



- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) CỤC SỞ HỮU TRÍ TUỆ
(51)^{2020.01} **G06F 16/21; G06F 16/28; G06Q 40/02; G06F 16/25** (13) **B**



1-0046114

-
- (21) 1-2020-05411 (22) 20/02/2020
(86) PCT/IB2020/051423 20/02/2020 (87) WO2020/170187 27/08/2020
(30) 1902413.2 22/02/2019 GB
(45) 26/05/2025 446 (43) 25/11/2021 404A
(73) CHANNEL TECHNOLOGIES FZE (AE)
Jebel Ali Free Zone, Office number FZJOA1813, Dubai, United Arab Emirates
(72) LELIS, Stelios (GR); CHATZISTAMATIOU, Antonios (GR).
(74) CÔNG TY TRÁCH NHIỆM HỮU HẠN TƯ VẤN ĐẦU TƯ VÀ SỞ HỮU TRÍ TUỆ
INTERFIVE (INTERFIVE CO., LTD)
-

- (54) PHƯƠNG PHÁP VÀ HỆ THỐNG XỬ LÝ CÁC BẢN GHI DỮ LIỆU Ở QUY MÔ LỚN

(21) 1-2020-05411

(57) Sáng chế đề cập đến phương pháp và hệ thống xử lý các bản ghi dữ liệu. Phương pháp này bao gồm bước xử lý các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện. Các bản ghi dữ liệu được lọc và biến đổi thành các định dạng tiêu chuẩn hoá. Các bản ghi dữ liệu đã biến đổi được lưu trữ vào bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian. Các bản ghi giao dịch thuộc các loại khác nhau được nhóm theo thông tin nhận dạng duy nhất của các thực thể. Các loại bản ghi giao dịch được lưu trữ vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất. Dữ liệu đã lưu trữ được sử dụng để tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

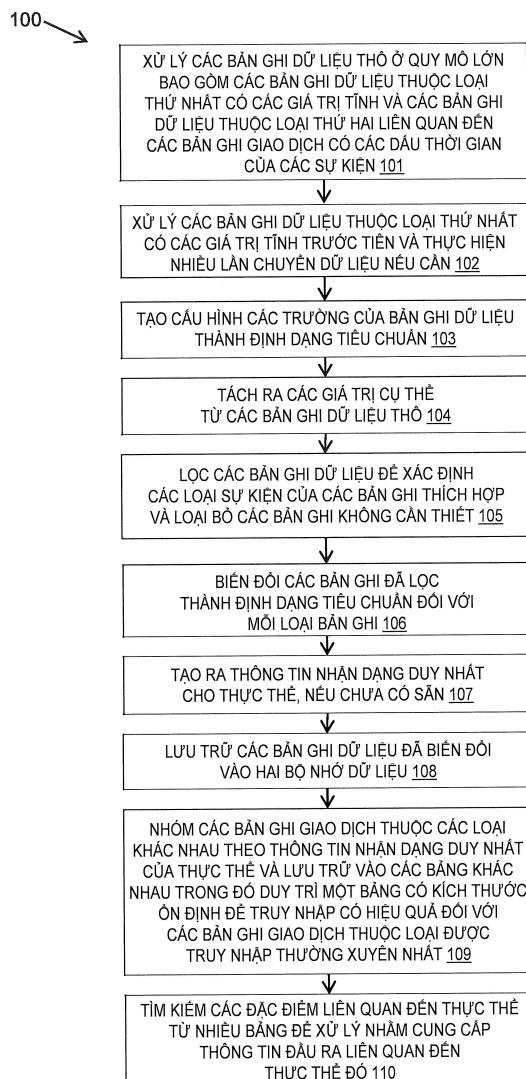


FIG. 1

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và hệ thống xử lý các bản ghi dữ liệu để tìm kiếm có hiệu quả. Cụ thể là, sáng chế đề cập đến phương pháp và hệ thống xử lý các bản ghi dữ liệu ở quy mô lớn trong lĩnh vực viễn thông.

Tình trạng kỹ thuật của sáng chế

Các thuê bao của nhà điều hành mạng di động (Mobile Network Operator, MNO) thường sử dụng các dịch vụ trước khi thanh toán và/hoặc nhận được các khoản vay thông qua nhà điều hành. Ví dụ, các thuê bao trả trước (tức là các thuê bao phải có số tiền trả trước ban đầu để sử dụng mạng), ở nhiều MNO, được quyền lựa chọn sử dụng các gói dịch vụ ứng tiền và dịch vụ di động trước khi thanh toán, và hoàn trả trong lần nạp tiền tiếp theo. Các loại khoản vay khác mà các thuê bao của MNO có thể nhận được, là số tiền ghi có vào ví di động của họ, cung cấp hàng hoá (ví dụ máy điện thoại di động) trước khi thanh toán, v.v..

Việc sử dụng trước các dịch vụ và các khoản vay, trong hầu hết các trường hợp, phát sinh chi phí cho thuê bao. Chi phí này là phí hoặc tiền lãi phải trả khi hoàn trả khoản ứng trước hoặc khoản vay. Chi phí này là khoản lãi cho bên cung cấp khoản ứng trước/khoản vay (MNO hoặc bên thứ ba). Mặt khác, nếu thuê bao không hoàn trả khoản ứng trước hoặc khoản vay, thì bên cung cấp đó nhận khoản lỗ.

Thực thể cung cấp các khoản ứng trước hoặc các khoản vay tìm cách tạo ra lợi nhuận bằng cách tối đa hoá các khoản lãi trong khi tối thiểu hoá các khoản lỗ. Một cách để đạt được mục tiêu này là thực hiện chức năng phân tích tín dụng đối với mỗi thuê bao, xác định khả năng trả được nợ của thuê bao, và cấp hạn mức tín dụng thích hợp cho mỗi thuê bao, bao gồm cả hạn mức tín dụng bằng không để chỉ báo rằng thuê bao sẽ không thể vay.

Thông thường, chức năng phân tích tín dụng là một trong số các chức năng của các tổ chức ngân hàng và dựa vào dữ liệu tài chính, như các giao dịch ngân hàng hoặc thông tin về tài sản do khách hàng nắm giữ. Nói chung, các MNO không có các dữ liệu như vậy

về các thuê bao trong mạng của họ. Các MNO chỉ có dữ liệu về việc sử dụng mạng của họ và trong một số trường hợp có dữ liệu nhân khẩu học cơ bản về các thuê bao của họ.

Dữ liệu mà các MNO có thể cung cấp, là các bản ghi chi tiết cuộc gọi điện thoại (Call Detail Record, CDR), còn được gọi là các bản ghi dữ liệu sự kiện (Event Data Record, EDR), và dữ liệu nhân khẩu học hạn chế ở dạng dữ liệu biết khách hàng của bạn (Know Your Customer, KYC) khó phân tích cho mục đích phân tích tín dụng và các mục đích sử dụng khác vì các dữ liệu này, CDR và KYC, thường không tuân theo một định dạng được xác định rõ ràng và đòi hỏi phương pháp xử lý ở quy mô lớn. Những đòi hỏi như vậy có thể khiến cho việc sử dụng dữ liệu này để cập nhật điểm tín dụng thường xuyên (ví dụ hằng ngày) là điều đặc biệt khó khăn.

Phần mô tả về tình trạng kỹ thuật của sáng chế trên đây chỉ nhằm mục đích tạo điều kiện thuận lợi cho việc hiểu rõ về sáng chế. Cần phải hiểu rằng nội dung này không phải là sự thừa nhận hay sự gợi ý rằng bất cứ tài liệu nào được đề cập đều thuộc phạm vi hiểu biết chung thông thường trong cùng một lĩnh vực cho đến ngày ưu tiên của đơn đăng ký sáng chế này.

Bản chất kỹ thuật của sáng chế

Theo khía cạnh thứ nhất của sáng chế, sáng chế đề xuất phương pháp được thực hiện bằng máy tính để xử lý các bản ghi dữ liệu ở quy mô lớn, bao gồm các bước: thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện; lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá; lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện; nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất; và tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

Các bản ghi dữ liệu thô có thể là các bản ghi dữ liệu thô ở quy mô lớn và có thể không có cùng một định dạng hoặc không có định dạng được xác định rõ ràng. Bộ nhớ dữ liệu thứ nhất có thể được tạo cấu hình để tìm kiếm có hiệu quả dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể. Bộ nhớ dữ liệu thứ hai có thể được tạo cấu hình để tìm kiếm có hiệu quả dựa vào các dấu thời gian của các sự kiện.

Bước tìm kiếm các đặc điểm liên quan đến thực thể có thể là tạo ra các đặc điểm có các số đo của hành động của thực thể ở thời điểm tham chiếu nhất định, và có thể bao gồm các bước: tìm kiếm các đặc điểm từ các bảng khác nhau; liên kết các đặc điểm; và biến đổi các đặc điểm đến thời điểm tham chiếu này.

Bộ nhớ dữ liệu thứ hai có thể được phân vùng cho các khoảng thời gian và/hoặc theo loại sự kiện. Các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất có thể là các giao dịch mở đang diễn ra. Phương pháp này có thể xử lý các bản ghi dữ liệu tĩnh trước tiên để thu được thông tin nhận dạng của các thực thể.

Theo một phương án thực hiện sáng chế, các bản ghi dữ liệu thô liên quan đến dữ liệu viễn thông và các thực thể là những người dùng, và trong đó các bản ghi dữ liệu thuộc loại thứ nhất là các bản ghi dữ liệu người dùng có các giá trị tĩnh của các thuộc tính của người dùng và các bản ghi dữ liệu thuộc loại thứ hai là các bản ghi chi tiết cuộc gọi điện thoại của các sự kiện. Phương pháp này có thể bao gồm bước tạo ra thông tin nhận dạng duy nhất cho mỗi người dùng là dạng kết hợp của số danh bạ thuê bao quốc tế của trạm di động (Mobile Station International Subscriber Directory Number, MSISDN) và ngày kích hoạt môđun nhận dạng thuê bao (Subscriber Identity Module, SIM).

Theo phương án này, bước lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá có thể bao gồm bước áp dụng các quy tắc lọc để tách ra các bản ghi thích hợp từ các bản ghi giao dịch có một hoặc nhiều nhóm gồm: các bản ghi liên quan đến các sự kiện như: các cuộc gọi điện thoại, các tin nhắn, các sự kiện giao dịch tiền tệ, các sự kiện sử dụng dữ liệu; các sự kiện liên quan đến chu kỳ sản phẩm; các sự kiện sử dụng khoản ứng trước; các sự kiện liên quan đến khoản vay; các sự kiện liên quan đến ví di động. Bước nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của thực thể có thể bao gồm bước nhóm các loại giao dịch ứng trước tín dụng mở và các loại giao dịch ứng trước tín dụng đóng, và các bản ghi giao dịch thuộc loại được

truy nhập thường xuyên nhất là các giao dịch ứng trước tín dụng mở có sẵn để sau đó phân tích. Thông tin đầu ra liên quan đến thực thể có thể là profin người dùng để sau đó phân tích rủi ro tín dụng.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất hệ thống xử lý các bản ghi dữ liệu ở quy mô lớn, hệ thống này bao gồm bộ nhớ để lưu trữ mã chương trình đọc được bằng máy tính và bộ xử lý để thực hiện mã chương trình đọc được bằng máy tính, hệ thống này bao gồm: bộ phận thu dữ liệu để thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện; bộ phận biến đổi để biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá; bộ phận lọc để lọc các bản ghi dữ liệu thích hợp; bộ phận lưu trữ vào bộ nhớ dữ liệu để lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện; bộ phận tạo nhóm để nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và bộ phận lưu trữ vào bảng để lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất; và bộ phận profin thực thể để tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

Bộ phận profin thực thể để tìm kiếm các đặc điểm liên quan đến thực thể có thể tạo ra các đặc điểm có các số đo của hành động của thực thể ở thời điểm tham chiếu nhất định, và có thể bao gồm: bộ phận tìm kiếm đặc điểm để tìm kiếm các đặc điểm từ các bảng khác nhau; bộ phận liên kết đặc điểm để liên kết các đặc điểm; và bộ phận biến đổi đặc điểm để biến đổi các đặc điểm đến thời điểm tham chiếu này.

