(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 648 587**

(51) Int. CI.:

H04L 29/06 (2006.01)
G06F 21/55 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **25.07.2014 E 14178515 (4)**

(97) Fecha y número de publicación de la concesión europea: **23.08.2017 EP 2838241**

(54) Título: **Procedimiento de detección de eventos sospechosos en un fichero de recopilación de informaciones relativas a un flujo de datos; soporte de registro y sistema asociados**

(30) Prioridad:

31.07.2013 FR 1301843

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
04.01.2018

(73) Titular/es:

THALES (100.0%)
45, rue de Villiers
92200 Neuilly Sur Seine, FR

(72) Inventor/es:

HUYOT, BENOÎT;
MABIALA, YVES;
SANS, STÉPHANE y
CHOLLON, LAURENT

(74) Agente/Representante:

SALVA FERRER, Joan

DESCRIPCIÓN

Procedimiento de detección de eventos sospechosos en un fichero de recopilación de informaciones relativas a un flujo de datos; soporte de registro y sistema asociados.

[0001] La invención tiene como campo el de la seguridad de las redes de comunicación y, más particularmente, de las redes de comunicación IP, tales como una red local de empresa.

[0002] En el presente documento, por «paquete de datos», se entiende un datagrama transmitido en la red de comunicación supervisada. Un paquete consta de una parte de encabezado, que contiene unos datos que permiten el envío del paquete desde un emisor hacia un destinatario a través de la red, así como una parte de carga útil, que contiene los datos de aplicación o mensaje, que el emisor desea compartir con el destinatario.

[0003] En el presente documento, por «sesión», se entiende el conjunto de los paquetes de datos intercambiados, según un mismo protocolo (FTP, HTTP, etc.), entre una primera máquina (identificada por una dirección IP y un puerto) y una segunda máquina (identificada por una dirección IP y un puerto).

[0004] En este campo, se conocen unos componentes informáticos de supervisión de la red, tales como unos sistemas de detección de intrusión, también denominados sistemas IDS (según el acrónimo inglés «Intrusion Détection System»). Tal componente es parametrazo por el editor que incorpora un cierto número de reglas de alerta de base, así como por el operador, que define unas reglas de alerta específicas. Cuando el componente detecta que un evento capturado en la red supervisada satisface al conjunto de los criterios de una regla de alerta, registra el evento sospechoso en una tabla de una base de datos. Un ejemplo de tal componente de supervisión es dado por el software libre SNORT.

[0005] Se conocen igualmente unos componentes informáticos de recopilación de informaciones relativas al flujo de datos que circulan en la red de comunicación supervisada.

[0006] Tal componente de recopilación genera periódicamente un fichero, del que cada línea corresponde a un evento.

[0007] Para cada par de máquinas que hayan intercambiado al menos un paquete de datos durante el período correspondiente al fichero, un evento representa el flujo elemental de la sesión de comunicación entre estas dos máquinas, durante este período. Así, un evento corresponde a la agregación de los paquetes transmitidos, según un mismo protocolo, entre dos máquinas durante el período de recopilación.

[0008] Un ejemplo de tal componente de recopilación es dado por el software NETFLOW desarrollado por la sociedad CISCO SYSTEMS. El fichero generado por este software respeta el protocolo NETFLOW, cuyas características de la versión 9 se publican en el documento RFC 3954 accesible en línea.

[0009] El IETF (según el acrónimo inglés «Internet Engineering Task Force») ha generalizado el protocolo NETFLOW en la norma IPFIX (según el acrónimo inglés «IP Flow Information Export»), definida en los documentos RFC 5101, RFC 5102 y RFC 5103 accesibles en línea. En lo sucesivo, se habla de «fichero NETFLOW» para todo fichero que responda a la norma IPFIX.

[0010] Un fichero NETFLOW consta así de unas informaciones relativas a un flujo elemental intercambiado entre dos máquinas, pero no consta de información que indica si un flujo elemental parece corresponder a un ataque o a una intrusión en la red supervisada.

