

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 844 399**

51 Int. Cl.:

G06F 21/00	(2013.01)
G06F 21/57	(2013.01)
H04L 29/06	(2006.01)
G06F 17/00	(2009.01)
G06Q 40/00	(2012.01)
G06Q 20/40	(2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **21.12.2016 PCT/US2016/068050**
- 87 Fecha y número de publicación internacional: **17.08.2017 WO17139035**
- 96 Fecha de presentación y número de la solicitud europea: **21.12.2016 E 16890117 (1)**
- 97 Fecha y número de publicación de la concesión europea: **21.10.2020 EP 3414670**

54 Título: **Detección automática de puntos de compromiso**

30 Prioridad:

10.02.2016 US 201662293535 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

22.07.2021

73 Titular/es:

**FEEDZAI - CONSULTADORIA E INOVAÇÃO TECNÓLOGICA, S.A. (100.0%)
Rua Pedro Nunes - IPN
3030-199 Coimbra, PT**

72 Inventor/es:

**RAMOS DE ARAUJO, MIGUEL;
BORGES ALMEIDA, MIGUEL, SOUSA y
BIZARRO, PEDRO, GUSTAVO SANTOS
RODRIGUES**

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 844 399 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Detección automática de puntos de compromiso

5 REFERENCIA CRUZADA A OTRAS SOLICITUDES

[0001] Esta solicitud reivindica prioridad con respecto a la Solicitud de Patente Provisional de Estados Unidos de América con n.º 62/293.535 que lleva por título DETECCIÓN AUTOMÁTICA DE PUNTOS DE COMPROMISO, presentada el 10 de febrero de 2016.

10 ANTECEDENTES DE LA INVENCION

[0002] Las medidas de seguridad electrónicas son técnicas implementadas por ordenador para proteger posesiones e información. Las medidas de seguridad electrónicas tienen como objetivo prevenir vulneraciones de seguridad, incluida la sustracción indebida de información y el uso posterior de información sustraída indebidamente. Un ejemplo de un punto de compromiso es una ubicación de transacción (por ejemplo, un cajero automático o terminal punto de venta (TPV)) que se ha visto comprometida y las cuentas (por ejemplo, instrumentos de pago) que han utilizado el punto de compromiso pueden haber sufrido una sustracción indebida de la información de cuenta en el punto de compromiso. Un ordenador intenta detectar una vulneración de seguridad analizando transacciones/datos relacionados. Sin embargo, con frecuencia es difícil detectar de manera eficaz, rápida y precisa una vulneración de seguridad con los sistemas informáticos tradicionales, debido a la potencia informática que se necesita para analizar grandes cantidades de datos de transacciones rápidamente. Cuanto más tiempo se tarda en detectar un punto de compromiso, más probable es que se utilice la información robada. Por lo tanto, existe la necesidad de una medida de seguridad electrónica que detecte de manera eficaz, rápida y precisa las fuentes de vulneraciones de seguridad para, entre otras cosas, prevenir el uso de la información robada.

[0003] En la patente estadounidense n.º 2011/307382 se describe un sistema, aparato y método para reducir el fraude en las transacciones de pago mediante la identificación de un posible punto de compromiso (PDC) que puede tener como consecuencia transacciones fraudulentas posteriores, o mediante la confirmación de que un punto de compromiso sospechoso es de hecho responsable de una vulneración de seguridad que provoca transacciones de pago fraudulentas. También se describen un sistema, aparato y método para identificar datos o características de transacciones que pueden usarse para identificar de manera más eficaz un punto de compromiso en un sistema de transacciones de pago.

[0004] En *Introduction to Parallel Programming – Parallel Methods for Matrix Multiplication* (“Introducción a la computación paralela: métodos paralelos para la multiplicación de matrices”), de V. V. Gergel, se describen tres algoritmos paralelos para la multiplicación de matrices.

[0005] En la patente estadounidense n.º 2005/055373 se describe un método y sistema de minería de datos y descubrimiento de conocimientos en los que se establece una base de datos en forma de matriz virtual para cada transacción de eventos transaccionales de gran escala. Los datos se archivan y registran de manera que se clasifican rápidamente y, dado un conjunto de identificadores para las transacciones comprendidas, solo es necesario acceder a un subconjunto limitado de la matriz para el descubrimiento de conocimientos específicos en la naturaleza del punto de compromiso. Se registra una puntuación para cada punto de compromiso potencial. Los puntos de compromiso potenciales se ordenan según la puntuación registrada. La puntuación indica la probabilidad creciente de un punto de compromiso de origen.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0006] Diversas realizaciones de la invención se divulgan en la siguiente descripción detallada y los dibujos adjuntos.

La Figura 1 es un diagrama de bloques en el que se ilustra una realización de un sistema para identificar uno o más puntos de compromiso (“PDC”).

La Figura 2 es un diagrama funcional en el que se ilustra un sistema informático programado para identificar uno o más PDC de acuerdo con algunas realizaciones.

La Figura 3 es un diagrama de flujo en el que se ilustra una realización de un proceso para identificar uno o más PDC probables.

La Figura 4 es un diagrama de flujo en el que se ilustra una realización de un proceso para calcular una o más puntuaciones de responsabilidad.

La Figura 5 es un diagrama de flujo en el que se ilustra una realización de un proceso para identificar uno o más PDC que incluye la distribución del cálculo de una o más puntuaciones de responsabilidad.

La Figura 6 es un diagrama conceptual de transacciones en las que están implicadas cuentas y ubicaciones.

La Figura 7A es un diagrama conceptual de un estado inicial de una matriz de cálculo informático para identificar uno o más PDC probables.

La Figura 7B es un diagrama conceptual de un estado de una matriz de cálculo informático para identificar uno o más PDC probables después de una primera iteración.

La Figura 7C es un diagrama conceptual de un estado de una matriz de cálculo informático para identificar uno o más PDC probables después de una segunda iteración.

5 La Figura 7D es un diagrama conceptual de un estado de una matriz de cálculo informático para identificar uno o más PDC probables después de una tercera iteración.

La Figura 7E es un diagrama conceptual de una matriz de cálculo informático para identificar uno o más PDC probables que incluyen puntuaciones de responsabilidad de compromiso globales convergentes.

10 DESCRIPCIÓN DETALLADA

[0007] La invención puede implementarse de numerosas maneras, entre las que figuran: como un proceso; como un aparato; como un sistema; como una composición de materia; como un producto de programa informático incorporado en un soporte de almacenamiento legible por ordenador; y/o como un procesador, como por ejemplo un procesador configurado para ejecutar instrucciones almacenadas en una memoria acoplada al procesador y/o provistas por dicha memoria. En esta memoria descriptiva, estas implementaciones, o cualquier otra forma que pueda tomar la invención, pueden denominarse técnicas. En general, el orden de los pasos de los procesos descritos puede modificarse dentro del ámbito de la invención. A menos que se indique lo contrario, un componente, como por ejemplo un procesador o una memoria, que se describe como configurado para realizar una tarea, puede implementarse como un componente general que se configura temporalmente para realizar la tarea en un momento determinado o un componente específico que se fabrica para realizar la tarea. Tal y como se usa en el presente, el término "procesador" se refiere a uno o más dispositivos, circuitos y/o núcleos de procesamiento configurados para procesar datos, por ejemplo, instrucciones de programas informáticos.

25 [0008] Más adelante se proporciona una descripción detallada de una o más realizaciones de la invención, junto con las figuras adjuntas que ilustran los principios de la invención. La invención se describe en relación con dichas realizaciones, pero la invención no está limitada a ninguna realización. El ámbito de la invención está limitado solo por las reivindicaciones y la invención abarca numerosas alternativas, modificaciones y equivalentes. En la siguiente descripción se exponen numerosos detalles específicos con el fin de proporcionar una comprensión exhaustiva de la invención. Estos detalles se proporcionan con fines ilustrativos y la invención se puede poner en práctica, de acuerdo con las reivindicaciones, sin algunos o sin la totalidad de estos detalles específicos. En aras de la claridad, no se ha descrito en detalle el material técnico que se conoce en los campos técnicos relacionados con la invención con el fin de no oscurecer innecesariamente la invención.

35 [0009] Las técnicas de la presente descripción detectan una vulneración de seguridad y proporcionan una paralelización del procesamiento de datos de transacciones para detectar de manera eficiente, rápida y precisa la vulneración de seguridad. Un tipo de vulneración de seguridad es el compromiso de datos en el que se obtienen los datos durante una transacción fraudulenta. La vulneración de seguridad puede haberse originado en un punto de compromiso. Un punto de compromiso ("PDC") es una ubicación de origen de compromisos de datos fraudulentos. Un solo PDC puede llevar a un grupo de víctimas de fraudes (por ejemplo, cuentas comprometidas) que proporcionaron información de sus cuentas en el mismo PDC. Una ubicación se convierte en un PDC cuando la información proporcionada en la ubicación es sustraída indebidamente (por ejemplo, es interceptada, robada, copiada y transmitida ilegalmente a través de una red). La ubicación puede ser una ubicación virtual, como por ejemplo una ubicación de Internet, o una ubicación física, como el terminal de una tienda. Por ejemplo, la ubicación tiene un dispositivo de transacción asociado y la ubicación es vulnerada cuando se utiliza un dispositivo no autorizado para leer la tarjeta de una víctima (por ejemplo, mediante un *skimmer* o dispositivo de robo de información de tarjetas de crédito o por un empleado malintencionado cuando la víctima no tiene bajo control visual su tarjeta de crédito). Mientras una ubicación está comprometida, cualquier información de la cuenta que se proporcionó en el PDC es vulnerable a una sustracción indebida. Por lo tanto, la detección temprana y precisa de un PDC previene y/o reduce el impacto de las vulneraciones de seguridad. El descubrimiento oportuno del PDC evita el uso fraudulento de otras cuentas que pueden haber sido comprometidas en la misma ubicación en un periodo de tiempo relacionado. La identificación de un PDC y de la información relacionada permite deshabilitar o reparar un dispositivo de transacción en el PDC antes de que otras cuentas se vean comprometidas, y permite que las cuentas utilizadas en el PDC durante el periodo de tiempo relacionado (por ejemplo, cuentas posiblemente comprometidas) se deshabiliten (por ejemplo, volviendo a emitir una nueva tarjeta de crédito al titular de la cuenta). Mediante el uso de la representación y división de datos y de las técnicas de procesamiento distribuido descritas en el presente, se pueden identificar los PDC probables de manera eficiente y rápida.

60 [0010] Dado un conjunto de transacciones, las técnicas de la presente divulgación detectan la ubicación en la que la cuenta correspondiente se ha visto comprometida. Es posible detectar varios tipos de compromisos, incluidas las vulneraciones de bases de datos y dispositivos de *skimming* o robo de información de tarjetas bancarias en cajeros automáticos y terminales TPV comprometidos. En algunas realizaciones, para cada cuenta de una pluralidad de cuentas, se calcula una puntuación de responsabilidad de elemento para cada PDC potencial de una pluralidad de PDC potenciales. El cálculo de la puntuación de responsabilidad de elemento incluye la representación de los cálculos como al menos una operación de matriz de cálculo informático, la división de una matriz de la operación de

matriz de cálculo informático en bloques de cálculo y la distribución de los bloques entre una pluralidad de nodos informáticos de cálculo para un cálculo informático distribuido. Basándose en las puntuaciones de responsabilidad de elemento determinadas, se calcula una puntuación de responsabilidad de compromiso global para cada PDC potencial. Se determina si las puntuaciones de responsabilidad de compromiso globales determinadas han convergido. Si las puntuaciones de responsabilidad de compromiso globales determinadas han convergido, se selecciona al menos uno de los PDC potenciales como un PDC probable basándose en la puntuación de responsabilidad de compromiso global correspondiente del PDC probable seleccionado. Si las puntuaciones de responsabilidad de compromiso globales determinadas no han convergido, las puntuaciones de responsabilidad de elemento se actualizan para cada PDC potencial de la pluralidad de PDC potenciales para cada cuenta de la pluralidad de cuentas.

[0011] En algunas realizaciones, se implementa la determinación de un PDC mediante un procesamiento distribuido escalable ejecutable por nodos de un clúster. Se logra una rápida velocidad de procesamiento dividiendo y distribuyendo los cálculos en múltiples sistemas. Por ejemplo, debido a que es necesario procesar una enorme cantidad de datos de transacciones de tarjetas de crédito para identificar patrones que identifiquen probables PDC, se han diseñado las técnicas de procesamiento, selección, división y distribución descritas en el presente para lograr resultados precisos de manera eficiente y precisa.

[0012] La Figura 1 es un diagrama de bloques en el que se ilustra una realización de un sistema para identificar uno o más PDC. En la Figura 1 se muestran los componentes de un entorno en el que uno o más nodos de un clúster 140 identifican un PDC. El entorno incluye uno o más dispositivos de transacción 102, 104 y 106, una puerta de enlace 110, una red 120, un emisor 130 y un clúster 140 formado por uno o más nodos 142.1 y 142.2.

[0013] Cada uno de los dispositivos de transacción 102-106 recibe información de cuenta y transmite los datos de transacción a través de la puerta de enlace 110 para verificar la información de la cuenta. En algunas realizaciones, los dispositivos de transacción facilitan una transacción. Una transacción incluye el intercambio de información. Por ejemplo, el procesamiento de una transacción que implica una compra incluye la recepción de información de cuenta (por ejemplo, la cuenta de la tarjeta de crédito) y detalles de la transacción (por ejemplo, la cantidad de la compra) en una ubicación de transacción y la determinación de si se aprobará la transacción. Una transacción aprobada puede significar que el pago de la cuenta se acepta a cambio de bienes o servicios. Una transacción denegada puede significar que se rechaza el pago de la cuenta.

[0014] En algunas realizaciones, la información de cuenta incluye la información de pago, como por ejemplo un número de cuenta de una tarjeta de crédito, un número de una cuenta de débito, un número de clasificación de cuenta bancaria y similares. En algunas realizaciones, los datos de transacción son generados por el dispositivo de transacción mediante el procesamiento/filtrado de la información de la cuenta. Por ejemplo, se puede cifrar o utilizar con hash un número de cuenta para proteger el número de cuenta. Un dispositivo de transacción puede ser implementado por un terminal, un dispositivo de punto de venta (TPV) o cualquier otro dispositivo que acepte información de cuenta. Por ejemplo, un terminal incluye un terminal de tarjeta de crédito que procesa el pago basándose en un número de cuenta de tarjeta de crédito recibido. El dispositivo de transacción puede recibir y analizar información de la cuenta usando una variedad de técnicas electrónicas como un lector de chip, un lector de banda magnética, un lector de códigos de barras, etc. En algunas realizaciones, un dispositivo de transacción está asociado con una ubicación y puede ser identificado por su ubicación asociada. Por ejemplo, un vendedor minorista con instalaciones físicas (BM, por sus siglas en inglés, *brick and mortar*) que tiene tres terminales de caja (1-3), cada uno de ellos equipado con uno de los dispositivos de transacción 102-106, puede ser identificado por los dispositivos de transacción BM1, BM2 y BM3. Otro ejemplo sería un dispositivo de transacción que es un sitio web que procesa el pago de bienes y servicios comprados a través de Internet.

[0015] Se asocia típicamente una ubicación de transacción con un dispositivo de transacción. Los ejemplos de dispositivos de transacción incluyen un cajero automático, un terminal TPV, un ordenador y dispositivos similares. Una ubicación de transacción es una ubicación donde se puede recibir información de la cuenta para iniciar una transacción. Una ubicación de transacción puede ser una ubicación física, una ubicación de un terminal, una ubicación web y similares. Los ejemplos de ubicaciones de transacciones incluyen uno o más terminales de caja, una o más tiendas, un grupo de tiendas o todo el sistema (por ejemplo, todo el comerciante de comercio electrónico) y similares.

[0016] Un dispositivo de transacción puede estar sujeto a vulnerabilidades de seguridad. Un tipo de vulneración de seguridad implica convertir el dispositivo de transacción en un PDC cuando una parte no autorizada desvía o copia la información de cuenta recibida. Por ejemplo, una parte no autorizada instala un *skimmer* de tarjetas de crédito en el dispositivo de transacción para almacenar la información de la cuenta proporcionada al dispositivo de transacción.