Theo một phương án thực hiện sáng chế, các bản ghi dữ liệu thô liên quan đến dữ liệu viễn thông và các thực thể là những người dùng, và trong đó các bản ghi dữ liệu thuộc loại thứ nhất là các bản ghi dữ liệu người dùng có các giá trị tĩnh của các thuộc tính của người dùng và các bản ghi dữ liệu thuộc loại thứ hai là các bản ghi chi tiết cuộc gọi điện thoại

của các sự kiện. Theo phương án như vậy, hệ thống này có thể bao gồm bộ phận tạo ra thông tin nhận dạng duy nhất để tạo ra thông tin nhận dạng duy nhất cho mỗi người dùng là dạng kết hợp của số danh bạ thuê bao quốc tế của trạm di động (MSISDN) và ngày kích hoạt môđun nhận dạng thuê bao (SIM). Bộ phận lọc có thể áp dụng các quy tắc lọc để tách ra các bản ghi thích hợp từ các bản ghi giao dịch có một hoặc nhiều nhóm gồm: các bản ghi liên quan đến các sự kiện như: các cuộc gọi điện thoại, các tin nhắn, các sự kiện giao dịch tiền tệ, các sự kiện sử dụng dữ liệu; các sự kiện liên quan đến chu kỳ sản phẩm; các sự kiện sử dụng khoản ứng trước; các sự kiện liên quan đến khoản vay; các sự kiện liên quan đến ví di động. Bộ phận tạo nhóm có thể nhóm các loại giao dịch ứng trước tín dụng mở và các loại giao dịch ứng trước tín dụng đóng, và các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất là các giao dịch ứng trước tín dụng mở có sẵn để sau đó phân tích.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất sản phẩm chương trình máy tính để xử lý các bản ghi dữ liệu ở quy mô lớn, sản phẩm chương trình máy tính này có vật ghi đọc được bằng máy tính lưu trữ mã chương trình đọc được bằng máy tính để thực hiện các bước: thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện; lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá; lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện; nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất; và tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

Các dấu hiệu khác của sáng chế đề xuất vật ghi đọc được bằng máy tính là vật ghi bất khả biến đọc được bằng máy tính và mã chương trình đọc được bằng máy tính được

thực hiện bằng mạch xử lý.

Theo một khía cạnh khác của sáng chế, sáng chế đề xuất phương pháp được thực hiện bằng máy tính để xử lý các bản ghi dữ liệu viễn thông ở quy mô lớn, bao gồm các bước: thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tính liên quan đến thực thể người dùng và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch của các sự kiện viễn thông có các dấu thời gian của các sự kiện; lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá; lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tính liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tính có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện; nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất liên quan đến các giao dịch ứng trước tín dụng mở cho thực thể; và tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra ở dạng điểm tín dụng liên quan đến thực thể đó.

Các phương án thực hiện sáng chế sẽ được mô tả dưới đây, chỉ để làm ví dụ, dựa vào các hình vẽ kèm theo.

Mô tả vắn tắt các hình vẽ

Trên các hình vẽ:

Fig.1 là lưu đồ thể hiện phương pháp theo phương án làm ví dụ để thực hiện sáng chế;

Fig.2 là lưu đồ thể hiện phương pháp trên Fig.1 theo một phương án làm ví dụ khác để thực hiện sáng chế;

Fig.3 là sơ đồ khối thể hiện hệ thống theo phương án làm ví dụ để thực hiện sáng chế;

Fig.4 là hình vẽ thể hiện thiết bị máy tính làm ví dụ trong đó các khía cạnh khác nhau của sáng chế có thể được thực hiện.

Mô tả chi tiết sáng chế

Phương pháp và hệ thống được mô tả trong sáng chế xử lý các bản ghi dữ liệu, thường là xử lý các bản ghi dữ liệu ở quy mô lớn, trong đó các bản ghi dữ liệu bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện được ghi. Phương pháp xử lý ở quy mô lớn có thể dùng để chỉ phương pháp xử lý từ hàng trăm triệu đến hàng tỷ bản ghi dữ liệu trở lên. Theo các phương án làm ví dụ được mô tả trong sáng chế, các bản ghi dữ liệu ở dạng các bản ghi dữ liệu viễn thông bao gồm các bản ghi dữ liệu thuộc loại thứ nhất là các bản ghi dữ liệu nhân khẩu học của khách hàng như các bản ghi biết khách hàng của bạn (KYC) và các bản ghi dữ liệu thuộc loại thứ hai là các bản ghi chi tiết cuộc gọi điện thoại (CDR). Tuy nhiên, phương pháp và hệ thống được mô tả trong sáng chế có thể áp dụng tương đương cho các dạng khác của các bản ghi dữ liệu trong đó có các bản ghi dữ liệu thuộc loại thứ nhất và loại thứ hai tương tự với nhau.

Việc áp dụng phương pháp và hệ thống được mô tả trong sáng chế là sử dụng dữ liệu viễn thông để chấm điểm tín dụng cho các thuê bao của các nhà điều hành mạng di động (MNO), và sau đó sử dụng điểm tín dụng này để cấp các hạn mức tín dụng thích hợp. Cách làm này có thể được áp dụng với mọi hình thức cho vay ở MNO bao gồm, nhưng không chỉ giới hạn ở, cung cấp các khoản ứng trước cho phép sử dụng mạng, các khoản tiền vay được ghi có trên ví di động, các khoản vay của khách hàng để mua hàng hoá (ví dụ máy điện thoại di động), v.v..

Dựa vào Fig.1, lưu đồ (100) thể hiện phương pháp được mô tả trong sáng chế theo phương án làm ví dụ để thực hiện sáng chế khi được thực hiện bằng hệ thống xử lý dữ liệu được thực hiện bằng máy tính. Phương pháp này được mô tả bằng các thuật ngữ chung có thể áp dụng cho các dạng khác nhau của các bản ghi dữ liệu, dựa vào các ví dụ được nêu trong lĩnh vực liên quan đến các bản ghi dữ liệu viễn thông. Phương pháp này đưa hoặc nhập các bản ghi dữ liệu thô đầu vào vào bộ nhớ dữ liệu có thể sử dụng để sau đó phân tích các bản ghi này.

Phương pháp này có thể thu và xử lý (101) các bản ghi dữ liệu thô đầu vào bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện. Các bản

ghi dữ liệu thô đầu vào có thể là các bản ghi dữ liệu thô đầu vào ở quy mô lớn. Các bản ghi dữ liệu đầu vào ở dạng các bản ghi dữ liệu viễn thông có thể bao gồm các bản ghi dữ liệu thuộc loại thứ nhất là các bản ghi dữ liệu nhân khẩu học của khách hàng như các bản ghi biết khách hàng của bạn (KYC) và các bản ghi dữ liệu thuộc loại thứ hai là các bản ghi chi tiết cuộc gọi điện thoại (CDR).

Dữ liệu viễn thông thô là dữ liệu đầu vào của hệ thống có thể được cung cấp ở dạng văn bản (ví dụ, các tệp phân cách bằng dấu phẩy) hoặc ở dạng nhị phân. Các bản ghi CDR có tính chất giao dịch, các bản ghi này mô tả sự kiện diễn ra ở một thời điểm cụ thể và các thuộc tính của nó (ví dụ, dấu thời gian) là bất biến. Mặt khác, các bản ghi KYC đề cập đến các thuộc tính của thuê bao phần lớn là có giá trị tĩnh (ví dụ, ngày mà thuê bao tham gia vào mạng di động là một thông tin tĩnh) trong khi các thuộc tính mới có thể được biết trong tương lai đối với các bản ghi mới.

Phương pháp này có thể xử lý (102) các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh trước tiên khi các bản ghi này chứa thông tin nhận dạng của các thực thể được liên hệ với các bản ghi dữ liệu này. Trong ví dụ mà trong đó các bản ghi dữ liệu liên quan đến các sự kiện viễn thông, các thực thể có thể là những người dùng ở dạng là các khách hàng hoặc các thuê bao đối với các dịch vụ viễn thông di động và các bản ghi dữ liệu thuộc loại thứ nhất có thể là các bản ghi dữ liệu của khách hàng (ví dụ, các bản ghi KYC) có các giá trị tĩnh của các thuộc tính của người dùng. Hệ thống này có thể thực hiện nhiều lần chuyển dữ liệu nếu cần và cũng có thể kết hợp dữ liệu từ các nguồn khác nhau và theo các định dạng khác nhau.

Phương pháp này có thể tạo cấu hình (103) các trường của bản ghi dữ liệu thành định dạng tiêu chuẩn. Cách làm này có thể được áp dụng cho các bản ghi dữ liệu thuộc cả hai loại thứ nhất và thứ hai để bảo đảm rằng các trường sử dụng định dạng tiêu chuẩn, nhất quán. Ví dụ, cấu hình có thể được xác định riêng cho các MNO khác nhau và các mục cấu hình như vậy có (các) múi giờ ở quốc gia mà MNO cung cấp dịch vụ, tiền tệ được sử dụng, mã quốc gia, và các mục cấu hình khác.

Phương pháp này có thể tách ra (104) các giá trị cụ thể từ dữ liệu thô theo định dạng sử dụng được trong khi loại bỏ thông tin không cần thiết khi phân tích sau đó. Mục đích của bước này là để tách ra, theo định dạng sử dụng được, tất cả các giá trị cần thiết từ các

bản ghi đầu vào. Các giá trị bao gồm MSISDN (hoặc thông tin nhận dạng khác) của thuê bao, ngày kích hoạt SIM của thuê bao, dấu thời gian của sự kiện, khoảng thời gian của sự kiện, nếu có thể, số tiền liên quan, và các giá trị khác dành riêng cho sự kiện. Bước này có thể tách ra thông tin nhận dạng duy nhất của thực thể, như thông tin nhận dạng người dùng, nếu có trong các bản ghi dữ liệu thuộc loại thứ nhất.

Phương pháp này có thể lọc (105) các bản ghi dữ liệu để xác định loại sự kiện của mỗi bản ghi đầu vào bằng cách áp dụng các quy tắc lọc. Các quy tắc này được MNO cung cấp cùng với dữ liệu hoặc các quy tắc này được tạo ra sau khi phân tích dữ liệu. Loại sự kiện trong dữ liệu viễn thông có thể là sự kiện nạp tiền, ứng trước cho phép sử dụng mạng, hoàn trả, mua gói dịch vụ, chuyển tiền giữa các cá nhân (Person to Person, P2P), v.v.. Các bản ghi dữ liệu thô đầu vào có thể bao gồm các bản ghi không cần thiết khi phân tích sau đó (như các bản ghi liên quan đến các hoạt động nội bộ trong các hệ thống của MNO). Mục đích của bước này là để phân loại mỗi bản ghi và loại bỏ các bản ghi không cần thiết bất kỳ.

