(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 568 602**

(51) Int. CI.:

H04L 12/26 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **23.11.2011 E 11796640 (8)**

(97) Fecha y número de publicación de la concesión europea: **03.02.2016 EP 2767037**

(54) Título: **Un método para minimizar el posprocesamiento del tráfico de red**

(30) Prioridad:

28.09.2011 US 201161540228 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
03.05.2016

(73) Titular/es:

TELEFÓNICA S.A. (100.0%)
C/ Gran Vía 28
28013 Madrid, ES

(72) Inventor/es:

MAESO MARTÍN-CARNERERO, ADRIÁN;
GARCÍA DE BLAS, GERARDO;
RAMÓN SALGUERO, FRANCISCO JAVIER y
MONTES MORENO, PABLO

(74) Agente/Representante:

ARIZTI ACHA, Monica

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

Un método para minimizar el posprocesamiento del tráfico de red.

Campo de la técnica

La presente invención se refiere en general, a un método para minimizar el posprocesamiento del tráfico de red, supervisado dicho tráfico de red por medio de metadatos descriptivos, producidos dichos metadatos descriptivos mediante un despliegue de una Interfaz de Metadatos Descriptivos de la Inspección a Fondo de Paquetes de una red, conteniendo dicha descripción campos de paquetes textuales e información contable, y más particularmente a un método que comprende correlacionar al menos parte de dichos metadatos descriptivos con información incluida en dichos metadatos descriptivos, firmas centralizadas y fuentes externas de datos para enriquecer dichos metadatos descriptivos.

Estado previo de la técnica

La supervisión de la red se ha convertido en una tarea importante en las modernas redes. Permite el mantenimiento de la estabilidad, disponibilidad y seguridad del sistema de la red y permite tomar buenas decisiones respecto a la capacidad y planificación de la red.

Estudiando el comportamiento del tráfico en diferentes momentos es posible deducir patrones en el crecimiento del tráfico que permitan la creación de modelos predictivos. Para ser precisos, estos modelos deben basarse no solamente en la cantidad de tráfico transferido, sino que deben considerar los diferentes protocolos y tipos de tráfico presentes en la red y cómo pueden ser afectados por cambios en la red o en los proveedores de servicio. Por ejemplo, si un proveedor de contenido de video incrementa la tasa de bits de sus videos, la misma cantidad de solicitudes de video producirán una cantidad de tráfico mayor.

Algunos productos comerciales tales como Sandvine, iPoque o Cisco SCE proporcionan una solución basada en un análisis de DPI y la detección de patrones de paquetes. Estos sistemas inspeccionan los paquetes que atraviesan un enlace y clasifican cada paquete como perteneciente a una clase específica de aplicación o se clasifican como desconocidos. Esta información se usa para proporcionar informes de tráfico que son el producto final del sistema. Es importante notar que cualquier tráfico que no se clasifique correctamente permanecerá en esa clasificación dado que los informes de tráfico no proporcionan suficiente información para aplicarles otros análisis. Una alternativa a estos sistemas de supervisión es el método descrito en la solicitud de patente WO 2011/051750 de supervisión del tráfico de red por medio de metadatos descriptivos. Este método es capaz de proporcionar una captura de tráfico reducida que pueda posprocesarse en una etapa posterior, desconectando de esta forma la captura del tráfico del análisis e incrementando grandemente la flexibilidad al tiempo que se minimiza el número de actualizaciones en el sistema de captura.

Una solución para la gestión de la red o el tráfico se describe en el documento EP 2262173 que describe un método para gestión de una red para controlar o influir en los flujos de tráfico usando un conocimiento a priori de las fuentes de tráfico y de las aplicaciones o servicios proporcionados a través de la red.

La mayor parte de las soluciones de supervisión del tráfico realizan un análisis del tráfico usando un enfoque de sistema monolítico mediante la comparación de los paquetes únicos o las transmisiones continuas de tráfico con patrones de tráfico almacenados y combinando la información obtenida con fuentes externas de datos. Estos dos tipos de información se procesan en el mismo sistema que capturó el tráfico produciendo una interpretación de lo que se observa en la red, tal como se mostrará en la Figura 1.

El método de supervisión del tráfico de red por medio de metadatos descriptivos introdujo una alternativa al procedimiento de DPI general, dividiendo el sistema de DPI en dos: detección del tráfico y posprocesamiento.

El componente de detección de tráfico en este modelo alternativo al de DPI consiste en la detección de paquetes relevantes y la extracción a partir de ellos de campos clave. Por ejemplo, un paquete relevante podría ser una solicitud HTTP y uno de sus campos clave el nombre del alojamiento. El resultado de la detección del tráfico es una transmisión continua de campos de paquetes textuales, que a partir de ahora se denominarán como metadatos. Añadiendo estos datos a una contabilidad de flujo agregado se forma la Interfaz de Metadatos Descriptivos, tal como se muestra en la Figura 2.

La Interfaz de Metadatos Descriptivos proporciona una descripción de todo el tráfico observado en la red. Esta descripción del tráfico, suficientemente general para permitir la detección de firmas en él, puede posprocesarse fuera de la caja de la DPI para generar informes de tráfico. De esta forma el resultado de la Interfaz de Metadatos Descriptivos, debido a su tamaño reducido, puede almacenarse y procesarse fuera de línea.

El procesamiento fuera de línea implica una gran ganancia en términos de análisis del tráfico. Dado que la interfaz de metadatos descriptivos proporciona un resumen del tráfico incluyendo campos claves de paquetes (metadatos), es posible usar firmas para detectar nuevos tipos de tráfico. En esta forma, el resultado de la Interfaz de Metadatos Descriptivos puede usarse varios meses después con nuevo análisis, por ejemplo para comprobar si un tipo nuevamente popular de tráfico estaba presente en el momento de la captura.

El posprocesamiento de la captura usa dos fuentes de información para procesar las capturas: las firmas instaladas y las fuentes externas de datos, por ejemplo datos RADIUS.

Las firmas para posprocesamiento no son estáticas, en ocasiones necesitan actualizarse. Esto es necesario cuando cambia un protocolo o si desea incluirse la detección de un nuevo tipo de tráfico.

Las fuentes externas de datos se modifican frecuentemente, por ejemplo, pueden actualizarse archivos que ajustan los intervalos IP a su localización geográfica, por ejemplo mejorando la resolución de países a ciudades.

Dado que los cambios en las firmas y fuentes externas pueden conducir a un mejor posprocesamiento es interesante procesar la captura de nuevo cuando esto sucede, teniendo en esta forma la capacidad de proporcionar informes de tráfico más completos y precisos.

