

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 755 763**

51 Int. Cl.:

H04L 29/06 (2006.01)

H04L 29/08 (2006.01)

H04L 12/26 (2006.01)

H04L 12/14 (2006.01)

H04L 29/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.12.2012 PCT/CN2012/087848**

87 Fecha y número de publicación internacional: **03.07.2014 WO14101112**

96 Fecha de presentación y número de la solicitud europea: **28.12.2012 E 12890875 (3)**

97 Fecha y número de publicación de la concesión europea: **02.10.2019 EP 2940954**

54 Título: **Dispositivo y método para identificar un sitio web**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.04.2020

73 Titular/es:
**HUAWEI TECHNOLOGIES CO., LTD. (100.0%)
Huawei Administration Building, Bantian,
Longgang District
Shenzhen, Guangdong 518129, CN**

72 Inventor/es:
YANG, WENHONG

74 Agente/Representante:
ELZABURU, S.L.P

ES 2 755 763 T3

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

DESCRIPCIÓN

Dispositivo y método para identificar un sitio web

Campo técnico

5 La presente invención se refiere al campo de las tecnologías de comunicaciones y, en particular, a un método, un aparato y un sistema de red para identificar un sitio web.

Antecedentes

10 En un proceso de funcionamiento de red existente, con el objetivo de gestionar el comportamiento de visita en la red de un usuario, por ejemplo, filtrar un sitio web visitado por el usuario, realizar una facturación según el sitio web visitado por el usuario, o implementar varios servicios de paquetes de sitio web promocionados por un operador, el sitio web visitado por el usuario tiene que ser identificado. Cuando un usuario visita un sitio web sobre un protocolo de transferencia de hipertexto (en inglés, hypertext transfer protocol - HTTP), el sitio web visitado por el usuario puede ser identificado analizando el contenido de un paquete HTTP del usuario, debido a que el paquete se transmite en formato de texto plano. Sin embargo, cuando el usuario visita un sitio web sobre protocolo seguro de transferencia de hipertexto (en inglés, hypertext transfer protocol secure - HTTPS), el sitio web visitado por el usuario no puede ser
15 identificado a partir de un paquete generado en un proceso en que el usuario visita a el sitio web, debido a que el paquete de la capa de aplicación es encapsulado en un túnel de transmisión cifrado.

20 El documento US 7,778,194 B1 se refiere a una clasificación de un tráfico de red cifrado. Un dispositivo de monitorización del tráfico puede estar dispuesto entre el primer dispositivo de red y el segundo dispositivo de red. El dispositivo de monitorización de tráfico comprende un motor de clasificación de tráfico, y el motor de clasificación de tráfico incluye un módulo de flujo cifrado. El motor de clasificación del tráfico puede utilizar atributos de apretón de manos para clasificar el flujo de datos en una clase de tráfico.

El documento US 2007/180 510 A1 se refiere un sistema para obtener información que se puede utilizar para filtrado de URL.

Compendio

25 Las realizaciones de la presente invención dar a conocer un método y un dispositivo de inspección profunda de paquetes para identificar un sitio web, de tal modo que cuando un usuario visita un sitio web sobre un protocolo HTTPS, el sitio web visitado por el usuario puede ser identificado.

En un primer aspecto, una realización de la presente invención da a conocer un método para identificar un sitio web, que incluye:

30 cuando un cliente visita un sitio web sobre protocolo seguro de transferencia de hipertexto HTTPS, adquirir, mediante un dispositivo de inspección profunda de paquetes, por sus siglas en inglés, DPI, un mensaje de certificado que es generado en un proceso de negociación de clave entre el cliente y un servidor del sitio web;

obtener, analizando sintácticamente mediante el dispositivo DPI el mensaje de certificado, un certificado de servidor del sitio web visitado por el cliente;

35 obtener un valor de clave del certificado de servidor calculando una síntesis digital del certificado de servidor utilizando una función preestablecida;

40 buscar, mediante el dispositivo DPI, en una tabla de valores de clave de sitio web el valor de clave del certificado de servidor, donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene calculando el contenido del certificado de servidor del sitio web utilizando el algoritmo preestablecido; y

si se encuentra en la tabla de valores de clave de sitio web un nombre de sitio web correspondiente al valor de clave del certificado de servidor, identificar, mediante el dispositivo DPI, según el nombre de sitio web, el sitio web visitado por el cliente.

45 En un primer posible modo de implementación del primer aspecto, el método para identificar un sitio web incluye además:

determinar un tipo de formato del certificado de servidor si el nombre de sitio web correspondiente al valor de clave del certificado de servidor no se encuentra en la tabla de valores de clave de sitio web;

50 obtener un nombre de dominio del sitio web a partir de un atributo de sujeto del certificado de servidor, si el tipo de formato del certificado de servidor es un formato X.509, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente; y

obtener el nombre de dominio del sitio web a partir de un atributo de identificador de usuario del certificado de servidor si el tipo de formato del certificado de servidor es un formato PGP, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente.

5 Haciendo referencia al primer aspecto, en un segundo posible modo de implementación del primer aspecto, el método incluye además:

no reunir estadísticas de facturación sobre tráfico del cliente si el nombre de sitio web correspondiente al valor de clave del certificado de servidor se encuentra en la tabla de valores de clave de sitio web; y

reunir estadísticas de facturación sobre el tráfico del cliente si el nombre de sitio web correspondiente al valor de clave del certificado de servidor no se encuentra en la tabla de valores de clave de sitio web

10 Haciendo referencia al primer aspecto, al primer posible modo de implementación del primer aspecto o al segundo posible modo de implementación del primer aspecto, en un tercer posible modo de implementación, adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente y un servidor incluye:

interceptar un paquete que se genera en un proceso de la negociación de clave entre el cliente y el servidor; e

identificar el mensaje de certificado según un campo Apretón de manos en el paquete.

15 En un segundo aspecto, una realización de la presente invención da a conocer un dispositivo de inspección profunda de paquetes, que incluye:

un módulo de adquisición, configurado para, cuando un cliente visita un sitio web sobre protocolo seguro de transferencia de hipertexto HTTPS, adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente y un servidor del sitio web;

20 un módulo de análisis sintáctico, configurado para obtener, mediante el análisis sintáctico del mensaje de certificado adquirido por el módulo de adquisición, un certificado de servidor del sitio web visitado por el cliente;

un módulo de cálculo, configurado para obtener un valor de clave del certificado de servidor mediante calcular una síntesis digital del certificado de servidor utilizando una función preestablecida;

25 un módulo de búsqueda, configurado para buscar en una tabla de valores de clave de sitio web el valor de clave del certificado de servidor obtenido mediante el módulo de cálculo, donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene calculando el contenido del certificado de servidor del sitio web utilizando el algoritmo preestablecido; y

30 un módulo de identificación, configurado para, si el módulo de búsqueda encuentra en la tabla de valores de clave de sitio web un nombre de sitio web correspondiente al valor de clave del certificado de servidor, identificar, según el nombre de sitio web, el sitio web visitado por el cliente.

En un primer posible modo de implementación del segundo aspecto, el dispositivo de inspección profunda de paquetes incluye además:

35 un módulo de ajuste, configurado para recibir y almacenar la tabla de valores de clave de sitio web enviada por un servidor de gestión, donde la tabla de valores de clave de sitio web incluye un nombre de sitio web de por lo menos un sitio web y un valor de clave de un certificado de servidor del sitio web.

Haciendo referencia al segundo aspecto o al primer posible modo de implementación del segundo aspecto, en un segundo posible modo de implementación del segundo aspecto, el dispositivo de inspección profunda de paquetes incluye además:

40 un módulo de determinación, configurado para, cuando el módulo de búsqueda no encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, determinar un tipo de formato del certificado de servidor obtenido mediante el módulo de análisis sintáctico, y activar el módulo de identificación, donde:

45 el módulo de identificación está configurado además para obtener un nombre de dominio del sitio web a partir de un atributo de sujeto del certificado de servidor cuando el módulo de determinación determina que el tipo de formato del certificado de servidor del sitio web es un formato X.509, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente; y obtener el nombre de dominio del sitio web a partir de un atributo de identificador de usuario del certificado de servidor cuando el módulo de determinación determina que el tipo de formato del certificado de servidor del sitio web es un formato PGP, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente.

50 Haciendo referencia al segundo aspecto o al primer posible modo de implementación del segundo aspecto, en un

tercer posible modo de implementación del segundo aspecto, el dispositivo de inspección profunda de paquetes incluye además:

5 un módulo de gestión de tráfico, configurado para no reunir estadísticas de facturación sobre tráfico del cliente cuando el módulo de búsqueda encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web; y para reunir estadísticas de facturación sobre el tráfico del cliente cuando el módulo de búsqueda no en cuenta el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web.

10 Haciendo referencia al segundo aspecto, al primer posible modo de implementación del segundo aspecto, al segundo posible modo de implementación del segundo aspecto o al tercer posible modo de implementación del segundo aspecto, en un cuarto posible modo de implementación del segundo aspecto, el módulo de adquisición está configurado específicamente para interceptar un paquete que se genera en un proceso de la negociación de clave entre el cliente y el servidor del sitio web, e identificar el mensaje de certificado según el campo Apretón de manos en el paquete.

15 Según el método para identificar un sitio web dado a conocer en las realizaciones de la presente invención, cuando un cliente visita un sitio web sobre un protocolo HTTPS, se adquiere un certificado de servidor del sitio web visitado por el cliente, se obtiene un valor de clave del certificado de servidor calculando una síntesis digital de certificado de servicio según un algoritmo preestablecido, se busca en una tabla de valores de clave de sitio web el valor de clave obtenido, y el sitio web visitado por el cliente es identificado utilizando un nombre de sitio web que corresponde al valor de clave del certificado de servidor y se encuentra en la tabla de valores de clave de sitio web. De este modo, el sitio web visitado por el cliente se puede identificar incluso cuando el cliente visita el sitio web sobre protocolo HTTPS.

Breve descripción de los dibujos

25 Para describir más claramente las soluciones técnicas de las realizaciones de la presente invención o de la técnica anterior, a continuación se introducen brevemente los dibujos adjuntos necesarios para describir las realizaciones o la técnica anterior. Evidentemente, los dibujos adjuntos en la siguiente descripción muestran tan sólo algunas realizaciones de la presente invención, y los expertos en la materia pueden obtener sin esfuerzos creativos otros dibujos a partir de estos dibujos adjuntos.

La figura 1 es un diagrama de un escenario de aplicación de un método para identificar un sitio web, según una realización de la presente invención;

30 La figura 2 es un diagrama de flujo de un método para identificar un sitio web, según una realización de la presente invención;

La figura 3 es un diagrama de flujo de otro método para identificar un sitio web, según una realización de la presente invención;

La figura 4 es un diagrama de señalización de una negociación de clave realizada entre un cliente y servidor de un sitio web, según una realización de la presente invención;

35 La figura 5 es un diagrama de flujo de un método para analizar un certificado de servidor en un método para identificar un sitio web, según una realización de la presente invención;

La figura 6 es un diagrama esquemático de una estructura física de un dispositivo de inspección profunda de paquetes, según una realización de la presente invención;

40 La figura 7 es un diagrama estructural esquemático de otro dispositivo de inspección profunda de paquetes, según una realización de la presente invención; y

La figura 8 es un diagrama estructural esquemático de un sistema para identificar un sitio web, según una realización de la presente invención.

