi cửa bán vé của các nhà ga, làm dữ liệu đặc trưng của các khách hàng.

Theo Fig.11A, dữ liệu đặc trưng của các cơ sở dữ liệu liên quan đến khách hàng 1535L đại diện cho điểm đặt bảng chỉ dẫn kỹ thuật số. Ngoài ra, dữ liệu đặc trưng đại diện cho giới tính, tuổi và ngày và giờ khách đi qua. Giới tính và tuổi của khách đi qua có thể được xác định từ một hình ảnh được chụp bằng máy ảnh lắp trên bảng chỉ dẫn kỹ thuật số hoặc được lắp ở vùng xung quanh của nó. Dữ liệu đặc trưng có thể đại diện cho số lượng khách đi qua trong mỗi múi giờ gần bảng chỉ dẫn kỹ thuật số, số lượng người xem bảng chỉ dẫn kỹ thuật số trong mỗi múi giờ, và dữ liệu đại diện cho thông tin người quảng cáo trong mỗi múi giờ.

Dữ liệu đặc trưng của mỗi khách đi qua trong cơ sở dữ liệu liên quan đến khách hàng 1535R đại diện cho giới tính, tuổi và ngày và giờ khách đi qua, cũng như một ID của cửa bán vé đi qua và ID của nhà ga có cửa bán vé.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1535L phân cụm phần lớn dữ liệu đặc trưng tương tự về đặc trưng của khách đi qua, ví dụ, trong cơ sở dữ liệu liên quan đến khách hàng 1535L. Sau đó, đối với mỗi phân cụm, phần lớn dữ liệu đặc trưng thuộc cùng phân cụm được xử lý thống kê để tạo ra dữ liệu đặc trưng của khách đi qua ảo trong đó phần lớn dữ liệu đặc trưng thuộc cùng phân cụm được tích hợp cho mỗi phân cụm. Trong cách này, thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi khách đi qua ảo, tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1535L.

Ví dụ, thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1535R phân cụm phần lớn dữ liệu đặc trưng tương tự về đặc trưng của khách đi qua trong cơ sở dữ liệu liên quan đến khách hàng 1535R. Dữ liệu đặc trưng được xử lý thống kê cho mỗi phân cụm để tạo ra cơ sở dữ liệu sau xử lý 155 bao gồm dữ liệu đặc trưng của mỗi khách đi qua ảo.

Thiết bị kết hợp dữ liệu 50 được cấu hình để tạo ra cơ sở dữ liệu kết hợp 557 bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1535L và cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1535R.

Ví dụ, thiết bị kết hợp dữ liệu 50 sử dụng các tham số đại diện cho giới tính và tuổi của khách đi qua được giữ chung bởi các cơ sở dữ liệu sau xử lý 155, cũng như tham số liên quan đến địa điểm làm dữ liệu tham chiếu, để kết hợp các cơ sở dữ liệu sau xử lý 155 được kết hợp để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện cho dữ liệu tham chiếu giữa các cơ sở dữ liệu sau xử lý 155 được kết hợp.

Theo cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1535L, “tham số liên quan đến địa điểm” tương ứng với một tham số là địa điểm của bảng chỉ dẫn kỹ thuật số của dữ liệu đặc trưng. Trong cơ sở dữ liệu hậu xử lý 155 mà tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1535R, “tham số liên quan đến địa điểm” tương ứng với các tham số là nhà ga và cửa bán vé.

Theo cơ sở dữ liệu kết hợp 557 được tạo ra như vậy, ví dụ, mối quan hệ tương ứng giữa khách đi qua một cửa bán vé và khách đi qua trước bảng chỉ dẫn kỹ thuật số có thể được phân tích. Do đó, theo sự thay đổi khách đi qua ở cửa bán vé, việc quảng cáo hiển thị trên bảng chỉ dẫn kỹ thuật số có thể được thay đổi thích hợp.

[Phương án thực hiện thứ tám]

Dưới đây, phần mô tả các cấu hình và các nội dung xử lý trong hệ thống xử lý thông tin 1 theo phương án thực hiện thứ 8 tương tự với cấu hình và nội dung xử lý của phương án thực hiện thứ nhất và phương án thực hiện thứ hai sẽ được bỏ qua. Hệ thống xử lý thông tin 1 theo phương án thực hiện này có các cơ sở dữ liệu liên quan đến khách hàng 1536L, 1536R thể hiện trên Fig.11B, làm cơ sở dữ liệu liên quan đến khách hàng 153.

Trong phần trái trên Fig.11B, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1536L gồm dữ liệu đặc trưng của mỗi khách đi qua trong khu vực giám sát được chụp bởi máy ảnh, làm dữ liệu đặc trưng về các khách hàng. Trong phần phải trên Fig.11B, thể hiện một ví dụ về cơ sở dữ liệu liên quan đến khách hàng 1536R gồm dữ liệu đặc trưng đại diện cho cách sử dụng máy bán đồ uống tự động, làm dữ liệu đặc trưng của các khách hàng.

Theo Fig.11B, dữ liệu đặc trưng của cơ sở dữ liệu liên quan đến khách hàng 1536L đại diện cho địa điểm (khu vực giám sát) nơi khách đi qua được chụp và mùi giờ. Ngoài ra, dữ liệu đặc trưng đại diện cho giới tính và tuổi của khách đi qua. Thêm

nữa, dữ liệu đặc trưng có thể đại diện cho thời gian giữ lại hoặc thời gian đi qua của khách.

Cơ sở dữ liệu liên quan đến khách hàng 1536R gồm dữ liệu đặc trưng của mỗi lần bán đồ uống trong mỗi máy bán hàng tự động. Dữ liệu đặc trưng tạo thành cơ sở dữ liệu liên quan đến khách hàng 1536R được cấu hình theo cách giống như trong dữ liệu đặc trưng của cơ sở dữ liệu liên quan đến khách hàng 1532R theo phương án thực hiện thứ tư.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1536L phân cụm phần lớn dữ liệu đặc trưng tương tự về đặc trưng khách đi qua, trong cơ sở dữ liệu liên quan đến khách hàng 1536L. Sau đó, đối với mỗi phân cụm, phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được xử lý thống kê để tạo ra dữ liệu đặc trưng của một khách đi qua ảo trong đó phần lớn dữ liệu đặc trưng thuộc cùng một phân cụm được tích hợp nhất cho mỗi phân cụm. Trong cách này, thiết bị xử lý dữ liệu 10 tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi khách đi qua ảo, tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1536L.

Thiết bị xử lý dữ liệu 10 có cơ sở dữ liệu liên quan đến khách hàng 1536R phân cụm phần lớn dữ liệu đặc trưng cho mỗi dữ liệu đặc trưng tương tự về đặc trưng người mua hàng, trong cơ sở dữ liệu liên quan đến khách hàng 1536R, theo đó tạo ra cơ sở dữ liệu sau xử lý 155 gồm dữ liệu đặc trưng của mỗi người mua hàng ảo.

