(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 700 245**

(51) Int. CI.:

H04L 29/06 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **05.10.2010 PCT/CN2010/077569**

(87) Fecha y número de publicación internacional: **16.06.2011 WO11069388**

(96) Fecha de presentación y número de la solicitud europea: **05.10.2010 E 10835423 (4)**

(97) Fecha y número de publicación de la concesión europea: **05.09.2018 EP 2482517**

(54) Título: **Método, aparato y sistema para identificación de protocolo**

(30) Prioridad:

10.12.2009 CN 200910225440

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
14.02.2019

(73) Titular/es:

HUAWEI TECHNOLOGIES CO., LTD. (100.0%)
Huawei Administration Building, Bantian, Longgang District
Shenzhen, Guangdong 518129, CN

(72) Inventor/es:

LIU, HUA

(74) Agente/Representante:

LEHMANN NOVO, María Isabel

DESCRIPCIÓN

Método, aparato y sistema para identificación de protocolo.

Campo de la invención

La presente invención se refiere al campo de la técnica de la comunicación y, en particular, a un método, un aparato y un sistema para la identificación de protocolo de capa de aplicación.

Antecedentes de la invención

Actualmente, la red soporta más y más aplicaciones y el requisito de ancho de banda se convierte en cada vez más alto. El operador espera asegurar la red de funcionamiento para soportar bien varios servicios cruciales (p.ej., navegación en la red), mientras se limita la ocupación excesiva de recursos de red por algunos servicios no cruciales (p.ej., Voz en Protocolo de Internet (VOIP, por sus siglas en inglés) y Entre Pares (P2P, por sus siglas en inglés)). Sobre dicha base, el operador necesita identificar el tipo de protocolo usado en la capa de aplicación por el mensaje (o al que se hace referencia como tren de datos) soportado por la red.

El método actual para la identificación de tipo de protocolo es identificar mediante interrogación todas las condiciones o algoritmos característicos de identificación que posiblemente concuerdan. La manera de interrogación puede ser una interrogación en secuencia en lotes. La implementación específica es, en general, de la siguiente manera: los protocolos del sistema se clasifican en protocolos de alta frecuencia, protocolos ordinarios y protocolos infrecuentes según las frecuencias de uso de los protocolos en la red. Cuando llega un tren de datos, se intenta identificar el tren de datos con todas las reglas o algoritmos de identificación de los protocolos de alta frecuencia. En caso de que la identificación fallara, se prueban todas las reglas o algoritmos de identificación de los protocolos ordinarios. Y si la identificación fallara otra vez, se prueban todas las reglas o algoritmos de identificación de los protocolos infrecuentes. El método actual para la identificación de tipo de protocolo puede también ser la concordancia de patrón y la manera de la concordancia de patrón es explorar todos los protocolos al mismo tiempo y buscar un protocolo que concuerde con el mensaje. Los dos métodos de más arriba requieren, ambos, explorar todos los protocolos y, por consiguiente, la eficacia es relativamente baja.

Con el fin de mejorar la eficacia de identificación de protocolo, se propone un nuevo método para la identificación de protocolo. La dirección de Protocolo de Internet (IP, por sus siglas en inglés) y un par de puertos de un mensaje se extraen. Una tabla de asociación preestablecida se busca para una entrada que incluye la dirección IP y el par de puertos, y la tabla de asociación almacena entradas de la correspondencia entre direcciones IP y pares de puertos y tipos de servicio (un tipo de servicio corresponde al protocolo usado y una vez que el tipo de servicio se determina, el protocolo usado por el mensaje puede determinarse también). Y cuando la entrada que incluye la dirección IP y el par de puertos se busca en la tabla de asociación preestablecida, el tipo de servicio del mensaje es el indicado por la entrada.

Durante el proceso de implementación de la presente invención, el inventor descubre que hay muchos errores de identificación cuando se usa la dirección IP y el par de puertos para determinar el tipo de servicio del tren de datos, dado que el fenómeno de multiplexación de puertos (a saber, una dirección IP y un par de puertos pueden usarse para diferentes tipos de servicios) ocurre con frecuencia en la red actual. Por lo tanto, el método no es apropiado para la ocasión que requiere una precisión alta de identificación de mensaje.

El documento *Extending the NetPDL Language to Support Traffic Classification* describe que: a pesar de la importancia de la clasificación del tráfico en redes modernas, el número de lenguajes adaptados a la presente tarea es extremadamente limitado. Dichos lenguajes pueden ser valiosos porque permiten la actualización de una aplicación (p.ej., cortafuegos) en términos de protocolos soportados simplemente mediante la actualización de su base de datos de descripción de protocolo, en lugar de recopilar la aplicación desde cero. El presente documento presenta un conjunto de extensiones del Lenguaje de Descripción de Protocolo de Red (NetPDL, por sus siglas en inglés) que permiten el soporte de la clasificación de tráfico del enlace de datos a protocolos de capa de aplicación. Un conjunto de resultados experimentales preliminares obtenidos con dichas nuevas extensiones se presenta también.

En la Publicación de Estados Unidos No. 2004/148417A1, se revisa, de forma crítica, la clasificación del tráfico mediante la realización de una evaluación exhaustiva de tres enfoques de clasificación, según los puertos de capa de transporte, comportamiento del anfitrión y características de flujo. La solidez de nuestro trabajo es el amplio rango de datos contra los cuales probamos los tres enfoques de clasificación: sietes trazas con carga útil recolectadas en Japón, Corea y Estados Unidos. Las diversas ubicaciones geográficas, características de enlace y mezcla de tráfico de aplicación en dichos datos nos han permitido evaluar los enfoques bajo una amplia variedad de condiciones. Analizamos las ventajas y limitaciones de cada enfoque, evaluamos métodos para superar las limitaciones y extraemos conocimiento y recomendaciones tanto para el estudio como para la aplicación práctica de la clasificación del tráfico. Ponemos nuestro software, clasificadores y datos a disposición de investigadores interesados en la validación o extensión del presente trabajo.

La publicación de FULVIO RISSO Y OTROS: "*Extending the NetPDL Language to Support Traffic Classification*", *GLOBAL TELECOMMUNICATIONS CONFERENCE*, 2007. *GLOBECOM '07. IEEE, IEEE*, PISCATAWAY, NJ, Estados Unidos, 1 noviembre 2007 (01-11-2007), páginas 22-27, XP031195940, ISBN: 978-1-4244-1042-2, describe mecanismos de identificación de protocolo de capa de aplicación según el uso de una 5-tupla comprendida en un mensaje recibido.