Descripción de las realizaciones

45 Para hacer más comprensibles para los expertos en la materia las soluciones de la presente invención, a continuación se describen de manera clara y completa las soluciones técnicas de las realizaciones de la presente invención, haciendo referencia a los dibujos adjuntos de las realizaciones de la presente invención. Evidentemente, las realizaciones a describir son tan sólo una parte y no la totalidad de las realizaciones de la presente invención.

50 Tal como se muestra en la figura 1, la figura 1 muestra un escenario de una realización de la presente invención. En el escenario de aplicación mostrado en la figura 1, un cliente 100 visita un sitio web sobre protocolo seguro de transferencia de hipertexto (protocolo seguro de transferencia de hipertexto, HTTPS). Un dispositivo de pasarela 105 puede llevar a cabo control de acceso de red sobre el comportamiento de visita en la red del cliente 100. Un dispositivo de inspección profunda de paquetes (inspección profunda de paquetes, DPI) 110 puede llevar a cabo inspección y control del tráfico basado en la capa de aplicación sobre el comportamiento de visita en la red del cliente 100, puede

5 estar configurado para inspeccionar y analizar tráfico de visita en la red del cliente 100 y puede implementar una función de cumplimiento de políticas y facturación (en inglés, Policy and Charging Enforcement Function - PCEF), solicitar una cuota de tráfico de un usuario a partir de un sistema de facturación en línea (en inglés, Online Charging System - OCS) 115, y notificar tráfico del cliente 100, de tal modo que la OCS puede facturar al usuario en función del tráfico de visita del cliente 100. El dispositivo DPI 110 se despliega en un enlace descendente del dispositivo de pasarela 105, y un paquete de datos generado en un proceso en que el cliente 100 visita el sitio web pasa a través del dispositivo de pasarela y del dispositivo DPI y llega a continuación a un servidor del sitio web 120 a visitar. El dispositivo DPI 110 analiza el tráfico del cliente 100 para identificar el sitio web visitado por el cliente 110, y controla el comportamiento de visita en la red del cliente 100 según el sitio web identificado, por ejemplo, si un determinado usuario participa en un servicio de paquetes de facturación, el dispositivo DPI 110 realiza la gestión de facturación sobre el tráfico de visita del cliente 100 del usuario según una política de facturación de paquetes establecida por un dispositivo de función de reglas de políticas y facturación (en inglés, Policy and Charging Rules Function - PCRF) 125 para el usuario. Si el sitio web visitado por el cliente 100 es un sitio web del paquete, el dispositivo DPI 110 no reúne estadísticas de tráfico sobre el tráfico de visita en la red del cliente 100, y si el sitio web visitado por el cliente 100 no es un sitio web del paquete, el dispositivo DPI 110 reúne estadísticas de tráfico sobre el tráfico de visita en la red del cliente 100 y notifica el resultado de las estadísticas de tráfico del cliente 100 al sistema de facturación en línea 115 para el proceso de facturación. Se puede comprender que la gestión de facturación realizada sobre el comportamiento de visita en la red del cliente 100 según el sitio web identificado en este escenario de aplicación, es solamente un tipo de gestión sobre el comportamiento de red del cliente 100, y en una aplicación práctica, el control de tráfico, el filtrado o similares pueden además ser realizados por el usuario según el sitio web identificado.

Se puede comprender que el cliente 100 puede ser un dispositivo que puede implementar visita en la red, tal como un teléfono móvil y un ordenador, y el dispositivo de pasarela 105 puede ser un nodo de soporte GPRS de pasarela (en inglés, Gateway GPRS Support Node - GGSN), una pasarela de datos de paquetes (en inglés, Packet Data Gateway - PDG) o similares, lo cual no se limita en la presente memoria.

25 La figura 2 es un diagrama de flujo de un método para identificar un sitio web, según una realización de la presente invención. El método se aplica a un escenario en el que un cliente visita un sitio web sobre un protocolo HTTPS y puede ser ejecutado por el dispositivo DPI 110 mostrado en la figura 1. El método incluye:

30 Etapa 200: cuando un cliente visita un sitio web sobre un protocolo HTTPS, adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente y un servidor del sitio web. A continuación, avanzar a la etapa 205.

Haciendo referencia a lo mostrado en la figura 1, cuando el cliente 100 visita un sitio web sobre un protocolo HTTPS, el cliente 100 realiza una negociación de clave con el servidor 120 del sitio web visitado, y el proceso de negociación de clave se puede llevar a cabo según un protocolo de seguridad de capa de transporte (en inglés, Transport Layer Security - TLS), donde el protocolo TLS es un protocolo de cifrado que proporciona un canal confidencial y seguro en internet, un certificado de servidor es un certificado que se utiliza para demostrar un tipo de utilización del servidor y es aplicado por el servidor a un centro de autoridad de certificación (en inglés, Certificate authority - CA), y el cliente confía en el servidor solamente cuando el certificado es utilizado en el servidor correspondiente. Debido a que el dispositivo DPI 110 está desplegado en el enlace descendente del dispositivo de pasarela 105, el dispositivo DPI 110 puede adquirir, mediante interceptación, un mensaje de certificado Certificado que es generado en un proceso de la negociación de clave entre el cliente 100 y el servidor del sitio web 120.

40 Etapa 205: obtener, mediante el análisis sintáctico del mensaje de certificado, un certificado de servidor del sitio web visitado por el cliente. A continuación, avanzar a la etapa 210.

45 En algunos casos, si el mensaje de certificado adquirido incluye una lista de certificados Lista de certificado, el certificado de servidor del sitio web visitado se puede obtener extrayendo un primer certificado de una parte de lista de certificados Lista de certificado en el mensaje de certificado.

Etapa 210: obtener un valor de clave del certificado de servidor, según un algoritmo preestablecido. A continuación, avanzar a la etapa 215.

50 En esta realización de la presente invención, el contenido del certificado de servidor puede ser calculado utilizando el algoritmo preestablecido para obtener el valor de clave del certificado de servidor. Por ejemplo, el contenido del certificado de servidor se puede calcular adoptando un algoritmo de síntesis digital para obtener el valor de clave del certificado de servidor, donde el algoritmo de síntesis digital puede ser un algoritmo de síntesis de mensajes 5 (en inglés, Message-Digest Algorithm - MD5), un algoritmo resumen ("hash") seguro SHA-1, un algoritmo de síntesis de mensaje de evaluación de primitivas de integridad RACE (en inglés, RACE Integrity Primitives Evaluation Message Digest - RIPEMD), o similares, lo que no se limita en la presente memoria.

55 Etapa 215: buscar una tabla de valores de clave de sitio web según el valor de clave del certificado de servidor, donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene según el algoritmo preestablecido. A continuación, avanzar a la etapa 220.

En esta realización de la presente invención, la tabla de valores de clave de sitio web puede estar preestablecida en el dispositivo DPI, y la tabla de valores de clave de sitio web preestablecida registra varios nombres de sitio web y valores de clave de certificados de servidor de sitios web, donde los valores de clave de los certificados de servidor en la tabla de valores de clave de sitio web se obtienen calculando el contenido de certificados de servidor de varios sitios web según el mismo algoritmo que se ha adoptado en la etapa 210, por ejemplo, adoptando el mismo algoritmo de síntesis digital. Por ejemplo, si en la etapa 215 los valores de clave de los certificados de servidor en la tabla de valores de clave de sitio web preestablecida se obtienen utilizando el algoritmo MD5, en la etapa 210, cuando un usuario visita un sitio web, un valor de clave de un certificado de servidor del sitio web visitado por el usuario tiene asimismo que ser calculado utilizando el algoritmo MD5. Definitivamente, se puede entender que la tabla de valores de clave de sitio web puede asimismo estar preestablecida en otro dispositivo, siempre que pueda ser accedida por el dispositivo DPI.

Etapa 220: si se encuentra un nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, identificar, según el nombre de sitio web, el sitio web visitado por el cliente.

Según el algoritmo de síntesis digital, tal como MD5, se puede obtener una síntesis digital (un valor resumen, es decir, el valor de clave en esta realización) después de que se realice el cálculo de resumen sobre la información de primitivas. Por lo tanto, se utiliza una característica del algoritmo de síntesis digital en esta realización de la presente invención, donde la característica del algoritmo de síntesis digital indica que diferentes primitivas se pueden convertir en diferentes valores resumen, el contenido de un certificado de servidor de un sitio web que tiene que ser identificado se calcula previamente utilizando un algoritmo de síntesis digital preestablecido para obtener un valor de clave del certificado de servidor, y el valor de clave del certificado de servidor obtenido y el nombre del sitio web se almacenan en consecuencia en la tabla de valores de clave de sitio web. Cuando el dispositivo DPI tiene que identificar el sitio web visitado por el usuario, un valor de clave de un certificado de servidor adquirido a partir de un flujo de datos generado en un proceso de visita del usuario se puede comparar con un valor preestablecido de clave del certificado de servidor, y si un valor de clave de un certificado de servidor de un sitio web, que es el mismo valor de clave del certificado de servidor del sitio web visitado por el cliente, se encuentra en la tabla de valores de clave de sitio web, el nombre de un sitio web correspondiente al valor de clave del certificado de servidor se puede obtener según la tabla de valores de clave de sitio web.

Según el método para identificar un sitio web dado a conocer en esta realización de la presente invención, cuando un cliente visita un sitio web sobre un protocolo HTTPS, se adquiere un certificado de servidor del sitio web visitado por el cliente, se obtiene un valor de clave del certificado de servidor según un algoritmo preestablecido, se busca en una tabla de valores de clave de sitio web el valor de clave obtenido, y el sitio web visitado por el cliente se identifica utilizando un nombre de sitio web que corresponde al valor de clave del certificado de servidor y se encuentra en la tabla de valores de clave de sitio web. De este modo, el sitio web visitado por el cliente se puede identificar incluso cuando el cliente visita el sitio web sobre protocolo HTTPS.

La figura 3 es un diagrama de flujo de otro método para identificar un sitio web, según una realización de la presente invención. El método se aplica a un escenario en el que un cliente visita un sitio web sobre un protocolo HTTPS y puede ser ejecutado por el dispositivo DPI 110 mostrado en la figura 1. Tal como se muestra en la figura 3, el método incluye:

Etapa 300: recibir y almacenar una tabla de valores de clave de sitio web enviada por un servidor de gestión, donde la tabla de valores de clave de sitio web incluye un nombre de sitio web de por lo menos un sitio web y un valor de clave de un certificado de servidor del sitio web. A continuación, avanzar a la etapa 305.

En una aplicación práctica, el servidor de gestión puede adquirir previamente un certificado de servidor de por lo menos un sitio web, calcular un valor de clave del certificado de servidor adquirido de dicho por lo menos un sitio web según un algoritmo de síntesis digital preestablecido, registrar el valor de clave del certificado de servidor obtenido por medio del cálculo y un nombre de sitio web en una tabla de valores de clave de sitio web, y enviar la tabla de valores de clave de sitio web al dispositivo DPI.