[0011] Se conoce por el documento US 8 180 886 B2 un procedimiento de detección de ciber-ataques propio para detectar una modificación de los parámetros de una comunicación con respecto a unos perfiles aceptables, resultando dichos perfiles de un histórico de la utilización de la red. El procedimiento de D1 consta de la recepción de una comunicación corriente, definida al menos por unos primer y segundo parámetros independientes, después la asignación de una primera probabilidad al primer parámetro (a partir de una comparación con un primer perfil) y una segunda probabilidad al segundo parámetro (a partir de una comparación con un primer perfil) y una segunda probabilidad al segundo parámetro (a partir de una comparación con un segundo perfil). La probabilidad final se determina por el simple producto de la primera y segunda probabilidades. La probabilidad final se compara a continuación con un umbral. Una acción se origina en función del resultado de esta comparación.

[0012] No obstante, este documento del estado de la técnica no describe un sistema subyacente supervisado que consta de parámetros correlacionados entre ellos.

[0013] La invención tiene por tanto como objeto responder a este problema proponiendo un procedimiento que permite, sobre la sola base de al menos un fichero NETFLOW, detectar si un evento es sospechoso y asociarle una alerta.

[0014] Para ello, la invención tiene como objeto un procedimiento de detección de eventos sospechosos, un soporte de registro de informaciones y un sistema que consta de una unidad de control programada para ejecutar dicho procedimiento según las reivindicaciones.

[0015] Otras características y ventajas de la invención se mostrarán más claramente de la descripción detallada que aparece a continuación de un modo de realización particular, dado a título indicativo y nulamente limitativo, y realizada en referencia al dibujo anexo que representa esquemáticamente, en forma de bloques, un modo de realización particular de un procedimiento de detección de eventos sospechosos en un fichero de eventos generado por un componente informático de recopilación de informaciones relativas a un flujo de datos que circulan en una red de comunicaciones IP supervisada.

[0016] Un componente NETFLOW se implanta en un punto de una red local de comunicación IP, por ejemplo al nivel de una pasarela de conexión de esta red local a una red de comunicación IP pública, tal como INTERNET.

[0017] El componente NETFLOW es apto para el análisis del flujo de datos que transitan en la pasarela.

[0018] El componente NETFLOW es apto para generar periódicamente un fichero NETFLOW. Por ejemplo, el componente NETFLOW genera un fichero NETFLOW cada quince minutos.

[0019] Un fichero NETFLOW consta de unas informaciones relativas al flujo de datos en un período de supervisión que corresponde a los quince minutos que preceden a la creación del fichero.

[0020] Las informaciones del fichero NETFLOW están fechadas con la fecha de generación del fichero NETFLOW.

[0021] Un fichero NETFLOW consta de una pluralidad de líneas, correspondiendo cada línea a un evento. Cada línea se identifica por un número entero n.

[0022] Cada evento $X(n)$ corresponde a una sesión en el período de supervisión, es decir a la agregación de los paquetes intercambiados, según un mismo protocolo, entre una primera máquina (identificada por su dirección IP y su puerto de comunicación) y una segunda máquina (identificada por su dirección IP y su puerto de comunicación), pudiendo una máquina, durante una misma sesión, funcionar en emisor de un paquete o en destinatario de un paquete.

[0023] En el fichero NETFLOW, cada evento $X(n)$ se caracteriza por una pluralidad de L parámetros $X_i(n)$.

[0024] Por ejemplo, en un modo de realización actual, un evento $X(n)$ se define por siete parámetros:

- $X_1(n)$ corresponde al protocolo de la cuarta capa del modelo OSI (en general igual TCP o UDP, pero otros protocolos son posibles);
- $X_2(n)$ corresponde a la dirección IP de la primera máquina;
- $X_3(n)$ corresponde a la dirección IP de la segunda máquina;
- $X_4(n)$ corresponde al puerto de la primera máquina;
- $X_5(n)$ corresponde al puerto de la segunda máquina;
- $X_6(n)$ corresponde al número de paquetes intercambiados durante la sesión; y
- $X_7(n)$ corresponde al número de octetos intercambiados durante la sesión.

