(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 787 262**

(51) Int. CI.:

G06Q 30/02 (2012.01)
H04L 9/32 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **28.11.2012 PCT/EP2012/004920**

(87) Fecha y número de publicación internacional: **05.06.2014 WO14082648**

(96) Fecha de presentación y número de la solicitud europea: **28.11.2012 E 12795339 (6)**

(97) Fecha y número de publicación de la concesión europea: **11.03.2020 EP 2926307**

(54) Título: **Método de anonimización mediante transmisión de un conjunto de datos entre diferentes entidades**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
15.10.2020

(73) Titular/es:

TELEFÓNICA GERMANY GMBH & CO. OHG (100.0%)
Georg-Brauchle-Ring 50
80992 München, DE

(72) Inventor/es:

UKENA, JONATHAN y
SCHÖPF, PHILIPP

(74) Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

DESCRIPCIÓN

Método de anonimización mediante transmisión de un conjunto de datos entre diferentes entidades

La invención se refiere a un método para la transmisión de un conjunto de datos desde al menos una primera entidad que suministra el conjunto de datos a al menos una segunda entidad que agrega el conjunto de datos en el que el conjunto de datos incluye al menos un identificador que identifica al menos un usuario de la primera entidad.

Los sistemas de comunicación permiten la comunicación entre dos o más entidades. Además de mover datos de carga útil entre estas entidades, los sistemas de comunicación deben generar, recopilar y procesar datos de gestión como direcciones, ubicaciones, descripciones de servicios, etc. Por ejemplo, para un servidor web que se comunica con un ordenador de cliente, el servidor web necesita procesar la dirección IP de un cliente, las URL solicitadas, información de encabezado HTTP y datos de sesión. En un sistema de comunicación móvil, datos adicionales tales como información de ubicación, tipo de servicios o identificadores que identifican el dispositivo móvil (IM-EI), las tarjetas SIM (IMSI) se procesan. Además, cada relación de comunicación crea datos adicionales a los que se hace referencia a continuación como un conjunto de datos.

Los operadores de sistemas de comunicación también registran datos relacionados con el cliente, tal como los datos de contacto y la información del contrato. La recopilación de estos datos es necesaria para fines de facturación o para estar disponible para las autoridades. A continuación, dichos datos se definen como datos de relación con el cliente (CRM). Los datos de CRM pueden agregarse para formar datos de clase de cliente. Los documentos WO2009/009505A1, WO2009/079407A2 o WO2009/158681A1 describen tales aplicaciones.

Debido al mantenimiento de estos sistemas de comunicación modernos de información, en particular, sistemas de comunicación móvil, ofrecen la posibilidad de proporcionar información sobre los hábitos de los usuarios, en particular, con respecto a sus datos de ubicación durante un intervalo de tiempo definido. Estos datos podrían usarse para crear perfiles de ubicación para sitios geográficos o para derivar patrones dinámicos de movimiento de multitudes. En este contexto, la información podría ser útil para una amplia gama de aplicaciones en el área de servicios de tráfico, servicios de ciudad inteligente, servicios de optimización de infraestructuras, servicios de información de minoristas, servicios de seguridad y muchos más. Por lo tanto, es deseable proporcionar la información generada en forma adecuada a las partes que se benefician de aplicaciones como las mencionadas anteriormente. Tales partes podrían incluir consejos locales, empresas de transporte público e infraestructura como proveedores de transporte público o proveedores de electricidad, minoristas, grandes organizadores de eventos u organismos de seguridad pública.

Sin embargo, es obligatorio proporcionar esta información de forma anónima para proteger la privacidad de cada individuo, en particular, cada usuario del sistema de comunicación móvil. En consecuencia, la primera entidad que proporciona esta información solo debe proporcionar información extraída de datos anónimos y agregados sin vender información personal. La divulgación de cualquier información personal está estrictamente prohibida, el seguimiento e identificación de individuos debe evitarse en cualquier circunstancia.

Es el objeto de la invención proporcionar un método para anonimizar conjuntos de datos recopilados en una primera entidad y que debe proporcionarse a una segunda identidad sin permitir una conclusión a un individuo asociado con los datos recopilados.

El objeto mencionado anteriormente se resuelve mediante un método según la combinación de características de la reivindicación 1. Realizaciones preferidas son la materia objeto de las reivindicaciones dependientes. De acuerdo con la reivindicación 1 de la invención, se propone un método para transmitir un conjunto de datos desde al menos una primera entidad que suministra el conjunto de datos a al menos una segunda entidad que agrega el conjunto de datos. El conjunto de datos incluye al menos un identificador que identifica al menos un usuario de la primera entidad. El conjunto de datos a transmitir se genera en la primera entidad. La segunda entidad está interesada en la información contenida en dicho conjunto de datos para cualquier propósito de aplicación, tal como análisis de datos para crear perfiles de ubicación para sitios geográficos o derivar patrones dinámicos de movimiento de multitudes.

Para mantener la privacidad del al menos un usuario de la primera entidad, se realizan las siguientes etapas del método para ofuscar la información personal sobre el usuario: En una primera etapa, se realiza un primer cifrado del al menos un identificador sobre la base de una clave de cifrado que solo es conocida por la primera entidad. El primer cifrado conduce a un único identificador personal cifrado.

En una segunda etapa, una serie de caracteres aleatorios que definen una cadena de caracteres, que comprende preferiblemente caracteres alfanuméricos, en particular, una serie de dígitos y/o letras y/o cualquier otro tipo de símbolos, se agrega al identificador cifrado único.

En una tercera etapa, se realiza un segundo cifrado del resultado de acuerdo con la segunda etapa ejecutando un método de cifrado de clave asimétrica. La primera entidad utiliza una clave pública de la segunda entidad para el

cifrado. El resultado de la tercera etapa es un identificador doblemente encriptado.

En una última etapa, el conjunto de datos que incluye el identificador personal cifrado doble se transmite al menos a una segunda entidad.

La invención describe un proceso de anonimización de niveles múltiples, implementado preferiblemente como parte de un sistema de comunicación en la primera entidad y un sistema informático genérico en la segunda entidad. Al realizar las etapas antes mencionadas, un identificador personal incluido en el conjunto de datos se ofusca para evitar cualquier inferencia en el individuo o usuario respectivo fuera de la primera entidad. Preferentemente, tan pronto como se inicia la ejecución del método, la primera entidad no puede leer y/o modificar y/o descifrar el identificador personal.

En una realización preferida de la invención, la decodificación del identificador encriptado doble transmitido en el segundo lado de la entidad se realiza desencriptando el identificador encriptado doble en función de la clave privada respectiva del mecanismo de encriptado asimétrico aplicado. Además, el número de caracteres incluidos en la cadena de caracteres aleatorios se conoce en la segunda entidad. Sin embargo, no es necesario que la segunda entidad esté informada sobre la cadena de caracteres concreta utilizada. El número de caracteres es suficiente. Además, la segunda entidad necesita saber de qué manera la cadena de caracteres se combina con la cadena del identificador de cifrado único. La segunda entidad llegará a un solo identificador cifrado borrando el número conocido de caracteres en las posiciones conocidas en la cadena resultante del descifrado asimétrico. El identificador de cifrado único difiere del identificador personal original dentro de las premisas de la primera entidad, así como el identificador de cifrado doble suministrado por esa entidad. Por lo tanto, la identidad real del usuario permanece ofuscada en el segundo lado de la entidad. Además, el identificador cifrado único resultante es conocido exclusivamente por la segunda entidad, dado que la primera entidad no tiene acceso a un resultado intermedio del mecanismo de cifrado según la invención.

En otro aspecto preferido de la invención, la al menos una primera entidad es una red de comunicación móvil. Si es así, al menos un identificador que identifica a un usuario de la primera entidad puede ser un identificador que identifica a un usuario/cliente y/o un dispositivo móvil en un sistema de comunicación móvil, tal como el IMEI, IMSI o identificadores similares proporcionados por el estándar técnico de comunicación móvil (por ejemplo, dirección MAC) o definidos por un software que utiliza dicha infraestructura técnica (por ejemplo, credenciales de cuenta en aplicaciones móviles o sistemas operativos).

En todos los aspectos de la invención, el conjunto de datos incluye al menos un atributo de datos asociado con el al menos un identificador/usuario y que se transmite dentro del conjunto de datos que incluye el identificador doblemente encriptado respectivo.

En detalle, en un sistema de comunicación móvil, un conjunto de datos puede comprender un IMSI como identificador y otros atributos de datos que caracterizan un evento actualmente activado por el usuario o dispositivo móvil respectivo, por ejemplo, una solicitud de servicio (voz, datos, etc.) o un proceso de búsqueda de posición ejecutado o cualquier otro posible evento activo o pasivo que pueda activarse en una red de comunicación móvil.

Por ejemplo, al menos un atributo del conjunto de datos incluye información de eventos del usuario que caracteriza el tipo de evento capturado actualmente y/o la marca de tiempo de un evento y/o la ubicación actual del usuario en dicha marca de tiempo. En detalle, en un sistema de comunicación móvil, el conjunto de datos puede comprender el IMSI como un identificador y el tipo de servicio, por ejemplo, SMS, voz de conmutador de circuito, datos de conmutación de paquetes, y/o la marca de tiempo y/o la posición exacta o estimada del usuario que exige el tipo de servicio respectivo.

Además, los operadores de la primera entidad registran datos relacionados con el usuario, tal como datos de contacto o información del contrato. La recopilación de esos datos es necesaria para fines de facturación o para estar disponible para las autoridades. Estos datos se refieren a datos relacionados con el cliente (usuario) (CRM).