Los sistemas de DPI tradicionales tienen varias desventajas:

- No son modulares dado que realizar la tarea de clasificación del tráfico y contabilidad del tráfico en un único equipo.

La información acerca de la clasificación del tráfico no puede exportarse para análisis posterior. Se están exportando formatos para contabilidad del tráfico (por ejemplo, Netflow realiza contabilidad de bytes por flujo), pero no hay forma de exportar las decisiones acerca de la clasificación del tráfico. Una vez se clasifica un paquete, el paquete se borra y no se exporta ninguna información acerca de esta clasificación. Esto tiene varios inconvenientes:

- No es posible volver a clasificar los paquetes de nuevo. Si algunos paquetes se clasifican como desconocidos, estos paquetes no se pueden volver a clasificar en otra categoría, aunque mejoren los métodos para identificar el tráfico.

Junto a ello, el equipo necesita actualizarse para mantener las firmas actualizadas, lo que permite la clasificación del tráfico en la categoría correcta. Dado que la información acerca de la clasificación del tráfico no se exporta y no es posible la reclasificación, esto fuerza al equipo a actualizarse frecuentemente.

La supervisión del tráfico de red por medio de metadatos descriptivos resuelve los inconvenientes mencionados, pero no acomete cómo analizar eficientemente el resultado de este método de supervisión.

El inconveniente principal de los sistemas de DPI tradicionales es su flexibilidad limitada para realizar nuevos tipos de análisis de tráfico. Esto se debe principalmente al hecho de que estos dispositivos trabajan como un sistema monolítico, generando directamente como resultado la información que se incluiría en un informe de tráfico, y por lo tanto si se requiere un nuevo tipo de análisis todo el sistema debe modificarse.

El método de supervisión del tráfico de red por medio de metadatos descriptivos permite la separación de la captura del tráfico del procesamiento del tráfico, incrementando en esta forma la flexibilidad del sistema. Básicamente este método permite guardar una captura del tráfico de pequeño tamaño, incluyendo piezas claves de información, que se posprocesan por separado. Esta separación entre la captura y el análisis incrementa significativamente la flexibilidad del sistema, dado que los cambios se aplicarían a la etapa de posprocesamiento y no a su adquisición.

El posprocesamiento incluye todos los tipos de operaciones a ser realizados en la captura para obtener los datos requeridos para un análisis del tráfico. Esto puede incluir la correlación con fuentes externas de datos, correlación con firmas de protocolo y el uso de heurística de tráfico entre otros métodos. Este procesamiento a ser aplicado a la captura es muy costoso en términos computacionales de modo que debería optimizarse, pero el posprocesamiento incluye también la aplicación de un procesamiento más simple que puede realizarse solo después de que se hayan realizado todas las correlaciones. Por ejemplo, la obtención de la cantidad total de bytes descargados desde servidores de YouTube en el Reino Unido con una tasa de bits específica requeriría la detección de la tasa de bits de los videos, correlación de las solicitudes de video con la cantidad total de bytes descargados, correlación con la localización geográfica y finalmente sumar los conceptos de los registros que coinciden con las restricciones de tráfico impuestas. En este ejemplo, todo el proceso pesado son todas las correlaciones, pero el análisis es simplemente la suma de bytes.

El objetivo final del posprocesamiento es ser capaz de generar un informe de tráfico a partir del que se puedan

deducir conclusiones acerca del tráfico. Estas conclusiones pueden ser acerca del tráfico en general o acerca de un protocolo o aplicación específica, y por lo tanto el posprocesamiento puede variar dependiendo del tipo de análisis de tráfico a realizar.

Descripción de la invención

Es necesario ofrecer una alternativa al estado de la técnica que cubra los huecos encontrados en ella, particularmente relativos a la carencia de propuestas que permitan realmente la definición de cómo analizar el resultado de una Interfaz de Metadatos Descriptivos permitiendo el uso de herramientas de análisis simples para crear informes de tráfico.

Con este fin, la presente invención propone un método para minimizar el posprocesamiento del tráfico de red, que comprende la correlación y procesamiento de al menos parte de un resultado compuesto de metadatos y datos de contabilidad del tráfico con información incluida en los metadatos, datos de contabilidad del tráfico, firmas de protocolo almacenadas centralmente y fuentes de datos externas de datos de localización geográfica relacionados con el tráfico de red, obtenidos dichos metadatos y dichos datos de contabilidad del tráfico a partir de una Interfaz de Metadatos Descriptivos de un despliegue de Inspección a Fondo de Paquetes (DPI, del inglés “Deep Packet Inspection”) de una red.

Al contrario de las propuestas conocidas, el método de la invención, en una forma característica, incluye un proceso de enriquecimiento que comprende correlación y nuevo procesamiento de dichos resultados correlacionados y procesados previamente compuestos de metadatos y datos de contabilidad del tráfico.

La invención, en un segundo aspecto, proporciona un sistema para minimizar el posprocesamiento del tráfico de red, que comprende

- medios para correlacionar y procesar al menos parte de un resultado compuesto de metadatos y datos de contabilidad del tráfico con información incluida en dichos metadatos, dichos datos de contabilidad del tráfico, firmas de protocolo almacenadas centralmente y fuentes de datos externas de datos de localización geográfica relacionados con el tráfico de la red;
- una Interfaz de Metadatos Descriptivos para proporcionar un resumen del tráfico de una red que incluye dichos metadatos, y
- un almacenamiento para dichas firmas de protocolo almacenadas centralmente.

En una forma característica el sistema de la invención comprende adicionalmente medios de correlación y procesamiento adaptados para realizar un enriquecimiento de dicha salida previamente correlacionada y procesada de metadatos y datos de contabilidad del tráfico.

Otras realizaciones del método de la invención se describen de acuerdo con las reivindicaciones adjuntas, y en una sección posterior relacionada con la descripción detallada de varias realizaciones.

Breve descripción de los dibujos

Las previas y otras ventajas y características se comprenderán más completamente a partir de la descripción detallada a continuación de las realizaciones, con referencia a los dibujos adjuntos (algunos de los cuales ya se han descrito en la sección de Estado previo de la técnica), que deben considerarse en una forma ilustrativa y no limitativa, en los que:

La Figura 1 muestra sistemas de Inspección a Fondo de Paquetes genéricos actuales.
La Figura 2 muestra sistemas de Inspección a Fondo de Paquetes actuales en base a la supervisión del tráfico de la red por medio de metadatos descriptivos.
La Figura 3 muestra la concatenación del Sistema de Enriquecimiento de Metadatos de la DPI con un módulo de generación de informes que produce informes de tráfico, de acuerdo con una realización de la presente invención.
La Figura 4 muestra los diferentes procesos a realizarse sobre los metadatos descriptivos para enriquecerlos, de acuerdo con una realización de la presente invención.
La Figura 5 ilustra el hecho de que el Sistema de Enriquecimiento de Metadatos de la DPI mantiene el formato de los datos como su salida, de acuerdo con una realización de la presente invención.

