(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **3 009 812**

(51) Int. CI.:

H04W 4/80 (2008.01)
G06N 5/022 (2013.01)
G06N 7/01 (2013.01)
G06N 20/20 (2009.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **05.09.2019 PCT/US2019/049630**

(87) Fecha y número de publicación internacional: **12.03.2020 WO20051265**

(96) Fecha de presentación y número de la solicitud europea: **05.09.2019 E 19858178 (7)**

(97) Fecha y número de publicación de la concesión europea: **11.12.2024 EP 3847508**

(54) Título: **Sistemas y métodos para resoluciones automáticas de señales inalámbricas**

(30) Prioridad:

06.09.2018 US 201862727871 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
01.04.2025

(73) Titular/es:

PWC PRODUCT SALES LLC (100.00%)
300 Madison Avenue
New York, NY 10017, US

(72) Inventor/es:

MARINOVIC, SRDJAN;
STANCULESCU, ERIC;
COHEN, REBECCA E.;
HERRING, KRISTOPHER E.;
MORROW, ANNE E. y
ROGERS, ROBERT T.

(74) Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

DESCRIPCIÓN

Sistemas y métodos para resoluciones automáticas de señales inalámbricas

Campo técnico

Esta solicitud se refiere en general a sistemas y métodos para resoluciones automáticas de señales inalámbricas, y más específicamente a la asignación de significados semánticos a señales inalámbricas.

Antecedentes

Las señales inalámbricas pueden representarse o identificarse mediante una cadena de números y letras que pueden no tener un significado semántico, tales como los nombres de las entidades o los dispositivos que generan las señales inalámbricas. Por ejemplo, las señales inalámbricas pueden identificarse mediante direcciones de código de acceso de máquina (MAC), direcciones de protocolo de Internet (IP) y/o identificadores de conjunto de servicios (SSID), que pueden ser cadenas sin significado semántico. Por lo tanto, incluso si el dispositivo electrónico de un usuario puede recibir/detectar diferentes señales inalámbricas de un entorno circundante, el usuario puede saber poco sobre el entorno y sobre los dispositivos que emiten las señales inalámbricas desde la cadena los números de letras. El dispositivo electrónico del usuario puede conocer simplemente acerca de la información de identificación no semántica de las señales inalámbricas recibidas.

Es técnicamente desafiante resolver las señales inalámbricas asignando significados semánticos a señales inalámbricas. La asignación de los significados semánticos puede incluir determinar el tipo de un dispositivo, negocio asociado con el dispositivo, marca, fabricante, función y entorno en el que se encuentra el dispositivo, y cualquier otra propiedad del dispositivo asociada con las señales inalámbricas. En la actualidad, la mayoría de las relaciones entre nombres de señales inalámbricas (por ejemplo, SSID) y nombres comerciales se resuelven mediante etiquetado manual. El etiquetado manual es lento, ineficiente y engorroso.

En consecuencia, se requiere una mejora significativa sobre la resolución automática implementada por ordenador de señales inalámbricas.

Sumario

Por lo tanto, lo que se desea son sistemas y métodos que resuelvan señales inalámbricas asignando significados semánticos a esas señales inalámbricas. Las realizaciones divulgadas en el presente documento proporcionan soluciones técnicas al problema técnico mencionado anteriormente y pueden proporcionar otras soluciones también. Un servidor analítico puede entrenar una pluralidad de clasificadores inductivos para generar un conjunto respectivo de reglas de deducción (por ejemplo, reglas lógicas probabilísticas y/o modelos de predicción lógica modal). Por ejemplo, el servidor analítico puede entrenar un primer clasificador inductivo para extraer un primer conjunto de reglas de deducción que coinciden con identificadores de conjunto de servicios (SSID) con entidades comerciales. Como otro ejemplo, el servidor analítico puede entrenar un segundo clasificador inductivo para extraer un segundo conjunto de reglas de deducción que coinciden con agrupaciones proximales de señales inalámbricas (también denominadas hipergrupos) con entidades comerciales. Al recibir una señal inalámbrica no resuelta, el servidor analítico puede aplicar al menos una del primer y segundo conjunto de reglas de deducción para asignar un significado semántico a la señal inalámbrica.

En una realización, un método implementado por ordenador comprende, en un modo de procesamiento por lotes, recibir, por un ordenador, datos de observación de una pluralidad de señales inalámbricas observadas por una pluralidad de dispositivos de observador; entrenar, por el ordenador, un primer clasificador inductivo para generar un primer conjunto de reglas deductivas para coincidencias sintácticas entre identificadores de conjunto de servicios de la pluralidad de señales inalámbricas y entidades asociadas con los identificadores de conjunto de servicios; entrenar, por el ordenador, un segundo clasificador inductivo para generar un segundo conjunto de reglas deductivas basándose en la proximidad espacial y la persistencia temporal de la pluralidad de señales inalámbricas; en un modo de deducción en tiempo real: recibir, por el ordenador, información de una señal inalámbrica no resuelta; generar, por el ordenador, una resolución semántica de la señal inalámbrica basada en la aplicación de al menos uno del primer y segundo conjunto de reglas deductivas; y mostrar, por el ordenador, la resolución semántica de la señal inalámbrica en una interfaz gráfica de usuario.

En otra realización, un sistema comprende un medio de almacenamiento no transitorio que almacena una pluralidad de instrucciones de programa informático; y un procesador acoplado eléctricamente al medio de almacenamiento no transitorio y configurado para ejecutar la pluralidad de instrucciones de programa informático para: en un modo de procesamiento por lotes: recibir datos de observación de una pluralidad de señales inalámbricas observadas por una pluralidad de dispositivos de observador; entrenar un primer clasificador inductivo para generar un primer conjunto de reglas deductivas para coincidencias sintácticas entre identificadores de conjunto de servicios de la pluralidad de señales inalámbricas y entidades asociadas con los identificadores de conjunto de servicios; entrenar un segundo clasificador inductivo para generar un segundo conjunto de reglas deductivas basándose en la proximidad espacial y

la persistencia temporal de la pluralidad de señales inalámbricas; en un modo de deducción en tiempo real: recibir información de una señal inalámbrica no resuelta; generar una resolución semántica de la señal inalámbrica basada en la aplicación de al menos uno del primer y segundo conjunto de reglas deductivas; y mostrar la resolución semántica de la señal inalámbrica en una interfaz gráfica de usuario.

Debe entenderse que tanto la descripción general anterior como la siguiente descripción detallada son ilustrativas y explicativas y están destinadas a proporcionar una explicación adicional de la invención según se reivindica. Un ejemplo del estado de la técnica se proporciona por Guo Xianan y col., en un documento de la conferencia del IEEE titulado "ShopProfiler: Profiling Shops with Crowdsourcing Data", publicado en abril de 2024 y que describe datos de entrenamiento de correlación cruzada alimentados en un clasificador SVM, con el fin de inferir reglas que se usarán para perfilar y clasificar lugares de compras. Los datos de entrenamiento comprenden SSID (u otros identificadores de las señales recibidas) junto con valores de intensidad de señal y marcas de tiempo de señal

Breve descripción de los dibujos

Los dibujos adjuntos constituyen una parte de esta memoria descriptiva e ilustran realizaciones de la materia objeto divulgada en el presente documento.

Las Figuras 1A-1C muestran entornos de red ilustrativos para resoluciones automáticas de señales inalámbricas, de acuerdo con una realización.
La Figura 2 muestra un diagrama de flujo ilustrativo para resoluciones automáticas de señales inalámbricas, de acuerdo con una realización.
La Figura 3 muestra una canalización de inducción ilustrativa, de acuerdo con una realización.
La Figura 4 muestra una canalización de deducción ilustrativa, de acuerdo con una realización.
La Figura 5 muestra un proceso de resolución ilustrativo, de acuerdo con una realización.
Las Figuras 6A-6B muestran una aplicación ilustrativa de la resolución de señal, de acuerdo con una realización.
La Figura 7 muestra otra aplicación ilustrativa de resolución de señal, de acuerdo con una realización.

Descripción detallada

A continuación, se hará referencia a las realizaciones ilustradas en los dibujos y aquí, se utilizará un lenguaje específico para describir las mismas. Se entenderá que, a pesar de todo, no por ello se pretende limitar el alcance de las reivindicaciones de esta divulgación. Las alteraciones y modificaciones adicionales de las características inventivas ilustradas en el presente documento, y aplicaciones adicionales de los principios de la materia objeto ilustrada en el presente documento, que se les ocurrirían a un experto en la materia y en posesión de esta divulgación, deben considerarse dentro del alcance de la materia objeto divulgada en el presente documento. La presente divulgación se describe aquí en detalle con referencia a las realizaciones ilustradas en los dibujos, que forman parte del presente documento.

Las realizaciones divulgadas en el presente documento describen sistemas y métodos para asignar significado semántico a señales inalámbricas. En una realización ilustrativa, un servidor analítico puede generar inductivamente un conjunto de reglas de deducción. El servidor analítico puede usar a continuación el conjunto de reglas de deducción para resolver las señales inalámbricas observadas por dispositivos de observador. Más particularmente, el servidor analítico puede generar y entrenar una pluralidad de clasificadores inductivos basándose en la información contenida dentro de las señales observadas por diversos dispositivos de observador. El servidor analítico puede utilizar los clasificadores inductivos para resolver o clasificar las señales entrantes, por ejemplo, para derivar un significado semántico de las señales entrantes.

El servidor analítico puede generar y entrenar clasificadores sintácticos de identificador de conjunto de servicios (SSID) basándose en los SSID en las señales observadas. El servidor analítico puede realizar un análisis de texto de los SSID para asociar el SSID (y en consecuencia las señales inalámbricas) con entidades como empresas y fabricantes. Por ejemplo, el servidor analítico puede asociar señales WiFi como "Wifi gratuito de Joe" y "Gestión segura de Joe" con un negocio denominado "Joe's Bar and Grill". Estos servidores analíticos pueden almacenar estas asociaciones, generalmente generado en el procesamiento por lotes de varios conjuntos de señales inalámbricas, como reglas de deducción para resoluciones de tiempo en tiempo real de señales inalámbricas.

El servidor analítico también puede generar y entrenar un clasificador de estructura de gráfico basándose en la proximidad espacial y la persistencia temporal de las señales inalámbricas observadas. Las señales inalámbricas pueden exhibir diferentes patrones basándose en sus asociaciones con diferentes establecimientos comerciales. Por ejemplo, un dispositivo de observador en un restaurante puede observar señales inalámbricas que permanecen durante un par de horas. En contraste, un dispositivo de observador en un hotel puede observar señales inalámbricas que llegan al momento del registro, se quedan durante la noche y salir a la hora de salida. El servidor analítico puede extraer estos patrones de comportamientos de las señales inalámbricas (o agrupaciones proximales de las mismas) para generar el clasificador de estructura de gráfico.

El servidor analítico también puede generar y entrenar un clasificador sintáctico de control de acceso de máquina

(MAC) basándose en las direcciones MAC de los dispositivos que generan las señales inalámbricas. Los prefijos de las direcciones MAC pueden venderse a granel a los fabricantes y, basándose en estos prefijos, el servidor analítico puede identificar al fabricante del dispositivo. Por ejemplo, los dispositivos Fitbit pueden tener un cierto prefijo MAC y, si el servidor analítico identifica que las señales inalámbricas generadas desde un grupo de dispositivos tienen el mismo prefijo, el servidor analítico puede determinar que estas señales inalámbricas provienen de dispositivos Fitbit.

Debería entenderse que los clasificadores anteriormente mencionados son meramente ilustrativos y el servidor analítico puede generar y entrenar otros clasificadores para extraer diferentes conjuntos de reglas de deducción para resolver señales inalámbricas. El servidor analítico puede generar y entrenar los clasificadores durante un modo de procesamiento por lotes, donde el servidor analítico puede recibir una pluralidad de señales inalámbricas generadas desde una pluralidad de dispositivos. El servidor analítico puede usar las reglas de deducción de los clasificadores para una resolución semántica en tiempo real de señales inalámbricas. Asimismo, debe entenderse que las reglas de deducción como se describen a lo largo de esta divulgación no se limitan a reglas probabilísticas estándar. Otros tipos de reglas de deducción, tales como cualquier modelo de predicción lógica modal estándar, también debe considerarse que están dentro del alcance de esta divulgación.

El servidor analítico también puede distinguir entre señales empresariales y comerciales (por ejemplo, enrutador asociado con una compañía o una ubicación de cadena) y señales personales (por ejemplo, una señal BLE emitida por el rastreador de actividad física de una persona). En particular, el servidor analítico puede incluir un clasificador entrenado para clasificar si una señal observada está asociada con una entidad comercial o una persona. Tal clasificación puede permitir que el servidor analítico realice operaciones de una manera que respete la expectativa de privacidad de los individuos y cumpla con las reglas de privacidad en diferentes jurisdicciones.

La Figura 1A muestra componentes de un entorno de red 100 para resoluciones automáticas de señales inalámbricas. El entorno de red 100 puede incluir un servidor analítico 102 y una base de datos 104 acoplada al servidor analítico 102, en donde el servidor analítico 102 puede recibir información sobre señales inalámbricas (o simplemente señales) detectadas por uno o más dispositivos de observador 106 a través de una red 116 para generar uno o más hipergrupos (también denominados "agrupaciones proximales de dispositivos electrónicos"). El servidor analítico 102 puede recibir información de identificación sobre señales inalámbricas detectadas por los dispositivos de observador 106a, 106b y almacenar la información de identificación en la base de datos 104 para procesamiento adicional.

El servidor analítico 102 puede funcionar como una interfaz para que un administrador establezca ajustes de configuración o proporcione instrucciones operativas a diversos componentes del entorno de red 100. El servidor analítico 102 puede ser cualquier dispositivo que comprenda un componente de comunicaciones capaz de comunicación por cable o inalámbrica con otros componentes del entorno de red 100, y un microprocesador configurado para transmitir y recibir ciertos tipos de datos desde los componentes del entorno de red 100. Ejemplos no limitantes del servidor analítico 102 pueden incluir un ordenador de sobremesa, un ordenador servidor, un ordenador portátil, un ordenador de tableta y similares. Para facilitar la explicación, la Figura 1A muestra un único dispositivo informático que funciona como el servidor analítico 102. Sin embargo, debería apreciarse que algunas realizaciones pueden comprender cualquier número de dispositivos informáticos que funcionan como el servidor analítico 102 y capaces de realizar las diversas tareas descritas en el presente documento.

El servidor analítico 102 puede estar conectado directa o indirectamente a los dispositivos de observador 106a, 106b y la base de datos 104. Por consiguiente, el servidor analítico 102 puede ser capaz de comunicación por cable o inalámbrica a través de una variedad de canales de comunicación con los dispositivos de observador 106a, 106b y la base de datos 104 a través de una red 116. Durante la comunicación por cable o inalámbrica entre el servidor analítico 102, los dispositivos de observador 106a, 106b, y la base de datos 104, cada uno de estos dispositivos puede ser capaz de transmitir y recibir datos entre sí. En algunas realizaciones, cada uno de estos dispositivos puede normalizar y formatear los datos de acuerdo con instrucciones previamente almacenadas antes de transmitir los datos a otros dispositivos. En algunas realizaciones, cada uno de estos dispositivos puede almacenar una copia local de los datos en su memoria antes de transmitir la copia original de los datos a otros dispositivos.

Ejemplos de una red 116 puede incluir, aunque sin limitación, LAN privada o pública, WLAN, MAN, WAN e Internet. La red 116 puede incluir canales de comunicaciones tanto alámbricos como inalámbricos de acuerdo con una o más normas y/o a través de uno o más medios de transporte. La comunicación a través de la red 116 entre los componentes del entorno de red 100 puede realizarse de acuerdo con diversos protocolos de comunicación tales como el protocolo de control de transmisión y el protocolo de internet (TCP/IP), protocolo de datagramas de usuario (UDP) y protocolos de comunicación IEEE. En un ejemplo, la red puede incluir comunicaciones inalámbricas de acuerdo con conjuntos de especificaciones de Bluetooth u otro protocolo de comunicación inalámbrica estándar o patentado. En otro ejemplo, la red también puede incluir comunicaciones a través de una red celular, incluyendo, por ejemplo, un GSM (sistema global para comunicaciones móviles), CDMA (acceso múltiple por división de código) y red EDGE (datos mejorados para la evolución global).