Compendio de la invención

Las realizaciones de la presente invención proveen un método (reivindicación independiente 1), un aparato (reivindicación independiente 6) y un sistema (reivindicación dependiente 9) para la identificación de protocolo de capa de aplicación para mejorar la precisión de identificación de mensaje.

Con el fin de resolver el problema técnico de más arriba, la realización del método para la identificación de protocolo provista por la presente invención puede implementarse por las siguientes soluciones técnicas.

Un método para la identificación de protocolo incluye extraer la tupla de identificación de un mensaje, en donde la tupla de identificación es una 1-tupla que consiste en una dirección de red, una 2-tupla que consiste en una dirección de red y un puerto, o una 3-tupla que consiste en una dirección de red, un puerto y un protocolo de capa de transporte; buscar en una tabla de identificación una lista de aplicación de algoritmo correspondiente a la tupla de identificación; y llevar a cabo la identificación de contenido para el mensaje con un algoritmo en la lista de aplicación de algoritmo para obtener el tipo de protocolo del mensaje usado en una capa de aplicación.

Un aparato para la identificación de protocolo incluye una unidad de extracción configurada para extraer la tupla de identificación de un mensaje, en donde la tupla de identificación es una 1-tupla que consiste en una dirección de red, una 2-tupla que consiste en una dirección de red y un puerto, o una 3-tupla que consiste en una dirección de red, un puerto y un protocolo de capa de transporte; una unidad de almacenamiento configurada para almacenar una tabla de identificación; una unidad de búsqueda de algoritmo configurada para buscar en la tabla de identificación una lista de aplicación de algoritmo correspondiente a la tupla de identificación; y una unidad de identificación de protocolo configurada para llevar a cabo la identificación de contenido para el mensaje con un algoritmo en la lista de aplicación de algoritmo para obtener el tipo de protocolo del mensaje usado en una capa de aplicación.

Un sistema para la identificación de protocolo incluye un aparato de recepción de mensaje y el aparato para la identificación de protocolo provisto por la realización de la presente invención, el aparato para la identificación de protocolo se configura para llevar a cabo la identificación de protocolo para el mensaje recibido por el aparato de recepción de mensaje.

Las soluciones técnicas de más arriba tienen el siguiente efecto beneficioso: mediante el uso de la tupla de identificación que al menos incluye una dirección de red y se extrae del mensaje para buscar la lista de aplicación de algoritmo correspondiente y luego llevar a cabo la identificación de contenido para el mensaje con un algoritmo en la lista de aplicación de algoritmo, la precisión de la identificación de mensaje puede mejorarse y el error de identificación puede reducirse.

Breve descripción de los dibujos

Con el fin de describir, de manera más clara, las soluciones técnicas de las realizaciones de la presente invención, los dibujos que se usarán en las descripciones de las realizaciones se introducirán brevemente de la siguiente manera. De manera evidente, los siguientes dibujos simplemente ilustran algunas realizaciones de la presente invención y una persona con experiencia en la técnica puede obtener otros dibujos a partir de dichos dibujos sin esfuerzos creativos.

La Figura 1 es un diagrama de flujo de un método según la realización 1 de la presente invención;

la Figura 2 es un diagrama de flujo de un método según la realización 1 de la presente invención;

la Figura 3 es un diagrama de flujo de un método según la realización 2 de la presente invención;

la Figura 4 es un diagrama de estructura de un aparato según la realización 3 de la presente invención;

la Figura 5 es un diagrama de estructura de un aparato según la realización 3 de la presente invención;

la Figura 6 es un diagrama de estructura de un aparato según la realización 3 de la presente invención;

la Figura 7 es un diagrama de estructura de un sistema según la realización 4 de la presente invención.

Descripción detallada de las realizaciones

Las soluciones técnicas de las realizaciones de la presente invención se describirán de forma clara y completa de la siguiente manera en conjunto con los dibujos.

Realización 1

Un método para la identificación de protocolo se ilustra en el diagrama de flujo de la Figura 1.

En la etapa 101, la tupla de identificación de un mensaje se extrae y la tupla de identificación al menos incluye una dirección de red.

La tupla de identificación puede ser N-tupla, por ejemplo, puede ser 1-tupla, a saber, dirección de red; 2-tupla, a saber, dirección de red y puerto; o 3-tupla, a saber, dirección de red, puerto y protocolo de capa de transporte. Por supuesto, el número de tuplas puede establecerse de forma concreta tras la solicitud de identificación de protocolo y no se encuentra limitado en la presente memoria. La dirección IP puede ser dirección IP de origen o destino del mensaje. Cuando se usa la dirección IP de origen, se usará un puerto de origen, mientras que cuando se usa la dirección IP de destino, se usará un puerto de destino.

En la etapa 102, la lista de aplicación de algoritmo correspondiente a la tupla de identificación se busca en la tabla de identificación.

De manera específica, la búsqueda en la tabla de identificación de la lista de aplicación de algoritmo correspondiente a la tupla de identificación incluye: tomar la tupla de identificación como un valor clave y usar un método de búsqueda de *hash* para buscar en la tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación. Como una manera de búsqueda concreta, el método de búsqueda de *hash* tiene una ventaja de búsqueda rápida. Se apreciará que el método de búsqueda de *hash* no es una implementación única de búsqueda en la tabla de una entrada y, por consiguiente, tomándolo como un ejemplo, el método de búsqueda de *hash* no se interpretará como una limitación a la realización de la presente invención. El método de búsqueda de más arriba puede buscar la lista de aplicación de algoritmo correspondiente a la tupla de identificación mediante la concordancia de la tupla de identificación con las entradas en la tabla de identificación. Además, el algoritmo puede también interpretarse como una regla de identificación de mensaje. La tabla de identificación puede preestablecerse. Además, un método para actualizar, de manera dinámica, la tabla de identificación se introducirá en la descripción de la realización subsiguiente.

En la etapa 103, la identificación de contenido para el mensaje se lleva a cabo con un algoritmo en la lista de aplicación de algoritmo para obtener el tipo de protocolo del mensaje.

La etapa 102 determina un algoritmo preferiblemente usado por el mensaje para la identificación y luego la etapa 103 identifica el mensaje con el algoritmo. Como puede verse a partir de la solución técnica de más arriba, después de adquirir el mensaje del tren de datos, la realización de la presente invención selecciona, primero, el algoritmo que se usará preferiblemente para la identificación con información del mensaje como, por ejemplo, la dirección IP, puerto y protocolo de capa de transporte. Por consiguiente, las posibilidades de éxito de concordancia son altas y no es necesario identificar el tren de datos con todos los algoritmos por medio de la interrogación o concordancia de patrón y, de esta manera, se logra la identificación de protocolo de alta velocidad. Además, mediante el uso de la tupla de identificación que al menos incluye la dirección IP y se extrae del mensaje para buscar el algoritmo correspondiente y luego llevar a cabo la identificación de contenido para el mensaje con el algoritmo buscado, la precisión de la identificación de mensaje puede mejorarse y el error de identificación puede reducirse.