Phương pháp này có thể biến đổi (106) các bản ghi dữ liệu đã lọc thành định dạng tiêu chuẩn đối với mỗi loại bản ghi dữ liệu. Sau khi biến đổi, tất cả các bản ghi thuộc một loại cụ thể (ví dụ, tất cả các bản ghi nạp tiền) có cùng một sơ đồ (tức là, cùng một tập hợp trường) và tất cả các giá trị tuân theo một định dạng tiêu chuẩn hoá. Cùng một dạng kết hợp sơ đồ/định dạng được sử dụng cho tất cả các bản ghi thuộc một loại cụ thể từ tất cả các MNO. Các dạng biến đổi làm ví dụ có thể là:

- mã quốc gia thích hợp được thêm vào trước MSISDN;
- các dấu thời gian được biến đổi thành định dạng ISO 8601 sử dụng giờ UTC để làm múi giờ;
- số tiền được biến đổi thành đơn vị tiền tệ thích hợp;
- dữ liệu dành riêng cho sự kiện được biến đổi thành các giá trị tiêu chuẩn hoá.

Phương pháp này có thể tạo ra (107) thông tin nhận dạng duy nhất cho mỗi thực thể, nếu thông tin nhận dạng duy nhất này không thu được từ các giá trị dữ liệu tĩnh trong các bản ghi dữ liệu thuộc loại thứ nhất ở bước (104) nêu trên. Ví dụ, thông tin nhận dạng duy

nhất cho mỗi thực thể người dùng có thể là dạng kết hợp của số danh bạ thuê bao quốc tế của trạm di động (MSISDN) và ngày kích hoạt môđun nhận dạng thuê bao (SIM). Một mình số MSISDN là không đủ, vì có thể xảy ra trường hợp số này được sử dụng lại (tức là, sau khi thuê bao chấm dứt đăng ký thuê bao của họ, số MSISDN này có thể được cấp cho một thuê bao mới). Vì vậy, phương pháp này có thể tạo ra dạng kết hợp của MSISDN và ngày kích hoạt SIM, nếu có, dạng kết hợp này nhận dạng duy nhất mỗi thuê bao. Ngày kích hoạt SIM phải là thông tin có sẵn từ bước xử lý trước đó. Thông tin nhận dạng duy nhất đã tạo ra được bổ sung vào các bản ghi dữ liệu được lưu trữ.

Phương pháp này có thể lưu trữ (108) các bản ghi dữ liệu giao dịch vào bộ nhớ dữ liệu thích hợp để hỗ trợ một số lượng lớn bản ghi, cỡ hàng tỷ bản ghi hoặc nhiều hơn nữa.

Các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh (ví dụ, các bản ghi dữ liệu KYC thứ nhất) có thông tin nhận dạng duy nhất được tạo ra hoặc tách ra của các thực thể có thể được lưu trữ trong bộ nhớ dữ liệu để thực hiện chức năng tìm kiếm có hiệu quả dựa vào thông tin nhận dạng duy nhất của thực thể (ví dụ, MSISDN ghép với ngày kích hoạt SIM).

Các bản ghi dữ liệu thuộc loại thứ hai ở dạng các bản ghi giao dịch (ví dụ, các bản ghi dữ liệu CDR) có thể được lưu trữ trong bộ nhớ dữ liệu để thực hiện chức năng tìm kiếm có hiệu quả dựa vào dấu thời gian của sự kiện. Các bản ghi có thể được phân chia theo loại bản ghi. Để tìm kiếm thông tin có hiệu quả, các bản ghi có thể còn được phân chia theo ngày bằng cách sử dụng các phương pháp phân chia được thực hiện bằng bộ nhớ dữ liệu (ví dụ, nếu bộ nhớ dữ liệu chứa các tệp phân cách bằng dấu phẩy trong hệ thống tệp thông thường, thì phương pháp phân chia có thể được thực hiện bằng cách sử dụng tệp hoặc thư mục khác nhau theo từng ngày).

Phương pháp này có thể nhóm (109) các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất. Một loại có thể là các bản ghi giao dịch mở, đang diễn ra và một loại khác có thể là các bản ghi giao dịch đóng. Các bản ghi giao dịch đóng sẽ liên tục tăng kích thước, trong khi các bản ghi giao dịch mở thường là sẽ có kích thước ổn định vì đóng một số giao dịch và mở một số giao

dịch mới.

Trong ví dụ liên quan đến dữ liệu viễn thông, bước này có thể kiểm tra tất cả các khoản ứng trước và các khoản hoàn lại tín dụng của thuê bao để xác định khoản ứng trước nào đã được hoàn trả, được gọi là “đóng”, và khoản ứng trước nào chưa được hoàn trả, được gọi là “mở”. Các khoản ứng trước và các khoản hoàn trả có thể được tìm kiếm từ bộ nhớ dữ liệu và được nhóm với nhau theo mỗi thuê bao (sử dụng thông tin nhận dạng duy nhất của thuê bao). Dữ liệu đầu vào của bước tạo nhóm này bao gồm dữ liệu từ những lần thực hiện phương pháp này trước đó, tức là các khoản ứng trước bất kỳ vẫn đang mở. Mỗi nhóm như vậy được sắp xếp theo dấu thời gian của nó, từ sớm nhất đến muộn nhất.

Sau đó, mỗi nhóm được kiểm tra riêng biệt. Nếu bắt gặp một khoản ứng trước thì khoản ứng trước đó được ghi lại để kiểm tra thêm, còn nếu bắt gặp một khoản hoàn trả thì khoản hoàn trả đó được so khớp, một phần hoặc toàn phần, với khoản ứng trước đã ghi, theo các quy tắc do MNO thiết lập. Một quy tắc có thể có là khoản hoàn trả được so khớp với khoản ứng trước mở đầu tiên. Khoản ứng trước được so khớp toàn phần với các khoản hoàn trả sẽ được đánh dấu là khoản ứng trước đóng.

Ở cuối quy trình này, các khoản ứng trước vẫn đang mở được lưu trữ vào một bảng trong bộ nhớ dữ liệu và các khoản ứng trước đã đóng được lưu trữ vào một bảng khác. Theo một phương án thực hiện sáng chế, có thể có một bảng lưu trữ tất cả các khoản ứng trước mở, và theo một phương án khác, có thể có một bảng chứng từ lưu trữ một chứng từ cho mỗi thuê bao. Lý do phải lưu trữ phân tách như vậy là để bảo đảm tìm kiếm có hiệu quả các khoản ứng trước mở để sau đó phân tích. Số đếm của các khoản ứng trước mở là khá ổn định vì trong hoạt động thông thường, ở mỗi lần thực hiện của hệ thống, số đếm của các khoản ứng trước mới có cùng mức độ lớn với số đếm của các khoản ứng trước cũ được đóng lại trong lần thực hiện đó. Vì vậy, bảng lưu trữ các khoản ứng trước mở duy trì kích thước ổn định (điều này có nghĩa là truy nhập hiệu quả hơn) trong khi bảng lưu trữ các khoản ứng trước đóng có kích thước tăng lên sau mỗi lần thực hiện của hệ thống. Một bảng riêng biệt được sử dụng để cập nhật có hiệu quả bảng này. Vì chỉ cập nhật các khoản ứng trước mở, cho nên việc tìm kiếm và cập nhật nhanh hơn nếu có một bảng chỉ lưu trữ các khoản ứng trước mở.

Phương pháp này có thể tìm kiếm (110) các đặc điểm liên quan đến thực thể từ một hoặc nhiều bảng trong số nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó. Bước tạo ra các đặc điểm được thực hiện với các số đo của hành động của thực thể ở một thời điểm nhất định bao gồm các bước: tìm kiếm các đặc điểm từ các bảng khác nhau, liên kết các đặc điểm, và biến đổi các đặc điểm đã tách ra đến thời điểm tham chiếu.

Thông tin đầu ra liên quan đến thực thể có thể liên quan đến người dùng và có thể có một danh sách chứa các thông tin nhận dạng của thực thể và các hạn mức tín dụng có thông tin bổ sung tùy chọn. Thông tin đầu ra có thể là thông tin tách ra từ các profin của người vay tiềm năng, sự phát triển của các mô hình chấm điểm tín dụng, và các phương pháp chấm điểm tín dụng cho các thuê bao và cấp các hạn mức tín dụng.

Dựa vào Fig.2, lưu đồ (200) thể hiện khía cạnh xử lý các bản ghi dữ liệu theo phương pháp trên Fig.1 theo phương án làm ví dụ để thực hiện sáng chế.

Phương pháp này có thể thu (201) bản ghi dữ liệu giao dịch kế tiếp từ các bản ghi dữ liệu thô đầu vào và có thể tách ra (202) các giá trị bản ghi. Các quy tắc lọc có thể được áp dụng (203) và loại bản ghi được xác định (204).

Có thể xác định (205) xem bản ghi có được lưu trữ vào bộ nhớ dữ liệu hay không. Nếu bản ghi không được lưu trữ vào bộ nhớ dữ liệu, thì phương pháp này có thể lặp lại quy trình để xử lý bản ghi kế tiếp (201). Nếu bản ghi được lưu trữ vào bộ nhớ dữ liệu, thì bản ghi có thể được biến đổi (206) thành định dạng tiêu chuẩn và thông tin nhận dạng duy nhất có thể được sử dụng (207) cho bản ghi này. Thông tin nhận dạng duy nhất có thể được tạo ra (207) từ bản ghi nếu thông tin nhận dạng duy nhất này không có sẵn từ các giá trị bản ghi đã tách ra. Bản ghi có thể được lưu trữ (208) vào bộ nhớ dữ liệu.

Có thể xác định (209) xem có bản ghi khác cần xử lý hay không. Nếu có, thì phương pháp này có thể lặp lại quy trình để xử lý bản ghi kế tiếp (201). Nếu không còn bản ghi cần xử lý, thì phương pháp này có thể cập nhật (210) trạng thái đầu ra như trạng thái khoản vay đối với các thuê bao viễn thông.

Phương pháp này có thể tách ra các bản ghi sự kiện sau đây từ dữ liệu viễn thông.

