(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 565 842**

(51) Int. CI.:

H04L 29/06 (2006.01)
G06F 21/62 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **27.12.2011 E 11815759 (3)**

(97) Fecha y número de publicación de la concesión europea: **23.12.2015 EP 2798809**

(54) Título: **Método de asignación de seudónimos dinámicos para redes de creación de perfiles de datos de usuarios, y red de creación de perfiles de datos de usuarios que implementa el método**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
07.04.2016

(73) Titular/es:

TELECOM ITALIA S.P.A. (100.0%)
Via Gaetano Negri, 1
20123 Milano, IT

(72) Inventor/es:

GOLIC, JOVAN

(74) Agente/Representante:

PONTI SALES, Adelaida

DESCRIPCIÓN

Método de asignación de seudónimos dinámicos para redes de creación de perfiles de datos de usuarios, y red de creación de perfiles de datos de usuarios que implementa el método

Sector de la técnica

La presente invención se refiere a la privacidad y al anonimato en redes informáticas y, particularmente, a los métodos de asignación de seudónimos capaces de brindar anonimia de datos sensibles, tales como perfiles de datos de usuarios, que se almacenan en redes informáticas.

Estado de la técnica anterior

Las técnicas de asignación de seudónimos se pueden utilizar para brindar privacidad de datos sensibles en redes de creación de perfiles de datos, en las cuales los datos se adquieren dinámicamente desde varias fuentes de datos y luego se procesan, almacenan y recuperan, en un lapso de tiempo. Típicamente, una red de creación de perfiles de datos se implementa como una red informática. Cada fuente de datos provee datos que provienen o se relacionan con diferentes entidades del mundo real denominadas “usuarios”. Por ejemplo, los usuarios pueden ser individuos (personas) o grupos de personas, empresas, organizaciones, sitios web de Internet, o dispositivos tales como computadoras personales o teléfonos móviles. La privacidad implica que las identidades del mundo real de los usuarios permanecen ocultas de los nodos de red que procesan y almacenan los datos sensibles. En el contexto de la presente descripción, y a los efectos de la presente invención, una “identidad del mundo real” de un usuario se define como un conjunto de identificadores, en donde cada identificador es una descripción de una propiedad lógica o física verificable de un usuario, que se asume como válida en un lapso de tiempo. Una identidad del mundo real es una representación única de un solo usuario o un conjunto relativamente pequeño de usuarios dentro de un dominio global (por ejemplo, el mundo o un estado) o un dominio local (por ejemplo, una empresa o ciudad).

Por ejemplo, una red de creación de perfiles de datos puede ser instalada para la creación de perfiles en línea por los proveedores de servicio de Internet (ISP, *Internet Service Provider*) o sitios web particulares que proveen varios servicios de Internet. El perfil de datos se refiere, entonces, al uso de Internet por parte de los usuarios individuales y está destinado a ser utilizado para proveer servicios nuevos o mejorados en Internet, por ejemplo, para la comercialización focalizada por parte de entidades autorizadas mediante anuncios publicitarios personalizados. En este caso, la identidad del mundo real de un usuario a ser protegida puede incluir como identificadores, por ejemplo, un localizador de recursos uniforme (URL, *Uniform Resource Locator*), una dirección de correo electrónico, una dirección IP, un número de teléfono, el nombre de una persona, o un domicilio.

A fin de derivar perfiles de datos acumulados en el tiempo para cualquier usuario particular, es intrínsicamente necesario enlazar en forma conjunta los diferentes datos relacionados con el mismo usuario en distintos momentos. En esencia, esta capacidad de formar enlaces es restringida por el hecho de que se relaciona solamente con los datos necesarios para computarizar los perfiles de datos en un determinado momento y no necesariamente los perfiles de datos en momentos diferentes. A fin de asegurar esta capacidad de formar enlaces restringida, un método convencional consiste en usar un seudónimo estático (en lo sucesivo, el seudónimo también se denominará “PID”) en lugar de una identidad (en lo sucesivo, también denominada “ID”), donde las asociaciones entre las ID y los PID debería permanecer ocultas de los nodos de red que procesan y almacenan los datos sensibles.

El problema principal con el uso de seudónimos estáticos es que la capacidad de formar enlaces provista no es restringida, ya que no está limitada en tiempo y también se relaciona con los perfiles de datos de un usuario en diferentes momentos. La capacidad de formar enlaces no restringida significa que los perfiles de datos en diferentes momentos se enlazan conjuntamente mediante el mismo seudónimo estático y pueden usarse, por ende, para obtener las curvas de los perfiles de datos en el tiempo para cualquier usuario destinado, independientemente de cómo cambió el perfil de datos en el tiempo. En consecuencia, la capacidad de formar enlaces no restringida que deriva del uso de seudónimos estáticos provoca la pérdida de la privacidad de avance/retroceso (*forward/backward*) e incrementa el riesgo de la recuperación de la identidad del usuario al analizar las curvas de los perfiles de datos. Una privacidad de avance/retroceso escasa o nula significa que, si la identidad de un usuario está comprometida en determinado momento, entonces todos los perfiles de datos correspondientes de avance y retroceso están comprometidos, lo que en sí mismo genera la total capacidad de seguimiento del usuario identificado.

Como una curva de perfiles de datos contiene mucha más información que un único perfil de datos en un determinado momento, el riesgo de que exista alguien capaz de recuperar la identidad del correspondiente usuario aumenta de manera significativa —según el perfil de datos— especialmente si es posible correlacionar la curva de perfiles de datos con los datos de la vida real. En general, este riesgo es, en consecuencia, mucho más alto que en el caso de utilizar perfiles de datos solamente en momentos únicos.

El documento de patente de los Estados unidos con el número US 7213032 B2 describe un método y un sistema

implementados por computadora para la creación de perfiles anónimos y la comercialización dirigida a usuarios anónimos en una red de datos, tal como Internet. La red de datos se divide en tres partes: la parte anónima fiable (ATP, *Anonymous Trusted Part*), la parte no anónima (NAP, *Non-Anonymous Part*), y la parte que no crea perfiles (NPP, *Non-Profiling Part*). Los perfiles de usuarios anónimos se computarizan, se mantienen y se usan en la ATP, donde las transacciones no anónimas requieren que las identidades de usuarios del mundo real se ejecuten dentro de la NAP, y los perfiles de usuarios anónimos tomados de la ATP también se usan dentro de la NPP. El anonimato de los perfiles de usuarios es asegurado al asignar un identificador único (UID, *Unique Identifier*) a cada usuario en la ATP y un UID posiblemente diferente en la NPP. Los perfiles de usuarios rotulados por el UID se almacenan en una base de datos de la ATP. Los usuarios se autentican en forma anónima en la ATP o la NPP mediante la utilización de nombres o seudónimos de usuarios virtuales elegidos por ellos mismos, junto con contraseñas, cuando acceden al sistema. El punto central del documento de patente de los Estados Unidos con el número US 7213032 B2 es que la identidad del usuario del mundo real solamente se utiliza en la NAP y nunca se describe a ninguna parte de la ATP o NPP, mientras que los perfiles de usuarios nunca se usan explícitamente en la NAP. No obstante, se permite que los denominados “valores de transacción representativos o simbolizados” puedan atravesar la barrera entre la NAP por un lado, y la NPP por otro. Dichos valores se definen como “cualquier información codificada que puede ser generada o emitida por un usuario y no contiene ni el perfil del usuario ni la identidad del usuario del mundo real”. Dichos valores desempeñan un rol importante al conectar las partes anónimas y no anónimas de la red y, por ende, permiten las transacciones no anónimas dentro de la NAP.

El documento de patente de los Estados Unidos con el número US 7844717 B2 describe un método para el intercambio de seudónimos de datos personales privados asociados con usuarios entre dos o más servidores de almacenamiento de datos o dentro de un único servidor de almacenamiento de datos, donde la privacidad de los usuarios y los servidores de almacenamiento de datos es protegida mediante la utilización de seudónimos en lugar de las identidades del mundo real. En el sistema, los usuarios y servidores son autenticados mediante métodos estándares que utilizan seudónimos y credenciales seguros validados (en particular, el método descripto por D. Chaum y J.-H. Evertse, *A secure and privacy-protecting protocol for transmitting personal information between organizations*, en *Proceedings of Crypto ’86, Lecture Notes in Computer Science*, vol. 263, páginas 118-167, 1987).

El punto central del método consiste en el uso de un servidor proxy fiable denominado el “servidor de seudónimos” para controlar el acceso a los datos privados a través de normas de control de acceso, en donde los usuarios y servidores están registrados y representados por los identificadores únicos asociados (UID) junto con el usuario y los tipos de servidores, respectivamente. También pueden almacenarse las identidades de usuarios del mundo real.

El documento de patente de los Estados Unidos con el número US 7610390 B2 describe un método para enlazar cuentas de usuarios almacenadas en diferentes nodos en una red de datos tales como Internet, donde cada cuenta de usuario contiene cierta información de identidad (ID) de la cuenta de usuario única a nivel local, compuesta por identificadores del mundo real, posiblemente parciales, elegidos a nivel local (que deberían considerarse como privados si especifican el usuario en forma exclusiva) o nombres de cuentas de usuarios locales elegidos en forma arbitraria, información auxiliar compuesta por los así denominados *handles* (referencia abstracta conocida como un tipo particular de punteros inteligentes) y, posiblemente, otros datos privados (por ejemplo, perfiles de usuarios, preferencias, políticas, servicios que cuentan con autorización para su acceso, derechos de control de acceso, etc.). Existen dos tipos básicos de nodos, denominados “proveedores de identidad” y “proveedores de servicio”. El papel principal de los primeros es autenticar los usuarios; por ende, las ID locales almacenadas necesariamente incluyen identificadores del mundo real. El papel principal de los últimos es proveer varios servicios y, por ende, pueden o no incluir identificadores del mundo real como partes de las ID locales almacenadas.

Los nodos de servicio y de identidad interactúan entre sí y, por lo tanto, proveen diferentes servicios a los usuarios de las redes. Esta interacción requiere que las cuentas del usuario almacenadas en diferentes nodos se enlacen juntas. El papel de los handles es permitir este enlace sin intercambiar las ID de las cuentas de usuario locales. Esto se alcanza cuando los dos nodos que se comunican entre sí comparten el mismo handle (como secreto en común). El mismo handle compartido determina entonces que las dos cuentas de usuario corresponden al mismo usuario. Cada handle correspondiente a un usuario consta de dos partes, que son respectivamente generadas por los dos nodos y enviadas una a la otra, en forma —posiblemente— encriptada. Si el mismo nodo se comunica con varios nodos adicionales, entonces la parte del handle generada por ese nodo es la misma para todas las conexiones; es decir, depende de la cuenta de usuario local en lugar de la conexión. En este sentido, se puede llamar “seudónimo de la cuenta de usuario local en un nodo determinado”. Un par de seudónimos asociados con dos nodos determina entonces, como un handle, la conexión entre las cuentas del mismo usuario, en los dos nodos. También se sugiere que, al elegir seudónimos dinámicos —es decir, seudónimos que cambian en el tiempo— “la visibilidad del nombre de la cuenta se puede reducir”.

La publicación de S. Fouladgar y H. Afifi, *A simple privacy protecting scheme enabling delegation and ownership transfer for RFID tags, Journal of Communications*, vol. 2, no. 6, páginas 6 a 13, 2007, se basa en un protocolo de comunicación para autenticación mutua en un sistema compuesto por etiquetas de identificación de frecuencia de radio (RFID, *Radio Frequency IDentification*) y lectoras de etiquetas a través de base de datos en línea fiables. El protocolo es de tipo desafío-respuesta que usa seudónimos dinámicos para la autenticación de etiquetas, donde los seudónimos son generados a partir de claves secretas pre-compartidas y nonces (números usados una sola vez) locales contra-generados mediante la utilización de funciones de encriptación o verificación de errores

criptográficos. Las ID de etiquetas y las claves secretas se almacenan en la base de datos en línea fiables, y solamente son reveladas por el protocolo a los lectores autorizados, mientras que los seudónimos dinámicos aseguran que la autenticación de etiquetas siga siendo imposible de rastrear por los lectores no autorizados.

Explicación de la invención

En relación con el documento de patente de los Estados Unidos con el número US 7213032 B2, el solicitante observa que, a pesar de la importancia para el método descripto, la definición y el rol de los denominados “valores de transacción representativos o simbolizados” permanecen un tanto vagos y poco claros. El solicitante también observa que los UID del usuario cumplen el rol de seudónimos estáticos únicos asociados con los usuarios y los perfiles de usuarios, y que la separación lógica o física requerida de la NAP debería asegurar que los UID de los usuarios nunca se almacenen en el equipo del usuario y nunca puedan enlazarse con las identidades de usuarios del mundo real a través de cualquier información asociada con los usuarios (por ejemplo, a través de direcciones IP del equipo del usuario en Internet). Esto puede ser difícil de implementar dado el amplio uso de los UID de los usuarios en la ATP y la NPP. En cualquier caso, tal como se explicó con anterioridad, el uso de seudónimos estáticos ocasiona la no deseada capacidad de formar enlaces no restringida en el tiempo, de los perfiles de usuarios.

En relación con el documento de patente de los Estados Unidos con el número 7844717 B2, el solicitante observa que los UID cumplen el rol de seudónimos estáticos únicos asociados con las entidades implicadas y, por ende, provocan la capacidad de formar enlaces no restringida en el tiempo de los intercambios de datos privados. Asimismo, si los seudónimos estáticos se usan para representar los datos privados almacenados en servidores de datos individuales, entonces la capacidad de formar enlaces no restringida en el tiempo resultante de estos datos privados puede ser indeseable en muchas aplicaciones.

En relación con el documento de patente de los Estados Unidos con el número US 7610390 B2, el solicitante observa que las reivindicaciones principales de dicha patente (reivindicaciones independientes 1, 5 y 12) presentan defectos, ya que no especifican cómo el nodo -que es el primero en recibir la parte del handle enviada desde el otro nodo- determina la cuenta de usuario a la cual esta parte del handle debería estar asociada. A saber, las ID de las cuentas de los usuarios no se transmiten y, sin cierta información en común que especifica la cuenta del usuario, las partes del handle no pueden compartirse; es decir, el enlace inicial no se puede establecer. Por la misma razón, las partes dinámicamente generadas del handle ya existente no pueden compartirse de la manera especificada por las reivindicaciones. Al examinar el texto del documento de patente de los Estados Unidos con el número US 7610390 B2, el solicitante encontró que los autores pueden haber asumido que el enlace requerido (tanto inicialmente como después, cuando se cambian las partes localmente generadas del handle) se puede establecer mediante la utilización de, para cada par de nodos, la presencia simultánea del usuario en ambos nodos a través de una dirección IP común, o bien el almacenamiento de las partes del handle en el mismo equipo del usuario empleado para acceder a ambos nodos (por ejemplo, en forma de cookie mediante un navegador de Internet del usuario). El solicitante observa que, en este último caso, es necesario realizar una autenticación de usuario no solamente en uno de los nodos sino en ambos nodos, tal como se especifica en las reivindicaciones 5 y 12. Esto se debe a que las cookies no pueden considerarse auténticas (incluso cuando están encriptadas, debido a los ataques de *replay*).

Otra observación del solicitante es que, a pesar de las partes dinámicamente generadas del handle propuestas, cada cuenta de usuario, con todos los datos privados almacenables, se puede enlazar en el tiempo en cada nodo, donde la capacidad de formar enlaces es determinada por la misma ID de la cuenta de usuario almacenado así como también por la parte o las partes inalteradas almacenadas del handle. Otro defecto más del método, tal como lo observó el solicitante, es que no provee privacidad de los datos privados almacenados a menos que exista total confianza en cada uno de los proveedores de identidad. A saber, las cuentas de los usuarios comprometidas de un proveedor de identidad permitirían enlazar las correspondientes ID de las cuentas de los usuarios con los datos privados almacenados en otros nodos conectados a este proveedor. No obstante, tener esta confianza parece ser irreal bajo las circunstancias planteadas.

El solicitante ha enfrentado el problema de diseñar un método para la asignación de seudónimos dinámicos de usuarios para redes de creación de perfiles de datos que provee la anonimia de los datos mediante el empleo de seudónimos dinámicos que cambian en el tiempo a fin de evitar la capacidad de formar enlaces no restringida de los perfiles de datos que, tal como se analizó anteriormente, es inherente al método de asignación de seudónimos estáticos convencional, y es desventajosa por las razones planteadas anteriormente. El método de la presente invención funciona en el caso general de los nodos de creación de perfiles de datos (en lo sucesivo, para que sea más conciso, también se los denominará “nodos de datos”) de una red de creación de perfiles de datos, en donde los nodos de creación de perfiles de datos reciben entradas bajo seudónimos a partir de múltiples fuentes de datos u otros nodos de creación de perfiles de datos.

El solicitante ha observado que es posible sincronizar una red de creación de perfiles de datos compuesta por nodos de datos para trabajar con datos bajo seudónimos dinámicos, donde cada seudónimo cambia en el tiempo en forma aleatoria o pseudoaleatoria o como una función bajo clave de la respectiva identidad, donde la clave varía en el tiempo. El método de acuerdo con la presente invención permite que cada nodo de datos en la red de

creación de perfiles de datos encuentre el registro de datos correcto correspondiente a los datos de entrada bajo seudónimos dinámicos, para procesar los respectivos datos de entrada en datos de salida, para asignar un seudónimo dinámico a los datos de salida, y para enviar datos de salida bajo seudónimos dinámicos a otros nodos de datos en la red de creación de perfiles de datos. El acceso a los perfiles de datos de usuarios almacenados por parte de las entidades autorizadas se habilita al utilizar identidades de usuarios o seudónimos de usuarios.