Thiết bị kết hợp dữ liệu 50 được cấu hình để tạo ra cơ sở dữ liệu kết hợp 557 bằng cách kết hợp cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1536L và cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1536R.

Ví dụ, thiết bị kết hợp dữ liệu 50 sử dụng các tham số đại diện cho giới tính và tuổi được giữ chung bởi các cơ sở dữ liệu sau xử lý 155, cũng như một tham số liên quan đến địa điểm làm dữ liệu tham chiếu, để kết hợp các cơ sở dữ liệu sau xử lý 155 được kết hợp để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện cho dữ liệu tham chiếu giữa các cơ sở dữ liệu sau xử lý 155 được kết hợp. Trong cơ sở dữ liệu sau xử lý 155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1536L, “tham số liên quan đến địa điểm” tương ứng với một tham số được giữ bởi dữ liệu đặc trưng và đại diện cho địa điểm nơi khách đi qua đã được chụp ảnh. Trong cơ sở dữ liệu sau xử lý

155 tương ứng với cơ sở dữ liệu liên quan đến khách hàng 1536R, “tham số liên quan đến địa điểm” tương ứng với một tham số là địa điểm của một máy bán hàng tự động.

Theo cơ sở dữ liệu kết hợp 557 được tạo ra như vậy, ví dụ, mối quan hệ tương ứng giữa sự phân bổ và thời gian khách giữ lại đi qua và hành vi mua sắm tại các máy bán hàng tự động có thể được phân tích. Các kết quả của việc phân tích này có thể giúp cho việc quản lý hàng tồn kho của các máy bán hàng tự động, việc phân loại, và hiển thị các đề xuất.

[Các phương án thực hiện khác]

Sáng chế không bị giới hạn ở các phương án thực hiện được mô tả trên đây, và có thể thực hiện các phương án khác nhau. Ví dụ, theo các phương án thực hiện trên đây, một ví dụ đã được mô tả các nhóm khách hàng giữa các cơ sở dữ liệu được kết hợp (các cơ sở dữ liệu sau xử lý 155 hoặc các cơ sở dữ liệu liên quan đến khách hàng 153) là không đồng nhất. Tuy nhiên, các kỹ thuật của sáng chế này có thể được sử dụng để kết hợp các cơ sở dữ liệu đại diện cho các loại đặc trưng khác nhau đối với các nhóm khách hàng giống nhau. Ví dụ, các kỹ thuật của sáng chế này có thể được sử dụng để kết hợp dữ liệu mua tại các cửa hàng tiện lợi và kết hợp dữ liệu mua tại các siêu thị đối với các nhóm khách hàng giống nhau. Bất kỳ phương án thực hiện nào trong ý tưởng kỹ thuật được xác định từ cách diễn đạt của các yêu cầu bảo hộ là các phương án thực hiện của sáng chế.

Cuối cùng, sự tương quan giữa các thuật ngữ sẽ được mô tả. Bộ phận tính toán 11 (đặc biệt, bộ phận phân cụm 110 và bộ phận xử lý 115) được bố trí trong thiết bị xử lý dữ liệu 10 tương ứng với một ví dụ về bộ phận phát sinh. Bộ phận tính toán 51 (đặc biệt, bộ xử lý tổng hợp dữ liệu 510) bố trí trong thiết bị kết hợp dữ liệu 50 tương ứng với một ví dụ về bộ phận tổ hợp. Bộ phận lưu trữ 55 bố trí trong thiết bị kết hợp dữ liệu 50 tương ứng với một ví dụ về bộ phận lưu trữ. Tập hợp dữ liệu đặc trưng tạo thành cơ sở dữ liệu (cơ sở dữ liệu liên quan đến khách hàng 153/ cơ sở dữ liệu sau xử lý 155) hoặc tập hợp người, các đồ vật (ví dụ, các thiết bị đầu cuối di động/ các máy bán hàng tự động/ các lò / các thẻ), hoặc các địa điểm (ví dụ, các cửa bán vé/ các cổng thu phí) tương ứng với một ví dụ của một nhóm. Các yếu tố của tập hợp tương ứng với một ví dụ của một phân tử.

Yêu cầu bảo hộ

1. Hệ thống xử lý thông tin gồm:

bộ phận phát sinh được cấu hình để phát sinh cơ sở dữ liệu thứ nhất, cơ sở dữ liệu thứ nhất bao gồm dữ liệu đặc trưng của mỗi phần tử ảo được phát sinh bằng cách tích hợp dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng bằng cách xử lý thống kê dựa vào dữ liệu đặc trưng của mỗi phần tử của nhóm thứ nhất, dữ liệu đặc trưng của mỗi phần tử ảo bao gồm một số thống kê được tính toán bằng việc xử lý thống kê; và

bộ phận kết hợp được cấu hình để kết hợp cơ sở dữ liệu thứ nhất được phát sinh bởi bộ phận phát sinh với cơ sở dữ liệu thứ hai bao gồm dữ liệu đặc trưng của mỗi phần tử ở nhóm thứ hai, nhóm thứ hai là giống hoặc khác so với nhóm thứ nhất,

mỗi dữ liệu đặc trưng của cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai bao gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp, dữ liệu tham chiếu đại diện cho kiểu đặc trưng chung giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai,

bộ phận kết hợp kết hợp cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai.

2. Hệ thống xử lý thông tin theo điểm 1,

trong đó bộ phận phát sinh phân cụm một tập hợp dữ liệu đặc trưng liên quan đến nhóm thứ nhất phần lớn dữ liệu đặc trưng giống hoặc tương tự về đặc trưng, và xử lý thống kê cho mỗi phân cụm phần lớn dữ liệu đặc trưng tương ứng với phân cụm để tính một số thống kê cho mỗi tham số của phần lớn dữ liệu đặc trưng, theo cách đó phát sinh dữ liệu đặc trưng cho mỗi phần tử ảo bao gồm một số thống kê làm giá trị tham số đại diện cho đặc trưng.

3. Hệ thống xử lý thông tin bao gồm:

bộ phận lưu trữ được cấu hình để lưu cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai; và

bộ phận kết hợp được cấu hình để kết hợp cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai được lưu trong bộ phận lưu trữ,

cơ sở dữ liệu thứ nhất là cơ sở dữ liệu bao gồm dữ liệu đặc trưng của mỗi phần tử ảo được phát sinh bởi việc tích hợp dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng bằng cách xử lý thống kê dựa vào dữ liệu đặc trưng của mỗi phần tử ở nhóm thứ nhất, dữ liệu đặc trưng của mỗi phần tử ảo bao gồm một số thống kê được tính toán bởi việc xử lý thống kê.

cơ sở dữ liệu thứ hai là cơ sở dữ liệu bao gồm dữ liệu đặc trưng của mỗi phần tử ở nhóm thứ hai, nhóm thứ hai là giống hoặc khác với nhóm thứ nhất,

mỗi dữ liệu đặc trưng của cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai bao gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp, dữ liệu tham chiếu đại diện cho loại đặc trưng chung giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai,

bộ phận kết hợp kết hợp cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu giữa cơ sở dữ liệu thứ nhất và cơ sở dữ liệu thứ hai.