Los dispositivos de observador 106a, 106b puede ser cualquier dispositivo informático y/o de telecomunicaciones que comprenda un procesador y que sea capaz de realizar diversas tareas y procesos descritos en el presente documento. Ejemplos no limitantes de los dispositivos de observador pueden incluir un teléfono 106a (por ejemplo, teléfono

inteligente), un ordenador de usuario 106b (por ejemplo, ordenador de escritorio, un portátil, servidor, tableta), o cualquier otro dispositivo informático o de telecomunicaciones capaz de realizar las diversas tareas y procesos descritos en el presente documento. Para facilitar la explicación, la Figura 1A muestra dos dispositivos que funcionan como los dispositivos de observador 106a, 106b. Sin embargo, debería apreciarse que algunas realizaciones pueden comprender cualquier número de dispositivos de observador capaces de realizar las diversas tareas descritas en el presente documento.

En algunas realizaciones, los dispositivos de observador 106a, 106b pueden ser dispositivos informáticos que funcionan como dispositivos sensores, y están asociados directa o indirectamente con un servidor analítico 102 y/o una base de datos 104. Los dispositivos sensores pueden ser capaces de observar señales en su zona de operación emitidas por diversos dispositivos tales como dispositivos de IoT. El dispositivo sensor puede incluir adicionalmente un procesador de sensor configurado para procesar las señales observadas y extraer información de identificación de las señales observadas. Ejemplos no limitantes de las tecnologías de sensores para los dispositivos sensores pueden incluir sensores LC resonantes, sensores capacitivos y sensores inductivos. Basándose en el tipo particular de las ondas de sensor utilizadas y los protocolos particulares asociados con las ondas de sensor, los dispositivos sensores pueden observar señales y luego generar datos de sensor, que pueden incluir información asociada con las señales observadas. El procesador de sensor puede recibir, interpretar y procesar datos de sensores, que el sensor puede proporcionar a continuación a un procesador del servidor analítico 102 y/o la base de datos 104.

Cada dispositivo de observador puede incluir información de identificación. La información de identificación puede incluir un nombre del dispositivo de observador, un tipo del dispositivo de observador, un número de modelo del dispositivo de observador, una información de ubicación del dispositivo de observador, y un ID del dispositivo de observador donde el ID puede ser un identificador pseudoaleatorio tal como un valor de troceo. En algunos casos, cada dispositivo de observador puede tener múltiples ID y los ID pueden cambiar en cualquier momento. Toda la información de identificación pasada y actual de cada uno del dispositivo de observador puede almacenarse en una base de datos 104. Por ejemplo, un dispositivo de observador dado puede tener un ID antiguo y un ID nuevo, y en tal caso, tanto los ID antiguos como los nuevos pueden almacenarse en la base de datos 104. El servidor analítico 102 puede tener acceso a la información de identificación de cada dispositivo de observador almacenada en una base de datos 104. El servidor analítico 102 puede generar una consulta y/o una solicitud y transmitir la consulta y/o la solicitud en cualquier momento a la base de datos 104 para recibir información de identificación de cualquier dispositivo de observador. En algunos casos, el servidor analítico 102 al recibir datos de señal desde el dispositivo de observador puede consultar la base de datos 104 para recibir información de identificación adicional con respecto al dispositivo de observador desde el que recibió los datos de señal.

El servidor analítico 102 puede establecer ajustes de configuración o proporcionar instrucciones operativas a los dispositivos de observador 106a, 106b para hacer observaciones de señales transmitidas por diversos dispositivos tales como dispositivos de Internet de las Cosas (IoT) y luego proporcionar análisis y datos sobre actividad de aplicación de observación de señal de vuelta al servidor analítico 102. En algunas realizaciones, el servidor analítico 102 puede generar y transmitir las instrucciones operativas a los dispositivos de observador 106a, 106b en cualquier momento para permitir que los dispositivos de observador 106a, 106b para hacer las observaciones de las señales transmitidas por diversos dispositivos tales como dispositivos de IoT y luego proporcionar análisis y datos sobre actividad de aplicación de observación de señal de vuelta al servidor analítico 102. En algunas realizaciones, el servidor analítico 102 puede generar y transmitir las instrucciones operativas a los dispositivos de observador 106a, 106b en cualquier punto de tiempo para deshabilitar a los dispositivos de observador 106 para que no realicen ninguna observación de las señales transmitidas por diversos dispositivos, tales como dispositivos de IoT, y luego notificar la desactivación exitosa de los dispositivos de observador 106a, 106b de vuelta al servidor analítico 102. En algunas realizaciones, el servidor analítico 102 también puede transmitir un enlace web de ajustes de configuración a los dispositivos de observador 106a, 106b, y los dispositivos de observador 106a, 106b pueden usar el enlace web para la instalación de los ajustes de configuración en su hardware y/o software. Los ajustes de configuración pueden habilitar o deshabilitar a los dispositivos de observador 106a, 106b para hacer las observaciones de las señales transmitidas por diversos dispositivos tales como dispositivos de IoT y luego proporcionar análisis y datos sobre actividad de aplicación de observación de señal de vuelta al servidor analítico 102. En algunos casos, los ajustes de configuración pueden habilitar a los dispositivos de observador 106a, 106b para que realicen las observaciones de las señales transmitidas por diversos dispositivos, tales como dispositivos de IoT, durante un período de tiempo limitado (tal como 2 horas al día) en el día, y los mismos ajustes de configuración también pueden deshabilitar a los dispositivos de observador 106a, 106b para que no realicen ninguna observación de las señales durante el resto del día. En algunos casos, los ajustes de configuración pueden deshabilitar a los dispositivos de observador 106a, 106b para que no realicen ninguna observación de las señales cuando su carga de batería esté por debajo de un umbral predeterminado. Para este fin, los ajustes de configuración pueden permitir que el servidor analítico 102 supervise constantemente la carga de la batería de los dispositivos de observador 106a, 106b y cuando la carga de la batería está por debajo de un umbral predeterminado, y entonces el servidor analítico 102 puede deshabilitar a los dispositivos de observador 106a, 106b para que no realicen ninguna observación de las señales. En algunos casos, los ajustes de configuración pueden deshabilitar algunas aplicaciones de los dispositivos de observador 106a, 106b cuando su carga de batería está por debajo de un umbral predeterminado para permitir que los dispositivos de observador 106a, 106b realicen observaciones de las señales. En algunas realizaciones, los ajustes de configuración pueden dar instrucciones a los dispositivos de observador 106a, 106b para enviar al servidor analítico 102 señales asociadas con tipos

específicos de dispositivos o entidades tales como señales asociadas con negocios u otras empresas. Los ajustes de configuración pueden dar instrucciones a los dispositivos de observador 106a, 106b para no enviar señales asociadas con los dispositivos personales de los individuos (por ejemplo, rastreadores de actividad física) al servidor analítico 102.

El servidor analítico 102 puede recibir datos que incluyen señales inalámbricas detectadas por los dispositivos de observador 106a, 106b. En algunas realizaciones, los dispositivos de observador 106a, 106b puede transmitir los datos que incluyen señales observadas al servidor analítico 102 tan pronto como el servidor analítico 102 detecta cualquier señal. En algunas realizaciones, los dispositivos de observador 106a, 106b puede transmitir las señales observadas al servidor analítico 102 después de un período de tiempo predeterminado. Por ejemplo, los dispositivos de observador 106a, 106b pueden programarse para transmitir periódicamente (por ejemplo, diariamente) datos que incluyen todas las señales observadas al servidor analítico 102. En algunas realizaciones, el servidor analítico 102 puede obtener datos que incluyen los datos de señales observadas de los dispositivos de observador 106a, 106b periódicamente (por ejemplo, a diario). En algunas realizaciones, el servidor analítico 102 puede obtener datos que incluyen los datos de señales observadas de los dispositivos de observador 106a, 106b basándose en una condición de activación (por ejemplo, actualizaciones periódicas basadas en el tiempo, actualizaciones en tiempo real). Los datos pueden incluir, pero sin limitación, todas las señales inalámbricas observadas, un punto de tiempo en el que se observó cada señal inalámbrica, unas coordenadas de latitud aproximadas de donde se registra el evento de observación, unas coordenadas de longitud aproximadas de donde se registra el evento de observación, entre otros datos e información de identificación.

El servidor analítico 102 puede almacenar todos los datos, tales como señales inalámbricas observadas, un punto de tiempo en el que se observó cada señal inalámbrica, unas coordenadas de latitud aproximadas de donde se registra el evento de observación, y coordenadas de longitud aproximadas de donde se registra el evento de observación en una base de datos 104 para su procesamiento adicional. En algunas realizaciones, el servidor analítico 102 puede almacenar todos los datos en la base de datos 104 en un formato en el que todos los datos fueron recibidos por el servidor analítico 102. En algunas realizaciones, el servidor analítico 102 puede en primer lugar normalizar y formatear todos los datos, y luego almacenar la versión normalizada y formateada de los datos en la base de datos 104. El servidor analítico 102 puede usar cualquier técnica de normalización y formateo adecuada para normalizar y formatear todos los datos dependiendo del contenido, formato recibido, estructura y tamaño de los datos. Tras la normalización y el formateo de los datos, el servidor analítico 102 puede ejecutar algoritmos tales como algoritmos de agrupamiento para generar uno o más hipergrupos de los conjuntos de datos de señales. Cada hipergrupo puede representar un conjunto de señales que han sido observadas juntas por los dispositivos de observador 106a, 106b dentro de un número de observaciones realizadas por los dispositivos de observador 106a, 106b. En algunos casos, por cada dos observaciones en el hipergrupo, pueden existir al menos dos observaciones superpuestas que contienen dichas dos observaciones.

Como se ilustra en la Figura 1A, un primer dispositivo de observador 106a puede detectar, en el punto_tiempo_1, la *señal_wifi_1* generada por un primer enrutador WiFi 108 y la *señal_wifi_2* generada por un segundo enrutador WiFi 110. Un segundo dispositivo de observador 106b puede detectar, en el punto_tiempo_2, la *señal_wifi_1* generada por el primer enrutador WiFi 108, *la señal_bluetooth_4* generada por el transmisor Bluetooth 114, *señal_BLE*_3 generada por un transmisor Bluetooth de baja energía (BLE) 112. Asimismo, el primer dispositivo de observador 106a puede detectar, en el punto_tiempo_3, la *señal_BLE_3* generada por el transmisor BLE 112. Cada una de las señales mencionadas anteriormente puede incluir una tupla de (*nombre, dirección_MAC, tipo*). Dos señales pueden ser equivalentes de los tres elementos son equivalentes.

Cada dispositivo de observador 106a, 106b puede transmitir a través de la red 116 información de las señales detectadas al servidor analítico 102 para almacenamiento en la base de datos 104 y para análisis adicional. Basándose en la persistencia temporal y la proximidad espacial de las señales observadas por los dispositivos de observador 106a, 106b y recibidas por el servidor analítico 102, el servidor analítico 102 puede definir uno o más hipergrupos (o agrupaciones proximales de dispositivos electrónicos) asociados con la ubicación desde donde se reciben las señales.

El modelo de datos empleado por el servidor analítico 102 para identificar los hipergrupos puede incluir un conjunto de señales S observadas por una población de dispositivos móviles de observador *U*. En el entorno de red ilustrativo 100a, *S*= {*señal_wifi_1, señal_wifi_2, señal_BLE_3, señal_bluetooth_4*} y *U*= {106a, 106b}. Como se ha descrito anteriormente, cada una de las señales en el conjunto de señales S puede incluir una tupla de *(nombre, dirección MAC, tipo).* El servidor analítico 102 puede identificar cada dispositivo de observador 106 con un respectivo identificador de publicidad móvil (o cualquier otro identificador asignado o asociado con la aplicación o dispositivo de observador 106), abreviado como *adid.* El servidor analítico 102 puede asociar cada *adid* de los dispositivos de observador 106a, 106b con una matriz de señal y puntos de tiempo. Más específicamente, el servidor analítico 102 puede construir una matriz booleana dispersa para indicar qué señaliza un observador *adid* observado en una ventana de tiempo dada. En otras palabras, la matriz booleana para el dispositivo de observador 106a, 106b puede indicar una presencia de (indicada por la entrada 1) o ausencia de (indicada por la entrada 0) una o más señales, según lo detectado por el dispositivo de observador 106a, 106b durante un período de tiempo particular. El servidor analítico 102 puede, sin embargo, descartar señales en puntos de tiempo obsoletos según lo informado por los dispositivos de observador 106a, 106b incluso aunque los puntos de tiempo obsoletos pueden no indicar un comportamiento nefasto.

Por ejemplo, si un dispositivo de observador 106a, 106b tiene una única observación que estira la credulidad (umbral establecido en más de cinco días de retraso), el servidor analítico 102 puede simplemente eliminar la observación. En algunas realizaciones, los dispositivos de observador 106a, 106b también pueden transmitir las respectivas coordenadas de latitud y longitud de los dispositivos de observador 106a, 106b.

Basándose en el análisis de las matrices asociadas con los dispositivos de observador 106, el servidor analítico 102 puede generar uno o más hipergrupos basándose en la persistencia temporal y la proximidad espacial de las señales recibidas. La Figura 1B muestra un entorno de red 100b que incluye el hipergrupo 118 generado por el servidor analítico 102 basándose en las señales inalámbricas detectadas por los dispositivos de observador 106. En esta ilustración, el hipergrupo 118 puede contener tres señales inalámbricas: *señal_wifi 1, señal_wifi_ 2, señal_bluetooth*_4. El servidor analítico 102 puede determinar la proximidad espacial de *señal_wifi_1, señal_wifi_2, señal_bluetooth*_4 basándose en el hecho de que estas señales fueron detectadas simultáneamente o casi simultáneamente por cada uno de los dispositivos de observador 106a, 106b. El servidor analítico 102 puede determinar la persistencia temporal de la *señal_wifi_1, señal_wifi_2, señal bluetooth_4* basándose en el hecho de que los dos dispositivos de observador 106a, 106b han observado estas señales en dos puntos de tiempo: el primer dispositivo de observador 106a observó estas señales en el *punto_tiempo*_1 y el segundo dispositivo de observador 106b observó estas señales en el *punto_tiempo*_2. Sin embargo, el servidor analítico 102 puede determinar que la *señal_BLE*_3, a pesar de tener proximidad espacial con la *señal_wifi_1, señal_wifi_2, señal bluetooth 4* puede no tener la persistencia temporal requerida. Por ejemplo, el primer observador 106a no detectó la *señal_BLE*_3 en el *punto_tiempo_1.*

Los dispositivos de observador 106a, 106b pueden conectarse directa o indirectamente al servidor analítico 102 y a una base de datos 104. Por consiguiente, los dispositivos de observador 106a, 106b pueden ser capaces de comunicación por cable o inalámbrica a través de una diversidad de canales de comunicación con el servidor analítico 102 y la base de datos 104 a través de una red 116. Durante la comunicación por cable o inalámbrica entre los dispositivos de observador 106a, 106b, el servidor analítico 102 y la base de datos 104, cada uno de estos dispositivos puede ser capaz de transmitir y recibir datos entre sí. En algunas realizaciones, los dispositivos de observador 106 pueden normalizar y formatear los datos de acuerdo con instrucciones previamente almacenadas antes de transmitir los datos al servidor analítico 102 y/o la base de datos 104. En algunas realizaciones, los dispositivos de observador 106a, 106b pueden almacenar una copia local de los datos en su memoria antes de transmitir la copia original de los datos al servidor analítico 102 y/o la base de datos 104.

El dispositivo de observador 106a, 106b puede configurarse para observar un evento. El evento puede contener todas las señales de que el dispositivo de observador 106a, 106b escanea alrededor de su zona de operación en un punto de tiempo dado. Por consiguiente, el evento puede incluir datos de señal observados y, en algunos casos, el evento también puede incluir valores aproximados o correctos de las coordenadas de latitud de donde el dispositivo de observador 106a registra el evento, 106b en un punto de tiempo dado. En algunos casos, el evento puede incluir además valores aproximados o correctos de las coordenadas de longitud de donde el dispositivo de observador 106a registra el evento, 106b en un punto de tiempo dado.