En todas las realizaciones, al menos un atributo de datos del conjunto de datos contiene datos relacionados con el usuario (CRM). De forma ventajosa, los datos CRM se agregan para formar datos de clase de usuario que combinan atributos de varios usuarios en una clase de usuario común. En particular, los atributos de datos pueden contener información sobre la edad del usuario o el género del usuario. En ese caso, es posible formar datos de clase de usuario para clasificar un número de usuarios en diferentes clases o grupos, en particular, un determinado grupo de edad o grupo de género.

En otra realización preferida de la presente invención, podría ser útil tener una separación física entre las transmisiones de diferentes tipos de atributos.

Preferentemente, el método se aplica por separado a los diferentes tipos de atributos de un conjunto de datos. Diferentes atributos pueden ser procesados por diferentes instancias de la primera entidad y/o diferentes primeras entidades. Los procesos separados pueden ejecutarse en paralelo o consecutivamente por diferentes instancias y/o entidades.

También es una característica de la invención separar la transmisión de atributos de clase de usuario, tal como el sexo del usuario y/o la edad del usuario, y los atributos de datos del evento, tal como el tipo de evento, marca de tiempo, ubicación, en diferentes procesos ejecutados por la primera entidad. Por ejemplo, en un primer proceso, el identificador se encripta aplicando el método inventivo en el que dicho identificador doble encriptado se transmite en combinación con atributos de datos de eventos a la segunda entidad. En un segundo proceso, el mismo identificador se cifra de acuerdo con la invención y se transmite junto con los datos de clase de usuario a la segunda entidad. Ambos procesos vuelven a la misma primera clave y/o clave pública para el cifrado. Sin embargo, ambos procesos se basan en una cadena de caracteres aleatorios diferente que da como resultado identificadores dobles encriptados únicos para cada trayectoria de transmisión. De esta manera, la combinación de datos de clase de cliente y datos de eventos fuera de la segunda entidad es imposible debido a diferentes identificadores de doble cifrado. Este punto es particularmente relevante, si se usa una red pública para la transmisión entre la primera y la segunda entidad.

Además, al menos un atributo de datos incluido en el conjunto de datos puede codificarse y/o cifrarse antes de la transmisión. Sin embargo, los atributos de datos también se pueden transmitir en texto plano.

De forma ventajosa, el primer cifrado puede basarse en un método de cifrado irreversible utilizando una clave unidireccional.

Preferentemente, la primera etapa de cifrado puede usar claves diferentes con vidas variables. Por ejemplo, un método puede volver a una clave a corto plazo y/o una clave a largo plazo que tenga una vida útil de 24 horas o un año. Es obvio que se pueden concebir otros intervalos de tiempo o combinaciones con más de dos teclas. Los métodos seleccionan preferiblemente el tiempo de vida apropiado para la primera clave de cifrado dependiendo del identificador y/o el atributo del conjunto de datos a transmitir. La vida útil define un intervalo de tiempo dentro de un único identificador cifrado que la segunda entidad considera válido.

Algunas aplicaciones pueden requerir la incorporación de atributos de datos adicionales proporcionados por una tercera entidad como un conjunto de datos adicional que se transmitirá a la segunda entidad donde se asociará a los conjuntos de datos anónimos proporcionados por la primera entidad. Por lo tanto, de acuerdo con un aspecto preferido de la invención, al menos un atributo de datos adicional asociado a al menos un identificador o usuario de la primera entidad es proporcionado por una tercera entidad y se transmite a un socio de confianza, en donde está asignado a un identificador doble encriptado válido, que es proporcionado por la primera entidad. Después de la asignación, el socio de confianza transmite un nuevo conjunto de datos combinados con un identificador doble encriptado válido y los atributos de datos adicionales de la tercera entidad a la segunda entidad. La segunda entidad descifra el identificador doblemente encriptado de acuerdo con el proceso descrito anteriormente. Basado en el identificador cifrado único resultante, la segunda entidad ahora puede incorporar los atributos adicionales de la tercera entidad a los conjuntos de datos proporcionados por la primera entidad.

Es particularmente preferible si la entidad asociada de confianza asigna al menos un atributo de datos adicional al conjunto de datos haciendo coincidir un segundo identificador cifrado insertado con los datos recibidos de la primera entidad con un segundo identificador cifrado insertado a los datos recibidos de la tercera entidad. Ambos segundos identificadores están encriptados con claves idénticas en la primera y tercera entidad. En caso de una coincidencia positiva, el socio de confianza reemplazará el segundo identificador cifrado recibido con los datos de la tercera entidad con un identificador doble cifrado válido recibido desde la primera entidad. Posteriormente, se envía un conjunto de datos modificado a la segunda entidad que incluye el primer identificador encriptado doble recibido desde la primera entidad y el al menos un atributo de datos adicional recibido desde la tercera entidad.

En otra realización preferida de la presente invención, un conjunto de datos transmitido a la segunda entidad solo es aplicable dentro de su vida útil definida que se logra usando diferentes claves y/o cadenas de caracteres aleatorias con diferentes vidas. Esta es una medida de precaución para evitar la derivación de patrones de un cierto número de conjuntos de datos almacenados que crean un historial de eventos detallado. Por lo tanto, las tendencias pueden extrapolarse de una colección de conjuntos de datos durante un intervalo de tiempo definido y almacenarse como índices estadísticos a largo plazo que sobrevivirán a la vida útil seleccionada del evento respectivo. En particular, el cálculo de índices estadísticos a largo plazo que permiten declaraciones de probabilidad sobre movilidad y actividad basadas en datos de eventos de ubicación es posible sin guardar ningún historial de eventos detallado. Los índices estadísticos pueden mantenerse y usarse dentro de un intervalo de tiempo largo, por ejemplo, más de un año. Los datos del evento de ubicación se pueden descartar después de un breve intervalo de tiempo, por ejemplo, después de 24 h, para minimizar el riesgo de que los patrones de movimiento se puedan derivar de los datos de eventos de ubicación almacenados. Al usar índices en su lugar, no se puede extraer ninguna conclusión sobre la identidad real de un usuario generador de datos de eventos de ubicación.

Los índices pueden incluir una probabilidad estadística para la ocurrencia de un evento definido en una ubicación definida. Adicionalmente o como alternativa, Los índices pueden incluir un valor numérico que representa el número de eventos producidos en una ubicación definida.

Preferentemente, puede ser importante reconocer y filtrar las irregularidades para generar declaraciones estadísticas

precisas. El filtrado puede realizarse integrando valores de comparación históricos durante un período de tiempo más largo. Los enfoques sugeridos permiten derivar declaraciones de probabilidad sobre movilidad y actividad basadas en datos anónimos reales sin guardar ningún historial de eventos detallado que pueda usarse para derivar patrones que permitan sacar conclusiones sobre la identidad de un individuo.

La invención está relacionada además con un sistema de comunicación para realizar el método según la invención o según una realización preferida de la invención. Es obvio que el sistema de comunicación se caracteriza por las propiedades y ventajas según el método de la invención. Por lo tanto, una descripción repetitiva se considera innecesaria.

Las ventajas y propiedades adicionales de la presente invención deberían describirse sobre la base de tres realizaciones preferidas mostradas en las figuras. Las figuras muestran

Figura 1: una descripción esquemática de las etapas básicas del método,

Figura 2a, 2b: dos versiones de una primera realización preferida de la presente invención,

Figura 3a, 3b: dos versiones de una segunda realización preferida de la presente invención, y

Figura 4a, 4b: dos versiones de una tercera realización preferida de la presente invención.

La figura 1 ilustra la idea fundamental de la presente invención. La idea básica de la presente invención se refiere a un procedimiento de anonimización de datos para permitir el uso de datos de ubicación masiva para aplicaciones de big data con pleno respeto de los estándares europeos de protección de datos. Los datos de ubicación masiva serán recopilados por proveedores de redes de comunicaciones móviles o inalámbricas, así como proveedores que recopilan información que se basa en otras tecnologías de ubicación como GPS, Galileo, GLONASS, Compass, redes de sensores, etc., que también pueden poseer información personal detallada y verificada sobre sus usuarios. Además, los proveedores de redes móviles pueden extraer datos de eventos de ubicación de sus usuarios. Toda la información se combina en conjuntos de datos anónimos que pueden ser de interés para diferentes aplicaciones ofrecidas por empresas de terceros.

Por ejemplo, los proveedores de redes móviles pueden vender o proporcionar los datos anónimos y agregados a los consejos locales, empresas de transporte público, empresas de infraestructuras como proveedores de transporte público o proveedores de electricidad, minoristas, los principales organizadores de eventos o cuerpos de seguridad pública que utilizan dicha información para mejorar sus procesos de toma de decisiones.

Los conjuntos de datos proporcionados también pueden analizarse para determinar cuántas personas visitan un área por tiempo, género y edad. Las organizaciones podrán analizar los movimientos de las multitudes en cualquier lugar por hora, día, semana o mes, y hacer comparaciones comparables por área, así como también comprender los patrones de captación.

Una aplicación particular de los datos podría ser la implementación de ciudades inteligentes. El análisis de datos antes mencionado podría usarse para analizar el volumen de tráfico en ciertos distritos de la ciudad. Por lo tanto, el ayuntamiento puede optimizar la ingeniería vial sobre la base de dicha información. Por supuesto, dicha información es útil para cada planificación de construcción teniendo en cuenta la cantidad de usuarios/visitantes potenciales.

Sin embargo, es obligatorio cuidar la privacidad de cada usuario y la información personal del usuario. Por lo tanto, el objetivo de la presente invención es definir un proceso que permita la anonimización real en lugar de solo la seudoanonimización. Al dividir el proceso en varias etapas del proceso que se ejecutaron dentro de diferentes premisas legales de entidades independientes, se evita la posibilidad de generar una tabla de asignación entre identificadores anónimos y no anónimos.