Específicamente, el servidor de gestión puede adquirir, previamente, un certificado de servidor de un sitio web que se tiene que involucrar, por ejemplo, cuando se tiene que llevar a cabo la gestión de facturación sobre tráfico de usuarios, se puede adquirir previamente un certificado de servidor de un sitio web que puede ser visitado según lo estipulado en un paquete de sitio web, tal como un certificado de servidor de un sitio web, como Facebook o YouTube; y un certificado de servidor de un sitio web en el que es necesario realizar control de visitas puede asimismo adquirirse previamente según un requisito, que se puede establecer específicamente según un requisito y no se limita en la presente memoria. Además, el servidor de gestión puede visitar previamente, utilizando un navegador, un sitio web que se tiene que involucrar, para descargar a local un certificado de servidor del sitio web, y realizar a continuación el cálculo sobre el certificado de servidor descargado utilizando un algoritmo preestablecido (por ejemplo, un algoritmo MD5) para adquirir un valor de clave del certificado de servidor, con el fin de obtener una tabla de valores de clave de sitio web. Una forma de la tabla de valores de clave de sitio web puede ser la mostrada en la tabla 1:

Tabla 1 Tabla de valores de clave de sitio web

Número de serie	Nombre de dominio de sitio web	Valor de comprobación del certificado
1	Facebook	2C89277EE96E8B15E2F9A90FED69B0B2
2	YouTube	5F43086FACDB301B74334183A0472584
...

5 Se puede comprender que cuando se actualiza un certificado de servidor de un sitio web o hay que añadir o eliminar un sitio web que requiere atención, es necesario actualizar temporalmente una tabla de valores de clave de sitio web preestablecida. Específicamente, cuando se actualiza un certificado de servidor de un sitio web que requiere atención, un valor de clave del certificado de servidor en la tabla de valores de clave de sitio web puede ser actualizado en tiempo según el certificado actualizado de servidor del sitio web; cuando se tiene que añadir un sitio web que requiere atención, se añade un valor de clave de una certificación de servidor y un nombre de sitio web a la tabla de valores de clave de sitio web según el certificado de servidor del sitio web recién añadido; o cuando se tiene que eliminar un sitio web que requiere atención, se puede eliminar de la tabla de valores de clave de sitio web un valor de clave de una certificación de servidor correspondiente al nombre del sitio web que se tiene que eliminar.

10 Etapa 305: adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre un cliente y un servidor, donde el mensaje de certificado lleva un certificado de servidor de un sitio web visitado por el cliente. A continuación, avanzar a la etapa 310.

15 Haciendo referencia a lo mostrado en la figura 1, cuando el cliente 100 visita un sitio web sobre un protocolo HTTPS, dado que el dispositivo DPI 110 está desplegado en el enlace descendente del dispositivo de pasarela 105, el dispositivo DPI puede adquirir, mediante interceptación, un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente 100 y el servidor de sitio web 120, donde el mensaje de certificado lleva un certificado de servidor del sitio web visitado por el cliente. Específicamente, un proceso de negociación de clave basado en un protocolo TLS, y entre el cliente y el servidor, se puede mostrar específicamente en la figura 4, y el proceso de negociación de clave puede incluir:

20 Etapa 400: un cliente envía un mensaje de solicitud de negociación de clave Saludo cliente a un servidor de un sitio web a visitar. A continuación, avanzar a la etapa 405.

25 En el mensaje Saludo cliente, el cliente notifica al servidor un modo de cifrado que puede ser soportado por el cliente. A continuación, avanzar a la etapa 405.

Etapa 405: el servidor devuelve un mensaje de respuesta de negociación de clave Saludo servidor al cliente. A continuación, avanzar a la etapa 410.

30 El servidor que recibe el mensaje de Saludo cliente devuelve un mensaje de Saludo servidor al cliente y especifica, en el mensaje de Saludo servidor, que se adopta un modo de cifrado entre los modos de cifrado que pueden ser soportados por el cliente.

Etapa 410: el servidor envía un mensaje de certificado Certificado al cliente, donde el mensaje de certificado incluye un certificado de servidor de servidor. A continuación, avanzar a la etapa 415.

35 Después de enviar el mensaje Saludo servidor al cliente, el servidor envía a continuación mensaje de certificado, donde el mensaje de certificado incluye un certificado de servidor de servidor, de tal modo que el cliente verifica el certificado de servidor. El certificado de servidor es un certificado que se utiliza para demostrar un tipo de utilización del servidor y es solicitado por el servidor a un centro de autoridad de certificación (autoridad de certificación, CA), y el cliente confía en el servidor solamente cuando el certificado es utilizado en el servidor correspondiente. Un formato del mensaje de certificado puede ser el descrito en la tabla 2:

Tabla 2 Formato de mensaje de certificado

Lista de certificado		
Longitud Hi	Tipo de apretón de manos	Longitud
Tipo de contenido	Versión	Longitud Lo
TCP		
IP		
Trama		

Que el valor de un campo Tipo de apretón de manos sea de 11 indica que el mensaje es un mensaje de certificado.

5 Se debe observar que el mensaje de certificado incluye generalmente una lista de certificados Lista de certificado, la lista de certificados puede incluir múltiples certificados, y entre los múltiples certificados, un certificado enumerado en el reverso se utiliza para verificar si un certificado enumerado en el anverso es un certificado sobre el que se realiza satisfactoriamente la autenticación de una autoridad. Por ejemplo, cuando se visita el sitio web de Google sobre HTTPS, se adquieren múltiples certificados, donde un primer certificado es un certificado de servidor de Google, un certificado en el reverso puede ser un certificado de una determinada autoridad de certificación, y el certificado de la autoridad de certificación se utiliza para demostrar que el certificado de servidor de Google es un certificado legal. Por lo tanto, cuando se obtiene una lista de certificados Lista de certificado, un primer certificado es un certificado de servidor de un sitio web a visitar.

15 Etapa 415: después de que la verificación del cliente en el certificado de servidor del servidor sea satisfactoria, el cliente envía un mensaje de intercambio de claves Intercambiador de claves cliente al servidor, donde el mensaje Intercambiador de claves cliente incluye una clave que se cifra adoptando una clave pública del servidor. A continuación, avanzar a la etapa 420.

Etapa 420: el cliente envía al servidor un mensaje de negociación de clave finalizada, que indica que la negociación se ha completado. A continuación, avanzar a la etapa 425.

20 Etapa 425: el servidor devuelve un mensaje de negociación de clave finalizada el cliente, que indica que la negociación se ha completado. A continuación, avanzar a la etapa 430.

Etapa 430: el cliente y el servidor intercambian un paquete de datos, donde el paquete de datos es un paquete de datos que está cifrado mediante la clave negociada.

25 Se puede entender que el dispositivo DPI se puede desplegar asimismo en una conexión de red a modo de despliegue de derivación, y cuando el dispositivo DPI es desplegado en un modo de derivación, el dispositivo DPI puede obtener, a modo de espejo o de copia, un paquete de negociación que se genera en un proceso de la negociación de clave entre el cliente 100 y el servidor de sitio web 120, para obtener el mensaje de certificado en el proceso de la negociación de clave entre el cliente 100 y el servidor de sitio web 120.

Etapa 310: obtener el certificado de servidor del sitio web visitado analizando sintácticamente el mensaje de certificado. A continuación, avanzar a la etapa 315.

30 Etapa 315: obtener un valor de clave del certificado de servidor del sitio web visitado según un algoritmo preestablecido. A continuación, avanzar a la etapa 320.

35 Específicamente, el contenido del certificado de servidor puede calcularse utilizando el algoritmo preestablecido para obtener el valor de clave del certificado de servidor. Por ejemplo, el contenido del certificado de servidor se puede calcular adoptando un algoritmo de síntesis digital para obtener un valor de clave del certificado de servidor, donde el algoritmo de síntesis digital puede incluir un algoritmo, tal como MD5 o SHA-1, y el algoritmo adoptado en esta etapa tiene que ser el mismo que el algoritmo adoptado cuando el servidor de gestión establece una tabla de valores de clave de sitio web preestablecida.

Etapa 320: buscar en la tabla de valores de clave de sitio web según el valor de clave del certificado de servidor. A continuación, avanzar a la etapa 325.

40 Etapa 325: determinar si se encuentra en la tabla de valores de clave de sitio web un nombre de sitio web correspondiente al valor de clave del certificado de servidor; si se encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor, avanzar a la etapa 330; y de lo contrario, avanzar a la etapa 335.

Etapa 330: identificar, según el nombre de sitio web, el sitio web visitado por el cliente. A continuación, avanzar a la etapa 340.

5 Etapa 335: analizar el certificado de servidor del sitio web para obtener un nombre de dominio del sitio web visitado por el cliente, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente. A continuación, avanzar a la etapa 340.

10 Si no se encuentra el mismo valor de clave del certificado de servidor en la tabla de valores de clave de sitio web preestablecida, según el valor de clave obtenido del certificado de servidor del sitio web, en algunos casos, con el objetivo de identificar el nombre del sitio web visitado por el cliente, el nombre del sitio web visitado por el cliente se puede identificar realizando análisis en línea del certificado de servidor del sitio web visitado obtenido en la etapa 310. Se puede mostrar un análisis detallado en la figura 5, que incluye:

Etapa 505: determinar un tipo de formato del certificado de servidor del sitio web visitado; si el certificado de servidor del sitio web visitado es un certificado de servidor en un formato X.509, avanzar a la etapa 510, y si el certificado de servidor del sitio web visitado es un certificado de servidor en un formato PGP, avanzar a la etapa 520.

15 Los expertos en la materia pueden saber que los certificados de servidor soportados actualmente por un protocolo TLS incluyen certificados de servidor en dos formatos: un certificado de servidor en un formato X.509 y un certificado de servidor en un formato PGP. Durante el análisis detallado del certificado de servidor, es necesario identificar primero el tipo de formato del certificado de servidor obtenido en la etapa 310, y a continuación se realiza un análisis detallado sobre el certificado de servidor según una estructura de datos de un certificado de un tipo de formato diferente. Específicamente, en un proceso de negociación, el formato del certificado de servidor se puede negociar utilizando un campo Tipo de certificado en el mensaje Saludo cliente y el mensaje Saludo servidor. Si el campo Tipo de certificado es 0, esto indica que el certificado de servidor es un certificado de servidor en un formato X.509, y si el campo Tipo de certificado es 1, esto indica que el certificado de servidor es un certificado en el formato PGP. X.509 es un estándar de certificado digital utilizado ampliamente y un estándar de infraestructura de clave pública (en inglés, Public Key Infrastructure - PKI) formulado por el sector de estandarización de telecomunicaciones de la unión internacional de telecomunicaciones (ITU-T) para autenticación única (en inglés, Single Sign-on - SSO) e infraestructura de gestión de privilegios (en inglés, Privilege Management Infrastructure - PMI); y PGP (en inglés, Pretty Good Privacy, privacidad bastante buena) es un software de cifrado de correo basado en un sistema de cifrado de clave pública RSA, y utiliza un par de claves relacionadas matemáticamente, donde una (una clave pública) se utiliza para cifrar información y la otra (una clave privada) se utiliza para descifrar información. Cada clave pública y clave privada en PGP van acompañadas por un certificado de clave, es decir, un certificado de clave en formato PGP, donde el formato PGP incluye un certificado en un formato abierto Abierto PGP.

Etapa 510: analizar sintácticamente el certificado de servidor del sitio web visitado, según una estructura de datos de un certificado en el formato X.509, para obtener un atributo de sujeto en el certificado de servidor. A continuación, avanzar a la etapa 515.