[0017] La puerta de enlace 110 recibe datos de transacción de uno o más dispositivos de transacción 102-106, enruta los datos de transacción a la red 120 y devuelve un aviso de aprobación o denegación basado en el proceso de aprobación de la red 120. La puerta de enlace 110 puede incluir un adquirente de pagos o proveedor de servicios de Internet. Por ejemplo, el adquirente de pagos puede ser un software alojado en un servidor de terceros que

gestiona las transmisiones entre un comerciante (representado por los dispositivos de transacción 102-106) y un emisor 130. En algunas realizaciones, una puerta de enlace está asociada con un banco adquirente (también denominado banco mercantil). El banco adquirente está registrado en una red 120, en donde la red representa una asociación de tarjetas (por ejemplo, Visa®, MasterCard®, American Express®, etc.). El banco adquirente contrata a los comerciantes para crear y mantener cuentas que permiten al comerciante aceptar cuentas, como por ejemplo tarjetas de crédito y débito. En algunas realizaciones, la puerta de enlace 110 procesa y cifra los datos de transacción antes de enrutar los datos de transacción. En algunas realizaciones, la puerta de enlace 110 agrupa una o más transacciones juntas y envía el lote de transacciones al emisor 130 a través de la red 120.

[0018] La red 120 es una plataforma para transmitir datos entre dispositivos con el fin de admitir el procesamiento de pagos y pagos electrónicos. En algunas realizaciones, la red 120 está asociada con una asociación de tarjetas de crédito (por ejemplo, Visa®, Mastercard®, American Express®, etc.) y admite comunicaciones entre los miembros de la asociación, como por ejemplo un banco adquirente (por ejemplo, la puerta de enlace 110) y un banco emisor (por ejemplo, el emisor 130). En algunas realizaciones, la red 120 implementa una cámara de compensación para proporcionar servicios de compensación y liquidación. La red 120 determina un destino apropiado para enrutar los datos de transacción. Por ejemplo, varios bancos emisores pueden ser miembros de la red. La red determina el emisor correspondiente a los datos de transacción y enruta la transacción al emisor correspondiente. Por simplicidad, solo se muestra un emisor 130 en la Figura 1. En algunos casos, la red 120 elige entre varios emisores a los que se envían los datos de la transacción. En algunas realizaciones, la red 120 filtra los datos de transacción recibidos. Por ejemplo, la red 120 puede tener conocimiento de cuentas fraudulentas y determinar si los datos de transacción recibidos incluyen una cuenta fraudulenta. La red 120 puede incluir uno o más servidores conectados a la red para procesar, enrutar y/o facilitar transacciones.

[0019] El emisor 130 recibe datos de transacción de la red 120 y determina si aprobará o denegará una transacción (por ejemplo, una cuenta proporcionada/un pago). Por ejemplo, el emisor 130 incluye uno o más servidores/sistemas de un banco emisor. En algunas realizaciones, el emisor está asociado con un banco adquirente a través de la red 120. En algunas realizaciones, determinar si aprobar o denegar una cuenta/método de pago incluye determinar si la transacción es fraudulenta. Por ejemplo, una transacción se considera fraudulenta si el método de pago (por ejemplo, la cuenta) utilizado en la transacción está involucrado en una vulneración de seguridad (por ejemplo, si el método de pago ha interactuado con un punto de compromiso). En algunas realizaciones, el emisor 130 llama a uno o más nodos del clúster 140 en tiempo real para determinar si existe una vulneración de seguridad. Por ejemplo, el emisor 130 consulta uno o más nodos del clúster 140 sobre una o más transacciones y solicita información de uno o más nodos del clúster 140 sobre si debe continuar con la transacción. Por ejemplo, si la transacción implica una posible vulneración de seguridad, la transacción es denegada. La identificación de si un método de pago está implicado en una vulneración de seguridad se describe en mayor detalle en el presente. En algunas realizaciones, el emisor 130 proporciona a uno o más nodos del clúster 140 una lista de transacciones (pasadas) que buscan identificación de uno o más nodos del clúster 140 sobre probables PDC y/o cuentas comprometidas basándose en las transacciones.

[0020] En algunas realizaciones, los nodos del clúster 140 son controlados y administrados por el emisor 130. Por ejemplo, dispositivos/sistemas del emisor retienen información de las transacciones y realizan análisis para identificar probables PDC. En algunas realizaciones, los nodos del clúster 140 son controlados y administrados por un tercero. Por ejemplo, el emisor 130 ha contratado a un tercero para realizar un análisis utilizando los datos proporcionados por el emisor (por ejemplo, información de la transacción) con el fin de identificar para el emisor PDC posibles.

[0021] Uno o más nodos del clúster 140 realizan los procesos descritos en el presente, por ejemplo, los procesos mostrados en las Figuras 3-5. Los procesos se pueden realizar en paralelo o se pueden distribuir entre uno o más nodos, tal y como se describe en mayor detalle en el presente, por ejemplo, con respecto a la Figura 5. En algunas realizaciones, uno o más nodos del clúster 140 reciben una lista de transacciones del emisor 130, analizan las transacciones y proporcionan una salida. El proceso mostrado en la Figura 3 es un ejemplo de cómo las transacciones son analizadas por el nodo o los nodos.

[0022] En algunas realizaciones, uno o más nodos del clúster 140 reciben información de transacción, identifican PDC potenciales a partir de la misma y determinan cuál de los PDC potenciales son PDC probables basándose en los PDC potenciales identificados. Un PDC potencial representa una característica que un subconjunto de transacciones puede tener en común. Entre los ejemplos de características figuran un ID de terminal de TPV específico, un ID de tienda, un nombre de comerciante (por ejemplo, todas las tiendas de una corporación), la hora y características similares. Un PDC potencial ha interactuado con una cuenta comprometida y, por lo tanto, podría ser identificado en última instancia como un PDC. De entre los PDC potenciales se identifican uno o más PDC probables. Un PDC probable es un PDC potencial que cumple un criterio (descrito en mayor detalle más adelante en el presente) y representa una ubicación de transacción que probablemente se ha visto comprometida. Por ejemplo, un PDC potencial que no es una fuente de compromiso, a pesar de haber interactuado con una cuenta comprometida, no es identificado como un PDC probable. En algunas realizaciones, la salida de los PDC probables incluye información asociada como la ubicación, la fecha/hora en la que el punto se ha comprometido y las cuentas

que interactuaron con el PDC probable durante esa fecha/hora. En algunas realizaciones, la información asociada se usa para identificar otras cuentas que pueden haber sido comprometidas basándose en las cuentas que interactuaron con el PDC probable durante un periodo de tiempo durante el cual la ubicación probablemente estuvo comprometida.

5 [0023] En algunas realizaciones, los PDC probables se identifican a partir de transacciones pasadas. Basándose en la información de que la ubicación de una transacción es un PDC probable, una transacción en una ubicación de transacción puede ser denegada o marcada durante el proceso de verificación de pago. Por ejemplo, si una transacción involucra un PDC probable previamente identificado, la transacción es denegada o marcada. Es probable que las cuentas involucradas en una transacción marcada estén comprometidas. En respuesta al posible compromiso, se puede emitir una nueva cuenta al titular de la cuenta o se puede emitir una nueva tarjeta o instrumento de pago para la misma cuenta.

15 [0024] En algunas realizaciones, uno o más nodos 142,1 y 142,2 en el clúster 140 están controlados por el emisor 130. Por ejemplo, el nodo o los nodos pueden ser proporcionados dentro del entorno informático del emisor 130. En algunas realizaciones, el nodo o los nodos están controlados por un tercero (no el emisor 130), y el emisor contrata al tercero para la realización del análisis del PDC.

20 [0025] Por ejemplo, el análisis del PDC se realiza como sigue. En una realización en la que el nodo o los nodos son controlados por el emisor, el emisor controla el nodo o los nodos de acuerdo con los procesos descritos en el presente, por ejemplo, el proceso mostrado en la Figura 3 para determinar los PDC probables. En una realización en la que el nodo o los nodos son controlados por un tercero, el emisor proporciona datos de transacciones al nodo o los nodos, y el nodo o los nodos identifican PDC probables, donde los nodos operan de acuerdo con los procesos descritos en el presente, por ejemplo, el proceso mostrado en la Figura 3 para determinar PDC probables. El nodo o los nodos devuelven al emisor resultados del análisis del PDC, como por ejemplo PDC probables.

25 [0026] Los resultados del análisis del PDC son útiles para, entre otras cosas, determinar si una transacción es permitida o denegada. Por ejemplo, es probable que cualquier cuenta que haya interactuado previamente con el PDC probable durante una franja de tiempo de vulneración se haya visto comprometida. Una transacción que intenta ser realizada por una cuenta identificada como probablemente comprometida es denegada. En otro ejemplo, la autorización de transacciones se gestiona de la siguiente manera. Los PDC probables previamente identificados se almacenan en el almacenamiento 144. Cuando realiza la autorización de la transacción basada en la información de transacción recibida, el emisor 130 accede al almacenamiento 144 para determinar si la información de la transacción recibida está asociada con un dispositivo/ubicación de transacción previamente identificado como un PDC probable, habiéndose almacenado el PDC probable en el almacenamiento 144. Por ejemplo, si la información de transacción coincide con un PDC probable, el emisor deniega la transacción o marca la transacción (por ejemplo, envía a un técnico para verificar si se ha instalado un *skimmer* de tarjetas en el dispositivo de transacción, emite una nueva tarjeta de crédito para una cuenta comprometida involucrada en la transacción, etc.).

40 [0027] El almacenamiento 144 almacena información sobre transacciones. En algunas realizaciones, el almacenamiento 144 almacena una lista de transacciones y/o una lista de cuentas robadas/fraudulentas. En algunas realizaciones, el almacenamiento 144 almacena una lista de cuentas. La información de transacción se puede proporcionar como una sola transacción o como una lista de transacciones. Se analizan ejemplos de información de transacciones con respecto a 302 en la Figura 3.

45 [0028] Un proceso de verificación del pago puede tener lugar dentro del entorno mostrado en la Figura 1. Durante el funcionamiento, un dispositivo de transacción (102, 104 y/o 106) recibe información de transacción como la cuenta, la hora, la cantidad, etc., como se describe en mayor detalle en el presente. En algunas realizaciones, el dispositivo de transacción procesa la información de transacción (por ejemplo, empaqueta los datos). El dispositivo de transacción envía los datos de transacción a la puerta de enlace 110. La puerta de enlace 110 enruta los datos de transacción recibidos a la red 120. La red 120 determina un emisor basándose en los datos de transacción y envía los datos de transacción al emisor. El emisor 130 determina si aprueba o deniega la transacción basándose en los datos de transacción y un proceso de seguridad realizado por uno o más nodos 142.1 y 142.2. El proceso de seguridad realizado por uno o más nodos incluye el análisis de los datos de transacciones recibidos para identificar PDC probables. El proceso mostrado en la Figura 3 es un ejemplo de un proceso de seguridad realizado por uno o más nodos. En algunas realizaciones, los PDC probables se almacenan de tal manera que las transacciones futuras son gestionadas basándose en los PDC probables previamente identificados.

60 [0029] En algunas realizaciones, los PDC probables son utilizados por el nodo o los nodos para determinar si se autorizará una transacción. En algunas realizaciones, se proporcionan los PDC probables al emisor 130 para que el emisor determine cómo gestionar la transacción. La transacción puede gestionarse de la siguiente manera. Si la transacción implica un PDC probable, la transacción es denegada o marcada. El PDC probable puede ser deshabilitado o reparado. Por ejemplo, se envía a un técnico de servicio para comprobar si se ha instalado un *skimmer* de tarjetas en el dispositivo de transacción. Las cuentas que han interactuado con el PDC probable pueden ser identificadas como comprometidas. Los titulares de cuentas de cuentas comprometidas pueden ser notificados,

se les puede proporcionar una nueva tarjeta de crédito, etc. Si la transacción no implica un PDC probable, la transacción es aprobada.

5 [0030] La red 120 y la puerta de enlace 110 retransmiten el aviso de aprobación o denegación al dispositivo de transacción. Si la transacción es aprobada, se acepta el pago y la transacción se realiza correctamente. Si la transacción es denegada, no se acepta el pago y la transacción es denegada.

10 [0031] La Figura 2 es un diagrama funcional en el que se ilustra un sistema informático programado para identificar uno o más PDC de acuerdo con algunas realizaciones. El sistema informático programado es un ejemplo de un nodo (142.1 o 142.2) mostrado en la Figura 1. Como será evidente, se pueden utilizar otras arquitecturas y configuraciones de sistemas informáticos para realizar las técnicas descritas en el presente. Por lo que respecta de nuevo a la Figura 2, el sistema informático, que incluye varios subsistemas como se describe a continuación, incluye al menos un subsistema de microprocesador (también denominado procesador o unidad central de procesamiento (CPU, por sus siglas en inglés, *Central Processing Unit*) 202). Por ejemplo, se puede implementar el procesador 202 mediante un procesador de un solo chip o mediante procesadores múltiples. En algunas realizaciones, el procesador 202 es un procesador digital de uso general que controla el funcionamiento del sistema informático. En algunas realizaciones, el procesador 202 también incluye uno o más coprocesadores o procesadores de uso especial (por ejemplo, un procesador de gráficos, un procesador de red, etc.). Usando instrucciones recuperadas de la memoria 210, el procesador 202 controla la recepción y manipulación de los datos de entrada recibidos en un dispositivo de entrada (por ejemplo, un dispositivo señalador 206, la interfaz del dispositivo de E/S 204), y la salida y visualización de datos en dispositivos de salida (por ejemplo, una pantalla 218). En algunas realizaciones, el procesador 202 se usa para realizar los procesos descritos en el presente con respecto a las Figuras 3-5.

25 [0032] El procesador 202 está acoplado bidireccionalmente con la memoria 210, lo que puede incluir, por ejemplo, una o más memorias de acceso aleatorio (RAM, por sus siglas en inglés, *Random Access Memory*) y/o una o más memorias de solo lectura (ROM, por sus siglas en inglés, *Read-Only Memory*). Como es bien conocido en la técnica, la memoria 210 se puede usar como un área de almacenamiento general, una memoria temporal (por ejemplo, *scratchpad*) y/o una memoria caché. También se puede utilizar la memoria 210 para almacenar datos de entrada y datos procesados, así como para almacenar instrucciones y datos de programación, en forma de objetos de datos y objetos de texto, además de otros datos e instrucciones para procesos que operan en el procesador 202. Asimismo, como es bien conocido en la técnica, la memoria 210 incluye normalmente instrucciones operativas básicas, código de programa, datos y objetos usados por el procesador 202 para realizar sus funciones (por ejemplo, instrucciones programadas). Por ejemplo, la memoria 210 puede incluir cualquier medio de almacenamiento legible por ordenador adecuado que se describe a continuación, dependiendo de si, por ejemplo, el acceso a los datos debe ser bidireccional o unidireccional. Por ejemplo, el procesador 202 también puede recuperar y almacenar directa y muy rápidamente los datos que se necesitan con frecuencia en una memoria caché incluida en la memoria 210.

40 [0033] Un dispositivo de almacenamiento masivo extraíble 212 proporciona una capacidad de almacenamiento de datos adicional para el sistema informático, y está opcionalmente acoplado ya sea bidireccionalmente (lectura/escritura) o unidireccionalmente (solo lectura) al procesador 202. Un almacenamiento masivo fijo 230 puede también, por ejemplo, proporcionar una capacidad de almacenamiento de datos adicional. Por ejemplo, los dispositivos de almacenamiento 212 y/o 230 pueden incluir medios legibles por ordenador como cinta magnética, memoria flash, PC-CARDS, dispositivos portátiles de almacenamiento masivo como discos duros (por ejemplo, unidades magnéticas, ópticas o de estado sólido), dispositivos de almacenamiento holográfico y otros dispositivos de almacenamiento. Los dispositivos de almacenamiento masivo 212 y/o 230 generalmente almacenan instrucciones de programación adicionales, datos y similares que típicamente no están en uso activo por el procesador 202. Se apreciará que la información retenida dentro de los almacenamientos masivos 212 y 230 se puede incorporar, si es necesario, de manera estándar como parte de la memoria 210 (por ejemplo, RAM) como memoria virtual.