4. Hệ thống xử lý thông tin theo điểm bất kỳ trong số các điểm từ 1 đến 3,

trong đó dữ liệu đặc trưng của mỗi phần tử của cơ sở dữ liệu thứ hai là dữ liệu đặc trưng của mỗi phần tử ảo được phát sinh bởi việc tích hợp dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng, dựa vào dữ liệu đặc trưng của mỗi phần tử ở nhóm giống hoặc khác với nhóm thứ nhất.

5. Hệ thống xử lý thông tin theo điểm bất kỳ trong số các điểm từ 1 đến 3,

trong đó cơ sở dữ liệu thứ hai là cơ sở dữ liệu gồm dữ liệu nguồn đơn làm dữ liệu đặc trưng.

6. Hệ thống xử lý thông tin theo điểm bất kỳ trong số các điểm từ 1 đến 5,

trong đó mỗi nhóm thứ nhất và nhóm thứ hai là một tập hợp của ít nhất một trong số người, các đồ vật, dịch vụ, và địa điểm, và phần tử là một yếu tố của tập hợp.

7. Hệ thống xử lý thông tin theo điểm 6,

trong đó một trong số nhóm thứ nhất và nhóm thứ hai là một tập hợp người, và số còn lại trong nhóm thứ nhất và nhóm thứ hai khác là một tập hợp gồm ít nhất một trong số các đồ vật, dịch vụ, và địa điểm liên quan đến người.

8. Hệ thống xử lý thông tin theo điểm 6,

trong đó mỗi nhóm trong số các nhóm thứ nhất và nhóm thứ hai là một tập hợp ít nhất một trong số các đồ vật, dịch vụ, và địa điểm liên quan đến người.

9. Hệ thống xử lý thông tin bao gồm:

phần lớn bộ phận phát sinh được cấu hình để phát sinh phần lớn cơ sở dữ liệu cho các nhóm giống hoặc khác nhau; và

bộ phận kết hợp được cấu hình để kết hợp phần lớn cơ sở dữ liệu được phát sinh bởi phần lớn bộ phận phát sinh,

mỗi bộ phận phát sinh được cấu hình để phát sinh cơ sở dữ liệu bao gồm dữ liệu đặc trưng của mỗi phần tử ảo được phát sinh bằng cách tích hợp dữ liệu đặc trưng của phần lớn phần tử giống nhau hoặc tương tự về đặc trưng bằng cách xử lý thống kê, làm cơ sở dữ liệu, dựa vào dữ liệu đặc trưng của mỗi phần tử của nhóm tương ứng, dữ liệu đặc trưng của mỗi phần tử ảo bao gồm một số thống kê được tính toán bởi việc xử lý thống kê,

mỗi dữ liệu đặc trưng của phần lớn cơ sở dữ liệu gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp, dữ liệu tham chiếu đại diện cho một kiểu đặc trưng chung của cơ sở dữ liệu được kết hợp,

bộ phận kết hợp kết hợp phần lớn cơ sở dữ liệu để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu trong số phần lớn cơ sở dữ liệu.

10. Hệ thống xử lý thông tin theo điểm 9,

trong đó mỗi bộ phận phát sinh phân cụm một tập hợp dữ liệu đặc trưng liên quan đến nhóm thành phần lớn dữ liệu đặc trưng giống hoặc tương tự về đặc trưng, và xử lý thống kê đối với mỗi phân cụm phần lớn dữ liệu đặc trưng tương ứng với phân cụm để tính toán một số thống kê cho mỗi tham số của phần lớn dữ liệu đặc trưng, theo cách đó phát ra dữ liệu đặc trưng của mỗi phần tử ảo bao gồm một số thống kê làm giá trị tham số đại diện cho đặc trưng.

11. Hệ thống xử lý thông tin bao gồm:

bộ phận lưu trữ được cấu hình để lưu phần lớn cơ sở dữ liệu; và

bộ phận kết hợp được cấu hình để kết hợp phần lớn cơ sở dữ liệu được lưu trong bộ phận lưu trữ,

mỗi cơ sở dữ liệu là cơ sở dữ liệu đối với nhóm giống nhau hoặc nhóm khác nhau, cơ sở dữ liệu bao gồm dữ liệu đặc trưng của mỗi phần tử ảo được phát sinh bởi việc tích hợp dữ liệu đặc trưng của phần lớn phần tử giống hoặc tương tự về đặc trưng bởi việc xử lý thống kê, dựa vào dữ liệu đặc trưng của mỗi phần tử của nhóm tương ứng, dữ liệu đặc trưng của mỗi phần tử ảo bao gồm một số thống kê được tính toán bởi việc xử lý thống kê,

mỗi dữ liệu đặc trưng của phần lớn cơ sở dữ liệu bao gồm dữ liệu tham chiếu để tham chiếu đến vào lúc kết hợp các cơ sở dữ liệu, dữ liệu tham chiếu đại diện cho kiểu đặc trưng chung của cơ sở dữ liệu được kết hợp,

bộ phận kết hợp kết hợp phần lớn cơ sở dữ liệu để kết hợp dữ liệu đặc trưng giống hoặc tương tự về đặc trưng được đại diện bởi dữ liệu tham chiếu trong số phần lớn cơ sở dữ liệu.

12. Hệ thống xử lý thông tin theo điểm bất kỳ trong số các điểm từ 9 đến 11,

trong đó bộ phận kết hợp được cấu hình để kết hợp phần lớn cơ sở dữ liệu thông qua một kiểu cơ sở dữ liệu cụ thể, kiểu cơ sở dữ liệu cụ thể gồm dữ liệu nguồn đơn làm dữ liệu đặc trưng cho mỗi phần tử của nhóm mà nhóm đó là nhóm giống hoặc nhóm khác với một trong các nhóm tương ứng với phần lớn cơ sở dữ liệu,

dữ liệu nguồn đơn bao gồm dữ liệu tham chiếu được tham chiếu đến vào lúc kết hợp các cơ sở dữ liệu, dữ liệu tham chiếu đại diện cho một kiểu đặc trưng chung của cơ sở dữ liệu được kết hợp, và

bộ phận kết hợp kết hợp kiểu cơ sở dữ liệu cụ thể và mỗi kiểu cơ sở dữ liệu khác để kết hợp dữ liệu đặc trưng tương tự về đặc trưng đại diện bởi dữ liệu tham chiếu, theo cách đó kết hợp phần lớn cơ sở dữ liệu qua kiểu cơ sở dữ liệu cụ thể.