El método de la presente invención también se puede aplicar a las redes de creación de perfiles de datos combinadas; es decir, las redes de creación de perfiles de datos compuestas por nodos de datos inconexos, pero que posiblemente comparten los mismos usuarios. En una forma de realización, el método de la presente invención permite intercambiar perfiles de datos de usuarios individuales correspondientes al mismo usuario en diferentes redes de creación de perfiles de datos. En otra forma de realización, el método de la presente invención permite la acumulación de perfiles de datos de usuarios correspondientes a los conjuntos especificados de usuarios (posiblemente todos ellos) en diferentes redes de creación de perfiles de datos.

De acuerdo con un aspecto de la presente invención, se provee un método de asignación de seudónimos dinámicos para una red de creación de perfiles de datos que comprende al menos un nodo de datos configurado para recibir datos de entrada relacionados con usuarios y transformar dichos datos de entrada en perfiles de datos de salida de usuarios relacionados con los usuarios, donde dicho nodo de datos comprende registros de datos de usuarios para almacenar datos de entrada relacionados con los usuarios junto con seudónimos dinámicos de entrada de los usuarios, y dicho nodo de datos está configurado para computarizar dichos perfiles de datos de salida de usuarios relacionados con un usuario a partir de dichos datos de entrada y para almacenar los perfiles de datos de salida computarizados en dichos registros de datos de usuarios del mismo. El método comprende lo siguiente:

recibir, en el nodo de datos, datos de entrada nuevos relacionados con un usuario junto con un seudónimo de usuario nuevo asociado y un seudónimo de usuario antiguo que estuvo asociado con datos de entrada previamente recibidos relacionados con el usuario en el pasado o un conjunto de seudónimos de usuarios antiguos candidatos;

en dicho nodo de datos, encontrar el registro de datos del usuario correspondiente a los datos de entrada nuevos recibidos como el registro de datos del usuario que tiene almacenado en él un seudónimo de usuario de entrada dinámico igual a dicho seudónimo de usuario antiguo recibido junto con dichos datos de entrada nuevos o a un seudónimo de usuario perteneciente al conjunto recibido de seudónimos de usuarios antiguos candidatos;

almacenar temporalmente, en el registro de datos del usuario encontrado, los datos de entrada nuevos;

establecer el seudónimo de usuario de entrada dinámico almacenado en dicho registro de datos de usuarios de dicho nodo de datos igual al último seudónimo de usuario nuevo recibido asociado con los datos de entrada relacionados con el usuario recibidos;

computarizar por momentos dichos perfiles de datos de salida de usuarios mediante la utilización de datos de entrada nuevos acumulados en el registro de datos de usuarios; almacenar los perfiles de datos de salida de usuarios computarizados en el registro de datos de usuarios; y luego borrar dichos datos de entrada nuevos acumulados del registro de datos de usuarios.

En una forma de realización de la invención, el método puede comprender lo siguiente:

en dicho nodo de datos, generar y almacenar en dicho registro de datos de usuarios un seudónimo de usuario de salida dinámico junto con dichos perfiles de datos de salida de usuarios computarizados;

enviar por momentos dichos perfiles de datos de salida de usuarios a al menos otro nodo de datos en dicha red de creación de perfiles de datos; en cada momento generar un valor nuevo de dicho seudónimo de usuario de salida dinámico; sustituir dicho valor nuevo de dicho seudónimo de usuario de salida dinámico por un valor antiguo previamente almacenado de dicho seudónimo de usuario de salida dinámico; y enviar a dicho al menos otro nodo de datos tanto el valor nuevo como el antiguo de dicho seudónimo de usuario de salida dinámico junto con dichos perfiles de datos de salida de usuarios.

Dichos datos de entrada pueden ser recibidos por el nodo de datos desde al menos un nodo de fuente de datos de la red de creación de perfiles de datos, o desde al menos otro nodo de datos de la red de creación de perfiles de datos.

En una forma de realización de la invención, el método puede comprender lo siguiente:

proveer, en dicha red de creación de perfiles de datos, al menos un nodo de asignación de seudónimos operable para realizar lo siguiente:

- recibir desde dicho nodo de fuente de datos —al menos uno— identidades de usuarios, donde dichas

identidades de usuarios identifican el usuario en dicha fuente de datos, y comprenden uno o más identificadores de usuarios conocidos para la fuente de datos;

- generar seudónimos de usuarios a partir de las identidades de usuarios recibidas;

- proveer al nodo de fuente de datos los seudónimos de usuarios generados.

Dichos seudónimos de usuarios se pueden generar como valores aleatorios o pseudoaleatorios, o valores bajo clave generados por una función bajo clave a partir de las identidades de usuarios y una clave secreta.

Dicho nodo de asignación de seudónimos puede ser operable para encriptar y autenticar los seudónimos de usuarios generados a ser provistos al nodo de fuente de datos.

Los seudónimos de usuarios aleatorios o pseudoaleatorios generados se pueden almacenar en el nodo de asignación de seudónimos en asociación con las correspondientes identidades de usuarios.

Las identidades de usuarios pueden ser diferentes para los nodos de fuentes de datos diferentes. El método puede comprender proveer, en dicha red de creación de perfiles de datos, al menos un nodo de administración de identidades de usuarios equivalentes, operable para administrar como equivalentes diferentes identidades de un mismo usuario correspondientes a distintas fuentes de datos.

En una forma de realización de la invención, el método puede comprender lo siguiente:

después de recibir, en dicho nodo de datos, datos de entrada nuevos relacionados con un usuario desde dicha al menos una fuente de datos o desde dicho al menos otro nodo de datos, si no se encuentra un registro de datos de usuarios que incluye dicho seudónimo de usuario antiguo recibido junto con los datos de entrada nuevos o un seudónimo de usuario perteneciente a dicho conjunto de seudónimos de usuarios antiguos candidatos recibidos junto con los datos de entrada nuevos, hacer que el nodo de datos determine, al explotar dicho seudónimo de usuario antiguo o dicho conjunto de seudónimos de usuarios antiguos candidatos recibidos junto con los datos de entrada nuevos, si un registro de datos de usuarios con respecto a ese usuario ya existe, en donde dicho registro de datos de usuarios se ha creado para almacenar datos de entrada relacionados con ese usuario recibidos desde al menos otra fuente de datos en el pasado.

En el caso de que los datos de entrada nuevos sean recibidos por dicho nodo de datos desde al menos una fuente de datos, dicha determinación puede comprender lo siguiente:

hacer que dicho nodo de datos envíe hacia la capa anterior (*send backwards*) una solicitud a dicha al menos una fuente de datos para obtener seudónimos equivalentes del usuario, donde dicha solicitud contiene dicho seudónimo de usuario antiguo o dicho conjunto de seudónimos de usuarios antiguos candidatos;

hacer que dicha al menos una fuente de datos recupere la identidad de usuario temporalmente almacenada en ella, envíe la identidad de usuario recuperada al nodo de administración de identidades de usuarios equivalentes, y solicite que dicho nodo de administración de identidades de usuarios equivalentes provea las identidades equivalentes del usuario al nodo de asignación de seudónimos;

hacer que el nodo de asignación de seudónimos recupere los seudónimos de usuarios equivalentes y luego los envíe a los nodos de fuentes de datos conectados a él;

realizar un proceso de inundación de reenvíos (*forward flooding*) que comprende lo siguiente:

- hacer que los nodos de fuentes de datos conectados al nodo de asignación de seudónimos reenvíen a todos los nodos de datos conectados a éste solicitudes que contienen los seudónimos equivalentes recibidos del usuario;

- cuando un nodo de datos en dicha red de creación de perfiles de datos recibe en sus entradas una o más solicitudes que contienen seudónimos equivalentes desde al menos otro nodo de datos conectados a éste, hacer que el nodo de datos busque el registro de datos de usuarios y almacene uno de los seudónimos equivalentes recibidos como seudónimos de usuarios de entrada;

- si dicho registro de datos de usuarios es encontrado, y el nodo de datos es dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos use el registro de datos de usuarios encontrado para almacenar los datos de entrada nuevos recibidos;

– si dicho registro de datos de usuarios es encontrado, y el nodo de datos es diferente de dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos continúe el proceso de inundación de reenvíos al reenviar los seudónimos de usuarios de salida almacenados en el registro de datos de usuarios a todos los otros nodos de datos conectados a sus salidas.

En caso de que los datos de entrada nuevos sean recibidos por dicho nodo de datos desde dicho al menos otro nodo de datos, dicha determinación puede comprender realizar un proceso de retroseguimiento, un proceso de recuperación de seudónimos equivalentes y un proceso de inundación de reenvíos, en donde se observa lo siguiente:

dicho proceso de retroseguimiento comprende lo que se indica a continuación:

- hacer que dicho nodo de datos envíe de regreso una solicitud a dicho al menos otro nodo de datos para obtener seudónimos de usuarios equivalentes, donde dicha solicitud contiene dicho seudónimo de usuario antiguo o dicho conjunto de seudónimos de usuarios antiguos candidatos;

- hacer que dicho al menos otro nodo de datos busque el registro de datos de usuarios que almacena uno de los seudónimos de usuarios antiguos recibidos como seudónimo de usuario de salida y luego envíe de regreso al menos una solicitud a cualesquiera otros nodos de datos conectados a sus entradas, donde dicha solicitud contiene el seudónimo de usuario de entrada almacenado en el registro de datos de usuarios;

- cuando un nodo de datos en dicha red de creación de perfiles de datos recibe en cualquiera de sus salidas una solicitud desde cualquier otro nodo de datos conectados a éste, hacer que el nodo de datos busque el registro de datos de usuarios que almacena el seudónimo de usuario recibido como seudónimo de usuario de salida y luego envíe de regreso al menos una solicitud a cualquiera de los otros nodos de datos o cualquiera de los nodos de fuentes de datos conectados a sus entradas, donde dicha solicitud contiene el seudónimo de usuario de entrada almacenado en el registro de datos de usuarios;

- cuando un nodo de fuente de datos en dicha red de creación de perfiles de datos recibe en su salida una solicitud desde cualquier nodo de datos conectados a éste, hacer que la fuente de datos reenvíe el seudónimo de usuario recibido al nodo de asignación de seudónimos con una solicitud para proveer los seudónimos de usuarios equivalentes;

dicho proceso de recuperación de seudónimos equivalentes comprende lo siguiente:

- hacer que el nodo de asignación de seudónimos reciba desde dicho nodo de fuente de datos un seudónimo de usuario y luego recupere los seudónimos de usuarios equivalentes, ya sea directamente mediante la utilización de una tabla almacenada de seudónimos o bien indirectamente mediante la utilización de una función bajo clave invertible para recuperar la identidad de usuario, luego enviar esta identidad de usuario al nodo de administración de identidades de usuarios equivalentes para proveer identidades de usuarios equivalentes, luego generar los seudónimos de usuarios equivalentes candidatos a partir de las identidades de usuarios equivalentes recibidas, y luego enviarlas a los nodos de fuentes de datos conectados a éste;

dicho proceso de inundación de reenvíos comprende lo siguiente:

- hacer que los nodos de fuentes de datos conectados al nodo de asignación de seudónimos reenvíe a todos los nodos de datos conectados a éste las solicitudes que contienen los seudónimos de usuarios equivalentes recibidos;

- cuando un nodo de datos en dicha red de creación de perfiles de datos recibe en sus entradas una o más solicitudes que contienen seudónimos equivalentes desde al menos otro nodo de datos conectado a éste, hacer que el nodo de datos busque el registro de datos de usuarios y almacene uno de los seudónimos equivalentes recibidos como seudónimos de usuarios de entrada;

- si dicho registro de datos de usuarios es encontrado, y el nodo de datos es dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos use el registro de datos de usuarios encontrado para almacenar los datos de entrada nuevos recibidos;

– si dicho registro de datos de usuarios es encontrado, y el nodo de datos es diferente de dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos continúe el proceso de inundación de reenvíos al reenviar los seudónimos de usuarios de salida almacenados en el registro de datos de usuarios a todos los otros nodos de datos conectados a sus salidas.

En una forma de realización de la invención, el método puede comprender lo siguiente:

en caso de que, en dicho nodo de datos, el registro de datos de usuarios correspondientes a los datos de entrada nuevos recibidos no se encuentre incluso después de recibir los seudónimos equivalentes de usuario, hacer que el nodo de datos cree un nuevo registro de datos de usuarios con respecto a dicho usuario, y almacene los datos de entrada nuevos recibidos en éste junto con el seudónimo de usuario nuevo recibido junto con los datos de entrada nuevos recibidos.

Dichos seudónimos de usuarios pueden cambiarse dinámicamente después de un tiempo predeterminado según un período de validez de los perfiles de datos de usuarios.

Preferentemente, se evita que dicho al menos un nodo de datos y dicho al menos un nodo de fuente de datos asignen, usen o almacenen, en asociación con los datos relacionados con un usuario, seudónimos de usuarios estáticos que no cambian en el tiempo o valores antiguos de seudónimos de usuarios dinámicos generados y usados en el pasado para el usuario.

En una forma de realización de la invención, el método puede comprender lo siguiente:

cuando una entidad que solicita un perfil de datos de usuario desea recuperar el perfil de datos de usuario almacenados en dicho al menos un nodo de datos, se realiza lo siguiente:

- hacer que la identidad que solicita el perfil de datos de usuario envíe una solicitud de perfil de datos de usuario a un nodo seleccionado de entre dicho al menos un nodo de datos de la red de creación de perfiles de datos, en donde dicha solicitud de perfil de datos de usuario contiene un identificador del nodo de datos que almacena los perfiles de datos de usuario solicitados a ser recuperados, y el seudónimo de usuario de entrada o de salida actualmente válido, respectivamente almacenados en el registro de datos de usuarios de dicho nodo seleccionado de entre dicho al menos un nodo de datos al cual la solicitud de perfil de datos de usuario es enviada;

- en el caso de que dicho nodo seleccionado de entre dicho al menos un nodo de datos que recibe la solicitud de perfil de datos de usuario sea el nodo de datos que almacena los perfiles de datos de usuario solicitados, hacer que el nodo de datos recupere los perfiles de datos de usuario solicitados almacenados en el registro de datos de usuarios en asociación con el seudónimo de usuario de entrada o de salida actualmente especificado, y provea a la entidad que solicita el perfil de datos de usuario los perfiles de datos de usuarios recuperados;

- en el caso de que dicho nodo seleccionado de entre dicho al menos un nodo de datos que recibe la solicitud de perfil de datos de usuario no sea el nodo de datos que almacena los perfiles de datos de usuario solicitados, se deberá realizar lo siguiente:

a) hacer que dicho nodo seleccionado de entre dicho al menos un nodo de datos identifique, en los registros de datos de usuarios almacenados en éste, el seudónimo de usuario de salida correspondiente al seudónimo de usuario de salida actualmente válido recibido, contenido en la solicitud de perfil de datos de usuario recibida, o el seudónimo de usuario de salida correspondiente al seudónimo de usuario de entrada actualmente válido, recibido en la solicitud de perfil de datos de usuario recibida, y reenviar el seudónimo de usuario de salida recuperado a todos los otros nodos de datos conectados a éste; y

b) repetir el paso a) mediante la utilización del seudónimo de usuario de salida recuperado recibido en lugar de dicho seudónimo de usuario de salida actualmente válido o dicho seudónimo de usuario de entrada actualmente válido, hasta llegar al nodo de datos que almacena los perfiles de datos de usuario solicitados y, entonces, hacer que el nodo de datos recupere los perfiles de datos de usuarios solicitados, almacenados en el correspondiente registro de datos de usuarios.

El método también puede comprender, antes de realizar los pasos anteriores, hacer que la entidad que solicita los perfiles de datos de usuarios solicite al nodo de asignación de seudónimos un seudónimo de usuario de entrada actualmente válido correspondiente a un identificador de usuario temporal o permanente determinado, contenido en dicha solicitud de perfil de datos de usuario.

La red de creación de perfiles de datos puede comprender al menos una primera y una segunda redes de creación de perfiles de datos inconexas, cada una de las cuales comprende respectivos nodos de datos, respectivas fuentes de datos que proveen datos de entrada sobre la base de las cuales los perfiles de datos de usuarios son calculados por los nodos de datos, respectivos nodos de asignación de seudónimos para generar seudónimos de usuarios a partir de identidades de usuarios, y en donde se provee un nodo de administración de identidades de usuarios equivalentes combinadas, operable para administrar como equivalentes diferentes identidades de un mismo usuario correspondientes a diferentes fuentes de datos en la primera y la segunda redes de creación de perfiles de datos, donde el nodo de administración de identidades de usuarios equivalentes combinadas es explotado para recuperar perfiles de datos de usuarios de un usuario en la segunda red de creación de perfiles de datos cuando los perfiles de datos de dicho usuario son solicitados a través de la primera red de creación de perfiles de datos.

Otro aspecto de la presente invención se refiere a una red de creación de perfiles de datos configurada para realizar el método anterior.

