

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 683 850**

51 Int. Cl.:

H04L 12/26 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.03.2014 PCT/EP2014/055745**

87 Fecha y número de publicación internacional: **25.09.2014 WO14147241**

96 Fecha de presentación y número de la solicitud europea: **21.03.2014 E 14712279 (0)**

97 Fecha y número de publicación de la concesión europea: **16.05.2018 EP 2976855**

54 Título: **Procedimiento para determinar el número y el posicionamiento de sondas DPI en una red de comunicación**

30 Prioridad:
22.03.2013 FR 1300673

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.09.2018

73 Titular/es:
**THALES (100.0%)
45, rue de Villiers
92200 Neuilly-sur-Seine, FR**

72 Inventor/es:
**BOUET, MATHIEU;
LEGUAY, JÉRÉMIE y
CONAN, VANIA**

74 Agente/Representante:
SALVÀ FERRER, Joan

ES 2 683 850 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

Procedimiento para determinar el número y el posicionamiento de sondas DPI en una red de comunicación

- 5 **[0001]** El objeto de la invención es el de los procedimientos para determinar el número y el posicionamiento de las sondas DPI en una red de comunicación.
- [0002]** En informática, y más particularmente en seguridad cibernética, un análisis por inspección de los paquetes en profundidad (DPI en lo sucesivo para «deep packet inspection» en inglés) es la actividad que consiste
10 en analizar el contenido (encabezamiento y carga útil) de un paquete de datos (tal como un datagrama IP y su contenido) interceptado en un flujo de datos que circula en una red de comunicación, de forma que se obtengan estadísticas de utilización de la red, filtrar los paquetes, detectar unas intrusiones, etc.
- [0003]** Para dicho análisis DPI, se conoce la conexión a un equipo de infraestructura de la red de
15 comunicación, como un router, un dispositivo apropiado para efectuar esta inspección de los paquetes en profundidad. Dicho dispositivo se denomina en lo sucesivo «sonda DPI».
- [0004]** Actualmente, una sonda DPI es un equipo dedicado a la realización de la función de análisis DPI que está conectado a un equipo de infraestructura de la red. A corto plazo, es igualmente posible que las sondas DPI
20 sean virtualizadas, es decir tomen la forma de aplicaciones informáticas ejecutadas en uno o varios de los equipos de infraestructura de la red.
- [0005]** Para analizar la totalidad de los flujos que circulan en la red, es entonces tentador proporcionar a la red un componente DPI, que permite aplicar el análisis DPI de los flujos, que consta de un número importante de
25 sondas DPI posicionadas a través de la red.
- [0006]** No obstante, una sonda DPI presenta un coste financiero importante ya sea para la compra de un equipo dedicado o para la licencia de utilización del software correspondiente.
- 30 **[0007]** Además, el funcionamiento de una sonda DPI presenta un coste en términos de recursos materiales (tiempos de ocupación de un procesador) y en términos de recursos energéticos (potencia eléctrica adaptada al equipo que realiza la función de análisis DPI).
- [0008]** Pero sobre todo, un análisis DPI debe efectuarse de manera totalmente transparente para el usuario
35 de la red de comunicación que por ejemplo, a partir de un terminal usuario, interroga a un terminal servidor.
- [0009]** No obstante, cuando un equipo de infraestructura posterior está equipado con una sonda DPI, los equipos de infraestructuras anteriores son reprogramados dinámicamente a fin de dirigir todo o parte de los paquetes de datos incidentes hacia el equipo de infraestructura posterior equipado con la sonda DPI. Así, la
40 aplicación de un análisis DPI modifica el reparto de los flujos de datos en la red de comunicación. Si la sonda está mal posicionada en la red o se utilizan muy pocas sondas, la redirección de los paquetes de datos puede conducir a sobrecargar ciertos enlaces de la red y, en unos casos extremos, a ir más allá de la capacidad en banda pasante de uno o varios enlaces de la red. Esto genera unos retrasos y unos errores en la transmisión de los paquetes. La calidad de servicio de la red de comunicación se encuentra deteriorada.
- 45 **[0010]** Además, al forzar el enrutamiento de un paquete a través de un equipo de infraestructura equipado con una sonda DPI, el camino seguido por un paquete entre el terminal emisor y el terminal destinatario se expande. Si la sonda está mal posicionada en la red o se utilizan muy pocas sondas, la modificación de los caminos de enrutamiento de los paquetes puede conducir a un sobreconsumo de ancho de banda y a un aumento importante
50 del tiempo de latencia (tiempo, evaluado al nivel de un terminal usuario, entre la emisión de un mensaje de solicitud con destino a un terminal servidor y la recepción de un mensaje de respuesta procedente de este terminal servidor).
- [0011]** Así, las sondas DPI deben estar posicionadas en buenos lugares en la red de comunicación a fin de permitir el análisis de un máximo de los flujos que circulan en la red sin por ello reducir además medir el ancho de
55 banda disponible o degradar la calidad de servicio.
- [0012]** Existe por tanto una necesidad de un procedimiento que permita determinar el número y el posicionamiento de las sondas DPI en una red de comunicación en función de un cierto número de criterios, tales como la reducción de los costes de despliegue y de explotación del componente DPI, el tratamiento eficaz de un

máximo de flujo, la minimización de la sobrecarga sobre la red de comunicación ocasionada por el análisis DPI, etc. Estos objetivos ortogonales son tanto más difíciles de respetar simultáneamente cuando el tamaño de la red de comunicación estudiada aumenta.