El evento se provoca cuando el dispositivo de observador 106a, 106b observa señales de diversos dispositivos tales como dispositivos de IoT. Las señales pueden ser una señal electromagnética emitida por los dispositivos de IoT. Se ha de observar que la señal puede ser cualquier tipo de señal emitida por los dispositivos de IoT sin salirse del alcance de las realizaciones desveladas. Las señales observadas por el dispositivo de observador 106a, 106b pueden representar valores discretos sobre las señales. En algunas realizaciones, los valores discretos de las señales pueden caracterizarse por un tipo de señal. El tipo de señal puede incluir, pero sin limitación, una señal Bluetooth®, señal de fidelidad inalámbrica (Wi-Fi) o señales Bluetooth de baja energía (BLE). En algunas realizaciones, los valores discretos de las señales pueden caracterizarse además por un nombre de señal. El nombre de la señal puede ser un SSID (identificador de conjunto de servicios) que identifica el dispositivo de IoT. El SSID puede ser un ID único que consiste en 32 caracteres y se usa para nombrar redes inalámbricas. En algunas realizaciones, los valores discretos de las señales pueden caracterizarse además por una dirección del dispositivo de IoT a través del cual el dispositivo comunica la señal. Cada dispositivo de IoT puede emitir múltiples señales.

Los componentes de red pueden efectuar comunicaciones de señales cableadas y/o inalámbricas a y desde diversos dispositivos. Los componentes de red pueden incluir transmisores, un primer enrutador WiFi 108, un segundo enrutador WiFi 110 y un transmisor Bluetooth de baja energía (BLE) 112. Estos componentes de red pueden ser un componente integrado de un dispositivo electrónico; y, en algunos casos, el componente de red puede conectarse al dispositivo electrónico a través de cualquier medio de comunicaciones alámbrico o inalámbrico. Los componentes de red, como el primer enrutador WiFi 108, el segundo enrutador WiFi 110 y el transmisor Bluetooth de baja energía (BLE) 112 pueden incluir componentes electromecánicos (por ejemplo, procesador, antena) que permiten que los componentes de red comuniquen diversos tipos de datos de señal con uno o más dispositivos electrónicos. En algunas implementaciones, estas señales pueden representar un canal distinto para alojar comunicaciones. Los datos pueden comunicarse usando señales, basándose en protocolos cableados o inalámbricos predeterminados y tecnología de hardware y software asociada. Los componentes de red pueden operar basándose en cualquier número de protocolos de comunicación, tal como Bluetooth®, fidelidad inalámbrica (Wi-Fi) y otros.

Las bases de datos 104 pueden conectarse directa o indirectamente a los dispositivos de observador 106a, 106b y un

servidor analítico 102. Por consiguiente, la base de datos 104 puede ser capaz de comunicación por cable o inalámbrica a través de una variedad de canales de comunicación con los dispositivos de observador 106a, 106b y el servidor analítico 102 a través de una red 116. Durante la comunicación por cable o inalámbrica entre el servidor analítico 102, los dispositivos de observador 106a, 106b, y la base de datos 104, la base de datos 104 es capaz de recibir datos desde el servidor analítico 102 y los dispositivos de observador 106. Los datos pueden incluir, pero sin limitación, todas las señales inalámbricas observadas, un punto de tiempo en el que cada señal inalámbrica fue observada por los dispositivos de observador 106a, 106b, coordenadas de latitud aproximadas de donde se registra el evento de observación por los dispositivos de observador 106a, 106b, coordenadas de longitud aproximadas de donde se registra el evento de observación por los dispositivos de observador 106a, 106b, entre otros datos e información de identificación. Para facilitar la explicación, la Figura 1A muestra una única base de datos 104. Sin embargo, debería apreciarse que algunas realizaciones pueden comprender cualquier número de bases de datos capaces de realizar las diversas tareas descritas en el presente documento.

La base de datos 104 puede tener una construcción lógica de archivos de datos que se almacenan en medios de almacenamiento legibles por máquina no transitorios, como un disco duro o una memoria, controlados por módulos de software de un programa de base de datos (por ejemplo, SQL) y un sistema de gestión de base de datos relacionado (DBMS) que ejecuta los módulos de código (por ejemplo, scripts de SQL) para diversas consultas de datos y otras funciones de gestión generadas por el servidor analítico 102 y los dispositivos de observador 106a, 106b. En algunas realizaciones, una memoria de las bases de datos 104 puede ser un dispositivo de almacenamiento no volátil. La memoria puede implementarse con una unidad de disco magnético, una unidad de disco óptico, un dispositivo de estado sólido o un accesorio a un almacenamiento de red. La memoria puede incluir uno o más dispositivos de memoria para facilitar el almacenamiento y la manipulación del código de programa, un conjunto de instrucciones, tareas, datos, PDK y similares. Ejemplos no limitantes de implementaciones de memoria pueden incluir, aunque sin limitación, una memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), una unidad de disco duro (HDD), una tarjeta digital segura (SD), una memoria magnetorresistiva de lectura/escritura, una memoria óptica de lectura/escritura, una memoria caché o una memoria magnética de lectura/escritura. En algunas realizaciones, una memoria de las bases de datos 104 puede ser una memoria temporal, lo que significa que la finalidad principal de la memoria no es el almacenamiento a largo plazo. Ejemplos de las memorias volátiles pueden incluir memorias dinámicas de acceso aleatorio (DRAM), memorias estáticas de acceso aleatorio (SRAM) y otras formas de memorias volátiles conocidas en la técnica. En algunas realizaciones, la memoria puede configurarse para almacenar mayores cantidades de información que la memoria volátil. La memoria puede configurarse adicionalmente para almacenamiento a largo plazo de información. En algunos ejemplos, la memoria puede incluir elementos de almacenamiento no volátiles. Ejemplos de tales elementos de almacenamiento no volátiles incluyen discos duros magnéticos, discos ópticos, disquetes, memorias flash, o formas de memorias eléctricamente programables (EPROM) o memorias eléctricamente borrables y programables (EEPROM).

En funcionamiento, el servidor analítico 102 puede utilizar la hipergrupo 118 (o agrupación proximal de dispositivos inalámbricos 108, 110, 114) para generar una semántica a una o más señales inalámbricas. Por ejemplo, si el hipergrupo 118 es observado por múltiples dispositivos a lo largo del día y no observado por ningún dispositivo durante la noche, el servidor analítico 102 puede determinar que el hipergrupo 108 puede estar dentro de una oficina y los dispositivos 108, 110, 114 pueden instalarse en una oficina. Si el hipergrupo 118 es observado persistentemente por unos pocos dispositivos, el servidor analítico 102 puede determinar que el hipergrupo 108 puede estar dentro de un hogar y los dispositivos 108, 110, 114 pueden instalarse en un hogar. El servidor analítico 102 también puede distinguir entre señales empresariales y comerciales (por ejemplo, enrutador asociado con una compañía o una ubicación de cadena) y señales personales (por ejemplo, una señal BLE emitida por el rastreador de actividad física de una persona). En particular, el servidor analítico 102 puede incluir un clasificador entrenado para clasificar si una señal observada está asociada con una entidad comercial o una persona. Tal clasificación puede permitir que el servidor analítico 102 realice operaciones de una manera que respete la expectativa de privacidad de los individuos y cumpla con las reglas de privacidad en diferentes jurisdicciones.

La Figura 1C muestra un diagrama de arquitectura ilustrativo 100c de un sistema de resolución automática de señales inalámbricas, de acuerdo con una realización ilustrativa. Los componentes mostrados en el diagrama de arquitectura 100c pueden ser módulos de hardware, módulos de software, o una combinación de módulos de hardware y software. También debe entenderse que los componentes mostrados en el diagrama de arquitectura 100c son meramente ilustrativos y adicionales, alternativos, y debería considerarse un número menor de componentes dentro del alcance de esta divulgación.

Una componente *Buscador de geohash* 120 puede tomar entradas tales como un geohash o un par de latitud/longitud para consultar una base de datos de mapas para recuperar entidades (por ejemplo, negocios) en ubicaciones del geohash de entrada o el par de latitud/longitud. Por ejemplo, una componente de *Descubrimiento de lugar* 142 puede descubrir o identificar un lugar basándose en un átomo de señal recibido de una componente de *Motor de deducción* 140 y proporcionar la información del lugar a la componente de *Buscador de geohash* 120. La componente de *Buscador de geohash* 120 puede usar a continuación la información del lugar para recuperar un geohash o la latitud/longitud para consultar la base de datos de mapas. La componente de *Buscador de geohash* 126 puede proporcionar los resultados de búsqueda a una componente de *Extracción de entidad* 126 que puede mantener registros de entidades asociadas con el geohash o el par de latitud/longitud.

Una componente de *Qué buscar* 122 puede recibir entradas de una componente de *Repositorio de reglas* 138 y/o una componente de *Métricas de Gráfico de Señales* 144. La componente de *Qué buscar* 122 puede determinar, basándose en ponderaciones preconfiguradas, qué tipos de entidades (por ejemplo, negocios) buscar. Por lo tanto, la componente de *Qué buscar* 122 puede permitir que el sistema centre los recursos informáticos en tipos particulares de negocios (por ejemplo, hoteles). Basándose en los tipos de negocios determinados por la componente de *Qué buscar* 122, una componente de *Coincidencia de lugar* 132 puede encontrar señales candidatas y puede intentar clasificar las señales candidatas. En otras palabras, la componente de *Coincidencia de lugar* 132 puede ser un clasificador inductivo mientras que la componente de *Qué buscar* 122 puede proporcionar una configuración a la componente de *Coincidencia de lugar* 132. Una componente de *Panel de control brillante* 124 puede proporcionar una interfaz administrativa para gestionar el sistema de resolución automática de señales inalámbricas. Una componente de *Clasificaciones de gráficos* 128 puede extraer reglas de resolución basándose en la existencia de hipergrupos (o agrupaciones proximales de dispositivos electrónicos). Una componente de *Prefijos de Mac* 130 puede extraer reglas de resolución basándose en las direcciones de control de acceso de máquina (MAC). Una componente de *Co-ocurrencia de SSID* 134 puede agregar las reglas extraídas por las componentes 126-132. Una componente de *Cargar reglas* 136 puede almacenar las reglas de deducción generadas por la componente de *Co-ocurrencia de SSID* 134 a la componente de *Repositorio de reglas* 138. Una componente de *Motor de deducción* 140 puede recuperar las reglas de deducción de la componente de *Repositorio de reglas* 138 para asignar átomos a señales recibidas de una componente de *Descubrimiento del lugar* 142. La componente de *Motor de deducción* 140 puede proporcionar además actualizaciones a la componente de *Métricas de Gráficos de Señales* 144.

La Figura 2 muestra un diagrama de flujo 200 de un método para la resolución semántica automática de señales inalámbricas, de acuerdo con una realización ilustrativa. Otras realizaciones pueden comprender etapas adicionales o alternativas, o pueden omitir algunas etapas por completo. Aunque múltiples sistemas informáticos y bases de datos pueden implementar una o más etapas del método, esta descripción detalla, para más brevedad, un servidor analítico que implementa las diversas etapas del método.

En la etapa 202, el servidor analítico puede recibir una solicitud para una resolución de señales inalámbricas. El servidor analítico puede recibir la solicitud de un dispositivo informático, tal como un servidor, asociado con una entidad cliente. La solicitud puede tener uno o más parámetros, tal como un parámetro que indica un área geográfica de la que tienen que resolverse las señales inalámbricas. El servidor analítico puede recibir la solicitud desde cualquier forma de comunicación por cable o inalámbrica.

En la etapa 204, el servidor analítico puede monitorizar señales inalámbricas detectadas por una pluralidad de dispositivos de observador que pueden ser dispositivos electrónicos configurados para detectar y recibir señales inalámbricas. El servidor analítico puede recopilar las señales inalámbricas detectadas por los dispositivos de observador en diferentes puntos de tiempo periódicamente. Por ejemplo, el servidor analítico puede consultar las señales inalámbricas detectadas desde los dispositivos de observador cada cinco minutos. El servidor analítico puede monitorizar los dispositivos de observador durante una ventana de tiempo predeterminada (por ejemplo, una ventana deslizante). Por ejemplo, el servidor analítico puede monitorizar los dispositivos de observador durante siete días o tres meses. En algunas realizaciones, el servidor analítico puede realizar procesamiento por lotes para entrenar continuamente uno o más clasificadores inductivos sin una solicitud explícita de resoluciones semánticas. Por tanto, el servidor analítico puede generar continuamente conjuntos de reglas de deducción que se usarán para una resolución en tiempo real de las señales inalámbricas.

En otras palabras, el servidor analítico puede monitorizar una población dada de dispositivos de observador. Cada dispositivo de observador puede notificar las señales detectadas al servidor analítico. Sea S un conjunto de señales observadas por la población dada de dispositivos de observador. Como se ha descrito anteriormente, una señal *s* puede ser una tupla (*nombre, dirección MAC, tipo*). El servidor analítico puede considerar que dos señales son equivalentes si los tres elementos para las dos señales son equivalentes. Cada dispositivo de observador puede identificarse con un identificador de publicidad móvil, a veces abreviado como *adid.* Diferente *adid* pueden representar diferentes dispositivos de observador. Cada *adid* puede asociarse con una matriz de señales y puntos de tiempo. Cada fila es una señal en *S*, mientras que los puntos de tiempo *T* son de precisión mínima y pueden cerrarse por una ventana de tiempo dada para el análisis. El servidor analítico puede construir una matriz booleana dispersa, $U \rightarrow Bool^{S \times T}$ para almacenar qué señales el *adid* u móvil ha observado en la ventana de tiempo dada. Si un dispositivo de observador observara una señal *s* a en el momento *t*, el servidor analítico puede establecer el elemento correspondiente en la matriz en 1; de lo contrario, establecer el elemento en 0.

En algunas realizaciones, los dispositivos electrónicos informan puntos de tiempo que pueden volverse obsoletos durante unos pocos días. Independientemente de si esto es indicativo de un comportamiento nefasto, hacer análisis de señal dependiente del tiempo en un dispositivo de observador (por ejemplo, las observaciones de señal del teléfono móvil) pueden ser difíciles para el servidor analítico si sus tiempos son demasiado obsoletos. Si un dispositivo de observador tuviera una observación de señal que estira la credulidad (el umbral establecido en más de cinco días de retraso), el servidor analítico puede eliminar esa observación (por ejemplo, las señales inalámbricas detectadas). En algunas realizaciones, el servidor analítico puede eliminar de la consideración un dispositivo de observador (por ejemplo, dispositivo móvil) con dos o más puntos de tiempo increíbles.

En la etapa 206, el servidor analítico puede analizar las señales inalámbricas recopiladas en diferentes puntos de tiempo desde diferentes dispositivos de observador para generar clasificadores inductivos. Un clasificador inductivo puede producir una asignación de una señal y observación a átomos (por ejemplo, significados sintácticos y/o semánticos). En otras palabras, dada una señal y su historial, el clasificador inductivo puede crear una asignación que describa cómo las señales con patrones similares deben clasificarse en términos de negocios, fabricantes y/u otra información sintáctica y/o semántica. Los clasificadores inductivos pueden, por lo tanto, generar reglas de deducción o reglas de resolución que forman una base de conocimiento para asignaciones de señales inalámbricas a significados sintácticos y/o semánticos.

Para una señal dada dentro de una observación dada, un átomo puede ser una constante que define alguna propiedad intrínseca de la señal dentro de la observación. Una resolución puede ser una función que para una señal dada *s* dentro de una observación dada *o*, asigna un átomo *a* a s. El servidor analítico puede enlazar átomos y señales a través de una relación de primer orden definida como: $S \times O \times Átomos \rightarrow Bool$. Por ejemplo, en una resolución *(s, o, teléfono*), el átomo *teléfono* puede indicar que la señal *s* es emitida por un dispositivo de teléfono inteligente. De manera similar, otra resolución *(s, o, restaurante*) puede indicar que *s* es emitida por enrutadores que están ubicados dentro de un restaurante; y otra resolución más *(s, o, burger_king*) puede indicar que *s* es emitida por enrutadores que están ubicados dentro de un Burger King. Por lo tanto, el servidor analítico puede parametrizar el proceso de resolución mediante una base de conocimiento *K* que contiene una pluralidad de reglas de deducción (clasificadores) que derivan las relaciones entre señales y átomos, como las resoluciones descritas anteriormente.

El servidor analítico puede usar una base de conocimiento *K* definida como un conjunto de reglas lógicas de primer orden para hacer inferencias probabilísticas sobre las relaciones. Formalmente, la resolución puede representarse por la regla de Horn:

$(\forall s.\forall o)\ es\ (s,o,átomo_k) \leftarrow tieneSintaxis(patrón_i,\ s),\ cons_0\ (s,\ o),\ \ldots\ ,\ cons_i\ (s,\ o)$

La relación *tieneSintaxis* indica si una señal s satisface alguna relación sintáctica o semántica con un patrón dado. Puede ser sobre su nombre o su dirección. Las relaciones *Cons* pueden ser restricciones sobre las propiedades de una observación dada.