[0025] Una vez que un fichero NETFLOW se ha almacenado en un disco duro, es accesible a un sistema que consta de una unidad de cálculo, que se programa para ejecutar un procedimiento 100 de detección de eventos sospechosos.

[0026] De manera general, el procedimiento 100 está concebido para permitir detectar, en un fichero

NETFLOW, los eventos *X(n)* anormales o sospechosos.

[0027] Un evento se considera como sospechoso cuando, de forma comparativa al histórico de la utilización de la red supervisada, este evento corresponde a una utilización marginal de la red.

[0028] Para detectar estos eventos sospechosos, se utiliza un enfoque probabilístico. En efecto, las probabilidades permiten cuantificar el nivel de normalidad de un evento. Un evento sospechoso aparece entonces como un evento raro, es decir cuya probabilidad de ocurrencia es reducida.

[0029] Así, el procedimiento 100 consta primero de una estimación de la probabilidad P de ocurrencia de un evento *X* a partir de una muestra constituida por unos eventos *X(n)* del fichero NETFLOW fuente.

[0030] El procedimiento 100 consta a continuación de un cálculo del valor de la probabilidad de ocurrencia de un evento *X(n)* particular, utilizando la estimación anterior de la probabilidad P.

[0031] Por último, el procedimiento 100 consta del etiquetado del evento *X(n)* particular como evento sospechoso cuando el valor calculado de la probabilidad de ocurrencia satisface a una regla de alerta.

[0032] De una manera general, un estimador de una magnitud *G* se calificará como $\hat{G}$. Un estimador se puede obtener a partir de una muestra de N valores de la magnitud *G*. Se calificará entonces como $\hat{G}_N$.

[0033] Desde un punto de vista matemático, un evento *X* es una variable aleatoria multi-variada. Cada parámetro X_i de un evento es una variable aleatoria mono-variada.

[0034] La estimación de la probabilidad *P* de un evento *X* consiste primero en estimar una probabilidad individual para cada parámetro X_i. Las probabilidades individuales P_i se denominan con frecuencia «probabilidades marginales» de la probabilidad *P*.

[0035] Cuando el parámetro X_i. es una variable aleatoria cualitativa, es decir con valor en un conjunto discreto y terminado de elementos calificados como x_i^k, entonces un estimador $\hat{P}_i$ de la probabilidad individual X_i. se da por la relación:

$$\hat{P}_{i,N}(X_i = x_i^k) = \frac{1}{N}\sum_{n=1}^{N} 1_{X_i(n)=x_i^k} \qquad (1)$$

en la que *N* es el número de eventos tenido en cuenta en la estimación; y $1_{X_i(n)=x_i^k}$ es la función indicadora, vale 1 cuando su condición se realiza, si no 0

[0036] Según la relación (1), se trata de contar el número de eventos del fichero NETFLOW para los que $X_i(n) = x_i^k$, es llevado a la longitud *N* del histórico.

[0037] En una aplicación incremental, que permite actualizar los diferentes estimadores para tener en cuenta un nuevo evento *X(N)* del fichero NETFLOW, la relación (1) resulta:

$$\hat{P}_{i,N}(X_i = x_i^k) = \frac{N-1}{N}\hat{P}_{i,N-1} + \frac{1}{N}1_{X_i(n)=x_i^k} \qquad (1')$$

[0038] Cuando el parámetro X_i es una variable aleatoria cuantitativa, es decir con valor en un conjunto continuo, el procedimiento que se utiliza para estimar la probabilidad individual P_i, que es entonces una función

continua de X_i, aplica un estimador denominado «con núcleo».

[0039] Más particularmente, en el presente modo de realización, el estimador con núcleo aplicado es el denominado de Parzen-Rozenblatt.