5 **[0013]** Por otro lado, el artículo de MIROSLAV, «an efficient genetic algorithm for solving the multilevel uncapacitated facility location problem», divulga un algoritmo genético que permite encontrar una solución óptima a un problema que consiste en determinar el enlace de implantación, a lo largo de una red de carreteras, de fábricas para a la vez minimizar los costes de implantación de estas fábricas (por ejemplo limitando su número) y minimizar los costes de transporte entre una fábrica y otra fábrica o un cliente. No obstante, este algoritmo se basa en la
10 utilización de una métrica predeterminada y constante en el transcurso de la ejecución del algoritmo.

[0014] La invención tiene por tanto el objetivo de paliar el problema precitado.

[0015] Con este fin, la invención tiene como objeto un procedimiento de determinación del número y del
15 posicionamiento de sondas DPI en una red de comunicación, constituyendo las sondas DPI un componente DPI que permite un análisis del tipo «inspección de los paquetes en profundidad» de paquetes de datos interceptados en la red, caracterizado porque consta de las etapas que consisten en:

- 20 a)- a partir de un grupo de individuos padres, generar una población de individuos niños, correspondiendo cada individuo a una configuración posible del componente DPI en términos de número y de posicionamiento de sondas DPI en la red de comunicación;
- b)- para cada individuo de la población de individuos niños, simular el comportamiento de la red de comunicación equipada con el componente DPI correspondiente a dicho individuo;
- c)- evaluar un valor de una función de coste para dicho individuo en función especialmente del comportamiento
25 simulado;
- d)- seleccionar en dicha población de individuos niños, un grupo de individuos niños que presenta los valores más reducidos de la función de coste; y, mientras no se respete una condición de convergencia, iterar las etapas a) a d) considerando como grupo de individuos padres para la iteración siguiente, el grupo de individuos niños de la iteración corriente, retener de lo contrario un individuo del grupo de individuos niños como individuo solución, dando
30 el individuo solución la configuración óptima del componente DPI.

[0016] Por la aplicación de un algoritmo genético, el procedimiento según la invención permite encontrar una solución aproximada a una solución que constituye un local óptimo para la necesidad expresada más arriba y que constituye de hecho un problema de optimización.
35

[0017] En efecto, este problema de optimización no puede ser resuelto de manera exacta por un procedimiento matemático que permite encontrar, en un cierto plazo, una solución exacta que respeta los diferentes criterios de interés. Se trata en efecto de un problema matemático que entra en la categoría de los problemas NP-completos.
40

[0018] Según unas realizaciones particulares, el procedimiento consta de una o varias de las características siguientes, tomadas aisladamente o según todas las combinaciones técnicamente posibles:

- 45 - la simulación del comportamiento de la red conlleva el cálculo de una matriz de enrutamiento para el individuo considerado;
- la función de coste es la suma ponderada de funciones de coste elementales que corresponden cada una a un criterio cuantitativo que debe respetar la configuración óptima del componente DPI;
- el criterio se escoge entre:

- 50 ◦ un primer criterio que consiste en minimizar el número de sondas DPI en el componente DPI;
- un segundo criterio que consiste en minimizar la modificación del ancho de banda utilizado con y sin el componente DPI;
- un tercer criterio que consiste en minimizar la modificación del tiempo de latencia con y sin el componente DPI;
- un cuarto criterio que consiste en minimizar el número de flujo que no son analizados por el componente DPI;

- 55 - la etapa de selección en la población de individuos niños de un grupo de individuos niños consiste en retener los K individuos que presentan los valores más reducidos de la función de coste;
- la etapa de generación de una población de individuos niños a partir de un grupo de individuos padres consiste en volver a copiar los individuos de dicho grupo en dicha población y en completar dicha población;

- la población se completa generando un individuo niño por hibridación de dos individuos padres escogido al azar en el grupo de individuos padres de manera que se obtenga un individuo niño intermedio y hacer mutar de manera aleatoria el individuo niño intermedio para obtener dicho individuo niño;
- estando representado un individuo por una lista ordenada de bits, una mutación consiste en bascular el valor de un bit a 0 con una primera probabilidad y en bascular el valor de un bit a 1 con una segunda probabilidad.

[0019] La invención y sus ventajas se comprenderán mejor con la lectura de la descripción que aparece a continuación, dada únicamente a título de ejemplo y realizada en referencia a los dibujos anexos, en los que:

- 10 - la figura 1 es una representación esquemática de una red de comunicación que debe ser equipada con un componente DPI;
- la figura 2 es una representación esquemática de una red modificada correspondiente a la red de la figura 1 equipada con un componente DPI;
- la figura 3 es una representación esquemática de una plataforma informática para la aplicación del procedimiento de determinación del número y de la posición de una pluralidad de sondas DPI en una red de comunicación dada; y
- 15 - la figura 4 es una representación en forma de bloques del procedimiento aplicado en la plataforma de la figura 3.

[0020] En la figura 1, se ha representado una red de comunicación 10 que se desea equipar con un componente DPI óptimo, constandingo este componente DPI de una o varias sondas DPI.

20

[0021] La red 10 está constituida por una pluralidad de N nudos y una pluralidad de L enlaces. Un enlace L_{ij} conecta entre ellos un nudo anterior N_i y un nudo posterior N_j . En lo sucesivo, los enlaces son bidireccionales ($L_{ij}=L_{ji}$).

25 **[0022]**

Cada nudo N_i está identificado por un número lógico que consiste por ejemplo en un número entero i .

[0023] Los nudos N_i son unos equipos informáticos de infraestructura que constituyen unas pasarelas que constituyen unos puntos de acceso a la red 10, como los nudos N_1 , N_4 , N_3 y N_7 o unos routers, como los nudos N_2 , N_5 y N_6 .