50 [0034] Además de proporcionar acceso al procesador 202 a subsistemas de almacenamiento, también se puede utilizar un bus 214 para proporcionar acceso a otros subsistemas y dispositivos. Como se muestra, estos pueden incluir una pantalla 218, una interfaz de red 216, una interfaz de dispositivo de entrada/salida (E/S) 204, un dispositivo señalador 206, así como otros subsistemas y dispositivos. Por ejemplo, el dispositivo señalador 206 puede incluir o ser operado junto con una cámara, un escáner, etc.; la interfaz del dispositivo de E/S 204 puede incluir una interfaz de dispositivo para interactuar con una pantalla táctil (por ejemplo, una pantalla táctil capacitiva que admite la interpretación de gestos), un micrófono, una tarjeta de sonido, un altavoz, un teclado, un dispositivo señalador (por ejemplo, un ratón, un lápiz óptico o un dedo humano), un receptor de un sistema de posicionamiento global (GPS), un acelerómetro y/o cualquier otra interfaz de dispositivo apropiada para interactuar con el sistema. Se pueden utilizar múltiples interfaces de dispositivos de E/S junto con el sistema informático. La interfaz del dispositivo de E/S puede incluir interfaces generales y personalizadas que permiten al procesador 202 enviar y, más normalmente, recibir datos de otros dispositivos como teclados, dispositivos señaladores, micrófonos, pantallas táctiles, lectores de tarjetas de transductores, lectores de cinta, reconocedores de voz o escritura a mano, lectores biométricos, cámaras, dispositivos portátiles de almacenamiento masivo y otros ordenadores.

65 [0035] La interfaz de red 216 permite al procesador 202 ser acoplado a otro ordenador, red de ordenadores o red

de telecomunicaciones mediante una conexión de red como se muestra. Por ejemplo, a través de la interfaz de red 216, el procesador 202 puede recibir información (por ejemplo, objetos de datos o instrucciones de programa) de otra red o enviar información a otra red en el curso de la realización de un proceso o de los pasos de un proceso. La información, a menudo representada como una secuencia de instrucciones que se ejecutarán en un procesador, se puede recibir de otra red y enviar a otra red. Pueden usarse una tarjeta de interfaz o un dispositivo similar y el software apropiado implementado por (por ejemplo, ejecutado/realizado en) el procesador 202 para conectar el sistema informático a una red externa y transferir datos de acuerdo con protocolos estándar. Por ejemplo, se pueden ejecutar varias realizaciones de proceso descritas en el presente en el procesador 202, o se pueden realizar a través de una red como Internet, redes de intranet o redes de área local, junto con un procesador remoto que comparte una parte del procesamiento. También se pueden conectar dispositivos de almacenamiento masivo adicionales (no mostrados) al procesador 202 a través de la interfaz de red 216.

[0036] Además, diversas realizaciones descritas en el presente se refieren también a productos de almacenamiento informático con un medio legible por ordenador que incluye código de programa para realizar diversas operaciones implementadas por ordenador. El medio legible por ordenador incluye cualquier dispositivo de almacenamiento de datos que pueda almacenar datos que puedan ser leídos posteriormente por un sistema informático. Entre los ejemplos de medios legibles por ordenador figuran, pero no se limitan a los mismos: medios magnéticos como discos y cintas magnéticas; medios ópticos como discos CD-ROM; medios magneto-ópticos como discos ópticos; y dispositivos de hardware especialmente configurados, como por ejemplo circuitos integrados de aplicaciones específicas (ASIC, por sus siglas en inglés, *Application-Specific Integrated Circuit*), dispositivos lógicos programables (PLD, por sus siglas en inglés, *Programmable Logic Devices*) y dispositivos ROM y RAM. Los ejemplos de código de programa incluyen tanto código de máquina producido, por ejemplo, por un compilador, como archivos que contienen código de nivel superior (por ejemplo, *script*) que se pueden ejecutar utilizando un intérprete.

[0037] El sistema informático que se muestra en la Figura 2 constituye solo un ejemplo de un sistema informático apropiado para su uso con las diversas realizaciones descritas en el presente. Otros sistemas informáticos apropiados para tal uso pueden incluir subsistemas adicionales o un número inferior de subsistemas. En algunos sistemas informáticos, los subsistemas pueden compartir componentes (por ejemplo, para dispositivos con pantalla táctil como teléfonos inteligentes, tabletas, etc., la interfaz del dispositivo de E/S 204 y la pantalla 218 comparten el componente de pantalla táctil, que detecta las entradas del usuario y muestra las salidas al usuario). Además, el bus 214 es ilustrativo de cualquier esquema de interconexión que sirve para enlazar los subsistemas. También se pueden utilizar otras arquitecturas informáticas que tengan diferentes configuraciones de subsistemas.

[0038] Un procesador como el procesador 202 mostrado en la Figura 2 puede implementar las técnicas descritas en el presente. El procesador se puede implementar como componentes de software, componentes de hardware como dispositivos lógicos programables (por ejemplo, microprocesadores, matrices de puertas lógicas programables en campo (FPGA, por sus siglas en inglés, *Field-Programmable Gate Arrays*), procesadores de señales digitales (DSP, por sus siglas en inglés, *Digital Signal Processors*), etc.), circuitos integrados de aplicaciones específicas (ASIC) diseñados para realizar determinadas funciones, o una combinación de los mismos. En este ejemplo, el controlador de entrega de aplicaciones (ADC, por sus siglas en inglés, *Application Delivery Controller*) se implementa como un dispositivo de red que administra el tráfico de la red y permite a los clientes solicitar recursos web dentro de una red de suministro de aplicaciones.

[0039] La Figura 3 es un diagrama de flujo en el que se ilustra una realización de un proceso para identificar uno o más PDC probables. El proceso de la Figura 3 puede implementarse al menos en parte en uno o más componentes de nodos en el clúster 140 mostrado en la Figura 1. En algunas realizaciones, el proceso mostrado en la Figura 3 es realizado por un procesador, por ejemplo, un procesador implementado por un solo chip o por múltiples procesadores, como por ejemplo el procesador 202 mostrado en la Figura 2.

[0040] En 302, se recibe información de la transacción. Basándose en la información de la transacción, se identifican los PDC potenciales. La información de la transacción (también denominada "valores de parámetros de la transacción") incluye atributos de una o más transacciones (por ejemplo, la hora, el nombre del comerciante, la ubicación, etc.) realizadas en una ubicación de transacción que implica una cuenta. En algunas realizaciones, una ubicación de transacción es una ubicación donde se puede recibir información de cuenta para iniciar una transacción. Una ubicación de transacción puede ser una ubicación física, una ubicación de un terminal, una ubicación en la Web y ubicaciones similares. Entre los ejemplos de ubicaciones de transacciones figuran uno o más terminales de cobro, una o más tiendas, un grupo de tiendas o todo el sistema (por ejemplo, todo el comerciante de comercio electrónico) y similares. Por ejemplo, la información de la transacción es recopilada y proporcionada por un emisor como el emisor 130 mostrado en la Figura 1 a los nodos 142.1 y 142.2.

[0041] Se puede proporcionar información de la transacción como una sola transacción o como una lista de transacciones. Por ejemplo, una lista de transacciones incluye, para cada transacción, uno o más campos asociados que almacenan características de la transacción. A modo de ejemplo no limitativo, los campos incluyen detalles sobre la cuenta y el pago implicados en la transacción, los dispositivos/entidades implicados en la transacción, cuándo se realiza la transacción y dónde se realiza la transacción. Los detalles sobre la cuenta/el pago implicados

en la transacción incluyen la fecha de caducidad de la cuenta, los detalles de la cuenta (por ejemplo, el número de cuenta de tarjeta de crédito/el código de Valor de Verificación de Tarjeta (CVV, por sus siglas en inglés, *Card Verification Value*) con hash para proteger la identidad del titular de la tarjeta), el tipo de moneda utilizada, el importe del pago, la categoría de la compra (por ejemplo, tienda de comestibles, combustible, entretenimiento, etc.), cómo se aceptó el pago (por ejemplo, pasando la tarjeta o por chip), si el pago es un cargo periódico (por ejemplo, una suscripción), si el pago es una preautorización y similares. La identificación de los dispositivos/entidades implicados en la transacción incluye la identificación del comerciante, como por ejemplo el nombre, código, dirección, ciudad, país y/o código postal, si la transacción se realiza en línea o en una tienda física, el identificador de la puerta de enlace desde la cual se recibe la transacción (por ejemplo, la identificación de una puerta de enlace como la puerta de enlace 110 mostrada en la Figura 1) y similares. La información sobre cuándo se realiza la transacción incluye una hora, una fecha y similares. La información sobre dónde se realiza la transacción incluye una ubicación geográfica (por ejemplo, un nombre de ciudad y estado, una región, las coordenadas GPS, etc.), un terminal específico donde se realiza la transacción y similares.

[0042] En algunas realizaciones, la información de transacción corresponde a transacciones que tienen lugar dentro de una franja de tiempo determinada. Por ejemplo, a medida que pasa el tiempo, se genera más información sobre las transacciones, que se proporciona en lotes. Un primer lote incluye información de transacciones recopilada durante una primera franja de tiempo, un segundo lote incluye información de transacciones recopilada durante una segunda franja de tiempo, donde la segunda franja de tiempo es posterior a la primera franja de tiempo, etc. En varias realizaciones, las franjas de tiempo son secuenciales, son franjas no superpuestas de la misma longitud, franjas de tiempo de varias longitudes determinadas dinámicamente, etc. En algunas realizaciones, la franja de tiempo se define basándose en una cuenta comprometida. Por ejemplo, la información de transacción se establece por un periodo de tiempo definido con respecto a una transacción que involucra a la cuenta comprometida, como, por ejemplo, que comienza en el momento en que se detecta que la cuenta está comprometida (por ejemplo, la primera vez que participó en una transacción fraudulenta). En algunas realizaciones, la información de transacción es recibida a petición de uno o más nodos (nodos 142.1 y 142.2 mostrados en la Figura 1) o por iniciativa del emisor. En algunas realizaciones, la información de transacción es recibida como parte de la operación de un emisor, como por ejemplo un proceso de verificación de pago. En algunas realizaciones, la información de transacción es recibida como respuesta a la finalización de una transacción, por ejemplo, después de que el emisor haya aprobado o denegado una transacción. En algunas realizaciones, la información de transacción se recibe periódicamente, por ejemplo, después de que se hayan completado varias transacciones o cada pocas horas o días.

[0043] Para mejorar la eficiencia de la identificación de un PDC probable, la información de transacción se filtra para identificar información pertinente. Por ejemplo, una ubicación que no ha interactuado con una cuenta comprometida probablemente no sea la fuente del compromiso. Por lo tanto, al excluir las transacciones asociadas con ubicaciones con las que una cuenta comprometida no ha interactuado, la información de transacción filtrada incluye solo información de transacción para ubicaciones que han interactuado con la cuenta comprometida. En otro ejemplo, si se conoce el momento de compromiso más antiguo (por ejemplo, la cuenta se vio comprometida en algún momento después de enero de 2016), se pueden excluir las transacciones que tienen lugar antes del momento de compromiso más antiguo.

[0044] En 304, se identifican las cuentas y los PDC potenciales basándose en la información de transacción recibida.

[0045] Las cuentas se identifican analizando la información de cuenta a partir de la información de transacción recibida. Por ejemplo, se identifica una cuenta mediante los detalles de la cuenta, como por ejemplo el número de cuenta de una tarjeta de crédito, que se puede recibir en combinación con otra información, como un código CVV y/o la fecha de caducidad de la cuenta. Se puede identificar una cuenta sin descifrar la información de la cuenta cifrada. Por ejemplo, un número de cuenta de tarjeta de crédito con hash es suficiente para identificar una cuenta. En algunas realizaciones, se recibe en 302 una lista predeterminada de cuentas con la información de la transacción.

[0046] Una cuenta comprometida es una cuenta que ya no es segura y cuya información puede ser utilizada indebidamente por un usuario no autorizado. Una cuenta se ve comprometida cuando la información de cuenta es sustraída (por ejemplo, interceptada, robada, copiada, información transmitida ilegalmente a través de una red, poseída por un usuario no autorizado, etc.) o utilizada indebidamente (por ejemplo, utilizada por un usuario no autorizado). En algunas realizaciones, se informa que una cuenta está comprometida por un cliente, un banco u otra entidad en el sistema mostrado en la Figura 1 (por ejemplo, por un comerciante que opera los dispositivos de transacción 102-106, la puerta de enlace 110, la red 120 o el emisor 130). En algunas realizaciones, la identificación de cuentas incluye la identificación de cuentas que se sabe que han sido comprometidas, como se describe más detalladamente en el presente, por ejemplo, con respecto a la Figura 3.

[0047] Los PDC potenciales incluyen ubicaciones que han interactuado con cuentas comprometidas. En algunas realizaciones, las ubicaciones que han interactuado con cuentas comprometidas se identifican como PDC potenciales porque tienen la probabilidad de ser un PDC. Por ejemplo, las ubicaciones anteriores donde una cuenta que se sabe que está comprometida ha realizado transacciones son ubicaciones potenciales donde existió una

oportunidad de interceptar la información de la cuenta. La probabilidad precisa de ser un PDC se determina de acuerdo con un proceso como el que se muestra en la Figura 4. La Figura 3 es un ejemplo de cómo se identifican las cuentas y los PDC potenciales basándose en la información de transacción recibida. En algunas realizaciones, las cuentas identificadas y/o los PDC potenciales se almacenan en el almacenamiento 144 que se muestra en la Figura 1. Cuando se realizan actualizaciones a cuentas y/o PDC potenciales, se puede actualizar esta información en el almacenamiento. Por ejemplo, cuando se reparan los PDC o se reemplazan los servicios o las cuentas, la identificación de que están comprometidos se puede eliminar del almacenamiento.

[0048] En 306, se calcula una puntuación de responsabilidad de compromiso global para cada uno de los PDC potenciales. En algunas realizaciones, se calcula la puntuación de responsabilidad de compromiso global para una combinación de cuenta y PDC potencial (por ejemplo, ubicación). Una puntuación de responsabilidad de compromiso global se asocia con una ubicación y representa la medida de certeza (por ejemplo, probabilidad) de que la ubicación sea un PDC, por ejemplo, en qué grado se responsabiliza la ubicación por comprometer cuentas. El cálculo de la puntuación de responsabilidad se describe adicionalmente en el presente con respecto a la Figura 4. En algunas realizaciones, el cálculo se realiza mediante un sistema informático distribuido, como por ejemplo uno o más nodos del clúster 140 mostrado en la Figura 1. En el presente se describe adicionalmente un cálculo distribuido, por ejemplo, con respecto a las Figuras 4 y 5.

[0049] En 308, se identifica un PDC probable basándose en las puntuaciones de responsabilidad de compromiso globales calculadas en 306. El PDC probable se selecciona de entre los PDC potenciales identificados en 304. Se calcula una puntuación de responsabilidad de compromiso global para cada uno de los PDC potenciales. Se identifican uno o más PDC potenciales con puntuaciones de responsabilidad de compromiso globales que cumplen con uno o más criterios como PDC probables. Por ejemplo, el criterio es una puntuación mínima de umbral. En algunas realizaciones, se identifica más de un PDC probable, por ejemplo, una pluralidad de PDC potenciales que cumplen con una puntuación mínima de umbral se identifican como PDC probables. Usando el ejemplo del Cuadro 704.6 que se muestra en la Figura 7E, las ubicaciones L1-L3 son PDC potenciales con puntuaciones de responsabilidad de compromiso globales respectivas de 0%, 0% y 66%. Supongamos que el criterio de identificación como PDC probable es del 50%. Entonces, la ubicación L3 se identifica como el PDC probable porque su puntuación (66%) es más alta que la puntuación mínima de umbral, mientras que las otras puntuaciones de responsabilidad globales (0%) no alcanzan la puntuación mínima de umbral.

[0050] Se puede identificar un PDC probable basándose en sus etiquetas/propiedades geográficas. Por ejemplo, se procesan las etiquetas geográficas de ubicaciones de transacciones para determinar patrones en las ubicaciones geográficas de transacciones. Los patrones en las ubicaciones geográficas pueden permitir la agrupación de ubicaciones de transacciones similares en varios niveles de especificidad con el fin de identificar que la agrupación es un PDC probable. Por ejemplo, una ubicación individual (por ejemplo, cada terminal individual dentro de una tienda) no cumple con un umbral para identificar un PDC probable, pero una agrupación de ubicaciones (por ejemplo, un grupo de todos los terminales dentro de una tienda) cumple con el umbral. Esto puede tener como resultado un grupo de ubicaciones de transacciones juntas como una sola entidad que será identificada como un PDC probable.