Como se ilustra en la Figura 2, según la implementación correspondiente a la Figura 1, el método además incluye las etapas de más abajo.

En la etapa 201, si la búsqueda en la tabla de identificación de la lista de aplicación de algoritmo correspondiente a la tupla de identificación fallara, o la identificación de contenido para el mensaje con el algoritmo en la lista de aplicación de algoritmo fallara, el mensaje puede identificarse a través de la interrogación o concordancia de patrón. Por supuesto, la identificación de otra manera no se encuentra limitada en la realización de la presente invención. La manera de interrogación puede ser en lotes y secuencia fija, o cualquier manera que intente llevar a cabo la concordancia de patrón de todas las reglas características para descubrir un protocolo usado por el mensaje, lo cual no se encuentra limitado en la realización de la presente invención. El resultado de la búsqueda de la lista de aplicación de algoritmo y si la identificación de contenido para el mensaje falla puede descubrirse en los resultados de la búsqueda e identificación. Por supuesto, la manera de apreciación también es viable, la cual no se encuentra limitada en la realización de la presente invención.

Además, la realización de la presente invención provee un método para actualizar, de forma dinámica, la tabla de identificación, lo cual puede llevarse a cabo concretamente después de la etapa 201.

En la etapa 202, la lista de aplicación de algoritmo correspondiente a la tupla de identificación en la tabla de identificación se actualiza con el algoritmo del mensaje. Como puede apreciarse, la tabla de identificación puede

actualizarse o no actualizarse cuando se preestablece. Y cuando la tabla de identificación no se encuentra preestablecida, la lista de aplicación de algoritmo correspondiente a la tupla de identificación puede rellenarse en la tabla de identificación llevando a cabo las etapas 201 y 202. La actualización de más arriba puede ser una incorporación a los datos originales o una creación de nuevas entradas para almacenar la tupla de identificación y la lista de aplicación de algoritmo correspondiente.

La incorporación dinámica tiene la ventaja de actualizar, de manera automática, la tabla de identificación mientras el entorno de red cambia y, de esta manera, se evita que la tabla de identificación se convierta en antigua y no apropiada para un nuevo entorno de red, y se evita el problema de la configuración manual de la tabla de identificación.

Realización 2

Tomando la extracción de información de N-tupla como un ejemplo, la tabla de identificación se busca a través de la concordancia y el método para la identificación de protocolo se ilustra a través de los ejemplos de determinación de si la búsqueda y la identificación tienen o no éxito en la manera de apreciación, según se ilustra en la Figura 3. En la etapa 301, N-tupla del mensaje se extrae. La N-tupla incluye dirección de red, puerto y protocolo de capa de transmisión del mensaje, etc.

La N-tupla puede ser 1-tupla (solo incluye la dirección de red), 2-tupla (incluye la dirección de red y puerto) o 3-tupla (incluye la dirección de red, puerto y protocolo de capa de transmisión como, por ejemplo, el Protocolo de Control de Transmisión (TCP, por sus siglas en inglés//Protocolo de Datagrama de Usuario (UDP, por sus siglas en inglés)/Protocolo de Transmisión de Control de Tren (SCTP, por sus siglas en inglés)).

En la red actual, muchos trenes de datos son tipo de aplicaciones cliente/servidor. En donde un anfitrión que sirve a un servidor normalmente solo provee pocos tipos de protocolo de servicios, y el anfitrión de servidor provee un servicio fijo en un puerto fijo, a saber, normalmente solo un tipo de protocolo de servicio se provee en un puerto dentro de cierto período. Teniendo en cuenta la multiplexación de puerto, el mismo puerto del anfitrión puede proveer otros tipos de protocolo de servicio y no puede determinarse si el mismo puerto del mismo anfitrión cambia para proveer otros tipos de protocolo de servicio adicionales. Por lo tanto, después de haber identificado que cierto tipo de protocolo de servicio se provee en un puerto específico de cierto anfitrión, la información de N-tupla (una combinación de la dirección de red, puerto y protocolo de capa de transmisión) del anfitrión se registra, y una regla/algoritmo para identificar el tipo de protocolo se corresponde con la N-tupla. Cuando llega un tren subsiguiente, si su N-tupla es igual a aquella previamente almacenada, la regla/algoritmo para la identificación de protocolo correspondiente a la N-tupla previamente almacenada se usará en primer lugar para la identificación. En la mayoría de las circunstancias, el tipo de protocolo usado por el tren puede identificarse con la regla/algoritmo correspondiente, y no es necesario intentar otras reglas/algoritmos. Entonces el rendimiento de la identificación de un mismo tren puede mejorarse varias veces a través de dicha identificación y no se provocará ningún error de identificación. La información de N-tupla en la realización de la presente invención puede constar de información de origen o destino de un tren. Dado que el mensaje incluye dos IP y dos puertos (a saber, IP de origen y destino y puertos de origen y destino), respectivamente, la información de origen incluye IP de origen y puerto de origen del mensaje, mientras que la información de destino incluye IP de destino y puerto de destino del mensaje. La información de origen o destino se necesita para construir la N-tupla, y el protocolo de capa de transmisión también se requiere para construir la 3-tupla.

En la etapa 302, las entradas de la N-tupla se buscan en la tabla de identificación mediante el uso de la N-tupla descrita más arriba a través de la concordancia, y la tabla de identificación almacena entradas de correspondencia entre la N-tupla y la lista de aplicación.

Las entradas en la tabla de identificación pueden obtenerse de la siguiente manera: según el resultado de la identificación de protocolo, determinar y almacenar la información de N-tupla del anfitrión que provee el servicio; dado que el puerto puede multiplexarse, y no puede determinarse cuándo cambiará el tipo de servicio provisto por el puerto, la información de N-tupla se mantiene para que corresponda a la lista de aplicación como, por ejemplo, la regla/algoritmo usados para la identificación. Durante el uso, extraer la información de N-tupla del mensaje que se identificará y llevar a cabo la concordancia en la tabla de identificación. La concordancia con entradas en la tabla de identificación es exitosa una vez que la información de N-tupla compuesta de una de la información de origen y destino del mensaje que concordará concuerda. Se sugiere adoptar la búsqueda de *hash* y tomar la N-tupla como un valor clave para buscar en la tabla de identificación.