30

[0024] En la figura 1 unos terminales clientes 12 y 14 están conectados respectivamente a las pasarelas N_1 y N_4 y unos terminales servidores están conectados respectivamente a las pasarelas N_3 y N_7 .

[0025] Un terminal usuario es apropiado para emitir un mensaje de solicitud con destino a un terminal servidor. Este, en respuesta, es apropiado para emitir un mensaje de respuesta con destino al terminal cliente que haya emitido el mensaje de solicitud.

35

[0026] Los diferentes mensajes emitidos así son transmitidos en la red 10 en forma de paquetes de datos. El protocolo de comunicación en la red 10 es por ejemplo el protocolo TCP en IP.

40

[0027] A grandes rasgos, un paquete consta de una parte de encabezamiento que contiene, entre otras cosas, las direcciones de los terminales emisor y destinatario del paquete, y una parte de carga útil, que contiene todo o parte del mensaje generado por el terminal emisor.

[0028] Un paquete se propaga en la red 10 de la manera siguiente. Un nudo anterior que recibe un paquete lo vuelve a enviar en un enlace que sale hacia un nudo posterior. El nudo anterior selecciona el nudo posterior hacia el que vuelve a enviar el paquete entre el conjunto de los nudos posterior al que está unido, en función de la dirección del terminal destinatario indicado en la parte de encabezamiento del paquete. El nudo anterior consta de un conjunto de reglas de enrutamiento que permite por ejemplo asociar a la dirección de un terminal destinatario, el

50 identificador lógico del nudo posterior hacia el que volver a enviar el paquete.

[0029] El conjunto de los paquetes que circulan en un enlace constituye un flujo elemental en este enlace. Se expresa en porcentaje del ancho de banda de este enlace.

[0030] En la figura 1, se representan de manera esquemática los flujos elementales en los enlaces de la red 10 por un trazo más o menos espeso. Así, cuando la red 10 no está equipada con sonda DPI, el tráfico entre los nudos N_1 y N_2 es reducido, el tráfico entre los nudos N_1 y N_5 es medio y el flujo entre los nudos N_4 y N_5 es importante.

55

[0031] La red 10 se describe por una topología T0. La topología consiste en definir la geometría de la red 10 que da cada nudo de la red y cada enlace que conecta dos nudos de la red y la capacidad de la red 10 que da las características técnicas de cada nudo – en términos de función (router, pasarela equipada o no de módulos aplicativos del tipo cortafuegos u otro, etc.) y de capacidad de cálculo – y las de cada enlace – en términos de anchos de banda.

[0032] Los flujos a través de la red 10 durante un periodo de funcionamiento de esta son descritos por una matriz de tráfico M0. La matriz de tráfico M0 describe los intercambios de paquetes entre terminales conectados a la red 10 durante el período de utilización característico de la red. Este período puede corresponder a una utilización normal de la red 10, a una utilización en caso de ataque cibernético de la red 10, a un periodo de sobrecarga de la red 10 a causa de una campaña de spams, etc.

[0033] En la figura 2, se representa una red modificada 11. La red 11 corresponde a la red 10 aumentada de un componente Dpi. Esta consiste en una sonda DPI S5 conectada al equipo de infraestructura que constituye el nudo N5.

[0034] Este componente DPI es identificado por una lista ordenada que asocia a cada número lógico de un nudo de la red, un número binario correspondiente a la presencia (bit igual a 1) o no (bit igual a 0) de una sonda DPI.

[0035] El componente DPI de la red 11 es así identificado por la lista siguiente: (0, 0, 0, 0, 1, 0, 0).

[0036] El tráfico en la red 11 se modifica con respecto al tráfico en la red 10 de la figura 1, en igualdad de condiciones. El flujo elemental entre los nudos N1 y N2, mientras que era reducido, es ahora medio. Además, el camino seguido por los paquetes entre el terminal usuario 12 y el terminal servidor 22, en lugar de pasar por los nudos N1, N2 y N3, se modifica ahora de manera que pase por los nudos N1, N5, N6 y N3. El flujo de datos entre el nudo N2 y el nudo N3 es reducido mientras que era medio antes de la implementación de la sonda DPI S5.

[0037] El componente DPI de la red 11 constituye el resultado de la aplicación del procedimiento descrito más abajo. El componente DPI constituye una solución que permite realizar de manera óptima un análisis DPI de los flujos de datos que circulan en la red 11.

[0038] En la figura 3, se representa esquemáticamente una plataforma informática 50 programada para aplicar el procedimiento de determinación del número y del posicionamiento de sondas DPI.

[0039] La plataforma 50 consta de un procesador 52, como medio de cálculo, y una memoria 54, como medio de almacenamiento.

[0040] La plataforma consta igualmente de una interfaz hombre/máquina 56 apropiada para ofrecer a un usuario la posibilidad de introducir unos datos.

[0041] La memoria 54 consta de diferentes parámetros de configuración. Son introducidos por ejemplo por el usuario por medio de la interfaz 56.

[0042] Los parámetros de configuración constan de:

- la topología T0 de la red 10, para la que se busca una configuración DPI óptima;
- una matriz de tráfico M0 correspondiente a una utilización particular de la red 10;
- una lista 60 que asocia a cada criterio C(k) retenido en la optimización, un peso W(k) y eventualmente un umbral S(k). Esta lista es definida por el usuario por medio por ejemplo de un menú desplegable en la interfaz 56. A título ilustrativo, en la figura 3, cuatro criterios han sido seleccionados por el usuario;
- un grupo GP1 de G individuos padres.

[0043] La memoria 54 consta de una zona memoria 70 que consta de diferentes variables intermedias utilizadas durante la ejecución del programa 100:

- un grupo GPi de G individuos padres;
- una población PEi de P individuos niños;
- un vector Vi que asocia a cada individuo de la población PEi el valor de una función de coste;
- un grupo GEI de G individuos niños.