35 El certificado en el formato X.509 incluye generalmente:

Certificado certificado;

Versión versión;

Número de serie número de serie;

Identificación de algoritmo identificación de algoritmo ;

40 Emisor emisor;

Validez validez;

Sujeto sujeto;

Información de clave pública de sujeto información de clave pública de sujeto;

Algoritmo de clave pública algoritmo de clave pública;

45 Clave pública de sujeto clave pública ;

Algoritmo de firma de certificado algoritmo de firma de certificado;

Información de firma de certificado, como firma de certificado

50 El atributo "Sujeto sujeto" tiene información de nombre de dominio del sitio web. Por lo tanto, si el certificado de servidor del sitio web visitado obtenido en la etapa 310 es un certificado en el formato X.509, se puede obtener un atributo de sujeto del certificado según una estructura de datos del certificado en el formato X.509. Definitivamente, se puede

entender que cuando se determina un formato digital del certificado de servidor, es necesario determinar además la versión adoptada por el certificado, y el certificado se analiza sintácticamente según un formato de versión diferente.

Etapa 515: obtener un nombre de dominio del sitio web visitado, analizando sintácticamente el atributo de sujeto.

5 Por ejemplo, el nombre de dominio del sitio web visitado por el cliente puede ser obtenido analizando sintácticamente el atributo de sujeto del certificado de servidor del sitio web visitado. Por ejemplo, en el certificado de servidor en el formato X.509, un formato del atributo de sujeto puede ser: www.facebook.com.

Etapa 520: obtener un atributo de identificador de usuario analizando sintácticamente un certificado en formato PGP. A continuación, avanzar a la etapa 525.

El certificado de clave del PGP incluye generalmente el siguiente contenido:

10 contenido de clave: generalmente es una clave indicada utilizando un número grande con la posición de las centenas; tipo de clave: indica si la clave es una clave pública o una clave privada;

longitud de la clave: se utiliza para indicar la longitud de la clave y se indica generalmente utilizando un dígito binario;

número de clave: se utiliza para identificar de manera única la clave;

tiempo de creación: se utiliza para indicar el tiempo en el que se creó certificado;

15 indicador de usuario: indica información sobre el creador de la clave, por ejemplo, el nombre o un correo electrónico del creador;

huella digital de la clave: es un número con 128 dígitos y es un resumen del contenido de la clave e indica una característica única de la clave; y

20 firma de intermediario: indica una firma digital de un intermediario y se utiliza para declarar la autenticidad de la clave y su propietario, e incluye un número de clave e información de identificador del intermediario.

La parte de "identificador de usuario" se utiliza para identificar la información sobre el creador de la clave, y la información se puede indicar utilizando el nombre, correo electrónico o similar, del creador. Por lo tanto, si el certificado de servidor del sitio web es un certificado de clave en el formato PGP, un identificador de un sitio web se puede aprender a partir de información en la parte de "identificador de usuario" en el certificado de servidor. De este modo, se puede identificar el sitio web visitado por el cliente.

25 Etapa 525: obtener el nombre de dominio del sitio web visitado, mediante analizar sintácticamente el atributo de identificador de usuario.

30 Específicamente, en el certificado de servidor en el formato PGP, el nombre de dominio del sitio web está incluido en el atributo de identificador de usuario, donde el identificador de usuario puede ser un identificador que puede identificar a un usuario, por ejemplo, una dirección de correo electrónico o un nombre. Por ejemplo, en el certificado de servidor en el formato PGP, el identificador de usuario se puede representar como: administrator@facebook.com.

Etapa 340: gestionar un flujo de datos generado en un proceso de visita del cliente, según el sitio web identificado visitado por el cliente.

35 Después de que se identifique el sitio web visitado por el cliente, se puede gestionar el comportamiento de visita en la red del usuario, según el sitio web identificado visitado por el cliente y con una política de control preestablecida, por ejemplo, si es necesario llevar a cabo gestión de facturación sobre el tráfico del cliente, o se pueden determinar las necesidades de control o de filtrado a llevar a cabo sobre el tráfico de visita del cliente según el nombre del sitio web visitado y con una política de paquetes de sitio web preestablecida, lo cual no se limita en la presente memoria.

40 Según el método para identificar un sitio web dado a conocer en esta realización de la presente invención, cuando un cliente visita un sitio web sobre un protocolo HTTPS y el sitio web visitado por el cliente no puede ser identificado utilizando una tabla de valores de clave de sitio web preestablecida, el sitio web visitado por el cliente se puede identificar a modo de realizar un análisis en línea de un certificado de servidor del sitio web visitado por el cliente. El método para identificar un sitio web dado a conocer en esta realización de la presente invención es más completo y la precisión de identificación del sitio web es superior. Además, según el método para identificar un sitio web dado a conocer en esta realización de la presente invención, cuando se identifica el sitio web visitado por el cliente, el comportamiento de visita en la red de un usuario se puede gestionar en tiempo según el nombre del sitio web identificado, proporcionando de ese modo para el usuario una gestión de red más oportuna, eficiente y elegante.

45 En un escenario de aplicación, cuando el método para identificar un sitio web descrito en esta realización de la presente invención se utiliza para llevar a cabo gestión de facturación de tráfico sobre un usuario, una tabla de valores de clave de sitio web preestablecida puede incluir un nombre de dominio de un sitio web que puede ser visitado según se

50

estipula en un paquete suscrito por el usuario y un valor de clave de un certificado de servidor del sitio web. Cuando se determina, en la etapa 325, que un valor de clave de un certificado de servidor de un sitio web, que es el mismo que el valor de clave del certificado de servidor del sitio web visitado, se encuentra en la tabla de valores de clave de sitio web, esto indica que el sitio web visitado por el cliente está en el paquete, y pueden no reunirse estadísticas de tráfico de facturación sobre la visita del cliente; y cuando se determina en la etapa 325 que un valor de clave de un certificado de servidor de un sitio web, que es el mismo que el valor de clave del certificado de servidor del sitio web visitado, no se encuentra en la tabla de valores de clave de sitio web, esto indica que el sitio web visitado por el cliente no está en el paquete, se pueden reunir directamente estadísticas de tráfico sobre el tráfico de visita del cliente en tiempo, y las estadísticas de tráfico reunidas son notificadas para facturación a un servidor de facturación. En este escenario de aplicación, cuando un valor de clave de un certificado de servidor de un sitio web, que es el mismo que el valor de clave del certificado de servidor del sitio web visitado, no se encuentra en la tabla de valores de clave de sitio web en la etapa 325, es innecesario avanzar a la etapa 335 para obtener, analizando el certificado de servidor del sitio web, el nombre de dominio del sitio web visitado por el usuario, y la gestión de control puede en cambio realizarse directamente en sobre el tráfico del cliente. Cuando el método para identificar un sitio web dado a conocer en esta realización de la presente invención se aplica a gestión de facturación, el nombre del sitio web visitado por el cliente se obtiene utilizando un método de comparación de valores de clave de certificados de servidor, de tal modo que la eficiencia de identificación es alta y el funcionamiento es simple y, por lo tanto, se puede realizar la gestión de control oportunamente sobre el tráfico de visita del cliente y se puede mejorar la calidad de servicio de un operador.

La figura 6 es un diagrama esquemático de una estructura física de un dispositivo de inspección profunda de paquetes (inspección profunda de paquetes, DPI), según una realización de la presente invención. Tal como se muestra en la figura 6, el dispositivo DPI 60 incluye:

un procesador (procesador) 610, una interfaz de comunicaciones (interfaz de comunicaciones) 620, una memoria (memoria) 630 y un bus de comunicaciones 640.

El procesador 610, la interfaz de comunicaciones 620 y la memoria 630 comunican entre sí a través del bus de comunicaciones 640.

La interfaz de comunicaciones 620 está configurada para comunicar con un elemento de red, tal como un dispositivo de pasarela, un sistema de facturación en línea (en inglés, sistema de facturación en línea - OCS) o un servidor de gestión.

El procesador 610 está configurado para ejecutar un programa 632, y puede ejecutar específicamente etapas relacionadas, en las realizaciones de método mostradas en la figura 2 y la figura 3.

Específicamente, el programa 632 puede incluir un código de programa, donde el código de programa incluye una instrucción de funcionamiento de ordenador.

El procesador 610 puede ser una unidad central de procesamiento CPU, o un circuito integrado de aplicación específica ASIC (en inglés, Application Specific Integrated Circuit), o uno o varios circuitos integrados configurados para implementar esta realización de la presente invención.

La memoria 630 está configurada para almacenar el programa 632. La memoria 630 puede incluir una memoria RAM de alta velocidad y puede incluir además una memoria no volátil (memoria no volátil), por ejemplo, por lo menos una memoria de disco.

Para la implementación específica de cada módulo funcional en el programa 632, se puede hacer referencia a un módulo correspondiente en la realización descrita en la figura 7, que no se vuelve a describir en este caso.

La figura 7 es un diagrama estructural esquemático de un dispositivo DPI, según una realización de la presente invención. Tal como se muestra en la figura 7, el dispositivo DPI 70 incluye:

un módulo de adquisición 700 está configurado para, cuando un cliente visita un sitio web sobre protocolo seguro de transferencia de hipertexto HTTPS, adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente y un servidor del sitio web.

Haciendo referencia a lo mostrado en la figura 1, cuando el cliente 100 visita un sitio web sobre un protocolo HTTPS, el cliente 100 realiza una negociación de clave con el servidor 120 del sitio web visitado, y el proceso de negociación de clave se puede llevar a cabo según un protocolo de seguridad de capa de transporte (en inglés, seguridad de capa de transporte - TLS), donde el protocolo TLS es un protocolo de cifrado que proporciona un canal confidencial y seguro en internet, un certificado de servidor es un certificado que se utiliza para demostrar un tipo de utilización del servidor y es aplicado por el servidor a un centro de autoridad de certificación (Certificate authority, CA), y el cliente confía en el servidor solamente cuando el certificado es utilizado en el servidor correspondiente. El módulo de adquisición 700 en el dispositivo DPI puede adquirir, mediante interceptación, un mensaje de certificado que se genera en un proceso de la negociación de clave entre el cliente y el servidor de sitio web, donde el mensaje de certificado lleva un certificado de servidor del sitio web visitado por el cliente. Específicamente, el módulo de adquisición 700 puede identificar el mensaje de certificado interceptando un paquete en el proceso de la negociación de clave entre el cliente y el servidor,

y según un campo Tipo de apretón de manos en el paquete.

Un módulo de análisis sintáctico 705 está configurado para obtener, analizando sintácticamente el mensaje de certificado adquirido por el módulo de adquisición 700, el certificado de servidor del sitio web visitado por el cliente.

5 En algunos casos, si el mensaje de certificado adquirido incluye una lista de certificados Lista de certificado, el certificado de servidor del sitio web visitado se puede obtener extrayendo un primer certificado en una parte de lista de certificados Lista de certificado en el mensaje de certificado.

Un módulo de cálculo 710 está configurado para obtener un valor de clave del certificado de servidor, según un algoritmo preestablecido.