Descripción detallada de varias realizaciones

El Sistema de Enriquecimiento de Metadatos de la DPI (DMES, del inglés “DPI Metadata Enrichment System”) propuesto en la presente invención se ha creado como una solución para optimizar el posprocesamiento para el método de supervisión del tráfico de red por medio de metadatos descriptivos. Este sistema realiza acciones de

posprocesamiento pesadas de una manera que permite la reducción del tiempo de procesamiento y el incremento de la flexibilidad.

El Sistema de Enriquecimiento de Metadatos de la DPI (DMES) complementa la técnica de supervisión del tráfico de red por medio de metadatos descriptivos definiendo cómo analizar el resultado de la interfaz de metadatos descriptivos y permitiendo el uso de herramientas de análisis simples para crear informes de tráfico en base a los resultados del DMES.

Básicamente, el DMES procesa el resultado de la Interfaz de Metadatos Descriptivos; esta es la interfaz que ofrece la captura de un sistema de supervisión del tráfico de red por medio de metadatos descriptivos. La captura se correlaciona con firmas, la propia información en la captura y los datos de fuentes externas, produciendo un resultado enriquecido que incluye toda la información de correlación y que se usará en una etapa posterior para análisis del tráfico, tal como se muestra en la Figura 3.

La presente invención consiste en un sistema capaz de minimizar los esfuerzos necesarios para procesar el resultado del sistema siguiendo el método descrito en el documento WO 2011/051750 de supervisión del tráfico de red por medio de metadatos descriptivos.

La característica clave del Sistema de Enriquecimiento de Metadatos de la DPI es que los datos de salida tienen el mismo formato que los datos de entrada. En esta forma es posible usar como entrada del DMES sus propios datos de salida.

El DMES es alimentado con datos tales como cómo interpretar metadatos, localizaciones geográficas, alojamientos de interés, intervalos IP de interés, etc. Dado que estos datos se actualizan frecuentemente, sería deseable tener la capacidad de actualizar también el resultado del sistema de enriquecimiento. Este enriquecimiento de unos datos previamente enriquecidos se realiza en el DMES simplemente por reprocesamiento.

El Sistema de Enriquecimiento de Metadatos de la DPI es capaz de enriquecimiento de datos de modo selectivo. Esto implica que es posible, por ejemplo, añadir simplemente la localización geográfica a las trazas o enriquecer simplemente ciertas aplicaciones. Esta capacidad es muy útil cuando es necesario el reprocesamiento, dado que es posible enriquecer solamente los datos afectados por actualizaciones en el DMES, ahorrando en esta forma tiempo de procesamiento.

Algunas características de la presente invención son:

- La salida del DMES sigue el mismo formato que los datos proporcionados por la interfaz de metadatos descriptivos.
- El uso del DMES permite minimizar la complejidad de las etapas de procesamiento posterior.
- Es posible usar el resultado del DMES como entrada cuando es necesario reprocesamiento.
- El DMES enriquece capturas usando información incluida en la captura, firmas centralizadas y fuentes externas de datos.
- El DMES permite especificar qué tipos de enriquecimiento deben aplicarse a las capturas, siendo posible por ejemplo aplicar solo una detección de firmas específica.
- Las firmas y las fuentes externas de datos para correlación cambian/se mejoran frecuentemente y cuando esto ocurre es conveniente reprocesar las capturas.
- Cuando se realiza el reprocesamiento, permitir solo el enriquecimiento afectado por cambios en el DMES implica que el tiempo de procesamiento se reduce drásticamente.

La Figura 4 muestra un ejemplo de una implementación posible de la invención. Tal como se observa en la figura, la información de la interfaz de metadatos va a través del sistema usando diferentes fuentes para enriquecimiento de los datos:

Caja 1 - Actualización de metadatos. Los metadatos se actualizan usando la información de firmas, por ejemplo, un mensaje de metadatos que contiene información de una transacción HTTP puede actualizarse para indicar que la transacción HTTP fue una descarga desde un servicio de alojamiento de archivos.
Caja 2 - Correlación de contabilidad con metadatos. La información de contabilidad se enriquece usando la información presente en los mensajes de metadatos. Por ejemplo, el uso de un mensaje de metadatos que informa que un flujo procede desde un servicio de alojamiento de archivos. Esto permite la inclusión de esa información en la contabilidad de ese flujo, determinando el número de bytes subidos/descargados para realizar la descarga del archivo.
Caja 3 - Correlación con fuentes de datos externas. La correlación de la información de contabilidad con las fuentes adicionales de datos. Por ejemplo, si los datos externos usados para correlacionar es un diccionario que permite asignar las IP a la localización geográfica esta caja permitiría determinar dónde está físicamente situado el servidor de una compañía de alojamiento de archivos desde donde se ha descargado un contenido.

Caja 4 - Detección de firmas. Una vez se ha enriquecido la captura en las cajas previas es posible realizar detección adicional de firmas. Por ejemplo un uso heurístico para determinar el tipo de tráfico de flujos desconocidos.

La posible implementación representada en la Figura 4, es solo un esquema funcional. Las funcionalidades de los diferentes módulos podrían agruparse en un único equipo o separarse en equipos diferentes.

La capacidad del DMES de generar una salida enriquecida, manteniendo el mismo formato que su entrada, se basa en la definición del formato de la Interfaz de Metadatos Descriptivos. Este formato incluye campos en la información de contabilidad dirigidos a almacenar información adicional del flujo, tal como el tipo de tráfico o la localización geográfica de servidor, y estos son los campos que el DMES rellena/actualiza para la correlación de la descripción del tráfico con diferentes fuentes de datos (definiciones de firmas, metadatos actualizados y fuentes externas de datos).