El servidor analítico puede asignar *átomo* para señalizar *sigif* $K \vDash es\ (sig,\ obs,\ átomo)$, donde $\vDash$ puede ser una inferencia lógica de primer orden (FOL). En algunas realizaciones, el servidor analítico puede ampliar el lenguaje de deducción con la relación *tieneSemántica* que facilitará la asignación de átomos basándose en propiedades gráficas semánticas que exhibe una señal.

El servidor analítico puede extender la lógica de primer orden $\vDash$ relación usando editar distancia de cadena. En particular, una cláusula Horn puede representarse como:

$$< p > es\ (s,o,a) \leftarrow tienePrefijoSintaxis(patrón,s,p),\ldots$$

$$< p > es\ (s,o,a) \leftarrow tieneSufijoSintaxis(patrón,s,p),\ldots$$

donde *tienePrefijoSintaxis* puede satisfacerse si *patrón* es un prefijo de un nombre de señal dado, y *tieneSufijoSintaxis* puede satisfacerse si un *patrón* es un sufijo del nombre de la señal. La variable *p* es entonces la *editar distancia* entre el nombre y el patrón. Por último, $\vDash$ puede ser un *argmax* sobre todas las reglas satisfechas. El servidor analítico puede implementar el sistema deductivo usando una estructura de datos *Tries*, donde el patrón de cada regla puede codificarse como la clave, mientras que la asignación de átomo puede ser el valor correspondiente.

Para generar una base de conocimiento *K* de las reglas de deducción, el servidor analítico puede generar y entrenar una pluralidad de clasificadores inductivos. Tres clasificadores inductivos ilustrativos: un clasificador de SSID, un clasificador de MAC y un clasificador de estructura de gráfico (también denominado clasificador de red) se detallan a continuación.

La Figura 3 ilustra una canalización de inducción 300 ilustrativa implementada por un servidor analítico, de acuerdo con una realización. El servidor analítico puede recibir información de una pluralidad de señales inalámbricas observadas por una pluralidad de dispositivos de observador. Una señal inalámbrica ilustrativa se muestra como 302. En algunas realizaciones, el servidor analítico puede generar un clasificador 304 de identificador de conjunto de servicios (SSID) asignando probabilísticamente uno o más átomos a la señal 302 basándose en el SSID de la señal 302. El clasificador de SSID 304 puede generar, por lo tanto, un conjunto de reglas de deducción basándose en un emparejamiento observado entre la señal 302 y su SSID. En otras palabras, el clasificador de SSID 304 puede determinar una asignación de un SSID de la señal dada 302 a un nombre comercial. Por ejemplo, si la señal 302 es una señal WiFi con un SSID "WiFi libre de Joe", el clasificador de SSID 304 puede determinar que tal señal es de un

negocio llamado "Joe's Bar and Grill". Debería entenderse que el clasificador de SSID 304 puede ser un clasificador sintáctico o semántico.

En algunas realizaciones, el servidor analítico, para generar el clasificador de SSID 304 puede determinar la función de coincidencia de SSID-negocio. En otras palabras, dado un SSID y un conjunto de nombres comerciales, el servidor analítico puede identificar el emparejamiento SSID-negocio correcto.

El servidor analítico puede eliminar primero las palabras claves de los SSID. Las palabras claves pueden ser varias palabras y expresiones que pueden aparecer con frecuencia en los SSID y en los nombres de las empresas. Estas palabras pueden no ser generalmente un identificador único del negocio y pueden servir como una distracción que puede reducir el grado potencial de solapamiento entre un SSID y un negocio que puede estar estrechamente relacionado.

Por ejemplo, puede haber un negocio llamado "Joe's Bar and Grill". Joe's Bar and Grill puede tener dos señales WiFi con los SSID "WiFi gratuito de Joe" y "Gestión Segura de Joe". En este caso, aunque el nombre completo del restaurante es "Joe's Bar and Grill", la palabra "Joe's" puede ser la palabra más significativa en cada uno de los SSID que describen la identidad comercial. "Bar and Grill" puede ser simplemente una descripción del negocio, mientras que "WiFi gratuito" y "Gestión Segura" pueden ser descripciones de las respectivas redes WiFi.

Puede haber formas de usar pistas de contexto para determinar qué fonemas en palabras particulares son los más importantes, pero en esta realización, el servidor analítico puede evitar distracciones comunes filtrando un conjunto de palabras claves que es poco probable que describan la identidad central del negocio. Puede haber tres tipos diferentes de palabras claves que pueden ser útiles para el servidor analítico: palabras claves WiFi generales, palabras claves de tipo empresarial, palabras claves locales.

Las palabras claves de WiFi generales pueden aparecer con frecuencia en señales de WiFi y tienden a describir propiedades de la señal o red. En el ejemplo anterior, "gratuito", "WiFi" "gestión", y "segura" pueden ser ejemplos no limitantes de palabras claves de WiFi. Algunas palabras incluidas en la lista de palabras claves de WiFi pueden ser los nombres de puntos de acceso o proveedores comunes, por ejemplo, "xfinitywifi" puede ser el SSID de los puntos de acceso inalámbricos de un proveedor. Esto puede entenderse con mayor precisión como una resolución distinta, en lugar de una palabra clave, pero desde un punto de vista pragmático, puede ser más eficiente para el servidor analítico filtrar estas resoluciones comunes y dejar una cadena vacía en lugar de intentar resolver señales del diccionario de resolución conocido antes de intentar extraer nuevas relaciones. Las palabras claves de tipo comercial (por ejemplo, "tienda", y "hotel") pueden aparecer tanto en señales inalámbricas como en nombres comerciales. Tales palabras pueden ser útiles para establecer la relación entre una señal y un negocio, pero puede ser insuficiente para demostrar una conexión a menos que haya tokens adicionales en común.

A diferencia de las dos categorías de palabras claves descritas anteriormente, las palabras claves locales pueden no ser un conjunto de datos estático y pueden tener que derivarse en línea para una ubicación específica antes de que el servidor analítico intente extraer resoluciones. Las palabras claves locales pueden ser útiles para identificar diferentes ubicaciones de negocios de cadena. Tales palabras pueden tener el nombre de la ciudad o barrio al final de sus SSID. Por ejemplo, si Joe's Bar and Grill fuera una cadena, puede haber "Joe's Denver", "Joe's Downtown", y/o "Joe's Union Station". Sin embargo, puede haber ocasiones en las que las palabras claves locales puedan ser parte del nombre de una empresa. Por ejemplo, hay una cadena de tiendas de bagels en el área de Washington, DC llamada "Bethesda Bagels". La primera ubicación fue en Bethesda, Maryland, y luego se extendieron a otros lugares de la ciudad. Al filtrar palabras claves locales, el servidor analítico puede correr el riesgo de perder la ubicación insignia mientras identifica correctamente otras ubicaciones. Al igual que con los tipos de negocio, una palabra clave local puede ser una parte importante del nombre de la empresa, y puede no ser ignorada por el servidor analítico. Todavía, la palabra clave local a veces puede ser evidencia insuficiente para establecer una relación entre dos señales.

Las palabras claves generales filtradas por el servidor analítico pueden incluir: "punto de paso", "twcwi", "twc", "att", "acceso a verizonwi", "gratis", "wigratuito", "oficina", "ipad", "iphone", "teléfono", "apple", "xfinitywi ", "linksys", "inalámbrico", "wifi", "wi ", "fi", "pública", "privada", "invitado", "2g", "5g", "g", "ghz", "hz", "2", "4", "5", "red", "nulo", "wlan", "visitante", "corporativo", "directo", "interno", "segura", "exprés", "personal", "empleados", "internet", "xfinidad", "centurylink", "netgear", "ssid", "voip", "nan", "doméstica", "the", "corp", "gestión", "y" "familia" y "súper". Las palabras claves de tipo negocio filtradas por el servidor analítico pueden incluir: "hotel", "restaurante", "cafetería", "pizza", "estadio", "arena", "resort", "supermercado", "colegio", "universidad", "deli", "gourmet" y "bar".

El servidor analítico puede usar el nombre del pueblo o ciudad y cualquier vecindad en las proximidades del punto donde se observó una señal para producir palabras claves locales. El servidor analítico puede obtener nombres de ciudades y pueblos consultando la API de geocodificación inversa de Google Maps (interfaz de programación de aplicaciones). Identificar vecindarios puede ser más desafiante ya que muchos vecindarios pueden ser informales, pueden tener límites mal definidos, a menudo pueden superponerse y pueden tener múltiples abreviaturas o seudónimos. Si el servidor analítico no puede obtener una lista completa de vecindades dentro de un cierto radio de un punto dado, el servidor analítico puede identificar algunas vecindades consultando la API de geocodificación inversa de Google Maps. Debido a que los límites del vecindario pueden ser complejos, el servidor analítico puede derivar los

resultados tomando una muestra aleatoria uniforme de puntos dentro de un círculo y agregando todos los resultados.

Como se ha descrito anteriormente, el servidor analítico puede tomar un conjunto de SSID y un conjunto de nombres comerciales y devolver una lista de coincidencias potenciales entre los SSID y los negocios. En lugar de realizar comparaciones por pares entre cada SSID específico y cada nombre comercial, el servidor analítico puede tomar los nombres y comprimirlos o reducirlos de diferentes maneras. El servidor analítico puede almacenar dos diccionarios: SSID comprimidos y nombres comerciales comprimidos. Para cada uno de los SSID y nombres comerciales, el servidor analítico puede aplicar una o más funciones de compresión y guardar cada versión comprimida como una clave que apunta de vuelta a los SSID y nombres comerciales originales. Una vez que el servidor analítico ha comprimido los SSID y los nombres comerciales, el servidor analítico puede comparar los SSID comprimidos con nombres comerciales comprimidos, encontrar cualquier clave compartida en común y buscar los nombres originales. Este enfoque tiene la ventaja sobre la comparación por pares de que tiene un tiempo de ejecución lineal en relación con el número de nombres enviados, en lugar de un tiempo de ejecución polinómico. Asimismo, también puede ser fácil añadir o eliminar funciones de compresión a medida que el servidor analítico busca patrones específicos en los datos.

La primera etapa para encontrar coincidencias entre SSID y nombres comerciales puede ser identificar coincidencias exactas. Por ejemplo, "Joe's Bar and Grill" puede tener una WiFi con SSID "Joe's Bar and Grill".

En algunas realizaciones, puede que no haya coincidencias exactas, y el servidor analítico puede tener que convertir en tokens los SSID (por ejemplo, las palabras restantes en los SSID después de que se hayan eliminado las palabras claves) y generar SSID convertidos en tokens para su procesamiento adicional. Los nombres comerciales de fuentes externas tienden a seguir un patrón convencional de palabras divididas por espacios. Algunos SSID también pueden seguir este patrón, pero muchos usan otros separadores, tales como guiones o signos de puntuación de subrayado. Uno de los separadores más comunes para los nombres de WiFi es el uso de letras mayúsculas al comienzo de las palabras, tal como "WeWorkGuest".

El servidor analítico puede convertir en tokens los nombres comerciales y los SSID dividiendo en cualquier lugar con un espacio, guion, guion bajo. Si una letra mayúscula sigue a una letra minúscula, el servidor analítico puede dividirse después de la letra minúscula, permitiendo que la mayúscula comience la nueva palabra. Si una letra minúscula sigue a una letra mayúscula, el servidor analítico puede dividirse antes de la letra mayúscula, permitiendo una vez más que la letra mayúscula comience la palabra. Sin embargo, como todas las mayúsculas pueden ser comunes en los SSID, el servidor analítico puede no dividir palabras entre letras mayúsculas a menos que una mayúscula vaya seguida de una minúscula.

El servidor analítico puede usar el mismo algoritmo de conversión en tokens para negocios y SSID, aunque ciertos patrones son más comunes en los SSID. Debido a que puede haber negocios ocasionales cuyos nombres oficiales están en mayúsculas y minúsculas (como WeWork y MakeOffices), el servidor analítico puede convertir en tokens tales nombres comerciales de una manera consistente con los SSID para reducir a las mismas formas estandarizadas. El servidor analítico también puede eliminar cualquier subcadena que contenga cuatro o más caracteres de hexadecimal. Esto puede ser importante porque muchos SSID generados automáticamente contendrán largas cadenas de hexadecimales. Estas cadenas pueden producir coincidencias falsas positivas con otras cadenas.

El servidor analítico también puede realizar una compresión estándar tomando el nombre convertido en token, convertir a minúsculas, eliminar cualquier token que sea una palabra clave y volver a unir los tokens restantes en el orden original. El servidor analítico puede realizar una compresión estándar varias veces usando un conjunto cada vez más amplio de palabras claves: el servidor analítico puede eliminar primero palabras claves WiFi generales únicamente, luego palabras claves locales y finalmente palabras que indican tipos de negocio.

El servidor analítico también puede realizar una función de prefijo de cadena, que puede devolver todos los prefijos de un nombre que tienen por encima de la longitud mínima de prefijo (por ejemplo, satisfacer la longitud mínima de prefijo). El servidor analítico puede ajustar, por defecto, los prefijos de cadena si la cadena está compuesta de múltiples tokens, y el prefijo más corto puede establecerse para que sea el máximo de la longitud del token más corto o cuatro. Esto significa que una cadena de múltiples tokens puede no estar desajustada basándose en las primeras pocas letras de la primera palabra, y una cadena con un primer token muy corto puede requerir aún unos pocos caracteres más para inducir una mayor especificidad.

El uso de prefijos de cadena puede tener una clara ventaja sobre el uso únicamente de prefijos de tokens. En cadenas que tienen solo un token, los prefijos pueden comenzar con la longitud de token por defecto de cuatro. Esto puede ser importante para los SSID, que pueden estar todo en minúsculas y pueden carecer de espaciado o puntuación entre palabras. Por ejemplo, usando la función de prefijo de cadena, el servidor analítico puede coincidir con "villageclub" como el SSID para "Village Club of Sands Points". Antes de calcular los prefijos de cadena, el servidor analítico puede eliminar espacios en blanco y palabras claves desde el comienzo de la cadena. Esto puede asegurar que no todos los negocios que comienzan con "hotel" o "restaurante" o "Nueva York" se identificarán como la misma cosa. Después de calcular cada prefijo de token, el servidor analítico puede comprobar que el prefijo no es una palabra clave antes de guardar. Puede ser importante usar todos los prefijos en lugar de solo el más corto. Si dos nombres comparten las

primeras diez letras, compartirán las primeras nueve letras, y las primeras ocho, por lo que este cálculo puede parecer redundante. Sin embargo, uno de estos prefijados más largos puede ser una coincidencia exacta con un SSID completo, o uno que se ha reducido con alguna otra forma de compresión.

Por defecto, el servidor analítico puede no guardar prefijos de token porque son un subconjunto estricto de prefijos de cadena. Sin embargo, si los prefijos de cadena se establecen en falso, los prefijos de token pueden ser una opción deseada. En comparación con los prefijos de cadena, los prefijos de token pueden tener una mayor precisión a costa de una menor recuperación. Por tanto, los prefijos de token pueden no identificar ciertas relaciones correctas que los prefijos de cadena captarán, pero también será mucho menos probable que produzcan falsos positivos.

Si está habilitado, los prefijos de tokens requerirán un valor predeterminado de dos o más tokens. El servidor analítico puede tomar los dos primeros tokens como un prefijo, luego los tres primeros, y así sucesivamente, hasta que el servidor analítico tenga cada prefijo más corto que el número total de tokens. El servidor analítico puede tomar los prefijos de token, convertirlos a minúsculas y unirlos, sin eliminar las palabras claves. El servidor analítico también puede comprobar para asegurarse de que el prefijo de token completo no está en la lista de palabras claves. Al igual que con los prefijos de cadena, es importante guardar todos los prefijos de tokens, no simplemente el más corto; de lo contrario, el servidor analítico puede perder la oportunidad de hacer coincidir prefijos más largos con resultados de otros algoritmos.

El servidor analítico también puede comparar las iniciales de los SSID y los nombres comerciales. Comparar iniciales puede ayudar a identificar coincidencias que no comparten prefijos o ngramas. Muchos negocios, esencialmente aquellos con nombres largos, pueden usar sus iniciales como un apodo, y muchos SSID incluyen las iniciales de la compañía que los posee. Por ejemplo, algunas ubicaciones de WeWork tienen "ww-miembros" y "ww-invitados" como sus SSID. El algoritmo de iniciales puede convertir en tokens el nombre y concatenar el primer carácter de cada token. Una vez más, las palabras claves pueden interferir cuando se intenta extraer las iniciales adecuadas. Algunas palabras claves pueden ser parte de un acrónimo, mientras que otros pueden no estar excluidos. Por tanto, el servidor analítico puede producir hasta cuatro conjuntos diferentes de iniciales, filtrando palabras claves generales, luego palabras claves locales, luego los tipos de negocio, luego todo lo anterior.