Breve descripción de los dibujos

Éstas y otras características y ventajas de la presente invención serán evidentes a partir de la siguiente descripción detallada de las formas de realización de la misma, a modo de ejemplo mas no de limitación, con referencia a las

figuras adjuntas, en las cuales:

La **FIG. 1** ilustra la estructura general de una red de creación de perfiles de datos de ejemplo en la cual es posible implementar el método de acuerdo con la presente invención, la red de creación de perfiles de datos que comprende nodos-D, un nodo-ID, un nodo-PID fiable así como también fuentes de datos (DS, *Data Source* o nodos-DS) que proveen entrada de datos sin procesar a la red (en el ejemplo considerado, 4 fuentes de datos y 8 nodos-D);

La **FIG. 2** ilustra la estructura general de un nodo-D de la red de creación de perfiles de datos (en el ejemplo considerado en la presente, un nodo-D con 3 canales de entrada y 2 canales de salida);

La **FIG. 3** ilustra la interacción entre una DS, el nodo-PID y un nodo-D, que reciben entrada desde un nodo-DS, y un cambio de seudónimos y datos en el nodo-D que reciben entrada desde el nodo-DS;

La **FIG. 4** ilustra el cambio de datos y seudónimos en un nodo-D que recibe entrada desde otro nodo-D;

La **FIG. 5** ilustra una cadena de seudónimos dinámicos formados en un determinado momento por 4 nodos-D consecutivos en la red de creación de perfiles de datos.

DESCRIPCIÓN DETALLADA DE LAS FORMAS DE REALIZACIÓN DE LA INVENCIÓN

Red de creación de perfiles de datos

De acuerdo con una forma de realización de la presente invención, una red de creación de perfiles de datos (en lo sucesivo, resumida como “DPN”, *Data Profiling Network*) es una red informática que comprende tres tipos de nodos lógicos: (1) nodos de identidad (en lo sucesivo, también denominados “nodos-ID”), (2) nodos de seudónimos (en lo sucesivo, también denominados “nodos-PID”), y (3) los nodos de creación de perfiles de datos o, simplemente, nodos de datos (en lo sucesivo, también denominados nodos-D). Los nodos lógicos de la DPN se pueden implementar físicamente en forma centralizada o distribuida mediante la utilización de tecnología informática.

Los datos de entrada a la DPN (en lo sucesivo, también denominada “datos de entrada sin procesar”) son suministrados por un número de fuentes de datos (DS o nodos-DS). Asimismo, los componentes asociados de la DPN son usuarios; es decir, entidades del mundo real que proveen datos de entrada sin procesar a través de varias fuentes de datos y utilizan el perfil de datos de salida producido por la DPN. Por ejemplo, los usuarios pueden ser individuos (personas) o grupos de personas, empresas, organizaciones, sitios web de Internet, o dispositivos tales como computadoras personales o teléfonos móviles.

Con referencia al ejemplo ilustrado en la **FIG. 1**, la DPN –en la presente bajo la referencia numérica **100**- comprende un nodo-ID **105**, un nodo-PID **110**, ocho nodos-D **115_1** - **115_8**, y cuatro fuentes de datos (nodos-DS) **120_1** - **120_4**. Los cuatro nodos-DS **120_1** - **120_4** pueden corresponder, por ejemplo, a cuatro sitios web de Internet distintos que recaban datos de entrada sin procesar sobre el uso del sitio web por parte de los usuarios, mientras que los ocho nodos-D **115_1** - **115_8** acumulan y procesan los datos de entrada recibidos relacionados con un mismo usuario en los correspondientes datos de salida (perfiles de datos) a ser usados para diferentes aplicaciones (por ejemplo, para la comercialización focalizada por parte de entidades autorizadas mediante anuncios publicitarios personalizados). En forma alternativa, los cuatro nodos-DS **120_1** - **120_4** pueden corresponder a diferentes redes (por ejemplo, fijas o móviles) de ISP iguales o distintos, en cuyo caso los datos de entrada sin procesar se extraen de los correspondientes datos de tráfico (por ejemplo, mediante sondas de red).

Los nodos-ID (como el nodo-ID **105** en el ejemplo considerado) administran identidades equivalentes (en lo sucesivo, también denominadas “ID”) de los usuarios de la DPN, en donde cada una de las identidades equivalentes de cierto usuario corresponde a una fuente de datos diferente. La hipótesis subyacente es que, en la DPN, todos los usuarios representados por una misma identidad con respecto a cualquier fuente de datos son efectivamente considerados como un único usuario. Un nodo-ID implementa, por ende, una tabla de equivalencia de identidades que almacena identidades equivalentes correspondientes al mismo usuario, donde cada identidad corresponde a una fuente de datos distinta y, en sí misma, está compuesta por un conjunto ordenado de identificadores del usuario. En la tabla de equivalencia de identidades mantenida por el nodo-ID, las identidades ID son indexadas por las respectivas fuentes de datos *i* (*i* es un índice que identifica una única fuente de datos: por ejemplo, **120_i**, con **i** = 1, 2, 3, 4), y la tabla de equivalencia de identidades consiste, por ende, en conjuntos de identidades equivalentes $\{(i, \mathrm{ID})\}$, un conjunto por cada usuario único.

Los nodos-ID tienen un modo de operación actualizado: en el modo de operación actualizado, un nodo-ID actualiza la respectiva tabla de equivalencia de identidades sobre la base de las identidades equivalentes recibidas en su entrada. Además, un nodo-ID también puede generar y almacenar una identidad única nueva para un usuario que puede ser explotada como una representación única (posiblemente estática) del conjunto de identidades equivalentes de ese usuario.

Los nodos-ID tienen un modo de operación de generación: en el modo de operación de generación, un nodo-ID recibe en su entrada, desde una fuente de datos *i*, una identidad ID –es decir (*i*, ID)- de un usuario, y provee el correspondiente conjunto de identidades equivalentes {(*i*, ID)} para ese usuario en su salida. Formalmente, en el modo de operación de generación, un nodo-ID realiza un mapeo (*i*, ID) → {(*i*, ID)}.

Para una determinada DPN, puede existir un único nodo-ID que administra las identidades de todos los usuarios y con respecto a todas las fuentes de datos, o es posible proveer alternativamente múltiples nodos-ID mutuamente conectados, cada uno de los cuales administra las identidades para los respectivos subconjuntos de usuarios y fuentes de datos. Cada nodo-ID puede implementarse como un servidor de computación en una red informática. No obstante, sin pérdida de generalidad, la pluralidad de nodos-ID que sirven una única DPN pueden considerarse como un único nodo-ID lógico; en lo sucesivo, se asumirá entonces que la DPN comprende un único nodo-ID.

Los identificadores y las identidades de usuarios en la DPN son administrados por las fuentes de datos y por los nodos-ID. Cada identificador de usuario se puede clasificar como temporal (es decir, a corto plazo) o permanente (es decir, a largo plazo) según si el período de validez del identificador de usuario es relativamente corto o relativamente largo, respectivamente. Un ejemplo de identificador permanente es el nombre de una persona (que no se espera que cambie durante la vida de la persona) o su domicilio (que puede cambiar esporádicamente), mientras que un ejemplo de identificador temporal es un identificador de localización que describe la ubicación temporal de una fuente de datos, tal como una dirección IP en Internet. Los nodos-ID pueden administrar identificadores de usuarios solamente permanentes y las correspondientes identidades equivalentes; no obstante, en algunos casos, los identificadores de usuarios temporales también pueden ser administrados por los nodos-ID conjuntamente con los identificadores permanentes.

Una identidad de usuario puede estar compuesta por uno o más identificadores permanentes y uno o más identificadores temporales; por ende, la asignación de seudónimos puede relacionarse ya sea con la identidad en un todo o bien con identificadores particulares (permanentes y/o temporales) o con subconjuntos de identificadores seleccionados entre los identificadores que componen la identidad. En particular, puede ser deseable asignar seudónimos a identificadores temporales individuales en forma separada de los identificadores permanentes de una identidad.

Un nodo-PID (como nodo-PID **110** en el ejemplo considerado) transforma (es decir, asigna seudónimos) un determinado identificador de usuario o una identidad de usuario en su totalidad, definido como un conjunto ordenado de identificadores de usuarios (permanentes y temporales), en un seudónimo correspondiente. De acuerdo con la presente invención, si bien el acrónimo ID se usa en forma equivalente como una denominación genérica para una identidad en un todo o para un identificador perteneciente al conjunto de identificadores que constituyen una identidad, el acrónimo PID se usa como denominación genérica para el seudónimo correspondiente. Un nodo-PID recibe una ID en su entrada y genera el correspondiente seudónimo PID en su salida. Formalmente, un nodo-PID realiza un mapeo inyectivo ID → PID, el cual puede depender de una fuente de datos particular y, posiblemente, de identificadores individuales o subconjuntos de identificadores para una determinada fuente de datos. En cualquier momento, cada fuente de datos usa la misma ID para el mismo usuario; es decir, la DS no puede usar dos ID diferentes para el mismo usuario al mismo tiempo; sin embargo, en el tiempo, la ID de un usuario usada por una DS puede cambiar (por ejemplo, si la ID es temporal). Debería destacarse que, si el mapeo ID → PID realizado por el nodo-PID no es inyectivo, entonces los subconjuntos de las ID generados en un mismo PID son efectivamente tratados como usuario único. Tal como se explicó con anterioridad, se asume implícitamente que las ID equivalentes (identidades o identificadores individuales o subconjuntos de identificadores que conforman las identidades) originadas desde diferentes fuentes de datos y los correspondientes seudónimos son indexados por las respectivas fuentes de datos. Por consiguiente, una identidad con respecto a cierta fuente de datos es transformada en un único PID si todos los identificadores comprendidos en la identidad son transformados conjuntamente. En forma alternativa, si los identificadores comprendidos en la identidad son transformados individualmente, entonces la identidad es transformada en un conjunto ordenado de PID de seudónimos individuales, cada uno de los cuales corresponde a los identificadores individuales o sus subconjuntos. Incluso más generalmente, si un nodo-PID recibe un conjunto {ID} de ID equivalentes correspondientes a una o más fuentes de datos indexadas en su entrada, entonces genera un correspondiente conjunto {PID} de seudónimos equivalentes en su salida. Para la protección de la privacidad, se asume que los nodos-PID son fiables, y no revelan las asociaciones ID → PID a entidades no autorizadas. Cada nodo-PID se puede implementar como un servidor informático seguro en una red informática o como módulo de seguridad de hardware (HSM, *Hardware Security Module*) inviolable.

Los nodos-D (como nodos-D **115_1** - **115_8** en el ejemplo considerado) lidian con los datos de usuario a ser perfilados dentro de la DPN; es decir, los nodos-D generan perfiles de datos de usuarios.

Los nodos-D de la DPN pueden ser nodos-D de entrada, nodos-D intermedios y nodos-D de salida.

Un nodo-D de entrada (como nodos-D **115_1** - **115_3** y **115_5** en el ejemplo considerado) recibe datos de entrada sin procesar directamente desde al menos una fuente de datos (en el ejemplo considerado, el nodo-D **115_1** de entrada recibe datos de entrada sin procesar directamente desde las fuentes de datos **DS_1** (**120_1**) y **DS_2** (**120_2**); el nodo-D de

entrada **115_2** recibe datos de entrada sin procesar directamente desde la fuente de datos **DS_3**; el nodo-D de entrada **115_3** recibe datos de entrada sin procesar directamente desde las fuentes de datos **DS_3** (**120_3**) y **DS_4** (**120_4**); y el nodo-D de entrada **115_5** recibe datos de entrada sin procesar directamente desde la fuente de datos **DS_2**) y, posiblemente, recibe datos de entrada también desde otros nodos-D (como, por ejemplo, el nodo-D de entrada **115_5**, que recibe datos de entrada sin procesar también desde el nodo-D **115_1**). Los nodos-D de entrada de primer nivel reciben datos de entrada sin procesar sólo directamente desde al menos una fuente de datos, y no desde otros nodos-D (es decir, por ejemplo, el caso de los nodos-D **115_1**, **115_2**, **115_3**).

Un nodo-D intermedio recibe datos de entrada solamente desde otros nodos-D (y no desde cualquier fuente de datos) y, ante el procesamiento, provee datos de salida a otros nodos-D, pero no directamente a los usuarios. Por consiguiente, basta que los nodos-D intermedios implementen sólo el almacenamiento temporal de datos de salida. En el ejemplo considerado, el nodo-D **115_6** es un nodo intermedio si no provee datos de salida directamente a los usuarios.

Un nodo-D de salida recibe datos de entrada desde las fuentes de datos u otros nodos-D y, ante el procesamiento, provee datos de salida directamente a los usuarios y, posiblemente, también a otros nodos-D. Por consiguiente, los nodos-D de salida implementan el almacenamiento permanente de datos de salida, por ejemplo, en forma de base de datos o cualquier otra memoria no volátil. En el ejemplo considerado, los nodos-D **115_4**, **115_7**, **115_8** son nodos-D de salida ya que no proveen datos de salida a otros nodos-D. Un tipo especial de nodos-D de salida está constituido por los nodos-D de usuarios, que almacenan perfiles de datos de usuarios en el equipo de los usuarios (por ejemplo, computadoras personales o teléfonos móviles).

La conexión orientada entre una DS genérica y un nodo-D genérico (como la conexión orientada **125** entre la DS **120_1** y el nodo-D **115_1** en el ejemplo considerado) y entre dos nodos-D genéricos en la DPN (como la conexión orientada **130_1** entre el nodo-D **115_1** y el nodo-D **115_4**, y la conexión orientada **130_2** entre el nodo-D **115_1** y el nodo-D **115_5** en el ejemplo considerado) se denomina “canal de datos” o simplemente “canal”. Todos los datos transmitidos entre las DS y los nodos-D y entre los nodos-D en los canales de datos reciben seudónimos; es decir, son rotulados mediante seudónimos dinámicos, que, de acuerdo con la presente invención, pueden utilizarse para enlazar de manera conjunta los datos correspondientes al mismo usuario en diferentes momentos, al aplicar el método de asignación de seudónimos dinámicos descripto con posterioridad.

Funcionalidad de los nodos-D

Un nodo-D en la DPN puede tener un número de entradas y un número de salidas conectadas a otros nodos-D mediante canales de salida (con referencia al ejemplo considerado, el nodo-D **115_1** tiene dos salidas, respectivamente conectadas al nodo-D **115_4** y al nodo-D **115_5** mediante un respectivo canal de salida **130_1** y **130_2**). Una salida genérica de un nodo-D puede estar conectada a una o más entradas de otros nodos-D mediante canales de salida individuales. Tal como se explicó con anterioridad, si un nodo-D es un nodo-D de salida, entonces pueden existir salidas del nodo-D sin ningún canal de salida: en este caso, los correspondientes datos de salida no se envían a otros nodos-D, pero deberían estar accesibles para los usuarios autorizados. Cada entrada a un nodo-D está conectada a una fuente de datos (DS) o a otro nodo-D mediante un canal de entrada, que suministra los datos de entrada correspondientes a los diferentes usuarios y rotulados mediante seudónimos.

Las entradas y salidas individuales para un nodo-D genérico pueden ser indexadas mediante *j* y *k*, respectivamente. Luego, los datos de entrada bajo seudónimos recibidos por un nodo-D pueden almacenarse temporalmente en un registro de datos de entrada $\{(j, Data_{in}, PID_{in})\}$, donde *j* es el índice que identifica la entrada *j*-ésima del nodo-D, $Data_{in}$ es el dato de entrada recibido por el nodo-D en la entrada *j*-ésima, y el seudónimo PID_{in} de entrada es un seudónimo de valor único o un seudónimo de valores múltiples; es decir, un conjunto ordenado de valores de seudónimos que representan el dato de entrada $Data_{in}$ para la entrada *j*-ésima. En forma similar, los datos de salida bajo seudónimos de un nodo-D se pueden almacenar en un registro de datos de salida $\{(k, Data_{out}, PID_{out})\}$, donde *k* es el índice que identifica la salida *k*-ésima del nodo-D, $Data_{out}$ es el dato de salida, y el seudónimo de salida PID_{out} es un seudónimo de valor único que representa el dato de salida $Data_{out}$ para la salida *k*-ésima. Se asume que el seudónimo de salida toma un valor único, para que sea más sencillo, mientras que el seudónimo de entrada puede tomar múltiples valores si los datos de entrada son recibidos directamente desde una DS. Para distinguirlos, los seudónimos de entrada de múltiples valores también se denominarán, en lo sucesivo, “(PID_{in})”. Debería destacarse que el seudónimo de salida PID_{out} de la salida *k*–ésima no es necesario si esta salida no tiene canales de salida asociados. Los registros de los datos de entrada y de salida se pueden almacenar juntos en un registro de datos conexo ($\{(j, Data_{in}, PID_{in})\}$; $\{(k, Data_{out}, PID_{out})\}$) correspondiente a un usuario anónimo. Si los registros de los datos de entrada y de salida se almacenan por separado, entonces deberían compartir los seudónimos de entrada y de salida en común, que apuntan al mismo usuario anónimo. La **FIG. 2** ilustra esquemáticamente un nodo-D **115** con tres entradas **205_1**, **205_2**, **205_3**, con los registros de datos de entrada asociados **210_1**, **210_2**, **210_3**, y dos salidas **215_1**, **215_2**, con los registros de datos de salida asociados **220_1**, **220_2**.

Una DPN puede estar representada por un gráfico dirigido que consiste en nodos-D y DS que actúan como nodos de entrada externos. La hipótesis subyacente es que el correspondiente gráfico no dirigido está conectado. De lo contrario, la DPN podría dividirse en un número de componentes inconexos que funcionan por separado unos de

los otros.