Las actualizaciones de las fuentes de información usadas por el DMES implica un mejor enriquecimiento de las capturas y por lo tanto es conveniente actualizar las capturas reprocesándolas con el DMES. Hay dos razones para reprocesar una captura ya procesada en lugar de usar directamente la salida de la Interfaz de Metadatos Descriptivos:

1. Reducción del almacenamiento. Dado que la salida del DMES puede usarse como entrada del sistema no es necesario almacenar la captura original (resultado de la interfaz de metadatos descriptivos).
2. Reducción del tiempo requerido para generar la nueva salida. Dado que el DMES permite enriquecer selectivamente datos mediante la desactivación de la correlación con fuentes de datos específicas, solo es necesario activar el enriquecimiento que afecta a los datos modificados, y por lo tanto reducir el tiempo necesario para el reprocesamiento. Por ejemplo, si se mejora una firma que permite reclasificar videos FLV en transmisión continua para indicar el proveedor de contenido, el enriquecimiento de datos debe aplicarse solamente a los flujos que se detectaron en iteraciones previas como videos FLV en transmisión continua.

La Figura 5 representa gráficamente la posibilidad de usar el DMES para analizar directamente el resultado de la Interfaz de Metadatos Descriptivos respecto a la posibilidad de analizar su propio resultado. El uso normal del Sistema de Requerimiento de Metadatos de la DPI seguiría estas etapas:

1. Procesar la captura de la Interfaz de Metadatos Descriptivos.
2. Eliminar la captura de la Interfaz de Metadatos Descriptivos.
3. Usar el resultado del DMES para realizar un análisis dirigido a generar informes de tráfico y mantener el resultado del DMES para reprocesamiento si es necesario.

Como puede observarse estas etapas no incluyen el reprocesamiento en el DMES. El reprocesamiento solo se realiza cuando es necesario para introducir cambios en los datos que usa para enriquecer capturas. Esto es muy útil para determinar rápidamente la presencia de nuevos protocolos en una captura, dado que los únicos protocolos que es interesante detectar son los más significativos en volumen y aquellos que son interesantes desde una perspectiva táctica.

Para ilustrar el Sistema de Enriquecimiento de Metadatos de la DPI, se obtuvieron algunos resultados mediante una implementación particular de la invención.

En esta implementación toda la información gestionada son datos binarios. Esto se ha realizado para optimizar el rendimiento y el espacio en disco necesario para guardar resultados. En cualquier caso, la representación de datos binarios no permitiría ilustrar el DMES de modo que se usarán datos de texto en su lugar.

Las siguientes tablas representan el resultado de la interfaz de metadatos descriptivos:

1396673130:49569	3269476872:80	TCP	4	1	5360	40	VLAN_Q 50	00	00
1394646482:50108	3174935809:1536	TCP	2	0	2680	0	VLAN_Q 50	00	00
1394625343:24735	1396297335:48384	UDP	0	1	0	1466	VLAN_Q 50	00	00
1343932984:55259	1396055224:21784	TCP	5	4	5748	160	VLAN_Q 50	00	00
1436034701:24076	1361312813:3565	TCP	0	1	0	1188	VLAN_Q 50	00	00
1395069195:12259	3181184896:12408	UDP	1	0	63	0	VLAN_Q 50	00	00
1394646123:3322	3174935809:1536	TCP	3	0	156	0	VLAN_Q 50	00	00
1343932535:16018	1592110395:80	UDP	1	0	129	0	VLAN_Q 50	00	00
1395791963:23415	1114410499:51413	UDP	0	1	0	165	VLAN_Q 50	00	00

1395069348:54768	3654843008:18669	TCP	1	2	1440	109	VLAN_Q 50	00	00
1334864840:56106	1334904428:22938	UDP	0	1	0	1430	VLAN_Q 50	00	00
1396672799:12612	1440435422:3243	TCP	3	1	4172	40	VLAN_Q 50	00	00

USR 1301459263.895152 1396673130:49569 > 3269476872:80 TCP 529
489 18 VLAN_Q 50 HTTP GET 07 GET /hprofile-ak-snc4/
187272_1369476520_2672812_q.jpg HTTP/1.1 Mozilla/5.0 (Windows; U; Windows NT 6.0; en-US) AppleWebKit/534.16 (KHTML, like Gecko) Chrome/10.0.648.204
Safari/534.16 profile.ak.fbcdn.net
RED 1301459263.895272 1599070452:31446 > 1395791963:23415 UDP
661 633 VLAN_Q 50 GET_PEERS_RESPONSE n_peers: 22
1406509822:42600 1595098503:15820 306631082:27234
1394625343:24735 3563778190:15738 773555601:21607
1476399821:14528 1414351873:17078 3654843008:18669
1429675938:22124 1417290889:24633 1435079862:23633
1396344029:24938 1417452430:17941 1477912353:25503
1436034701:24076 411762532:16881 3192951174:58539
USR 1301459263.895332 1343932535:16018 > 1592110395:80 TCP 1030
990 18 VLAN_Q 50 GET_HTTP 07 GET
/opt/icons/icon_error_with_bg.gif HTTP/1.1 Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US) AppieWebKit/534.16 (KHTML, like Gecko) Chrome/10.0.648.151
Safari/534.16 es.madbid.com
RED 1301459263.895371 3174935809:1536 > 1396672799:12612 TCP
62 22 18 VLAN_Q 50 EM_54 file_hash:
586DAB21FE587A7204FFA03149E73525

Más concretamente, la primera tabla representa la información de contabilidad para un cierto número de flujos. Las últimas dos columnas de cada fila representan el tipo de tráfico y la localización geográfica. Como esta es la captura previamente a pasar a través del DMES, estas columnas tienen el valor de 00.

La segunda tabla representa la información de metadatos asociada con el mismo periodo que la información de contabilidad representada en la primera tabla. En esta tabla el tipo de cada paquete está marcado en gris:

- HTTP_GET → Solicitud HTTP
- GET_PEERS_RESPONSE → Mensaje de señalización para Bittorrent. Indica la IP y puerto de otras máquinas que ejecutan esta aplicación.
- EM_54 → Mensaje de señalización de eMule.

Después de correlacionar los metadatos con la base de datos interna de firmas es posible determinar que uno de los mensajes HTTP_GET puede volver a categorizarse a un tipo mejor (FACEBOOK) que indica que los metadatos representan una solicitud HTTP a un servidor de Facebook.