Como se puede ver en la figura 1, un proveedor de datos tal como una primera entidad y al que se hace referencia como DS está conectado comunicativamente a un agregador de datos como una segunda entidad y se hace referencia como DA a través de una red pública o privada virtual. La entidad proveedora de datos DS puede ser cualquier proveedor de datos de movimiento y/o personales. Esto incluye, por ejemplo, proveedores de redes móviles o inalámbricas, así como proveedores que recopilan información que se basa en otras tecnologías de ubicación como GPS, Galileo, GLONASS, Compass, redes de sensores, La siguiente argumentación se basará en el caso ejemplar de un sistema de red móvil que proporciona los conjuntos de datos antes mencionados que contienen datos personales, así como datos de eventos de ubicación sobre sus usuarios. Cada usuario individual de la red general de DS se identifica mediante un identificador personal PID. Para tener una anonimización real de acuerdo con las normas europeas de protección de datos, es necesario, entre otras cosas, separar el PID inicial y su contraparte, el O-PID (identificador personal ofuscado). En este contexto, el esfuerzo de reunir estos dos identificadores tiene que ser desproporcionado en comparación con el rendimiento que se podría obtener con tal acción. Este requisito se cumple,

si la separación se realiza físicamente dentro de las premisas de dos entidades legales independientes, donde una entidad solo conoce el PID y la otra solo el O-PID. Por lo tanto, es necesario encriptar y transmitir dicho identificador a una tercera parte denominada como el agregador de datos DA. Dicho identificador personal se combina con un conjunto de datos con atributos de datos adicionales que describen un determinado evento de ubicación. Por ejemplo, estos atributos de datos de eventos caracterizan una acción de un usuario en un lugar determinado. La ofuscación de los datos sensibles debe realizarse mediante un proceso de anonimización de múltiples niveles (MAP) realizado en el DS para proteger la privacidad del usuario.

En una primera etapa 1, un cifrado de primer nivel se ejecuta cifrando el PID sobre la base de una primera clave (clave DS) que solo el proveedor de datos DS conoce. Este mecanismo de cifrado podría ser un algoritmo hash simple con una clave unidireccional (clave DS). Diferentes claves DS pueden estar disponibles en el lado DS con diferentes vidas como ST/IT (corto tiempo/largo tiempo), por ejemplo. La salida de la primera etapa del método es un único PID ofuscado al que se hace referencia como O-PID. La vida útil de dicho O-PID depende del intervalo en que se cambia la clave DS. Es decir, si la clave DS es, por ejemplo, constante durante 24 horas, el DA obtendrá un identificador ofuscado estático durante exactamente ese período de tiempo.

En una segunda etapa 2, se agrega un número aleatorio de varios dígitos al final del O-PID de salida del cifrado de primer nivel en la etapa 2. Se observa que cualquier otra cadena de caracteres generada aleatoriamente y cualquier otro procedimiento de combinación de las dos cadenas podría ser apropiado. Las longitudes del intervalo del número aleatorio utilizado también podrían ser variables, pero tiene que ser conocido por el DA. La salida de la segunda etapa se marca como O-PID + RN.

En la última etapa 3, se ejecuta un cifrado de segundo nivel sobre la base de un mecanismo de cifrado asimétrico que utiliza la clave pública DA-Pub-Clave de la segunda entidad DA. El cifrado asimétrico se aplica al resultado de la etapa 2 O-ID+RN resultante en un resultado que está marcado como OO-PID. En consecuencia, el PID está doblemente ofuscado para proteger la privacidad del usuario.

La vida útil del identificador encriptado doble OO-PID PID solo depende del intervalo en el que se cambia el número aleatorio utilizado en la etapa 2. Esto significa que el OO-PID es constante siempre que el RN sea constante, lo cual es importante para los cálculos realizados en el OO-PID por un socio de confianza (por ejemplo, que construye índices estadísticos como se describirá más adelante). Por el contrario, el valor real del número aleatorio no es necesario para la decodificación del OO-PID en el DA.

Las etapas 1 a 3 se implementan en una sola sección de código. Es imposible que el proveedor de datos DS lea o escriba cualquier información generada entre las etapas individuales.

En el lado del agregador de datos, el descifrado DA se ejecuta en el segundo nivel utilizando su clave privada DA-Priv-Clave para descifrar el identificador cifrado recibido OO-PID. El resultado O-PID+RN se procesará aún más borrando el número conocido de dígitos al final de la cadena que representa el número aleatorio. El resultado resultante es el O-PID. La duración de este identificador cifrado único O-PID en el lado del agregador de datos DA se define por la longitud del intervalo de la clave DS generada. Si la longitud del intervalo de la Clave-DS ha transcurrido, se generará una nueva Clave-DS y, por lo tanto, un nuevo O-PID en el DS.

El PID original solo es visible en el lado del proveedor de datos DS ya que el lado del agregador de datos DA solo conoce el identificador cifrado único O-PID. Por lo tanto, es imposible construir un catálogo (una tabla que asigna cada PID no anónimo a su parte contraria anónima, el O-PID) dentro de las instalaciones de una sola parte.

Una primera realización de la presente invención se proporciona en la figura 2a. Describe una solución técnica para el anonimato de diferentes conjuntos de datos entregados por un único proveedor de datos DS. La anonimización y la transmisión de estos conjuntos de datos a un único agregador de datos DA se procesan mediante procesos completamente separados que se ejecutan en el proveedor de datos DS. Los diferentes conjuntos de datos se pueden combinar en función de sus identificadores iguales en el agregador de datos DA. La primera realización es apropiada, si el proveedor de datos DS está sujeto a restricciones legales o de otro tipo con respecto a la combinación de conjuntos de datos específicos en forma no anónima. Con respecto a los estándares europeos de protección de datos, esto se aplica a la combinación de datos de eventos de ubicación con datos personales de clientes, por ejemplo.

Por lo tanto, todo el proceso se subdivide en dos procesos independientes de anonimización de niveles múltiples (MAP) donde los identificadores personales PID (como elementos únicos entre los conjuntos de datos) se anonimizan por separado y se transmiten al agregador de datos DA junto con sus respectivos conjuntos de datos. De ese modo, el primer proceso MAP es responsable de transmitir los llamados atributos de evento de ubicación, incluido el tipo de evento, una marca de tiempo y la ubicación del usuario. El segundo proceso es responsable de transmitir los atributos que clasifican a los usuarios/identificadores en diferentes grupos de clases de usuarios, por ejemplo, género o grupos de edad.

Como se puede ver en la figura 2a, ambos procesos ejecutan el cifrado de primer nivel sobre la base de una clave DS idéntica. La vida útil de esta clave se ha establecido a modo de ejemplo en 24 horas. Para distinguir este tipo de clave DS de otras claves requeridas para otras aplicaciones, además de tener un nombre general para una mayor argumentación, esta clave DS se define como clave a corto plazo o clave ST (en referencia a su vida útil relativamente corta). En una segunda etapa, se agregan números aleatorios individuales RN a los resultados de las primeras etapas. Los números aleatorios se cambian para cada procedimiento de encriptación de cualquier conjunto de datos nuevo. Por lo tanto, los números aleatorios RN diferirán dentro del primer MAP para cada evento de ubicación individual y entre los propios dos procesos de MA.

En la tercera etapa, se ejecuta un cifrado de segundo nivel en el O-PID + RN mediante el uso de DA-Pub-Clave A. Esta clave se genera a partir del agregador de datos DA como parte de un par de claves asimétricas A. El DA proporciona el DA-Pub-Clave A al DS con el fin de realizar el cifrado de segundo nivel. Más tarde, esto podría descifrarse mediante el uso de la DA-Priv-Clave A, que solo DA conoce. En este contexto, la "A" mayúscula se entiende como un contador. Al tener más de un par de claves asimétricas, se pueden distinguir por la mayúscula (par de claves A, B, C ...). En la realización dada, ambos MAP que realizan el cifrado de segundo nivel emplean la misma clave pública DA-Pub-Clave A. Los resultados de la tercera etapa son PID de doble cifrado. Dado que estos PID dobles encriptados se basan en O-PID generados mediante el uso de una clave ST a corto plazo como se define en la etapa uno, el identificador doble encriptado OO-PID debe llamarse ST-OO-PID. Se transmiten diferentes ST-OO-PID en combinación con sus respectivos atributos a través de diferentes rutas de transmisión por cada MAP.

Debido al hecho de que el número aleatorio es diferente dentro de cada procedimiento de cifrado de segundo nivel, los ST-OO-PID cifrados resultantes son únicos. Es decir, un ST-OO-PID específico siempre pertenece a un conjunto de datos específico, por lo tanto, la vida útil de un ST-OO-PID se limita a un solo evento (que en este contexto incluye la generación de datos de eventos de ubicación, así como datos de clase de cliente). En consecuencia, no es posible combinar ninguna clase de clientes (género, grupo de edad) y datos del evento (tipo de evento, marca de tiempo, ubicación) ni varios conjuntos de datos de eventos de ubicación dentro de las instalaciones de DS del proveedor de datos o en la ruta de transmisión al agregador de datos DA.

La combinación de datos mencionada anteriormente solo puede ser realizada por el agregador de datos DA. Por lo tanto, el agregador de datos primero descifra el ST-OO-PID con la clave privada respectiva (DA-Priv-Clave A) correspondiente a la DA-Pub-Clave A proporcionada al proveedor de datos DS. La cadena de salida incluye el O-PID además del número aleatorio RN. Como DA sabe el número de dígitos bloqueados para el RN al final de la cadena, el agregador de datos simplemente elimina el número de dígitos para obtener el O-PID. Basado en este elemento unificador, el DA puede combinar los datos durante un período de tiempo correspondiente a la vida útil de la clave ST utilizada por el DS para generar el O-PID. Por lo tanto, el DA podría combinar varios eventos de ubicación con datos de clase de cliente para un O-PID estático durante un período de tiempo de 24 horas.