- Từ các bản ghi CDR:

- Các sự kiện cuộc gọi điện thoại, đó là các sự kiện mà trong đó thuê bao của MNO thực hiện hoặc thu nhận cuộc gọi điện thoại.
- Các sự kiện nhắn tin theo dịch vụ thông báo ngắn (Short Message Service, SMS), đó là các sự kiện mà trong đó thuê bao của MNO gửi hoặc nhận SMS.
- Các sự kiện nạp tiền (còn được gọi là “nạp tiền”), đó là các sự kiện mà trong đó thuê bao dịch vụ di động trả trước dùng một số tiền để tăng số dư tài khoản thuê bao trả cho nhà điều hành mạng di động và làm như vậy để tiếp tục sử dụng dịch vụ trong tương lai.
- Chuyển tiền giữa các cá nhân (Person to Person, P2P), đó là các sự kiện mà trong đó thuê bao chuyển một phần số dư tài khoản thuê bao cho một thuê bao khác.
- Mua gói dịch vụ, đó là các sự kiện mà trong đó thuê bao mua sản phẩm được nhà điều hành chào bán để kết hợp nhiều dịch vụ (điện thoại, dữ liệu, v.v.) với dung lượng cụ thể của mỗi sản phẩm. Các sản phẩm như vậy thường được gọi là các gói dịch vụ.
- Kích hoạt gói dịch vụ, đó là các sự kiện mà trong đó thuê bao mua với số tiền từ tài khoản chính của thuê bao (và do đó số tiền này được ghi có cho MNO) sản phẩm được nhà điều hành chào bán để kết hợp nhiều dịch vụ (điện thoại, dữ liệu, v.v.) với dung lượng cụ thể của mỗi sản phẩm.
- Sự kiện sử dụng dữ liệu, đó là các sự kiện tiêu thụ dữ liệu.
- Các sự kiện liên quan đến chu kỳ sản phẩm, đó là các sự kiện về những thay đổi trạng thái của các thuê bao ở MNO.
- Các sự kiện ứng tiền trước sử dụng dịch vụ ứng tiền qua mạng, đó là các sự kiện mà trong đó dịch vụ ứng tiền/các gói dịch vụ được cấp cho thuê bao vượt quá số dư tài khoản hoặc hợp đồng của thuê bao. Các khoản ứng trước này được thuê bao trả lại ở sự kiện nạp tiền trong tương lai.
- Các sự kiện hoàn trả khoản ứng trước cho phép sử dụng mạng, đó là các sự kiện mà trong đó các khoản ứng trước cho phép sử dụng mạng được hoàn trả.

- Các giao dịch ví điện tử, đó là các sự kiện ghi nợ và ghi có trên ví di động của thuê bao.
- Các khoản vay trên ví di động, đó là các sự kiện mà trong đó một số tiền được ghi có vào ví di động của thuê bao với điều kiện được thanh toán sau thường kèm theo một khoản phí.
- Các sự kiện hoàn trả khoản vay trên ví di động, đó là các sự kiện mà trong đó các khoản vay trên ví di động được hoàn trả.
- Loại bản ghi bất kỳ khác có thể được tách ra từ các bản ghi CDR.
- Từ các bản ghi biết khách hàng của bạn: thông tin nhân khẩu học liên quan đến các thuê bao, như tên, ngày sinh, hoặc địa chỉ của thuê bao, và dữ liệu về mối quan hệ giữa thuê bao và mạng, như bảng giá được chọn hoặc ngày thuê bao.

Dựa vào dữ liệu viễn thông đầu vào, thông tin đầu ra của phương pháp này bao gồm một danh sách mà trong đó mỗi mục trong danh sách có một thông tin nhận dạng để nhận dạng duy nhất mỗi thuê bao liên quan đến hạn mức tín dụng của thuê bao đó, và thông tin bổ sung chứa đựng kết quả phân tích dẫn đến hạn mức tín dụng cụ thể. Phương pháp này có thể cung cấp một danh sách có độ dài bất kỳ, nằm trong khoảng từ một thuê bao đến hàng triệu thuê bao, và cung cấp thông tin đầu ra theo thời gian thực đối với các danh sách nhỏ và ngoại tuyến đối với các danh sách lớn.

Phương pháp được thể hiện trên Fig.1 mô tả phương pháp nhập dữ liệu có hiệu quả bao gồm bước đưa các bản ghi dữ liệu thô đầu vào được cung cấp vào trong bộ nhớ dữ liệu có thể sử dụng để phân tích. Phương pháp này có thể được sử dụng cho các ứng dụng khác nhau, một trong số đó là để tạo ra profin thực thể (bao gồm tạo ra profin người dùng), chấm điểm tín dụng và thiết lập hạn mức cho các thuê bao viễn thông.

Phương pháp tạo ra profin thực thể được mô tả trong ngữ cảnh là việc tạo ra profin người dùng để đưa ra các profin người dùng để sau đó phân tích rủi ro tín dụng và cấp hạn mức có thể được thực hiện bằng phương pháp được mô tả trong sáng chế. Profin người dùng là một tập hợp của các giá trị tĩnh và gộp chung thu được từ dữ liệu người dùng được xử lý trong phương pháp nhập dữ liệu đã được mô tả dựa vào Fig.1. Các giá trị này được gọi là các đặc điểm và mô tả hành động của người dùng ở thời điểm tham

chiếu nhất định. Phương pháp này có thể tạo ra các profin người dùng theo định dạng cần thiết để phân tích rủi ro tín dụng và cấp hạn mức.

Phương pháp tạo ra hàng ngàn đặc điểm mô tả hành động của những người dùng ở thời điểm tham chiếu nhất định. Các đặc điểm là các tập hợp gộp chung của các sự kiện giao dịch (nạp tiền, ứng tiền trước, v.v.), các đặc điểm theo phân loại và cơ bản dựa vào KYC, và một số dạng kết hợp của các đặc điểm này. Các đặc điểm này mô tả một số số đo của hành động của người dùng trong các khung thời gian khác nhau. Các số đo của hành động của người dùng là lần thực hiện gần đây nhất, yếu tố trung thành, tần suất, tổng số lần và thời lượng thực hiện. Các số đo còn là giá trị tối thiểu, giá trị tối đa, số đếm, giá trị trung bình, độ lệch chuẩn, giá trị trung vị, điểm tứ phân vị và xu hướng của số đếm, số lần và thời lượng thực hiện. Ngoài ra, các số đo là tỷ số, và số lượng trung bình và mức độ hoạt động. Các khung thời gian, mà ở đó các số đo được biểu hiện trước thời điểm tham chiếu là tuần, hai tuần, và tháng gần đây. Các khung thời gian còn là quý, sáu tháng và năm. Ngoài ra, một số đặc điểm mô tả các số đo trong suốt toàn bộ khoảng thời gian sử dụng dịch vụ/sản phẩm của thuê bao.

Bảng dưới đây thể hiện một tập hợp con nhỏ của các đặc điểm phù hợp với mô hình nạp tiền của người dùng.

Tên viết tắt	Mô tả
lrtd	Ngày nạp tiền cuối cùng (trước RD, hoặc ngày cuối của RM nếu không có RD)
lrdsTnr	Số ngày đã qua tính từ ngày nạp tiền cuối cùng
rvXma	Giá trị tiền nạp trong X tháng dương lịch trước
rvXtoYma	Tổng giá trị tiền nạp trong khoảng từ X đến Y tháng dương lịch trước
mrvXtoYma	Giá trị trung bình của các giá trị tiền nạp hằng tháng (dương lịch) trong khoảng từ X đến Y tháng trước
rvstdXtoYma	Độ lệch chuẩn của các giá trị tiền nạp hằng tháng (dương lịch) trong khoảng từ X đến Y tháng trước
rnXma	Số lần nạp tiền trong X tháng dương lịch trước
mrnXtoYma	Giá trị trung bình của số lần nạp tiền hằng tháng (dương lịch) trong khoảng từ X đến Y tháng trước
rnstdXtoYma	Độ lệch chuẩn của số lần nạp tiền hằng tháng (dương lịch) trong khoảng từ X đến Y tháng trước
mrxXma	Số tiền nạp tối đa trong X tháng dương lịch trước
mrxXtoYma	Số tiền nạp tối đa trong khoảng từ X đến Y tháng dương lịch trước
arvXma	Giá trị tiền nạp trung bình (giá trị nạp tiền trung bình) trong X tháng trước
arvXtoYma	Giá trị tiền nạp trung bình (giá trị nạp tiền trung bình) trong khoảng từ X đến Y tháng trước
mwrXtoYma	Các tháng có nạp tiền trong khoảng từ X đến Y tháng trước

Hầu hết các đặc điểm trên profin người dùng có số đo theo thời gian; ví dụ rv1ma, giá trị tiền nạp trong một tháng trước. Việc tìm kiếm các đặc điểm gộp chung ở chế độ khối (tức là đối với nhiều người dùng) sẽ nhanh hơn khi dữ liệu được phân chia (tức là được gán chỉ số theo thời gian) như đã được mô tả dựa vào Fig.1.

Phương pháp tạo ra profin người dùng bao gồm bước tạo cấu hình dành riêng cho mỗi nguồn dữ liệu, như MNO. Các mục cấu hình bao gồm múi giờ, danh sách gồm các đặc điểm có thể được tìm kiếm, và bộ nhớ dữ liệu và các bảng được tìm kiếm từ bộ nhớ dữ liệu. Để thực hiện tối ưu, một số đặc điểm có thể được tìm kiếm từ các bảng khác nhau.

Sau đó, phương pháp này tìm kiếm các đặc điểm từ dữ liệu đã được lưu trữ trong phương pháp nhập dữ liệu và liên kết các đặc điểm tìm được từ tất cả các bộ nhớ dữ liệu và các bảng khác nhau. Sau đó, các đặc điểm đã được tách ra được biến đổi đến thời điểm tham chiếu này.

Phương pháp chấm điểm tín dụng và thiết lập hạn mức tín dụng là một ví dụ về việc sử dụng các profin người dùng được tạo ra trên đây để phát triển các mô hình chấm điểm tín dụng, để chấm điểm tín dụng cho các thuê bao, và cấp các hạn mức tín dụng.

Các mô hình chấm điểm tín dụng dự báo với độ chính xác cao về xác suất xảy ra trường hợp thuê bao sẽ không trả được nợ hoặc dịch vụ tín dụng. Các mô hình chấm điểm tín dụng có thể được phát triển theo phương pháp thống kê, phương pháp máy tự học và trí tuệ nhân tạo và được chọn dựa vào hiệu suất của mô hình. Một mô hình có thể được phát triển cho toàn bộ cơ sở của MNO, hoặc các mô hình khác nhau có thể được phát triển cho các đoạn cơ sở khác nhau.

Sau đó, mô hình chấm điểm tín dụng có thể được sử dụng để tính điểm tín dụng của mỗi thuê bao và cấp hạn mức tín dụng thích hợp. Dựa vào điểm tín dụng, nhiều cách có thể được sử dụng để cấp hạn mức tín dụng, với cơ sở hợp lý là, mọi thứ khác đều bình đẳng, các thuê bao có điểm tín dụng cao hơn (và vì vậy khả năng trả được nợ cao hơn) sẽ nhận được hạn mức tín dụng cao hơn, trong khi các thuê bao có điểm tín dụng thấp hơn sẽ nhận được hạn mức tín dụng thấp hơn, và các thuê bao có điểm tín dụng thấp nhất có thể không được phép vay.

Ví dụ về phương pháp cấp hạn mức tín dụng dựa vào điểm tín dụng là phương pháp dựa vào ngưỡng. Phương pháp này có một danh sách sắp xếp theo thứ tự gồm các ngưỡng điểm số và các hạn mức tín dụng liên quan. Điểm số của thuê bao được kiểm tra dựa vào danh sách này và hạn mức tín dụng liên quan đến ngưỡng cao nhất nhỏ hơn so với điểm số đó được sử dụng để làm hạn mức tín dụng của thuê bao này.