[0040] El estimador $\hat{\grave{P}}_i$ se da por la relación siguiente:

$$\hat{P}_{i,N}(X_i = x_i) = \frac{1}{N\hat{h}_{i,N}} \sum_{n=1}^{N} K\left(\frac{[x_i - x_i(n)]}{\hat{h}_{i,N}} \right) \qquad (2)$$

En la que N es el número de eventos tenido en cuenta en la estimación; $x_i(n)$, el valor del parámetro X_i para la enésima observación, $X_i(n)$; K es una función de núcleo; y $\hat{\grave{h}}_{i,N}$, una estimación sobre una muestra de N eventos, de un parámetro de alisado h_i.

[0041] O incluso, en una aplicación incremental por la relación:

$$\hat{P}_{i,N}(X_i = x_i) = \frac{N-1}{N} \hat{P}_{i,N-1} + \frac{1}{N\hat{h}_{i,N}} K\left(\frac{[x_i - x_i(N)]}{\hat{h}_{i,N}} \right) \qquad (2')$$

[0042] La función de núcleo K se escoge como una función gaussiana estándar, es decir de esperanza nula y de varianza unitaria:

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$$

[0043] El parámetro de alisado h_i se obtiene utilizando la relación usualmente asociada a este núcleo:

$$h_i = 1{,}06.\sigma_i.N^{-\frac{1}{5}}$$

en la que σ_i es la desviación tipo de la distribución de los $X_i(n)$, que es igual a la raíz cuadrada de la varianza V_i de esta distribución.

[0044] El estimador $\hat{\grave{h}}_{i,N}$ se determina así por medio de una estimación de la varianza $\hat{\grave{V}}_{i,N}$, dada por la relación siguiente, que corresponde a un estimador sin inclinación:

$$\hat{V}_{i,N} = \frac{1}{N-1} \left(\sum_{n=1}^{N} \left[X_i(n) - \hat{\bar{X}}_{i,N} \right]^2 \right) \qquad (3)$$

en la que $\bar{\grave{X}}_i$ es el valor medio de la variable aleatoria X_i, cuya estimación es dada por:

$$\hat{\overline{X}}_{i,N} = \frac{1}{N}\sum_{n=1}^{N} X_i(n)$$

[0045] Para una actualización incremental, la relación (3) resulta:

$$\hat{V}_{i,N} = \frac{N-2}{N-1}\left(\hat{V}_{i,N-1} + (X_i(N) - \overline{X}_i)^2\right) \qquad (3')$$

[0046] Una vez que los estimadores $\hat{\grave{P}}_{i,N}$ de las probabilidades individuales se han obtenido, se trata de estimar la probabilidad *P* global.

[0047] Las variables aleatorias X_i no son independientes entre ellas, de modo que la probabilidad *P* de la variable *X* no es el simple producto de las probabilidades individuales P_i.

[0048] Para estimar la dependencia entre las variables aleatorias X_i, se utiliza la noción de par.

[0049] De manera general, para una variable *X* dos variada en las variables X_1 y X_2 la función de par *C* permite escribir:

$$F(X) = C(F_1(X_1); F_2(X_2))$$

[0050] En la que, de manera general, F designa la función de distribución de la variable *X* de probabilidad *P*, definida por:

$$F(X) = \int_0^X P(X')dX'$$

[0051] En el modo de realización actualmente considerado, se utiliza el par de Farlie-Gumbel-Morgenstern, cuya expresión matemática es dada por la relación siguiente:

$$C(F_1(X_1); F_2(X_2)) = F_1(X_1).F_2(X_2).(1+\theta(1-F_1(X_1))(1-F_2(X_2)))$$

en la que el acoplamiento entre las variables X_1 y X_2 es reducido, y el par corresponde a una aproximación al orden uno de la ecuación en derivadas parciales que establecen el modelo de la dependencia entre las dos variables. Toda la información relativa a la dependencia entre las variables X_1 y X_2 se encuentra en el parámetro θ.