Una versión ligeramente diferente de la realización descrita anteriormente se muestra en la figura 2b. Este modelo de proceso difiere del que se muestra en la figura 2b en el hecho de que los dos procesos de MA son realizados por dos agregadores de datos diferentes DS1 y DS2. En el caso de un operador de red móvil, esto podría ser necesario, por ejemplo, si partes de la infraestructura (como la red móvil, por ejemplo) se subcontratan a otras empresas.

Para generar O-PID iguales, ambos agregadores de datos deben usar la misma clave DS y la misma técnica con respecto a la adición de números aleatorios a este resultado. El agregador de datos podría proporcionar diferentes claves públicas de diferentes conjuntos de claves. Para que sea simple, en este ejemplo, ambos DS funcionan con las mismas claves públicas, a saber, DA-Pub-Clave A.

Una segunda realización es el proceso de indexación anónima a largo plazo (ALIP) (figura 3a). Esta realización debería permitir el cálculo de índices estadísticos a largo plazo que permitan declaraciones de probabilidad sobre movilidad y actividad basadas en datos de eventos de ubicación, sin guardar ningún historial de eventos detallado. El desafío general se plantea por el hecho de que la precisión de tales declaraciones estadísticas depende directamente de la cantidad de datos disponibles para la derivación de las declaraciones. Si, por ejemplo, un operador de red móvil desea calcular el código postal donde probablemente vive un identificador personal anónimo (O-PID), una probabilidad respectiva es estadísticamente derivable contando el número de eventos asignados a este O-PID en diferentes ubicaciones entre las 7 pm y las 7 am en días hábiles. En este contexto, es muy importante reconocer y filtrar las irregularidades (vacaciones, viajes de negocios, etc.) para generar declaraciones precisas. El filtrado puede realizarse integrando valores de comparación históricos durante un período de tiempo más largo. Este argumento podría extenderse a muchas otras aplicaciones de movilidad y actividad (por ejemplo, código postal de trabajo, comportamiento de uso promedio de varios servicios, ...). En general, la precisión de las declaraciones estadísticas calculadas depende de la cantidad de datos de eventos de ubicación recopilados dentro de un intervalo de tiempo dado. Cuantos más datos se usen para el cálculo, mejor será la precisión. Además, la precisión también aumentará si se extiende el intervalo de tiempo definido para la recopilación de datos.

Sin embargo, un número creciente de datos de eventos de ubicación recopilados y almacenados aumenta proporcionalmente la probabilidad de una identificación exitosa de patrones de movimiento únicos a partir de los datos almacenados. Estos patrones permiten llegar a una conclusión sobre la identidad real de una persona o al revés.

Según las normas europeas de protección de datos, es obligatorio minimizar el riesgo de identificación de patrones. A la luz de lo anterior, el proceso ALI de construir índices a largo plazo sin guardar el historial detallado de eventos es otro componente central de esta invención.

Las siguientes variaciones de procesos de indexación a largo plazo anónimos (ALIP) (figura 3a y figura 3b) describen soluciones técnicas para tal problema basadas en la técnica general del proceso de anonimización multinivel (MAP) como se presentó anteriormente. La idea básica prevé dividir diferentes partes del proceso, así como la extensión de los datos visibles entre las diferentes entidades participantes. Además de los proveedores de datos conocidos y los agregadores de datos, se presenta un socio de confianza (TP) como una nueva instancia entre DS y DA. El socio de confianza TP crea índices estadísticos basados en conjuntos de datos de eventos de ubicación reales durante un período a corto plazo y reenvía solo estos índices (pero ningún dato de eventos de ubicación reales) a un DA. El DA respectivo asocia estos índices con valores de comparación históricos durante un período a largo plazo.

La siguiente explicación describe una posible variación del proceso visualizada en la figura 3a. En esta variación, un único DA es responsable del cálculo de los índices a largo plazo, así como de la agregación de datos. Dado que la agregación de datos se basa en un identificador personal ofuscado a corto plazo (ST-O-PID) y los índices a largo plazo se identifican con un identificador personal ofuscado a largo plazo (LT-O-PID), el DA no puede combinar ambos conjuntos de datos directamente dentro de sus instalaciones. Una combinación siempre requiere una traducción a través del socio de confianza. No obstante, en algunos casos, puede desearse distribuir el cálculo de índices a largo plazo y las tareas de agregación de datos en dos agregadores de datos independientes (DA 1 y DA 2). Dicha versión modificada respectiva de la realización de acuerdo con la figura 3a se visualiza en la figura 3b y funciona de forma análoga a la versión descrita anteriormente y a continuación.

En el proveedor de datos, los datos del evento de ubicación del lado DS consisten en un identificador personal PID y el tipo de evento de atributos de datos, marca de tiempo, la ubicación se aplica a dos tipos diferentes de procesos de anonimización basados en la lógica MAP como se describió anteriormente. El primer MAP funciona con la misma DS-Clave a corto plazo que también se usa para cifrar datos de eventos de ubicación como se describe en la primera realización de esta invención. Por lo tanto, la clave también cambia a modo de ejemplo cada 24 horas (ST-Clave). El ST-O-PID saliente se agrega mediante un número aleatorio cambiante RN que cambia cada vez que este primer MAP se realiza con un número conocido de dígitos y luego se encripta por segunda vez mediante el uso de la clave pública proporcionada por el DA (DA-Pub- Clave A). El primer MAP se activa una vez cada vez que el segundo MAP genera un nuevo identificador personal doblemente ofuscado a largo plazo LT-OO-PID (véase más adelante). El resultado del primer proceso es un único ST-OO-PID que el DA puede descifrar para obtener un ST-O-PID como un identificador único para la combinación con otras fuentes de datos en un momento posterior. La vida útil de este ST-O-PID en el DA será de 24 horas en el ejemplo dado.

El segundo MAP opera con una DS-Clave a largo plazo (clave LT). De acuerdo con la realización ilustrada de la figura 3a, la clave LT se cambia una vez al año. El cifrado de primer nivel del PID con la clave LT resulta en un LT-O-PID. En la segunda etapa, se agrega un número aleatorio (con un número conocido de dígitos) al LT-O-PID. En el ejemplo dado, el RN cambia en los mismos períodos de tiempo que la clave ST. Por lo tanto, la RN es constante durante 24 horas. Dentro del cifrado de segundo nivel, la combinación de LT-O-PID + RN se ofusca al usar otra clave pública del DA (DA-Pub-Clave B). Por lo tanto, el LT-OO-PID resultante es constante para todos los eventos transmitidos al socio de confianza TP dentro de las 24 horas. Después de que el agregador de datos decodifica el LT-OO-PID, un DA LT-O-PID constante durante un año está disponible en el DA.

Como se mencionó anteriormente, el primer MAP para generar el ST-OO-PID se activa solo uno cada la vez, cuando el segundo MAP genera un nuevo LT-OO-PID (en el ejemplo descrito en el presente documento, esto es cada 24 horas cuando cambia el RN del segundo MAP). En este momento, el ST-OO-PID (sin más atributos) y el respectivo LT-OO-PID (fuera de un conjunto completo de datos de ubicación que incluye todos los atributos) se transmiten al socio de confianza TP. Al mismo tiempo, el LT-OO-PID (sin más atributos) también se reenvía al DA. Aquí se utiliza para devolver los valores de los índices a largo plazo, ya que se han calculado hasta el final de la última vida útil de LT-OO-PID (esto se describe en detalle más adelante).

El TP guarda la asignación entre ST-OO-PID y LT-OO-PID en una tabla de traducción ST/IT. Luego, el socio de confianza crea índices estadísticos basados en los conjuntos de datos de eventos de ubicación que entrega el DS a través del segundo MAP dentro de la vida útil del LT-OO-PID (aquí: 24 horas). Como estos índices se calculan dentro de un intervalo a corto plazo, se denominan índices a corto plazo. En las figuras 3a y 3b muestra un único índice (ST-Índice 1) como representante de algunos índices potenciales que podrían calcularse en este punto. El término índice estadístico en esta descripción debe incluir valores de frecuencia simples (por ejemplo, número de envíos de SMS en el corto plazo), así como valores de probabilidad (por ejemplo, el 80 % de los eventos entre las 7 p.m. y las 7 a.m. ocurrieron en una región geográfica que tiene el código postal 80639). Los eventos de ubicación originales se descartan después de haber sido procesados para calcular varios índices estadísticos. Antes del final de la vida útil de LT-OO-PID, un nuevo conjunto de datos con el LT-OO-PID, así como todos los índices ST construidos para este identificador dentro del período de corto plazo en el TP, se envían al agregador de datos DA.

El DA descifra el LT-OO-PID utilizando la clave privada apropiada (DA-Priv-Clave B). El próximo identificador a largo plazo LT-O-PID permite al DA combinar los nuevos índices a corto plazo recibidos del TP con valores históricos para los mismos índices. Por lo tanto, el DA primero guarda los nuevos valores en su base de datos (historial de índices a corto plazo) y luego calcula nuevos índices a largo plazo (por ejemplo, LT-Índice 1) basándose en todos los valores (nuevos e históricos) en la base de datos. Estos índices a largo plazo se combinan en un nuevo conjunto de datos y se guardan con el LT-O-PID como un identificador hasta el comienzo de un nuevo período de LT-OO-PID.
Al comienzo del nuevo período, se envía un nuevo LT-OO-PID desde el DS al DA como se describió anteriormente. Este LT-OO-PID se descifra mediante el uso de la DA-Priv-Clave B para encontrar el LT-O-PID apropiado. Posteriormente, los índices a largo plazo que se han almacenado para el LT-O-PID en la base de datos del historial de índices a corto plazo hasta ese momento se combinan con el LT-OO-PID recién recibido. Si no hay valores históricos disponibles en la base de datos, se aplica un valor ficticio "n.a." a todos los atributos del conjunto de datos recién generado antes de retransmitir dicho conjunto de datos al TP. A medida que el TP obtiene este nuevo conjunto de datos con LT-OO-PID y los índices a largo plazo calculados por el DA, busca el ST-OO-PID apropiado que guardó para el LT-OO-PID en la tabla de traducción ST/IT anteriormente. Después de cambiar el LT-OO-PID con el ST-OO-PID, se reenvía el conjunto de datos al DA nuevamente. Para asegurar, que el DA solo podría usar la tabla de traducción dentro de un período de vida LT-OO-PID, la asignación ST/IT utilizada se puede eliminar después.