FIG. 1

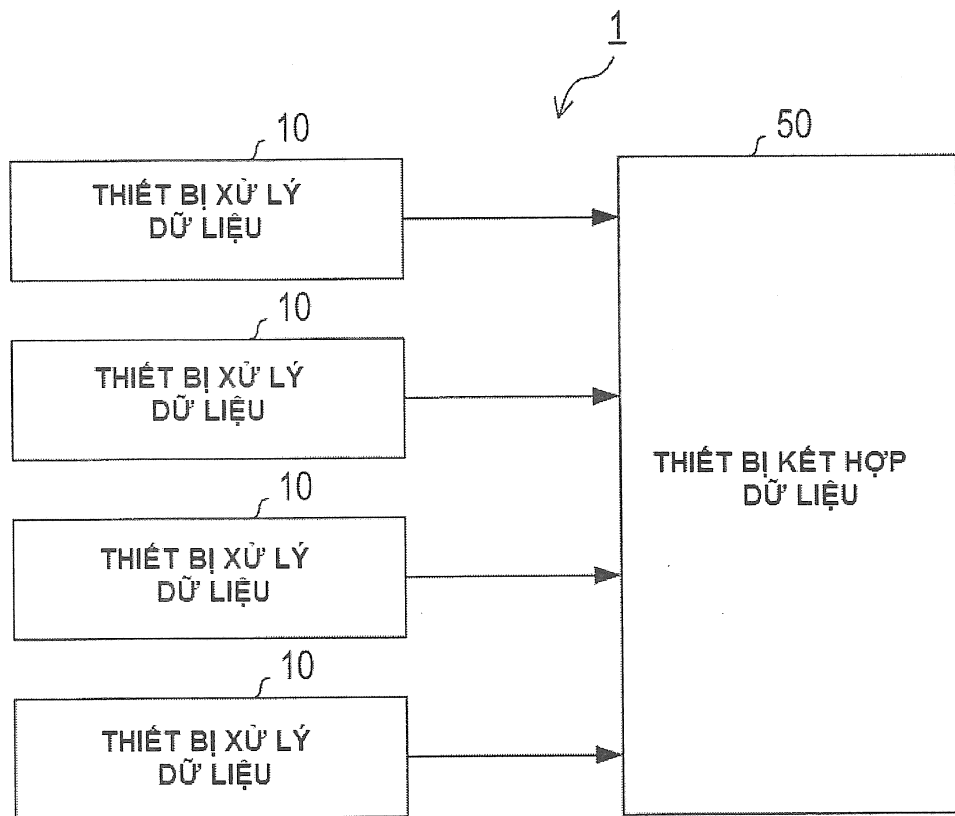


FIG. 2A

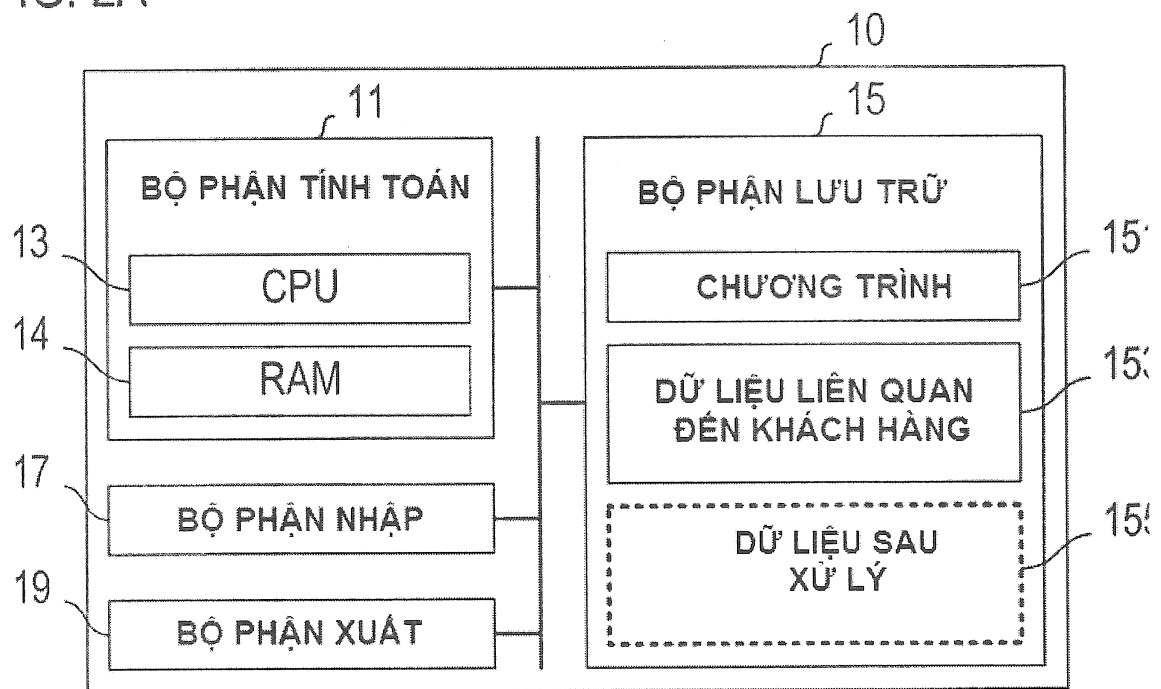


FIG. 2B

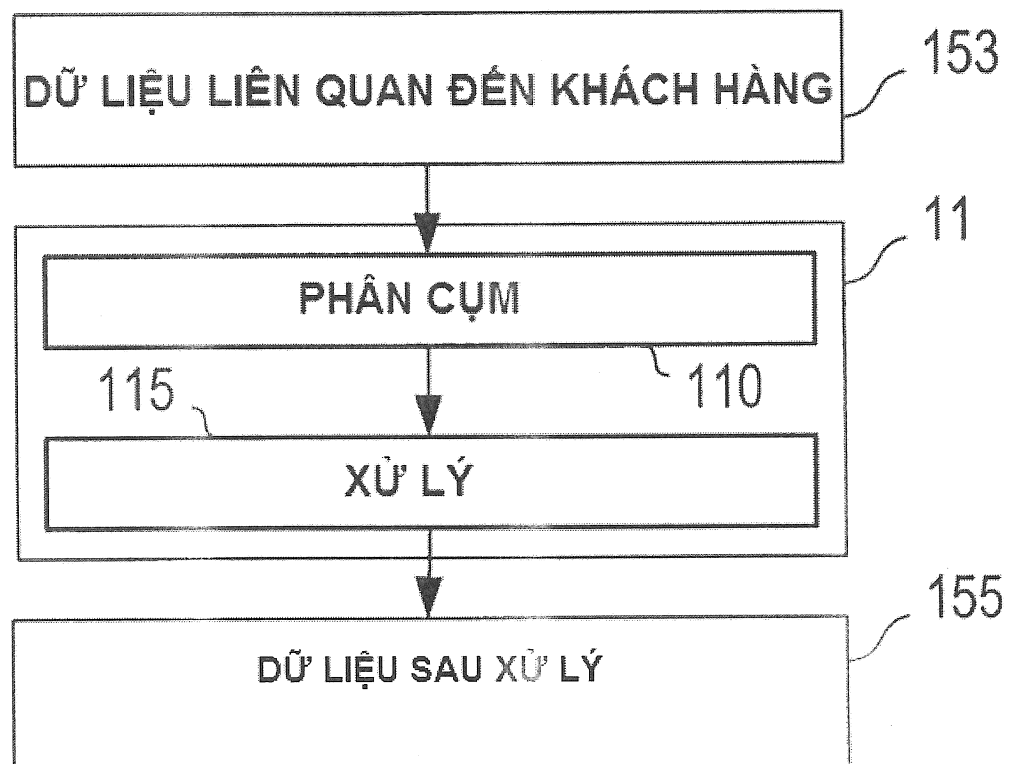


FIG. 3A

GIỚI TÍNH	TUỔI	NƠI MUA HÀNG	MÃ JAN	SỐ LƯỢNG MUA	GIÁ ĐƠN VỊ
...	...	...	...	...	...