También puede haber una compensación en términos del número de tokens necesarios para producir iniciales significativas. Si hay un único token, esa única inicial puede ser insuficiente para hacer una coincidencia. Reducir el número de tokens necesarios de 3 a 2 puede introducir 3 nuevos positivos verdaderos y 154 nuevos falsos positivos. Por tanto, el servidor analítico puede establecer el número predeterminado de tokens en tres. Sin embargo, en caso de que haya una gran supervisión manual o verificaciones automatizadas adicionales además de las iniciales (por ejemplo, usando la estructura gráfica subyacente), el servidor analítico puede establecer el número de tokens en dos en lugar de tres.

El servidor analítico también puede calcular ngramas de token. Basándose en las señales que se han resuelto, se puede ver que existe un patrón fuerte de que es más probable que se produzcan coincidencias al comienzo de una cadena, mientras que la divergencia se vuelve más frecuente hacia el final. Por tanto, puede desearse el almacenamiento en caché y la comparación de los grandes números de prefijos. Sin embargo, si el servidor analítico se centra exclusivamente en prefijos, el servidor analítico puede perder casos en los que el comienzo no es una coincidencia exacta. Por ejemplo, si un nombre comercial comienza con "The" y el SSID no lo tiene. Si esta es la única diferencia, entonces la compresión estándar eliminaría el "The" como palabra clave e identificaría una coincidencia exacta. Sin embargo, si hay otro token diferente más adelante en el nombre, el servidor analítico puede fallar al identificar la coincidencia.

Para resolver un problema de este tipo, el servidor analítico puede aplicar ngrama. Calcular el ngrama de token es una forma de encontrar coincidencias donde el comienzo está desactivado pero la mayor parte de los nombres coinciden. El servidor analítico puede calcular ngramas de entre dos y cuatro tokens de longitud. Al igual que con los prefijos, el servidor analítico puede mantener palabras claves que pueden ser parte del ngrama, pero si el ngrama completo es una palabra clave, el servidor analítico puede excluir la palabra clave.

Los ngramas de cadenas son la base de la arquitectura de resolución. Las cadenas se dividen en ngramas de entre 3 y 6 caracteres. Para establecer una coincidencia, debe compartirse una fracción suficiente de ngramas entre un SSID y un nombre comercial. Las comparaciones de ngramas de cadena introducen un gran número de resultados falsos positivos, dependiendo de la longitud exacta y el número de ngramas que se utilicen. El servidor analítico puede determinar las longitudes y umbrales óptimos y comparar diferentes enfoques para procesar palabras claves.

El servidor analítico puede usar un modelo de ngrama relativamente laxo diseñado para tolerar una alta tasa de falsos positivos para obtener más coincidencias verdaderas. El servidor analítico puede identificar como una coincidencia cualquier par que haya compartido al menos cuatro ngramas de longitud tres a cinco. Esto significa que dos cadenas cualesquiera que comparten una subsecuencia de seis caracteres se identifican como una coincidencia, o dos cualesquiera con una subsecuencia de cinco caracteres y una subsecuencia de tres caracteres en común, o dos subsecuencias de cuatro caracteres. El servidor analítico puede no filtrar ninguna palabra clave.

Con las etapas implementadas que incluyen coincidencia exacta, compresión estándar, prefijo de cadena, prefijo de token, iniciales y ngramas de token, el servidor analítico puede identificar más resultados positivos verdaderos que un modelo de ngrama relativamente generoso que identifica como una coincidencia cualquier par que comparta al menos cuatro ngramas de longitud tres a seis.

El servidor analítico puede generar y entrenar el clasificador de código de acceso de máquina (MAC) 306 asignando probabilísticamente átomos a la señal observada 302 basándose en la dirección MAC de la señal 302 para generar otro conjunto de reglas de deducción. El clasificador de MAC 306 puede asignar el fabricante del dispositivo que emite la señal 302 basándose en el identificador único de organización (OUI) de la dirección de MAC. Por ejemplo, el clasificador de MAC 306 puede determinar que la señal 302 con OUI a4:77:33 es de un dispositivo de Google.

El servidor analítico puede entrenar el clasificador de estructura de gráfico 308 (también denominado clasificador de análisis de red) para asignar átomos a la señal 302 basándose en un gráfico de red de la señal 302, tal como un hipergrupo de la señal 302. En otras palabras, el clasificador de estructura de gráfico 308 puede considerar la proximidad espacial y la persistencia temporal de una o más señales alrededor de la señal 302 para determinar los átomos y generar otro conjunto más de reglas deductivas. Por ejemplo, la señal 302, si desde un hotel puede observarse por muchos dispositivos diferentes, mientras que si desde una vivienda solo puede observarse por un número limitado de dispositivos. En otras palabras, el clasificador de estructura de gráfico 308 puede extraer un conjunto de reglas de deducción basándose en las agrupaciones proximales de las señales inalámbricas.

El servidor analítico puede ejecutar un proceso de conjunto 302 que reúne las diferentes reglas del clasificador sintáctico de SSID 304, clasificador sintáctico de MAC 306, clasificador de estructura de gráficos 308 en un proceso de conjunto 310 para producir un conjunto agregado de reglas de deducción 312 a partir de los conjuntos de reglas de deducción anteriormente mencionados. Asimismo, debe entenderse que las reglas de deducción 312 y como se describe a lo largo de esta divulgación no se limitan a reglas probabilísticas estándar. Otros tipos de reglas de deducción 312, tal como cualquier modelo de predicción lógica modal estándar, también debe considerarse que está dentro del alcance de esta divulgación.

En la etapa 208, el servidor analítico puede usar un motor de deducción (también denominado como un agente de deducción) para clasificar una señal dada aplicando las reglas de deducción. Las reglas de deducción pueden asignarse al patrón de señal dado dentro de algunos átomos relacionados con el patrón de observación. Para aplicar reglas de deducción a una señal, el servidor analítico puede tener que extraer información de la señal, tal como, SSID, prefijo MAC. El servidor analítico puede usar el motor de deducción para una resolución en tiempo real de la señal inalámbrica dada.

La Figura 4 ilustra una canalización de deducción ilustrativa 400, de acuerdo con una realización. El servidor analítico puede recibir una observación en tiempo de ejecución 402 desde un dispositivo de observador. La observación en tiempo de ejecución 402 puede comprender una o más señales inalámbricas detectadas/observadas por el dispositivo de observador. El servidor analítico puede ejecutar un proceso de extracción de patrones 404 para obtener el SSID, OUI de dirección MAC y/o la estructura gráfica asociada con las señales. El servidor analítico puede ejecutar un proceso de deducción 408 para aplicar reglas de deducción 406 para determinar los átomos 410 para las una o más señales inalámbricas en la observación de tiempo de ejecución 402.

Resolución de red ilustrativa

El objetivo de una resolución de red puede ser introducir un identificador único de organización (OUI) y devolver un vector de probabilidad para diferentes resoluciones. El vector de probabilidad puede incluir probabilidades correspondientes de que una o más entidades estén asociadas con una señal inalámbrica. Un identificador único organizativo (OUI) es un número de 24 bits que identifica de manera única a un proveedor, fabricante u otra organización. Los OUI pueden usarse como una primera porción de identificadores derivados para identificar de manera única una pieza particular de equipo como direcciones MAC. En las direcciones MAC, el OUI se combina con un número de 24 bits (asignado por el propietario o 'cesionario' del OUI) para formar la dirección. Los primeros tres octetos de la dirección son el OUI.

Por ejemplo, dado el OUI de a4:77:33, un vector de resolución puede ser como sigue:

	compañía	probabilidad
resolución(a4:77:33)	Google	0,544809
	Frontier Silicon	0.141.183
	Samsung	0,067101
	Bose	0,042528
	Denon	0,035264
	...	...

Como se muestra en el ejemplo, un servidor analítico puede determinar que el OUI a4:77:33 es un OUI de Google, porque la probabilidad de que el OUI pertenezca a Google es la más alta.

El servidor analítico puede hacer un gráfico bipartito entre los OUI y las palabras que aparecen en los SSID basándose en el número de direcciones MAC distintas desde su intersección, tomar palabras que aparecen en los SSID de la lista de SSID resueltos, y tratar la palabra como resuelta si solo aparece en los SSID resueltos de una compañía.

La Figura 5 ilustra un proceso de resolución de una red 500, de acuerdo con una realización. Como se muestra en la figura, puede haber tres tipos de nodos en la red 500: un nodo de SSID 502, un nodo de OUI 504 y un nodo de palabra 506. En algunas realizaciones, el nodo de SSID 502 puede incluir SSIDS resueltos (por ejemplo, emparejados con una empresa y/o un fabricante). El nodo de palabra 506 puede incluir palabras que son subcadenas de SSID que han sido identificadas por un servidor analítico como similares a palabras. Los bordes dirigidos 508, 510, 512, 514 (también mostrados como diagramas de Venn) pueden medir la influencia de un nodo al otro, basándose en frecuencias conjuntas de los contenidos de los nodos asociados con los nodos.

El borde dirigido 508 puede mostrar la influencia del nodo de SSID 502 al nodo de palabra 506. Por ejemplo, el servidor analítico puede determinar que el 40 % de las veces cuando se identifica la palabra "Serie", se identifica en el SSID "Samsung Serie 8". El servidor analítico puede indicar entonces que el peso del borde desde un nodo de SSID 502 que comprende "Samsung Serie 8" a un nodo de palabra 506 que comprende la palabra "Serie" puede ser 0,4. El borde dirigido 510 puede mostrar la influencia del nodo de SSID 502 al nodo de OUI 506. Por ejemplo, el servidor analítico puede determinar el 10 % del tiempo cuando se identifica un OUI; el OUI se empareja con un SSID específico. El servidor analítico puede indicar entonces que el peso de borde 510 puede ser 0,1. El borde dirigido 512 puede indicar una frecuencia en la que un SSID emparejado con un OUI contiene una palabra, cuando el SSID aún no es un nodo resuelto, a través de la frecuencia total en la que se identifica la palabra. El borde dirigido 514 puede indicar una frecuencia en la que un SSID emparejado con un OUI contiene una palabra, cuando el SSID aún no es un nodo resuelto, a través de la frecuencia total para el OUI.

El servidor analítico puede asociar cada uno de los nodos 502, 504, 506 con un vector de resolución. Cuando el servidor analítico inicia la red 500, todos esos vectores pueden ser vectores cero excepto para los nodos de SSID (resueltos), que pueden ser todos 0 excepto uno único que representa una resolución. El servidor analítico puede atravesar la red a través de una serie de etapas, para calcular vectores para todas las palabras y OUI en los vectores en la etapa anterior.

El servidor analítico puede calcular las resoluciones como sigue:

$$vector\ de\ resolución\ de\ nodo = \sum_{bordes\ entrantes} (peso\ de\ borde) \cdot (vector\ de\ nodo\ de\ vecindario)$$

De esta forma, la resolución de los SSID resueltos puede fluir a través de la red, resolver palabras y OUI. Con el tiempo, todos los nodos pueden tener vectores cuyos valores suman uno.

Las realizaciones divulgadas en el presente documento pueden utilizarse para correcciones de lugar y descubrimiento basándose en agrupaciones proximales de señales asociadas con los lugares. Pueden desearse correcciones de lugar cuando un mapa digital convencional incluye múltiples entradas del mismo lugar. Las correcciones de lugar pueden desearse adicionalmente para identificar negocios cercanos recientes o verificar negocios introducidos manualmente. Puede desearse el descubrimiento de lugares para localizar negocios recién establecidos.

La Figura 6A muestra un mapa digital 600a que contiene tres ubicaciones posibles para un restaurante denominado Zocalo. Sin embargo, basándose en una distribución de señales asociadas con el restaurante, un servidor analítico puede identificar una ubicación correcta del restaurante como se muestra en un mapa digital actualizado 600b en la Figura 6B. La ubicación correcta del restaurante se muestra en un círculo en el mapa digital actualizado 600b.

La Figura 7 muestra un mapa digital 700 de una ciudad. Como se muestra, un servidor analítico que usa las realizaciones divulgadas en el presente documento puede descubrir un restaurante The Way Back recientemente abierto basándose en las observaciones de las señales WiFi del restaurante, tales como "thewayback", "Wayback Invitados". La ubicación descubierta del restaurante se muestra que está dentro de un círculo del mapa digital 700.

Las descripciones de método anteriores y los diagramas de flujo de proceso se proporcionan simplemente como ejemplos ilustrativos y no pretenden requerir o implicar que las etapas de las diversas realizaciones deben realizarse en el orden presentado. Las etapas en las realizaciones anteriores pueden realizarse en cualquier orden. Palabras como "después", "siguiente", etc. no pretenden limitar el orden de las etapas; estas palabras se usan simplemente para guiar al lector a través de la descripción de los métodos. Aunque los diagramas de flujo de proceso pueden describir las operaciones como un proceso secuencial, muchas de las operaciones se pueden realizar en paralelo o simultáneamente. Además, el orden de las operaciones puede reorganizarse. Un proceso puede corresponder a un método, una función, un procedimiento, una subrutina, un subprograma y similares. Cuando un proceso corresponde a una función, la terminación del proceso puede corresponder a un retorno de la función a una función de llamada o una función principal.

Los diversos bloques lógicos, módulos, circuitos y etapas de algoritmo ilustrativas descritos en relación con las realizaciones divulgadas en el presente documento pueden implementarse como hardware electrónico, software informático o combinaciones de ambos. Para ilustrar claramente esta intercambiabilidad de hardware y software, diversos componentes, bloques, módulos, circuitos y etapas ilustrativas se han descrito anteriormente en general en términos de su funcionalidad. Si dicha funcionalidad se implementa como hardware o software depende de la aplicación particular y las restricciones de diseño impuestas en el sistema global. Los expertos en la materia pueden implementar la funcionalidad descrita de diversas maneras para cada aplicación particular, pero tales decisiones de implementación no deben interpretarse como que provocan una desviación del alcance de esta divulgación o de las reivindicaciones.

Las realizaciones implementadas en software informático pueden implementarse en software, firmware, software intermedio, microcódigo, lenguajes de descripción de hardware o cualquier combinación de los mismos. Un segmento de código o instrucciones ejecutables por máquina pueden representar un procedimiento, una función, un subprograma, un programa, una rutina, una subrutina, un módulo, un paquete de software, una clase, o cualquier combinación de instrucciones, estructuras de datos o declaraciones del programa. Un segmento de código puede acoplarse a otro segmento de código o un circuito de hardware pasando y/o recibiendo información, datos, argumentos, parámetros o contenidos de memoria. La información, argumentos, parámetros, datos, etc. pueden pasar, reenviarse o transmitirse a través de cualquier medio adecuado, incluido el uso compartido de memoria, paso de mensajes, paso de tokens, transmisión de red, etc.

El código de software real o hardware de control especializado usado para implementar estos sistemas y métodos no es limitante de las características reivindicadas o de esta divulgación. Por tanto, la operación y el comportamiento de los sistemas y métodos se describieron sin referencia al código de software específico, entendiéndose que el software y el hardware de control pueden diseñarse para implementar los sistemas y métodos basándose en la descripción en el presente documento.

Cuando se implementan en software, las funciones pueden almacenarse como una o más instrucciones o código en un medio de almacenamiento no transitorio legible por ordenador o legible por procesador. Las etapas de un método o algoritmo divulgado en el presente documento pueden incorporarse en un módulo de software ejecutable por procesador, que puede residir en un medio de almacenamiento legible por ordenador o legible por procesador. Un medio legible por ordenador o legible por procesador no transitorio incluye tanto medios de almacenamiento informático como medios de almacenamiento tangibles que facilitan la transferencia de un programa informático de un lugar a otro. Un medio de almacenamiento legible por procesador no transitorio puede ser cualquier medio disponible al que pueda acceder un ordenador. A modo de ejemplo, y sin limitación, tales medios legibles por procesador no transitorios pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento en disco óptico, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio de almacenamiento tangible que pueda usarse para almacenar el código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse por un ordenador o procesador. Disco, como se usa en el presente documento, incluye disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), disquete y disco Blu-ray donde los discos generalmente reproducen datos magnéticamente, mientras que los discos reproducen datos ópticamente con láseres. Las combinaciones de los anteriores deben incluirse también dentro del alcance de los medios legibles por ordenador. Adicionalmente, las operaciones de un método o algoritmo pueden residir como una o cualquier combinación o conjunto de códigos y/o instrucciones en un medio legible por procesador y/o medio legible por ordenador no transitorio, que puede incorporarse en un producto de programa informático.