Nuevamente en el DA, el ST-OO-PID se descifra con la DA-Priv-Clave A. El ST-O-PID saliente como un identificador único finalmente permite la combinación de los índices a largo plazo con otras fuentes de datos como el evento de ubicación y datos de clase de cliente. Al transmitir los índices a largo plazo que se han calculado hasta el final del último período de vida de LT-OO-PID al comienzo del nuevo período de vida de LT-OO-PID, el ALIP garantiza la disponibilidad de índices a largo plazo (o valores ficticios) en cada momento de la vida útil de ST-O-PID.

La solución técnica dada brinda la posibilidad de calcular índices estadísticos anónimos a largo plazo sin guardar datos de eventos de ubicación durante un período de tiempo más largo. Como ya se mencionó anteriormente, la segunda versión de esta solución (figura 3b) funciona de la misma manera, pero difunde el cálculo de índices a largo plazo y la combinación de varias fuentes de datos a dos agregadores de datos diferentes.

Otra realización de la presente invención se muestra en las figuras 4a y 4b. Esta realización proporciona una solución para una aplicación de terceros para proporcionar atributos de datos adicionales (Atributo 1...Atributo n) identificados por un identificador personal secundario (SID), que también es conocido por el DS (variante 1 según la figura 4a), o por el PID original del DS (variante 2; figura 4b). En este contexto, la tercera parte actúa como una entidad de entrega de datos similar al DS. Una diferencia importante entre el DS y la entidad de tercera parte es que la tercera parte no realiza el MAP por sí misma.

Según una primera variante (figura 4a), un socio de confianza TP compara el conjunto de datos recibido de la tercera parte con un identificador personal doblemente encriptado a corto plazo (ST-OO-PID) generado (a través de un MAP) y proporcionado por el DS. Tal escenario podría ser razonable si la tercera parte no está dispuesta o no puede proporcionar sus datos directamente al DS. La siguiente argumentación describe este proceso basado en un identificador secundario (SID). También es posible realizar el proceso de correspondencia del socio de confianza utilizando el identificador personal PID. Por lo tanto, el PID simplemente necesita ser encriptado de la misma manera descrita para el SID. Es decir, debe encriptarse una segunda vez en paralelo al MAP en un proceso de un solo nivel.

Además de la generación del ST-OO-PID, el DS también define una nueva clave para el cifrado del identificador secundario SID (Clave SI). Por un lado, el DS utiliza esta clave para cifrar todos los SID en su base de datos con un cifrado simple (de un nivel) para obtener el O-SID. Por otro lado, la clave se proporciona a la tercera parte para habilitar el mismo proceso en sus conjuntos de datos. El resultado en el DS es una tabla de asignación que vincula todos los O-SID con el ST-OO-PID apropiado. La tabla se transmite al socio de confianza TP.

La tercera parte también aplica el cifrado de un nivel al identificador SID incluido en sus conjuntos de datos con atributos adicionales. Los conjuntos de datos resultantes, incluido el O-SID como identificador, se transmiten al TP.

El TP ahora realiza una búsqueda en la base de datos en el DB de coincidencia de identificador para encontrar el ST-OO-PID apropiado para el O-SID de cada conjunto de datos recibido de la tercera parte. Después de reemplazar el identificador, el nuevo conjunto de datos, incluido el ST-OO-PID y los atributos adicionales de la tercera parte, se reenvía al agregador de datos DA.

El DA descifra el ST-OO-PID (de acuerdo con la lógica MAP descrita anteriormente) y obtiene el ST-O-PID. En función de este identificador único, el DA puede realizar la combinación con otras fuentes de datos, como datos de eventos de ubicación, datos de clase de clientes o índices a largo plazo.

En la segunda variante (figura 4b) no se desea ocultar la información de la tercera parte del DS. En este caso, la tercera parte podría simplemente transferir sus datos al DS, donde se cifra a través de un MAP de la misma manera que cualquier fuente de datos interna del DS. Después de descifrar el ST-OO-PID en el DA, es posible una

combinación de datos de la manera común.

REIVINDICACIONES

1. Un método para la transmisión de un conjunto de datos desde al menos una primera entidad que suministra el conjunto de datos a al menos una segunda entidad que agrega el conjunto de datos en el que el conjunto de datos incluye al menos un identificador que identifica al menos un usuario de la primera entidad, al menos un atributo de datos de evento que contiene información de eventos del usuario y al menos un atributo de clase de usuario que contiene datos relacionados con el usuario, comprendiendo el método las etapas de:

a. un primer cifrado del al menos un identificador sobre la base de una clave de cifrado que solo la primera entidad conoce para obtener un único identificador cifrado,
b. añadir un número de caracteres aleatorios que definen una cadena de caracteres al identificador cifrado único,
c. un segundo cifrado del resultado de acuerdo con la etapa b. ejecutando un método de cifrado de clave asimétrica para obtener un identificador doble cifrado, y

en el que el método se realiza por separado para los atributos de clase de usuario y los atributos de datos de eventos mediante diferentes procesos, de modo que en un primer proceso el identificador se encripta aplicando las etapas a-c, en el que dicho identificador doble encriptado se transmite en combinación con atributos de datos de eventos a al menos una segunda entidad y en un segundo proceso se cifra el mismo identificador aplicando las etapas a-c y se transmiten junto con los atributos de clase de usuario a la al menos una segunda entidad, en el que ambos procesos se basan en una cadena de caracteres aleatorios diferente pero vuelven a la misma clave de cifrado utilizada para el primer cifrado.

2. El método de acuerdo con la reivindicación 1, en el que la decodificación del identificador doble encriptado en la segunda entidad se realiza decodificando el identificador doble encriptado usando la clave privada respectiva del cifrado asimétrico y eliminando la cadena de caracteres que da como resultado un único identificador cifrado.

3. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que al menos una primera entidad es una red de comunicación móvil.

4. El método de acuerdo con cualquiera de las reivindicaciones 1 a 3, en el que el atributo de datos de evento está relacionado con el tipo de evento y/o la marca de tiempo y/o la ubicación del usuario.

5. El método de acuerdo con cualquiera de las reivindicaciones 1 a 4, en el que el atributo de clase de usuario está relacionado con la edad del usuario y/o el género del usuario.

6. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que el método se realiza por separado para atributos de clase de usuario y atributos de datos de eventos por diferentes instancias de la primera entidad y/o diferentes primeras entidades.

7. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que el al menos un atributo de datos de evento y/o atributo de clase de usuario incluido en el conjunto de datos se transmite en forma de texto sin formato.

8. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que el conjunto de datos se transfiere desde al menos una primera entidad a al menos una segunda entidad a través de una red pública.

9. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que el primer cifrado se basa en un cifrado irreversible.

10. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que la clave utilizada para el primer cifrado varía en su vida útil, por ejemplo, una vida útil a corto y largo plazo, en particular, una vida útil de 24 horas o 1 año, en el que la vida útil seleccionada depende ventajosamente del identificador y/o del atributo cifrado del conjunto de datos.

11. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que se aplica una cadena de caracteres individual que incluye caracteres aleatorios para cada conjunto/evento de datos y/o una determinada cadena de caracteres se aplica a todos los conjuntos/eventos de datos dentro de un cierto intervalo de tiempo, por ejemplo, dentro de 24 h.

12. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que al menos un atributo de datos adicional asociado a al menos un identificador/usuario de la primera entidad es proporcionado por una tercera entidad y asignado e incorporado al conjunto de datos transmitidos por una entidad asociada de confianza que está ubicado en la ruta de transmisión entre la primera y la segunda entidad.

13. El método de acuerdo con la reivindicación 13, en el que la entidad asociada de confianza asigna al menos un atributo de datos adicional al conjunto de datos haciendo coincidir un segundo identificador cifrado insertado con los datos recibidos desde la primera entidad con un segundo identificador cifrado insertado con los datos recibidos desde la tercera entidad en la que ambos segundos identificadores están encriptados con claves idénticas en la primera y segunda entidad.

14. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que uno o más índices estadísticos a largo plazo se derivan de varios conjuntos de datos y/o atributos de datos de eventos recopilados dentro de un cierto intervalo de tiempo.

15. El método de acuerdo con la reivindicación 15, en el que la vida útil de los índices derivados excede la vida útil del número de conjuntos de datos y/o atributos de datos de eventos utilizados para la derivación de los índices.

16. El método de acuerdo con una cualquiera de las reivindicaciones 15 o 16, en el que irregularidades de los conjuntos de datos recopilados se filtran mediante la integración de valores de comparación históricos durante un período de tiempo más largo.