FIG. 3B

GIỚI TÍNH	TUỔI	NƠI MUA HÀNG	GIÁ MUA
...	...	...	...

FIG. 3C

GIỚI TÍNH	TUỔI	NGHỀ NGHIỆP	TRẠNG THÁI ĐĂNG KÝ BÁO F	TRẠNG THÁI ĐĂNG KÝ TẠP CHÍ G	...
...	...	...	...	...	...

FIG. 3D

GIỚI TÍNH	TUỔI	KHU VỰC	TRẠNG THÁI THẨM ĐỊA ĐIỂM L	TRẠNG THÁI THẨM ĐỊA ĐIỂM M	...
...	...	...	...	...	...

FIG. 4

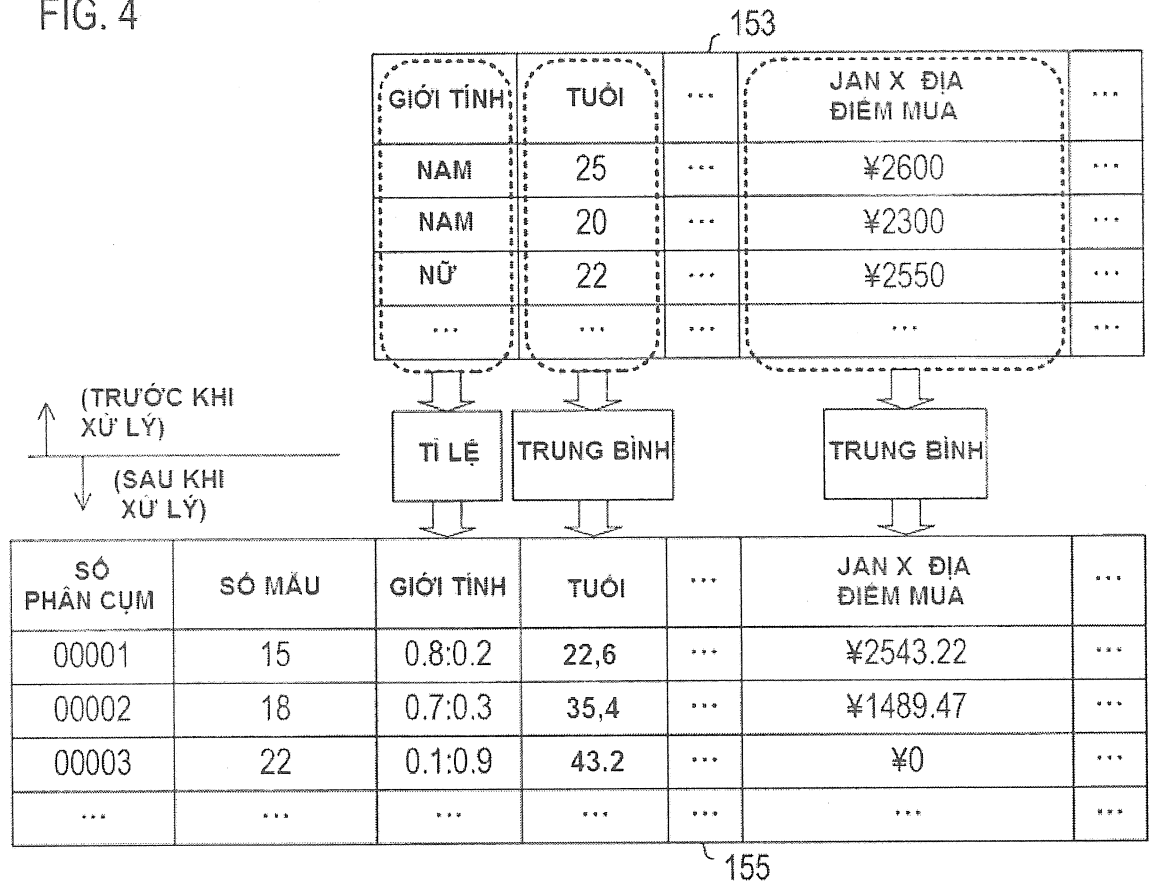


FIG. 5A

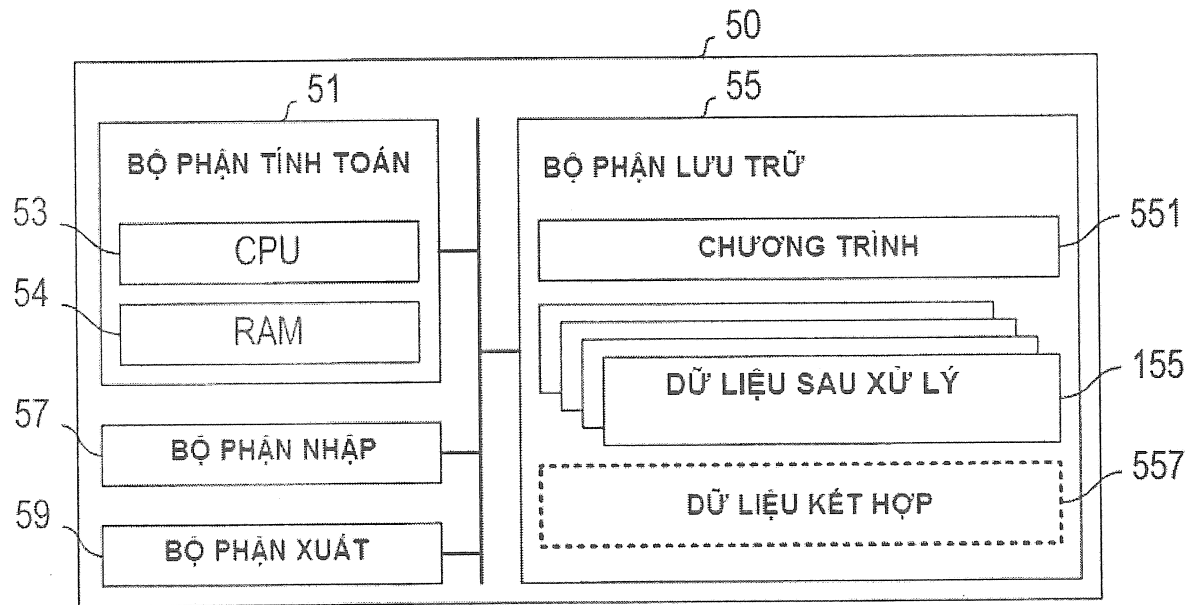


FIG. 5B

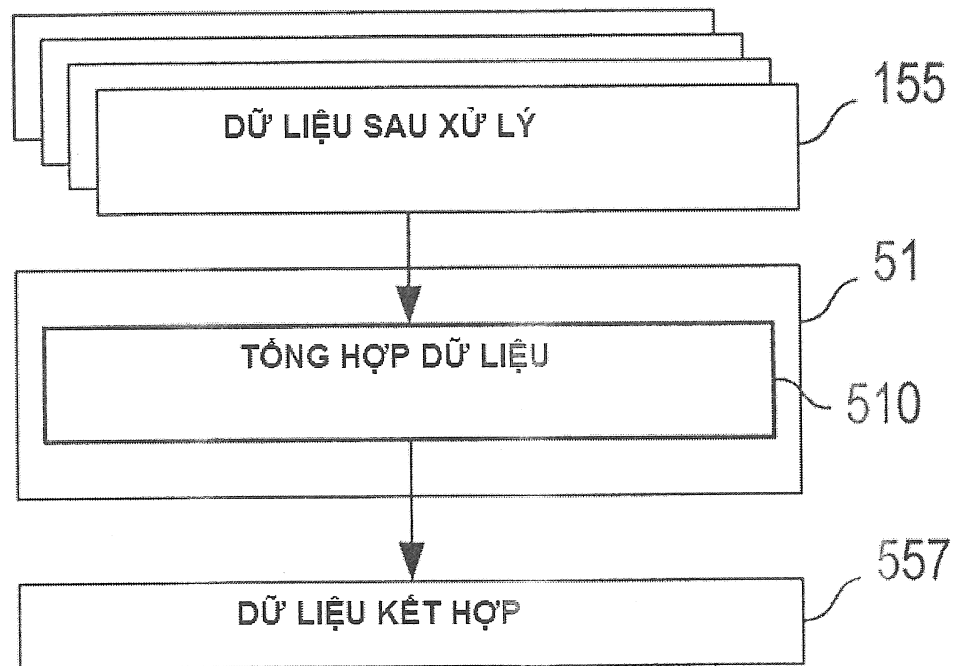


FIG. 6A

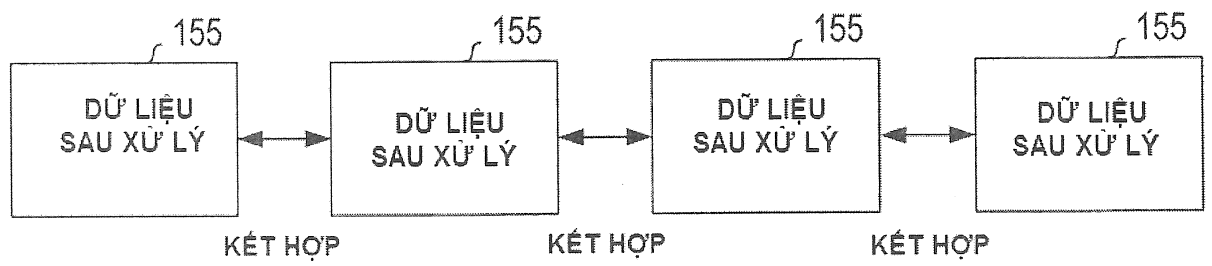


FIG. 6B

155A						
A	SỐ LƯỢNG MẪU	GIỚI TÍNH	TUỔI	MUA P1	MUA P2	...
00001	20	...	...	...	...	...
00002	23	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
155B						
B	SỐ LƯỢNG MẪU	GIỚI TÍNH	TUỔI	MUA R1	MUA R2	...
00001	18	...	...	...	...	...
...	...	...	...	...	...	...
00234	17	...	...	...	...	...
...	...	...	...	...	...	...

FIG. 6C

557

DỮ LIỆU SAU XỬ LÝ (155B)	DỮ LIỆU SAU XỬ LÝ (155B)	TRỌNG SỐ (SỐ LƯỢNG MẪU)
A00001	B20056	5,3
A00001	B00234	8,2
A00001	B01123	6,5
A00002	B00234	8,8
...	...	...