De acuerdo con la forma de realización de ejemplo aquí descripta del método de asignación de seudónimos dinámicos de la presente invención, los seudónimos de entrada y de salida almacenados en los registros de datos ({(j, $Data_{in}$, PID_{in})}; {(k, $Data_{out}$, PID_{out})}) de los nodos-D en la DPN cambian en el tiempo de tal manera que permiten que cada nodo-D encuentre, en cada momento, el registro de datos correcto correspondiente a los datos de entrada bajo seudónimos recibidos. El principal paradigma que permite esta funcionalidad es que, en cada momento, el seudónimo de entrada nuevo de un nodo-D (nodo-D receptor) es definido como el último seudónimo de salida nuevo recibido desde el correspondiente nodo-D anterior (nodo-D emisor) en la DPN. En consecuencia, el seudónimo de entrada antiguo del nodo-D receptor almacenado en su registro de datos de entrada {(j, $Data_{in}$, PID_{in})} es, por ende, igual al seudónimo de salida antiguo del nodo-D receptor, almacenado en el registro de datos de salida {(k, $Data_{out}$, PID_{out})} del nodo-D emisor, recibido por el nodo-D receptor en el pasado. Esto permite que el nodo-D receptor encuentre el registro de datos de usuarios correcto mediante la utilización del seudónimo de salida antiguo, que necesita ser enviado por el nodo-D emisor junto con el seudónimo de salida nuevo. El seudónimo de salida antiguo recibido coincidirá entonces con el valor almacenado del seudónimo de entrada antiguo en el respectivo registro de datos de entrada del nodo-D receptor, correspondiente al usuario correcto, pero anónimo, siempre que este valor ya exista en la tabla de registros de datos almacenados del nodo-D receptor. Después de encontrar el registro de datos correcto, el seudónimo de salida nuevo recibido es entonces sustituido por el seudónimo de entrada antiguo almacenado (es decir, el seudónimo de salida nuevo recibido se transforma en el seudónimo de entrada nuevo).

Un nodo-D acumula datos de entrada y actualiza datos de salida (es decir, perfiles de usuarios) al procesar los datos de entrada acumulados en determinados momentos (por ejemplo, periódicamente). Entre dos actualizaciones sucesivas, los datos de entrada se acumulan solamente y, por ende, el $Data_{in}$ denota todos los datos de entrada acumulados para cierto usuario desde la última actualización, para cualquier entrada de nodo-D determinada. Por otro lado, tal como se describió con anterioridad, para cualquier entrada determinada, los seudónimos de entrada no se acumulan sino que solamente se renuevan en los valores nuevos recibidos y, por ende, en cada momento, el PID_{in} denota el último seudónimo de entrada (desde la ultima actualización) del nodo-D receptor, que es igual al último valor de seudónimo de salida recibido del respectivo nodo-D emisor.

Los datos de salida de un nodo-D son actualizados (es decir, reemplazados por un valor computarizado nuevo) al procesar los datos de entrada acumulados con estado o sin estado; un nodo-D que actualiza los datos de salida sin estado se refiere a un “nodo-D sin estado”, mientras que un nodo-D que actualiza los datos de salida con estado se refiere a un “nodo-D con estado”. Por “actualización sin estado” se hace referencia a que, por cada salida de nodo-D, el dato de salida $Data_{out}$ se computariza como una función de {(j, $Data_{in}$)}, mientras que por “actualización con estado” se hace referencia a que, por cada salida de nodo-D, el dato de salida $Data_{out}$ se computariza como una función de {(j, $Data_{in}$)} y la variable adicional S se refiere al estado correspondiente a esa salida de nodo-D. Asimismo, por cada salida de nodo-D, el estado S se actualiza como una función del estado anterior y {(j, $Data_{in}$)}. Esto implica que el registro de datos de salida de un nodo-D con estado contiene {(k, S, $Data_{out}$, PID_{out})}, donde S es el estado para la salida k-ésima del nodo-D. Por ejemplo, si los datos de salida se definen como el valor promedio de los datos de entrada actuales y los datos de entrada anteriores en un período pasado, entonces el estado comprende todos los datos de entrada anteriores que se necesitan para la computación del valor promedio. Se asume que la actualización ocurre al mismo tiempo para todas las salidas de los nodos-D (si no, cuando un nodo-D puede dividirse lógicamente en un número de nodos-D). Después de cada actualización, el dato de entrada (temporal) $Data_{in}$ se borra del registro de datos de entrada, que luego se transforma simplemente en {(j, PID_{in})}, es decir, solamente quedan por almacenarse los seudónimos de entrada.

Si un nodo-D tiene al menos un canal de salida conectado a sus salidas, entonces, en determinados momentos, el dato de salida $Data_{out}$ es emitido a través del (los) respectivo(s) canal(es) de salida al (los) correspondiente(s) nodo(s)-D en la DPN (por ejemplo, el nodo-D **115_1** emite el dato de salida $Data_{out}$ al nodo-D **115_4** a través del canal de salida **130_1**). En particular, esto puede ocurrir luego de cada actualización de los datos de salida. No es necesario enviar los datos de salida al mismo tiempo para todas las salidas de los nodos-D. Para que se emita el dato de salida $Data_{out}$, el seudónimo de salida antiguo PID_{out}^{old} se define como el valor (almacenado) anterior de PID_{out} y el seudónimo de salida PID_{out} se actualiza al computarizar el valor del seudónimo de salida nuevo como una función de los seudónimos de entrada actuales {(PID_{in})} de todas las entradas (al momento del envío). Preferentemente, esta función debería ser sensible a los cambios de una o más entradas individuales. En particular, si los seudónimos de entrada actuales {(PID_{in})} contienen solamente un valor de seudónimo PID_{in}, entonces el seudónimo de salida nuevo PID_{out} se puede definir como $PID_{out}=PID_{in}$. Más generalmente, el seudónimo de salida PID_{out} se puede definir como XOR en bits de los seudónimos de entrada individuales representados como cadenas binarias. El dato de salida $Data_{out}$ luego es emitido junto con el valor antiguo y el valor nuevo del correspondiente seudónimo de salida PID_{out}^{old}, PID_{out}, es decir, como el triple ($Data_{out}$, PID_{out}^{old}, PID_{out}). Tal como se explicó con anterioridad, el registro de datos de usuarios correcto en el nodo-D receptor se puede encontrar entonces mediante la utilización del seudónimo de salida antiguo transmitido PID_{out}^{old}, ya que, en el nodo-D receptor, este valor coincide con el valor almacenado del seudónimo de entrada en el respectivo registro de datos. Una vez encontrado el registro de datos de usuarios correcto, el valor del seudónimo de entrada en el registro de datos de entrada del nodo-D receptor se renueva entonces al valor nuevo PID_{out}.

El proceso descripto con anterioridad se esquematiza en la **FIG. 4**, en la cual se ilustran dos nodos-D interconectados genéricos **115$_a$** y **115$_b$**, en donde el nodo-D **115$_a$** es el nodo-D emisor y el nodo-D **115$_b$** es el nodo-D receptor. La figura ilustra la evolución de los dos nodos-D **115$_a$** y **115$_b$**, en términos de los datos y los valores de PID almacenados en los registros de entrada y de salida de los mismos, antes y después de que los datos de salida sean emitidos por el nodo-D **115$_a$** emisor y antes y después de una actualización en el nodo-D receptor **115$_b$**.

Si un nodo-D es un nodo-D intermedio, entonces el dato de salida (temporal) $Data_{out}$ se borra del registro de datos de salida una vez que es emitido a través del(los) respectivo(s) canal(es) de salida, mientras que el seudónimo de salida permanece almacenado (a fin de ser emitido junto con el futuro seudónimo de salida nuevo en el próximo envío del dato de salida actualizado $Data_{out}$). La correspondiente parte del registro de datos de salida para la salida *k*–ésima de un nodo-D intermedio se transforma luego en (*k*, *S*, PID_{out}). Para que sea posible borrar, se asume que, si un nodo-D es intermedio y existen múltiples canales de salida conectados a la misma salida del nodo-D, entonces la emisión de los datos de salida se lleva a cabo al mismo tiempo a través de todos estos canales de salida.

Si un nodo-D es un nodo-D de salida, entonces el dato de salida $Data_{out}$ se almacena a fin de ser accesible por parte de los usuarios autorizados. En este caso, pueden o no existir canales de salida y, si existen los canales de salida, entonces el dato de salida no se borra una vez que se emitió a través del(los) canal(es) respectivo(s) de salida. En particular, los canales de salida pueden conducir a los nodos-D de usuarios. Si no existen canales de salida para una determinada salida, entonces el seudónimo de salida PID_{out} en realidad no es necesario para esa salida, porque es posible encontrar y acceder al dato de salida $Data_{out}$ mediante la utilización de cualquiera de los correspondientes seudónimos de entrada actuales desde {(PID_{in})}, tal como se describe con posterioridad.

La funcionalidad de un nodo-D de entrada con respecto a los datos de entrada bajo seudónimos recibidos directamente desde una DS en lugar de otro nodo-D es análoga, con la única diferencia de que el seudónimo de salida nuevo recibido desde una DS puede ser un seudónimo de valor único (como en el caso de cualquier nodo-D que no es un nodo de entrada) o un seudónimo de valores múltiples, es decir, un conjunto ordenado de seudónimos correspondientes a diferentes identificadores o subconjuntos de identificadores del mismo usuario con respecto a la DS determinada. Este seudónimo de salida nuevo de valor único o valores múltiples PID_{out} es generado, ante la solicitud desde la DS, por un nodo-PID (como el nodo-PID **110** en el ejemplo considerado) y enviado a la DS junto con el seudónimo antiguo de valor único o valores múltiples o un conjunto de seudónimos de salida antiguos candidatos. Por consiguiente, cuando el nodo-D de entrada receptor encuentra el registro de datos correcto, entonces el seudónimo de entrada PID_{in} del nodo-D de entrada receptor se establece igual al seudónimo de salida nuevo de valor único o valores múltiples PID_{out} recibido desde la DS emisora, respectivamente. Por otro lado, tal como se describió previamente, cuando el dato de entrada bajo seudónimo es recibido desde otro nodo-D, entonces el seudónimo de salida recibido es siempre un seudónimo de valor único PID_{out} generado por el nodo-D emisor y, por ende, el correspondiente seudónimo de entrada PID_{in} del nodo-D receptor es entonces un único seudónimo.

Funcionalidad de nodos-PID

Un nodo-PID (como el nodo-PID **110** en la DPN de ejemplo aquí considerado) opera en la DPN como una entidad de servicio con respecto a las fuentes de datos y los nodos-ID. El nodo-PID puede recibir una ID como un único identificador o una única identidad (conjunto de identificadores) desde una DS que emite una solicitud (como una de las cuatro DS **120$_1$** - **120$_4$**, en el ejemplo considerado) en su entrada, en cuyo caso el nodo-PID genera el correspondiente PID de valor único nuevo en su salida y luego envía de regreso este PID generado nuevo a la DS que emite solicitudes, junto con el conjunto de posible valores antiguos del PID. Si un nodo-PID recibe en su entrada desde una DS una ID como un conjunto ordenado de identificadores o subconjuntos de éstos a ser transformados en forma individual, entonces el nodo-PID genera el correspondiente PID de valores múltiples nuevo como un conjunto ordenado de valores de PID en su salida, y luego envía de regreso este PID de valores múltiples nuevo a la DS, junto con el conjunto de posibles valores antiguos del PID. En un caso especial, si los valores del seudónimo nuevo coinciden con los valores del seudónimo antiguo, entonces el seudónimo efectivamente no es alterado. De manera alternativa, si un nodo-PID recibe un conjunto {ID} de ID equivalentes desde el nodo-ID (como el nodo-ID **105** en la FIG. 1), donde las ID equivalentes son indexadas por las respectivas DS, entonces el nodo-PID reproduce el correspondiente conjunto indexado actual {PID} en su salida, y envía este conjunto {PID} al nodo nodo-ID.

La hipótesis subyacente es que cada identificador o identidad bajo seudónimo es una representación única de un usuario en el medio determinado (global o local), con probabilidad 1 o con una muy alta probabilidad. En una forma de realización de la presente invención, los PID antiguos y nuevos generados por el nodo-PID se transmiten en forma conjunta con una DS que emite solicitudes, preferentemente en forma encriptada y autenticada, que luego han de ser reenviadas por la DS a los correspondientes nodos-D de entrada junto con los datos de entrada sin procesar relacionados con un usuario anónimo. Para la privacidad, la DS no almacena ningún dato, identidad, identificador, ni el correspondiente seudónimo. Se asume que los nodos-D de entrada comparten con el nodo-PID las correspondientes claves necesarias para la desencriptación y autenticación.

Un nodo-PID puede generar un PID dinámico en forma aleatoria, pseudoaleatoria, o como una función bajo clave de la ID de entrada, donde la clave secreta usada es dinámica; es decir, cambia en el tiempo. La función bajo clave se puede definir en términos de encriptación y funciones de transformación (*hash*) criptográfica. Los seudónimos dinámicos varían en el tiempo; posiblemente, de una manera que dependa de una DS particular. Es posible cambiar los seudónimos periódicamente, en momentos pre-definidos, o de manera consecuente con un evento, ante la aparición de ciertos eventos. En particular, los eventos que desencadenan el cambio del pseudónimo pueden ser definidos por las DS (por ejemplo, inicio del proceso de adquisición de datos tales como sesión IP en Internet o acceso a un determinado sitio web) o por el nodo-PID propiamente dicho (por ejemplo, en términos de los períodos de validez y los tiempos de expiración asociados con los seudónimos generados). Los PID aleatorios y pseudoaleatorios necesitan almacenarse de manera segura en la tabla (ID, PID) del nodo-PID, mientras que los PID bajo clave no necesitan almacenarse puesto que se puede generar un PID nuevo en cualquier momento mediante la utilización de la clave secreta actual, mientras que un conjunto de posibles PID antiguos se pueden reproducir mediante la utilización de un conjunto de posibles claves secretas antiguas que pueden haberse utilizado en el pasado, con respecto a un determinado período de validez. Por consiguiente, en cualquier momento, la tabla (ID, PID) del nodo-PID almacena los últimos PID generados. Debería destacarse que, en la práctica, como el espacio de todas las posibles ID usualmente no es muy grande, almacenar la tabla (ID, PID) en un medio de almacenamiento es factible a través de las tecnologías actuales. Por “almacenamiento seguro” se hace referencia a que las entradas del PID deberían ser almacenadas encriptadas y que la clave de encriptación en la memoria debería ser almacenada en forma segura en el *hardware* o el *software*.

En cualquier momento, el mapeo (ID, PID) debería ser inyectivo sobre un conjunto de posibles ID; es decir, se deberían mapear diferentes ID en diferentes PID. Tanto la función bajo clave como su inversa deberían ser difíciles de computarizar con las tecnologías actuales si la clave es desconocida. En particular, esto implica que la clave debería ser prácticamente impredecible (por ejemplo, una clave de 128 bits). La inyectividad puede garantizarse ya sea teóricamente, con probabilidad 1, o prácticamente, con una muy alta probabilidad. Por ejemplo, la inyectividad se garantiza teóricamente si se elige una función de encriptación para la función bajo clave y se garantiza con una muy alta probabilidad si la función bajo clave es definida como una transformación criptográfica de la salida de una función de encriptación o una transformación criptográfica de una concatenación de la ID de entrada y la clave secreta. Como, en la práctica, el espacio de todas las posibles ID usualmente no es muy grande, la inyectividad se puede chequear en cualquier momento al separar los PID generados para todas las posibles ID. Cada PID aleatorio o pseudoaleatorio nuevo puede entonces generarse repetidamente hasta que se obtenga la inyectividad. De manera alternativa, si los PID se eligen aleatoriamente a partir de un espacio suficientemente grande, entonces la inyectividad puede satisfacerse con una muy alta probabilidad y no necesita chequearse.

Un nodo-PID puede implementar la funcionalidad de asignación de seudónimos dinámicos para una DS individual o para un conjunto de DS (posiblemente todas ellas, como el nodo-PID **110** en la DPN de ejemplo de la **FIG. 1**). En el caso de los seudónimos bajo clave, la clave puede depender de una DS particular. Con respecto a una determinada DS, un nodo-PID puede implementar la funcionalidad de asignación de seudónimos dinámicos para un identificador individual o para un conjunto de identificadores (posiblemente todos ellos). Los valores de los seudónimos, junto con las respectivas ID, correspondientes a un único usuario (para múltiples DS y múltiples identificadores para una única DS) se almacenan conjuntamente en el mismo registro de datos de la tabla (ID, PID) de un determinado nodo-PID. Se asume que un nodo-DS está conectado a todos los nodos-PID que generan los seudónimos necesarios. Tal como se mencionó con anterioridad, sin pérdida de generalidad, la multitud de nodos-PID que sirven una única DPN se pueden considerar como un único nodo-PID lógico, denominado “nodo-PID” en la presente invención.