17. Un sistema de comunicación para realizar el método de acuerdo con una cualquiera de las reivindicaciones anteriores.

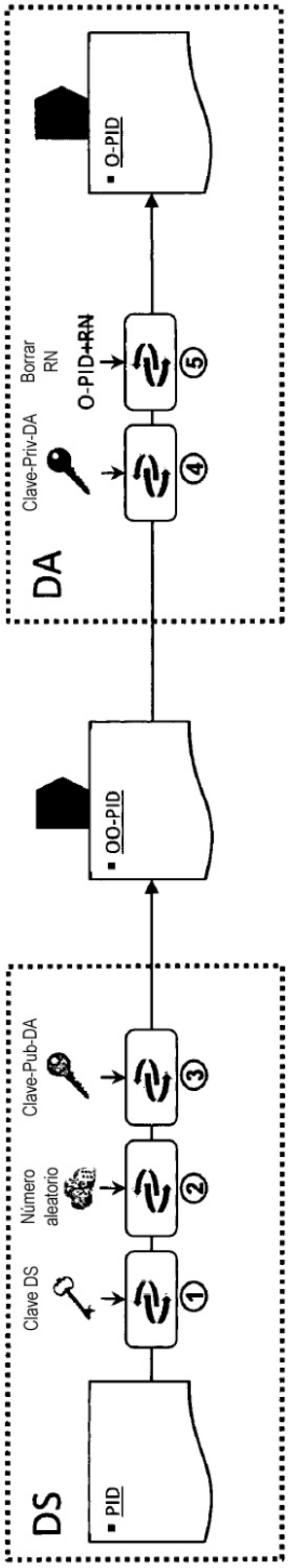

DS
PID
Clave DS
Número aleatorio
Clave-Pub-DA
1
2
3
OO-PID
DA
Clave-Priv-DA
Borrar RN
O-PID+RN
4
5
O-PID

Figura 1

Figura 2a

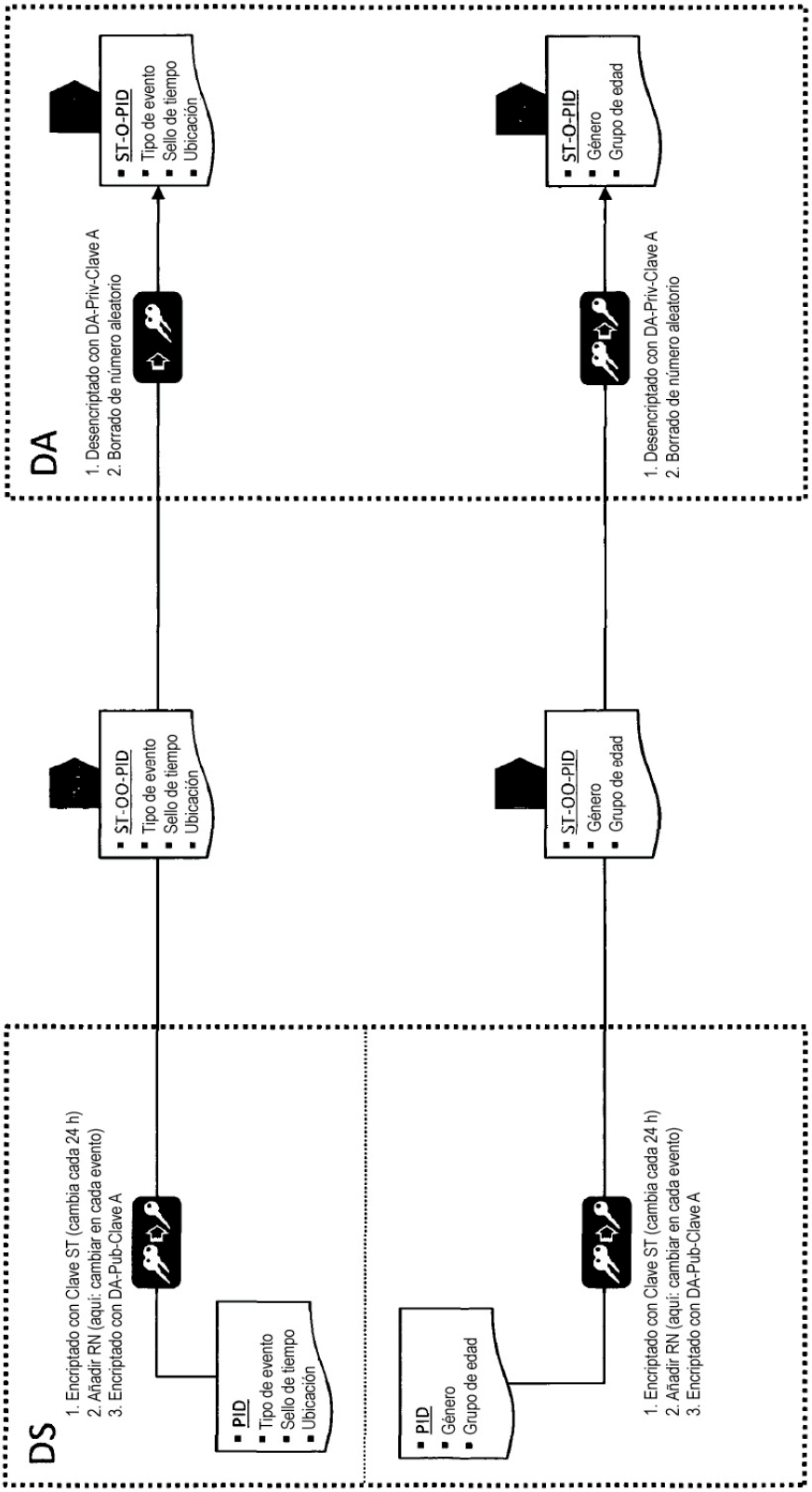

DS
1. Encriptado con Clave ST (cambia cada 24 h)
2. Añadir RN (aquí: cambiar en cada evento)
3. Encriptado con DA-Pub-Clave A
PID
Tipo de evento
Sello de tiempo
Ubicación
ST-OO-PID
Tipo de evento
Sello de tiempo
Ubicación
PID
Género
Grupo de edad
ST-OO-PID
Género
Grupo de edad
1. Encriptado con Clave ST (cambia cada 24 h)
2. Añadir RN (aquí: cambiar en cada evento)
3. Encriptado con DA-Pub-Clave A
DA
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
Tipo de evento
Sello de tiempo
Ubicación
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
Género
Grupo de edad

Figura 2b

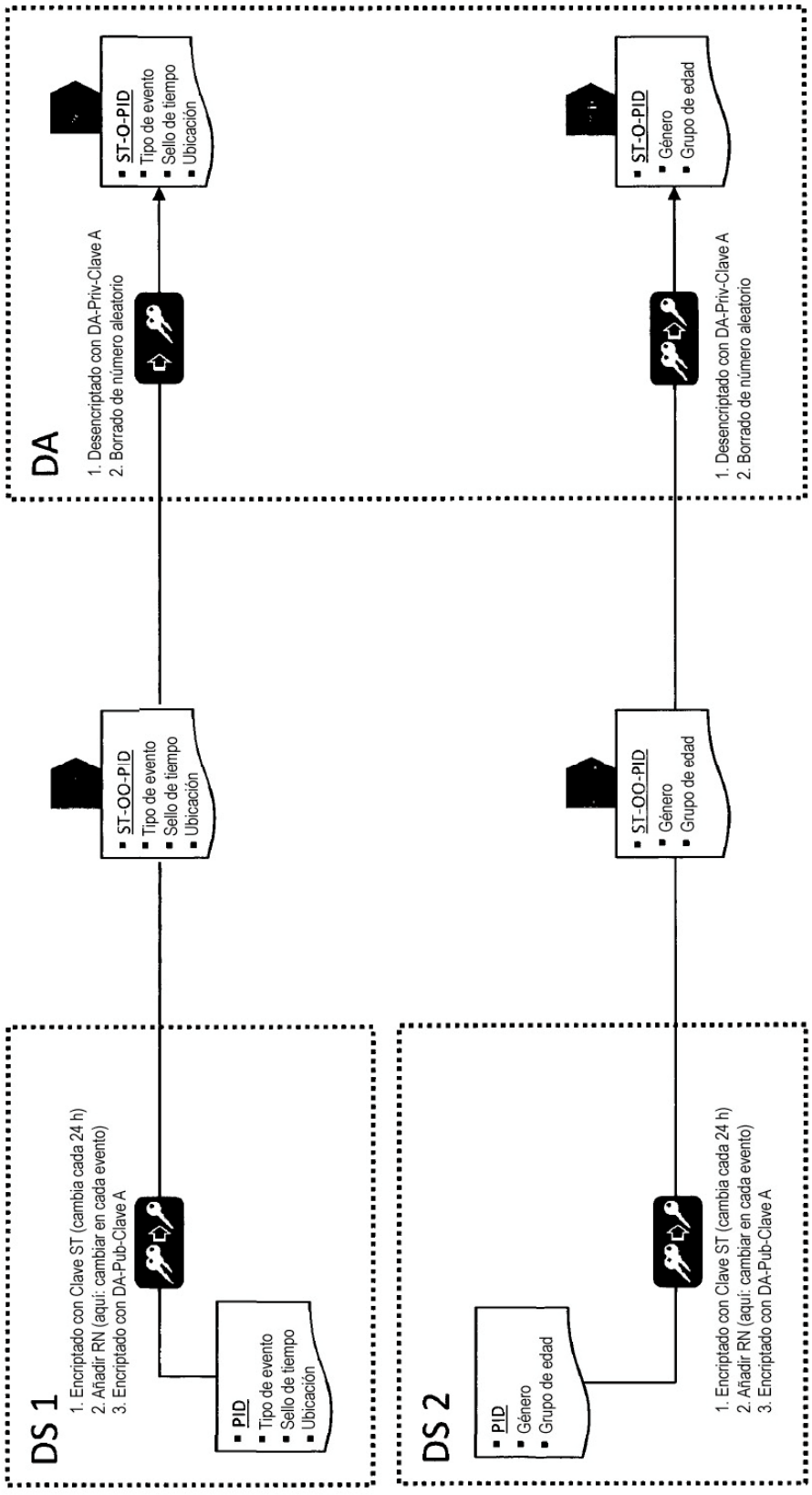

DS 1
1. Encriptado con Clave ST (cambia cada 24 h)
2. Añadir RN (aquí: cambiar en cada evento)
3. Encriptado con DA-Pub-Clave A
PID
Tipo de evento
Sello de tiempo
Ubicación
ST-OO-PID
Tipo de evento
Sello de tiempo
Ubicación
DA
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
Tipo de evento
Sello de tiempo
Ubicación
DS 2
PID
Género
Grupo de edad
1. Encriptado con Clave ST (cambia cada 24 h)
2. Añadir RN (aquí: cambiar en cada evento)
3. Encriptado con DA-Pub-Clave A
ST-OO-PID
Género
Grupo de edad
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
Género
Grupo de edad