[0051] Por ejemplo, supongamos que una puntuación de responsabilidad de compromiso global de cada uno de veinte terminales es del 1%. Cuando se observan individualmente, los terminales no cumplen con una puntuación mínima de umbral para reunir las condiciones necesarias para ser considerados PDC. Sin embargo, cuando son observados en su conjunto, la puntuación de responsabilidad de compromiso global es la suma de las puntuaciones individuales, es decir, un 20%. Esta puntuación de responsabilidad de compromiso global podría alcanzar un umbral mínimo (por ejemplo, si el umbral es del 15%), haciendo así que el grupo de terminales reúna las condiciones necesarias para ser considerado un PDC probable. Por lo tanto, se puede identificar todo el grupo de terminales como un PDC probable. Por ejemplo, cuando una fuente de una vulneración de seguridad es un empleado de tienda que ocasionalmente roba información de la tarjeta de crédito de varios terminales diferentes en la tienda, se puede identificar la fuente de esta manera.

[0052] En algunas realizaciones, la agrupación de ubicaciones de transacción basándose en una ubicación geográfica compartida permite llegar a conclusiones sobre una fuente externa de compromiso basada en la agrupación. Por ejemplo, una agrupación de cajeros automáticos en una ubicación compartida se ha identificado como un PDC probable, pero la fuente real del compromiso no es en realidad ninguno de los cajeros automáticos, ya que un *skimmer* de tarjetas instalado en una puerta para acceder a los cajeros automáticos es la fuente real del compromiso.

[0053] En 310, se proporciona información acerca del PDC probable. Esta información permite que se tomen medidas para reparar/eliminar el PDC probable, de modo que la ubicación ya no sea un punto de compromiso, identificar otras cuentas que se han visto comprometidas y/o evitar que otras cuentas se vean comprometidas en el PDC probable. La información sobre el PDC probable puede incluir un identificador del PDC probable, identificadores de ubicación de terminal/transacción asociados, identificadores de ubicación geográfica, identificadores de cuentas que pueden haber sido comprometidas en el PDC probable y/o un identificador de un

periodo de tiempo en el que el PDC probable puede haber sido comprometido (por ejemplo, terminales uno y tres en *Mom and Pop Shop* en Mountain View, California, en el que están involucradas tarjetas de crédito A-F entre el 1 de julio de 2015 y el 4 de julio de 2015). Puede proporcionarse la información sobre el PDC probable (y la información adjunta) a un emisor como el emisor 130 que se muestra en la Figura 1, representado en una GUI o proporcionado de otro modo a un solicitante de la información.

[0054] La información sobre el tiempo/periodo de tiempo proporcionado asociado con el PDC probable incluye el tiempo en términos de horas/minutos/segundos y/o fecha(s) de cuándo es probable que se vio comprometido. Por ejemplo, el tiempo/periodo de tiempo proporciona una identificación de la franja de tiempo durante la que se instaló un *skimmer* de tarjetas de crédito en un dispositivo de transacción. El periodo de tiempo proporcionado se identifica mediante una hora de inicio y/o una hora de finalización.

[0055] En algunas realizaciones, el tiempo de inicio se asocia con un tiempo de una primera transacción/interacción en el PDC probable (por ejemplo, el más antiguo dentro de una franja de tiempo de las transacciones que se analizaron) por una cuenta que se sabe que ha sido comprometida (por ejemplo, una cuenta identificada como comprometida en 302 y/o 304). Por ejemplo, si se supone que el PDC probable ha sido la fuente de la sustracción indebida de la información de la cuenta, también es probable que se vean comprometidas las cuentas que interactuaron con el PDC probable después de la primera cuenta que se sabe que fue comprometida. Debido a que puede existir una primera cuenta real desconocida comprometida en el PDC probable, puede ser insuficiente establecer el tiempo de inicio como la hora de la interacción/transacción en el PDC probable por parte de una cuenta comprometida conocida. Sin embargo, en algunos casos, puede ser deseable hacer una inclusión insuficiente en un esfuerzo por minimizar los impactos en las cuentas que pueden no haber sido comprometidas. En algunas realizaciones, se establece el tiempo de inicio para que sea una cantidad de tiempo establecida antes de la transacción/interacción más antigua en el PDC probable por la cuenta que se sabe que ha sido comprometida. Esta cantidad de tiempo establecida proporciona un periodo de margen en caso de que la primera cuenta comprometida conocida no sea la primera cuenta comprometida en el PDC probable. La cantidad de tiempo establecida puede ser ajustable para permitir el ajuste dinámico de una inclusión excesiva/insuficiente de la franja de tiempo comprometida asociada con el PDC probable.

[0056] En algunas realizaciones, la hora de finalización identifica un punto en el tiempo cuando se ha verificado que el PDC probable ya no es una fuente de compromiso. La hora de finalización es opcional porque el PDC probable identificado aún puede ser una fuente de compromiso en el momento de la detección. Por ejemplo, un *skimmer* de tarjetas de crédito que permanece operativo en un dispositivo de transacción identificado como el PDC probable.

[0057] Se puede utilizar la información proporcionada en 310 para reparar dispositivos de transacción asociados con el PDC probable. Por ejemplo, el emisor 130 mostrado en la Figura 1 puede inspeccionar los dispositivos de transacción asociados con el PDC probable para determinar si se han instalado en los mismos dispositivos maliciosos.

[0058] Se puede utilizar la información proporcionada en 310 para identificar las cuentas adicionales que fueron potencialmente comprometidas en el PDC probable. Por ejemplo, es probable que las cuentas que interactuaron con el PDC probable durante un periodo de tiempo en el que el PDC probable estaba probablemente comprometido tuvieran la información de su cuenta sustraída indebidamente en el PDC probable. Los titulares de estas cuentas desconocen potencialmente el compromiso y es posible que aún no hayan informado sobre la cuenta que está comprometida y la identificación de estas cuentas adicionales permite tomar medidas proactivas para evitar el uso no autorizado adicional de dichas cuentas (por ejemplo, la cancelación de estas cuentas y la emisión de nuevas tarjetas de crédito).

[0059] La Figura 4 es un diagrama de flujo en el que se ilustra una realización de un proceso para calcular una o más puntuaciones de responsabilidad. El proceso de la Figura 4 puede implementarse, al menos en parte, en uno o más componentes de los nodos 142.1 y/o 142.2 mostrados en la Figura 1. En algunas realizaciones, el proceso mostrado en la Figura 4 se realiza mediante un procesador, por ejemplo, un procesador implementado por un solo chip o por múltiples procesadores como el procesador 202 mostrado en la Figura 2. En algunas realizaciones, al menos una parte del proceso de la Figura 4 está incluida en el paso 306 de la Figura 3.

[0060] Se explicará el proceso mostrado en la Figura 4 utilizando el ejemplo de la Figura 6. La Figura 6 es un diagrama conceptual de transacciones que involucran las cuentas C1-C4 y las ubicaciones L1-L3. Se sabe que las cuentas C3 y C4 están comprometidas, pero se desconoce qué ubicación o ubicaciones de entre L1-L3 son un PDC. La interacción entre una cuenta y una ubicación se representa mediante una línea entre la cuenta y la ubicación. Una interacción incluye proporcionar la cuenta a la ubicación, como por ejemplo al pasar una tarjeta de crédito por un TPV. En el ejemplo que se muestra en la Figura 6, la cuenta C1 ha interactuado con las ubicaciones L1 y L2; la cuenta C2 ha interactuado con las ubicaciones L1, L2 y L3; la cuenta C3 ha interactuado con las ubicaciones L1, L2 y L3; y la cuenta C4 ha interactuado con las ubicaciones L2 y L3. Dada la información transaccional que representa el escenario de la Figura 6, la Figura 4 identifica un PDC probable de entre las ubicaciones L1, L2 y L3.

[0061] Si nos fijamos de nuevo en la Figura 4, en 402, se genera una matriz de cálculo informático que representa cálculos de puntuación de responsabilidad. Por ejemplo, la matriz de cálculo informático no solo va a servir como un punto de partida de los cálculos para identificar los PDC probables, sino también para almacenar valores de cálculo intermedios y ayudar a distribuir el procesamiento entre los nodos de un clúster. Un ejemplo de esta matriz incluye una matriz de cálculo informático con filas correspondientes a cuentas y columnas correspondientes a ubicaciones. Los valores de la matriz de ejemplo incluyen una medida de certeza basada en si la cuenta ha interactuado con la ubicación. Las operaciones (por ejemplo, 406/408) se realizan utilizando datos de la matriz de ejemplo para identificar en última instancia un patrón que sugiere una ubicación que es una fuente de compromiso. El cuadro 702.1 mostrado en la Figura 7A es un ejemplo de la matriz de cálculo informático relleno con valores iniciales.

[0062] En algunas realizaciones, la matriz de cálculo informático se inicializa con puntuaciones de responsabilidad de elemento predeterminadas. Una puntuación de responsabilidad de elemento es la medida de certeza de que una cuenta en particular asigna a una ubicación en particular, y la puntuación representa el grado de responsabilidad que la cuenta en particular asigna a la ubicación en particular por ser la fuente del compromiso de la cuenta. En algunas realizaciones, la matriz incluye puntuaciones de responsabilidad de compromiso globales para ubicaciones respectivas. Una puntuación de responsabilidad de compromiso global para una ubicación en particular representa una medida de certeza de que la ubicación en particular es un PDC, por ejemplo, la probabilidad de que la ubicación es una fuente de compromisos de cuentas.

[0063] La matriz de cálculo informático se rellena basándose en los datos de transacciones, por ejemplo, la información transaccional recibida en 302. Las cuentas identificadas a partir de la información transaccional están representadas por filas en el cuadro 702.1 y las ubicaciones identificadas a partir de la información transaccional están representadas por las columnas en el cuadro 702.1. Si una cuenta en particular realiza una transacción en una ubicación en particular, se realiza un seguimiento de esta información. Para una cuenta determinada, las ubicaciones se resumen y se asigna la responsabilidad por igual a cada ubicación basándose en el número total de ubicaciones que han interactuado con la cuenta.

[0064] Como se describe en el presente, el procesamiento de la información transaccional produce cuentas y PDC potenciales. La Figura 7A es un diagrama conceptual de un estado inicial de una matriz de cálculo informático para identificar uno o más PDC probables. La Figura 7A corresponde al escenario de transacción mostrado en la Figura 6 en el que las tarjetas C3 y C4 se identifican como tarjetas comprometidas y las tarjetas C1-C4 han interactuado con al menos una de las ubicaciones L1-L3 durante un periodo de tiempo determinado. El escenario de transacción mostrado en la Figura 6 está representado por información de transacciones que incluye una lista de cuentas C1-C4, una lista de ubicaciones L1-L3, tiempos de transacciones y otra información descrita con respecto a 302.

[0065] En algunas realizaciones, una matriz de cálculo informático se construye como se muestra en el cuadro 702.1 de la Figura 7A basándose en la información de transacción recibida. El Cuadro 1 muestra un ejemplo de información de transacción recibida en forma de cuadro.

Cuadro 1

ID de transacción	Hora	Ubicación	Cuenta	Compra	...
00001	13:05	L2	C1	Artículo A x 1	...
00002	13:06	L1	C2	Artículo B x 5	...

La información de la transacción se analiza para identificar las cuentas y ubicaciones involucradas en las transacciones. Usando el escenario de la Figura 6 y el ejemplo del cuadro 702.1, las cuentas se identifican a partir de la información de transacción recibida como las cuentas C1-C4. Las ubicaciones se identifican a partir de la información de transacción recibida como las ubicaciones L1-L3. Para cada combinación de cuenta y ubicación (por ejemplo, representada por una celda en el cuadro 702.1), se determina si la cuenta y la ubicación han interactuado entre sí. Por ejemplo, proporcionar detalles de la cuenta a la ubicación es una interacción entre la cuenta y la ubicación. Las cuentas y ubicaciones que no han interactuado son marcadas en consecuencia. En el ejemplo que se muestra en el cuadro 702.1, la cuenta C4 y la ubicación L1 no han interactuado y la cuenta C1 y la ubicación L3 no han interactuado. Esto se marca con un guion “-” en el cuadro 702.1. Si una ubicación no ha interactuado con ninguna tarjeta, esa ubicación se elimina del cuadro. Es decir, la columna que representa la ubicación se elimina del cuadro. En otras palabras, en el ejemplo mostrado, las ubicaciones que han interactuado con al menos una de las cuentas C1-C4 son las ubicaciones L1-L3.

[0066] En algunas realizaciones, la matriz de cálculo informático se inicializa de manera uniforme con responsabilidad asignada por igual entre las ubicaciones con la que una cuenta ha interactuado. En el ejemplo en el que una cuenta comprometida ha interactuado con tres ubicaciones, es muy probable que una de las tres ubicaciones sea un PDC. En este ejemplo, la matriz de cálculo informático que representa las puntuaciones de responsabilidad de elemento se inicializa con el 33% de la responsabilidad asignada a cada una de las tres ubicaciones. Alternativamente, la matriz de cálculo informático que representa las puntuaciones de responsabilidad de elemento se inicializa con medidas de certeza que se ponderan basándose en una serie de transacciones que se

han procesado en una ubicación determinada. Supongamos que la ubicación L1 ha procesado dos transacciones de la cuenta C3, la ubicación L2 ha procesado una transacción de la cuenta C3 y la ubicación L3 ha procesado una transacción de la cuenta C3. Las puntuaciones de responsabilidad de elemento pueden ponderarse para asignar, para la cuenta C3, más responsabilidad a la ubicación L1 porque han existido más interacciones con la ubicación L1. Por ejemplo, la puntuación de responsabilidad de elemento para C3/L1 es del 50%, la puntuación de responsabilidad de elemento para C3/L2 es del 25% y la puntuación de responsabilidad de elemento para C3/L3 es del 25%, lo que es proporcional al número de interacciones entre cada ubicación y cada cuenta.

[0067] En el cuadro 702.1 mostrado en la Figura 7A se muestran puntuaciones de responsabilidad de elemento en un estado inicial. Las puntuaciones de responsabilidad de elemento representan, desde la perspectiva de cada cuenta, el grado de responsabilidad que cada cuenta C1-C4 asigna a una ubicación en particular (L1-L3). En este ejemplo, el cuadro 702.1 se inicializa uniformemente. Por ejemplo, en el caso de la cuenta C3, que está comprometida, es muy probable que una de las ubicaciones L1-L3 sea la responsable de su compromiso. Inicialmente, la cuenta C3 asigna la responsabilidad por igual entre cada una de las ubicaciones L1-L3. Esto se muestra en el cuadro como un 33% de responsabilidad para cada una de las ubicaciones L1-L3 para la tarjeta C3. De manera similar, en el caso de la cuenta C4, que está comprometida, es muy probable que una de las ubicaciones L2 y L3 sea la responsable de su compromiso. Inicialmente, la cuenta C4 asigna la responsabilidad por igual entre cada una de las ubicaciones L2 y L3. Esto se muestra en el cuadro como un 50% de responsabilidad para cada una de las ubicaciones L2 y L3 para la tarjeta C4. Las tarjetas C1 y C2 no están comprometidas. Por consiguiente, no implican responsabilidad para ninguna de las localizaciones (0%).

[0068] Por lo que respecta de nuevo a la Figura 4, en 404, la matriz de cálculo informático está preprocesada. El preprocesamiento simplifica los cálculos posteriores. Por ejemplo, el preprocesamiento elimina los cálculos para las cuentas no comprometidas. En otras palabras, no se realizan cálculos para tarjetas que no se han comprometido, lo que ahorra recursos de cálculo y produce una determinación más rápida. El cuadro 703.1 mostrado en la Figura 7A es un ejemplo de una matriz preprocesada. Las filas para las cuentas C1 y C2 se eliminan del cuadro 702.1 para generar el cuadro preprocesado 703.1 porque las cuentas C1 y C2 no están comprometidas en este ejemplo (véase la discusión relacionada con la Figura 6) y no es necesario calcular las puntuaciones de responsabilidad de elemento para esas cuentas. El número de cuentas con las que interactúa una ubicación se registra para los cálculos. Por ejemplo, el número es útil para determinar una puntuación de responsabilidad de compromiso, como se describe en mayor detalle en el presente. La entrada para C4/L1 está marcada con un guion porque la cuenta C4 y la ubicación L1 no interactuaron entre sí y no es necesario realizar ningún cálculo para esa combinación.