FIG. 7

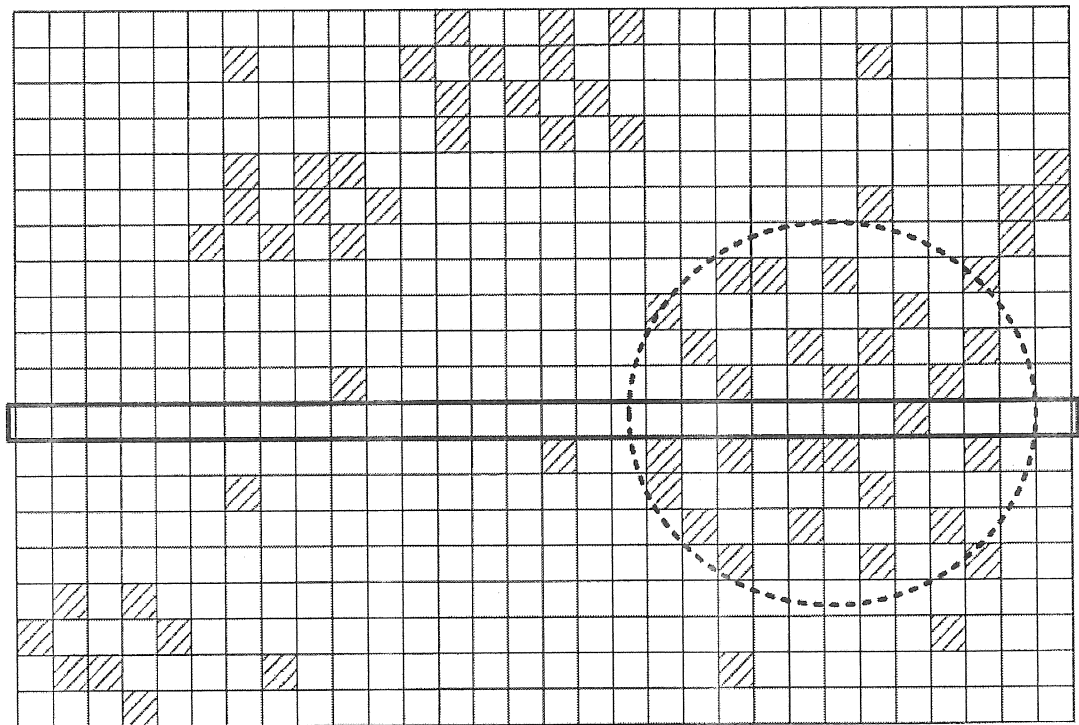


FIG. 8A

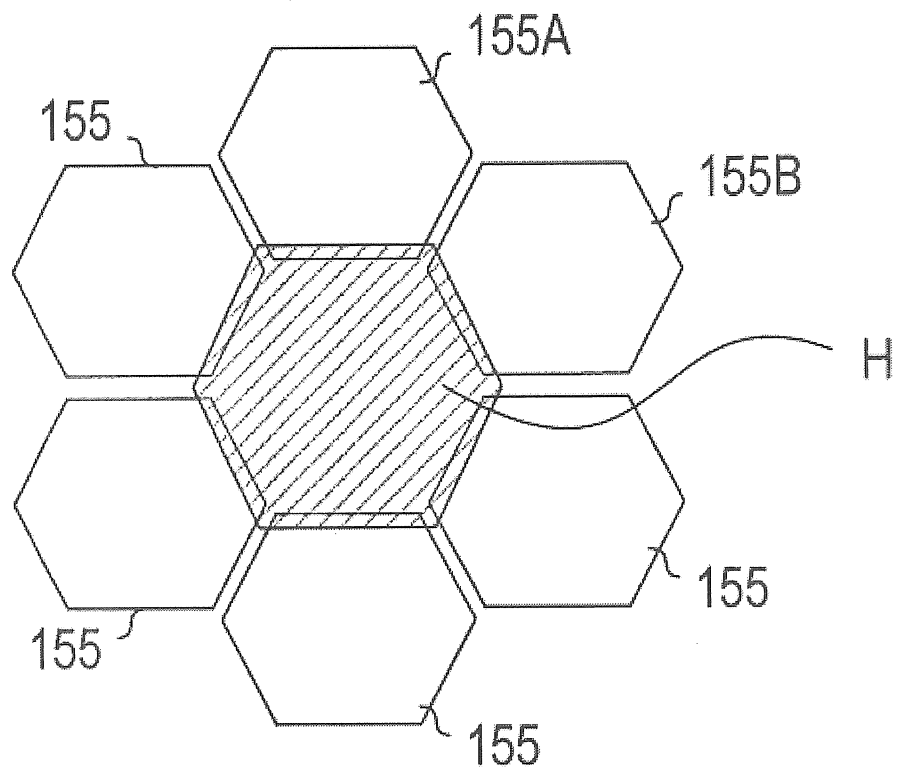


FIG. 8B

557

DỮ LIỆU HUB (H)	DỮ LIỆU SAU XỬ LÝ (155B)	TRỌNG SỐ (SỐ LƯỢNG MẪU)	DỮ LIỆU SAU XỬ LÝ (155B)	TRỌNG SỐ (SỐ LƯỢNG MẪU)	...
H00001	A12345	5.5	B00044	7.0	...
H00001	A00456	6.2	B02345	9.4	...
H00001	A11004	9.3	-	-	...