Dựa vào Fig.3, sơ đồ khối thể hiện hệ thống được mô tả trong sáng chế (300) để xử lý các bản ghi dữ liệu theo phương án làm ví dụ để thực hiện sáng chế. Hệ thống (300) bao gồm ít nhất một bộ xử lý (301) hoặc nhiều bộ xử lý chạy song song và bộ nhớ (302) để lưu trữ mã chương trình đọc được bằng máy tính ở dạng các lệnh máy tính (303).

Hệ thống (300) bao gồm hệ thống nhập dữ liệu (310), bộ phận profin thực thể (330) có các thông tin đầu ra đến các ứng dụng như bộ phận phát triển mô hình (340) và bộ phận chấm điểm tín dụng và cấp hạn mức tín dụng (350).

Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận thu dữ liệu (311) để thu và xử lý các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện. Các bản ghi dữ liệu thô có thể là các bản ghi dữ liệu thô ở quy mô lớn, ví dụ các bản ghi có thể được thu với số lượng lớn (ví dụ từ hàng trăm triệu đến hàng tỷ bản ghi, hoặc nhiều hơn). Theo phương án làm ví dụ để thực hiện sáng chế, các bản ghi dữ liệu thô liên quan đến dữ liệu viễn thông. Các bản ghi dữ liệu thô có thể không có cùng một định dạng hoặc không có định dạng được xác định rõ ràng vì các bản ghi dữ liệu khác nhau có thể được tạo định dạng khác nhau cho nên các kiểu hoặc các loại dữ liệu tương tự hoặc các đặc điểm tương tự có trong các bản ghi dữ liệu khác nhau được ghi khác nhau khiến cho việc tách trực tiếp một loại hoặc đặc điểm của dữ liệu không thực hiện được.

Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận tạo cấu hình (312) để tạo cấu hình các trường của bản ghi dữ liệu theo các định dạng tiêu chuẩn và bộ phận tách (313) để tách ra các giá trị cụ thể từ dữ liệu thô theo định dạng sử dụng được.

Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận lọc (315) để lọc các bản ghi dữ liệu thích hợp bằng cách áp dụng các quy tắc lọc để tách ra các bản ghi thích hợp từ các bản ghi giao dịch và bộ phận biến đổi (314) để biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá.

Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận lưu trữ vào bộ nhớ dữ liệu (316) để lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau (320, 321) để hỗ trợ một số lượng lớn bản ghi. Bộ nhớ dữ liệu thứ nhất (320) dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai (321) dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện.

Bộ nhớ dữ liệu thứ nhất có thể được tạo cấu hình để tìm kiếm có hiệu quả dựa vào các giá trị tĩnh (như thông tin nhận dạng duy nhất). Ví dụ, bộ nhớ dữ liệu thứ nhất có thể được

tạo cấu hình để lưu trữ thông tin bằng cách sử dụng thông tin nhận dạng duy nhất của thực thể làm từ khoá để cho phép tìm kiếm có hiệu quả dữ liệu tĩnh của các thực thể.

Bộ nhớ dữ liệu thứ hai có thể được tạo cấu hình để tìm kiếm có hiệu quả dựa vào các dấu thời gian của các sự kiện. Ví dụ, theo một phương án thực hiện sáng chế, bộ nhớ dữ liệu thứ hai có thể được tạo cấu hình để lưu trữ thông tin bằng cách sử dụng dấu thời gian của các dữ liệu giao dịch làm từ khoá. Việc lưu trữ thông tin chỉ sử dụng dấu thời gian làm từ khoá có thể cho phép tìm kiếm có hiệu quả dữ liệu giao dịch của một nhóm thực thể, hoặc tất cả các thuê bao. Ví dụ, theo một phương án khác để thực hiện sáng chế, bộ nhớ dữ liệu thứ hai có thể được tạo cấu hình để lưu trữ thông tin bằng cách sử dụng dấu thời gian của các dữ liệu giao dịch và thông tin nhận dạng duy nhất của thực thể làm từ khoá (ví dụ từ khoá kết hợp). Điều này có thể cho phép tìm kiếm có hiệu quả dữ liệu giao dịch của mỗi thực thể. Trong mọi trường hợp, việc tìm kiếm dữ liệu giao dịch của mỗi thực thể, hoặc các thực thể, ở các thời điểm khác nhau và khi dữ liệu giao dịch mới được bổ sung liên tục, và việc tạo ra các đặc điểm sau đó dựa vào dữ liệu giao dịch này, có thể là rất hiệu quả.

Các bộ nhớ dữ liệu có thể được thực hiện dưới dạng là các cơ sở dữ liệu (ví dụ các cơ sở dữ liệu quan hệ) và các từ khoá có thể tạo ra cơ chế sử dụng phần mềm để nhận dạng, truy nhập và cập nhật thông tin trong bảng cơ sở dữ liệu.

Bộ nhớ dữ liệu thứ hai (321) có thể được phân vùng cho các khoảng thời gian và/hoặc theo loại sự kiện.

Cấu hình của bộ nhớ dữ liệu có thể là sơ đồ sắp xếp theo mức kiến trúc để hoạt động bất kể dữ liệu thực tế đang được xử lý (nhưng chỉ yêu cầu là dữ liệu này phải có hai loại hoặc đặc trưng khác nhau, là dữ liệu tĩnh và dữ liệu giao dịch). Nói cách khác, không quan trọng các giá trị dữ liệu cụ thể đang được xử lý. Bằng cách tạo ra hai bộ nhớ dữ liệu khác nhau được tạo cấu hình khác nhau để tìm kiếm có hiệu quả các loại dữ liệu khác nhau, hệ thống có sự cải tiến về hiệu quả xử lý dữ liệu bất kể các loại dữ liệu cụ thể có thể được cung cấp. Do đó, các khía cạnh của sáng chế có thể được áp dụng trong việc xử lý dữ liệu chung liên quan đến quy trình xử lý dữ liệu tĩnh và dữ liệu giao dịch.

Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận tạo nhóm (318) để nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của thực thể và bộ phận lưu trữ vào bảng (319) có thể được tạo ra để lưu trữ các loại bản ghi giao dịch

vào các bảng khác nhau trong đó duy trì một bảng (322) có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất. Ví dụ, một loại có thể là các bản ghi giao dịch mở, đang diễn ra và một loại khác có thể là các bản ghi giao dịch đóng. Các bản ghi giao dịch đóng sẽ liên tục tăng kích thước, trong khi các bản ghi giao dịch mở thường là sẽ có kích thước ổn định vì đóng một số giao dịch và mở một số giao dịch mới.

Số lượng giao dịch có trong dữ liệu giao dịch có thể là rất lớn và có thể tăng với tốc độ hàng trăm triệu giao dịch mỗi ngày. Việc duy trì các bảng có kích thước tương đối ổn định và không tăng với tốc độ đó, dành cho các giao dịch được tìm kiếm riêng biệt và thường xuyên nhất, cho phép truy nhập các giao dịch đó có hiệu quả hơn. Ví dụ, theo một phương án thực hiện sáng chế, khi các khoản ứng trước sử dụng dịch vụ ứng tiền qua mạng và các khoản hoàn trả được nhập vào từ dữ liệu viễn thông, thì việc tách ra dữ liệu giao dịch và sau đó tạo ra các đặc điểm, đòi hỏi phải tìm kiếm có hiệu quả rất cao các khoản ứng trước chưa được hoàn trả. Việc duy trì các bảng riêng biệt dành cho các khoản ứng trước đã được hoàn trả và các khoản ứng trước chưa được hoàn trả, cho phép tìm kiếm rất nhanh các khoản ứng trước chưa được hoàn trả, vì kích thước của chúng tăng với tốc độ rất chậm là nhờ việc duy trì một cách có hiệu quả một bảng có kích thước tương đối ổn định để lưu trữ các khoản ứng trước sử dụng dịch vụ ứng tiền qua mạng chưa được hoàn trả.

Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận tạo ra thông tin nhận dạng duy nhất (317) để tạo ra thông tin nhận dạng duy nhất của các thực thể trong các bản ghi dữ liệu, nếu thông tin nhận dạng duy nhất này không được tách ra bằng bộ phận tách (313). Ví dụ, các thực thể có thể là các thuê bao viễn thông. Hệ thống nhập dữ liệu (310) có thể bao gồm bộ phận cập nhật trạng thái (360) để cập nhật các kết quả trạng thái để tiếp tục xử lý các bản ghi dữ liệu.

Bộ phận profin thực thể (330) có thể tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng trong các bộ nhớ dữ liệu (320, 321). Bộ phận profin thực thể (330) có thể tạo ra các profin người dùng khi thực thể là người dùng. Bộ phận profin thực thể (330) để tìm kiếm các đặc điểm liên quan đến thực thể tạo ra các đặc điểm có các số đo của hành động của thực thể ở thời điểm nhất định và, theo một phương án thực hiện sáng chế, bao gồm bộ phận tìm kiếm đặc điểm (331) để tìm kiếm các đặc điểm từ các bảng khác nhau, bộ phận

liên kết đặc điểm (332) để liên kết các đặc điểm, và bộ phận biến đổi đặc điểm (333) để biến đổi các đặc điểm đến thời điểm tham chiếu.

Bộ phận profin thực thể (330) có thể cung cấp thông tin đầu ra từ bộ phận đầu ra (334) liên quan đến thực thể. Thông tin đầu ra liên quan đến thực thể có thể là profin người dùng để sau đó phân tích rủi ro tín dụng bằng cách sử dụng bộ phận phát triển mô hình (340) và bộ phận chấm điểm tín dụng và cấp hạn mức tín dụng (350).

Fig.4 là hình vẽ thể hiện thiết bị máy tính làm ví dụ (400) trong đó các khía cạnh khác nhau của sáng chế có thể được thực hiện. Thiết bị máy tính (400) có thể được thực hiện dưới dạng bất kỳ của thiết bị xử lý dữ liệu bao gồm thiết bị máy tính cá nhân (ví dụ máy tính xách tay hoặc máy tính để bàn), máy chủ (có thể là máy chủ độc lập, được phân bố ở nhiều vị trí vật lý), máy khách, hoặc thiết bị truyền thông, như máy điện thoại di động (ví dụ điện thoại di động), máy điện thoại vệ tinh, máy tính dạng bảng, thiết bị kỹ thuật số hỗ trợ cá nhân hoặc các thiết bị khác. Thiết bị máy tính theo các phương án khác nhau có thể được hiểu là có hoặc không có nhiều bộ phận hoặc hệ thống con khác nhau như được mô tả dưới đây.

Thiết bị máy tính (400) có thể phù hợp để lưu trữ và thực hiện mã chương trình máy tính. Nhiều thành phần và bộ phận khác nhau trong các sơ đồ thể hiện hệ thống đã được mô tả trên đây có thể sử dụng các bộ phận hoặc hệ thống con với số lượng bất kỳ trong thiết bị máy tính (400) để tạo điều kiện thuận lợi cho các chức năng được mô tả trong sáng chế. Thiết bị máy tính (400) có thể có các bộ phận hoặc hệ thống con được liên kết với nhau thông qua cơ sở hạ tầng truyền thông (405) (ví dụ, bus truyền thông, mạng, v.v.). Thiết bị máy tính (400) có thể bao gồm một hoặc nhiều bộ xử lý (410) và ít nhất một bộ phận lưu trữ ở dạng vật ghi đọc được bằng máy tính. Một hoặc nhiều bộ xử lý (410) có thể có một hoặc nhiều bộ xử lý trong số: bộ xử lý trung tâm (Central Processing Unit, CPU), bộ xử lý đồ họa (Graphical Processing Unit, GPU), bộ vi xử lý, mảng cửa lập trình được bằng trường (Field Programmable Gate Array, FPGA), mạch tích hợp chuyên dụng (Application Specific Integrated Circuit, ASIC) và các bộ xử lý khác. Trong một số cấu hình, nhiều bộ xử lý có thể được bố trí và có thể được sắp xếp để thực hiện các phép tính đồng thời. Theo một số phương án thực hiện sáng chế, nhiều bộ phận và hệ thống con khác nhau trong thiết bị máy tính (400) có thể được phân bố ở nhiều vị trí vật lý (ví dụ

theo cấu hình phân tán, theo nhóm hoặc điện toán đám mây) và các bộ phận phần mềm thích hợp có thể được sắp xếp để quản lý và/hoặc xử lý dữ liệu thay mặt cho các thiết bị ở xa.