La siguiente tabla representa a los metadatos en la salida del DMES:

```
USR 1301459263.895152 1396673130:49569 3269476872:80 TCP 529
489 18 VLAN_Q 50 | FACEBOOK | 07 GET /hprofile-
Windows snc4/187272_1369476520_2672812_q.jpg HTTP/1.1 Mozilla/5.0 (Windows; U; Windows NT 6.0; en-US) AppleWebKit/534.16 (KHTML, like Gecko)
Chrome/10.0.648.204 Safari/534.16 profile.ak.fbcdn.net
RED 1301459263.895272 1599070452:31446 1395791963:23415 UDP
661 633 VLAN_Q 50 GET_PEERS_RESPONSE n_peers: 22
1406509822:42600 1595098503:15820 3066631082:27234
1394625343:24735 3563778190:15738 773555601:21607
1476399821:14528 1414351873:17078 3654843008:18669
1429675938:22124 1417290889:24633 1435079862:23633
1396344029:24938 1417452430:17941 1477912353:25503
1436034701:24076 411762532:16881 3192951174:58539
USR 1301459263.895332 1343932535:16018 > 1592110395:80 TCP 1030
990 18 VLAN_Q 50 GET_HTTP 07 GET
/opt/icons/icon_error_with_bg.gif HTTP/1.1 Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US) AppleWebKit/534.16 (KHTML, like Gecko)
Chrome/10.0.648.151 Safari/534.16 es.madbid.com
RED 1301459263.895371 3174935809:1536 1396672799:12612 TCP
62 22 18 VLAN_Q 50 EM_54 file_hash:
586DAB21 FE587A7204FFA03149E73525
```

La información de contabilidad, cuando se correlaciona con estos metadatos actualizados adquiere el tipo de tráfico que es cada flujo. Adicionalmente, correlacionando las IP de los flujos con el diccionario de localización geográfica es posible determinar la localización geográfica de los servidores.

La siguiente tabla representa la información de contabilidad en la salida del DMES:

1396673130:49569	3269476872:80	TCP	4	1	5360	40	VLAN_Q 50	FACEBOOK	396
1394646482:50108	3174935809:1536	TCP	2	0	2680	0	VLAN_Q 50	EMULE	32
1394625343:24735	1396297335:48384	UDP	0	1	0	1466	VLAN_Q 50	BITTORRENT	396
1343932984:55259	1396055224:21784	TCP	5	4	5748	160	VLAN_Q 50	00	00
1436034701:24076	1361312813:3565	TCP	0	1	0	1188	VLAN_Q 50	BITTORRENT	396
1395069195:12259	3181184896:12408	UDP	1	0	63	0	VLAN_Q 50	00	145
1394646123:3322	3174935809:1536	TCP	3	0	156	0	VLAN_Q 50	EMULE	439
1343932535:16018	1592110395:80	UDP	1	0	129	0	VLAN_Q 50	HTTP GET	439
1395791963:23415	1114410499:51413	UDP	0	1	0	165	VLAN_Q 50	00	439
1395069348:54768	3654843008:18669	TCP	1	2	1440	109	VLAN_Q 50	BITTORRENT	00
1334864840:56106	1334904428:22938	UDP	0	1	0	1430	VLAN_Q 50	00	396
1396672799:12612	1440435422:3243	TCP	3	1	4172	40	VLAN_Q 50	EMULE	354

Puede observarse que se han rellenado al menos dos columnas. La primera de ellas tiene el tipo de tráfico y la segunda un código numérico que identifica un país. Como puede observarse, en este ejemplo algunos flujos aún tienen el código 00 para el tipo de tráfico y/o la localización geográfica. Esto significa que el DMES no tuvo suficiente información para enriquecer todos los flujos, de modo que la actualización de las firmas y el reprocesamiento darían como resultado la identificación total del tráfico. Cuando se reprocesa, solo se analizarían por el DMES los flujos que no se enriquecieron previamente, ahorrando en esta forma tiempo de procesamiento.

Ventajas de la invención

Las características principales del Sistema de Enriquecimiento de Metadatos de la DPI son que mantiene el formato de datos, que está dirigido al procesamiento de pesadas correlaciones de datos y que las tareas realizadas por el DMES pueden seleccionarse previamente al inicio del análisis. Estas características implican algunos importantes beneficios:

- El DMES no necesita modificarse cuando se requieren cambios de análisis. Esto es debido a que las correlaciones se realizan siempre de la misma manera, siendo las fuentes de los datos en sí mismas (fuentes de datos externas, interpretación de metadatos y firmas) las que cambian, pero no el sistema.
- La realización del enriquecimiento por separado del análisis del tráfico permite que este último sea mucho más simple de modo que pueda realizarse usando lenguajes de scripting, que son mucho más fáciles de programar y están orientadas específicamente a seguir el procesamiento.
- El resultado del Sistema de Enriquecimiento de Metadatos de la DPI tiene el mismo formato que su entrada. Esto implica que cualquier análisis que pudiera realizarse usando directamente la salida de la Interfaz de Metadatos Descriptivos puede realizarse también a la salida del DMES, asegurando en esta forma la compatibilidad.
- Que el DMES mantenga el formato de datos implica que la salida del sistema puede usarse como su entrada para una nueva iteración. Esto implica que tras el procesamiento de una captura, la captura original puede borrarse dado que, en caso de que se requiera reprocesamiento en el DMES, puede usarse la salida previa, reduciendo de esta forma necesidades de almacenamiento.
- El DMES puede enriquecer los datos selectivamente. Esto significa que si es necesario un reprocesamiento debido a que ha cambiado la información que afecta a un cierto protocolo o a una correlación específica es posible aplicar el posprocesamiento solo a la parte del análisis que cambió, ahorrando en esta forma tiempo de procesamiento.

Un experto en la materia podría introducir cambios y modificaciones en las realizaciones descritas sin apartarse del alcance de la invención tal como se define en las reivindicaciones adjuntas.

Acrónimos

DMES Sistema de Enriquecimiento de Metadatos de la DPI
DPI Inspección a Fondo de Paquetes
FLV Video Flash
HTTP Protocolo de Transferencia de Hipertexto

REIVINDICACIONES

1. Un método para minimizar el posprocesamiento del tráfico de red, que comprende la correlación y procesamiento de al menos parte de un resultado compuesto de metadatos y datos de contabilidad del tráfico con información incluida en dichos metadatos, dichos datos de contabilidad del tráfico, firmas de protocolo almacenadas centralmente y fuentes de datos externas de datos de localización geográfica relacionados con el tráfico de red, obtenidos dichos metadatos y dichos datos de contabilidad del tráfico a partir de una Interfaz de Metadatos Descriptivos de un despliegue de Inspección a Fondo de Paquetes (DPI) de una red, **caracterizado por que** incluye un proceso de enriquecimiento que comprende correlación y reprocesamiento de dichos resultados correlacionados y procesados previamente compuestos de metadatos y datos de contabilidad del tráfico.

2. Un método de acuerdo con la reivindicación 1, en el que solo se proporcionan a dicho proceso de enriquecimiento una parte de dichos metadatos y/o una parte de dichos datos de contabilidad del tráfico.