[0069] En 406, se calcula una puntuación de responsabilidad de compromiso global para cada PDC potencial. La puntuación de responsabilidad de compromiso global es la medida de certeza de que una ubicación es un PDC. La puntuación de responsabilidad de compromiso global se calcula basándose en las puntuaciones de responsabilidad de elemento. El cálculo de una puntuación de responsabilidad de compromiso global incluye la determinación de una responsabilidad combinada para cada ubicación. La responsabilidad combinada es la suma de las puntuaciones de responsabilidad de elemento (responsabilidades asignadas por las cuentas) para la ubicación determinada. Se obtiene la puntuación de responsabilidad de compromiso global para una ubicación determinada (p_j) dividiendo la responsabilidad combinada ($\sum_i b_{ij}$) por el número (n_j) de cuentas con las que interactuó la ubicación, donde p es una medida de certeza, b es una puntuación de responsabilidad de elemento (que se describe en mayor detalle en el presente), j es una ubicación (columna en el cuadro 702.1) e i es una cuenta (fila en el cuadro 702.1). La puntuación de responsabilidad de compromiso global para una ubicación determinada viene dada por:

$$p_j = \frac{\sum_i b_{ij}}{n_j} \quad (1)$$

[0070] La Figura 7B es un diagrama conceptual de un estado de una matriz de cálculo informático para identificar uno o más PDC probables después de una primera iteración. El cuadro 704.2 es un ejemplo de una matriz de cálculo informático de puntuaciones de responsabilidad de compromiso globales en la que las puntuaciones de responsabilidad calculadas se redondean al número entero más cercano. El cuadro 704.2 enumera las ubicaciones L1-L3 y las puntuaciones de responsabilidad de compromiso globales correspondientes (probabilidades). Las puntuaciones de responsabilidad de compromiso globales se generan a partir de las puntuaciones de responsabilidad de elemento iniciales que se muestran en el cuadro 703.1 de la siguiente manera. La ubicación L1 tiene una responsabilidad combinada del 33%, la ubicación L2 tiene una responsabilidad combinada del 83% y la ubicación L3 tiene una responsabilidad combinada del 83%. Para la ubicación L1, la puntuación de responsabilidad de compromiso global es del 11% porque la ubicación L1 tiene una responsabilidad combinada del 33% y tres cuentas interactuaron con la ubicación L1. Para la ubicación L2, la puntuación de responsabilidad de compromiso global es del 21% porque la ubicación L2 tiene una responsabilidad combinada del 83% e interactuó con cuatro tarjetas. Para la ubicación L3, la puntuación de responsabilidad de compromiso global es del 28% porque la ubicación L3 tiene una responsabilidad combinada del 83% e interactuó con tres tarjetas.

[0071] Si nos fijamos de nuevo en la Figura 4, en 408, se actualizan las puntuaciones de responsabilidad de elemento. Las puntuaciones de responsabilidad de elemento se actualizan basándose en las puntuaciones de responsabilidad de compromiso globales calculadas en 406. Las puntuaciones de responsabilidad de elemento se

actualizan para reflejar proporcionalmente la puntuación de responsabilidad de compromiso global. Se obtiene una puntuación de responsabilidad de elemento para una cuenta determinada y una ubicación determinada (b_{ij}) mediante la proporción de la puntuación de responsabilidad de compromiso global para la ubicación determinada (p_j) con respecto a la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que la cuenta ha interactuado ($\sum_{k \in B} p_k$), donde b es una medida de certeza para un elemento, p es una puntuación de responsabilidad de compromiso global (que se describe en mayor detalle en el presente), k es una cuenta, B es un conjunto de PDC potenciales (por ejemplo, $k \in B$ significa que una cuenta ha interactuado con un PDC potencial), j es una ubicación (columna en el cuadro 702.1) e i es una cuenta (fila en el cuadro 702.1). La puntuación del elemento para una celda (i, j) en el cuadro 702.2 viene dada por:

$$b_{ij} = \frac{p_j}{\sum_{k \in B} p_k} \quad (2)$$

[0072] Si nos fijamos de nuevo en el ejemplo de la Figura 7B, tras la determinación de las puntuaciones de responsabilidad de compromiso globales mostradas en el cuadro 704.2 de la Figura 7B, se actualizan las puntuaciones de responsabilidad de elemento para reflejar que se responsabiliza más a la ubicación L3 (28%) que a la ubicación L2 (21%) y la ubicación L1 (11%). Las puntuaciones de responsabilidad de elemento actualizadas se muestran en el cuadro 702.2 y se generan de la siguiente manera. Para la cuenta C3, la puntuación de responsabilidad de elemento para la ubicación L1 es del 18% porque la puntuación de responsabilidad de compromiso global para L1 es del 11% y la suma de las puntuaciones de responsabilidad de compromiso globales para ubicaciones con las que C3 ha interactuado es del 60% (11% + 21% + 28%), la puntuación de responsabilidad de elemento para la ubicación L2 es del 35% porque la puntuación de responsabilidad de compromiso global para L2 es del 21% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es del 60% y la puntuación de responsabilidad de elemento para la ubicación L3 es del 47% porque la puntuación de responsabilidad de compromiso global para L3 es del 28% y la suma de las puntuaciones de responsabilidad de compromiso globales para ubicaciones con las que C3 ha interactuado es del 60%. Para la cuenta C4, la puntuación de responsabilidad de elemento para la ubicación L2 es del 43% porque la puntuación de responsabilidad de compromiso global para L2 es del 21% y la suma de las puntuaciones de responsabilidad de compromiso globales para ubicaciones con las que C4 ha interactuado es del 49% (21% + 28%), y la puntuación de responsabilidad de elemento para la ubicación L3 es del 57% porque la puntuación de responsabilidad de compromiso global para L3 es del 28% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C4 ha interactuado es del 49%.

[0073] Para resumir, 406 y 408 representan un procedimiento alternativo en el que las puntuaciones de responsabilidad de elemento y las puntuaciones de responsabilidad de compromiso globales se actualizan en sucesión hasta la convergencia. En cada iteración, las puntuaciones de responsabilidad de elemento y las puntuaciones de responsabilidad de compromiso globales se actualizan sucesivamente. Es decir, en una sola iteración, se realizan 406 y 408. En el ejemplo en el que las puntuaciones de responsabilidad de elemento se inicializan de manera sustancialmente uniforme, cada iteración actualiza las puntuaciones de responsabilidad de elemento para diferenciar entre los PDC potenciales que tienen más probabilidades de ser un PDC probable y los PDC potenciales que tienen menos probabilidades de ser un PDC probable. Después de una sucesión de iteraciones, el PDC probable se hace cada vez más evidente a medida que las diferencias de valor de puntuación de responsabilidad entre sí se amplifican hasta la convergencia. En la convergencia, hay suficiente certeza (según lo define un umbral) de que un PDC potencial es un PDC probable y se devuelve el PDC probable.

[0074] En 410, se determina si las puntuaciones de responsabilidad de compromiso globales han convergido. La repetición de cálculos (por ejemplo, la repetición de 406 y 408) para cada iteración minimiza el número de PDC probables identificados y aumenta la certeza de que los PDC probables identificados sean correctos. A medida que se repiten las iteraciones, las tendencias se hacen evidentes. Por ejemplo, las puntuaciones de responsabilidad pueden tener una tendencia más baja para algunos PDC potenciales, mientras que las puntuaciones de responsabilidad pueden tener una tendencia más alta para otros PDC potenciales después de cada iteración para amplificar las diferencias de puntuación de responsabilidad entre los PDC potenciales. Los cálculos se pueden repetir hasta que los valores se estabilizan. A medida que se realizan más iteraciones, normalmente la magnitud de los cambios entre iteraciones disminuye. Los valores se estabilizan cuando la magnitud de los cambios alcanza un umbral, lo que indica que los valores cambian de forma relativamente mínima entre iteraciones. Los valores estables indican que se han identificado los PDC más probables con un nivel de certeza relativamente alto. En otras palabras, a medida que avanzan las iteraciones de los cálculos, los resultados de los cálculos se vuelven más precisos (por ejemplo, es más probable que los PDC probables identificados sean correctos). La convergencia incluye determinar si los valores se han estabilizado. En algunas realizaciones, la determinación se realiza determinando la diferencia entre puntuaciones de responsabilidad de compromiso globales sucesivas. Cuando las puntuaciones de responsabilidad de compromiso globales ya no cambian demasiado (medidas en función de un criterio de convergencia), entonces las puntuaciones han convergido.

[0075] La convergencia se basa en criterios predefinidos, como por ejemplo un umbral. Por ejemplo, cuando la diferencia entre una puntuación de responsabilidad de compromiso global actual y una puntuación de responsabilidad de compromiso global para una iteración anterior está por debajo de un umbral de convergencia, se

determina que las puntuaciones de responsabilidad de compromiso globales han convergido. En otras palabras, si la diferencia entre las puntuaciones de responsabilidad de compromiso globales sucesivas alcanzan un umbral de convergencia, se determina que las puntuaciones de responsabilidad de compromiso globales han convergido. En otro ejemplo, los criterios predefinidos incluyen un periodo de tiempo y/o una serie de iteraciones de modo que se determine que las puntuaciones de responsabilidad de compromiso globales han convergido después de que haya transcurrido un periodo de tiempo o se haya realizado un número predefinido de iteraciones. La convergencia indica que se ha identificado un posible punto de compromiso.

[0076] Si las puntuaciones de responsabilidad de compromiso globales han convergido, el control pasa a 412, en donde se proporcionan las puntuaciones de responsabilidad convergentes. Si las puntuaciones de responsabilidad de compromiso globales no han convergido, el control pasa a 406, en donde se calculan las puntuaciones de responsabilidad de compromiso globales. Se pueden realizar 406 y 408 hasta que se determine que las puntuaciones de responsabilidad de compromiso globales han convergido (410). En el ejemplo de la Figura 7B, las puntuaciones de responsabilidad de compromiso globales no han convergido. Por tanto, se realiza una segunda iteración en la que se calculan/actualizan las puntuaciones de responsabilidad de compromiso globales basándose en las puntuaciones de elemento actualizadas mostradas en el cuadro 704.3 y el cuadro 702.3 de la Figura 7C.

[0077] En algunos casos, los valores pueden no converger nunca o pueden requerir demasiadas iteraciones de cálculos (por ejemplo, iteraciones de 406 y 408 de la Figura 4) para llegar a una convergencia, y puede ser ineficiente realizar iteraciones hasta que los valores se estabilizan. Se puede imponer un límite de tiempo, de modo que se determine que la convergencia ha ocurrido después de un umbral de tiempo. El umbral puede definirse mediante una serie de iteraciones, un periodo de tiempo desde que comenzaron los cálculos, etc. En algunas realizaciones, los criterios de convergencia incluyen un umbral de tiempo. Por ejemplo, un PDC potencial con la puntuación más alta en relación con otros PDC después de un periodo de tiempo de umbral desde que comiencen los cálculos satisface los criterios de convergencia. En otro ejemplo, después de que haya transcurrido un periodo de tiempo desde que comenzaron los cálculos, se determina que un PDC potencial ha convergido si se produce un cambio en la puntuación (por ejemplo, una puntuación generada por una segunda iteración en comparación con una puntuación generada por una primera iteración) que satisface una puntuación mínima de umbral. Por ejemplo, el umbral de tiempo define un periodo de tiempo que ha transcurrido desde que se calculó una primera puntuación de responsabilidad de compromiso global. En algunos casos, la adición del umbral de tiempo a los criterios de convergencia permite determinar un resultado antes de lo que se determinaría sin el umbral de tiempo. Por ejemplo, se determina que las puntuaciones han convergido si, 5 ms después de que comenzaron los cálculos, se alcanza un umbral de convergencia. En la Figura 7B se muestran puntuaciones después de una primera iteración. En este ejemplo, las puntuaciones aún no han convergido porque no se alcanza un umbral de convergencia.

[0078] La Figura 7C es un diagrama conceptual de un estado de una matriz de cálculo informático para identificar uno o más PDC probables después de una segunda iteración. El cuadro 704.3 muestra las puntuaciones de responsabilidad de compromiso globales para cada una de las ubicaciones L1-L3 en las que las puntuaciones de responsabilidad calculadas se redondean al número entero más cercano. La ubicación L1 tiene una responsabilidad combinada del 18%, la ubicación L2 tiene una responsabilidad combinada del 78% y la ubicación L3 tiene una responsabilidad combinada del 104%. Para la ubicación L1, la puntuación de responsabilidad de compromiso global es del 6% porque la ubicación L1 tiene una responsabilidad combinada del 18% y tres cuentas interactuaron con la ubicación L1. Para la ubicación L2, la puntuación de responsabilidad de compromiso global es del 20% porque la ubicación L2 tiene una responsabilidad combinada del 78% e interactuó con cuatro tarjetas. Para la ubicación L3, la puntuación de responsabilidad de compromiso global es del 35% porque la ubicación L3 tiene una responsabilidad combinada del 104% e interactuó con tres tarjetas. A continuación, las puntuaciones de los elementos se actualizan basándose en las puntuaciones de responsabilidad de compromiso globales, como se muestra en el cuadro 702.3, y de acuerdo con los procedimientos descritos en el presente.

[0079] La Figura 7D es un diagrama conceptual de un estado de una matriz de cálculo informático para identificar uno o más PDC probables después de una tercera iteración. El cuadro 704.4 muestra las puntuaciones de responsabilidad de compromiso globales para cada una de las ubicaciones L1-L3 en las que las puntuaciones de responsabilidad calculadas se redondean al número entero más cercano. Después, las puntuaciones de los elementos se actualizan basándose en las puntuaciones de responsabilidad de compromiso globales, como se muestra en el cuadro 702.4, y de acuerdo con los procedimientos descritos en el presente. Se pueden repetir el cálculo de las puntuaciones de responsabilidad de compromiso globales y el de las puntuaciones de responsabilidad de elemento hasta alcanzar la convergencia.

[0080] La Figura 7E es un diagrama conceptual de una matriz de cálculo informático para identificar uno o más PDC probables que incluyen puntuaciones de responsabilidad de compromiso globales convergentes. El cuadro 704.5 muestra las puntuaciones de responsabilidad de compromiso globales (redondeadas al número entero más cercano) que no han convergido. Los procedimientos descritos en el presente se repiten (representados por “...” en la Figura 7E) hasta la convergencia. El cuadro 702.5 muestra las puntuaciones de responsabilidad de elemento en una iteración inmediatamente antes de la convergencia. Es decir, las puntuaciones de responsabilidad de compromiso globales del cuadro 704.6 se generan a partir de las puntuaciones de responsabilidad de elemento del

cuadro 702.5.

[0081] Se determina que las puntuaciones de responsabilidad de compromiso globales del cuadro 704.6 han convergido porque la diferencia entre la puntuación de responsabilidad de compromiso global actual (66% para la ubicación L3) y una puntuación de responsabilidad de compromiso global de una iteración anterior (no mostrada) es inferior a un umbral. En este ejemplo, las diferencias entre la puntuación de responsabilidad de compromiso global para la ubicación L3 en iteraciones anteriores no alcanzan este umbral (7% entre la iteración 704.2 y 704.3, 5% entre la iteración 704.3 y 704.4 y 6% entre la iteración 704.4 y 704.5).

[0082] Si nos fijamos de nuevo en la Figura 4, en 412, se proporcionan puntuaciones de responsabilidad de compromiso globales convergentes. Las puntuaciones de responsabilidad de compromiso globales convergentes reflejan el PDC probable, donde el PDC probable tiene una puntuación de responsabilidad de compromiso global que alcanza un umbral. Se pueden proporcionar las puntuaciones de responsabilidad de compromiso globales como una lista de puntuaciones. Usando el ejemplo de la Figura 7E, se proporcionan las puntuaciones de responsabilidad de compromiso globales correspondientes a las probabilidades de que una ubicación sea un PDC probable (66% para la ubicación L3 y 0% para cada una de las ubicaciones L1 y L2). En algunas realizaciones, las puntuaciones de responsabilidad globales convergentes se utilizan en 308 de la Figura 3 para identificar un PDC probable. Usando el cuadro 704.6 mostrado en la Figura 7E, las ubicaciones L1-L3 son PDC potenciales con puntuaciones de responsabilidad de compromiso globales respectivas de 0%, 0% y 66%. Supongamos que el criterio de PDC probable requiere superar un umbral del 50%. Entonces, la ubicación L3 se identifica como el PDC probable porque su puntuación (66%) supera el umbral, mientras que las puntuaciones de las otras ubicaciones (0%) no superan el umbral.