10 En esta realización de la presente invención, el contenido del certificado de servidor obtenido por el módulo de análisis sintáctico 705 se puede calcular utilizando el algoritmo preestablecido para obtener el valor de clave del certificado de servidor. Por ejemplo, el contenido del certificado de servidor se puede calcular adoptando un algoritmo de síntesis digital para obtener el valor de clave del certificado de servidor, donde el algoritmo de síntesis digital puede ser un algoritmo de síntesis de mensajes 5 (algoritmo de síntesis de mensajes, MD5), un algoritmo resumen ("hash") seguro SHA-1, un algoritmo de síntesis de mensaje de evaluación de primitivas de integridad RACE (algoritmo de síntesis de mensaje de evaluación de primitivas de integridad RACE, RIPEMD), o similares, lo que no se limita en la presente memoria.

15 Un módulo de búsqueda 715 está configurado para buscar una tabla de valores de clave de sitio web según el valor de clave del certificado de servidor obtenido por el módulo de cálculo 710, donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene según el algoritmo preestablecido.

20 En esta realización de la presente invención, la tabla de valores de clave de sitio web puede estar preestablecida en el dispositivo DPI, donde el valor de clave del certificado de servidor en la tabla de valores de clave de sitio web adopta el mismo algoritmo que el adoptado durante el cálculo que lleva a cabo el módulo de cálculo 710, por ejemplo, adoptando el mismo algoritmo de síntesis digital. Por ejemplo, si el valor de clave del certificado de servidor en la tabla de valores de clave de sitio web preestablecida se obtiene utilizando el algoritmo MD5, cuando un usuario visita un sitio web, el módulo de cálculo 710 tiene asimismo que calcular, utilizando el algoritmo MD5, un valor de clave de un certificado de servidor del sitio web visitado por el usuario. Definitivamente, se puede entender que la tabla de valores de clave de sitio web puede asimismo ser preestablecida en otro dispositivo, siempre que pueda ser accedida por el dispositivo DPI.

25 Un módulo de identificación 720 está configurado para, si el módulo de búsqueda 715 encuentra un nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, identificar, según el nombre de sitio web, el sitio web visitado por el cliente.

30 Según el algoritmo de síntesis digital, tal como MD5, se puede obtener una síntesis digital (un valor resumen, es decir, el valor de clave en esta realización) después de que se realice el cálculo de resumen sobre la información de primitivas. Por lo tanto, se utiliza una característica del algoritmo de síntesis digital en esta realización de la presente invención, donde la característica del algoritmo de síntesis digital indica que diferentes primitivas se pueden convertir en diferentes valores resumen, el contenido de un certificado de servidor de un sitio web que tiene que ser identificado se calcula previamente utilizando un algoritmo de síntesis digital preestablecido para obtener un valor de clave del certificado de servidor, y el valor de clave del certificado de servidor obtenido y el nombre del sitio web se almacenan en consecuencia en la tabla de valores de clave de sitio web. Cuando el dispositivo DPI tiene que identificar el sitio web visitado por el usuario, un valor de clave de un certificado de servidor adquirido a partir de un flujo de datos generado en un proceso de visita del usuario se puede comparar, a través del medio del módulo de búsqueda 715, con un valor preestablecido de clave del certificado de servidor, y si se encuentra en la tabla de valores de clave de sitio web un valor de clave de un certificado de servidor de un sitio web, que es igual que el valor de clave del certificado de servidor del sitio web visitado por el cliente, el módulo de identificación 720 puede obtener, según la tabla de valores de clave de sitio web, el nombre de un sitio web correspondiente al valor de clave del certificado de servidor.

35 Cuando un cliente visita un sitio web sobre un protocolo HTTPS, el dispositivo DPI en esta realización de la presente invención adquiere un certificado de servidor del sitio web visitado por el cliente, obtiene un valor de clave del certificado de servidor según un algoritmo preestablecido, busca en una tabla de valores de clave de sitio web el valor de clave obtenido, e identifica, utilizando un nombre de sitio web que corresponde al valor de clave y se encuentra en la tabla de valores de clave de sitio web, en nombre de un sitio web visitado por el cliente. De este modo, el nombre del sitio web visitado por el cliente se puede identificar incluso cuando el cliente visita el sitio web sobre protocolo HTTPS.

40 En otro caso, el dispositivo DPI 70 dado a conocer en esta realización de la presente invención puede incluir además: un módulo de ajuste 725, configurado para recibir y almacenar una tabla de valores de clave de sitio web enviada por un servidor de gestión, donde la tabla de valores de clave de sitio web incluye un nombre de sitio web de por lo menos

un sitio web y un valor de clave de un certificado de servidor del sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene según el algoritmo preestablecido.

5 Se puede comprender que, cuando un certificado de servidor de un sitio web es actualizado o un sitio web que requiere de atención tiene que ser añadido o eliminado, el módulo de ajuste 725 tiene que actualizar en tiempo una tabla de valores de clave de sitio web preestablecida. Específicamente, cuando se actualiza un certificado de servidor de un sitio web que requiere atención, el módulo de ajuste 725 puede actualizar en tiempo un valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, según el certificado actualizado de servidor del sitio web; cuando se tiene que añadir un sitio web que requiere atención, el módulo de ajuste 725 puede añadir un valor de clave de una certificación de servidor y un nombre de sitio web a la tabla de valores de clave de sitio web según el certificado de servidor del sitio web recién añadido; o cuando se tiene que eliminar un sitio web que requiere atención, el módulo de ajuste 725 puede eliminar, de la tabla de valores de clave de sitio web, un valor de clave de una certificación de servidor correspondiente al nombre del sitio web que tiene que ser eliminado.

En otro caso más, el dispositivo DPI 70 dado a conocer en la anterior realización de la presente invención puede incluir además:

15 un módulo de determinación 730, configurado para, cuando el módulo de búsqueda 715 no encuentra un nombre de sitio web correspondiente al valor de clave del certificado de servidor del sitio web visitado en la tabla de valores de clave de sitio web, determinar un tipo de formato del certificado de servidor obtenido por el módulo de análisis sintáctico 705, y activar el módulo de identificación 720.

20 El módulo de identificación 720 está configurado además para analizar sintácticamente el certificado de servidor del sitio web visitado, según un certificado de un formato diferente determinado por el módulo de determinación 730, para identificar el sitio web visitado por el cliente.

25 Específicamente, si el módulo de determinación 730 determina que el tipo de formato del certificado de servidor del sitio web es un certificado en un formato X.509, el módulo de identificación 720 obtiene un nombre de dominio del sitio web a partir de un atributo de sujeto de certificado e identifica, según el nombre de dominio del sitio web, el sitio web visitado por el cliente; y si el módulo de determinación 730 determina que el tipo de formato del certificado de servidor del sitio web es un certificado en un formato PGP, el módulo de identificación 720 obtiene el nombre de dominio del sitio web a partir del atributo de identificador de usuario del certificado e identifica, según el nombre de dominio del sitio web, el sitio web visitado por el cliente.

30 En otro caso más, el dispositivo DPI 70 dado a conocer en la anterior realización de la presente invención puede incluir además:

un módulo de gestión de tráfico 735, configurado para gestionar, según el sitio web que es visitado por el cliente e identificado por el módulo de identificación 720, un flujo de datos generado en un proceso de visita del cliente.

35 Después de que se identifique el sitio web visitado por el cliente, el módulo de gestión de tráfico 735 puede gestionar el comportamiento de visita en la red del usuario según el sitio web que es visitado por el cliente e identificado por el módulo de identificación 720 y con una política de control preestablecida, por ejemplo, si es necesario llevar a cabo gestión de facturación sobre el tráfico del cliente, o es necesario realizar control o filtrado sobre el tráfico de visita del cliente, se puede determinar según el nombre del sitio web visitado y con una política de paquetes de sitio web preestablecida, lo que no se limita en la presente memoria.

40 Cuando un cliente visita un sitio web sobre un protocolo HTTPS y el sitio web visitado por el cliente no puede ser identificado utilizando una tabla de valores de clave de sitio web preestablecida, el dispositivo DPI en esta realización de la presente invención puede identificar, en una manera de llevar a cabo análisis en línea sobre un certificado de servidor del sitio web visitado por el cliente, el sitio web visitado por el cliente. De este modo, la identificación del sitio web es más completa y la precisión de la identificación del sitio web es mayor. Además, cuando identifica el sitio web visitado por el cliente, el dispositivo DPI dado a conocer en esta realización de la presente invención puede gestionar en tiempo el comportamiento de visita en la red de un usuario, según el nombre del sitio web identificado, proporcionando de ese modo una gestión de red más oportuna, eficiente y elegante para el usuario.

45 En un escenario de aplicación, cuando el dispositivo DPI descrito en esta realización de la presente invención se utiliza para llevar a cabo gestión de facturación de tráfico sobre un usuario, una tabla de valores de clave de sitio web preestablecida puede incluir un nombre de dominio de un sitio web que se puede visitar según lo estipulado en un paquete suscrito por el usuario y un valor de clave de un certificado de servidor del sitio web. Cuando el módulo de búsqueda 715 encuentra en la tabla de valores de clave de sitio web un valor de clave de un certificado de servidor de un sitio web, que es igual que el valor de clave del certificado de servidor del sitio web visitado, esto indica que el sitio web visitado por el cliente está en el paquete, y el módulo de gestión de tráfico 735 puede no reunir estadísticas de tráfico de facturación sobre la visita del cliente y puede permitir directamente el paso del cliente; y cuando el módulo de búsqueda 715 no encuentra en la tabla de valores de clave de sitio web un valor de clave de un certificado de servidor de un sitio web, que es igual que un valor de clave del certificado de servidor del sitio web visitado, esto indica que el sitio web visitado por el cliente no está en el paquete, y el módulo de gestión de tráfico 735 puede en tiempo reunir directamente estadísticas de tráfico sobre el tráfico de visita del cliente y notificar las estadísticas de tráfico

reunidas a un servidor de facturación, para la facturación. En este escenario de aplicación, cuando el módulo de búsqueda 715 no encuentra en la tabla de valores de clave de sitio web un valor de clave de certificado de servidor de un sitio web, que es igual que el valor de clave del certificado de servidor del sitio web visitado, es innecesario activar el módulo de determinación 730 y el módulo de identificación 720 para obtener, analizando el certificado de servidor del sitio web, el nombre de dominio del sitio web visitado por el usuario, y puede en cambio realizarse directamente gestión de control sobre el tráfico del cliente. Cuando el dispositivo DPI dado a conocer en esta realización de la presente invención se aplica a gestión de facturación, el dispositivo DPI puede obtener, comparando valores de clave de certificados de servidor, el nombre del sitio web visitado por el cliente, de tal modo que la eficiencia de la identificación es alta y el funcionamiento es simple y, por lo tanto, se puede llevar a cabo oportunamente una gestión de control sobre el tráfico de visita del cliente, y se puede mejorar la calidad de servicio de un operador.