En la etapa 303, se determina si la concordancia tiene éxito. El flujo puede proceder a la etapa 304 en caso de éxito y, de lo contrario, proceder a la etapa 306.

En la etapa 304, la identificación de contenido se lleva a cabo con una lista de aplicación de algoritmo/regla de la entrada buscada para obtener el tipo de protocolo del mensaje.

Dado que la N-tupla en la tabla de identificación tiene una lista de aplicación de regla/algoritmo de identificación correspondiente, cuando la N-tupla del mensaje que se identificará concuerda en la tabla de identificación, la lista de aplicación de regla/algoritmo de identificación correspondiente puede usarse para llevar a cabo una identificación de contenido para el mensaje para obtener el tipo de protocolo del mensaje.

En la etapa 305, se determina si la identificación de contenido tiene éxito. El flujo finaliza en caso de éxito y, de lo contrario, procede a la etapa 306.

En la etapa 306, el mensaje se identifica llevando a cabo la concordancia de patrón de todas las reglas características/llevando a cabo todos los algoritmos de identificación.

La etapa 306 puede usar cualquier método actual para llevar a cabo la identificación o concordancia, p.ej., intentar la interrogación o concordancia de patrón de todos los algoritmos de identificación para llevar a cabo la identificación de contenido para el mensaje que se identificará. La manera concreta no se encuentra limitada en la realización de la presente invención.

En la etapa 307, después de que la identificación tiene éxito, si se determina que el protocolo usado por el mensaje es un protocolo tipo cliente/servidor, la tabla de identificación se actualiza con la correspondencia entre la N-tupla del anfitrión en el extremo de servidor y la regla/algoritmo de identificación usados. De manera específica, la implementación de la etapa de actualización puede ser: determinar y extraer información (p.ej., dirección de red, puerto, protocolo de capa de transmisión) del anfitrión que sirve como el servidor en el mensaje según la regla/algoritmo de identificación para construir la N-tupla, y luego buscar en una tabla de asociación rápida la N-tupla correspondiente, y cuando la N-tupla correspondiente se descubre, reemplazar la aplicación de la regla/algoritmo de identificación correspondiente a aquella N-tupla por la aplicación de la nueva regla/algoritmo actualmente usados, o añadir la nueva regla/algoritmo a la lista de aplicación que existía. En caso de que la tupla de identificación use 1-tupla que solamente incluye la dirección de red, el procesamiento de actualización puede llevarse a cabo en la manera de incorporación. En caso de que se use la 2-tupla o 3-tupla, el procesamiento de actualización puede llevarse a cabo en la manera de reemplazo.

En la red real, trenes de datos que usan protocolos tipo cliente/servidor (principalmente P2P, Protocolo de Transferencia de Hipertexto (HTTP, por sus siglas en inglés), otros protocolos WEB, etc.) ocupan la mayor parte del tráfico de red. La solución implementada por la presente invención puede usar dicho protocolo que se basa en la identificación exitosa del tren de datos y mejora los rendimientos de identificación de trenes subsiguientes y, de esta manera, se mejora el rendimiento de todo el aparato de identificación. Las pruebas muestran que mediante dicho procesamiento, los rendimientos de identificación de trenes subsiguientes mejoran ampliamente en comparación con el primer tren con respecto a la identificación de protocolos tipo cliente/servidor, y el rendimiento de todo el aparato para la identificación de protocolo mejora en más del 50%.

Realización 3

Como se muestra en la Figura 4, la realización de la presente invención además provee un aparato para la identificación de protocolo, el cual incluye una unidad de extracción 401, una unidad de almacenamiento 402, una unidad de búsqueda de algoritmo 403 y una unidad de identificación de protocolo 404. La unidad de extracción 401 se configura para extraer una tupla de identificación de un mensaje, y la tupla de identificación al menos incluye una dirección de red. La unidad de almacenamiento 402 se configura para almacenar una tabla de identificación. La unidad de búsqueda de algoritmo 403 se configura para buscar en la tabla de identificación una lista de aplicación de algoritmo correspondiente a la tupla de identificación. La unidad de identificación de protocolo 404 se configura para llevar a cabo una identificación de contenido para el mensaje con el algoritmo en la lista de aplicación de algoritmo para obtener el tipo de protocolo del mensaje.

De manera específica, la unidad de extracción 401 se configura para extraer la dirección de red del mensaje, o la dirección de red y puerto, o la dirección de red, puerto y protocolo de capa de transmisión.

Según se ilustra en la Figura 5, el aparato además incluye una unidad de búsqueda 501. La unidad de búsqueda 501 se configura para identificar el mensaje a través de la interrogación o concordancia de patrón cuando la búsqueda en la tabla de identificación una lista de aplicación de algoritmo correspondiente a la tupla de identificación falla, o la identificación de contenido para el mensaje con el algoritmo en la lista de aplicación de algoritmo falla.

Según se ilustra en la Figura 6, el aparato además incluye una unidad de actualización 601. La unidad de actualización 601 se configura para actualizar la lista de aplicación de algoritmo correspondiente a la tupla de identificación en la tabla de identificación con el algoritmo del mensaje, después de que el mensaje se identifica a través de la interrogación o concordancia de patrón.

De manera específica, la unidad de búsqueda de algoritmo 403 se configura para buscar en la tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación mediante el uso del método de búsqueda de *hash* y tomando la tupla de identificación como un valor clave.

Como puede verse a partir de la solución técnica provista por la realización de la presente invención, después de adquirir el mensaje del tren de datos, la realización de la presente invención selecciona, primero, el algoritmo que se usará preferiblemente para la identificación con información del mensaje como, por ejemplo, la dirección IP, puerto y protocolo de capa de transporte. Por consiguiente, las posibilidades de éxito de concordancia son altas y no es necesario identificar el tren de datos con todos los algoritmos por medio de la interrogación o concordancia de patrón y, de esta manera, se logra la identificación de protocolo de alta velocidad. Además, el algoritmo correspondiente se busca con la tupla de identificación que al menos incluye la dirección IP y se extrae del mensaje, y luego una identificación de contenido para el mensaje se lleva a cabo con el algoritmo buscado. La identificación de contenido puede mejorar la precisión de identificación de mensaje y reducir el error de identificación.

La implementación que usa la incorporación dinámica tiene la ventaja de actualizar, de manera automática, la tabla de identificación mientras el entorno de red cambia y, de esta manera, se evita que la tabla de identificación se convierta en antigua y no apropiada para un nuevo entorno de red, y se evita el problema de la configuración manual de la tabla de identificación.