3. Un método de acuerdo con la reivindicación 1 o 2, que comprende la realización de dicho reprocesamiento solamente a dichos metadatos enriquecidos y a información de contabilidad del tráfico enriquecida afectada por actualizaciones aplicadas a dichas firmas de protocolo centralizadas y/o dichas fuentes de datos externas.

4. Un sistema para minimizar el posprocesamiento del tráfico de red, que comprende

- medios para correlacionar y procesar al menos parte de un resultado compuesto de metadatos y datos de contabilidad del tráfico con información incluida en dichos metadatos, dichos datos de contabilidad del tráfico, firmas de protocolo almacenadas centralmente y fuentes de datos externas de datos de localización geográfica relacionados con el tráfico de la red;
- una Interfaz de Metadatos Descriptivos para proporcionar un resumen del tráfico de una red que incluye dichos metadatos, y
- un almacenamiento para dichas firmas de protocolo almacenadas centralmente,

caracterizado por que comprende adicionalmente medios de correlación y procesamiento adaptados para realizar un enriquecimiento de dicha salida correlacionada y procesada previamente compuesta de metadatos y datos de contabilidad del tráfico.

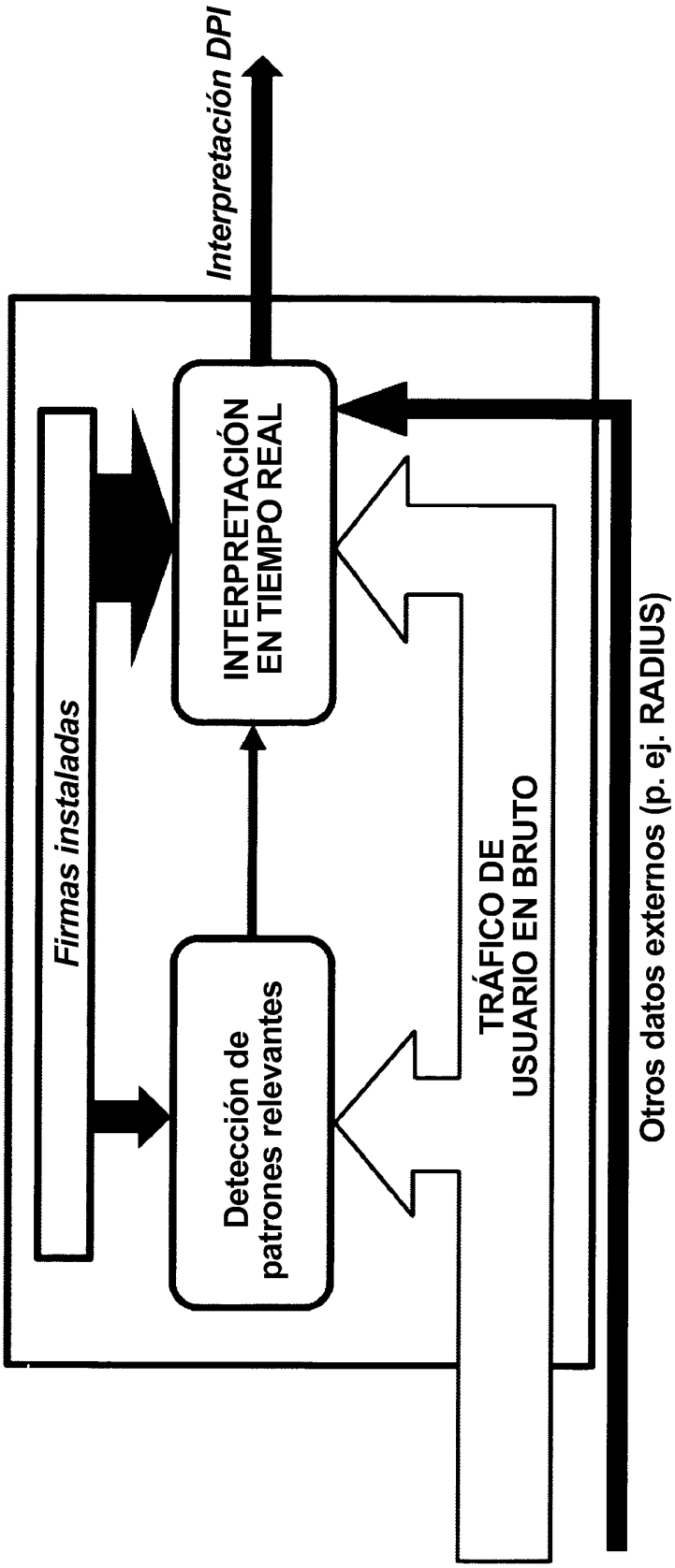
Interpretación DPI
INTERPRETACIÓN EN TIEMPO REAL
Firmas instaladas
Detección de patrones relevantes
TRÁFICO DE USUARIO EN BRUTO
Otros datos externos (p. ej. RADIUS)