La figura 8 es un diagrama estructural esquemático de un sistema de red acorde con una realización de la presente invención. Tal como se muestra en la figura 8, el sistema de red incluye: un cliente 100, un dispositivo de inspección profunda de paquetes DPI 110, y por lo menos un servidor de sitio web 120, y el dispositivo DPI está desplegado en una conexión de comunicación entre un cliente 100 y el servidor de sitio web 120 a modo de despliegue en trayecto, donde:

el cliente 100 está configurado para enviar una solicitud de visita en base a un protocolo HTTPS a dicho por lo menos un servidor de sitio web 120, y llevar a cabo negociación de clave con dicho por lo menos un servidor de sitio web 120;

el dispositivo de inspección profunda de paquetes DPI 110 está configurado para, cuando el cliente visita un sitio web sobre protocolo seguro de transferencia de hipertexto HTTPS, adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente y un servidor del sitio web; obtener, analizando sintácticamente el mensaje de certificado, un certificado de servidor del sitio web visitado por el cliente; obtener un valor de clave del certificado de servidor según un algoritmo preestablecido; buscar en una tabla de valores de clave de sitio web según el valor de clave del certificado de servidor, donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene según el algoritmo preestablecido; y si se encuentra un nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, identificar, según el nombre de sitio web, el sitio web visitado por el cliente; y

el servidor de sitio web 120 está configurado para recibir la solicitud de visita de red enviada por el cliente 100, y realizar la negociación de clave con el cliente según la solicitud de visita de red.

En otro caso, el dispositivo DPI 110 está configurado además para determinar un tipo de formato del certificado de servidor si no se encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web; obtener un nombre de dominio del sitio web a partir de un atributo de sujeto del certificado de servidor cuando el tipo de formato del certificado de servidor es un formato X.509; obtener el nombre de dominio del sitio web a partir de un atributo de identificador de usuario del certificado de servidor cuando el tipo de formato del certificado de servidor es un formato PGP; e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente.

En otro caso más, el dispositivo DPI 110 está configurado además para gestionar, según el nombre del sitio web identificado visitado por el cliente, un flujo de datos generado en un proceso de visita del cliente.

Específicamente, para una descripción detallada del dispositivo DPI 110 en esta realización de la presente invención, se puede hacer referencia a la anterior realización relacionada, que no se vuelve a describir en este caso.

Cuando un cliente visita un sitio web sobre un protocolo HTTPS, el sistema de red en esta realización de la presente invención adquiere un certificado de servidor del sitio web visitado por el cliente, obtiene un valor de clave del certificado de servidor según un algoritmo preestablecido, y busca en una tabla de valores de clave de sitio web el valor de clave obtenido, y determina, utilizando un nombre de sitio web que corresponde al valor de clave del certificado de servidor y se encuentra en la tabla de valores de clave de sitio web, el nombre de un sitio web visitado por el cliente. De este modo, el nombre del sitio web visitado por el cliente se puede identificar incluso cuando el cliente visita el sitio web sobre protocolo HTTPS.

Los expertos en la materia pueden comprender que parte o la totalidad de las etapas de las anteriores realizaciones de métodos se pueden implementar mediante un programa que instruya al hardware pertinente. El programa mencionado puede estar almacenado en un medio de almacenamiento legible por ordenador. Cuando el programa se ejecuta, se llevan a cabo las etapas de las anteriores realizaciones de método. El anterior medio de almacenamiento puede incluir cualquier medio que pueda almacenar código de programa, tal como una ROM, una RAM, un disco magnético o un disco óptico.

Los expertos en la materia pueden comprender claramente que, para una descripción cómoda y breve, para un proceso de trabajo detallado del dispositivo y el módulo anteriores, se puede hacer referencia a un correspondiente proceso en las anteriores realizaciones de método, lo que no se vuelve a describir en este caso.

En las diversas realizaciones dadas a conocer en la presente solicitud, se debe entender que el dispositivo y el método

dados a conocer se pueden implementar de otras maneras. Por ejemplo, las realizaciones de aparato dadas a conocer en lo anterior son tan sólo a modo de ejemplo. Por ejemplo, la división de módulos es tan sólo una división de funciones lógicas y puede ser otra división en una implementación real. Por ejemplo, una serie de módulos o componentes se pueden combinar o integrar en otros dispositivos, o algunas características pueden ser ignoradas o no llevadas a cabo. Además, los acoplamientos mutuos o acoplamientos directos o conexiones de comunicación mostradas o explicadas se pueden implementar por medio de algunas interfaces. Los acoplamientos indirectos o conexiones de comunicación entre aparatos o unidades se pueden implementar de forma eléctrica, mecánica u otras.

5

Los módulos descritos como partes independientes pueden o no ser físicamente independientes, y las partes mostradas como módulos pueden o no ser unidades físicas, pueden estar localizadas en una posición o pueden estar distribuidas en una serie de unidades de red. Se puede seleccionar una parte de los módulos o su totalidad en función de las necesidades reales, para conseguir los objetivos de las soluciones de las realizaciones.

10

Además, los módulos de función en las realizaciones de la presente invención pueden estar integrados en un módulo de procesamiento, o cada uno de los módulos puede existir por separado físicamente, o dos o más módulos pueden estar integrados en un módulo.

15

Finalmente, se debe observar que las anteriores realizaciones de método están destinadas tan sólo a describir las soluciones técnicas de la presente invención, y no a limitar la presente invención. Aunque la presente invención se ha descrito en detalle haciendo referencia a las realizaciones anteriores, los expertos en la materia deberán comprender que pueden seguir realizando modificaciones a las soluciones técnicas descritas en lo anterior sin apartarse del alcance de la invención definida por las reivindicaciones adjuntas.

20

REIVINDICACIONES

1. Un método para identificar un sitio web, que comprende:
- 5 cuando un cliente visita un sitio web sobre protocolo seguro de transferencia de hipertexto HTTPS, adquiriendo (200, 305), mediante un dispositivo de inspección profunda de paquetes, DPI, (110) un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente (100) y el servidor (120) del sitio web;
- obtener (205, 305) analizando sintácticamente mediante el dispositivo DPI (110) el mensaje de certificado, un certificado de servidor del sitio web visitado por el cliente (100); e
- identificar (220, 330) el sitio web visitado por el cliente (100),
- caracterizado por
- 10 obtener (210, 315), mediante el dispositivo DPI (110), un valor de clave del certificado de servidor, mediante calcular una síntesis digital del certificado de servidor utilizando una función preestablecida;
- buscar (215, 320), mediante el dispositivo DPI (110), en una tabla de valores de clave de sitio web el valor de clave del certificado de servidor, donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la
- 15 tabla de valores de clave de sitio web se obtiene calculando el contenido del certificado de servidor del sitio web utilizando el algoritmo preestablecido; y
- si se encuentra un nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, identificar (220, 330), mediante el dispositivo DPI (110), según el nombre de sitio web, el sitio web visitado por el cliente (100).
- 20 2. El método para identificar un sitio web según la reivindicación 1, que comprende además:
- determinar (505) un tipo de formato del certificado de servidor si el nombre de sitio web correspondiente al valor de clave del certificado de servidor no se encuentra en la tabla de valores de clave de sitio web;
- obtener (515) un nombre de dominio del sitio web a partir de un atributo de sujeto del certificado de servidor, si el tipo de formato del certificado de servidor es un formato X.509, e identificar, según el nombre de dominio del sitio web, el
- 25 sitio web visitado por el cliente (100); y
- obtener (525) el nombre de dominio del sitio web a partir de un atributo de identificador de usuario del certificado de servidor si el tipo de formato del certificado de servidor es un formato PGP, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente (100).
3. El método para identificar un sitio web según la reivindicación 1, que comprende además:
- 30 no reunir estadísticas de facturación sobre tráfico del cliente si el nombre de sitio web correspondiente al valor de clave del certificado de servidor se encuentra en la tabla de valores de clave de sitio web; y
- reunir estadísticas de facturación sobre el tráfico del cliente si el nombre de sitio web correspondiente al valor de clave del certificado de servidor no se encuentra en la tabla de valores de clave de sitio web
4. El método para identificar un sitio web según cualquiera de las reivindicaciones 1 a 3, en el que la adquisición (200, 305) de un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente (100) y un
- 35 servidor (120) del sitio web comprende:
- interceptar un paquete que se genera en un proceso de la negociación de clave entre el cliente y el servidor del sitio web; e
- identificar el mensaje de certificado según un campo Tipo de apretón de manos en el paquete.
- 40 5. Un dispositivo de inspección profunda de paquetes (70), que comprende:
- un módulo de adquisición (700), configurado para, cuando un cliente (100) visita un sitio web sobre protocolo seguro de transferencia de hipertexto, HTTPS, adquirir un mensaje de certificado que se genera en un proceso de negociación de clave entre el cliente (100) y un servidor (120) del sitio web;
- 45 un módulo de análisis sintáctico (705), configurado para obtener, analizando sintácticamente el mensaje de certificado adquirido por el módulo de adquisición (700), un certificado de servidor del sitio web visitado por el cliente (100);
- un módulo de cálculo (710), configurado para obtener un valor de clave del certificado de servidor, mediante calcular una síntesis digital del certificado de servidor utilizando una función preestablecida;

- 5 un módulo de búsqueda (715), configurado para buscar en una tabla de valores de clave de sitio web el valor de clave del certificado de servidor obtenido mediante el módulo de cálculo (710), donde la tabla de valores de clave de sitio web registra un nombre de sitio web y un valor de clave de un certificado de servidor de un sitio web, y el valor de clave del certificado de servidor del sitio web en la tabla de valores de clave de sitio web se obtiene calculando el contenido del certificado de servidor del sitio web utilizando el algoritmo preestablecido; y
- un módulo de identificación (720), configurado para, si el módulo de búsqueda (715) encuentra un nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, identificar, según el nombre de sitio web, el sitio web visitado por el cliente (100).
6. El dispositivo de inspección profunda de paquetes (70) según la reivindicación 5, que comprende además:
- 10 un módulo de ajuste (725), configurado para recibir y almacenar la tabla de valores de clave de sitio web enviada por un servidor de gestión, donde la tabla de valores de clave de sitio web comprende un nombre de sitio web de por lo menos un sitio web y un valor de clave de un certificado de servidor del sitio web.
7. El dispositivo de inspección profunda de paquetes (70) según la reivindicación 5 o 6, que comprende además:
- 15 un módulo de determinación (730) configurado para, cuando el módulo de búsqueda (715) no encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web, determinar un tipo de formato del certificado de servidor obtenido mediante el módulo de análisis sintáctico (705), y activar el módulo de identificación (720), en el que:
- 20 el módulo de identificación (720) está configurado además para obtener un nombre de dominio del sitio web a partir de un atributo de sujeto del certificado de servidor cuando el módulo de determinación determina que el tipo de formato del certificado de servidor del sitio web es un formato X.509, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente (100); y obtener el nombre de dominio del sitio web a partir de un atributo de identificador de usuario del certificado de servidor cuando el módulo de determinación (730) determina que el tipo de formato del certificado de servidor del sitio web es un formato PGP, e identificar, según el nombre de dominio del sitio web, el sitio web visitado por el cliente (100).
- 25 8. El dispositivo de inspección profunda de paquetes según la reivindicación 5 o 6, que comprende además:
- un módulo de gestión de tráfico (735), configurado para no reunir estadísticas de facturación sobre el tráfico del cliente (100) cuando el módulo de búsqueda (715) encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web; y reunir estadísticas de facturación sobre el tráfico del cliente (100) cuando el módulo de búsqueda (715) no encuentra el nombre de sitio web correspondiente al valor de clave del certificado de servidor en la tabla de valores de clave de sitio web.
- 30 9. El dispositivo de inspección profunda de paquetes (70) según cualquiera de las reivindicaciones 5 a 8, en el que:
- el módulo de adquisición (700) está configurado específicamente para interceptar un paquete que se genera en un proceso de la negociación de clave entre el cliente (100) y el servidor (120) del sitio web, e identificar el mensaje de certificado según un campo Tipo de apretón de manos en el paquete.
- 35 10. El dispositivo de inspección profunda de paquetes (70) según cualquiera de las reivindicaciones 5 a 9, en el que:
- el algoritmo preestablecido comprende un algoritmo MD5 o un algoritmo SHA-1.
11. Un dispositivo de inspección profunda de paquetes (60) según cualquiera de las reivindicaciones 5 a 10, en el que:
- 40 el dispositivo de inspección profunda de paquetes comprende un procesador (610), una interfaz de comunicaciones (620), una memoria (630) y un bus de comunicaciones (640);
- el procesador (610) y la interfaz de comunicaciones (620) llevan a cabo comunicación a través del bus de comunicaciones (640);
- la interfaz de comunicaciones (620) está configurada para llevar a cabo una comunicación.
- la memoria (630) está configurada para almacenar un programa; y
- 45 el procesador (610) está configurado para ejecutar el programa para implementar los módulos definidos en cualquiera de las reivindicaciones 5 a 10.