Si el nodo-PID almacena una tabla (ID, PID), entonces el valor antiguo de un PID es directamente recuperado de la tabla, antes de que el PID nuevo generado sea sustituido por este valor de PID antiguo. Para un PID bajo clave, que no está almacenado en una tabla, el valor antiguo –que ha sido generado mediante la utilización de una clave antigua- no puede ser reproducido si el valor usado de la clave antigua es desconocido. En este caso, el nodo-PID genera un conjunto de PID antiguos candidatos, a partir de la misma ID, mediante la utilización de un conjunto de claves antiguas que pueden haber sido usadas en un período de validez adoptado para el correspondiente perfil de datos, y luego envía este conjunto de PID antiguos candidatos a la DS que emite solicitudes. Por ejemplo, si T denota el período de validez y T_K denota el período de cambio de clave, entonces existen valores de claves antiguos $\lceil T/T_K \rceil$ que pueden haber sido usados en el último período de validez. En forma alternativa, en una solución con almacenamiento, el nodo-PID almacena una tabla (ID, I_K), donde I_K denota el índice de la clave usada para generar el último PID, junto con una tabla segura (mucho más pequeña) (I_K, K) que mapea los índices de las claves en los valores de las claves. La ventaja es que la tabla relativamente grande (ID, I_K) no tiene que ser almacenada en forma segura. En este caso, el PID antiguo puede ser reproducido en cualquier momento al recuperar I_K de la primera tabla y el correspondiente K de la segunda tabla almacenada de manera segura.

Operación de la DPN

Tal como se mencionó con anterioridad, una DPN —como la DPN de ejemplo **100** en la **FIG. 1**— consiste en nodos-D (como los nodos-D **115$_1$** - **115$_8$**), nodos-PID (como el nodo-PID **110**), y nodos-ID (como el nodo-ID **105**), así como también en DS (como las DS **120$_1$** - **120$_4$**) que suministran entrada de datos sin procesar a la DPN, donde las

pluralidades de nodos-ID y nodos-PID pueden considerarse respectivamente como un único nodo-ID lógico y un único nodo-PID lógico, en la presente denominados “nodo-ID” y “nodo-PID”, respectivamente. Tal como se describió con anterioridad, los nodos-D de entrada de primer nivel (como los nodos-D **115_1**, **115_2**, **115_3** en la **FIG. 1**) tienen entradas conectadas a los canales de entrada sólo provenientes de las DS, mientras que otros nodos-D de entrada pueden recibir canales de entrada tanto desde las DS como de otros nodos-D. Los nodos-D están conectados de manera conjunta por canales de datos en una red, que puede estar representada por un gráfico acícliclo o cíclico. Los nodos-D de salida almacenan y proveen datos de salida a los usuarios autorizados, mientras que los nodos-D intermedios temporalmente almacenan y proveen datos de salida solamente a otros nodos-D.

Tal como se esquematiza en la FIG. 3, cada DS (una DS genérica **120** se ilustra en la **FIG. 3**) ocasionalmente envía los datos de entrada sin procesar bajo seudónimos relacionados con un usuario anónimo a los nodos-D de entrada conectados a ésta, por ejemplo, en forma simultánea (en la **FIG. 3**, se ilustra el nodo-D **115** genérico). Más precisamente, el dato de entrada sin procesar $Data_{raw}$ es emitido junto con el valor antiguo y el valor del correspondiente seudónimo de salida, que la DS recibió desde el nodo-PID, es decir, como el triple ($Data_{raw}$, $\{PID_{out}^{old}\}$, PID_{out}^{new}). Todos los seudónimos preferentemente son encriptados y autenticados en forma conjunta; las claves de encriptación y verificación se almacenan de manera segura en los respectivos nodos-D de entrada (por ende, tal como se ilustra en la **FIG. 3**, el triple emitido por el nodo-DS **120** en realidad es ($Data_{raw}$, $Enc(\{PID_{out}^{old}\}, PID_{out}^{new})$). El nodo-DS temporalmente almacena los datos emitidos junto con la ID, como el registro de datos (ID, $Data_{raw}$, $\{PID_{out}^{old}\}$, PID_{out}), que posteriormente se borra después de recibir una confirmación desde el nodo-D de entrada, que confirma que el dato fue recibido y almacenado de manera exitosa.

Si los seudónimos o índices de las claves usadas se almacenan en el nodo-PID, entonces el conjunto de posibles seudónimos antiguos $\{PID_{out}^{old}\}$ se reduce a un único seudónimo antiguo PID_{out}^{old}, que en sí mismo puede ser de valor único o de valores múltiples, según si los identificadores reciben seudónimos por separado o en forma conjunta, respectivamente. El campo del PID_{out}^{old} está vacío si la correspondiente entrada en la tabla (ID, PID) en el nodo-PID no existe; es decir, si el seudónimo PID_{out} nuevo es generado por el nodo-PID por primera vez (porque es la primera vez que, para ese usuario, una DS solicita al nodo-PID la generación de un seudónimo). Si el campo de PID_{out}^{old} no está vacío, entonces –debido a una entrega de datos anterior- existe un registro de datos en cada uno de los nodos-D de entrada conectados que contiene PID_{out}^{old} como la entrada del PID_{in} para esa DS. El registro de datos correcto en cada nodo-D receptor se encuentra entonces como el (único) registro que contiene PID_{out}^{old} (PID^{old} en la FIG. 3) como la respectiva entrada de PID_{in}. Después de encontrar el registro de datos, el seudónimo PID_{in} de entrada encontrado se renueva luego al PID_{out} de valor nuevo recibido (PID^{new} en la **FIG. 3**) y el dato sin procesar $Data_{raw}$ se acumula en el $Data_{in}$ (en la **FIG. 3**, el $Data_{in}^{new}$ se obtiene al acumular el $Data_{in}^{old}$ y el $Data_{raw}$).

Si no existe una tabla (ID, PID) almacenada en el nodo-PID, entonces el conjunto de posibles seudónimos antiguos $\{PID_{out}^{old}\}$ por lo general contiene múltiples seudónimos candidatos, obtenidos a partir de la misma ID mediante la utilización de diferentes claves antiguas, con respecto a un período de validez adoptado. El registro de datos correcto en cada nodo-D receptor –si existe- es luego encontrado como el (único) registro que contiene uno de los seudónimos antiguos candidatos de $\{PID_{out}^{old}\}$ como la entrada de PID_{in} para esa DS. Una vez encontrado el registro de datos correcto —si existe— el seudónimo de entrada PID_{in} encontrado se renuevan entonces al PID_{out} de valor nuevo recibido y el $Data_{raw}$ se acumula en el $Data_{in}$.

Cada DS se puede implementar como una familia de DS constitutivas, todas conectadas al nodo-PID, donde cada DS constitutiva emite a los nodos-D sólo una parte del dato sin procesar $Data_{raw}$ y, posiblemente, sólo una parte del PID_{out} de valores múltiples, que determina de manera única un usuario anónimo, y puede o no ser actualizada a un valor nuevo. El registro de datos correcto se encuentra luego mediante la utilización de sólo la parte del PID_{out} de valores múltiples para la respectiva familia de DS. Las DS constituyentes pueden ser lógica o físicamente separadas.

En una forma de realización, una DS puede emitir a los nodos-D conectados a ella sólo un seudónimo de salida PID_{out} nuevo actualizado a un valor nuevo, sin estar acompañado por datos sin procesar $Data_{raw}$ (para un seudónimo PID_{out} de valores múltiples, solamente se pueden actualizar algunos de los valores de los seudónimos constitutivos). El seudónimo es alterado ya sea porque uno de los identificadores temporales (por ejemplo, un localizador tal como una dirección IP) ha cambiado, o porque es necesario cambiar el seudónimo correspondiente a uno de los identificadores permanentes a fin de evitar que el mismo seudónimo sea usado en períodos prolongados (por ejemplo, esto puede ocurrir cuando se emite el dato sin procesar $Data_{raw}$ en alto volumen).

En otra forma de realización, el dato sin procesar $Data_{raw}$ es emitido junto con sólo una parte de PID_{out}, que en sí misma no es actualizada a un valor nuevo, pero es suficiente para identificar un usuario. Por ejemplo, un identificador temporal, tal como un localizador (por ejemplo, una dirección IP en Internet) se puede usar para emitir los datos sin procesar, pero no es necesario cambiar el correspondiente seudónimo cada vez que los datos sin procesar se emiten siempre que el identificador temporal permanezca igual.

En el caso en que, en un nodo-D, el registro de datos buscado —donde el seudónimo PID_{in} almacenado coincide con el seudónimo recibido PID_{out}^{old} o $\{PID_{out}^{old}\}$ recibido desde una DS emisora— no se encuentra, esto significa que dicha DS está emitiendo los datos sin procesar relacionados con el respectivo usuario por primera vez. No

obstante, es posible que el registro de ese mismo usuario ya exista, debido al hecho de que los datos sin procesar relacionados con ese usuario han sido enviados al nodo-D considerado en el pasado desde otras DS o los datos de salida relacionados con ese usuario han sido enviados al nodo-D considerado en el pasado desde otros nodos-D (este último caso se aplica a un nodo-D de entrada que no es un nodo-D de entrada de primer nivel; es decir, un nodo-D de entrada que también puede recibir datos de entrada desde otro nodo-D). Esto puede ocurrir ya sea si el campo de PID_{out}^{old} está vacío (en el caso de los seudónimos almacenados) o si no está vacío (en el caso de los seudónimos bajo clave que no están almacenados). En cualquier caso, inicialmente se intenta encontrar el registro de datos buscado mediante la utilización de los seudónimos actuales candidatos correspondientes a los identificadores temporales (por ejemplo, un localizador) que pueden ser compartidos por varias DS u otros nodos-D, bajo la hipótesis de que los correspondientes PID también son iguales. A saber, si el mismo localizador (por ejemplo, una dirección IP) se usa simultáneamente para diferentes DS (por ejemplo, dos sitios web en Internet), entonces el registro posiblemente puede ser encontrado mediante la utilización del mismo seudónimo correspondiente para una cualquiera entre las dos DS.

Si este chequeo inicial de identificadores temporales comunes no arroja un registro de datos, entonces se intenta que el registro de datos buscado se encuentre al invocar un procedimiento que en lo sucesivo se denominará “protocolo de correlación-ID”, que comprende la respectiva DS y el nodo-ID, mediante la utilización de los seudónimos actuales candidatos enviados por otras DS o por otros nodos-D en el pasado. En consecuencia, el protocolo de correlación-ID se lleva a cabo cada vez que no se encuentra el registro de datos buscado en el respectivo nodo-D de entrada, después de recibir los datos de entrada sin procesar bajo seudónimos desde una DS, mediante la utilización del (los) seudónimo(s) antiguo(s) candidato(s) enviado(s) por la DS junto con los datos de entrada sin procesar, siempre que existan otras entradas al nodo-D de entrada y que el chequeo inicial mencionado con anterioridad de identificadores temporales comunes no logre encontrar un registro de datos que corresponda al usuario considerado. En particular, ni el chequeo inicial ni el protocolo de correlación-ID se efectúan si existe solamente una entrada en el nodo-D de entrada, proveniente de la DS considerada. Antes de realizar el protocolo de correlación-ID, el nodo-D de entrada almacena el triple recibido ($Data_{raw}$, $\{PID_{out}^{old}\}$, PID_{out}). Se asume que a cada nodo-D de entrada en la DPN se asigna un único identificador estático (por ejemplo, un índice del nodo-D).

El protocolo de correlación-ID es iniciado por el nodo-D de entrada receptor, que genera, almacena y envía de regreso a la DS emisora un mensaje que inicia el protocolo de correlación-ID. El mensaje de inicio contiene un identificador de mensaje que comprende un número usado una sola vez o nonce (*number used once*) localmente generado por el nodo-D de entrada (por ejemplo, un sello con la hora o un número de serie) y el identificador (por ejemplo, el índice del nodo-D) del nodo-D de entrada, que debería ser único en la DPN. Este mensaje de inicio también contiene el (los) seudónimo(s) antiguo(s) candidato(s) recibido(s) PID_{out}^{old} o $\{PID_{out}^{old}\}$. Una vez recibido este mensaje desde el nodo-D de entrada, la DS encuentra la correspondiente ID desde el registro de datos temporalmente almacenados (mediante la utilización del los seudónimos antiguos candidatos) y lo envía al nodo-ID solicitando los PID actuales candidatos correspondientes a las ID equivalentes del usuario, con respecto a todas las DS en el DPN, a ser reproducidas. El nodo-ID genera entonces el correspondiente conjunto $\{(i, ID)\}$ de ID equivalentes del usuario, indexadas por las respectivas DS, y envía este conjunto al nodo-PID. Debería destacarse que, si el nodo-ID almacena solamente los identificadores permanentes, entonces cada ID generada –aunque consista solamente en identificadores permanentes- también determina el usuario en forma exclusiva. Sobre la base del conjunto recibido $\{(i, ID)\}$, el nodo-PID reproduce el correspondiente conjunto de seudónimos actuales candidatos $\{(i, \{PID_{out}^{old}\})\}$ en su salida y los envía –preferentemente encriptados y autenticados- a las respectivas DS junto con el identificador de mensaje de inicio. En el caso de los seudónimos almacenados, este conjunto de seudónimos actuales candidatos $\{(i, \{PID_{out}^{old}\})\}$ también puede incluir los seudónimos ya almacenados para los identificadores temporales, que han sido recibidos directamente desde las DS en el pasado. En caso de que el nodo-PID no almacene los seudónimos, este conjunto solamente contiene los seudónimos para los identificadores permanentes recibidos desde el nodo-ID. El protocolo de correlación-ID entonces continúa mediante una fase de inundación de reenvíos, descripta con posterioridad.

Cada DS, excepto la DS emisora que envío los datos de entrada sin procesar al nodo-D que inicia el protocolo de correlación-ID, reenvía el conjunto encriptado y autenticado recibido de seudónimos actuales candidatos a los nodos-D de entrada coexistentes en la DPN, a la cual está conectada, junto con el identificador de mensaje de inicio recibido, sin ningún dato de entrada sin procesar. Cada nodo-D de entrada coexistente luego desencripta y autentica los seudónimos actuales candidatos recibidos. Ahora, si uno de los nodos-D de entrada es el nodo-D que inició el protocolo de correlación-ID (es decir, el nodo-D con el mismo identificador de nodo-D que el contenido en el identificador de mensaje de inicio), entonces este nodo-D de entrada almacena el mensaje recibido siempre que el identificador de mensaje de inicio en el mensaje recibido coincida con uno de los identificadores de mensaje de inicio almacenados; de lo contrario, borra el mensaje si no existe coincidencia (debido a las fallas). De lo contrario, si el identificador de mensaje de inicio recibido contiene un identificador de nodo-D diferente, entonces el nodo-D de entrada procede a reenviar un mensaje modificado a todos los subsiguientes nodos-D coexistentes en la DPN, a través de cada una de sus salidas, en donde la modificación consiste en reemplazar los seudónimos actuales candidatos recibidos por los correspondientes seudónimos de salida actuales, siempre que se encuentre un registro de datos. Esto se alcanza cuando se encuentra el registro de datos –si existe- mediante la utilización de cada seudónimo actual candidato recibido como seudónimo de entrada y al extraer el correspondiente seudónimo de salida actual de la respectiva salida. Cada nodo-D en la DPN (con un identificador de nodo-D diferente) luego

procede de la misma manera al reenviar el mensaje que contiene el mismo identificador de mensaje y el seudónimo actual sustituido a los nodos-D coexistentes subsiguientes en la DPN. En el caso de los nodos-D cíclicamente conectados, un mensaje con el mismo identificador de mensaje de inicio no es reenviado por el mismo nodo-D más de dos veces.

En consecuencia, el nodo-D de entrada que inició el protocolo de correlación-ID recibirá y almacenará, dentro de un marco de tiempo adoptado, un conjunto de mensajes con el mismo identificador de mensaje de inicio, que coincide con uno de los identificadores de mensaje almacenados, que contiene los seudónimos actuales candidatos para todas las entradas conectadas a las DS u otros nodos-D. El registro de datos buscado luego se encuentra —si existe— mediante la utilización de este conjunto de seudónimos actuales candidatos y buscar una coincidencia con un seudónimo de entrada para la respectiva entrada. Si el registro de datos buscado no se encuentra, entonces el registro directamente no existe y el nodo-D de entrada crea entonces un registro de datos nuevo mediante la utilización del seudónimo PID_{out} nuevo almacenado para la respectiva DS como el correspondiente seudónimo de entrada.

En una situación más general, si un nodo-D receptor (en particular, un nodo-D de entrada) está recibiendo datos de entrada bajo seudónimos desde otro nodo-D en lugar de otra DS, y si, en el nodo-D receptor, el registro de datos buscado no se encuentra, entonces se usa un procedimiento que es una versión modificada del protocolo de correlación-ID, en lo sucesivo denominado el “protocolo de correlación-PID”. A saber, en el protocolo de correlación-ID, el nodo de la DS que emite los datos es capaz de emitir la correspondiente ID (es decir, la ID correspondiente al seudónimo que es enviada al nodo-D junto con los datos) al nodo-ID, debido a que la DS temporalmente almacena un registro de datos que contiene la ID. No obstante, un nodo-D que envía los datos no conoce la ID del usuario para cualquiera de las DS, sino solamente los seudónimos de entrada y de salida actuales y un valor nuevo generado del seudónimo de salida, donde el valor antiguo y el valor nuevo del seudónimo de salida son enviados al nodo-D considerado, junto con los datos de salida.