- [0044]** En lo sucesivo, un individuo corresponde a un componente DPI posible en la red 10. En consecuencia, un individuo es una lista ordenada que asocia a cada número lógico de un nudo de la red un bit que toma el valor 1 si el nudo considerado está conectado a una sonda DPI y el valor 0 si el nudo considerado no está conectado a una sonda DPI.
- [0045]** La memoria 54 almacena las instrucciones de un programa de ordenador 100 que, cuando son ejecutadas por el procesador 52, permiten la aplicación del procedimiento.
- 10 **[0046]** El programa 100 consta de un subprograma 110 de evaluación de una función de coste H para un individuo corriente l_j .
- [0047]** El subprograma 110 consta de un módulo 112 de modificación de la topología, que es apropiado para tomar como entradas la topología T_0 y las informaciones sobre el número y la posición de las sondas DPI indicadas por el individuo corriente l_j , para generar en salida una topología modificada T_j . En particular, la topología modificada T_j tiene en cuenta las modificaciones de las características del o de los equipos equipados de sondas DPI.
- 15 **[0048]** El subprograma 110 consta de un módulo 114 de simulación del tráfico en la red modificado definido por la topología T_j .
- 20 **[0049]** El módulo 114 toma en entrada la topología T_j generada en salida del módulo 112 y la matriz de tráfico M_0 contenido en los parámetros de configuración almacenados en la memoria 54.
- [0050]** El módulo 114 determina una matriz de enrutamiento R_j , es decir el conjunto de las reglas de enrutamiento aplicadas por cada uno de los nudos de la red modificada T_j para enrutar un flujo que se va a analizar hacia una sonda DPI.
- 25 **[0051]** Para ello, el módulo 114 aplica unos algoritmos complejos que permiten repartir el tráfico definido por la matriz de tráfico M_0 en la red modificada a la vez que se respeta toda una serie de limitaciones. Estas limitaciones pueden ser parecidas a ciertos de los criterios utilizados en la función de coste H.
- 30 **[0052]** Así, la matriz de enrutamiento R_j se calcula para definir el camino más corto entre un terminal emisor y un terminal destinatario, que pasa igualmente por un equipo de infraestructura equipado con una sonda DPI.
- 35 **[0053]** El subprograma 110 consta de una pluralidad de módulos 116(1) a 116(4) de cálculos parciales de la función de costes. Cada módulo 116(k) está asociado a un criterio $C(k)$ seleccionable por el usuario para construir la función de coste H. Cada módulo 116(k) es apropiado para tomar en entrada la matriz de enrutamiento R_j generada por el módulo 114 así como la topología modificada T generada en salida del módulo 112. Cada módulo 116(k) permite calcular un coste elemental $H_j(k)$ para el individuo corriente l_j .
- 40 **[0054]** Así, el módulo 116(1) corresponde al criterio $C(1)$ que consiste en minimizar el número total de sonda en el componente DPI. La función de coste parcial $H(1)$ es por ejemplo proporcional al número de sondas indicado en el individuo corriente l_j . Eventualmente, la función de coste parcial $H(1)$ es infinita cuando el número de sondas supera un número umbral $S(1)$ definido por el usuario.
- 45 **[0055]** Así, el módulo 116(2) corresponde al criterio $C(2)$ que consiste en minimizar la modificación de la banda pasante utilizada con y sin el componente DPI representado por el individuo corriente l_j .
- [0056]** Así, el módulo 116(3) corresponde al criterio $C(3)$ que consiste en minimizar la modificación del tiempo de latencia con y sin el componente DPI representado por el individuo corriente l_j .
- 50 **[0057]** Así, el módulo 116(4) corresponde al criterio $C(4)$ que consiste en minimizar el número de flujos que no son analizados el componente DPI representado por el individuo corriente l_j .
- 55 **[0058]** El subprograma 110 consta de un módulo 118 de agregación apropiado para calcular un coste global H, para el individuo corriente l_j , sumando los diferentes costes elementales $H_j(k)$ generados por los módulos 116(k) correspondientes a los criterios $C(k)$ mencionados en la lista 60.
- [0059]** Eventualmente, la suma se pondera utilizando los pesos $W(k)$ definidos por el usuario en la lista 60.