FIG. 9A

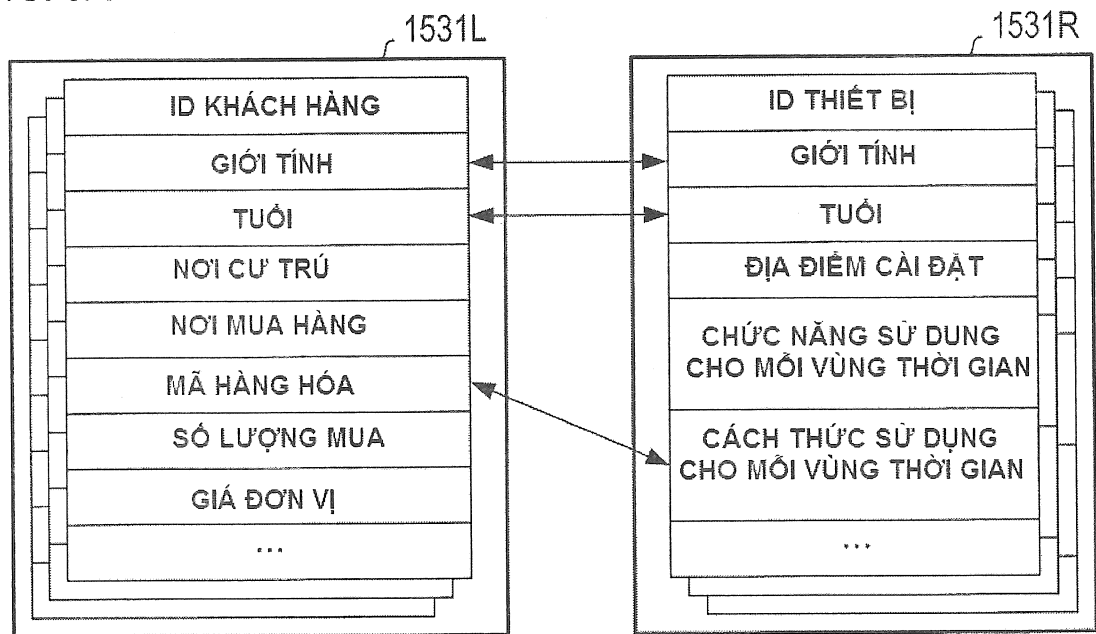


FIG. 9B

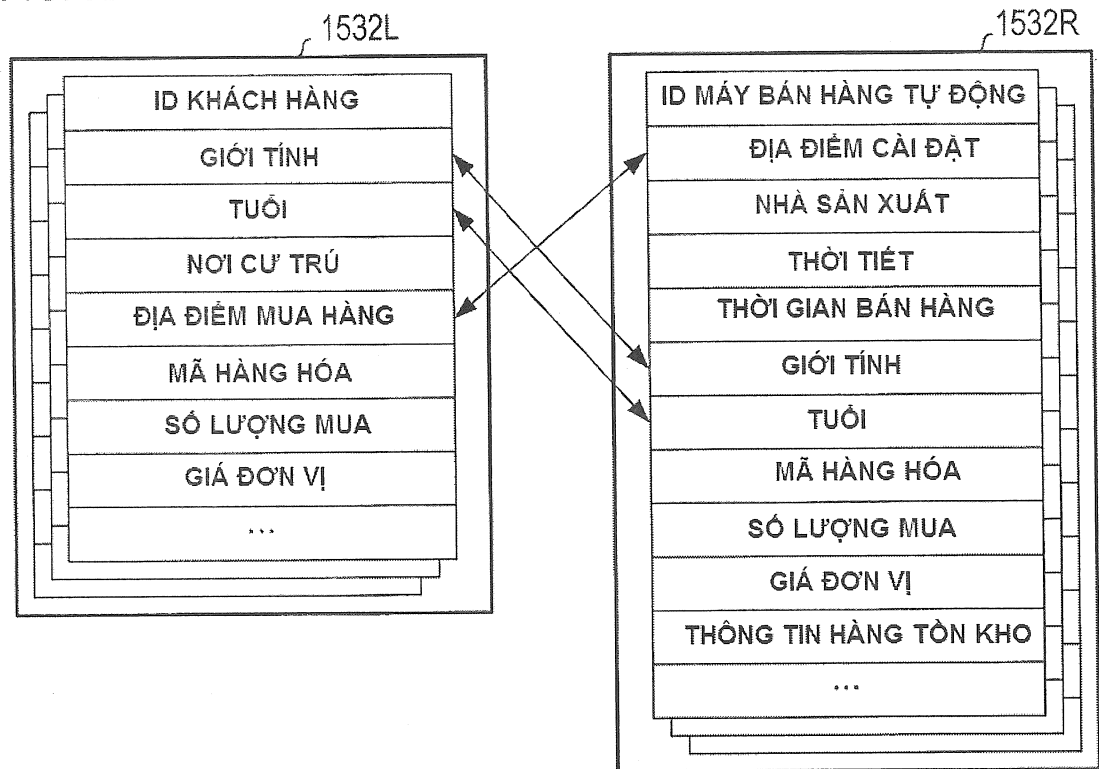


FIG. 10A

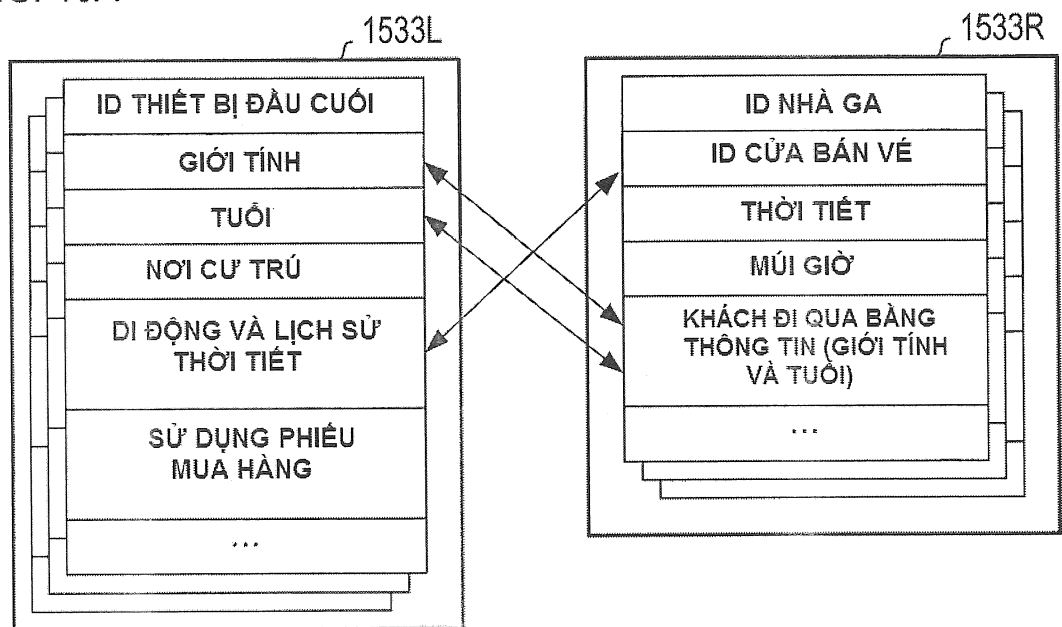