El protocolo de correlación-PID se inicia si el registro de datos buscado en un nodo-D receptor que recibe datos (perfil de datos relacionados con cierto usuario) desde otro nodo-D emisor y que tiene al menos otra entrada no se encuentra mediante la utilización del seudónimo antiguo enviado por dicho nodo-D emisor, siempre que ese chequeo inicial de identificadores temporales comunes mencionado con anterioridad no logre encontrar un registro de datos. El objetivo del protocolo de correlación-PID es encontrar el registro de datos buscado mediante la utilización de los seudónimos actuales enviados por otros nodos-D emisores o los seudónimos actuales candidatos enviados por las DS al nodo-D receptor bajo consideración. El protocolo de correlación-PID consiste en las fases de avance y retroceso descriptas en relación con el protocolo de correlación-ID, donde la fase de avance consiste en la inundación de reenvíos como en el protocolo de correlación-ID. No obstante, la fase de retroceso es más compleja que en el protocolo de correlación-ID, porque el mensaje que inicia el protocolo de correlación-PID no se envía directamente a una DS que envía los datos de entrada sin procesar y que temporalmente almacena la correspondiente ID del usuario. En cambio, el mensaje de inicio del protocolo de correlación-ID se envía a un nodo-D que emite los datos de salida al nodo-D receptor, y el nodo-D emisor no tiene conocimiento de la correspondiente ID del usuario.

La fase de retroceso del protocolo de correlación-PID consiste en el retroseguimiento del nodo-D receptor que recibió los datos de salida junto con el valor antiguo y el valor nuevo del seudónimo de salida desde otro nodo-D emisor, en la forma ($Data_{out}$, PID_{out}^{old}, PID_{out}). La fase de retroceso empieza cuando el nodo-D receptor emite el mensaje que inicia el protocolo de correlación-PID a este nodo-D emisor, donde el mensaje contiene PID_{out}^{old}. Este nodo-D emisor puede entonces reenviar el mensaje hacia atrás, a través de al menos uno de sus canales de entrada (por ejemplo, todos ellos) al suministrar los datos de entrada que contribuyen con su $Data_{out}$ asociado con PID_{out}^{old}, donde, para cada una de las entradas que participan, el seudónimo de entrada PID_{in} actual correspondiente a esa entrada está contenido en el mensaje enviado de regreso a través de esa entrada al correspondiente nodo-D emisor anterior en la DPN. Como el PID_{in} enviado coincide con el seudónimo de salida PID_{out} del nodo-D anterior, este nodo-D puede encontrar el respectivo registro de datos correspondientes al PID_{out} y recuperar los correspondientes seudónimos de entrada PID_{in} para las entradas existentes. El proceso de retroseguimiento, por ende, continúa hacia atrás, desde cualquier nodo-D alcanzado hasta al menos un nodo-D anterior conectado a ese nodo-D alcanzado, en donde cada vez se sustituye el correspondiente seudónimo de entrada actual del nodo-D actual para el seudónimo de entrada recibido en el mensaje hasta que, por último, el mensaje alcanza al menos una DS (por ejemplo, todas ellas) por lo que contribuye con el dato de salida $Data_{out}$ recibido por el nodo-D que inicia el protocolo de correlación-PID. Por lo tanto, este mensaje contiene el correspondiente seudónimo de entrada PID_{in} como el último seudónimo de salida PID_{out} emitido por la DS y generado por el nodo-PID (para el respectivo usuario). Este proceso de retroseguimiento es, por ende, esencialmente permitido por las cadenas existentes de seudónimos almacenados (tal como se esquematiza en la FIG. 5), donde, de acuerdo con el método propuesto en la presente invención, en cada momento, el seudónimo de entrada actual de cualquier nodo-D en la DPN es igual al seudónimo de salida actual del nodo-D anterior conectado a este nodo-D.

La DS alcanzada no puede reconstruir la ID desde el seudónimo PID_{in} recibido, que es igual al último seudónimo PID_{out} emitido por esa DS, debido a que cualquier DS almacena sólo temporalmente la ID y la borra después de recibir la confirmación de un nodo-D receptor conectado a ésta, pero la DS alcanzada puede reenviar el mensaje al

nodo-PID. Alternativamente, el proceso de retroseguimiento puede detenerse cuando el mensaje alcanza al menos un nodo-D de entrada, que recibe al menos una entrada desde una DS. Este nodo-D de entrada luego produce el mismo mensaje y lo envía directamente al nodo-PID, en lugar de reenviarlo a través de la respectiva DS. Ahora, como en el protocolo de correlación-ID, el nodo-PID debería reproducir el correspondiente conjunto de seudónimos actuales candidatos equivalentes $\{(i, \{PID_{out}^{old}\})\}$ en su salida y enviarlos -encriptados y autenticados- a las respectivas DS, junto con el identificador de mensaje, a fin de iniciar la fase de inundación de reenvíos del protocolo de correlación-PID, que es la misma que en el protocolo de correlación-ID. Como esto no se puede realizar a través del uso del conjunto de ID equivalentes producidas por el nodo-ID como en el protocolo de correlación-ID, el nodo-PID debería reproducir directamente el conjunto de todos los PID equivalentes a los PID_{in} recibidos, lo que es válido para la DS desde la cual se recibieron los PID_{in}. En principio, existen dos soluciones al problema según si el nodo-PID almacena la tabla (ID, PID) o no.

En caso de que el nodo-PID implemente un almacenamiento de los PID, el nodo-PID busca a través de la tabla almacenada mediante el PID_{in} como clave de búsqueda, y luego recupera el conjunto $\{(i, PID_{out}^{old})\}$ desde el registro encontrado. Este conjunto también puede contener los seudónimos almacenados para los identificadores temporales. En el caso de los seudónimos bajo clave sin almacenamiento, la función bajo clave usada debería cumplir con el requisito adicional de que su inversa debería ser fácil de computarizar cuando la clave es conocida (por ejemplo, puede definirse como una función de encriptación). Por ende, la correspondiente ID puede ser reconstruida por el nodo-PID a partir del PID_{in} y enviada al nodo-ID en forma encriptada y autenticada. Como ocurre en el protocolo de correlación-ID, el nodo-ID genera entonces el correspondiente conjunto $\{(i, ID)\}$ de ID equivalentes indexadas por las respectivas DS y envía este conjunto generado de regreso al nodo-PID en forma encriptada y autenticada. El nodo-PID luego reproduce el correspondiente conjunto de seudónimos actuales candidatos $\{(i, \{PID_{out}^{old}\})\}$ en su salida, mediante la utilización de las claves antiguas candidatas usadas en el pasado, como en el protocolo de correlación-ID. Este conjunto sólo contiene los seudónimos para los identificadores permanentes recibidos desde el nodo-ID. En ambos casos, el nodo-PID envía entonces el conjunto recuperado o reproducido de seudónimos -encriptados y autenticados- a las respectivas DS junto con el identificador de mensaje, para iniciar la fase de inundación de reenvíos como en el protocolo de correlación-ID.

En una forma de realización de la presente invención, a fin de controlar la capacidad de formar enlaces temporal de los perfiles de datos, la ley de privacidad, implementada por autoridades legales, debería especificar los tiempos máximos de capacidad de formar enlaces para varios perfiles de datos de usuarios de interés. Más precisamente, la ley de privacidad debería especificar los tiempos máximos permitidos para mantener los seudónimos dinámicos inalterados, para varios perfiles de datos, donde estos tiempos pueden depender del período de validez de los perfiles de datos particulares. Como estos tiempos pueden ser efectivamente prolongados al memorizar los valores anteriores de los seudónimos dinámicos, la ley de privacidad debería prohibir que cualquier entidad que lidia con datos sin procesar o perfiles de datos en una DPN (es decir, una DS o un nodo-D) pueda asignar, almacenar y usar cualesquiera seudónimos estáticos asociados con los registros de datos de los usuarios individuales o almacenar los valores antiguos de los seudónimos dinámicos generados y usados en el pasado. En particular, para evitar que las direcciones físicas en la memoria sean efectivamente usadas como seudónimos estáticos, debería actualizarse un registro de datos en los nodos-D (especialmente, en los nodos-D de salida) al borrar el registro anterior y almacenar el contenido actualizado en un nuevo registro.

Acceso a perfiles de datos in DPN

Cada nodo-D de salida en una DPN almacena los perfiles de datos de usuarios anónimos individuales como los datos de salida, donde los perfiles de datos almacenados están marcados por los seudónimos dinámicos de valor único o de valores múltiples correspondientes a las entradas y salidas individuales del nodo-D. En cualquier momento, el valor actual de cada seudónimo de entrada es el último seudónimo recibido desde un nodo-D anterior o, directamente, desde una DS. Los seudónimos recibidos desde una DS pueden ser PID de valor único, correspondientes a la identidad de usuario ID en su totalidad, o PID de valores múltiples, correspondientes a los subconjuntos de identificadores individuales que comprenden la identidad. En cualquier momento, el valor actual de cada seudónimo de salida es el último seudónimo enviado a nodos-D subsiguientes, que se generaron como un seudónimo de valor único a partir de los seudónimos de entrada al momento de la última emisión. Los seudónimos de salida se generan y se usan solamente para los datos de salida que son emitidos a otros nodos-D. Los datos de salida se actualizan al procesar datos de entrada acumulados en determinados tiempos de actualización.

Sobre la base de los perfiles de datos de usuarios anónimos individuales, cada nodo-D de salida también puede almacenar cualquier dato estadístico computarizado a partir de los perfiles de datos individuales (por ejemplo, mediante técnicas de cálculo del promedio o conteo) y relacionado con ciertos subconjuntos de usuarios, posiblemente todos ellos. En una forma de realización de la presente invención, en cada nodo-D de salida, los perfiles de datos y los datos estadísticos se almacenan en forma encriptada que puede ser desencriptada sólo por los usuarios autorizados.

En cualquier momento, los usuarios autorizados pueden tener acceso a los perfiles de datos de salida y datos estadísticos almacenados en los nodos-D de salida, donde la autorización se puede definir mediante reglas de control de acceso y deberían cumplir con la ley de privacidad. Tanto para los perfiles de datos individuales como para los datos estadísticos, se debería especificar el nodo-D de salida y la salida específica de ese nodo-D de

salida que almacena los perfiles de datos deseados. Para tener acceso a los datos estadísticos, es suficiente tener la autorización necesaria aceptada por el nodo-D de salida. Para tener acceso a los perfiles de datos individuales de los usuarios anónimos, además de la autorización, se deberían especificar los perfiles de datos solicitados mediante la utilización de los respectivos seudónimos (dinámicos) de entrada y/o salida de manera significativa a nivel práctico. Los seudónimos se pueden especificar ya sea directamente o, en forma alternativa, mediante los identificadores de usuarios de acuerdo con la ley de privacidad, en cuyo caso el nodo-PID necesita participar para reproducir los seudónimos sobre la base de los identificadores de usuarios especificados. Los seudónimos propiamente dichos pueden relacionarse con el nodo de salida especificado o con cualquier otro nodo-D en la DPN tal como, por ejemplo, cualquier nodo-D de entrada. En el último caso, el nodo-D de salida especificado puede ser alcanzado por el proceso de inundación de reenvíos descripto con anterioridad. Algunos casos de ejemplo de la especificación de perfiles de datos individuales a los cuales ha de accederse se describen con posterioridad. Debería destacarse que los casos no son necesariamente inconexos.

En un primer caso de ejemplo, cualquier seudónimo de entrada actual de un nodo-D de entrada elegido que ha sido recibido directamente desde una DS en el pasado se usa para especificar el perfil de datos a los cuales ha de accederse. Dicho seudónimo de entrada puede ser un seudónimo de valor único correspondiente a una DS individual o cualquier componente o subconjunto de componentes de un seudónimo de valores múltiples correspondiente a una DS. Más generalmente, el usuario autorizado también puede usar cualquier subconjunto de dichos seudónimos.

A fin de tener acceso al perfil de datos solicitado, el usuario autorizado envía el seudónimo elegido al nodo-D de entrada elegido junto con el identificador de un nodo-D de salida que contiene el perfil de datos deseado a ser recuperado. El nodo-D de entrada luego inicia el proceso de inundación de reenvíos con el mensaje que contiene el identificador de nodo-D de salida. El nodo-D de salida correspondiente a dicho identificador luego encuentra el perfil de datos solicitado mediante la utilización del último seudónimo de entrada correspondiente recibido desde un nodo-D anterior o directamente desde una DS, al final de la cadena subyacente de seudónimos dinámicos (tal como se esquematiza en la **FIG. 5**).

En un segundo caso de ejemplo, el seudónimo de entrada usado en el primer caso de ejemplo para lograr acceso a un perfil de datos deseado se obtiene mediante el nodo-PID desde un identificador de usuario temporal especificado, de acuerdo con la ley de privacidad. En particular, el identificador de usuario temporal especificado puede definir la localización actual de un usuario anónimo destinado (por ejemplo, la dirección IP actual en Internet). En este caso, el perfil de datos recuperado del usuario se puede usar para emitir información comercial (por ejemplo, anuncios publicitarios personalizados) al usuario en una determinada ubicación.

En un tercer caso de ejemplo, el seudónimo de entrada usado del primer caso de ejemplo se obtiene mediante el nodo-PID desde un identificador de usuario permanente especificado, de acuerdo con la ley de privacidad. La ley de privacidad debería especificar los identificadores, los perfiles de datos, los usuarios autorizados, y las condiciones relacionadas para tener acceso y usar los perfiles de datos de usuarios. En particular, si el identificador de usuario permanente especificado identifica el usuario en forma exclusiva (global o localmente), entonces la ley de privacidad debería especificar la granularidad mínima de los datos como el número mínimo de usuarios por perfil de datos solicitado para que el acceso sea permitido.

En un cuarto caso de ejemplo, el perfil de datos al cual ha de accederse es directamente especificado por un seudónimo de entrada del nodo-D de salida elegido que almacena el perfil de datos deseado. Dicho seudónimo de entrada puede ser recibido ya sea desde otro nodo-D o desde un DS, como en el primer caso de ejemplo. En particular, el nodo-D de salida elegido puede ser un nodo-D de usuario que almacena un perfil de datos de un usuario particular en el equipo del usuario, donde el perfil de datos junto con el correspondiente seudónimo de entrada ha sido recibido desde un nodo-D de base de datos de salida, que almacena los perfiles de datos de todos los usuarios, durante la última actualización (aperiódica o periódica) de los perfiles de datos para ese usuario. En este caso, un usuario autorizado puede recuperar directamente el perfil de datos (antiguos) actualmente almacenado desde el equipo de usuario en una determinada ubicación –en cuyo caso, el seudónimo de entrada no se utiliza- o bien puede tener acceso al perfil de datos del usuario más reciente en el nodo-D de base de datos mediante la utilización del seudónimo de entrada desde el equipo del usuario como el seudónimo de salida en el nodo-D de base de datos. El perfil de datos del usuario más reciente recuperado, junto con el seudónimo de entrada recibido más reciente, se almacenan luego en el equipo del usuario y se pueden usar para emitir información comercial (por ejemplo, anuncios publicitarios personalizados) al usuario en una determinada ubicación.

Funcionalidad de DPN combinada

Una DPN combinada es un sistema compuesto por un número (dos o más) de DPN inconexas. Cada DPN funciona por separado, pero no es necesario que los correspondientes conjuntos de usuarios estén inconexos. Las DPN individuales pueden entonces compartir los usuarios en común. El principal objetivo de una DPN combinada de acuerdo con la presente invención es permitir el acceso conexo a los perfiles de datos de un mismo usuario en diferentes DPN, sin cambiar las DPN individuales. Se asume que los perfiles de datos se almacenan en los nodos-D de salida, que son especificados por sus identificadores, únicos para las respectivas DPN. La combinación

deseada de las DPN se pueden alcanzar al introducir un nodo-ID combinado formado al fusionar, es decir, acumular los nodos-ID de las DPNs individuales. Por consiguiente, el nodo-ID combinado implementa una tabla combinada que almacena todas las identidades equivalentes del mismo usuario, para diferentes DPN y para diferentes fuentes de datos de cada DPN, donde a cada DPN se asigna un índice diferente para su distinción.

En una forma de realización, un usuario autorizado que accede a un perfil de datos en una salida especificada (en lo sucesivo, denominada "Out_1") de un nodo-D de salida especificado (en lo sucesivo, denominado "D_1") en cierta DPN en un determinado momento desea acceder al perfil de datos del mismo usuario –si existe- en una salida especificada (en lo sucesivo, denominada "Out_2") de un nodo-D de salida especificado (en lo sucesivo, denominado "D_2") de otra DPN, que en sí misma está especificada por el correspondiente índice. Las dos DPN, en lo sucesivo, se denominan "DPN_1" y "DPN_2", respectivamente. El usuario autorizado especifica el perfil de datos solicitado en D_1/Out_1 de DPN_1 mediante un seudónimo o mediante un identificador de usuario, tal como se explicó con anterioridad. A fin de recuperar el perfil de datos solicitado en DPN_2, es necesario recuperar los seudónimos de entrada o salida actuales correspondientes al mismo usuario en D_2 de DPN_2.

Si el perfil de datos solicitado en DPN_1 es especificado por un identificador de usuario, entonces los correspondientes seudónimos en DPN_2 se pueden recuperar mediante una variante del protocolo de correlación-ID, descripto con posterioridad. El protocolo es iniciado por el nodo-D D_1 que almacena y envía un mensaje al nodo-ID combinado donde el identificador de mensaje consiste en un nonce localmente generado (por ejemplo, un sello con la hora o un número de serie), el identificador de D_1/Out_1 en DPN_1, el identificador de D_2/Out_2 en DPN_2, y los índices de DPN_1 y DPN_2. El mensaje contiene un identificador de usuario en DPN_1. Junto con el mensaje, el nodo-D D_1 también almacena localmente el perfil de datos recuperado en D_1/Out_1 y un identificador del usuario autorizado. El nodo-ID combinado recupera entonces desde el identificador de usuario recibido todas las identidades equivalentes del mismo usuario en DPN_2 y las envía al nodo-PID en DPN_2. Como en el protocolo de correlación-ID, el nodo-PID recupera el (los) seudónimo(s) (candidatos) correspondiente(s) a las identidades de usuarios equivalentes para todas las DS en DPN_2 y luego envía el mensaje con el mismo identificador de mensaje a cada DS en DPN_2 al sustituir los) seudónimo(s) (candidatos) recuperado(s) para el identificador de usuario originalmente enviado en el contenido del mensaje.