Según se ilustra en la Figura 7, la realización de la presente invención además provee un sistema para la identificación de protocolo, el cual incluye un aparato de recepción de mensaje 701 y cualquier aparato 702 para la identificación de protocolo provista por la realización de la presente invención. El aparato 702 para la identificación de protocolo se configura para llevar a cabo una identificación de protocolo para el mensaje recibido por el aparato de recepción de mensaje 701. La implementación de la identificación de protocolo para el mensaje puede referirse a la realización del método o a la realización del aparato para la identificación de protocolo.

Como puede verse a partir de las soluciones técnicas provistas por la realización de la presente invención, después de adquirir el mensaje del tren de datos, la realización de la presente invención selecciona, primero, el algoritmo que se usará preferiblemente para la identificación con información del mensaje como, por ejemplo, la dirección IP, puerto y protocolo de capa de transporte. Por consiguiente, las posibilidades de éxito de concordancia son altas y no es necesario identificar el tren de datos con todos los algoritmos por medio de la interrogación o concordancia de patrón y, de esta manera, se logra la identificación de protocolo de alta velocidad. Además, el algoritmo correspondiente se busca con la tupla de identificación que al menos incluye la dirección IP y se extrae del mensaje, y luego una identificación de contenido para el mensaje se lleva a cabo con el algoritmo buscado. La identificación de contenido puede mejorar la precisión de identificación de mensaje y reducir el error de identificación.

Una persona con experiencia en la técnica puede apreciar que todas o una parte de las etapas en los métodos según las realizaciones de más arriba se implementan mediante la instrucción al hardware relevante a través de un programa que puede almacenarse en un medio de almacenamiento legible por ordenador (p.ej., ROM, disco magnético, disco óptico, etc.).

El método, aparato y sistema para la identificación de protocolo provistos por las realizaciones de la presente invención se describen en detalle según se establece más arriba y ejemplos específicos se usan para ilustrar el principio y las realizaciones de la presente invención. Las realizaciones de más arriba solo se describen para ayudar a comprender los métodos y conceptos principales de la presente invención.

El alcance de la invención se define por las reivindicaciones anexas.

REIVINDICACIONES

1. Un método para la identificación de protocolo, que comprende:

extraer (101) la tupla de identificación de un mensaje, en donde la tupla de identificación es una 1-tupla que consiste en una dirección de red, una 2-tupla que consiste en una dirección de red y un puerto, o una 3-tupla que consiste en una dirección de red, un puerto y un protocolo de capa de transporte;

buscar (102) en una tabla de identificación una lista de aplicación de algoritmo correspondiente a la tupla de identificación; y

llevar a cabo (103) la identificación de contenido para el mensaje con un algoritmo en la lista de aplicación de algoritmo para obtener el tipo de protocolo del mensaje usado en una capa de aplicación.

2. El método según la reivindicación 1, en donde cuando la búsqueda en la tabla de identificación de la lista de aplicación de algoritmo correspondiente a la tupla de identificación falla, o la identificación de contenido para el mensaje con el algoritmo en la lista de aplicación de algoritmo falla, el tipo de protocolo del mensaje usado en la capa de aplicación se identifica a través de la concordancia de patrón o mediante la interrogación de todos los algoritmos de identificación.

3. El método según la reivindicación 2, en donde después de identificar el tipo de protocolo del mensaje a través de la concordancia de patrón o interrogación de todos los algoritmos de identificación, el método además comprende:

actualizar la lista de aplicación de algoritmo correspondiente a la tupla de identificación en la tabla de identificación con el algoritmo para identificar el tipo de protocolo del mensaje usado en la capa de aplicación.

4. El método según la reivindicación 1, en donde la búsqueda en la tabla de identificación de la lista de aplicación de algoritmo correspondiente a la tupla de identificación comprende tomar la tupla de identificación como un valor clave y usar un método de búsqueda de *hash* para buscar en la tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación.

5. Un aparato para la identificación de protocolo, que comprende:

una unidad de extracción (401) configurada para extraer la tupla de identificación de un mensaje, en donde la tupla de identificación es una 1-tupla que consiste en una dirección de red, una 2-tupla que consiste en una dirección de red y un puerto, o una 3-tupla que consiste en una dirección de red, un puerto y un protocolo de capa de transporte;

una unidad de almacenamiento (402) configurada para almacenar una tabla de identificación;

en donde el aparato además comprende una unidad de búsqueda de algoritmo (403) configurada para buscar en la tabla de identificación una lista de aplicación de algoritmo correspondiente a la tupla de identificación; y

una unidad de identificación de protocolo (404) configurada para llevar a cabo la identificación de contenido para el mensaje con un algoritmo en la lista de aplicación de algoritmo para obtener el tipo de protocolo del mensaje usado en una capa de aplicación.

6. El aparato según la reivindicación 5, que además comprende:

una unidad de búsqueda (501) configurada para identificar el tipo de protocolo del mensaje usado en la capa de aplicación en la manera de interrogación de todos los algoritmos de identificación o concordancia de patrón cuando la búsqueda en la tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación falla, o la identificación de contenido para el mensaje con el algoritmo en la lista de aplicación de algoritmo falla.

7. El aparato según la reivindicación 6, que además comprende:

una unidad de actualización (601) configurada para actualizar la lista de aplicación de algoritmo correspondiente a la tupla de identificación en la tabla de identificación con el algoritmo para identificar el tipo de protocolo del mensaje usado en la capa de aplicación, después de que el tipo de protocolo del mensaje usado en la capa de aplicación se identifica en la manera de interrogación de todos los algoritmos de identificación o concordancia de patrón.

8. El aparato según la reivindicación 5, en donde la unidad de búsqueda de algoritmo se configura para buscar en la tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación tomando la tupla de identificación como un valor clave y mediante el uso del método de búsqueda de *hash*.

9. Un sistema para la identificación de protocolo, que comprende un aparato de recepción de mensaje (701) y el aparato (702) para la identificación de protocolo según cualquiera de las reivindicaciones 5 a 8, en donde el aparato

para la identificación de protocolo se configura para llevar a cabo la identificación de protocolo para el mensaje recibido por el aparato de recepción de mensaje.