[0083] En algunas realizaciones, se identifica más de un PDC probable. Por ejemplo, múltiples transacciones/ubicaciones en el mismo comerciante pueden verse comprometidas. En este caso, las puntuaciones de responsabilidad de los elementos se actualizan de la siguiente manera. Se obtiene una puntuación de responsabilidad de elemento para una cuenta determinada y una ubicación determinada (b_{ij}) mediante la proporción de la puntuación de responsabilidad de compromiso global para la ubicación determinada (p_j) con respecto a la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que la cuenta ha interactuado y las ubicaciones no corresponden al mismo comerciante ($\sum_{k \in B \text{ y } merch(j) \neq merch(k)} p_k$), donde b es una medida de certeza para un elemento, p es una puntuación de responsabilidad de compromiso global (que se describe en mayor detalle en el presente), k es una cuenta, B es un conjunto de PDC potenciales y $merch()$ es un comerciante (por ejemplo, $k \in B$ y $merch(j) \neq merch(k)$ significa que una cuenta ha interactuado con un PDC potencial y el PDC potencial no corresponde al mismo comerciante), j es una ubicación (columna en el cuadro 702.1) e i es una cuenta (fila en el cuadro 702.1). La puntuación del elemento para una celda (i, j) en el cuadro 702.2 viene dada por:

$$b_{ij} = \frac{p_j}{p_j + \sum_{k \in B \text{ y } merch(j) \neq merch(k)} p_k} \quad (3)$$

[0084] En una realización alternativa, una puntuación de responsabilidad de compromiso global se basa en las probabilidades previamente calculadas y la certeza relativa entre una puntuación de responsabilidad de compromiso global y una o más puntuaciones de responsabilidad de compromiso globales previamente calculadas. Por ejemplo, la puntuación de responsabilidad de compromiso global se calcula ponderando (por los factores α y β) las probabilidades previas. La puntuación de responsabilidad de compromiso global para una ubicación dada (p_j) se obtiene dividiendo (i) la suma de una responsabilidad combinada ($\sum_i b_{ij}$) y un factor de ponderación (α) por (ii) la suma de un número (n_j) de cuentas con las que interactuó la ubicación, el factor de ponderación (α) y un segundo factor de ponderación (β), donde p es una medida de certeza, b es una puntuación de responsabilidad del elemento (que se describe en mayor detalle en el presente), j es una ubicación (columna en el cuadro 702.1) e i es una cuenta (fila en el cuadro 702.1). La puntuación de responsabilidad de compromiso global para una ubicación determinada viene dada por:

$$p_j = \frac{\sum_i b_{ij} + \alpha}{n_j + \alpha + \beta} \quad (4)$$

En algunas realizaciones, se calcula una puntuación de responsabilidad de compromiso global de acuerdo con la ecuación (4) en lugar de la ecuación (1). En algunas realizaciones, se pueden seleccionar α y β para alterar la tasa de convergencia (por ejemplo, hacer que la convergencia se produzca antes). α y β definen la certeza en cálculos de puntuación anteriores (por ejemplo, procedentes de iteraciones anteriores) en comparación con una iteración actual. Esto afecta la cantidad de iteraciones que se realizan para tener la certeza de que una ubicación se ha visto comprometida, por ejemplo, para alcanzar la convergencia. α es funcionalmente una cuenta comprometida "virtual" que interactuó con una ubicación (j) y β es funcionalmente una cuenta no comprometida "virtual" que interactuó con una ubicación (j). Por consiguiente, dos ubicaciones con una misma proporción de tarjetas comprometidas y no comprometidas tendrán sus medidas de certeza ajustadas de acuerdo con la ponderación del grado de diferencia con respecto a una distribución anterior (iteración previa).

[0085] La Figura 5 es un diagrama de flujo en el que se ilustra una realización de un proceso para identificar uno o

más PDC que incluyen la distribución del cálculo de una o más puntuaciones de responsabilidad. El proceso de la Figura 5 puede implementarse al menos en parte en uno o más nodos del clúster 140 mostrado en la Figura 1. En algunas realizaciones, el proceso mostrado en la Figura 5 es realizado por un procesador, por ejemplo, un procesador implementado por un solo chip o por múltiples procesadores, como por ejemplo el procesador 202
5 mostrado en la Figura 2. En algunas realizaciones, el cálculo se puede realizar utilizando una unidad de procesamiento gráfico (GPU, por sus siglas en inglés, *Graphics Processing Unit*). En algunas realizaciones, al menos una parte del proceso de la Figura 5 se produce antes de 406 en la Figura 4.

[0086] La identificación de los PDC probables se realiza idealmente de manera eficiente, rápida y precisa. Cuanto más tiempo se tarda en detectar un punto de compromiso, más probable es que se utilice la información robada. Debido a la gran cantidad de datos de transacciones que debe procesarse y cálculos que deben realizarse, se requieren mejoras en el procesamiento de datos y el funcionamiento de los sistemas informáticos para lograr el rendimiento deseado.

[0087] En algunas realizaciones, el cálculo de una o más puntuaciones de responsabilidad se distribuye entre una pluralidad de máquinas. Es decir, se utiliza más de una máquina para calcular las puntuaciones de responsabilidad de elemento y/o las puntuaciones de responsabilidad de compromiso globales. Cada máquina actualiza una parte de la matriz de cálculo informático y los cálculos distribuidos se sincronizan como se explica a continuación. El uso de una pluralidad de nodos para gestionar estos cálculos implica la división estratégica y la distribución de los cálculos entre los nodos.

[0088] De acuerdo con el proceso de la Figura 5, los cálculos se gestionan de manera eficiente. El trabajo de realizar los cálculos se puede dividir entre múltiples nodos de un clúster basado en elementos de matriz en lugar de simplemente asignar un cálculo particular a un nodo particular basándose en cada cálculo individual. Por ejemplo, se asigna un nodo informático de un clúster para gestionar todos los cálculos necesarios y completar uno o más elementos asignados de la matriz, como se describe en mayor detalle en el presente.

[0089] En 502, una matriz de cálculo informático se divide en bloques de cálculo, donde cada bloque o grupo de bloques se asigna a un nodo respectivo para el cálculo. A modo de ejemplo no limitativo, las formas de dividir la matriz de cálculo informático incluyen las basadas en transacciones y las basadas en elementos. Cada nodo realiza un proceso como el que se muestra en la Figura 4 para su parte de la matriz de cálculo informático. En algunas realizaciones, la división de la matriz de cálculo informático corresponde a un algoritmo de distribución como MapReduce, Spark, etc.

[0090] Una forma de dividir la matriz de cálculo informático es dividir la matriz de cálculo en bloques de cálculo basados en elementos de la matriz de cálculo. Por ejemplo, las celdas/elementos de la matriz de cálculo deben llenarse con valores calculados y las celdas/elementos de la matriz de cálculo se dividen entre los nodos de manera que el nodo asignado a una celda/elemento particular debe realizar los cálculos necesarios para determinar el valor calculado y llenar la celda/elemento particular. La agrupación de una o más celdas/elementos de matriz en un bloque de cálculo puede incluir elementos adyacentes de una fila particular de la matriz de cálculo, una columna particular de la matriz de cálculo y/o un bloque de elementos adyacentes de una parte de una columna y/o fila particular.

[0091] En el ejemplo del cuadro mostrado en 702,1 en la Figura 7A, un ejemplo de división basada en elementos asigna cada una de las ubicaciones L1-L3 a un nodo respectivo, de manera que cada nodo gestiona los cálculos para una ubicación particular. Es decir, un nodo gestiona los cálculos para la columna L1, otro nodo gestiona los cálculos para la columna L2 y otro nodo gestiona los cálculos para la columna L3. Por supuesto, los cálculos se pueden distribuir de manera desigual. Por ejemplo, un nodo gestiona los cálculos para la columna L1 y otro nodo gestiona los cálculos para las columnas L2 y L3. Otro ejemplo de división basada en elementos para el cuadro de la Figura 7A asigna cada una de las cuentas C1-C4 a un nodo respectivo, de manera que cada nodo gestiona los cálculos para una cuenta particular. Es decir, un primer nodo gestiona los cálculos para la fila C1, un segundo nodo gestiona los cálculos para la fila C2, un tercer nodo gestiona los cálculos para la fila C3 y un cuarto nodo gestiona los cálculos para la fila C4. Por supuesto, los cálculos se pueden distribuir de manera desigual. Por ejemplo, un nodo gestiona los cálculos para las cuentas C1-C3 y otro nodo gestiona los cálculos para la cuenta C4.

[0092] Otra forma de dividir la matriz de cálculo informático se basa en la división de los datos de entrada (por ejemplo, los datos de transacción) necesaria para calcular los elementos de la matriz de cálculo. Cada bloque de cálculo puede incluir un subconjunto de datos de entrada necesarios para ser procesados para la matriz de cálculo. Por ejemplo, la división basada en datos de transacciones asigna un subgrupo de transacciones de la lista de transacciones recibidas (por ejemplo, las transacciones recibidas en 302 de la Figura 3) a un nodo para el cálculo. En la división basada en datos, cada nodo realiza cálculos individuales basados en su parte de datos asignados y los resultados de los nodos se combinan para generar la matriz resultante. En contraste con la división basada en elementos, la división basada en datos puede tener como resultado que múltiples nodos realicen cálculos para determinar un elemento/celda de matriz de cálculo.

[0093] Supongamos que la información de la transacción recibida está representada por el Cuadro 1 (que se reproduce a continuación):

5

Cuadro 1

ID de transacción	Hora	Ubicación	Cuenta	Compra	...
00001	13:05	L2	C1	Artículo A x 1	...
00002	13:06	L1	C2	Artículo B x 5	...

[0094] Un ejemplo de división basada en datos de transacción asigna la Transacción 00001 a un primer nodo y asigna la Transacción 00002 a un segundo nodo. En otro ejemplo, un nodo puede administrar varias transacciones (por ejemplo, un primer nodo gestiona la Transacción 00001 y la Transacción 00002). En otro ejemplo, los datos de transacción se dividen según la ubicación, la cuenta, etc. Un primer nodo gestiona todas las transacciones de la ubicación L1. Cada nodo realiza el proceso mostrado en la Figura 4 para su respectiva transacción. Usando el ejemplo de 702.1 que se muestra en la Figura 7A, varios nodos son responsables de cada elemento y las puntuaciones de responsabilidad de elemento se obtienen combinando los resultados determinados por cada nodo. Para la cuenta C3 y la ubicación L1, se obtiene la medida de certeza, 33%, sumando la medida de certeza determinada por los nodos que son responsables de los cálculos de las transacciones que involucran la cuenta C3 y L1. Supongamos que dos transacciones involucran la cuenta C3 y la ubicación L1. La primera transacción se asigna a un primer nodo y la segunda transacción se asigna a un segundo nodo. La medida de certeza basada en la primera transacción se combina con la medida de certeza basada en la segunda transacción para obtener la medida global de certeza. En algunas realizaciones, cada nodo genera su propia matriz resultante. Se combinan múltiples matrices resultantes (generadas por cada nodo) al final de cada paso (por ejemplo, 406/408) para producir un único resultado para cada paso de cálculo de la matriz.

[0095] En 504, los bloques se distribuyen entre los nodos de cálculo para la computación distribuida. Por ejemplo, los nodos del clúster 140 mostrados en la Figura 1 colaboran para calcular valores de una matriz de cálculo informático. Los bloques se pueden distribuir entre nodos de acuerdo con la división determinada en 502 de diversas formas, por ejemplo, por turnos (*round-robin*), basándose en una carga de un nodo, basándose en la localidad de datos y similares.

[0096] En un esquema de distribución por turnos, los bloques se distribuyen a cada nodo en un orden circular. Usando el ejemplo de un clúster de tres nodos, un primer bloque se distribuye a un primer nodo, un segundo bloque se distribuye a un segundo nodo y un tercer bloque se distribuye a un tercer nodo. Se gestiona un cuarto bloque volviendo al primer nodo, por ejemplo, el cuarto bloque se distribuye al primer nodo, un quinto bloque se distribuye al segundo nodo y un sexto bloque se distribuye al tercer nodo.

[0097] En una distribución basada en carga, se utiliza el equilibrio de carga para distribuir bloques a los nodos basándose en la carga del nodo. Utilizando el ejemplo de un clúster de tres nodos para equilibrar las cargas en tres nodos de un clúster, se asigna un bloque al nodo con la carga más ligera de los tres nodos.

[0098] En una distribución basada en la localidad de datos, los bloques almacenados más cerca de un nodo están asignados a ese nodo en lugar de a un nodo más lejos de los bloques. Usando el ejemplo de un clúster de tres nodos, se asigna un bloque al nodo más cercano.

[0099] La división de una matriz de cálculo informático en bloques y la distribución de los bloques entre nodos para el cálculo se explicarán ahora con respecto a las Figuras 7A-7E. El cuadro 702.1 representa una lista de información de transacciones recibidas. En algunas realizaciones, después del ensamblado de la matriz de cálculo informático representada por el cuadro 702.1, las filas C1 y C2 se eliminan mediante preprocesamiento (como se describe en el presente).

[0100] En un ejemplo en el que la división se basa en elementos por ubicación, la división y la distribución son como sigue. Los cálculos se dividen por ubicación de modo que cada una de las ubicaciones L1-L3 se asigna a un nodo respectivo. La matriz de cálculo informático se divide entre los tres nodos, de manera que cada nodo gestiona los cálculos para una columna del cuadro. En este ejemplo, un primer nodo gestiona los cálculos para la columna L1, un segundo nodo gestiona los cálculos para la columna L2 y un tercer nodo gestiona los cálculos para la columna L3. Usando el ejemplo de 703.1 en la Figura 7A, un primer nodo gestiona los cálculos para la ubicación L1, un segundo nodo gestiona los cálculos para la ubicación L2 y un tercer nodo gestiona los cálculos para la ubicación L3. Al final de cada iteración, los resultados se sincronizan para actualizar la matriz de cálculo informático global. Por ejemplo, una iteración produce un resultado generado por el primer nodo para la ubicación L1, un resultado generado por el segundo nodo para la ubicación L2 y un resultado generado por el tercer nodo para la ubicación L3. Estos resultados se sincronizan para actualizar el cuadro 703.1 antes de proceder a una segunda iteración de cálculos.

[0101] Si nos fijamos de nuevo en el cuadro 704.2, las puntuaciones de responsabilidad de compromiso globales se calculan como sigue. El primer nodo calcula las puntuaciones de responsabilidad de compromiso globales para la

ubicación L1, el segundo nodo calcula las puntuaciones de responsabilidad de compromiso globales para la ubicación L2 y el tercer nodo calcula las puntuaciones de responsabilidad de compromiso globales para la ubicación L3. El primer nodo determina que la puntuación de responsabilidad global para la ubicación L1 es del 11% al dividir la responsabilidad combinada para L1 (33%) por el número de cuentas con las que interactuó la ubicación L1 (3). El segundo nodo determina que la puntuación de responsabilidad global para la ubicación L2 es del 21% al dividir la responsabilidad combinada para L2 (83%) por el número de cuentas con las que interactuó la ubicación L2 (4). El tercer nodo determina que la puntuación de responsabilidad global para la ubicación L3 es del 28% al dividir la responsabilidad combinada para L3 (83%) por el número de cuentas con las que interactuó la ubicación L1 (3).

[0102] Si nos fijamos ahora en el cuadro 702.2 en la Figura 7B, las puntuaciones de responsabilidad de elemento se calculan de la siguiente manera. El primer nodo determina que la puntuación de responsabilidad de elemento para la ubicación L1 es del 18% porque la puntuación de responsabilidad de compromiso global para L1 es del 11% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es 60% (11% + 21% + 28%). El segundo nodo determina que la puntuación de responsabilidad de elemento para la cuenta C3/ubicación L2 es del 35% porque la puntuación de responsabilidad de compromiso global para L2 es del 21% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es 60%. El segundo nodo determina que la puntuación de responsabilidad de elemento para la cuenta C4/ubicación L2 es del 43% porque la puntuación de responsabilidad de compromiso global para L2 es del 21% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C4 ha interactuado es del 49% (21% + 28%). El tercer nodo determina que la puntuación de responsabilidad de elemento para la cuenta C3/ubicación L3 es del 47% porque la puntuación de responsabilidad de compromiso global para L3 es del 28% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es 60%. El tercer nodo determina que la puntuación de responsabilidad de elemento para la cuenta C4/ubicación L3 es del 57% porque la puntuación de responsabilidad de compromiso global para L3 es del 28% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C4 ha interactuado es 49%.