FIG. 10B

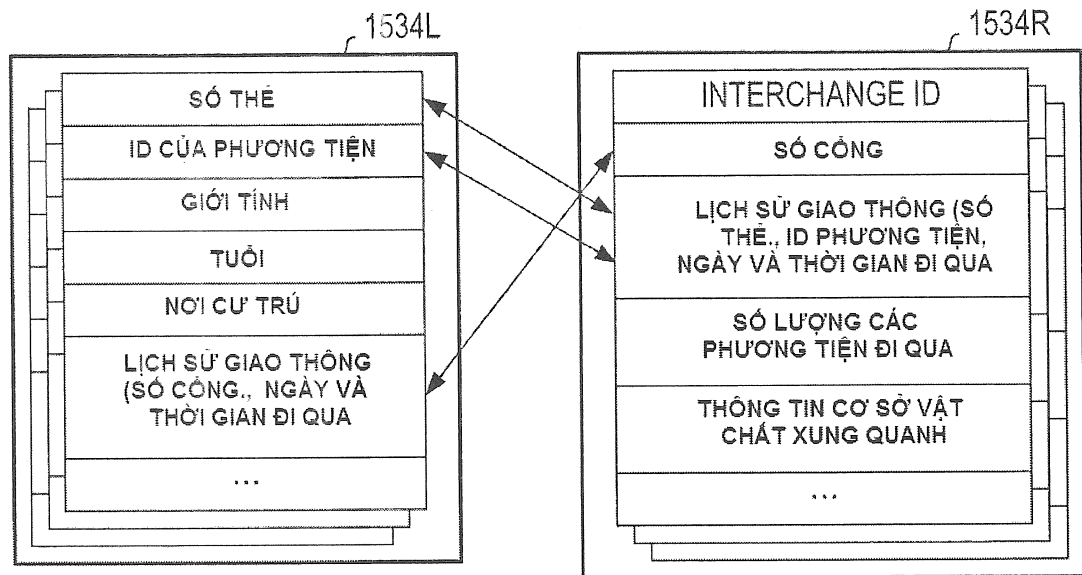


FIG. 11A

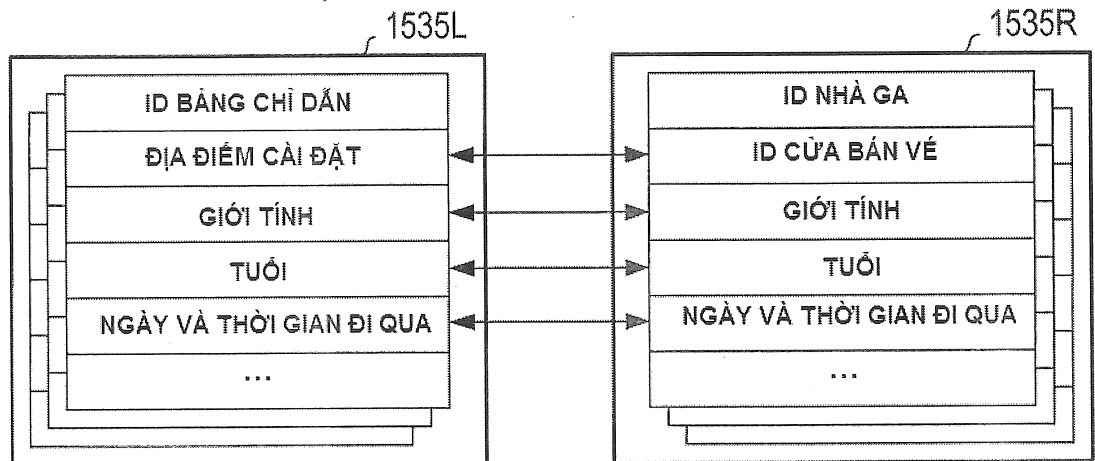


FIG. 11B