- [0060]** El programa 100 consta de un subprograma 120 de gestión de poblaciones de individuos.
- [0061]** El subprograma 120 consta de un módulo 122 de generación de una población de individuos niños PEi a partir de un grupo de individuos padres GPi.
- [0062]** El módulo 122 consta de diferentes submódulos:
- un submódulo 124 de copia que permite copiar los G individuos del grupo GPi como G primeros individuos de la población PEi;
 - un submódulo 126 de hibridación que permite crear un individuo intermedio a partir de un primer individuo padre IP1 y de un segundo individuo padre IP2 elegidos de manera aleatoria en el grupo GPi. Para ello, la lista asociada al individuo intermedio se construye a partir de los q primeros bits de la lista asociada al primer individuo padre IP1 y de los P-g últimos bits del segundo individuo padre IP2. El número entero q puede ser predeterminado o definido aleatoriamente a cada llamada del submódulo 126;
 - un submódulo 128 de mutación aleatoria, que permite alterar un individuo intermedio para generar un individuo niño IE, convirtiendo cada bit a 0 del individuo intermedio a 1 con una probabilidad p1 y cada bit a 1 del individuo intermedio a 0 con una probabilidad p2. De preferencia, p2 es superior a la probabilidad p1 de manera que tienda hacia un número reducido de sondas DPI en la red.
- [0063]** El subprograma 120 consta de un módulo 130 de cálculo de la función de coste para cada uno de los individuos Ij de una población de individuos niños PEi. El módulo 130 es de hecho apropiado para llamar al subprograma 110 para calcular el valor Hj de la función de coste de cada uno de los individuos Ij de la población de individuos niños PEi.
- [0064]** El subprograma 120 consta de un módulo 140 de selección apropiado para seleccionar los G individuos Ij que presentan los valores Hj de la función de coste más reducidos en una población de individuos niños PEi. El módulo 140 es apropiado para suministrar un grupo de individuos niños GEi.
- [0065]** El subprograma 120 consta de un módulo 150 de asignación apropiado para asignar un grupo de individuos niños GEi en salida del módulo 140 como grupo de individuos padres GPi en entrada del módulo 122, para la iteración siguiente del subprograma 120.
- [0066]** El subprograma 120 consta por último de un módulo 160 de convergencia que permite verificar el respeto de una condición de convergencia.
- [0067]** Mientras esta condición no es respetada, la población de individuos evoluciona iterando la ejecución de los subprogramas 110 y 120.
- [0068]** Cuando esta condición es respetada, la iteración de la ejecución de los subprogramas 110 y 120 es interrumpida y el individuo solución del problema de optimización es el individuo que tiene el valor Hj más reducido del grupo de individuos niños GEi obtenido durante la última iteración de los subprogramas 110 y 120.
- [0069]** Este individuo solución Is corresponde al componente DPI óptimo que trata de implementar en la red estudiada.
- [0070]** En la figura 4, el procedimiento consta de una etapa inicial 1000 de configuración durante la que el usuario informa sobre la topología inicial T0 de la red 10 que se va a estudiar, la matriz de tráfico M0, selecciona los criterios C(k) que se van a considerar en la función de coste H así como eventualmente los pesos W(k) y los umbrales S(k) que se van a asignar a cada uno de los criterios seleccionados (configuración de la lista 60), define un grupo GP1 de G individuos padres.
- [0071]** Después, el procedimiento propiamente dicho de optimización del número y del posicionamiento de las sondas DPI en la red 10 se prosigue por un bucle indexado por el número entero j.
- [0072]** En una etapa 1100, de ejecución del módulo 122, se genera una población de individuos niños PEi. Esta población consta de P individuos. P vale por ejemplo 50.
- [0073]** La generación se efectúa a partir de un grupo de individuos padres GPi leído en la zona 70. Este

grupo consta de G individuos. G vale por ejemplo 10.

- 5 **[0074]** En la primera iteración ($i=1$), el grupo GP_1 corresponde al grupo de individuos padres introducido por el usuario durante la etapa 1000 de configuración.
- [0075]** En una subetapa 1124, de ejecución de un submódulo 124, los G individuos del grupo GP_i son copiados, sin modificación, en la población PE_i . Esto permite seguir la evolución de la población de individuos entre dos iteraciones sucesivas y la convergencia del procedimiento.
- 10 **[0076]** La población PE_i se completa con $P-K$ individuos niños (o por ejemplo 10 individuos niños), que se obtienen respectivamente:
- ejecutando (etapa 1126) el submódulo 126 de hibridación para crear a partir de un primer individuo padre IP_1 y de un segundo individuo padre IP_2 , escogidos al azar entre los $K(K-1)/2$ pares posibles en el grupo GP_i , un individuo niño intermedio;
 - 15 - ejecutando (etapa 1128) el submódulo 128 de mutación aleatoria, que permite crear a partir del individuo niño intermedio, un individuo niño.
- [0077]** En la etapa 1200, correspondiente a la ejecución del módulo 130, el valor H_j de la función de coste H se calcula para cada individuo lj de la población PE_i .
- 20 **[0078]** Para ello, en la etapa 1212, de ejecución del módulo 112, la topología T_0 es modificada teniendo en cuenta el número y el posicionamiento de las sondas DPI correspondiente al individuo lj . Se obtiene una topología modificada T_j .
- 25 **[0079]** En la etapa 1214, de ejecución del módulo 114, la matriz de enrutamiento R_j de la red definida por la topología modificada T_j se calcula a partir de la topología modificada T_j y de la matriz de tráfico M_0 definida durante la etapa 1000 de configuración.
- 30 **[0080]** En la etapa 1216, los módulos 116(k) de los criterios $C(k)$ de la lista 60 se ejecutan sucesivamente. Así, en la etapa 1216(k), el cálculo de un valor $H_j(k)$ de un coste elemental $H(k)$ se efectúa a partir de la matriz de enrutamiento R_j y de la topología modificada T_j .
- 35 **[0081]** En la etapa 1218, correspondiente a la ejecución del módulo 118 de agregación, el valor H_j del coste global para el individuo lj se calcula sumando los diferentes valores $H_j(k)$ generados a las etapas 1216(k). Eventualmente, esta suma se pondera utilizando los pesos $W(k)$ definidos por el usuario durante la etapa 1000 y leídos en la lista 60.
- 40 **[0082]** Después de la etapa 1200, se obtiene un vector V_i . Un elemento V_{ij} corresponde al valor H_j de la función de coste para el individuo lj de la población PE_i .
- 45 **[0083]** En la etapa 1300, correspondiente a la ejecución del módulo 140, los G individuos lj de la población PE_i que presentan los valores H_j más reducidos (tal como se indican en el vector V_i) son seleccionados en un grupo de individuos niños GE_i .
- [0084]** En la etapa 1400, correspondiente a la ejecución del módulo 160, se efectúa una verificación de la condición de convergencia.
- 50 **[0085]** Si la distancia entre el valor mayor H_{jmax} y el valor más reducido H_{jmin} de los valores H_j de los individuos del grupo GE_i es superior a una distancia predefinida, el procedimiento continúa por la etapa 1500.
- [0086]** En la etapa 1500, de ejecución del módulo 150, el grupo de individuos niños GE_i es asignado como grupo de individuos padres GP_{i+1} para la iteración siguiente de las etapas 1100, 1200, 1300 y 1400.
- 55 **[0087]** Si, por el contrario, la distancia entre el valor mayor H_{jmax} y el valor más reducido H_{jmin} de los valores H_j de los individuos del grupo GE_i es inferior a la distancia predefinida, el procedimiento sale del bucle de iteración de las etapas 1100, 1200, 1300 y 1400.
- [0088]** Como variante, la condición de convergencia probada en la etapa 1400 consiste simplemente en