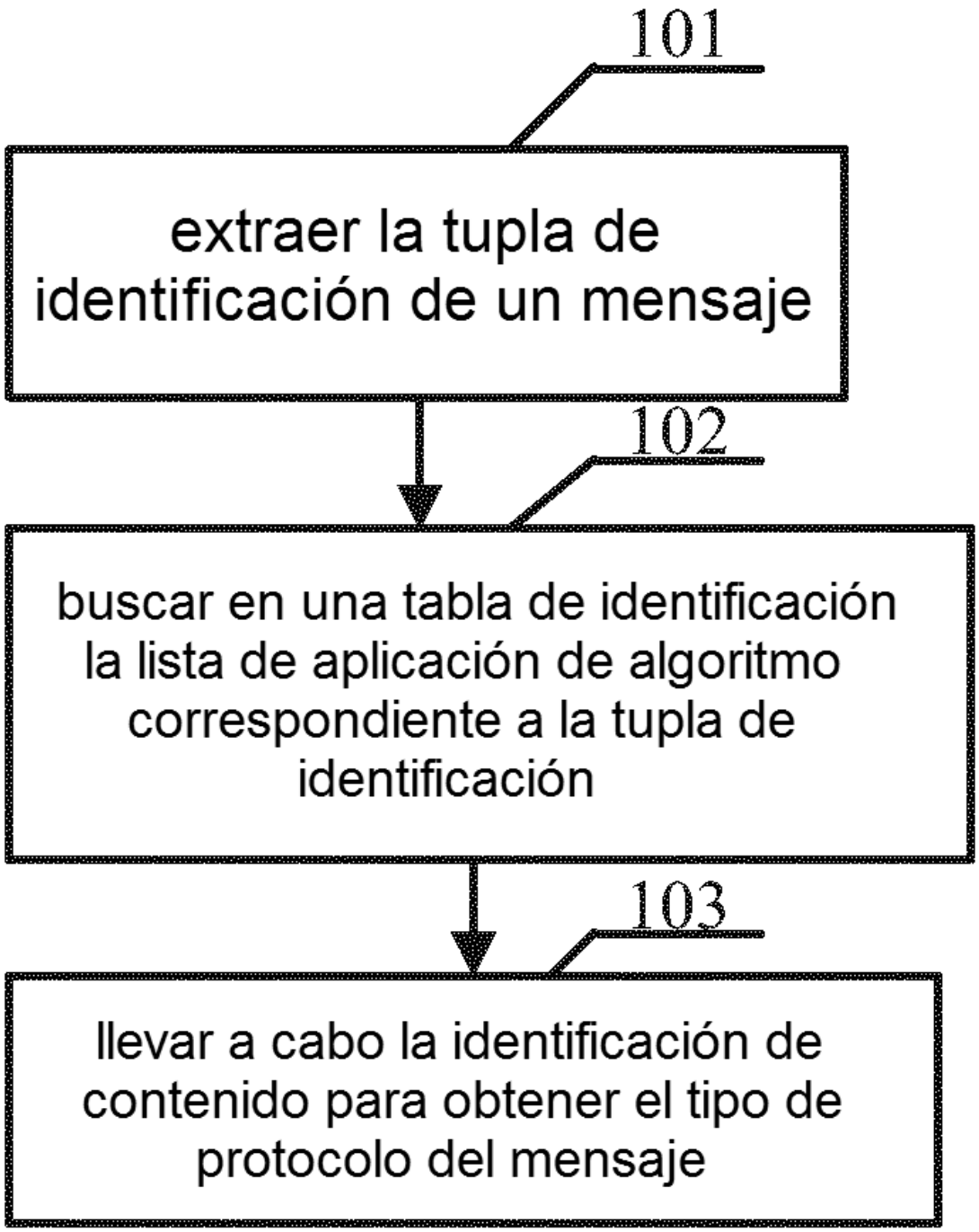
101
extraer la tupla de identificación de un mensaje
102
buscar en una tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación
103
llevar a cabo la identificación de contenido para obtener el tipo de protocolo del mensaje

Fig.1

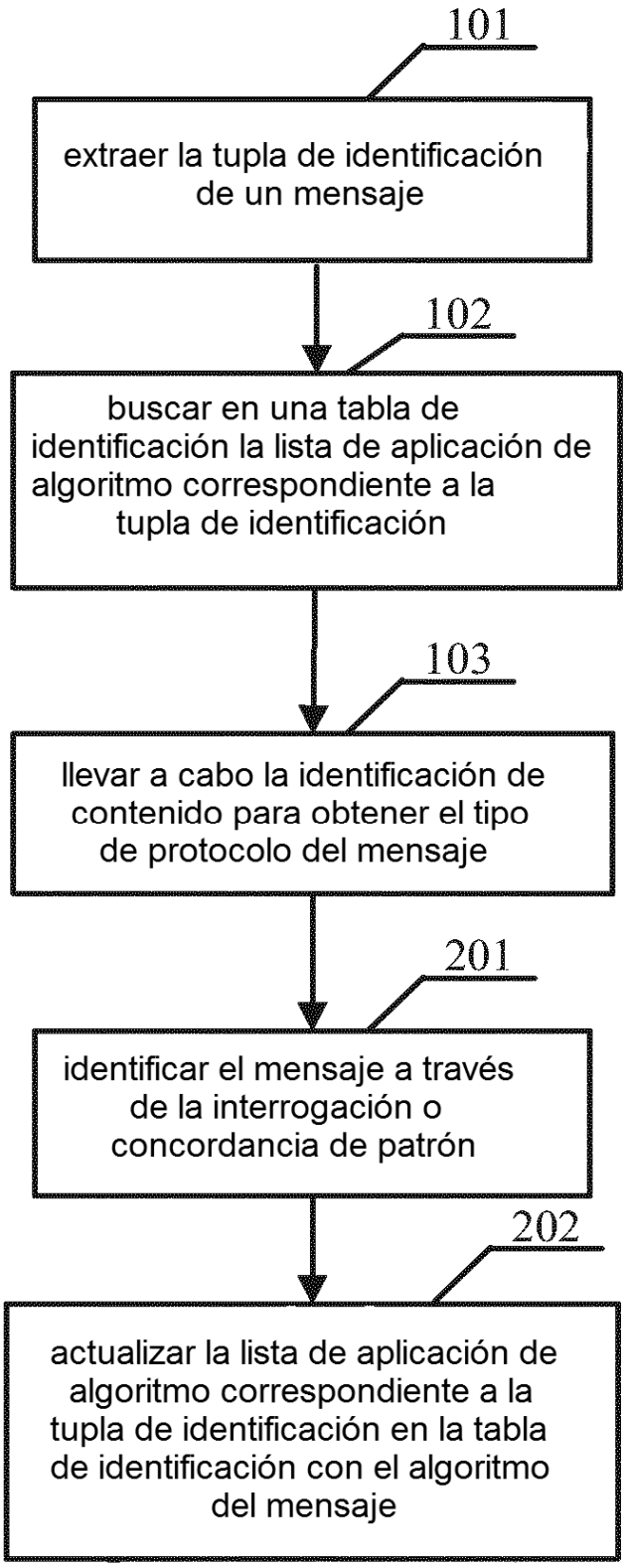
101
extraer la tupla de identificación de un mensaje
102
buscar en una tabla de identificación la lista de aplicación de algoritmo correspondiente a la tupla de identificación
103
llevar a cabo la identificación de contenido para obtener el tipo de protocolo del mensaje
201
identificar el mensaje a través de la interrogación o concordancia de patrón
202
actualizar la lista de aplicación de algoritmo correspondiente a la tupla de identificación en la tabla de identificación con el algoritmo del mensaje