Các bộ phận lưu trữ có thể bao gồm bộ nhớ hệ thống (415), bộ nhớ hệ thống này có thể có bộ nhớ chỉ đọc (Read Only Memory, ROM) và bộ nhớ truy nhập ngẫu nhiên (Random Access Memory, RAM). Hệ thống nhập/xuất cơ sở (Basic Input/Output System, BIOS) có thể được lưu trữ trong bộ nhớ ROM. Phần mềm hệ thống có thể được lưu trữ trong hệ thống bộ nhớ (415) có phần mềm hệ điều hành. Các bộ phận lưu trữ cũng có thể bao gồm bộ nhớ phụ (420). Bộ nhớ phụ (420) có thể có đĩa cố định (421), như ổ đĩa cứng, và, tùy ý có, một hoặc nhiều giao diện bộ nhớ (422) để giao diện với các bộ phận lưu trữ (423), như bộ phận lưu trữ tháo lắp được (ví dụ băng từ, đĩa quang, ổ nhớ tác động nhanh, ổ cứng bên ngoài, chip nhớ tháo lắp được, v.v.), các bộ phận lưu trữ được kết nối qua mạng (ví dụ các ổ đĩa NAS), các bộ phận lưu trữ từ xa (ví dụ bộ nhớ đám mây) hoặc các loại bộ nhớ khác.

Thiết bị máy tính (400) có thể bao gồm giao diện truyền thông bên ngoài (430) cho sự hoạt động của thiết bị máy tính (400) trong môi trường nối mạng cho phép truyền dữ liệu giữa nhiều thiết bị máy tính (400) và/hoặc mạng internet. Dữ liệu được truyền qua giao diện truyền thông bên ngoài (430) có thể ở dạng các tín hiệu, có thể là tín hiệu điện tử, tín hiệu điện từ, tín hiệu quang học, tín hiệu vô tuyến, hoặc các loại tín hiệu khác. Giao diện truyền thông bên ngoài (430) có thể cho phép truyền thông dữ liệu giữa thiết bị máy tính (400) và các thiết bị máy tính khác gồm có các máy chủ và các phương tiện lưu trữ bên ngoài. Các dịch vụ web có thể truy nhập được bằng và/hoặc từ thiết bị máy tính (400) thông qua giao diện truyền thông (430).

Giao diện truyền thông bên ngoài (430) có thể được tạo cấu hình để kết nối với các kênh truyền thông không dây (ví dụ, mạng điện thoại di động, mạng cục bộ không dây (ví dụ sử dụng liên kết truyền thông theo tiêu chuẩn Wi-Fi™), mạng điện thoại vệ tinh, mạng internet vệ tinh, v.v.) và có thể có bộ phận truyền dẫn không dây kèm theo, như anten và mạch kèm theo.

Vật ghi đọc được bằng máy tính ở dạng nhiều bộ phận lưu trữ khác nhau có thể thực hiện chức năng lưu trữ các lệnh thực hiện được bằng máy tính, các cấu trúc dữ liệu, các

môđun chương trình, các bộ phận phần mềm và các dữ liệu khác. Sản phẩm chương trình máy tính có thể được cung cấp thông qua vật ghi đọc được bằng máy tính lưu trữ mã chương trình đọc được bằng máy tính thực hiện được bằng bộ xử lý trung tâm (410). Sản phẩm chương trình máy tính có thể được cung cấp thông qua vật ghi bất khả biến đọc được bằng máy tính, hoặc có thể được cung cấp thông qua tín hiệu hoặc phương tiện chuyển tiếp khác thông qua giao diện truyền thông (430).

Sự kết nối với nhau thông qua cơ sở hạ tầng truyền thông (405) cho phép một hoặc nhiều bộ xử lý (410) truyền thông với mỗi bộ phận hoặc hệ thống con và điều khiển việc thực hiện các lệnh từ các bộ phận lưu trữ, cũng như trao đổi thông tin giữa các bộ phận hoặc hệ thống con. Các thiết bị ngoại vi (như máy in, máy quét, camera, hoặc các thiết bị ngoại vi khác) và các thiết bị nhập/xuất (Input/Output, I/O) (như chuột, vùng phím, bàn phím, micrô, màn hình cảm ứng, các nút nhập, loa và các thiết bị nhập/xuất khác) có thể kết nối với hoặc được chế tạo liền khối với thiết bị máy tính (400) một cách trực tiếp hoặc thông qua bộ điều khiển nhập/xuất (435). Một hoặc nhiều màn hình (445) (có thể là các màn hình cảm ứng) có thể được kết nối với hoặc được chế tạo liền khối với thiết bị máy tính (400) thông qua màn hình (445) hoặc bộ phận thích ứng video (440).

Phần mô tả sáng chế trên đây được mô tả nhằm mục đích làm ví dụ; phần này không được coi là nhằm mục đích loại trừ hoặc giới hạn phạm vi của sáng chế ở các phương án cụ thể được mô tả trong sáng chế. Người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng có thể hiểu rằng nhiều phương án cải biến và thay đổi có thể được thực hiện dựa vào phần mô tả sáng chế trên đây.

Mọi bước, thao tác, bộ phận hoặc quy trình được mô tả trong sáng chế có thể được thực hiện hoặc thể hiện bằng một hoặc nhiều bộ phận phần cứng hoặc phần mềm, ở dạng một mình hoặc kết hợp với các thiết bị khác. Theo một phương án thực hiện sáng chế, bộ phận phần mềm được thể hiện bằng sản phẩm chương trình máy tính có vật ghi bất khả biến đọc được bằng máy tính chứa mã chương trình máy tính, có thể được thực hiện bằng bộ xử lý để thực hiện một hoặc tất cả các bước, thao tác, hoặc quy trình được mô tả trong sáng chế. Các bộ phận phần mềm hoặc các chức năng được mô tả trong sáng chế này có thể được thể hiện dưới dạng mã chương trình máy tính sử dụng ngôn ngữ máy tính phù hợp bất kỳ, ví dụ như Java™, C++, hoặc Perl™ sử dụng, ví dụ, kỹ thuật lập trình thông

thường hoặc lập trình hướng đối tượng. Mã chương trình máy tính có thể được lưu trữ dưới dạng là một dãy lệnh, hoặc các lệnh điều khiển trên vật ghi bất khả biến đọc được bằng máy tính, như bộ nhớ truy nhập ngẫu nhiên (RAM), bộ nhớ chỉ đọc (ROM), vật ghi từ tính như ổ đĩa cứng, hoặc vật ghi quang học như đĩa compac-bộ nhớ chỉ đọc (Compact Disc-Read Only Memory, CD-ROM). Mọi vật ghi đọc được bằng máy tính như vậy cũng có thể nằm ở trên hoặc ở trong một thiết bị máy tính, và có thể nằm ở trên hoặc ở trong các thiết bị máy tính khác nhau trong một hệ thống hoặc mạng.

Các lưu đồ và các sơ đồ khối thể hiện các phương pháp, các hệ thống, và các sản phẩm chương trình máy tính theo các phương án thực hiện sáng chế được sử dụng trong sáng chế. Mỗi khối trong các lưu đồ và/hoặc các sơ đồ khối, và các dạng kết hợp của các khối trong các lưu đồ và/hoặc các sơ đồ khối, có thể thực hiện các chức năng có thể thực hiện được bằng các lệnh chương trình đọc được bằng máy tính. Theo một số phương án thực hiện khác, các chức năng được thể hiện bằng các khối có thể được thực hiện theo thứ tự khác với thứ tự được thể hiện trên các lưu đồ.

Một số phần trong bản mô tả sáng chế này mô tả các phương án thực hiện sáng chế dựa vào các thuật toán và các dạng biểu diễn ký hiệu của các thao tác trên thông tin. Các thuật toán và các dạng biểu diễn dùng để mô tả như vậy thường được người có hiểu biết trung bình về lĩnh vực kỹ thuật xử lý dữ liệu sử dụng để truyền tải nội dung công việc của họ một cách có hiệu quả đến những người có hiểu biết trung bình về lĩnh vực kỹ thuật tương ứng. Các thao tác này, dù được mô tả ở dạng chức năng, tính toán, hay logic, được hiểu là sẽ được thực hiện bằng các chương trình máy tính hoặc các mạch điện tương đương, vi mã, hoặc các dạng tương tự khác. Các thao tác được mô tả trong sáng chế có thể được thể hiện ở dạng phần mềm, phần sụn, phần cứng, hoặc các dạng kết hợp bất kỳ của các loại nêu trên.

Từ ngữ được sử dụng trong bản mô tả sáng chế này được chọn theo nguyên tắc nhằm mục đích hướng dẫn và dễ hiểu, và từ ngữ có thể không được chọn để mô tả hoặc giới hạn phạm vi của đối tượng yêu cầu bảo hộ của sáng chế. Vì vậy cần phải hiểu rằng phạm vi của sáng chế không bị giới hạn ở phần mô tả chi tiết sáng chế, mà được xác định dựa vào các điểm yêu cầu bảo hộ kèm theo của đơn đăng ký sáng chế này. Do đó, phần mô tả các phương án thực hiện sáng chế được hiểu là nhằm mục đích làm ví dụ, chứ

không nhằm mục đích giới hạn phạm vi của sáng chế, phạm vi của sáng chế được xác định dựa vào các điểm yêu cầu bảo hộ kèm theo.

Cuối cùng, trong toàn bộ phần mô tả sáng chế và phần yêu cầu bảo hộ, trừ trường hợp nội dung có yêu cầu khác, từ ‘bao gồm’ hoặc các biến thể của từ này như ‘có’ hoặc ‘gồm có’ sẽ được hiểu là đề cập đến sự có mặt của một phần tử nguyên hoặc một nhóm phần tử nguyên được nêu nhưng không loại trừ phần tử nguyên hoặc nhóm phần tử nguyên bất kỳ khác.

YÊU CẦU BẢO HỘ

1. Phương pháp được thực hiện bằng máy tính để xử lý các bản ghi dữ liệu ở quy mô lớn, bao gồm các bước:

thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện;

lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá;

lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện;

nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất; và

tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

2. Phương pháp theo điểm 1, trong đó bước tìm kiếm các đặc điểm liên quan đến thực thể là tạo ra các đặc điểm có các số đo của hành động của thực thể ở thời điểm tham chiếu nhất định, và bao gồm các bước:

tìm kiếm các đặc điểm từ các bảng khác nhau;

liên kết các đặc điểm; và

biến đổi các đặc điểm đến thời điểm tham chiếu này.