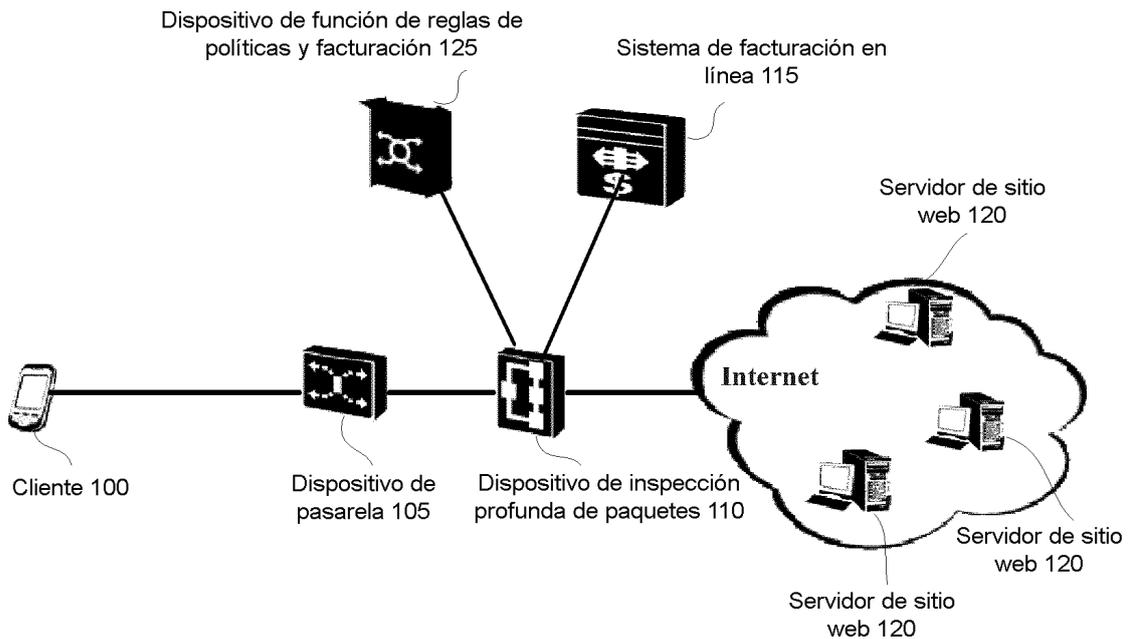


FIG. 1

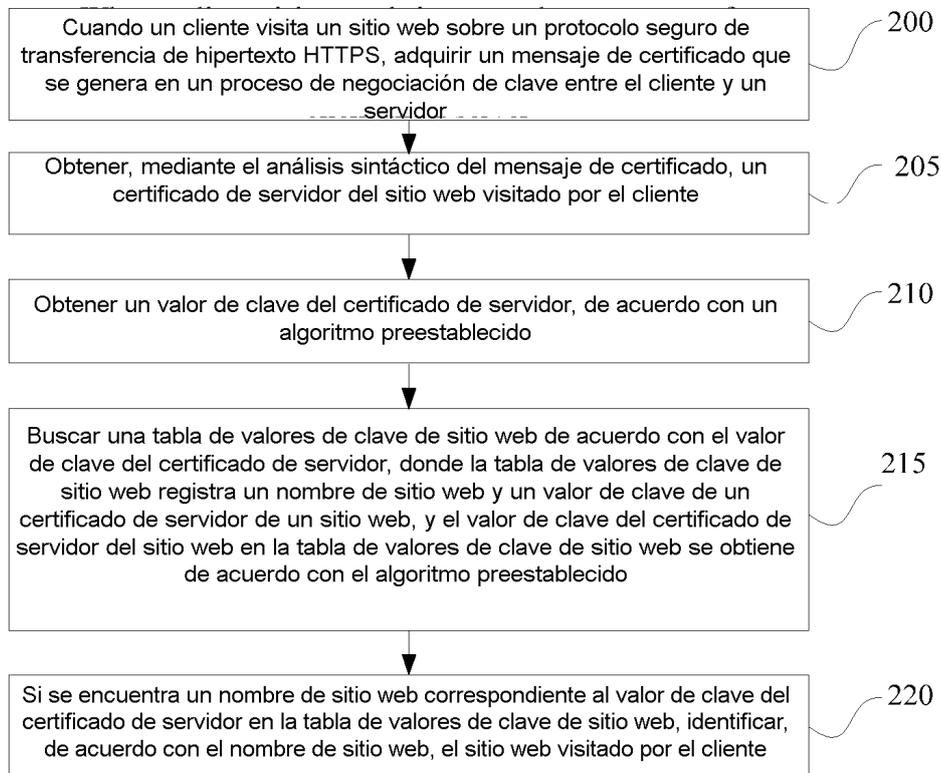


FIG. 2

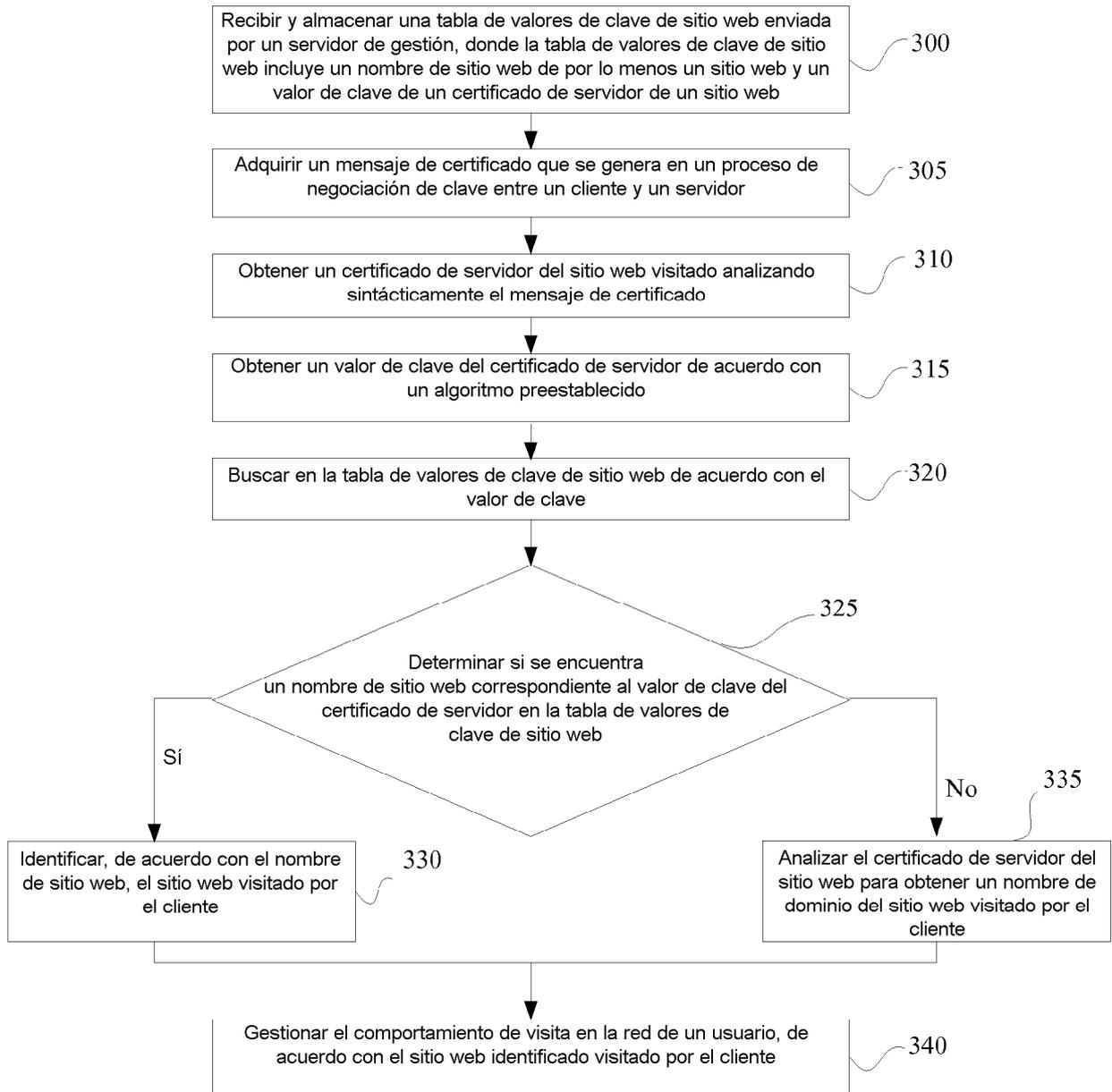


FIG. 3

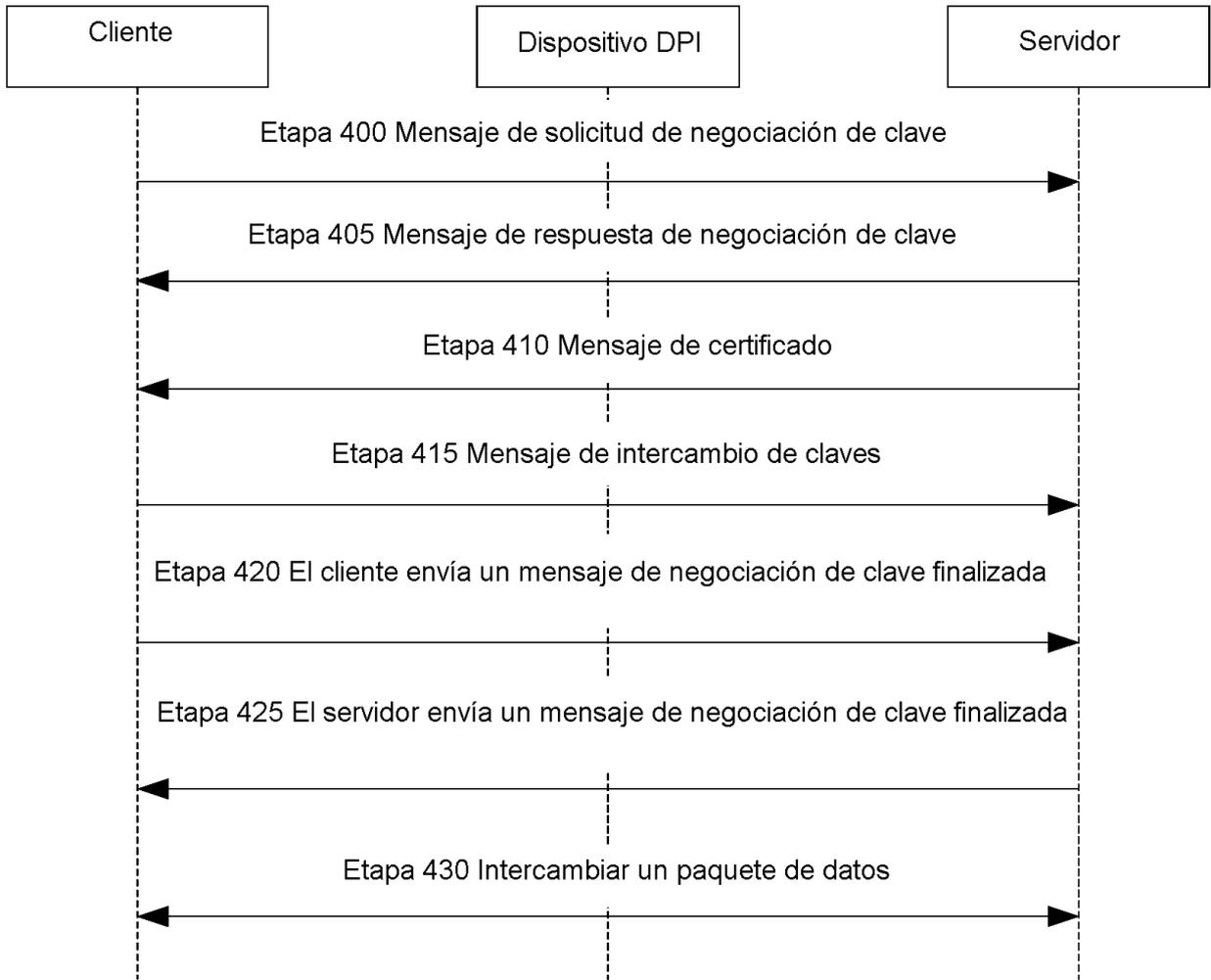


FIG. 4