Fig.2

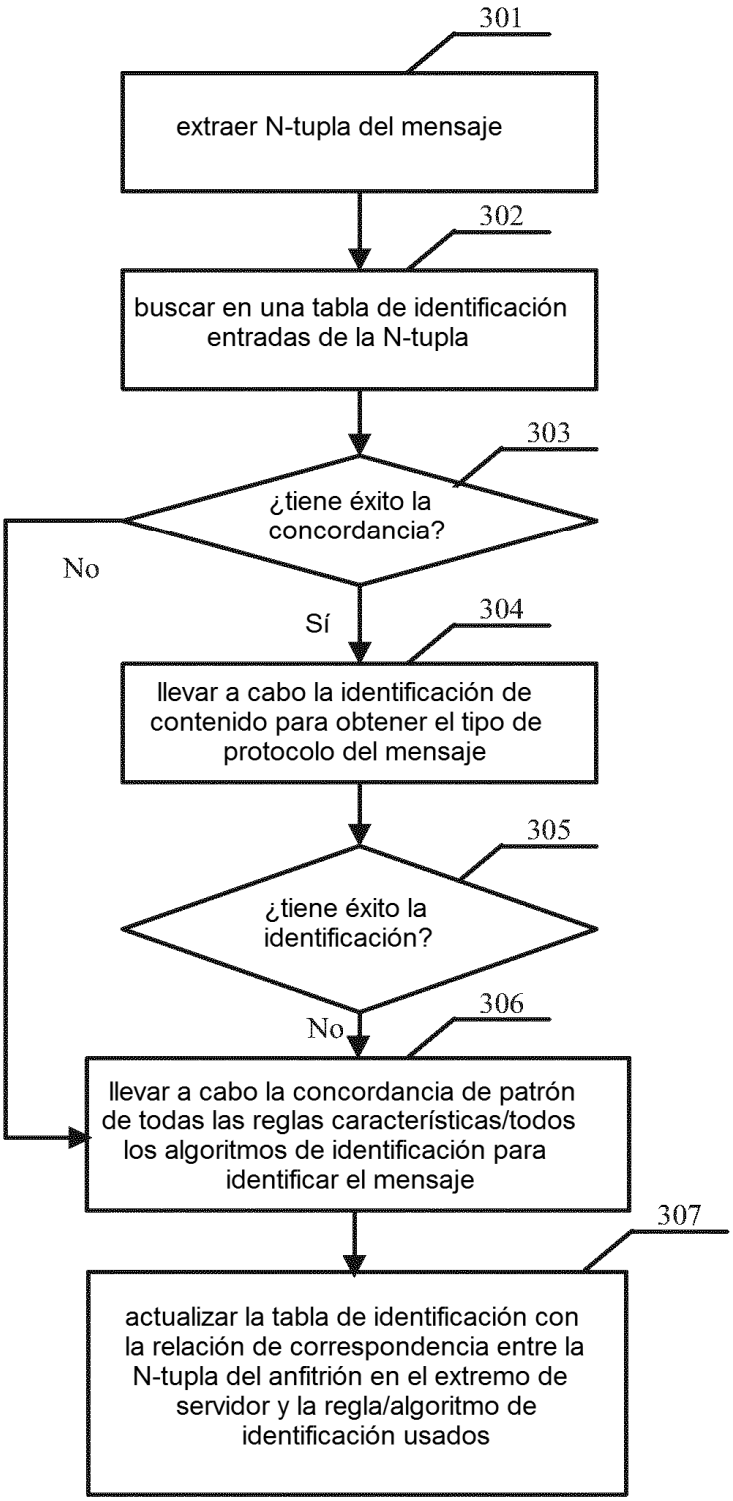
301
extraer N-tupla del mensaje
302
buscar en una tabla de identificación entradas de la N-tupla
303
¿tiene éxito la concordancia?
No
Sí
304
llevar a cabo la identificación de contenido para obtener el tipo de protocolo del mensaje
305
¿tiene éxito la identificación?
No
306
llevar a cabo la concordancia de patrón de todas las reglas características/todos los algoritmos de identificación para identificar el mensaje
307
actualizar la tabla de identificación con la relación de correspondencia entre la N-tupla del anfitrión en el extremo de servidor y la regla/algoritmo de identificación usados

Fig.3

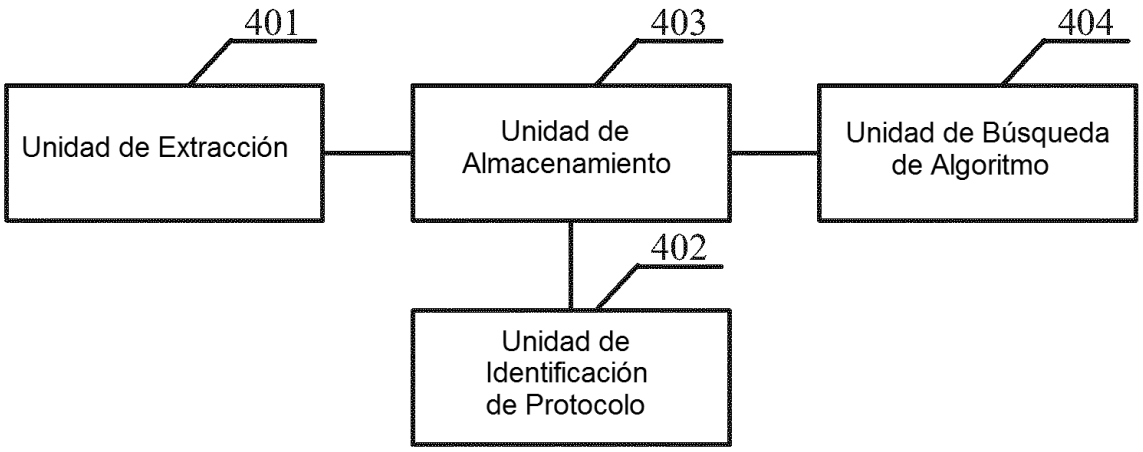
401
403
404
Unidad de Extracción
Unidad de Almacenamiento
Unidad de Búsqueda de Algoritmo
402
Unidad de Identificación de Protocolo