La descripción precedente de las realizaciones divulgadas se proporciona para permitir que cualquier persona experta en la materia haga o use las realizaciones descritas en el presente documento y sus variaciones. Diversas modificaciones a estas realizaciones serán fácilmente evidentes para los expertos en la materia, y los principios genéricos definidos en el presente documento se pueden aplicar a otras realizaciones sin apartarse del espíritu o alcance de la materia objeto divulgada en el presente documento. Por tanto, la presente divulgación no pretende limitarse a las realizaciones mostradas en el presente documento, sino que debe otorgarse el alcance más amplio consistente con las siguientes reivindicaciones y los principios y características novedosas divulgadas en el presente documento.

Si bien se han desvelado diversos aspectos y realizaciones, se contemplan otros aspectos y realizaciones. Los

diversos aspectos y realizaciones divulgados tienen la finalidad de ilustrar y no pretenden ser limitantes, con el verdadero alcance indicado por las siguientes reivindicaciones.

REIVINDICACIONES

1. Un método implementado por ordenador que comprende:

en un modo de procesamiento por lotes:

recibir, por un ordenador, datos de observación de una pluralidad de señales inalámbricas observadas por una pluralidad de dispositivos de observador;
entrenar, por el ordenador, un primer clasificador inductivo para generar un primer conjunto de reglas deductivas para coincidencias sintácticas entre identificadores de conjunto de servicios de la pluralidad de señales inalámbricas y entidades asociadas con los identificadores de conjunto de servicios;
entrenar, por el ordenador, un segundo clasificador inductivo para generar un segundo conjunto de reglas deductivas basándose en la proximidad espacial y la persistencia temporal de la pluralidad de señales inalámbricas, en donde el segundo clasificador inductivo está entrenado para asignar un átomo, que es una constante que define alguna propiedad intrínseca de la señal dentro de una observación dada, a una señal inalámbrica dada basándose en un gráfico de red de la señal inalámbrica dada;

en un modo de deducción en tiempo real:

recibir, por el ordenador, información de una señal inalámbrica no resuelta;
generar, por el ordenador, una resolución semántica de la señal inalámbrica basada en la aplicación de al menos uno del primer y segundo conjunto de reglas deductivas; y
mostrar, por el ordenador, la resolución semántica de la señal inalámbrica en una interfaz gráfica de usuario.

2. El método implementado por ordenador de la reivindicación 1, que comprende además:

en el modo de procesamiento por lotes:
entrenar, por el ordenador, un tercer clasificador inductivo para generar un tercer conjunto de reglas deductivas para coincidencias sintácticas entre códigos de acceso de máquina de dispositivos que generan la pluralidad de señales inalámbricas y fabricantes de los dispositivos; y
en el modo de deducción en tiempo real:
generar, por el ordenador, la resolución semántica de la señal inalámbrica basada en la aplicación de al menos uno del primer, segundo y tercer conjunto de reglas deductivas.

3. El método implementado por ordenador de la reivindicación 1, en donde la resolución semántica incluye al menos uno de un fabricante de un dispositivo que genera la señal inalámbrica o una entidad asociada con la señal inalámbrica.

4. El método implementado por ordenador de la reivindicación 3, en donde la resolución semántica incluye una puntuación de probabilidad que indica una probabilidad de que la señal inalámbrica esté asociada con el fabricante o la entidad.

5. El método implementado por ordenador de la reivindicación 3, que comprende además:
asociar, por el ordenador, la resolución semántica de la señal inalámbrica con una ubicación.

6. El método implementado por ordenador de la reivindicación 5, que comprende además:
mostrar, por el ordenador, la resolución semántica de la señal inalámbrica y la ubicación asociada dentro de un mapa digital en la interfaz gráfica de usuario.

7. El método implementado por ordenador de la reivindicación 1, en donde el entrenamiento del primer clasificador inductivo comprende:

retirar, por el ordenador, palabras claves de los identificadores de conjunto de servicios de la pluralidad de señales inalámbricas;
convertir en tokens, por el ordenador, las palabras restantes de los identificadores de conjunto de servicios de la pluralidad de señales inalámbricas para generar identificadores de conjunto de servicios convertidos en tokens; y
entrenar, por el ordenador, el primer clasificador inductivo utilizando los identificadores de conjunto de servicios convertidos en tokens.

8. El método implementado por ordenador de la reivindicación 7, en donde el entrenamiento del primer clasificador inductivo comprende:

identificar, por el ordenador, una pluralidad de tokens de tipo palabra en los identificadores de conjunto de servicios convertidos en tokens; y
asociar probabilísticamente, por el ordenador, una o más palabras que corresponden a uno o más tokens similares a palabras para cada identificador de conjunto de servicios.

9. El método implementado por ordenador de la reivindicación 8, en donde la generación de la resolución semántica de la señal inalámbrica comprende:
identificar, por el ordenador, al menos uno de un fabricante de un dispositivo que genera la señal inalámbrica o una entidad asociada con la señal inalámbrica utilizando la asociación probabilística de las una o más palabras a cada identificador de conjunto de servicios.

10. El método implementado por ordenador de la reivindicación 1, en donde la información de la señal inalámbrica no resuelta comprende al menos uno de identificador de conjunto de servicios de la señal inalámbrica o un código de acceso de máquina de un dispositivo que genera la señal inalámbrica.

11. Un sistema que comprende:

un medio de almacenamiento no transitorio que almacena una pluralidad de instrucciones de programa informático; y
un procesador acoplado eléctricamente al medio de almacenamiento no transitorio y configurado para ejecutar la pluralidad de instrucciones de programa informático para:

en un modo de procesamiento por lotes:

recibir datos de observación de una pluralidad de señales inalámbricas observadas por una pluralidad de dispositivos de observador;
entrenar un primer clasificador inductivo para generar un primer conjunto de reglas deductivas para coincidencias sintácticas entre identificadores de conjunto de servicios de la pluralidad de señales inalámbricas y entidades asociadas con los identificadores de conjunto de servicios;
entrenar un segundo clasificador inductivo para generar un segundo conjunto de reglas deductivas basándose en la proximidad espacial y la persistencia temporal de la pluralidad de señales inalámbricas, en donde el segundo clasificador inductivo está entrenado para asignar un átomo, que es una constante que define alguna propiedad intrínseca de la señal dentro de una observación dada, a una señal inalámbrica dada basándose en un gráfico de red de la señal inalámbrica dada;

en un modo de deducción en tiempo real:

recibir información de una señal inalámbrica no resuelta;
generar una resolución semántica de la señal inalámbrica basada en la aplicación de al menos uno del primer y segundo conjunto de reglas deductivas; y
mostrar la resolución semántica de la señal inalámbrica en una interfaz gráfica de usuario.

12. El sistema de la reivindicación 11, en donde el procesador está configurado para ejecutar adicionalmente la pluralidad de las instrucciones de programa informático para:

en el modo de procesamiento por lotes:
entrenar a un tercer clasificador inductivo para generar un tercer conjunto de reglas deductivas para coincidencias sintácticas entre códigos de acceso de máquina de dispositivos que generan la pluralidad de señales inalámbricas y fabricantes de los dispositivos; y
en el modo de deducción en tiempo real:
generar la resolución semántica de la señal inalámbrica basada en la aplicación de al menos uno del primer, segundo y tercer conjunto de reglas deductivas.

13. El sistema de la reivindicación 11, en donde la resolución semántica incluye al menos uno de un fabricante de un dispositivo que genera la señal inalámbrica o una entidad asociada con la señal inalámbrica.

14. El sistema de la reivindicación 13, en donde la resolución semántica incluye una puntuación de probabilidad que indica una probabilidad de que la señal inalámbrica esté asociada con el fabricante o la entidad.

15. El sistema de la reivindicación 13, en donde el procesador está configurado para ejecutar adicionalmente la pluralidad de instrucciones de programa informático para:
asociar la resolución semántica de la señal inalámbrica con una ubicación.

16. El sistema de la reivindicación 15, en donde el procesador está configurado para ejecutar adicionalmente la pluralidad de instrucciones de programa informático para:
mostrar la resolución semántica de la señal inalámbrica y la ubicación asociada dentro de un mapa digital en la interfaz gráfica de usuario.

17. El sistema de la reivindicación 11, en donde, para el entrenamiento del primer clasificador inductivo, el procesador está configurado para ejecutar adicionalmente las instrucciones de programa informático para:

eliminar palabras claves de los identificadores de conjunto de servicios de la pluralidad de señales inalámbricas;
convertir en tokens las palabras restantes de los identificadores de conjunto de servicios de la pluralidad de señales inalámbricas para generar identificadores de conjunto de servicios convertidos en tokens; y
entrenar el primer clasificador inductivo utilizando los identificadores de conjunto de servicios convertidos en tokens.

18. El sistema de la reivindicación 17, en donde, para entrenar el primer clasificador inductivo, el procesador está configurado para ejecutar adicionalmente la pluralidad de instrucciones de programa informático para:

identificar una pluralidad de tokens de tipo palabra en los identificadores de conjunto de servicios convertidos en tokens; y
asociar probabilísticamente una o más palabras que corresponden a uno o más tokens similares a palabras para cada identificador de conjunto de servicios.

19. El sistema de la reivindicación 18, en donde, para generar la resolución semántica de la señal inalámbrica, el procesador está configurado para ejecutar adicionalmente las instrucciones de programa informático para:
identificar al menos uno de un fabricante de un dispositivo que genera la señal inalámbrica o una entidad asociada con la señal inalámbrica que utiliza la asociación probabilística de las una o más palabras a cada identificador de conjunto de servicios.

20. El sistema de la reivindicación 11, en donde la información de la señal inalámbrica no resuelta comprende al menos uno de identificador de conjunto de servicios de la señal inalámbrica o un código de acceso de máquina de un dispositivo que genera la señal inalámbrica.

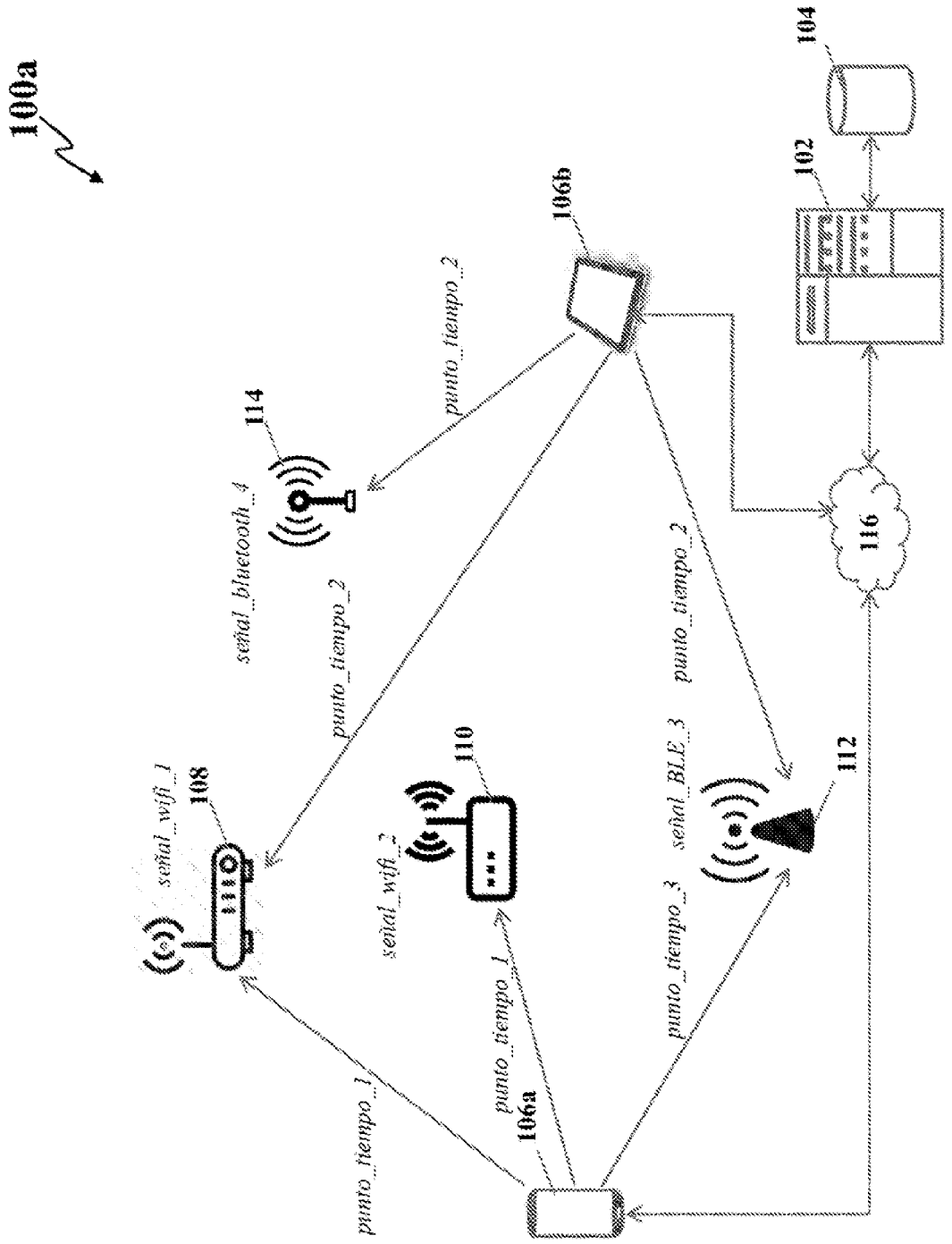
100a
104
102
106b
116
114
señal_bluetooth_4
punto_tiempo_2
punto_tiempo_2
punto_tiempo_2
señal_wifi_1
108
110
señal_wifi_2
112
señal_BLE_3
punto_tiempo_1
punto_tiempo_3
punto_tiempo_1
106a

FIG. 1A

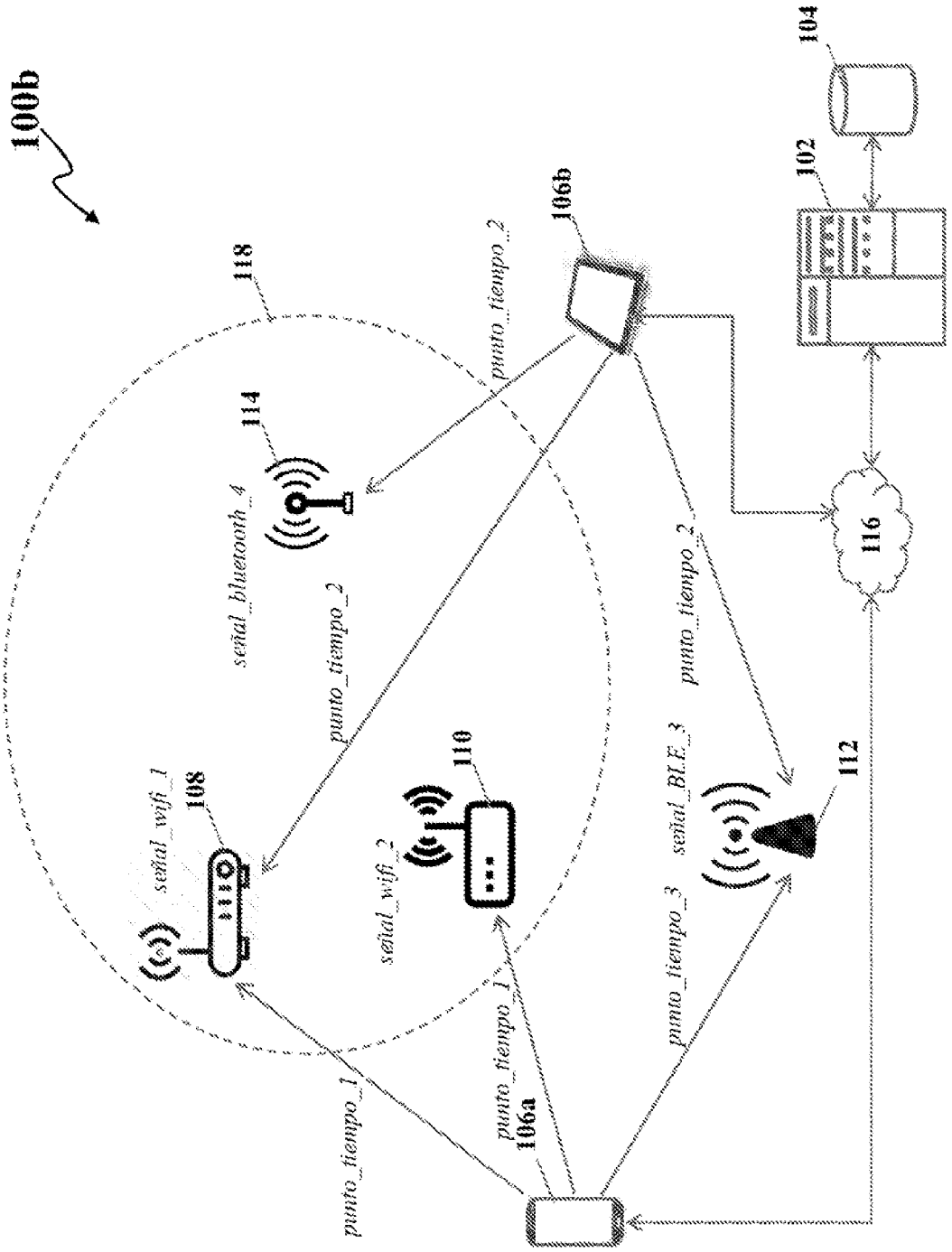
100b
104
102
106b
118
punto_tiempo_2
114
señal_bluetooth_4
punto_tiempo_2
116
punto_tiempo_2
señal_wifi_1
108
110
señal_wifi_2
señal_BLE_3
112
punto_tiempo_3
punto_tiempo_1
106a
punto_tiempo_1