FIG. 1

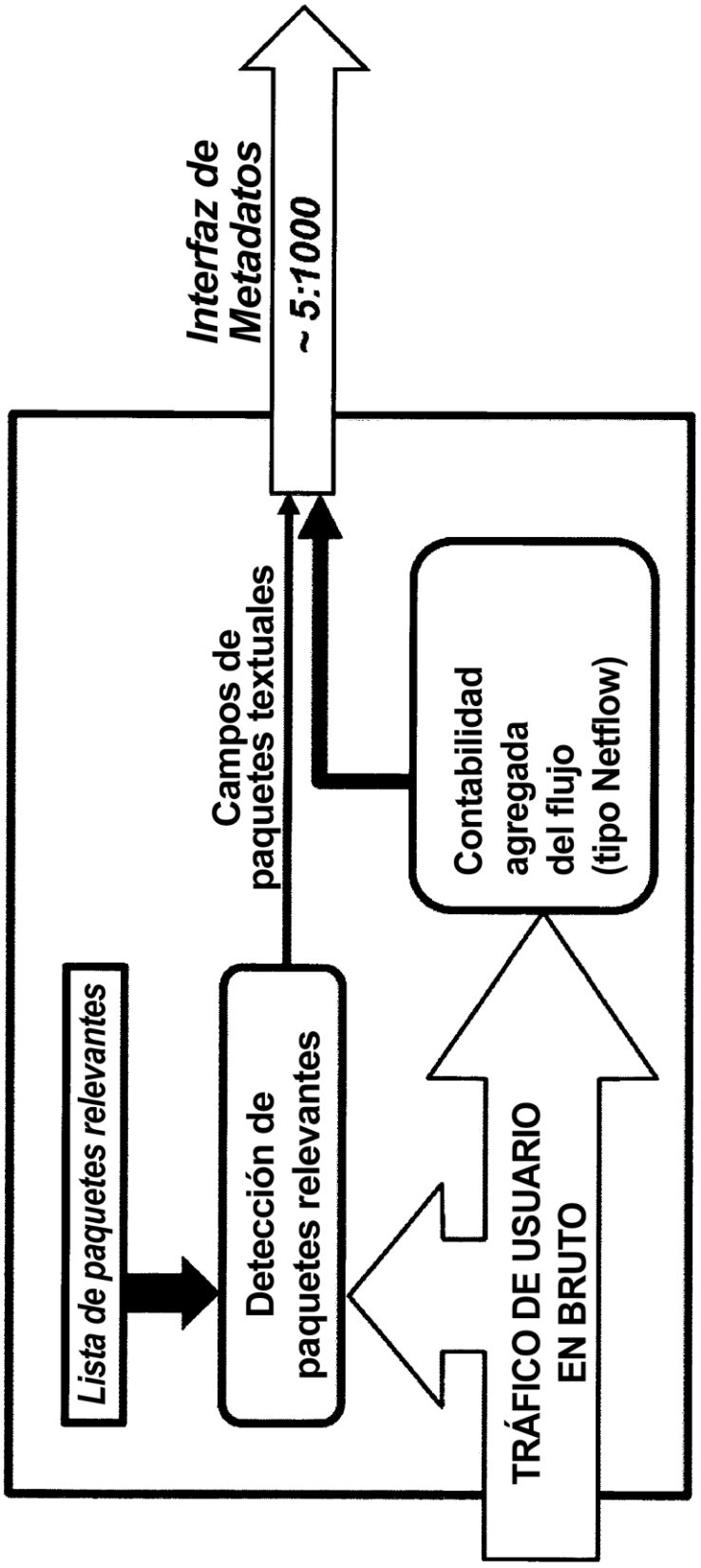
Interfaz de Metadatos
~ 5:1000
Campos de paquetes textuales
Contabilidad agregada del flujo (tipo Netflow)
Lista de paquetes relevantes
Detección de paquetes relevantes
TRÁFICO DE USUARIO EN BRUTO

FIG. 2

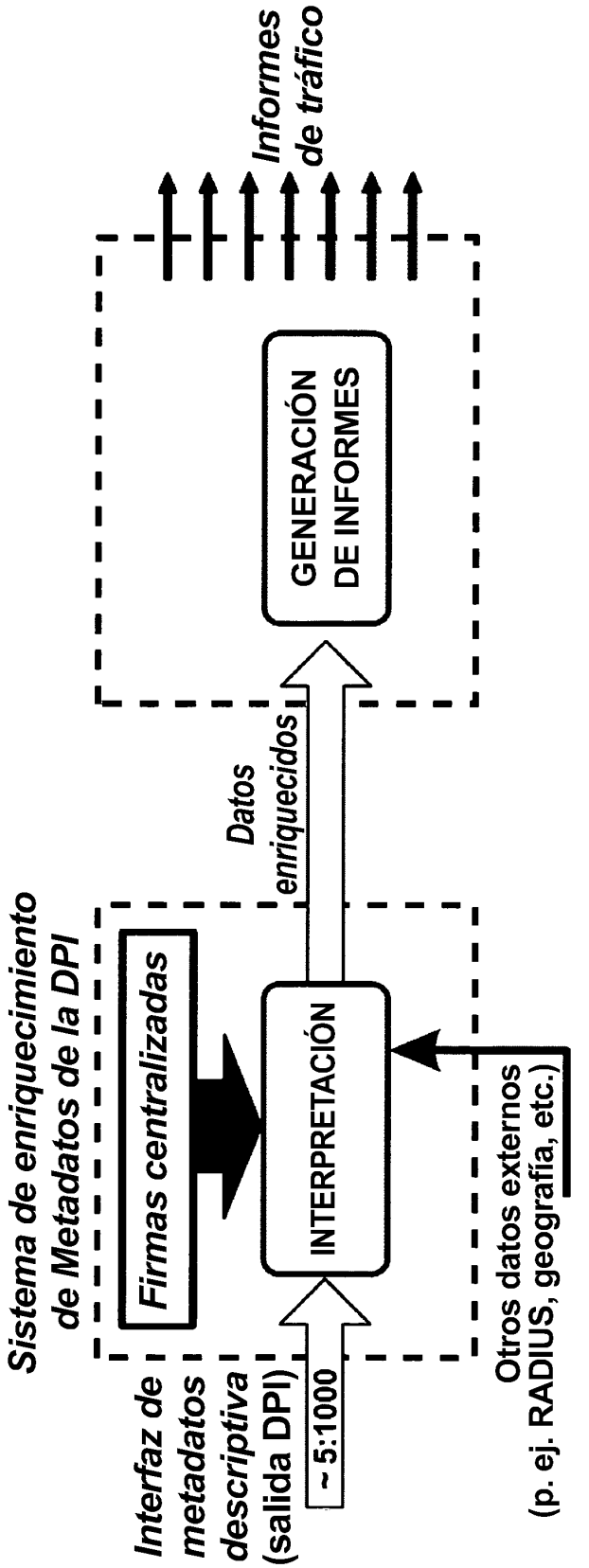
Informes de tráfico
GENERACIÓN DE INFORMES
Datos enriquecidos
Sistema de enriquecimiento de Metadatos de la DPI
Firmas centralizadas
INTERPRETACIÓN
Interfaz de metadatos descriptiva (salida DPI)
~ 5:1000
Otros datos externos (p. ej. RADIUS, geografía, etc.)