3. Phương pháp theo điểm 1, trong đó bộ nhớ dữ liệu thứ hai được phân vùng cho các khoảng thời gian và/hoặc theo loại sự kiện.

4. Phương pháp theo điểm 1, trong đó các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất là các giao dịch mở đang diễn ra.

5. Phương pháp theo điểm 1, trong đó phương pháp này xử lý các bản ghi dữ liệu tĩnh trước tiên để thu được thông tin nhận dạng của các thực thể.

6. Phương pháp theo điểm 1, trong đó các bản ghi dữ liệu thô liên quan đến dữ liệu viễn thông và các thực thể là những người dùng, và trong đó các bản ghi dữ liệu thuộc loại thứ nhất là các bản ghi dữ liệu người dùng có các giá trị tĩnh của các thuộc tính của người dùng và các bản ghi dữ liệu thuộc loại thứ hai là các bản ghi chi tiết cuộc gọi điện thoại của các sự kiện.

7. Phương pháp theo điểm 6, trong đó phương pháp này bao gồm bước tạo ra thông tin nhận dạng duy nhất cho mỗi người dùng là dạng kết hợp của số danh bạ thuê bao quốc tế của trạm di động (Mobile Station International Subscriber Directory Number, MSISDN) và ngày kích hoạt môđun nhận dạng thuê bao (Subscriber Identity Module, SIM).

8. Phương pháp theo điểm 6, trong đó bước lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá bao gồm bước áp dụng các quy tắc lọc để tách ra các bản ghi thích hợp từ các bản ghi giao dịch có một hoặc nhiều nhóm gồm: các bản ghi liên quan đến các sự kiện như: các cuộc gọi điện thoại, các tin nhắn, các sự kiện giao dịch tiền tệ, các sự kiện sử dụng dữ liệu; các sự kiện liên quan đến chu kỳ sản phẩm; các sự kiện sử dụng khoản ứng trước; các sự kiện liên quan đến khoản vay; các sự kiện liên quan đến ví di động.

9. Phương pháp theo điểm 6, trong đó bước nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của thực thể bao gồm bước nhóm các loại giao dịch ứng trước tín dụng mở và các loại giao dịch ứng trước tín dụng đóng, và các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất là các giao dịch ứng trước tín dụng mở có sẵn để sau đó phân tích.

10. Phương pháp theo điểm 1, trong đó thông tin đầu ra liên quan đến thực thể là profin người dùng để sau đó phân tích rủi ro tín dụng.

11. Hệ thống xử lý các bản ghi dữ liệu ở quy mô lớn, hệ thống này bao gồm bộ nhớ để lưu trữ mã chương trình đọc được bằng máy tính và bộ xử lý để thực hiện mã chương trình đọc được bằng máy tính, bao gồm:

bộ phận thu dữ liệu để thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện;

bộ phận biến đổi để biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá;

bộ phận lọc để lọc các bản ghi dữ liệu thích hợp;

bộ phận lưu trữ vào bộ nhớ dữ liệu để lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện;

bộ phận tạo nhóm để nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và bộ phận lưu trữ vào bảng để lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất; và

bộ phận profin thực thể để tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

12. Hệ thống theo điểm 11, trong đó bộ phận profin thực thể để tìm kiếm các đặc điểm liên quan đến thực thể tạo ra các đặc điểm có các số đo của hành động của thực thể ở thời điểm tham chiếu nhất định, bao gồm:

bộ phận tìm kiếm đặc điểm để tìm kiếm các đặc điểm từ các bảng khác nhau;

bộ phận liên kết đặc điểm để liên kết các đặc điểm; và

bộ phận biến đổi đặc điểm để biến đổi các đặc điểm đến thời điểm tham chiếu này.

13. Hệ thống theo điểm 11, trong đó bộ nhớ dữ liệu thứ hai được phân vùng cho các khoảng thời gian và/hoặc theo loại sự kiện.

14. Hệ thống theo điểm 11, trong đó các bản ghi giao dịch thuộc loại được truy nhập

thường xuyên nhất là các giao dịch mở đang diễn ra.

15. Hệ thống theo điểm 11, trong đó các bản ghi dữ liệu thô liên quan đến dữ liệu viễn thông và các thực thể là những người dùng, và trong đó các bản ghi dữ liệu thuộc loại thứ nhất là các bản ghi dữ liệu người dùng có các giá trị tính của các thuộc tính của người dùng và các bản ghi dữ liệu thuộc loại thứ hai là các bản ghi chi tiết cuộc gọi điện thoại của các sự kiện.

16. Hệ thống theo điểm 15, trong đó hệ thống này bao gồm bộ phận tạo ra thông tin nhận dạng duy nhất để tạo ra thông tin nhận dạng duy nhất cho mỗi người dùng là dạng kết hợp của số danh bạ thuê bao quốc tế của trạm di động (MSISDN) và ngày kích hoạt môđun nhận dạng thuê bao (SIM).

17. Hệ thống theo điểm 15, trong đó bộ phận lọc áp dụng các quy tắc lọc để tách ra các bản ghi thích hợp từ các bản ghi giao dịch có một hoặc nhiều nhóm gồm: các bản ghi liên quan đến các sự kiện như: các cuộc gọi điện thoại, các tin nhắn, các sự kiện giao dịch tiền tệ, các sự kiện sử dụng dữ liệu; các sự kiện liên quan đến chu kỳ sản phẩm; các sự kiện sử dụng khoản ứng trước; các sự kiện liên quan đến khoản vay; các sự kiện liên quan đến ví di động.

18. Hệ thống theo điểm 15, trong đó bộ phận tạo nhóm nhóm các loại giao dịch ứng trước tín dụng mở và các loại giao dịch ứng trước tín dụng đóng, và các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất là các giao dịch ứng trước tín dụng mở có sẵn để sau đó phân tích.

19. Vật ghi đọc được bằng máy tính lưu trữ mã chương trình đọc được bằng máy tính để thực hiện các bước:

thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tính và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch có các dấu thời gian của các sự kiện;

lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá;

lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm

bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện;

nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất; và

tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra liên quan đến thực thể đó.

20. Phương pháp được thực hiện bằng máy tính để xử lý các bản ghi dữ liệu viễn thông ở quy mô lớn, bao gồm các bước:

thu các bản ghi dữ liệu thô bao gồm các bản ghi dữ liệu thuộc loại thứ nhất có các giá trị tĩnh liên quan đến thực thể người dùng và các bản ghi dữ liệu thuộc loại thứ hai liên quan đến các bản ghi giao dịch của các sự kiện viễn thông có các dấu thời gian của các sự kiện;

lọc và biến đổi các bản ghi dữ liệu thành các định dạng tiêu chuẩn hoá;

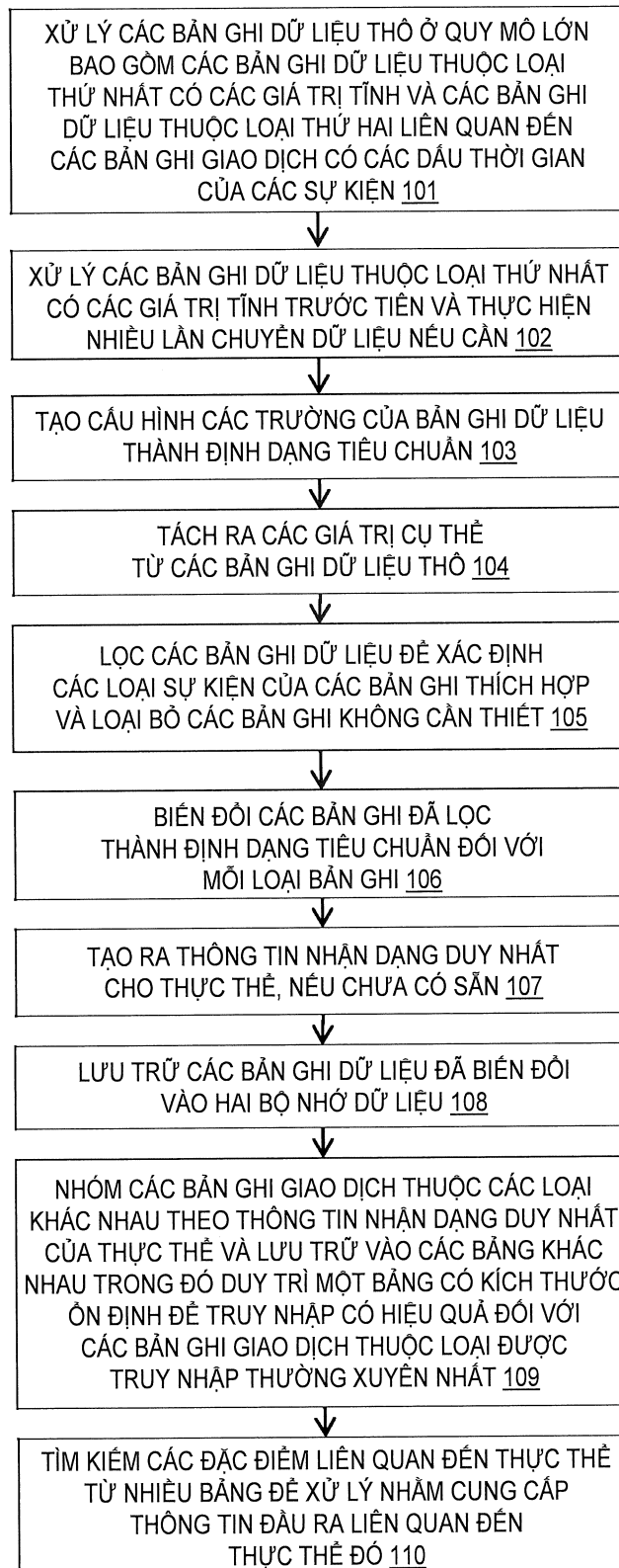
lưu trữ các bản ghi dữ liệu đã biến đổi vào hai bộ nhớ dữ liệu khác nhau, bao gồm bộ nhớ dữ liệu thứ nhất dùng cho dữ liệu tĩnh liên quan đến các thực thể để tìm kiếm dựa vào các giá trị tĩnh có thông tin nhận dạng duy nhất của thực thể, và bộ nhớ dữ liệu thứ hai dùng cho dữ liệu giao dịch để tìm kiếm dựa vào các dấu thời gian của các sự kiện;

nhóm các bản ghi giao dịch thuộc các loại khác nhau theo thông tin nhận dạng duy nhất của các thực thể và lưu trữ các loại bản ghi giao dịch vào các bảng khác nhau trong đó duy trì một bảng có kích thước ổn định để truy nhập có hiệu quả đối với các bản ghi giao dịch thuộc loại được truy nhập thường xuyên nhất liên quan đến các giao dịch ứng trước tín dụng mở cho thực thể; và

tìm kiếm các đặc điểm liên quan đến thực thể từ nhiều bảng để xử lý nhằm cung cấp thông tin đầu ra ở dạng điểm tín dụng liên quan đến thực thể đó.