FIG. 1B

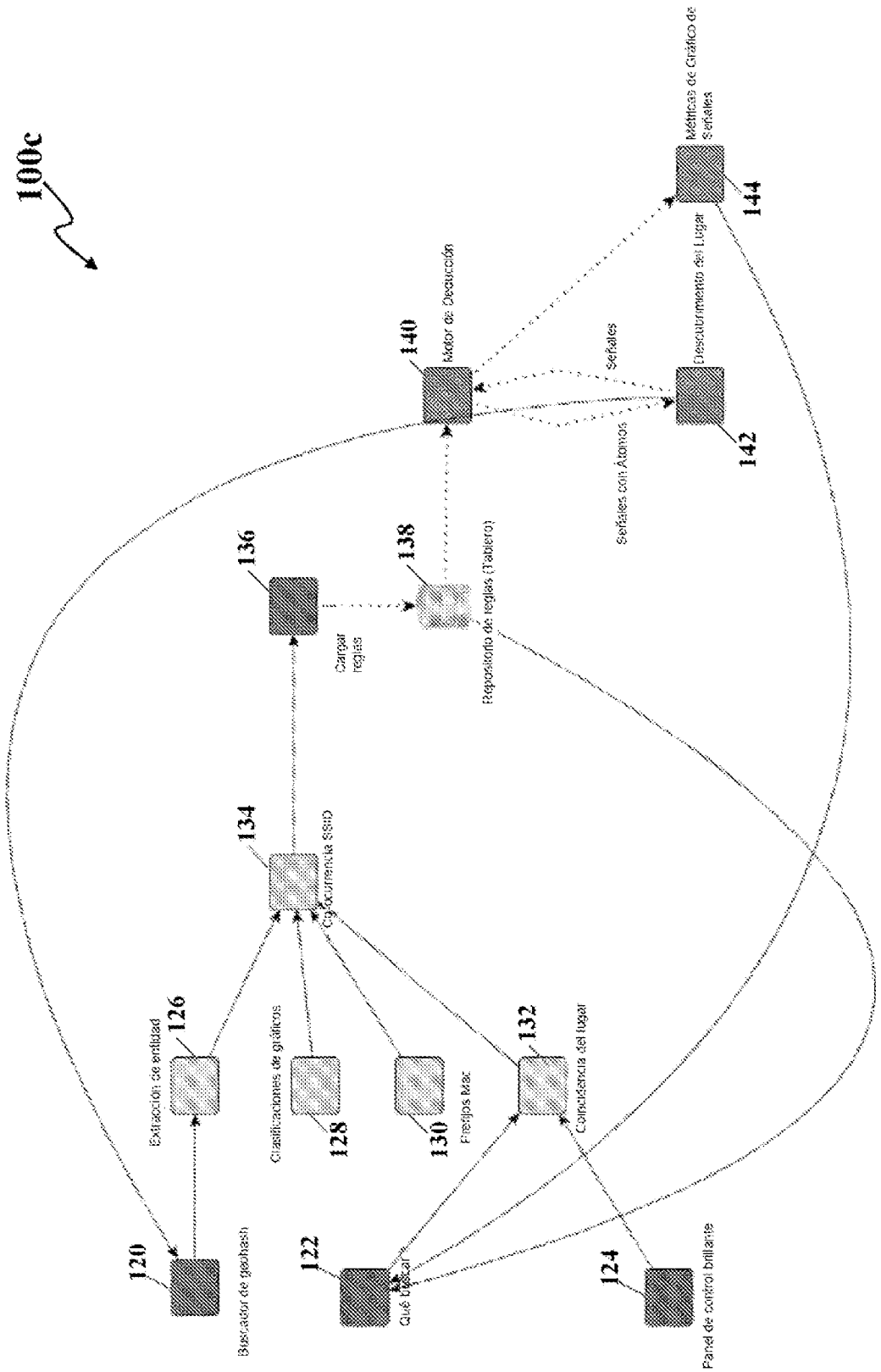
100c
120
Buscador de geohash
122
Qué buscar
124
Panel de control brillante
126
Extracción de entidad
128
Clasificaciones de gráficos
130
Prefijos Mac
132
Coincidencia del lugar
134
Coincidencia SSID
136
Cargar reglas
138
Repositorio de reglas (Tablero)
140
Motor de deducción
142
Descubrimiento del Lugar
144
Métricas de Gráfico de Señales
Señales
Señales con Atomos

FIG. 1C

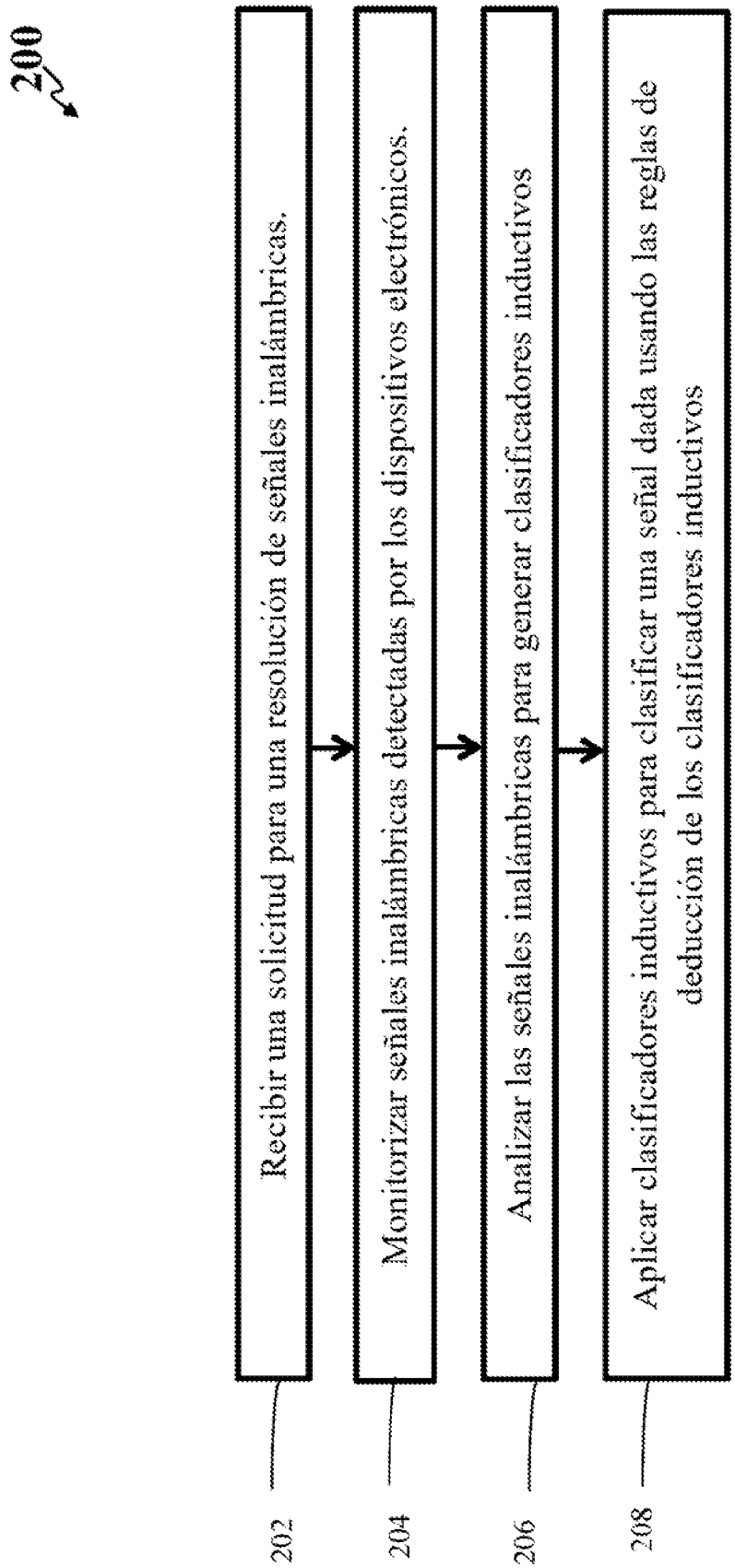
200
Recibir una solicitud para una resolución de señales inalámbricas.
Monitorizar señales inalámbricas detectadas por los dispositivos electrónicos.
Analizar las señales inalámbricas para generar clasificadores inductivos
Aplicar clasificadores inductivos para clasificar una señal dada usando las reglas de deducción de los clasificadores inductivos
202
204
206
208

FIG. 2

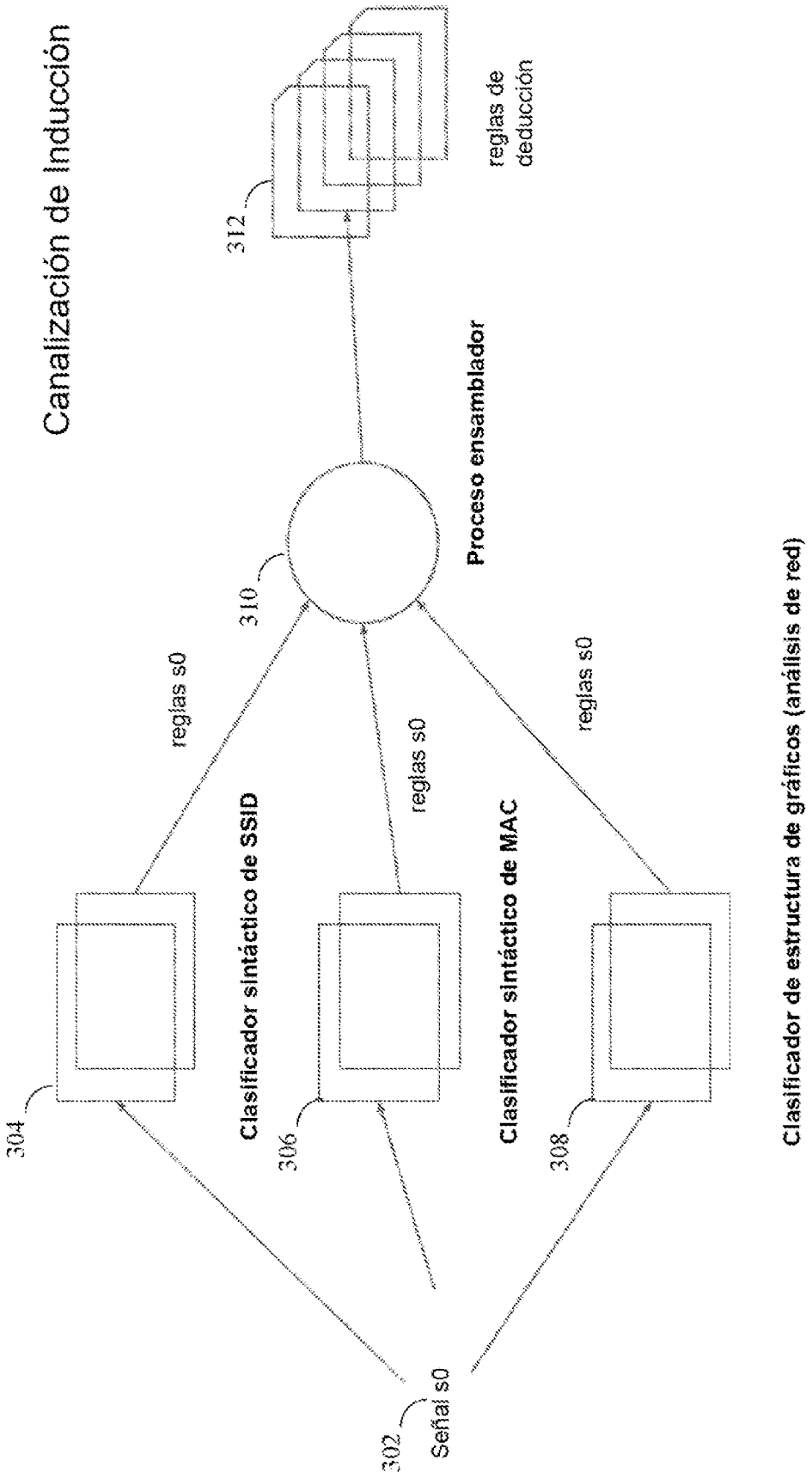
300
Canalización de Inducción
312
reglas de deducción
Proceso ensamblador
310
reglas s0
reglas s0
reglas s0
Clasificador sintáctico de SSID
Clasificador sintáctico de MAC
304
306
308
Señal s0
302
Clasificador de estructura de gráficos (análisis de red)

FIG. 3

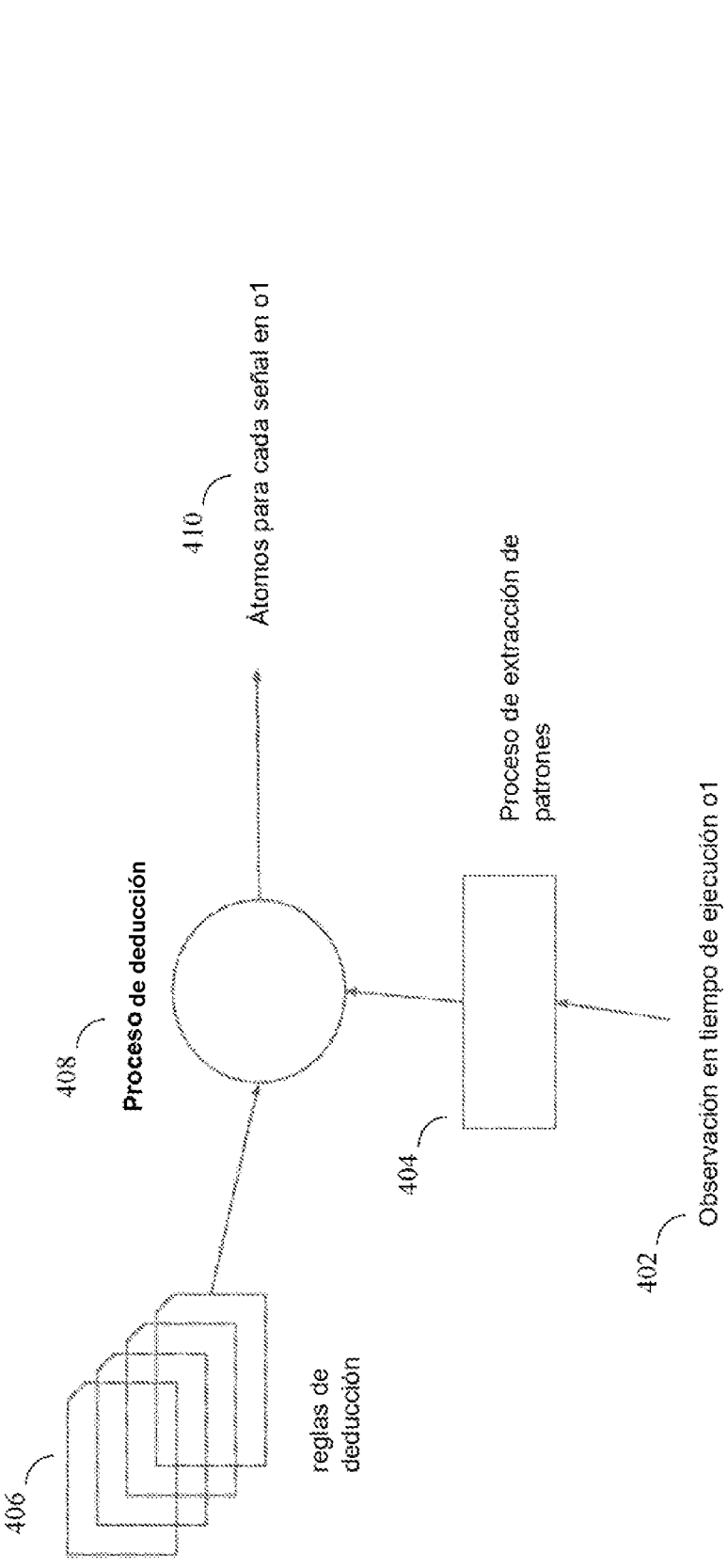
400
Canalización de deducción
410
Átomos para cada señal en o1
408
Proceso de deducción
Proceso de extracción de patrones
404
406
reglas de deducción
402
Observación en tiempo de ejecución o1