verificar si un número máximo de iteraciones ha sido superado (por ejemplo 100 iteraciones).

[0089] El individuo del grupo GE_i que presenta el valor H_j más reducido se retiene como individuo solución I_s.

5 **[0090]** Este individuo solución I_s constituye la solución buscada del procedimiento de optimización. Indica el número y el posicionamiento de las sondas DPI que se van a colocar en la red 10 para obtener un análisis DPI eficaz en sentido en que respeta los diferentes criterios C(k) seleccionados.

10 **[0091]** El procedimiento aplica un algoritmo genético que permite, por iteraciones sucesivas, hacer evolucionar una población de individuos que representan unas configuraciones posibles del componente DPI y seleccionar en esta población los individuos que se aproximan a la solución óptima buscada, es decir que presentan un valor de la función de coste más reducida teniendo en cuenta unos criterios establecidos.

15 **[0092]** El procedimiento permite converger rápidamente hacia una solución óptima.

REIVINDICACIONES

1. Procedimiento de determinación del número y del posicionamiento de sondas DPI, para «Deep Paquet Inspection», en una red de comunicación (10), constituyendo las sondas DPI un componente denominado DPI, que permite un análisis del tipo «inspección de los paquetes en profundidad» de paquetes de datos interceptados en la red,

caracterizado porque consta de las etapas que consisten en:

- 10 a- a partir de un grupo de individuos padres (G_{Pi}), generar una población de individuos niños (PE_i), correspondiendo cada individuo a una configuración posible del componente DPI en términos de número y de posicionamiento de sondas DPI en la red de comunicación;
- b- para cada individuo (lj) de la población de individuos niños (PE_i), simular el comportamiento de la red de comunicación equipada con el componente DPI correspondiente a dicho individuo;
- 15 c- evaluar un valor (H_j) de una función de coste (H) para dicho individuo en función especialmente del comportamiento simulado;
- d- seleccionar en dicha población de individuos niños (PE_i), un grupo de individuos niños (GE_i) que presenta los valores más reducidos de la función de coste; y,
- 20 mientras no se respete una condición de convergencia, iterar las etapas a a d considerando como grupo de individuos padres para la iteración siguiente (G_{Pi+1}), el grupo de individuos niños de la iteración corriente (GE_i), retener de lo contrario un individuo del grupo de individuos niños como individuo solución (ls), dando el individuo solución la configuración óptima del componente DPI.

2. Procedimiento según la reivindicación 1, en el que la simulación del comportamiento de la red conlleva el cálculo de una matriz de enrutamiento (R_j) para el individuo considerado (lj).

3. Procedimiento según la reivindicación 1 o la reivindicación 2, en el que la función de coste (H) es la suma ponderada de funciones de coste elementales ($H(k)$) correspondiente cada una a un criterio cuantitativo que debe respetar la configuración óptima del componente DPI.

30

4. Procedimiento según la reivindicación 3, en el que el criterio se escoge entre:

- un primer criterio que consiste en minimizar el número de sondas DPI en el componente DPI;
- un segundo criterio que consiste en minimizar la modificación del ancho de banda utilizado con y sin el componente DPI;
- 35 DPI;
- un tercer criterio que consiste en minimizar la modificación del tiempo de latencia con y sin el componente DPI;
- un cuarto criterio que consiste en minimizar el número de flujos que no son analizados por el componente DPI.

5. Procedimiento según cualquiera de las reivindicaciones 1 a 4, en el que la etapa de selección en la población de individuos niños (PE_i) de un grupo de individuos niños (GE_i) consiste en retener los K individuos que presentan los valores más reducidos de la función de coste (H_j).

6. Procedimiento según cualquiera de las reivindicaciones 1 a 4, en el que la etapa de generación de una población de individuos niños (PE_i) a partir de un grupo de individuos padres (G_{Pi}) consiste en copiar los individuos de dicho grupo en dicha población y en completar dicha población.

7. Procedimiento según la reivindicación 6, en el que la población se completa generando un individuo niño por hibridación de dos individuos padres escogido al azar en el grupo de individuos padres de manera que se obtenga un individuo niño intermedio y que se haga mutar de manera aleatoria el individuo niño intermedio para obtener dicho individuo niño.

8. Procedimiento según la reivindicación 7, en el que estando un individuo representado por una lista ordenada de bits, una mutación consiste en bascular el valor de un bit a 0 con una primera probabilidad ($p1$) y en bascular el valor de un bit a 1 con una segunda probabilidad ($p2$).