Figura 3a

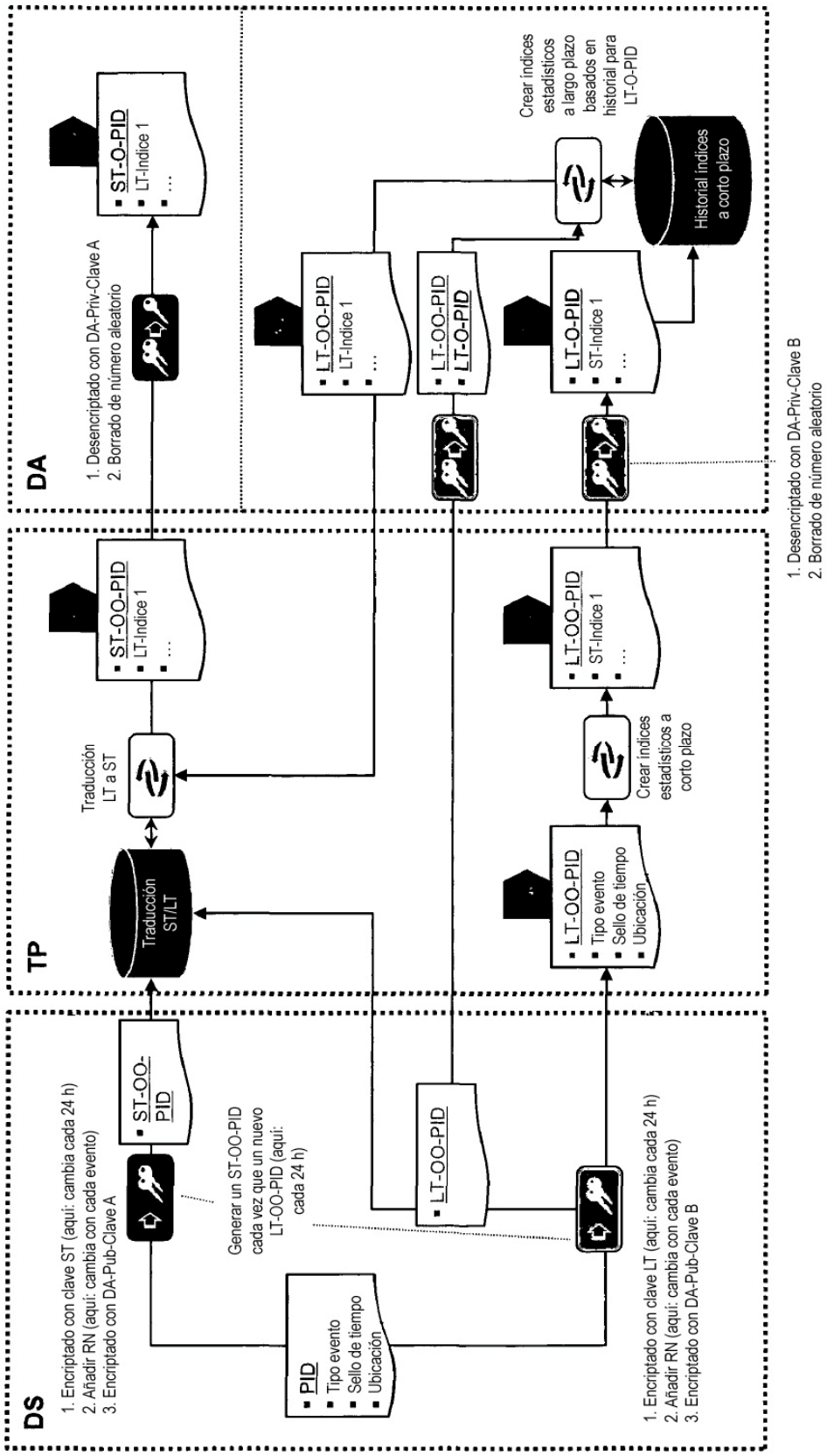

DS
1. Encriptado con clave ST (aquí: cambia cada 24 h)
2. Añadir RN (aquí: cambia con cada evento)
3. Encriptado con DA-Pub-Clave A
ST-OO-PID
PID
Tipo evento
Sello de tiempo
Ubicación
Generar un ST-OO-PID cada vez que un nuevo LT-OO-PID (aquí: cada 24 h)
LT-OO-PID
1. Encriptado con clave LT (aquí: cambia cada 24 h)
2. Añadir RN (aquí: cambia con cada evento)
3. Encriptado con DA-Pub-Clave B
TP
Traducción ST/LT
Traducción LT a ST
ST-OO-PID
LT-Indice 1
...
LT-OO-PID
Tipo evento
Sello de tiempo
Ubicación
Crear índices estadísticos a corto plazo
LT-OO-PID
ST-Indice 1
...
DA
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
LT-Indice 1
...
LT-OO-PID
LT-Indice 1
...
LT-OO-PID
LT-O-PID
LT-O-PID
ST-Indice 1
...
Crear índices estadísticos a largo plazo basados en historial para LT-O-PID
Historial índices a corto plazo
1. Desencriptado con DA-Priv-Clave B
2. Borrado de número aleatorio

Figura 3b

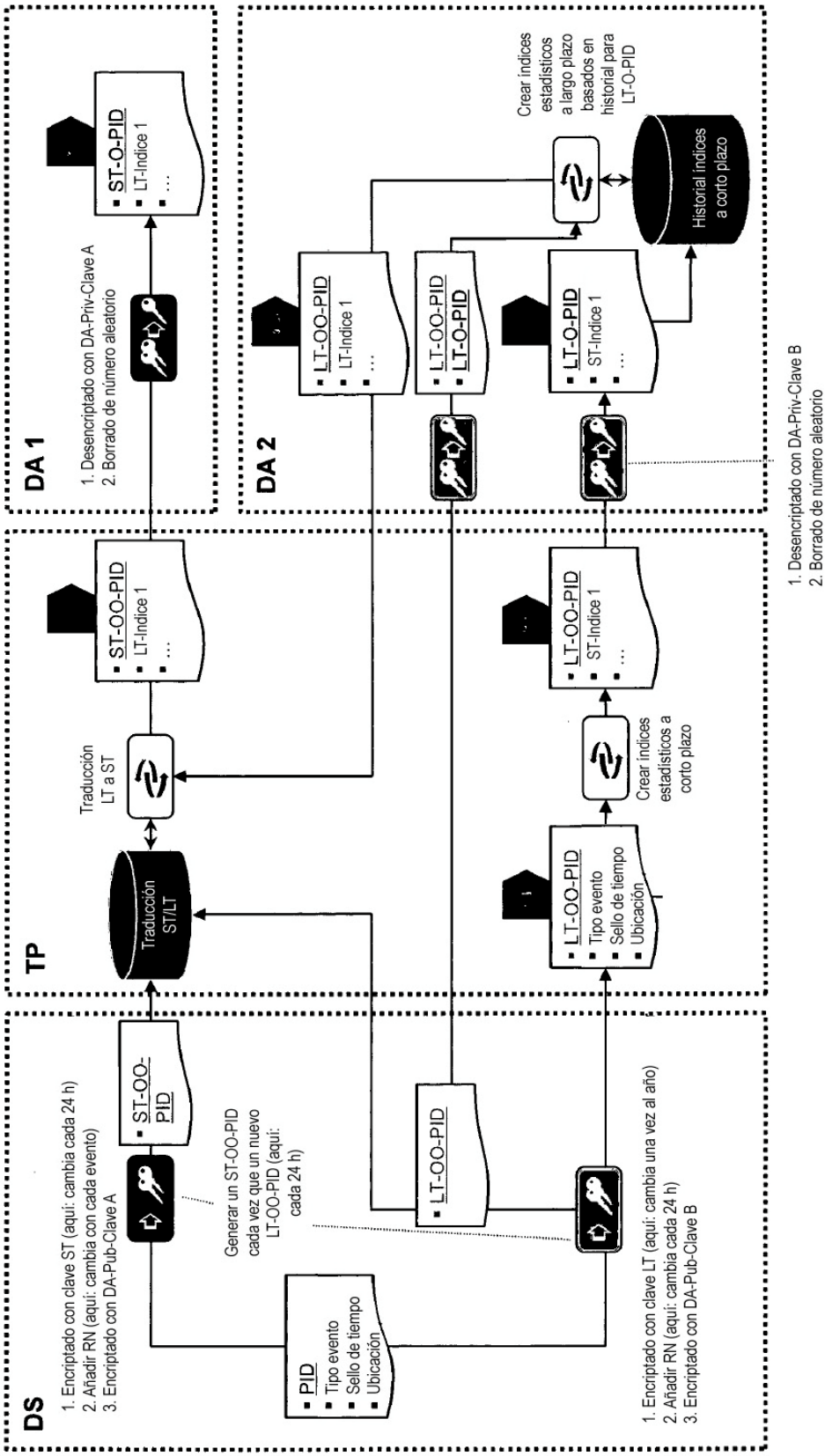

DS
1. Encriptado con clave ST (aquí: cambia cada 24 h)
2. Añadir RN (aquí: cambia con cada evento)
3. Encriptado con DA-Pub-Clave A
ST-OO-PID
PID
Tipo evento
Sello de tiempo
Ubicación
Generar un ST-OO-PID cada vez que un nuevo LT-OO-PID (aquí: cada 24 h)
LT-OO-PID
1. Encriptado con clave LT (aquí: cambia una vez al año)
2. Añadir RN (aquí: cambia cada 24 h)
3. Encriptado con DA-Pub-Clave B
TP
Traducción ST/LT
Traducción LT a ST
ST-OO-PID
LT-Indice 1
...
LT-OO-PID
Tipo evento
Sello de tiempo
Ubicación
Crear índices estadísticos a corto plazo
LT-OO-PID
ST-Indice 1
...
DA 1
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
LT-Indice 1
...
DA 2
LT-OO-PID
LT-Indice 1
...
LT-OO-PID
LT-O-PID
LT-O-PID
ST-Indice 1
...
Crear índices estadísticos a largo plazo basados en historial para LT-O-PID
Historial índices a corto plazo
1. Desencriptado con DA-Priv-Clave B
2. Borrado de número aleatorio