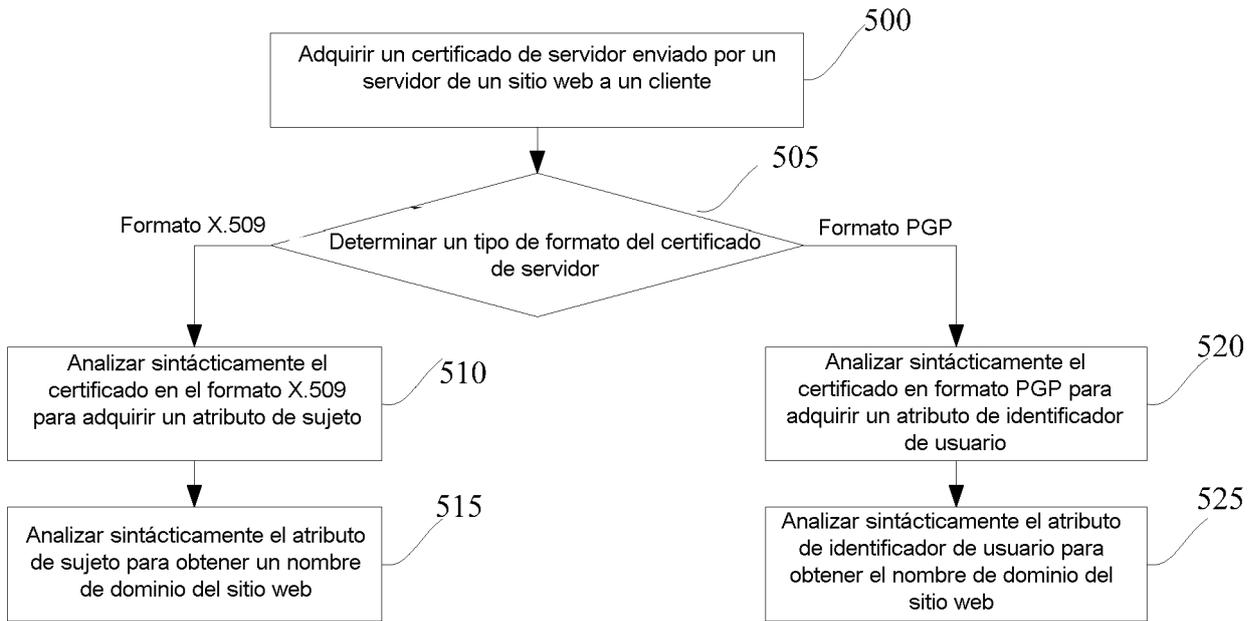


FIG. 5

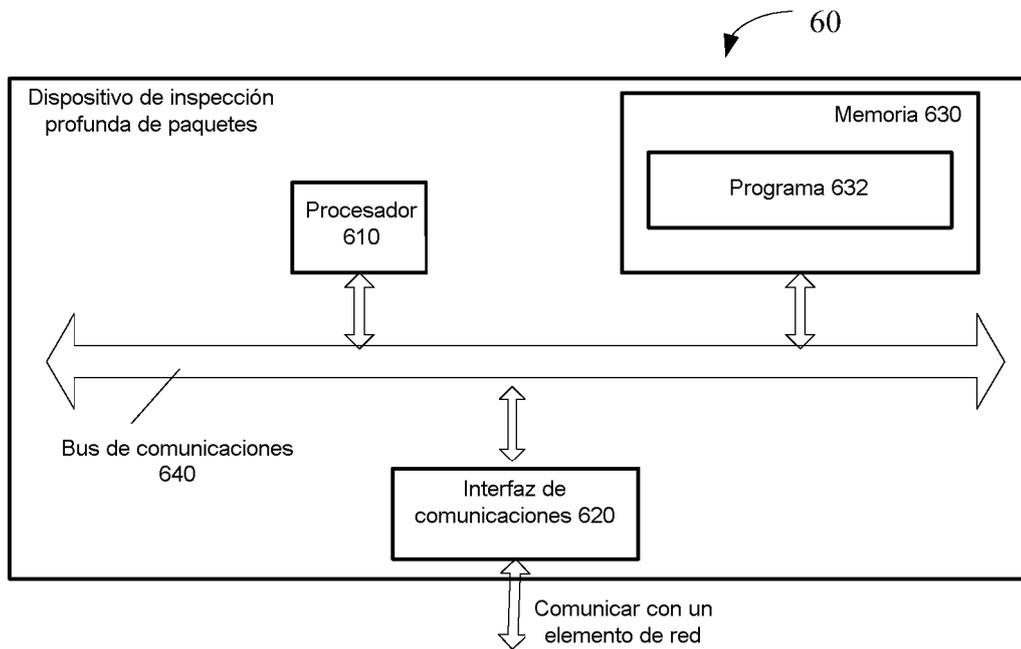


FIG. 6

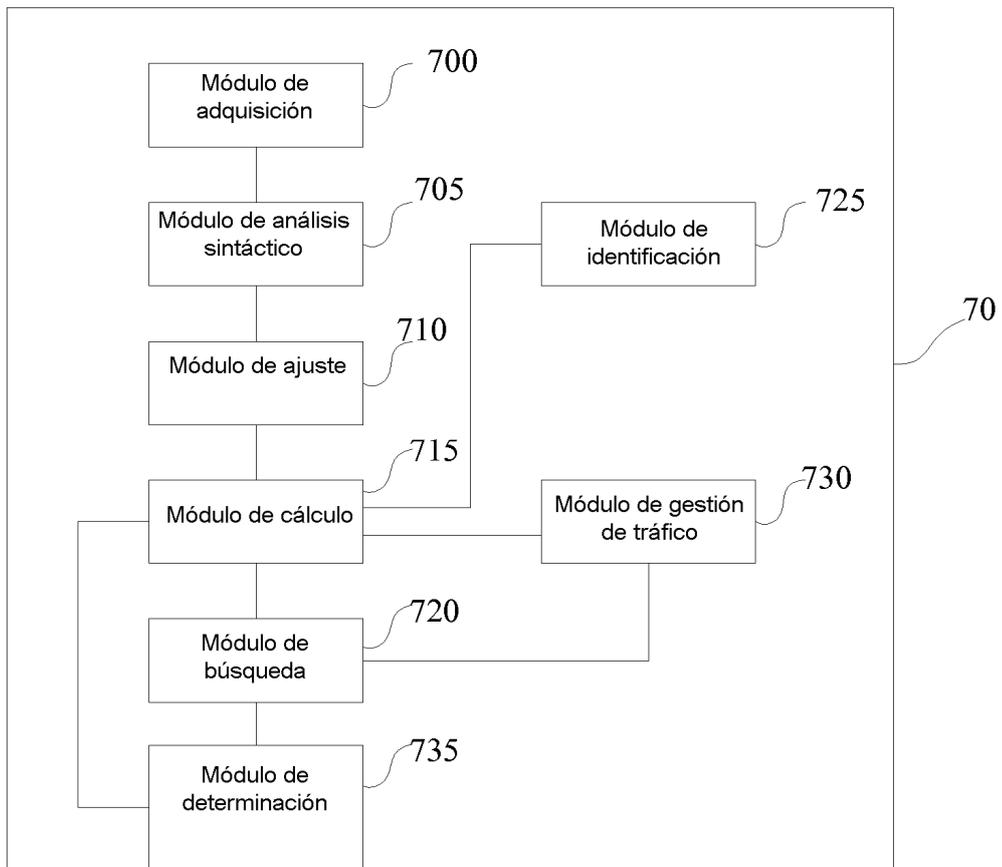


FIG. 7

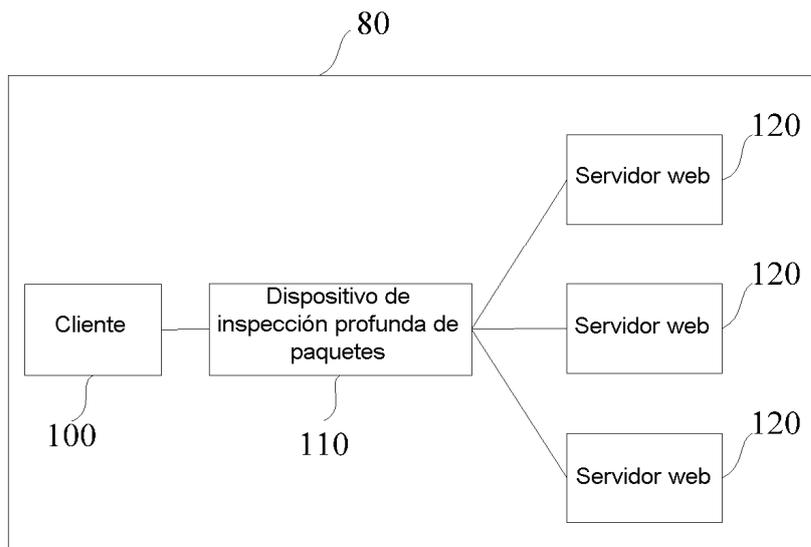


FIG. 8