100 →

1/4

**FIG. 1**

2/4

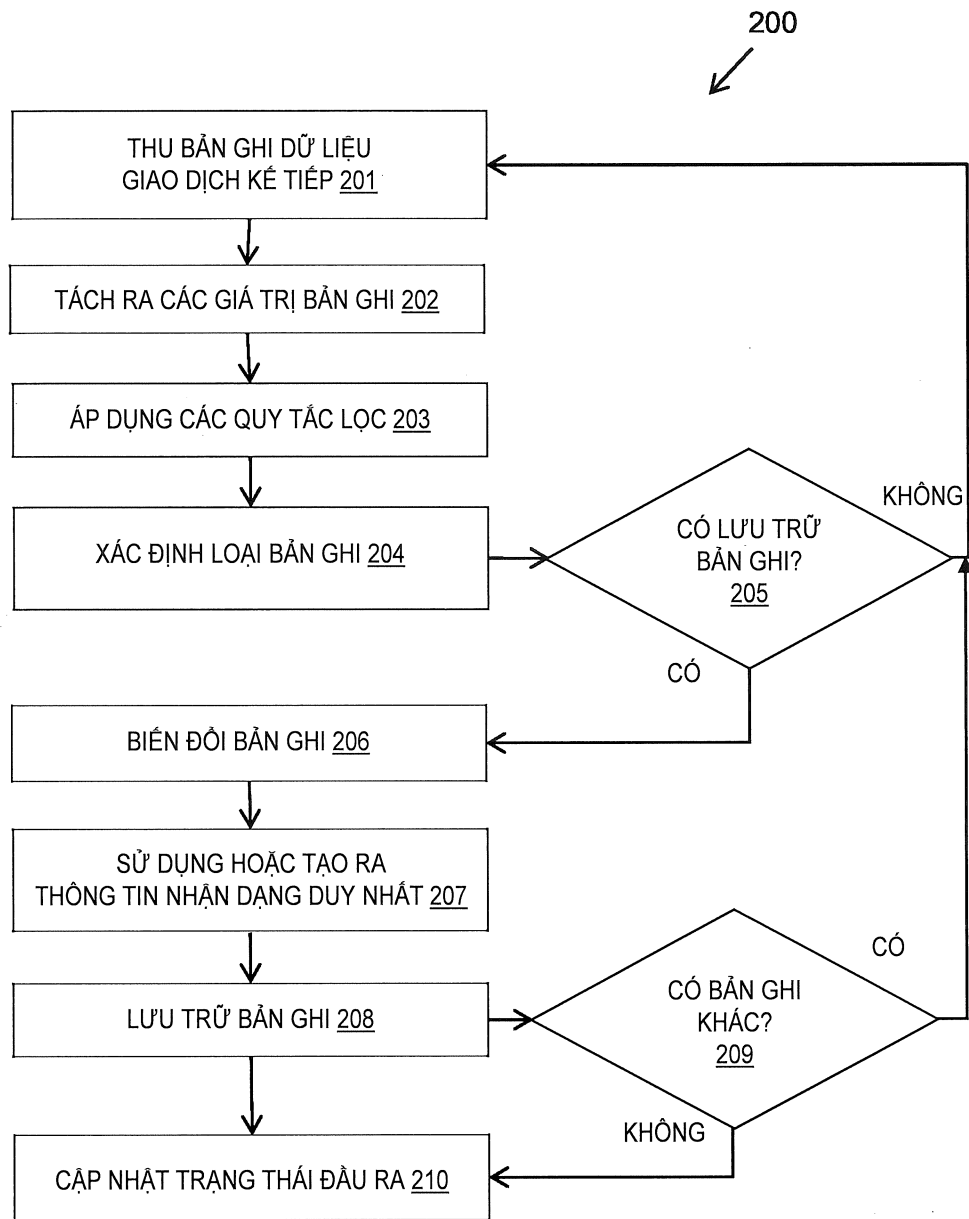
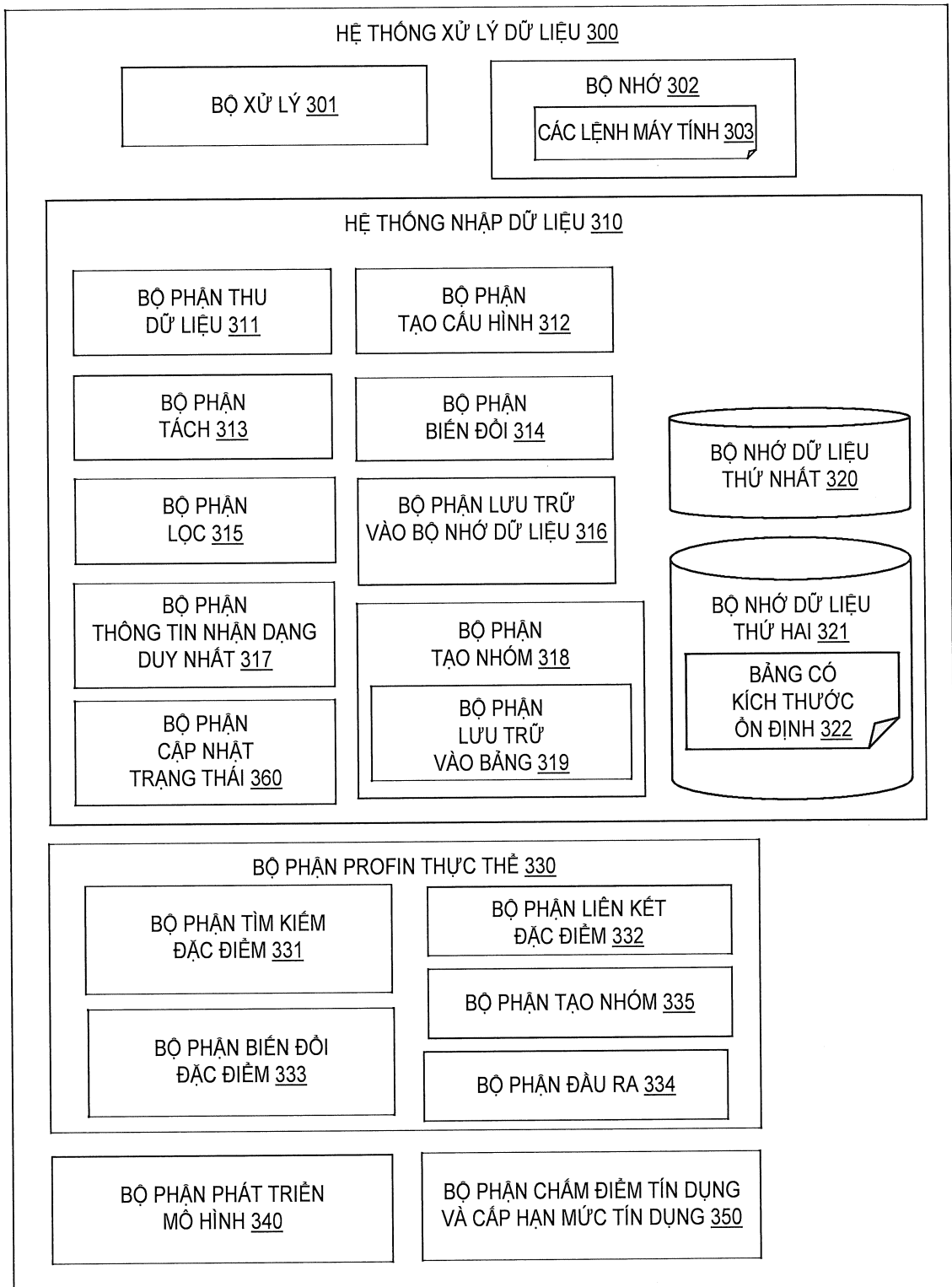
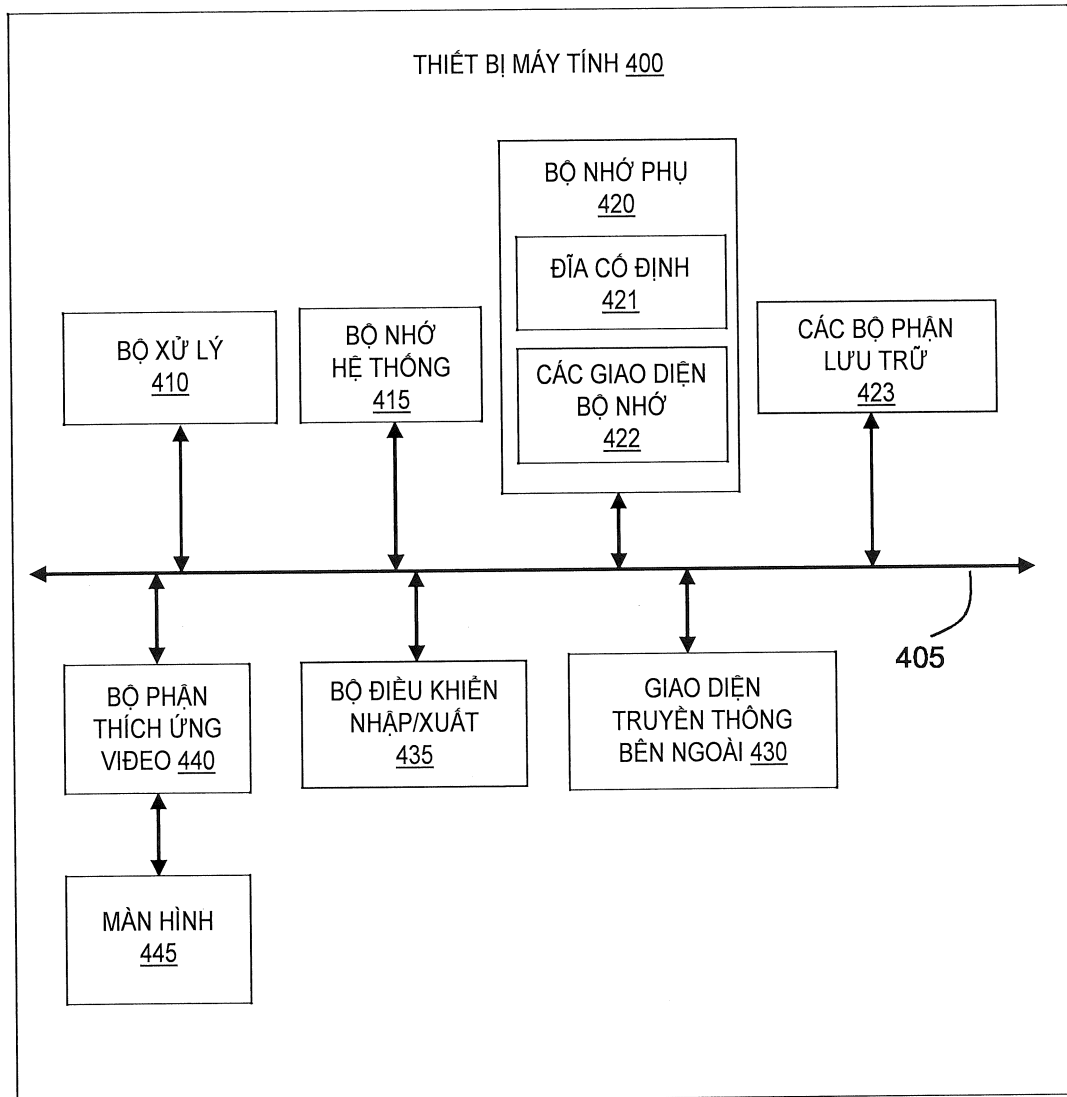


FIG. 2

3/4

**FIG. 3**

4/4

**FIG. 4**