Al reenviar el mensaje recibido a los nodos-D coexistentes en DPN_2, cada DS inicia entonces la fase de inundación de reenvíos del protocolo con el objetivo de enviar el identificador de mensaje, junto con el seudónimo de entrada actual para el mismo usuario, al nodo D_2 de salida de destino. Por consiguiente, en forma similar a lo que sucede con el protocolo de correlación-ID descripto anteriormente, cada nodo-D distinto de D_2 procede a reenviar el mensaje con el contenido modificado en el cual el o los seudónimos de entrada (candidatos) son reemplazados por los correspondientes seudónimos de salida en los subsiguientes nodos-D coexistentes en la DPN_2. Cuando el mensaje alcanza el D_2 a través de una de las entradas, el D_2 encuentra el registro de datos –si existe- mediante la utilización del seudónimo de entrada recibido o los seudónimos de entrada candidatos, y extrae el perfil de datos de la salida especificada Out_2, sustituye este perfil por el(los) seudónimo(s) recibido(s) en el contenido del mensaje, y luego envía este mensaje al D_1 (por ejemplo, directamente, mediante la utilización de una red de comunicación común, tal como Internet). El D_1 recupera el perfil de datos previamente almacenado de D_1/Out_1 mediante la utilización del identificador de mensaje recibido, adjunta el perfil de datos recibido desde D_2/Out_2, y luego envía el perfil de datos conexo al usuario autorizado que emitió la solicitud. En forma alternativa, en lugar de la inundación de reenvíos, los nodos-D en la DPN_2 pueden usar las tablas de encaminamiento (con localizadores apropiados) para reenviar el mensaje solamente a uno de los subsiguientes nodos-D inconexos, en lugar de todos ellos, como los protocolos de encaminamiento comunes en las redes de comunicaciones estándares.

Si el perfil de datos solicitado en la DPN_1 es especificado por un seudónimo, entonces los correspondientes seudónimos en la DPN_2 pueden ser recuperados mediante una variante del protocolo de correlación-PID descripto con posterioridad. La fase de retroceso del protocolo consiste en un retroseguimiento, que comienza desde la D_1 y que almacena y envía un mensaje de inicio de regreso a través de al menos un canal de entrada (por ejemplo, todos ellos) que provee los datos de entrada que contribuyen con el perfil de datos solicitado en D_1/Out_1. Tal como se mencionó anteriormente, el identificador de mensaje consiste en un nonce localmente generado (por ejemplo, un sello con la hora o un número de serie), el identificador de D_1/Out_1 en DPN_1, el identificador de D_2/Out_2 en DPN_2, y los índices de DPN_1 y DPN_2. Junto con el mensaje, el D_1 también almacena localmente el perfil de datos recuperado en D_1/Out_1 y un identificador del usuario autorizado. La diferencia es que el mensaje ahora contiene el seudónimo de entrada correspondiente a una entrada elegida en lugar de un identificador de usuario para especificar un perfil de datos en D_1/Out_1. Debería destacarse que todos los seudónimos de entrada se pueden recuperar a partir de cualquier entrada o seudónimo de salida que especifica de manera única el perfil de datos en D_1/Out_1.

Como en el protocolo de correlación-PID, en el proceso de retroseguimiento, cualquier nodo-D anterior alcanzado reenvía el mensaje hacia atrás, a al menos un nodo-D anterior conectado a ese nodo-D, sustituyendo cada vez el correspondiente seudónimo de entrada actual para el seudónimo de entrada recibido en el mensaje hasta que, por último, el mensaje alcanza al menos un nodo-D de entrada, que recibe al menos una entrada desde una DS en DPN_1. Este nodo-D de entrada luego recupera del registro de datos encontrado el correspondiente seudónimo de entrada PID_{in} recibido desde esa DS como el último seudónimo de salida PID_{out} emitido por la DS y generado por el nodo-PID (para el respectivo usuario). Al final de la fase de retroceso, cualquier nodo-D de entrada de este tipo envía un mensaje con el mismo identificador de mensaje que contiene un seudónimo (PID_{out}) al nodo-PID en la

DPN_1, ya sea directamente o a través de la respectiva DS. Ahora, como en el protocolo de correlación-PID, el nodo-PID recupera (en el caso con almacenamiento) o reproduce (en el caso sin almacenamiento) la correspondiente ID y la envía al nodo-ID combinado, que luego recupera desde la ID recibida todas las identidades equivalentes del mismo usuario en la DPN_2 y las envía al nodo-PID en la DPN_2. La fase de avance del protocolo en la DPN_2 es entonces igual a la descripta con anterioridad.

En otra forma de realización, es posible que un nodo-D de salida de la DPN_1, en un determinado momento, desee fusionar un subconjunto de perfiles de datos (por ejemplo, todos ellos) almacenados en una o más de sus salidas con el subconjunto de perfiles de datos correspondiente a los mismos usuarios almacenados en una o más salidas de un nodo-D de salida especificado de la DPN_2. El subconjunto de perfiles de datos en la DPN_1 es especificado por seudónimos. Luego, cada seudónimo correspondiente en la DPN_2 puede ser recuperado individualmente por la variante del protocolo de correlación-PID descripto con anterioridad. A tal fin, cada identificador de mensaje creado debería contener un nonce diferente. Como los nonces efectivamente juegan el papel de seudónimos nuevos creados para fusionar o intercambiar los perfiles de datos de los nodos-D en la DPN_1 y la DPN_2, pueden introducir capacidad de formar enlaces temporal indeseada de los perfiles de datos intercambiados, especialmente si los perfiles de datos de los mismos subconjuntos de usuarios (anónimos) son repetidamente fusionados, por ejemplo, si los perfiles de datos de todos los usuarios son fusionados. En consecuencia, los nonces deberían ser generados, preferentemente, en forma aleatoria o pseudoaleatoria como seudónimos usados una sola vez.

En otra forma de realización más, en lugar de recuperar los seudónimos y los perfiles de datos individualmente, en forma sucesiva, es más simple hacerlo de manera conjunta mediante la utilización de una lista de seudónimos y nonces en su totalidad. Más precisamente, es posible que el PID_1 y el PID_2 denoten seudónimos genéricos en D_1/Out_1 de la DPN_1 y D_2/Out_2 de la DPN_2, respectivamente, y que el $\{PID_1'\}$ y el $\{PID_2'\}$ denoten los correspondientes subconjuntos de seudónimos en el origen de las cadenas de seudónimos en la DPN_1 y la DPN_2 que terminan con el PID_1 y el PID_2, respectivamente. Es posible que el PID′ denote un nonce genérico como un seudónimo usado una sola vez para fusionar los perfiles de datos. El D_1 prepara entonces la forma inicial de un mensaje conexo con el identificador de mensaje que comprendel identificador de D_1/Out_1 en la DPN_1, el identificador de D_2/Out_2 en la DPN_2, y los índices de DPN_1 y DPN_2. El mensaje contiene una lista {(PID_1, PID′)}, donde cada PID_1 especifica de manera única un perfil de datos, $Data_{out,1}$, en D_1/Out_1 de la DPN_1, y PID' es un seudónimo generado una sola vez en forma aleatoria o pseudoaleatoria. Entonces, el D_1 prepara la forma final del mensaje que contiene la lista {($\{PID_1'\}$, PID′)}, que se obtiene al sustituir $\{PID_1'\}$ por PID_1, para cada PID_1, donde $\{PID_1'\}$ se obtiene a través del proceso de retroseguimiento, ya sea desde las DS o directamente desde los correspondientes nodos-D de entrada en la DPN_1.

El D_1 envía el mensaje con la lista {($\{PID_1'\}$, PID′)} al nodo-PID en la DPN_1. Por cada PID_1', el nodo-PID recupera (en el caso con almacenamiento) o reproduce (en el caso sin almacenamiento) la correspondiente ID y la envía al nodo-ID combinado, que luego recupera del ID recibido todas las identidades equivalentes del mismo usuario en la DPN_2 y las envía al nodo-PID en el DPN_2. El nodo-PID luego genera el correspondiente subconjunto $\{PID_2'\}$ e inicia el proceso de inundación de reenvíos en la DPN_2 que por último genera el seudónimo de entrada PID_2 recibido por el D_2 junto con el PID′. El D_2 recupera entonces el perfil de datos en D_2/Out_2 mediante la utilización del PID_2 y asocia este perfil de datos, $Data_{out,2}$, con el PID′ recibido correspondiente al PID_2. El D_2 puede, por ende, preparar un archivo que contiene {(PID′, $Data_{out,2}$)} y envía este archivo al D_1. El D_1 puede entonces fusionar el $Data_{out,2}$ con el $Data_{out,1}$ mediante la utilización del mismo PID′, para cada PID′. Si fusionar los perfiles de datos es mutuo, entonces D_1 envía {(PID′, $Data_{out,1}$)} a D_2, que entonces puede fusionar los perfiles de datos de manera análoga.

La presente invención ha sido descripta en esta memoria en términos de algunas posibles formas de realización de ella. Los expertos en la técnica entenderán fácilmente que son posibles varias modificaciones y diferentes formas de realización, sin alejarse del alcance de protección definido en las siguientes reivindicaciones adjuntas.

REIVINDICACIONES

1. Método de asignación de seudónimos dinámicos para una red de creación de perfiles de datos (**100**) que comprende al menos un nodo de datos (**115_1-115_8**) configurado para recibir datos de entrada relacionados con usuarios y transformar dichos datos de entrada en perfiles de datos de salida de usuarios relacionados con usuarios, en donde dicho nodo de datos comprende registros de datos de usuarios (**210_1-210_3**) para almacenar datos de entrada relacionados con usuarios junto con seudónimos dinámicos de entrada de los usuarios, y dicho nodo de datos está configurado para computarizar dichos perfiles de datos de salida de usuarios relacionados con un usuario a partir de dichos datos de entrada y para almacenar los perfiles de datos de salida computarizados en dichos registros de datos de usuarios (**220_1-220_2**) del mismo; donde dicho método comprende:

recibir, en el nodo de datos, datos de entrada nuevos relacionados con un usuario junto con un seudónimo de usuario nuevo asociado y un seudónimo de usuario antiguo que estuvo asociado con datos de entrada previamente recibidos relacionados con el usuario en el pasado o un conjunto de seudónimos de usuarios antiguos candidatos;

en dicho nodo de datos, encontrar el registro de datos de usuarios correspondiente a los datos de entrada nuevos recibidos como el registro de datos de usuarios que tiene almacenado en él un seudónimo de usuario de entrada dinámico igual a dicho seudónimo de usuario antiguo recibido junto con dichos datos de entrada nuevos o a un seudónimo de usuario perteneciente al conjunto recibido de seudónimos de usuarios antiguos candidatos;

almacenar temporalmente, en el registro de datos de usuarios encontrado, los datos de entrada nuevos;

establecer el seudónimo de usuario de entrada dinámico almacenado en dicho registro de datos de usuarios de dicho nodo de datos igual al último seudónimo de usuario nuevo recibido, asociado con los datos de entrada recibidos relacionados con el usuario;

computarizar, en determinados momentos, dichos perfiles de datos de salida de usuarios mediante la utilización de datos de entrada nuevos acumulados en el registro de datos de usuarios; almacenar los perfiles de datos de salida de usuarios computarizados en el registro de datos de usuarios; y luego borrar dichos datos de entrada nuevos acumulados del registro de datos de usuarios.

2. El método de acuerdo con la reivindicación 1, que comprende lo siguiente:

dentro de dicho nodo de datos, generar y almacenar en dicho registro de datos de usuarios un seudónimo de usuario de salida dinámico junto con dichos perfiles de datos de salida de usuarios computarizados;

enviar, por momentos, dichos perfiles de datos de salida de usuarios a al menos otro nodo de datos en dicha red de creación de perfiles de datos; en cada momento generar un valor nuevo de dicho seudónimo de usuario de salida dinámico; sustituir dicho valor nuevo de dicho seudónimo de usuario de salida dinámico por un valor antiguo previamente almacenado de dicho seudónimo de usuario de salida dinámico; y enviar a dicho al menos otro nodo de datos tanto el valor nuevo como el valor antiguo de dicho seudónimo de usuario de salida dinámico junto con dichos perfiles de datos de salida de usuarios.

3. El método de acuerdo con la reivindicación 1 ó 2, en donde dichos datos de entrada son recibidos por el nodo de datos desde al menos un nodo de fuente de datos (**120_1-120_4**) de la red de creación de perfiles de datos, o desde al menos otro nodo de datos de la red de creación de perfiles de datos.

4. El método de acuerdo con la reivindicación 3, en donde la red de creación de perfiles de datos (100) comprende además:

al menos un nodo de asignación de seudónimos (**110**) configurado para recibir identidades de usuarios y transformar dichas identidades de usuarios en seudónimos de usuarios, y

en donde el método comprende además:

- recibir, en dicho nodo de asignación de seudónimos, desde dicho al menos un nodo de fuente de datos, identidades de usuarios, donde dichas identidades de usuarios identifican el usuario en dicha fuente de datos, y las identidades de usuarios comprenden uno o más identificadores de usuarios conocidos para la fuente de datos;

- generar, en dicho nodo de asignación de seudónimos, seudónimos de usuarios a partir de las identidades de usuarios recibidas;

- proveer al nodo de fuente de datos los seudónimos de usuarios generados.

5. El método de acuerdo con la reivindicación 4, en donde dichos seudónimos de usuarios se generan como valores aleatorios o pseudoaleatorios, o valores bajo clave generados por una función bajo clave a partir de identidades de usuarios y una clave secreta.

6. El método de acuerdo con la reivindicación 5, en donde dicho nodo de asignación de seudónimos es operable para encriptar y autenticar los seudónimos de usuarios generados a ser provistos al nodo de fuente de datos.

7. El método de acuerdo con la reivindicación 5 ó 6, en donde los seudónimos de usuarios aleatorios o pseudoaleatorios generados se almacenan en el nodo de asignación de seudónimos en asociación con las correspondientes identidades de usuarios.

8. El método de acuerdo con una cualquiera de las reivindicaciones 4 a 7, en donde las identidades de usuarios son diferentes para distintos nodos de fuentes de datos; y en donde dicha red de creación de perfiles de datos comprende al menos un nodo de administración de identidades de usuarios equivalentes (**105**), y en donde el método comprende además administrar, en dicho nodo de administración de identidades de usuarios equivalentes, como equivalentes identidades diferentes de un mismo usuario correspondientes a diferentes fuentes de datos.

9. El método de acuerdo con la reivindicación 8, que comprende:

después de recibir, en dicho nodo de datos, datos de entrada nuevos relacionados con un usuario desde la al menos una fuente de datos o desde el al menos otro nodo de datos, si no se encuentra el registro de datos de usuarios que incluye dicho seudónimo de usuario antiguo recibido junto con los datos de entrada nuevos o un seudónimo de usuario perteneciente a dicho conjunto de seudónimos de usuarios antiguos candidatos recibido junto con los datos de entrada nuevos, hacer que el nodo de datos determine, al explotar dicho seudónimo de usuario antiguo o dicho conjunto de seudónimos de usuarios antiguos candidatos recibido junto con los datos de entrada nuevos, si ya existe un registro de datos de usuarios con respecto a ese usuario, en donde dicho registro de datos de usuarios ha sido creado para almacenar datos de entrada relacionados con ese usuario recibidos desde al menos otra fuente de datos en el pasado.

10. El método de acuerdo con la reivindicación 9, en el que, en caso de que los datos de entrada nuevos sean recibidos por dicho nodo de datos desde dicha al menos una fuente de datos, dicha determinación comprende:

hacer que dicho nodo de datos envíe de regreso una solicitud a dicha al menos una fuente de datos para obtener seudónimos de usuarios equivalentes, donde dicha solicitud contiene dicho seudónimo de usuario antiguo o dicho conjunto de seudónimos de usuarios antiguos candidatos;

hacer que dicha al menos una fuente de datos recupere la identidad de usuario temporalmente almacenada en ella, envíe la identidad de usuario recuperada al nodo de administración de identidades de usuarios equivalentes, y solicite a dicho nodo de administración de identidades de usuarios equivalentes que provea las identidades equivalentes del usuario al nodo de asignación de seudónimos;

hacer que el nodo de asignación de seudónimos recupere los seudónimos equivalentes del usuario y luego los envíe a los nodos de fuentes de datos conectados a éste;

realizar un proceso de inundación de reenvíos que comprende:

- hacer que los nodos de fuentes de datos conectados al nodo de asignación de seudónimos reenvíen a todos los nodos de datos conectados a éste solicitudes que contienen los seudónimos equivalentes recibidos del usuario;

- cuando un nodo de datos utilizado en dicha red de creación de perfiles de datos recibe en sus entradas una o más solicitudes que contienen seudónimos equivalentes desde al menos otro nodo de datos conectado a éste, hacer que el nodo de datos busque el registro de datos de usuarios y almacene uno de los seudónimos equivalentes recibidos como seudónimos de usuarios de entrada;

- si dicho registro de datos de usuarios es encontrado, y el nodo de datos es dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos use el registro de datos de usuarios encontrado para almacenar los datos de entrada nuevos recibidos;

– si dicho registro de datos de usuarios es encontrado, y el nodo de datos es diferente de dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos continúe el proceso de inundación de reenvíos mediante el reenvío de los seudónimos de usuarios de salida almacenados en el registro de datos de usuarios a todos los otros nodos de datos conectados a sus salidas.