55

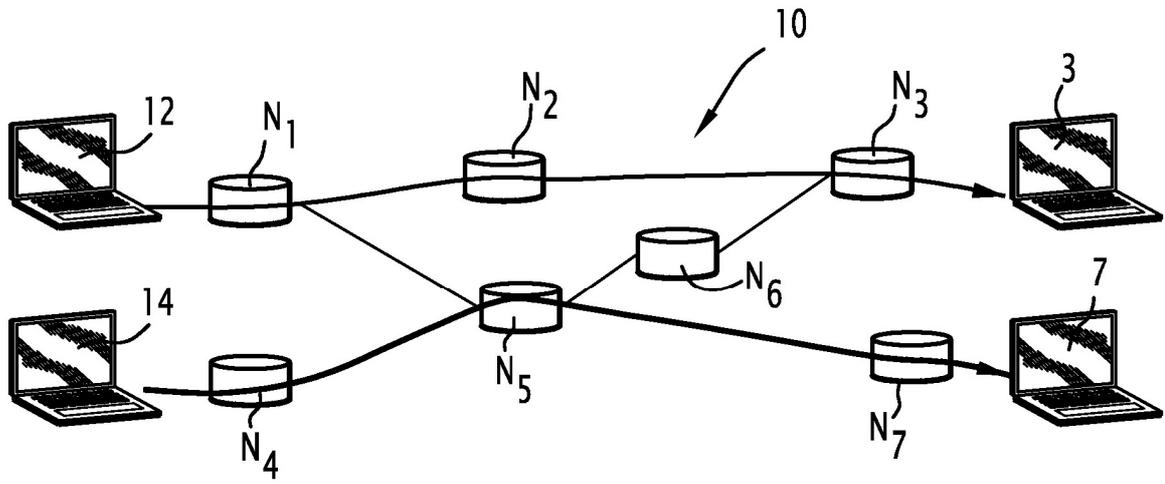


FIG.1

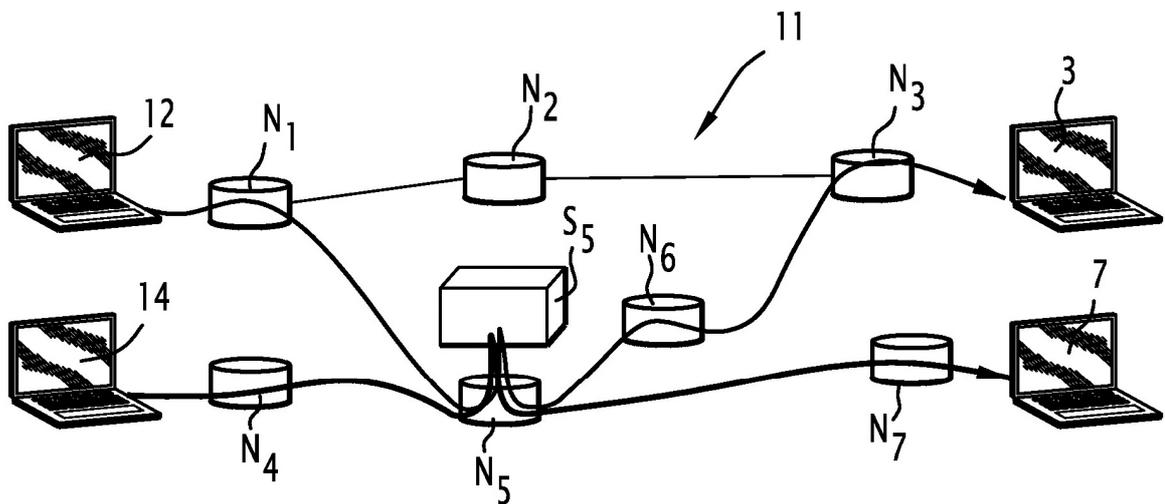


FIG.2