[0103] En un ejemplo en el que la división se basa en las transacciones, la división y distribución son como se indica a continuación. Los cálculos se dividen por transacción entre dos nodos para el cuadro 702.1 en la Figura 7A. En algunas realizaciones, después del ensamblado de la matriz de cálculo informático representada por el cuadro 702.1, las filas C1 y C2 se eliminan mediante preprocesamiento (como se describe en el presente), lo que tiene como resultado el cuadro 703.1, mientras que se retiene un recuento del número de cuentas no comprometidas para los cálculos, como se describe en mayor detalle en el presente. La eliminación de filas del cuadro reduce el número de cálculos que se realizan porque se rellenan menos entradas del cuadro en comparación con el cuadro 702.1. La matriz de cálculo informático se divide entre los dos nodos, de manera que cada nodo gestiona los cálculos para una fila del cuadro 703.1. En este ejemplo, un primer nodo gestiona los cálculos para la fila C3 y un segundo nodo gestiona los cálculos para la fila C4.

[0104] Si nos fijamos ahora en el cuadro 704.2, se calculan las puntuaciones de responsabilidad de compromiso globales basándose en las puntuaciones de responsabilidad de elemento mostradas en el cuadro 703.1 como se muestra a continuación. El primer nodo calcula las puntuaciones de responsabilidad de compromiso globales para la cuenta C3 y el segundo nodo calcula las puntuaciones de responsabilidad de compromiso globales para la cuenta C4. En algunas realizaciones, el primer nodo genera una primera matriz y el segundo nodo genera una segunda matriz y las matrices se combinan. En otras palabras, se combinan los resultados de los cálculos respectivos realizados por cada nodo para obtener puntuaciones de responsabilidad de compromiso globales como se indica a continuación.

[0105] Para la ubicación L1, el primer nodo determina que la puntuación de responsabilidad global para esta ubicación es del 11% al dividir la responsabilidad combinada de L1 (33%) por el número de cuentas con las que interactuó la ubicación L1 (3). Para la ubicación L2, el primer nodo determina que la puntuación de responsabilidad de compromiso global para esta ubicación es del 8,25% al dividir la responsabilidad combinada (33%) por el número de cuentas con las que interactuó la ubicación L2 (4). Para la ubicación L3, el primer nodo determina que la puntuación de responsabilidad de compromiso global para esta ubicación es del 11% al dividir la responsabilidad combinada (33%) por el número de cuentas con las que interactuó la ubicación L3 (3).

[0106] Para la ubicación L1, el segundo nodo no hace un cálculo porque sus transacciones (cuenta C4) no implican la interacción entre la ubicación L1 y la cuenta C4. Para la ubicación L2, el segundo nodo determina que la puntuación de responsabilidad global para esta ubicación es del 12,5% al dividir la responsabilidad combinada (50%) por el número de cuentas con las que interactuó la ubicación L2 (4). Para la ubicación L3, el segundo nodo determina que la puntuación de responsabilidad global para esta ubicación es del 16,67% al dividir la responsabilidad combinada (50%) por el número de cuentas con las que interactuó la ubicación L3 (3).

[0107] Después de que los nodos individuales completen los cálculos asignados al nodo, los resultados se combinan para generar un cuadro final. El cuadro final es un resultado para determinar la convergencia o un punto de partida para cálculos posteriores. En este ejemplo, los cálculos del primer nodo y el segundo nodo se combinan

al sumar los resultados generados por los nodos. Para la ubicación L1, la puntuación de responsabilidad de compromiso global es del 11%, como se muestra en el cuadro 704.2 de la Figura 7B. Para la ubicación L2, la puntuación de responsabilidad de compromiso global es del 21% (suma del 8,25% calculado por el primer nodo y del 12,5% calculado por el segundo nodo). Para la ubicación L3, la puntuación de responsabilidad de compromiso global es del 28% (suma del 11% calculado por el primer nodo y del 16,67% calculado por el segundo nodo).

[0108] Si nos fijamos ahora en el cuadro 702.2 en la Figura 7B, las puntuaciones de responsabilidad de elemento se calculan de la siguiente manera. El primer nodo determina que la puntuación de responsabilidad de elemento para la cuenta C3/ubicación L1 es del 18% porque la puntuación de responsabilidad de compromiso global para L1 es del 11% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es 60% (11% + 21% + 28%). El primer nodo determina que la puntuación de responsabilidad de elemento para la cuenta C3/ubicación L2 es del 35% porque la puntuación de responsabilidad de compromiso global para L2 es del 21% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es 60%. El primer nodo determina que la puntuación de responsabilidad de elemento para la cuenta C3/ubicación L3 es del 47% porque la puntuación de responsabilidad de compromiso global para L3 es del 28% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C3 ha interactuado es 60%.

[0109] El segundo nodo determina que la puntuación de responsabilidad de elemento para la cuenta C4/ubicación L2 es del 43% porque la puntuación de responsabilidad de compromiso global para L2 es del 21% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C4 ha interactuado es 49% (21% + 28%). El segundo nodo determina que la puntuación de responsabilidad de elemento para la cuenta C4/ubicación L3 es del 57% porque la puntuación de responsabilidad de compromiso global para L3 es del 28% y la suma de las puntuaciones de responsabilidad de compromiso globales para las ubicaciones con las que C4 ha interactuado es 49%.

[0110] Los resultados del primer nodo y el segundo nodo se combinan (por ejemplo, el primer nodo contribuye la tercera fila y el segundo nodo contribuye la cuarta fila) para obtener las puntuaciones de responsabilidad de elemento mostradas en el cuadro 702.2 de la Figura 7B.

[0111] Las técnicas de la presente divulgación encuentran aplicación en la prevención del uso de información sustraída indebidamente, incluidos viajes fraudulentos (que hacen uso de información), robo de números de cuentas de tarjetas de crédito, suplantación de identidad en cajeros automáticos, captura de PIN, etc. Los estafadores pueden robar y usar información. Por ejemplo, un estafador obtiene ilegalmente una cuenta, como por ejemplo una tarjeta de crédito y una tarjeta de débito. A veces, el estafador utiliza la tarjeta de crédito o de débito para comprar tarjetas de regalo en un intento de ocultar la fuente de la tarjeta de crédito/débito. Posteriormente, el estafador emprende un viaje fraudulento para gastar el dinero obtenido ilegalmente. Por ejemplo, el estafador utiliza la tarjeta de crédito o débito robadas o la tarjeta de regalo en una o más tiendas para comprar artículos. En otro ejemplo, en un viaje, el estafador puede visitar uno o más cajeros automáticos para retirar dinero. Los sistemas informáticos convencionales no pueden detectar eficazmente este tipo de fraude. Para simplificar, en el ejemplo de la Figura 6 se muestra una situación con relativamente pocas tarjetas y depósitos y se asumen solo dos PDC: las cuentas C3 y C4. En un gran número de redes, existen millones de nodos, lo que hace que su cálculo sea costoso informáticamente.

[0112] Aunque las realizaciones anteriores han sido descritas con cierto grado de detalle para facilitar su comprensión, la invención no se limita a los detalles proporcionados. Existen muchas formas alternativas de implementar la invención. Las realizaciones descritas son ilustrativas y no limitativas.

REIVINDICACIONES

1. Un método que comprende:
- 5 para cada cuenta de una pluralidad de cuentas, determinar (408) una puntuación de responsabilidad de componente para cada punto de compromiso (PDC) potencial de una pluralidad de puntos de compromiso potenciales, en donde la determinación de la puntuación de responsabilidad de componente incluye:
- representar las determinaciones como al menos una operación de matriz de cálculo informático, donde las filas de la matriz de cálculo informático corresponden a cuentas y las columnas corresponden a ubicaciones,
- 10 dividir (502) una matriz de la operación de matriz de cálculo informático en bloques de cálculo, y distribuir (504) los bloques entre una pluralidad de nodos informáticos de cálculo para un cálculo informático distribuido;
- determinar (406) una puntuación de responsabilidad de compromiso global para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales basándose en la puntuación de responsabilidad de componente determinada, incluyendo mediante la división de una suma de las puntuaciones de responsabilidad de componente determinadas por un número de cuentas con las que el PDC potencial interactuó, ponderada por al menos un cálculo anterior de la puntuación de responsabilidad de compromiso global;
- 15 actualizar alternativamente un valor de las puntuaciones de responsabilidad de componente determinadas para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales y un valor de la puntuación de responsabilidad de compromiso global determinada para cada punto de compromiso potencial a través de iteraciones sucesivas hasta que las puntuaciones de responsabilidad de compromiso globales determinadas han convergido, en donde la convergencia incluye determinar si los valores se han estabilizado; y
- 20 en caso de que las puntuaciones de responsabilidad de compromiso globales determinadas hayan convergido, seleccionar al menos uno de la pluralidad de puntos de compromiso potenciales como un punto de compromiso probable basándose al menos en la puntuación de responsabilidad de compromiso global correspondiente del punto de compromiso probable seleccionado.
- 30 2. El método de la reivindicación 1, que comprende además determinar la pluralidad de cuentas basándose en los datos de transacción recibidos, incluyendo los datos de transacción los valores de los parámetros de transacción.
- 35 3. El método de la reivindicación 1, en donde la determinación (408) de la puntuación de responsabilidad de componente incluye la comparación de un criterio de convergencia con:
- i) una primera puntuación de responsabilidad de componente generada por una primera combinación de valores de parámetros de transacción identificados; y
- ii) una segunda puntuación de responsabilidad de componente generada por una segunda combinación de valores de parámetros de transacción identificados,
- 40 en donde los criterios de convergencia comprenden un umbral que, cuando la diferencia entre una puntuación de responsabilidad de compromiso global actual y una puntuación de responsabilidad de compromiso global para una iteración anterior es inferior a dicho umbral, se determina que las puntuaciones de responsabilidad de compromiso globales han convergido.
- 45 4. El método de la reivindicación 1, en donde la determinación (410) de si las puntuaciones de responsabilidad de compromiso globales determinadas han convergido incluye la determinación de que al menos una de las puntuaciones de responsabilidad de compromiso globales cumple con un criterio de convergencia, en donde la determinación de si se cumple con dichos criterios de convergencia comprende al menos una de las siguientes etapas:
- 50 determinar si una diferencia entre las puntuaciones de responsabilidad globales sucesivas es inferior a un umbral de convergencia;
- determinar si ha transcurrido un periodo de tiempo umbral desde que comenzaron los cálculos y, si así se determina, determinar que se ha producido la convergencia; y
- 55 determinar si se ha producido un número umbral de iteraciones y, si así se determina, determinar que se ha producido la convergencia.
5. El método de la reivindicación 1, que comprende además identificar dinámicamente la pluralidad de cuentas y la pluralidad de puntos de compromiso potenciales, en donde la identificación de la pluralidad de cuentas y la pluralidad de puntos de compromiso potenciales se basa en al menos una interacción entre al menos una de la pluralidad de cuentas y al menos uno de la pluralidad de puntos de compromiso potenciales dentro de un periodo de tiempo predefinido.
- 60 6. El método de la reivindicación 1, que comprende además recibir una lista de cuentas que se sabe que están comprometidas.
- 65

7. El método de la reivindicación 1, que comprende además proporcionar un identificador del punto de compromiso probable seleccionado; y opcionalmente, en donde la selección del punto de compromiso probable incluye determinar un patrón en ubicaciones geográficas de la pluralidad de puntos de compromiso potenciales, y el identificador identifica una agrupación de la pluralidad de puntos de compromiso potenciales seleccionados basándose al menos en parte en el patrón.
8. El método de la reivindicación 1, que comprende además determinar la pluralidad de puntos de compromiso potenciales basándose en que cada punto de compromiso potencial haya interactuado al menos una vez con una cuenta comprometida.
9. El método de la reivindicación 1, en donde la determinación (408) de la puntuación de responsabilidad de componente incluye el uso de al menos una probabilidad calculada previamente, con un factor de ponderación aplicado, y una certeza relativa entre una puntuación de responsabilidad de compromiso global actual y la probabilidad o las probabilidades calculadas previamente para determinar la puntuación de responsabilidad de compromiso global; u opcionalmente, en donde la actualización incluye dividir una responsabilidad proporcionalmente basándose en una medida de certeza.
10. El método de la reivindicación 1, en donde después de que se actualicen la puntuación de responsabilidad de compromiso global para cada punto de compromiso potencial y las puntuaciones de responsabilidad de componente para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales en caso de que las puntuaciones de responsabilidad de compromiso globales determinadas no hayan convergido, se realiza una determinación de si las puntuaciones de responsabilidad de compromiso globales determinadas actualizadas han convergido.
11. El método de la reivindicación 1, que comprende además el preprocesamiento de la operación de matriz de cálculo informático para eliminar aquellas cuentas que se determina que no están comprometidas, de modo que no se realizan cálculos para las cuentas eliminadas.
12. El método de la reivindicación 1, en donde cada uno de los bloques de cálculo está asociado con un subconjunto diferente de datos de entrada que se utilizarán para rellenar la operación de la matriz de cálculo informático; u opcionalmente, en donde cada uno de los bloques de cálculo incluye un subconjunto diferente de elementos de la operación de la matriz de cálculo informático.
13. El método de la reivindicación 1, que comprende además inicializar las puntuaciones de responsabilidad de componente con medidas de certeza que son ponderadas basándose en una serie de transacciones que se han procesado en el PDC respectivo.
14. Un sistema para identificar un punto de compromiso, el cual comprende:
un procesador (202) configurado para:
para cada cuenta de una pluralidad de cuentas, determinar (408) una puntuación de responsabilidad de componente para cada punto de compromiso (PDC) potencial de una pluralidad de puntos de compromiso potenciales, en donde la determinación de la puntuación de responsabilidad de elemento incluye:
representar las determinaciones como al menos una operación de matriz de cálculo informático, en donde las filas de la matriz de cálculo informático corresponden a cuentas y las columnas corresponden a ubicaciones,
dividir (502) una matriz de la operación de matriz de cálculo informático en bloques de cálculo, y
distribuir (504) los bloques entre una pluralidad de nodos informáticos de cálculo para un cálculo informático distribuido;
determinar (406) una puntuación de responsabilidad de compromiso global para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales basándose en la puntuación de responsabilidad de componente determinada, incluyendo mediante la división de una suma de las puntuaciones de responsabilidad de componente determinadas por un número de cuentas con las que el PDC potencial interactuó, ponderada por al menos un cálculo anterior de la puntuación de responsabilidad de compromiso global;
actualizar alternativamente un valor de las puntuaciones de responsabilidad de componente determinadas para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales y un valor de la puntuación de responsabilidad de compromiso global determinada

- para cada punto de compromiso potencial a través de iteraciones sucesivas hasta que las puntuaciones de responsabilidad de compromiso globales determinadas han convergido, en donde la convergencia incluye determinar si los valores se han estabilizado; y
- 5 en caso de que las puntuaciones de responsabilidad de compromiso globales determinadas hayan convergido, seleccionar al menos uno de la pluralidad de puntos de compromiso potenciales como un punto de compromiso probable basándose al menos en la puntuación de responsabilidad de compromiso global correspondiente del punto de compromiso probable seleccionado; y una memoria (210) acoplada al procesador y configurada para proporcionar instrucciones al procesador.
- 10 15. Un producto de programa informático para identificar un punto de compromiso, estando incorporado el producto de programa informático en un medio de almacenamiento legible por ordenador no transitorio y que comprende instrucciones informáticas para:
- 15 para cada cuenta de una pluralidad de cuentas, determinar (408) una puntuación de responsabilidad de componente para cada punto de compromiso (PDC) potencial de una pluralidad de puntos de compromiso potenciales, en donde la determinación de la puntuación de responsabilidad de componente incluye:
- 20 representar las determinaciones como al menos una operación de matriz de cálculo informático, en donde las filas de la matriz de cálculo informático corresponden a cuentas y las columnas corresponden a ubicaciones,
- 20 dividir (502) una matriz de la operación de matriz de cálculo informático en bloques de cálculo, y distribuir (504) los bloques entre una pluralidad de nodos informáticos de cálculo para un cálculo informático distribuido;
- 25 determinar (406) una puntuación de responsabilidad de compromiso global para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales basándose en la puntuación de responsabilidad de componente determinada, incluyendo mediante la división de una suma de las puntuaciones de responsabilidad de componente determinadas por un número de cuentas con las que el PDC potencial interactuó, ponderada por al menos un cálculo anterior de la puntuación de responsabilidad de compromiso global;
- 30 actualizar alternativamente un valor de las puntuaciones de responsabilidad de componente determinadas para cada punto de compromiso potencial de la pluralidad de puntos de compromiso potenciales y un valor de la puntuación de responsabilidad de compromiso global determinada para cada punto de compromiso potencial a través de iteraciones sucesivas hasta que las puntuaciones de responsabilidad de compromiso globales determinadas han convergido, en donde la convergencia incluye determinar si los valores se han estabilizado; y
- 35 en caso de que las puntuaciones de responsabilidad de compromiso globales determinadas hayan convergido, seleccionar al menos uno de la pluralidad de puntos de compromiso potenciales como un punto de compromiso probable basándose al menos en la puntuación de responsabilidad de compromiso global correspondiente del punto de compromiso probable seleccionado.