FIG. 3

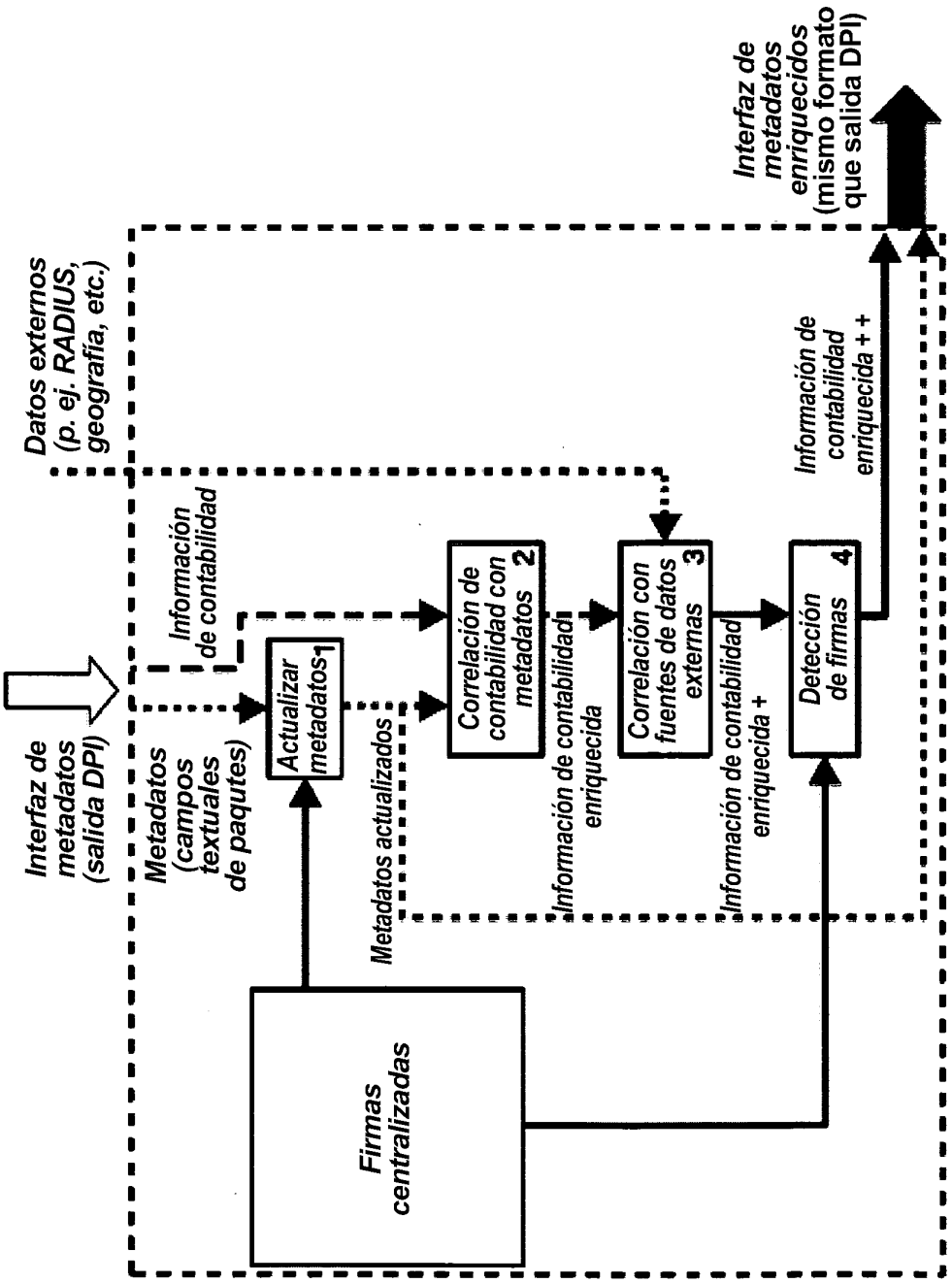
Interfaz de metadatos (salida DPI)
Metadatos (campos textuales de paquetes)
Información de contabilidad
Datos externos (p. ej. RADIUS, geografía, etc.)
Actualizar metadatos 1
Metadatos actualizados
Correlación de contabilidad con metadatos 2
Información de contabilidad enriquecida
Correlación con fuentes de datos externas 3
Información de contabilidad enriquecida +
Detección de firmas 4
Información de contabilidad enriquecida + +
Interfaz de metadatos enriquecidos (mismo formato que salida DPI)
Firmas centralizadas

FIG. 4

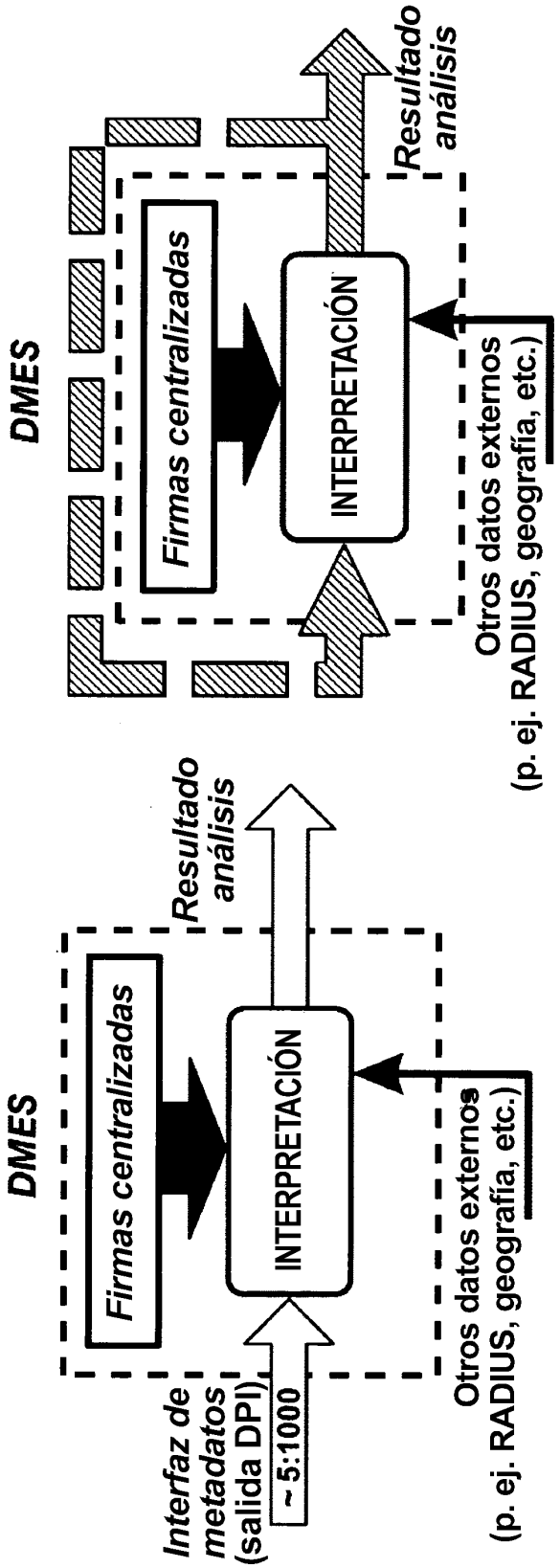
DMES
Firmas centralizadas
INTERPRETACIÓN
Resultado análisis
Otros datos externos (p. ej. RADIUS, geografía, etc.)
Interfaz de metadatos (salida DPI)
~ 5:1000
DMES
Firmas centralizadas
INTERPRETACIÓN
Resultado análisis
Otros datos externos (p. ej. RADIUS, geografía, etc.)

FIG. 5