11. El método de acuerdo con la reivindicación 9, en el que, en el caso de que los datos de entrada nuevos sean recibidos por dicho nodo de datos desde el al menos otro nodo de datos, dicha determinación comprende realizar un proceso de retroseguimiento, un proceso de recuperación de seudónimos equivalentes, y un proceso de

inundación de reenvíos, en donde:

dicho proceso de retroseguimiento comprende:

- hacer que dicho nodo de datos envíe de regreso una solicitud a dicho al menos otro nodo de datos para obtener seudónimos equivalentes del usuario, donde dicha solicitud contiene dicho seudónimo de usuario antiguo o dicho conjunto de seudónimos de usuarios antiguos candidatos;

- hacer que dicho al menos otro nodo de datos busque el registro de datos de usuarios que almacena uno de los seudónimos de usuarios antiguos recibidos como seudónimo de usuario de salida y luego envíe de regreso al menos una solicitud a cualesquiera otros nodos de datos conectados a sus entradas, donde dicha solicitud contiene el seudónimo de usuario de entrada almacenado en el registro de datos de usuarios;

- cuando un nodo de datos utilizado en dicha red de creación de perfiles de datos recibe en cualquiera de sus salidas una solicitud desde cualquier otro nodo de datos conectados a éste, hacer que el nodo de datos busque el registro de datos de usuarios que almacena el seudónimo de usuario recibido como seudónimo de usuario de salida y luego envíe de regreso al menos una solicitud a cualquiera de los otros nodos de datos o a cualquiera de los nodos de fuentes de datos conectados a sus entradas, donde dicha solicitud contiene el seudónimo de usuario de entrada almacenado en el registro de datos de usuarios;

- cuando un nodo de fuente de datos utilizado en dicha red de creación de perfiles de datos recibe en su salida una solicitud desde cualquier nodo de datos conectado a éste, hacer que la fuente de datos reenvíe el seudónimo de usuario recibido al nodo de asignación de seudónimos con una solicitud para proveer los seudónimos equivalentes del usuario;

dicho proceso de recuperación de seudónimos equivalentes comprende:

- hacer que el nodo de asignación de seudónimos reciba desde dicho nodo de fuente de datos un seudónimo de usuario y luego recupere los seudónimos equivalentes del usuario, ya sea directamente mediante la utilización de una tabla almacenada de seudónimos o bien indirectamente mediante la utilización de una función bajo clave invertible para recuperar la identidad de usuario, luego enviar esta identidad de usuario al nodo de administración de identidades de usuarios equivalentes para proveer las identidades equivalentes del usuario, luego generar los seudónimos equivalentes candidatos del usuario a partir de las identidades equivalentes recibidas del usuario, y luego enviarlas a los nodos de fuentes de datos conectados a éste;

dicho proceso de inundación de reenvíos comprende:

- hacer que los nodos de fuentes de datos conectados al nodo de asignación de seudónimos reenvíe a todos los nodos de datos conectados a éste las solicitudes que contienen los seudónimos equivalentes recibidos del usuario;

- cuando un nodo de datos utilizado en dicha red de creación de perfiles de datos recibe en sus entradas una o más solicitudes que contienen seudónimos equivalentes desde al menos otro nodo de datos conectado a éste, hacer que el nodo de datos busque el registro de datos de usuarios y almacene uno de los seudónimos equivalentes recibidos como seudónimos de usuarios de entrada;

- si dicho registro de datos de usuarios es encontrado, y el nodo de datos es dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos use el registro de datos de usuarios encontrado para almacenar los datos de entrada nuevos recibidos;

– si dicho registro de datos de usuarios es encontrado, y el nodo de datos es diferente de dicho nodo de datos que recibe los datos de entrada nuevos, hacer que el nodo de datos continúe el proceso de inundación de reenvíos al reenviar los seudónimos de usuarios de salida almacenados en el registro de datos de usuarios a todos los otros nodos de datos conectados a sus salidas.

12. El método de acuerdo con la reivindicación 10 ó 11, que comprende:

en caso de que dentro de dicho nodo de datos, no se encuentre el registro de datos de usuarios correspondiente a los datos de entrada nuevos recibidos incluso después de recibir los seudónimos de usuarios equivalentes, hacer que el nodo de datos cree un registro de datos de usuarios nuevo con respecto a dicho usuario, y almacene los datos de entrada nuevos recibidos en él junto con el seudónimo de usuario nuevo recibido junto con los datos de entrada nuevos recibidos.

13. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde dichos seudónimos de usuarios se cambian dinámicamente después de un tiempo predeterminado según un período de validez de los perfiles de datos de usuarios.

14. El método de acuerdo con una cualquiera de las reivindicaciones 2 a 13, que comprende:

cuando una entidad que solicita perfiles de datos de usuarios desea recuperar perfiles de datos de usuarios almacenados en el al menos un nodo de datos:

- hacer que la entidad que solicita perfiles de datos de usuarios envíe una solicitud de perfiles de datos de usuarios a un nodo seleccionado de entre dicho al menos un nodo de datos de la red de creación de perfiles de datos, en donde dicha solicitud de perfiles de datos de usuarios contiene un identificador del nodo de datos que almacena los perfiles de datos de usuarios solicitados a ser recuperados, y el seudónimo de usuario de entrada o el de salida actualmente válidos, respectivamente almacenados en el registro de datos de usuarios de dicho nodo seleccionado de entre al menos un nodo de datos al cual se envía la solicitud de perfiles de datos de usuarios;

- en caso de que dicho nodo seleccionado de entre el al menos un nodo de datos que recibe la solicitud de perfiles de datos de usuarios sea el nodo de datos que almacena los perfiles de datos de usuarios solicitados, hacer que el nodo de datos recupere los perfiles de datos de usuarios solicitados almacenados en el registro de datos de usuarios en asociación con el seudónimo de usuario de entrada o el de salida actualmente válidos especificados, y provea a la entidad que solicita los perfiles de datos de usuarios los perfiles de datos de usuarios recuperados;

- en caso de que dicho nodo seleccionado de entre dicho al menos un nodo de datos que recibe la solicitud de perfiles de datos de usuarios no sea el nodo de datos que almacena los perfiles de datos de usuarios solicitados:

a) hacer que dicho nodo seleccionado de entre dicho al menos un nodo de datos identifique, en los registros de datos de usuarios almacenados en él, el seudónimo de usuario de salida correspondiente al seudónimo de usuario de salida actualmente válido recibido, contenido en la solicitud de perfiles de datos de usuarios recibida o el seudónimo de usuario de salida correspondiente al seudónimo de usuario de entrada actualmente válido recibido, contenido en la solicitud de perfiles de datos de usuarios recibida, y reenvíe el seudónimo de usuario de salida recuperado a todos los otros nodos de datos conectados a éste; y

b) repetir el paso a) mediante la utilización del seudónimo de usuario de salida recuperado recibido en lugar de dicho seudónimo de usuario de salida actualmente válido o dicho seudónimo de usuario de entrada actualmente válido, hasta alcanzar el nodo de datos que almacena los perfiles de datos de usuarios solicitados y, entonces, hacer que el nodo de datos recupere los perfiles de datos de usuarios solicitados, almacenados en el correspondiente registro de datos de usuarios.

15. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que la red de creación de perfiles de datos (**100**) comprende al menos una primera y una segunda redes de creación de perfiles de datos inconexas, cada una de las cuales comprende respectivos nodos de datos, respectivas fuentes de datos que proveen datos de entrada sobre la base de las cuales los perfiles de datos de usuarios son calculados por los nodos de datos, respectivos nodos de asignación de seudónimos para generar seudónimos de usuarios a partir de las identidades de usuarios, y en donde se provee un nodo de administración de identidades de usuarios equivalentes combinado, operable para administrar como equivalentes diferentes identidades de un mismo usuario correspondiente a diferentes fuentes de datos en la primera y la segunda redes de creación de perfiles de datos, donde el nodo de administración de identidades de usuarios equivalentes combinado es explotado para recuperar perfiles de datos de usuarios de un usuario en la segunda red de creación de perfiles de datos cuando los perfiles de datos de dicho usuario son solicitados a través de la primera red de creación de perfiles de datos.

16. Una red de creación de perfiles de datos (**100**) que está configurada para llevar a cabo el método de acuerdo con una cualquiera de las reivindicaciones anteriores.

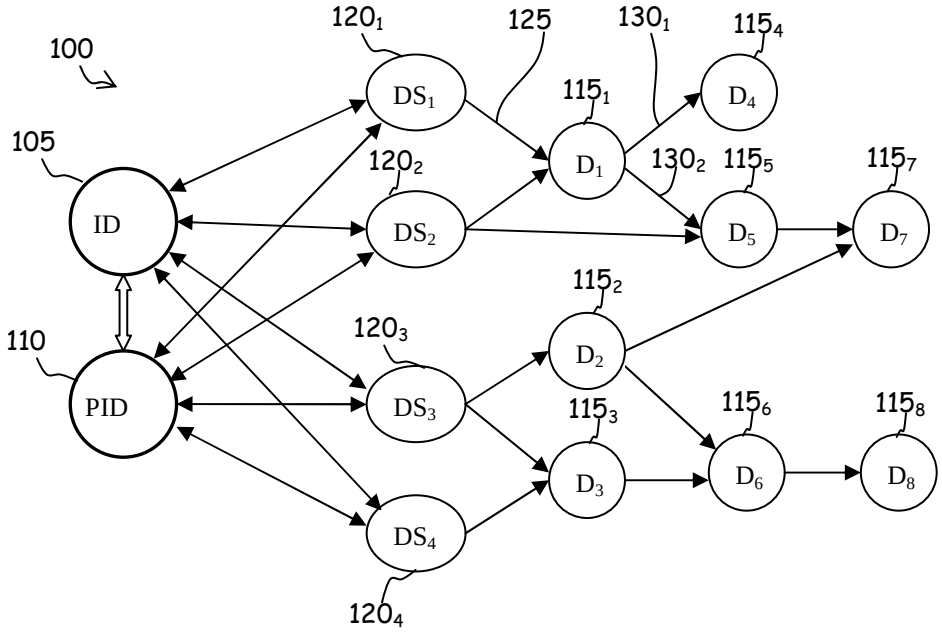
100
105
ID
110
PID
120_1
DS_1
125
130_1
115_4
D_4
115_1
D_1
120_2
DS_2
130_2
115_5
D_5
115_7
D_7
115_2
D_2
120_3
DS_3
115_3
D_3
115_6
D_6
115_8
D_8
DS_4
120_4

FIG. 1

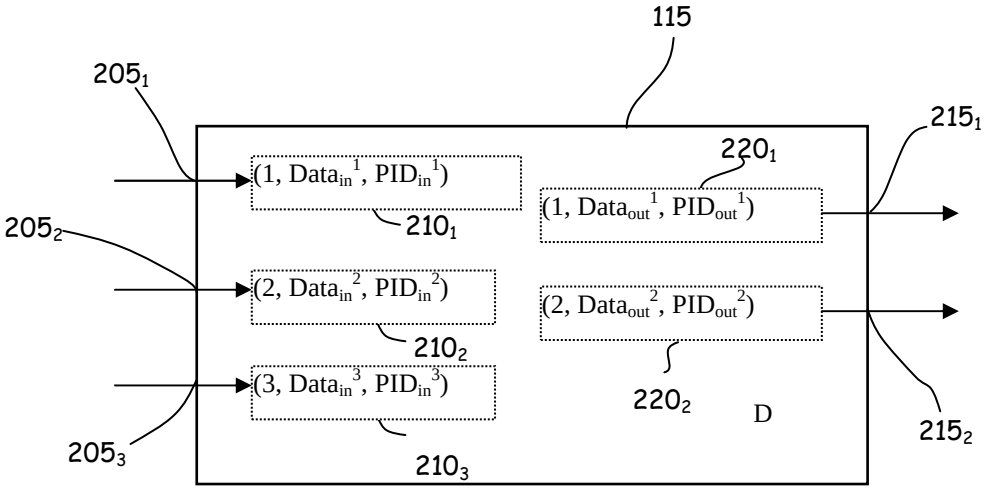
115
205_1
205_2
205_3
215_1
215_2
220_1
220_2
$(1, Data_{in}^1, PID_{in}^1)$
210_1
$(2, Data_{in}^2, PID_{in}^2)$
210_2
$(3, Data_{in}^3, PID_{in}^3)$
210_3
$(1, Data_{out}^1, PID_{out}^1)$
$(2, Data_{out}^2, PID_{out}^2)$
D

FIG. 2

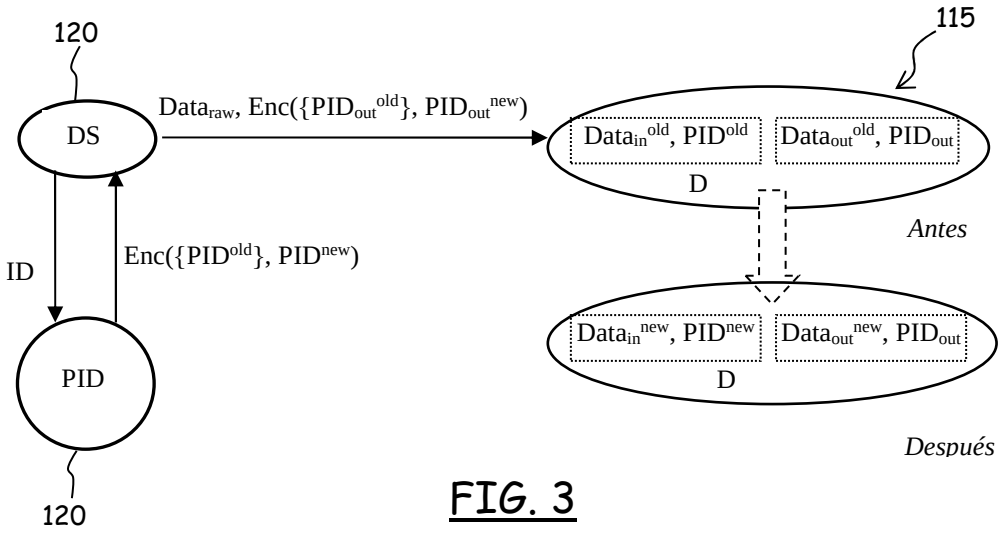

FIG. 3

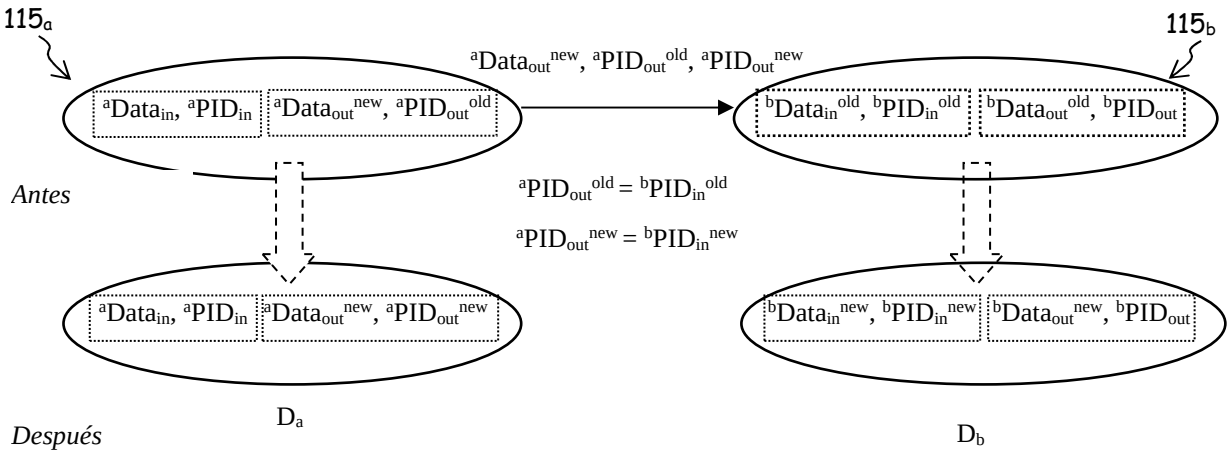

FIG. 4

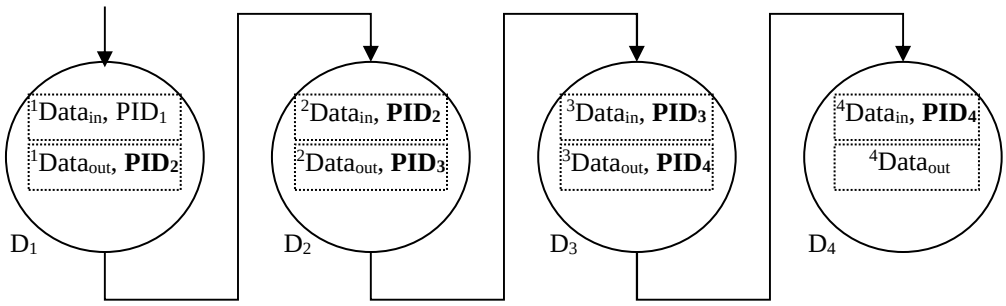

FIG. 5