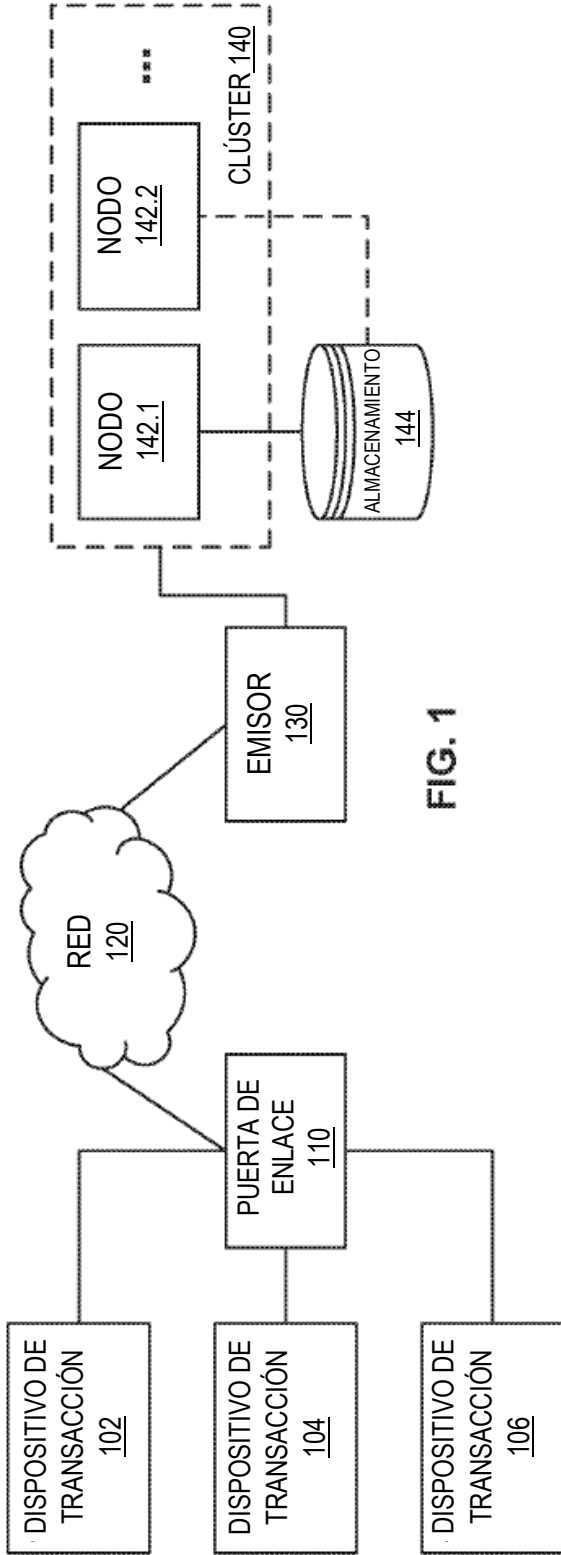


FIG. 1

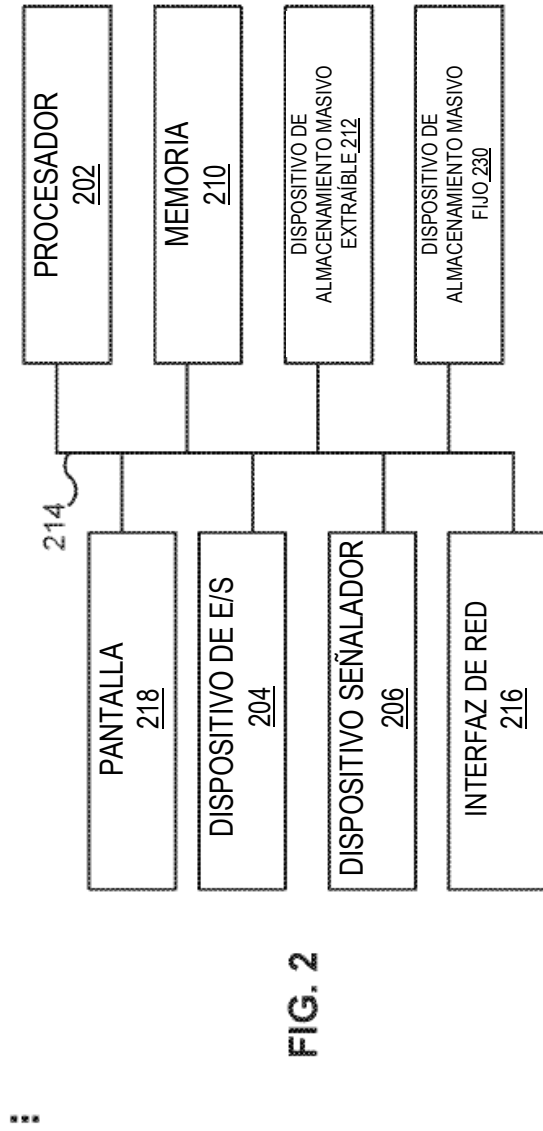


FIG. 2

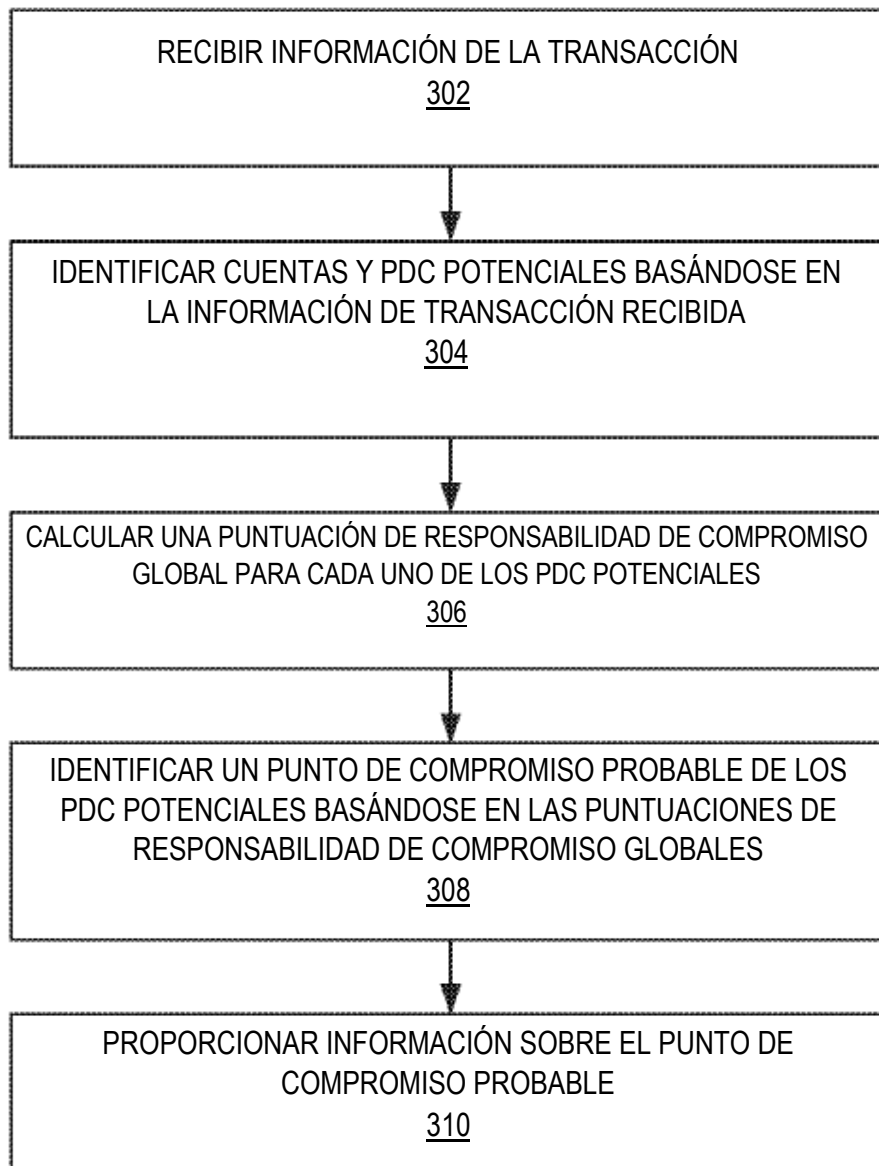


FIG. 3

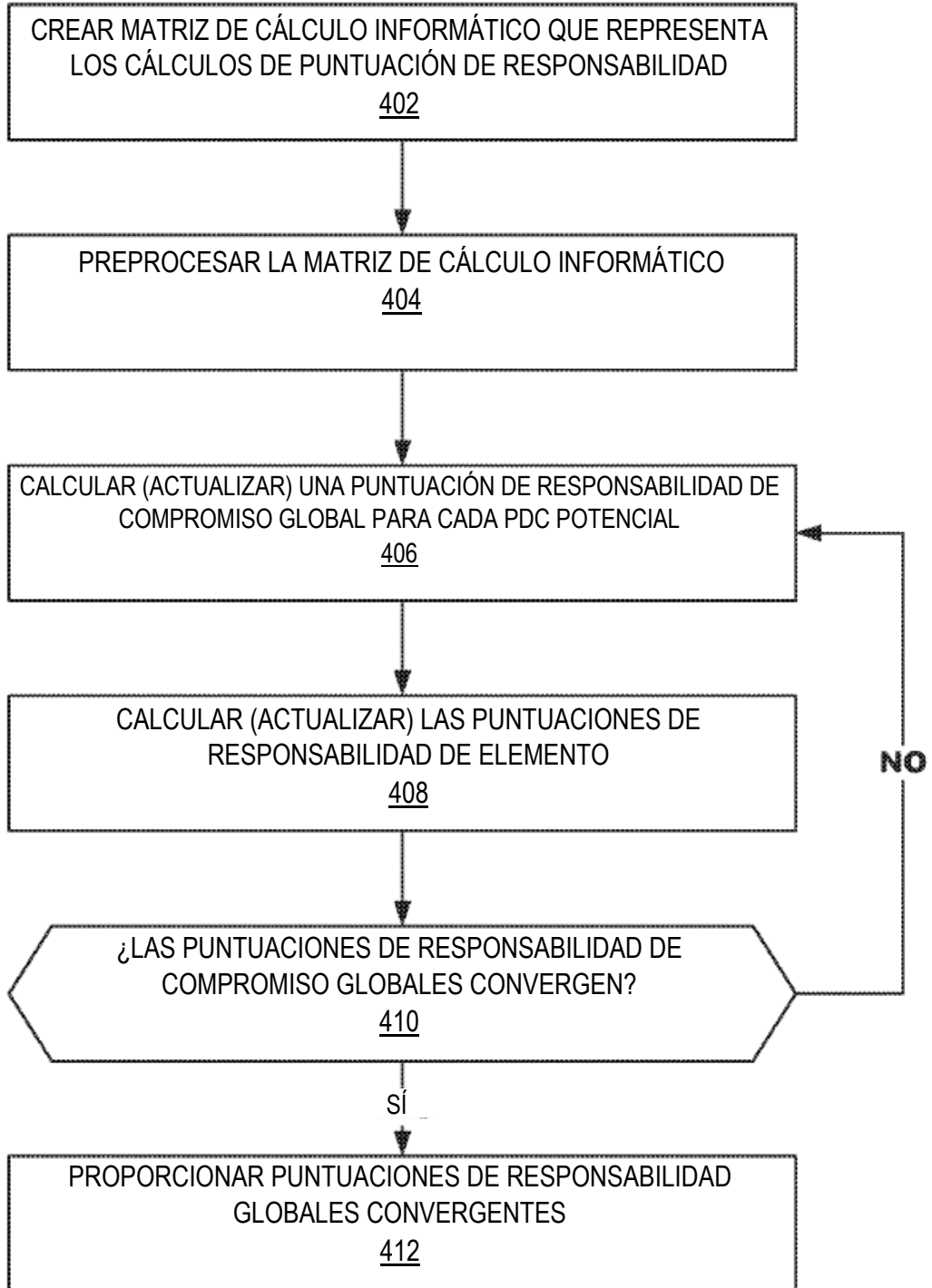


FIG. 4

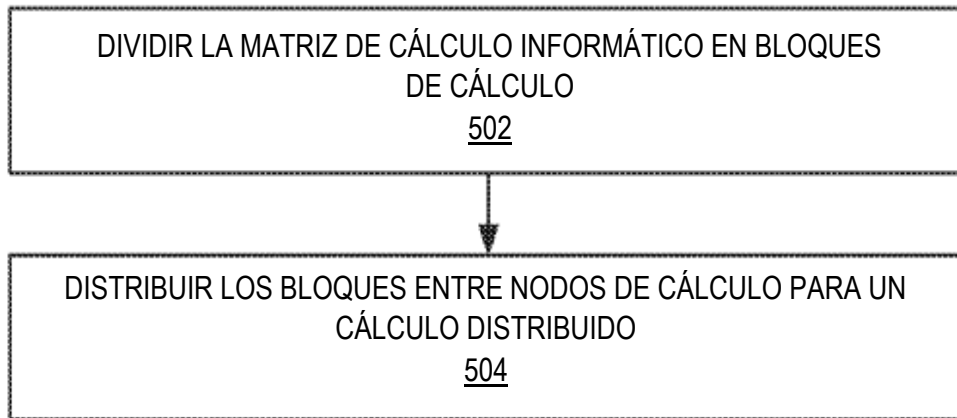


FIG. 5

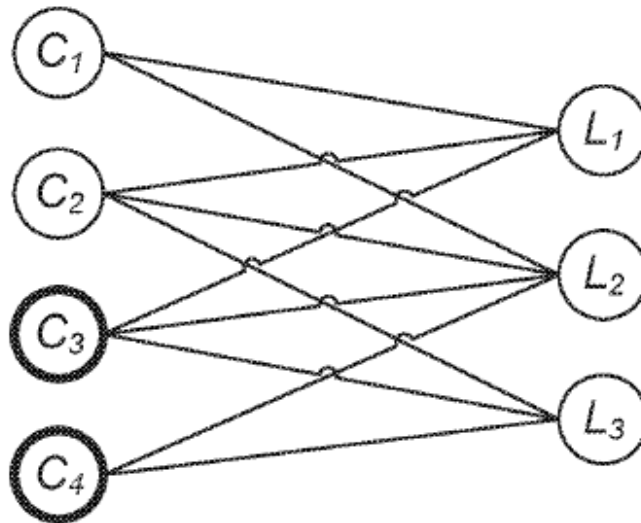


FIG. 6

FIG. 7A

	L1	L2	L3
C1	0%	0%	--
C2	0%	0%	0%
C3	33%	33%	33%
C4	--	50%	50%

702.1

	L1	L2	L3
C3	33%	33%	33%
C4	--	50%	50%

703.1

FIG. 7B

UBICACIÓN	L1	L2	L3
PROBABILIDAD	11%	21%	28%

704.2

	L1	L2	L3
C1	0%	0%	--
C2	0%	0%	0%
C3	18%	35%	47%
C4	--	43%	57%

702.2

FIG. 7C

UBICACIÓN	L1	L2	L3
PROBABILIDAD	6%	20%	35%

704.3

FIG. 7D

UBICACIÓN	L1	L2	L3
PROBABILIDAD	3%	17%	40%

704.4

	L1	L2	L3
C1	0%	0%	--
C2	0%	0%	0%
C3	10%	33%	57%
C4	--	36%	64%

702.3

	L1	L2	L3
C1	0%	0%	--
C2	0%	0%	0%
C3	5%	28%	67%
C4	--	30%	70%

702.4

FIG. 7E

UBICACIÓN	L1	L2	L3
PROBABILIDAD	2%	15%	46%

704.5

⋮

CUENTA	UBICACIÓN 1	UBICACIÓN 2	UBICACIÓN 3
C1	0%	0%	--
C2	0%	0%	0%
C3	0%	0%	100%
C4	--	0%	100%

702.5

UBICACIÓN	L1	L2	L3
PROBABILIDAD	0%	0%	66%

704.6