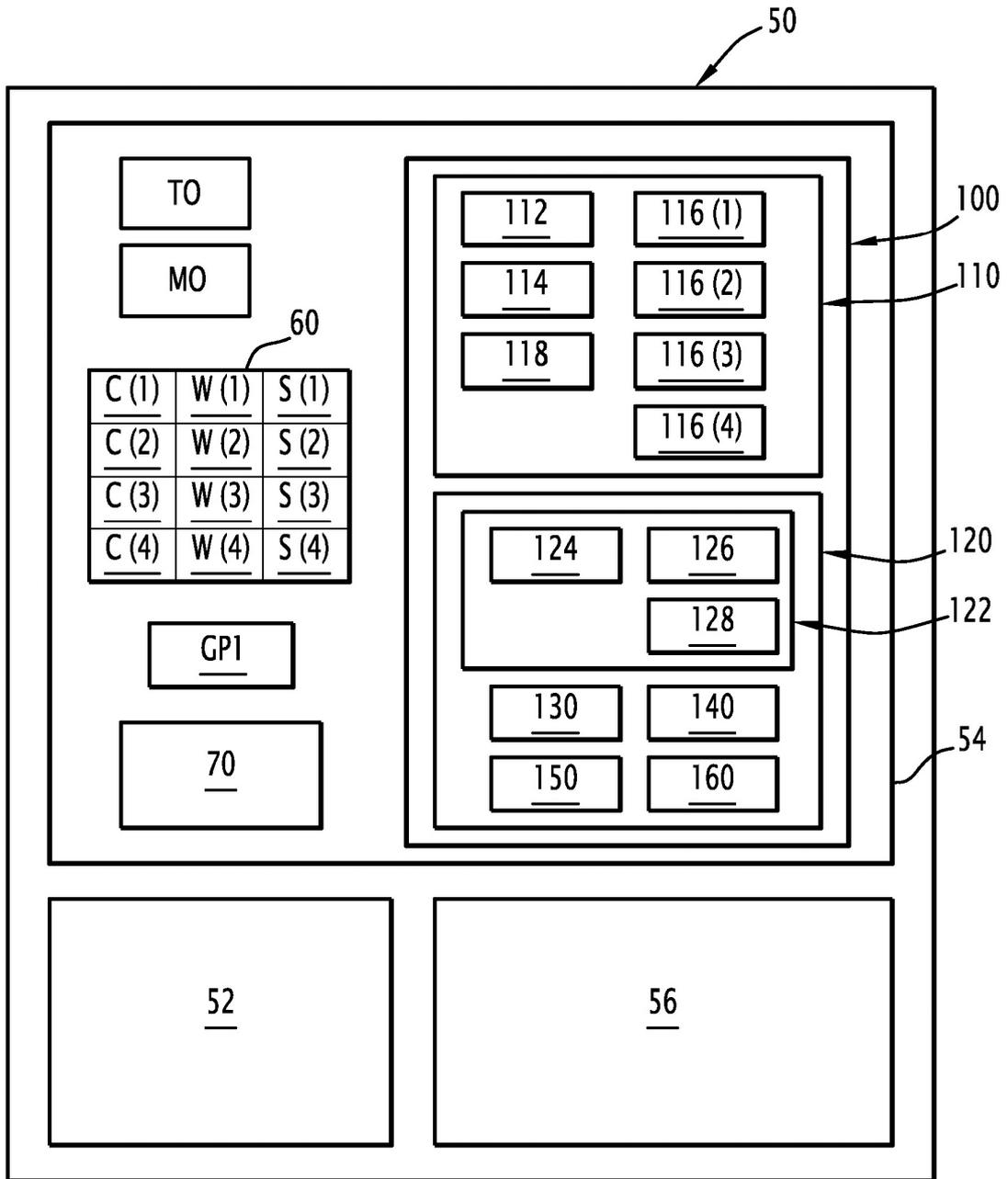


FIG.3

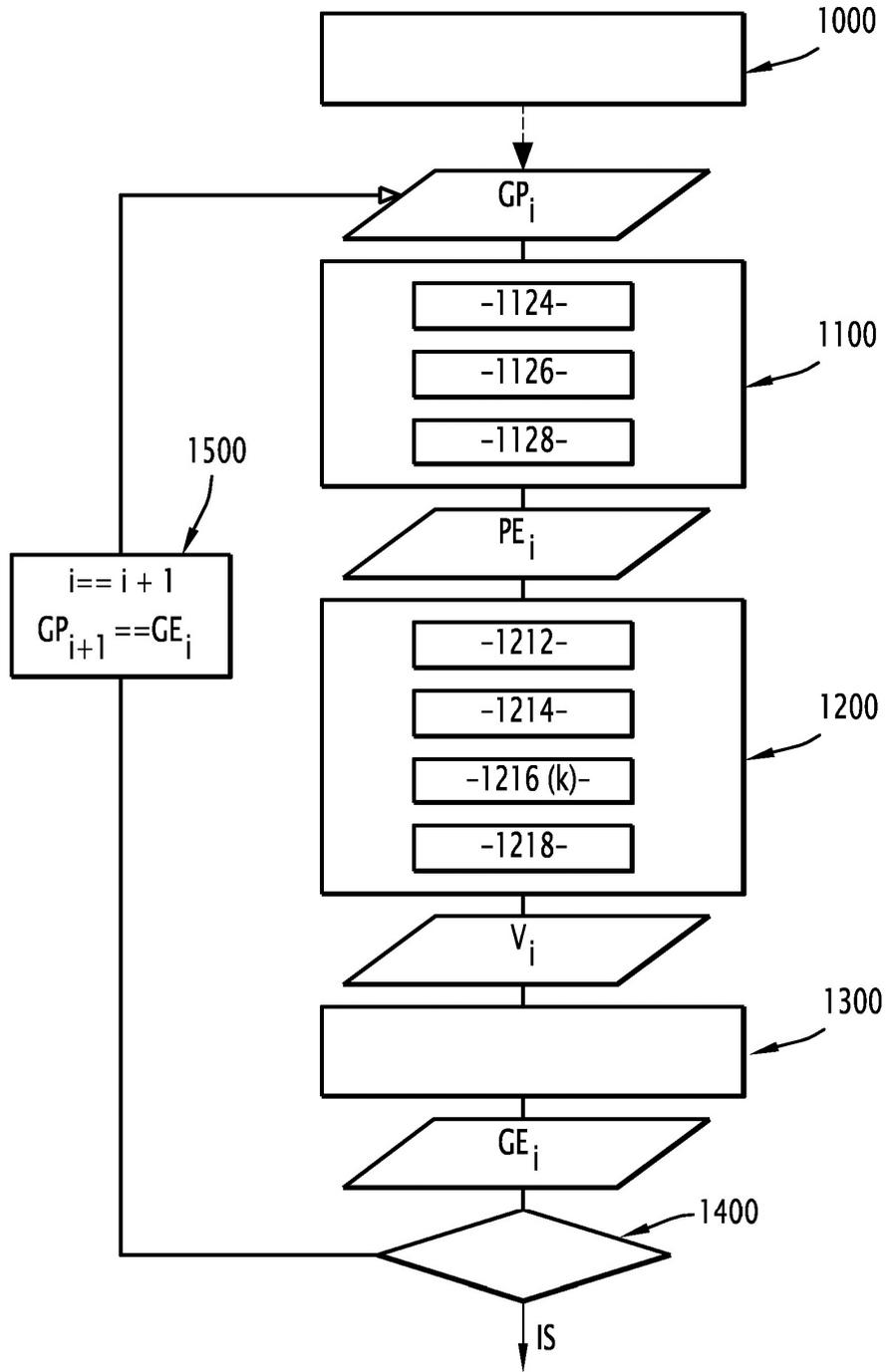


FIG.4