Figura 4a

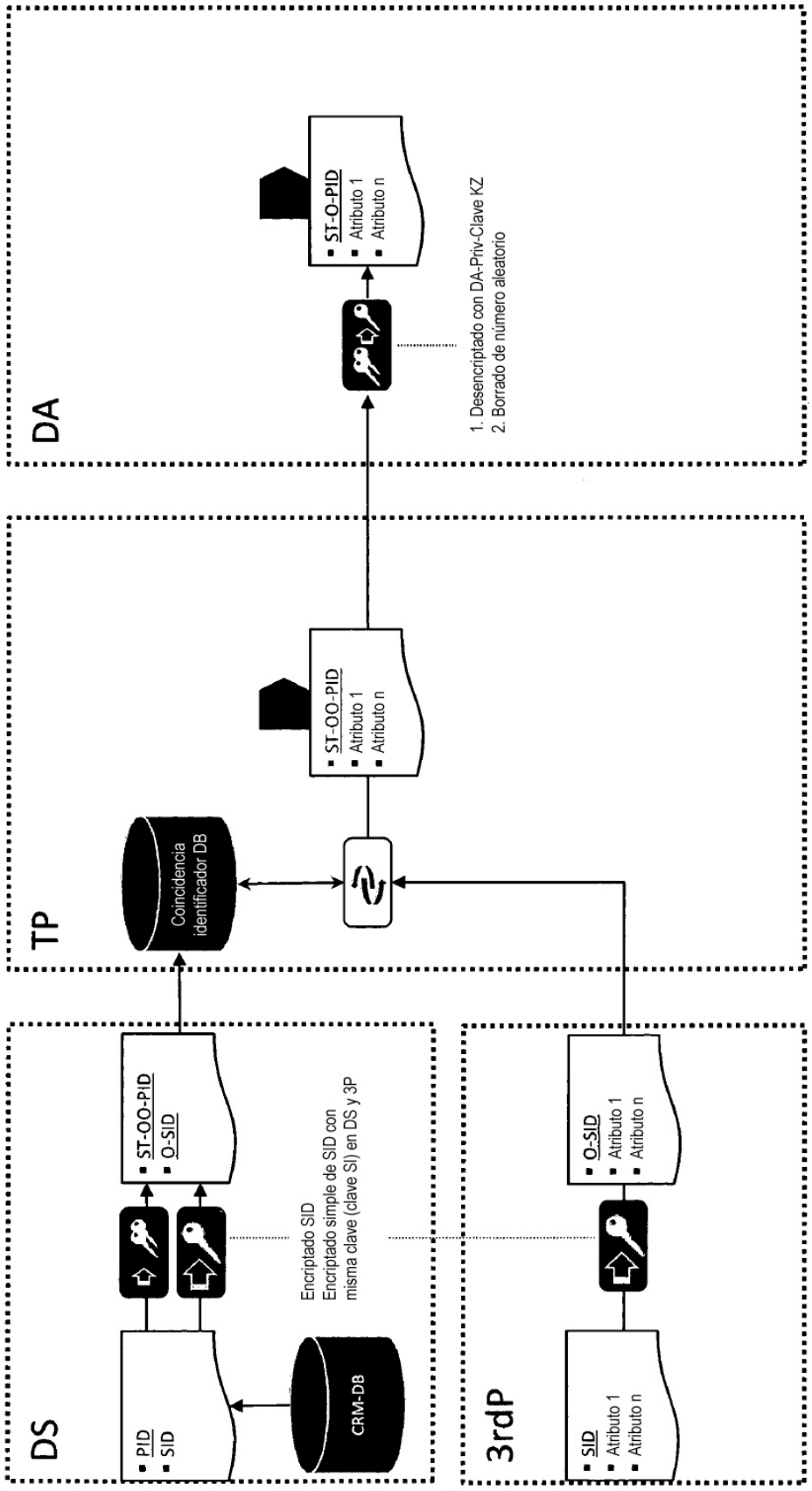
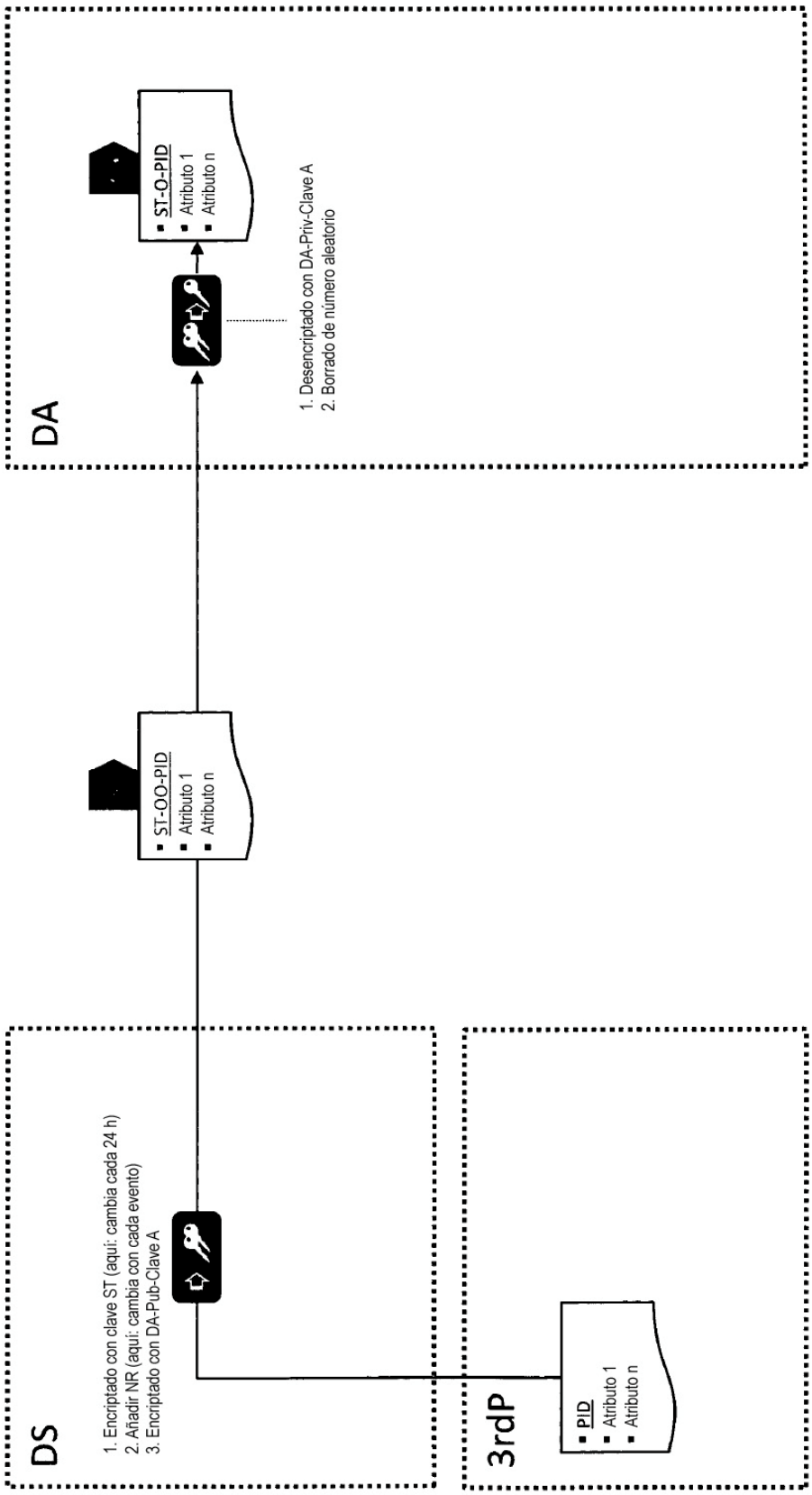

DA
ST-O-PID
Atributo 1
Atributo n
1. Desencriptado con DA-Priv-Clave KZ
2. Borrado de número aleatorio
TP
ST-OO-PID
Atributo 1
Atributo n
Coincidencia identificador DB
DS
PID
SID
ST-OO-PID
O-SID
Encriptado SID
Encriptado simple de SID con misma clave (clave SI) en DS y 3P
CRM-DB
3rdP
O-SID
Atributo 1
Atributo n
SID
Atributo 1
Atributo n

Figura 4b


DS
1. Encriptado con clave ST (aquí: cambia cada 24 h)
2. Añadir NR (aquí: cambia con cada evento)
3. Encriptado con DA-Pub-Clave A
ST-OO-PID
Atributo 1
Atributo n
3rdP
PID
Atributo 1
Atributo n
DA
1. Desencriptado con DA-Priv-Clave A
2. Borrado de número aleatorio
ST-O-PID
Atributo 1
Atributo n