Fig.4

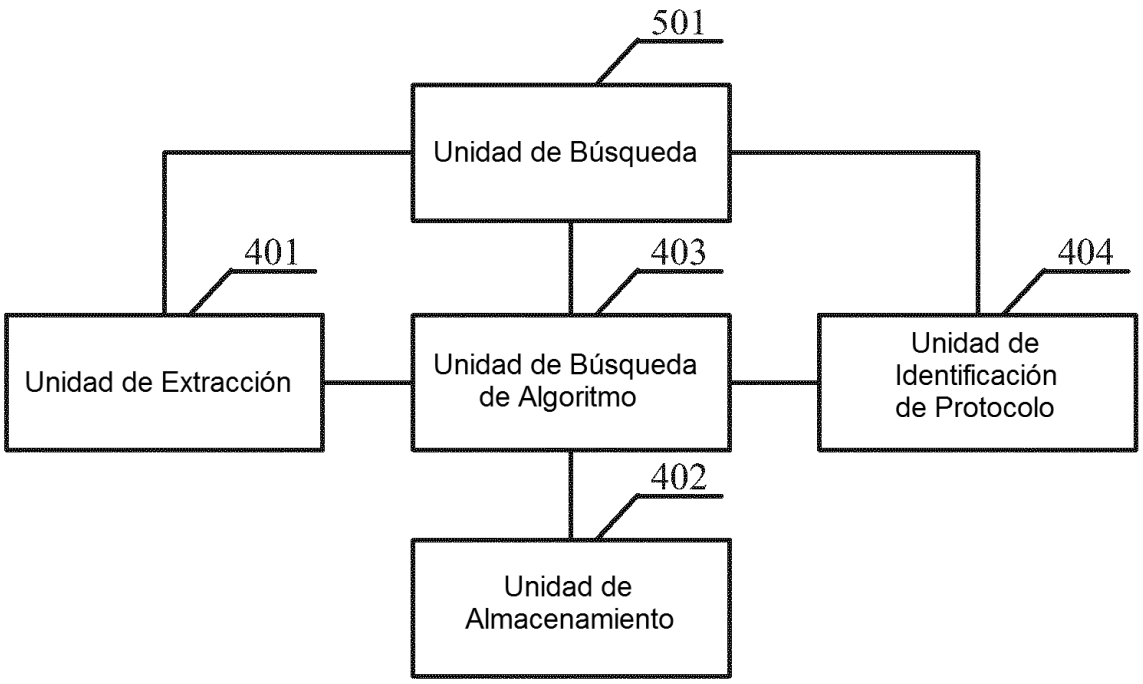
501
Unidad de Búsqueda
401
403
404
Unidad de Extracción
Unidad de Búsqueda de Algoritmo
Unidad de Identificación de Protocolo
402
Unidad de Almacenamiento

Fig.5

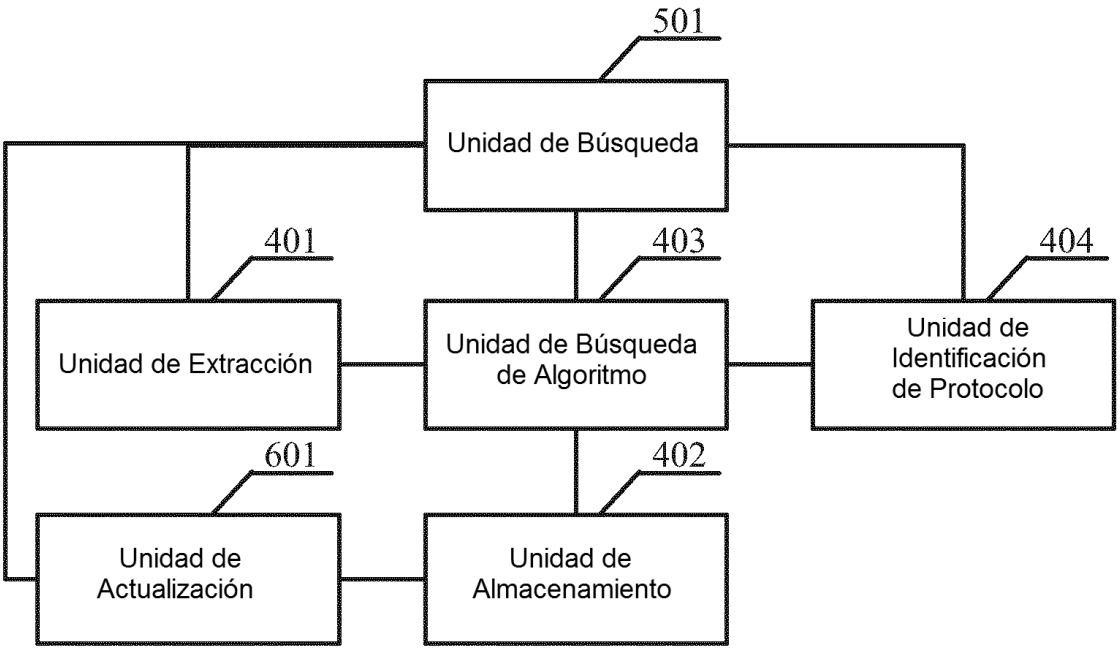
501
Unidad de Búsqueda
401
403
404
Unidad de Extracción
Unidad de Búsqueda de Algoritmo
Unidad de Identificación de Protocolo
601
402
Unidad de Actualización
Unidad de Almacenamiento

Fig.6

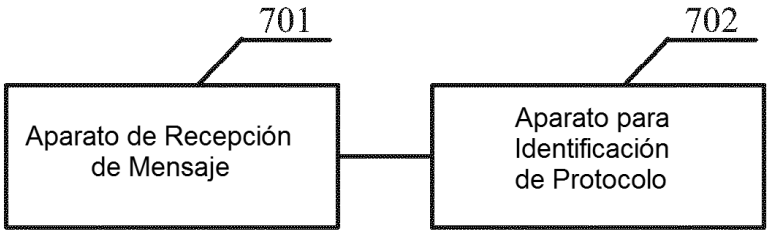
701
702
Aparato de Recepción de Mensaje
Aparato para Identificación de Protocolo

Fig.7