FIG. 4

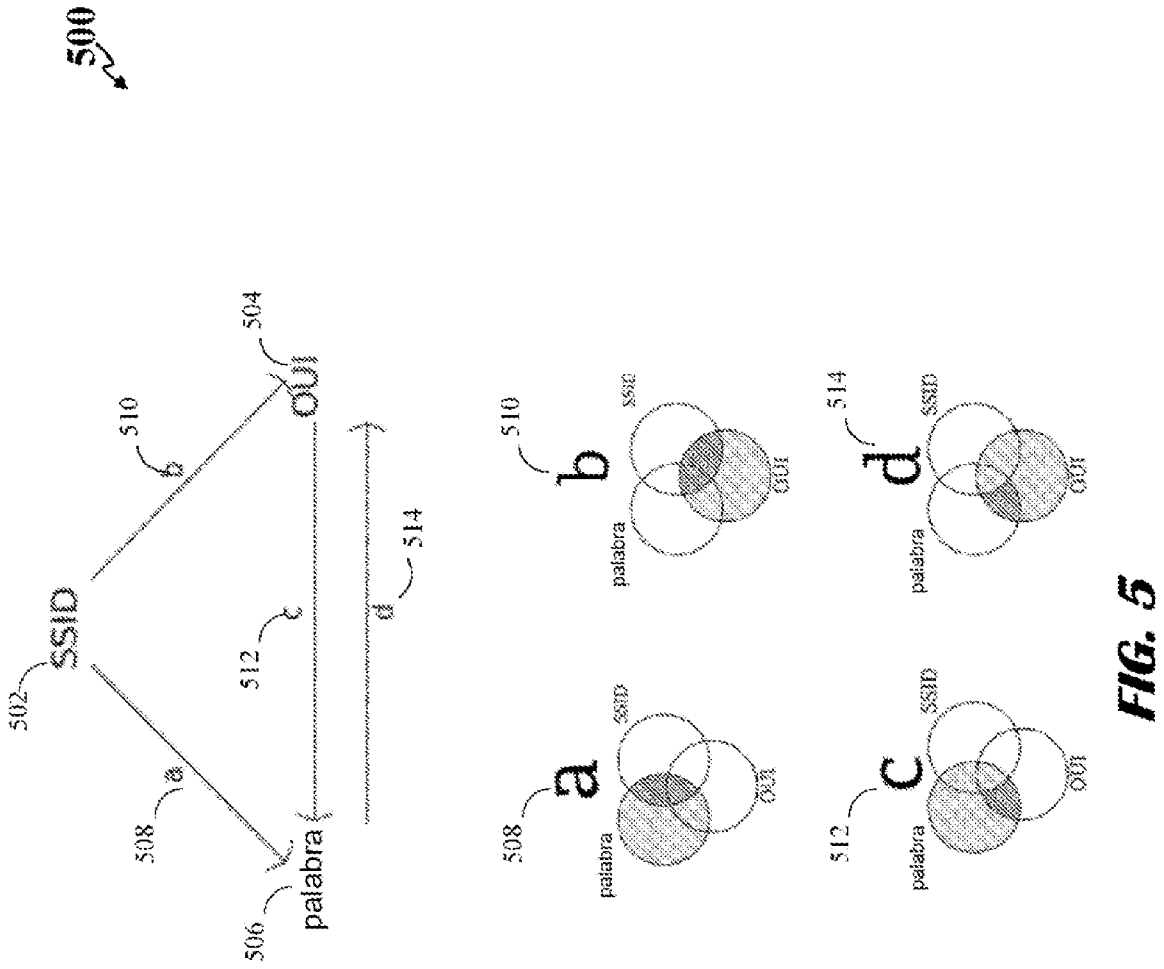

FIG. 5

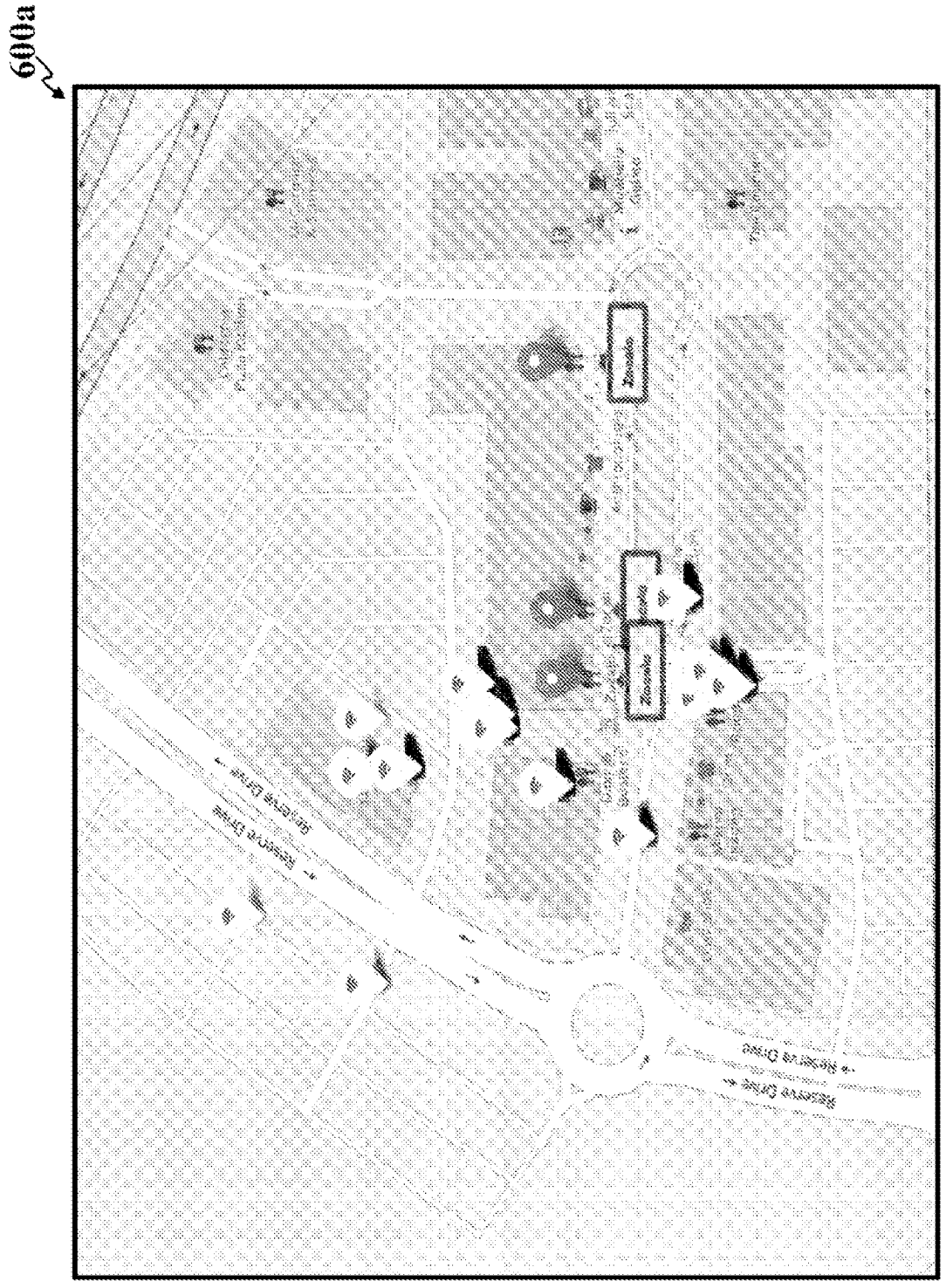

FIG. 6A

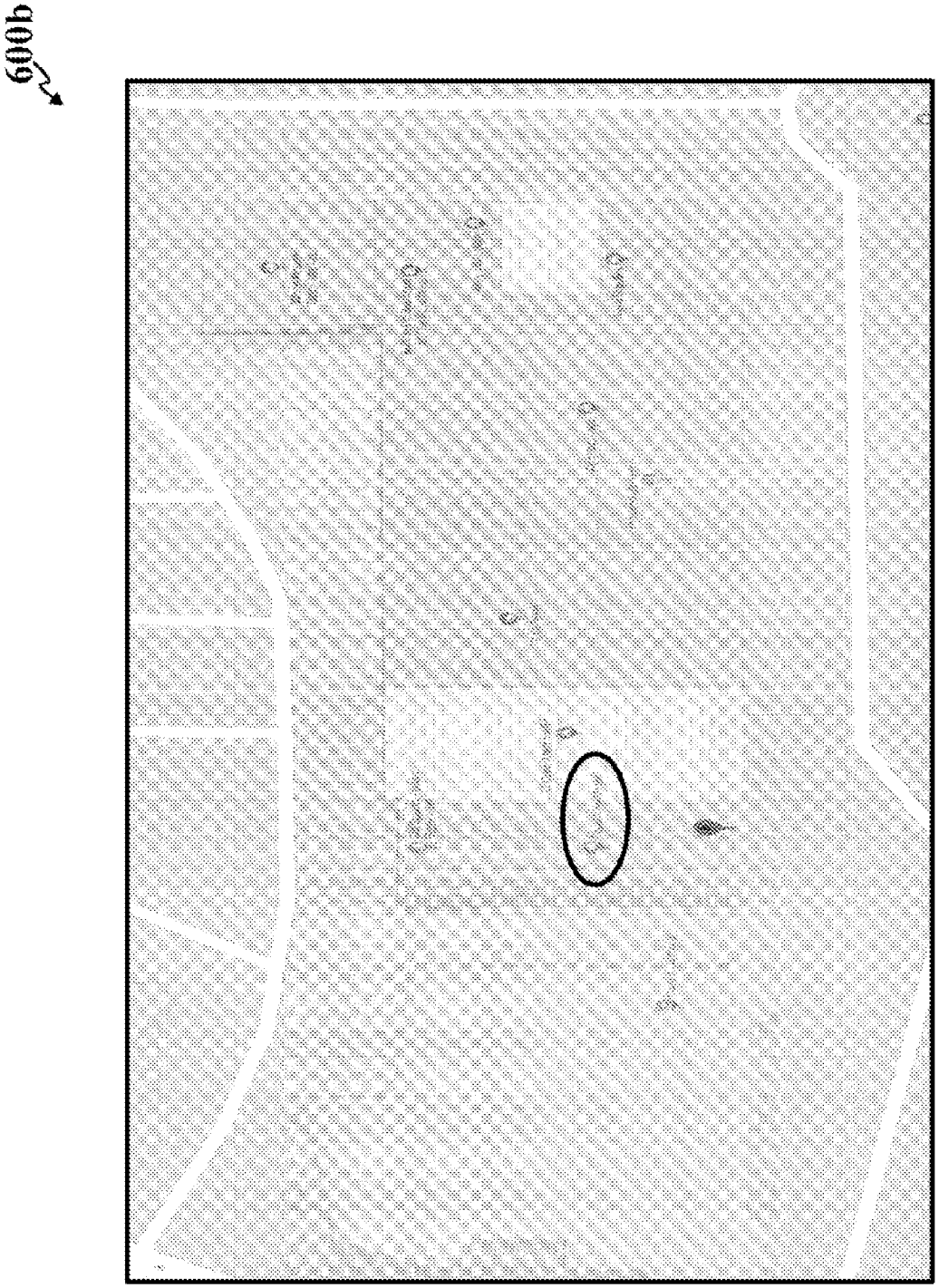

FIG. 6B

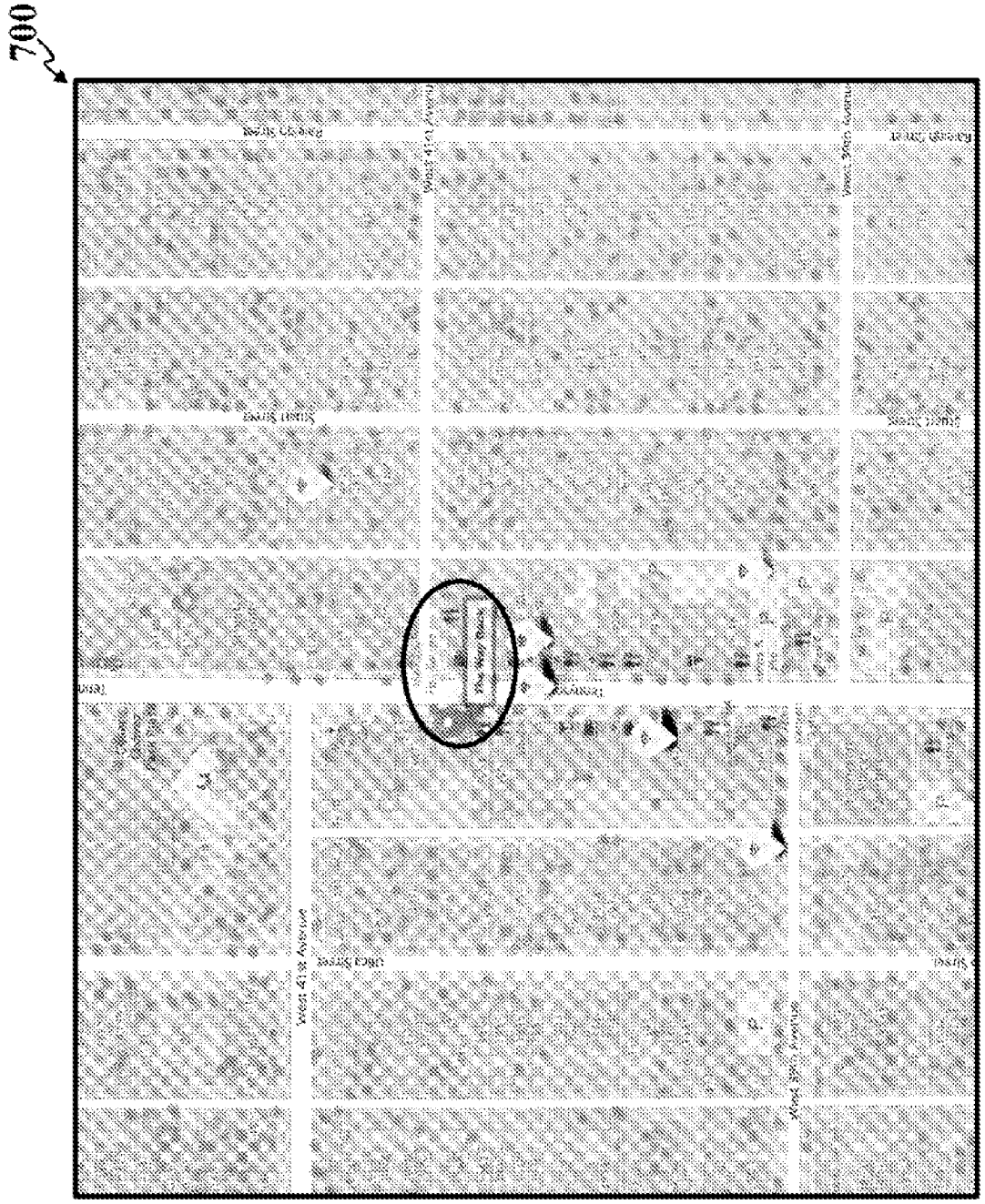

FIG. 7