[0052] En el caso presente, en el que la variable aleatoria *X* es L-variada (con L que vale 7), la aproximación anterior conduce a la expresión:

$$F(X) = C(F_1(X_1); F_2(X_2); ...; F_i(X_i); ...; F_L(X_L))$$

O

$$F(X) = \left(\prod_{i=1}^{L} F_i(X_i)\right)\left(1 + \sum_{i,j\neq i} \theta_{ij}(1-F_i(X_i))(1-F_j(X_j))\right) \qquad (4)$$

[0053] Esta aproximación permite simplificar los cálculos y resumir la información de dependencia entre los parámetros X_i de un evento *X* por un pequeño número de parámetros θ_{ij} (valiendo L 7, se deben calcular 21

parámetros θ_{ij}).

[0054] El experto en la técnica sabe que cada parámetro θ_{ij} está vinculado a la noción de correlación ρ_{ij} en el sentido de los rangos de Spearman, según la relación:

$$\theta_{ij} = 3\rho_{ij}$$

en la que ρ_{ij} es un coeficiente de correlación lineal, no entre las variables X_i y X_j, sino entre las funciones de distribución F_i y F_j de las variables X_i y X_j.

[0055] Así, teniendo en cuenta la relación conocida entre un coeficiente lineal y la varianza, tenemos:

$$\theta_{ij} = 3\frac{V(F_i(X_i)+F_j(X_j))-V(F_i(X_i))-V(F_j(X_j))}{2\sqrt{V(F_i(X_i))V(F_j(X_j))}} \qquad (5)$$

[0056] Para poder actualizar una estimación de cada parámetro θ_{ij} para tener en cuenta un nuevo evento *X(N)* del fichero NETFLOW, se utiliza la fórmula incremental de la estimación de la varianza V dada por la relación (3') anterior.

[0057] En referencia ahora a la figura, se va a presentar el modo de realización del procedimiento 100 actualmente considerado.

[0058] En la etapa 110 se lee una nueva línea del fichero NETFLOW. Esta línea corresponde a un nuevo evento *X(N)*.

[0059] En la etapa 112, para cada parámetro X_i, el estimador de la probabilidad individual, $\hat{P}_i$ se actualiza teniendo en cuenta el valor $X_i(N)$ del parámetro X_i para el nuevo evento *X(N)*.

[0060] Si el parámetro X_i es una variable cualitativa, se utiliza la relación (1'). Si el parámetro X_i es una variable cuantitativa, se utiliza la relación (2'). Se obtiene una función $\hat{P}_{i,N}$.

[0061] En la etapa 114, para cada parámetro X_i, un estimador de la función de distribución individual, $\hat{F}_i$ se actualiza a partir de $\hat{P}_{i,N}$ calculadas en la etapa 112. Se obtiene una función $\hat{F}_{i,N}$.

[0062] En la etapa 116, para cada par de parámetros X_i y X_j, se actualiza un estimador del parámetro θ_{ij} utilizando la relación (5) bajo una forma incremental y las $\hat{F}_{i,N}$ calculadas en la etapa 114. Se obtienen los veintiún coeficientes $\hat{\theta}_{i,N}$.

[0063] En la etapa 118, el valor de la probabilidad de ocurrencia *p(X(N))* del evento *X(N)* se calcula utilizando la relación (4) y los $\hat{F}_{i,N}$ y $\hat{\theta}_{i,N}$ determinados respectivamente en las etapas 114 y 116.

[0064] Después, en la etapa 120 se verifica si el evento *X(N)* respeta una regla de alerta.

[0065] En un modo de realización, la regla de alerta consiste en considerar que el evento *X(N)* es

sospechoso si el valor de la probabilidad de ocurrencia calculado, *p(X(N))*, es inferior a una probabilidad de referencia $p_{0,N}$, que de preferencia se determina dinámicamente. Por ejemplo, $p_{0,N}$ se determina de manera que los eventos sospechosos corresponden al 5% de los eventos que tienen las probabilidades de ocurrencia más reducidas. Tal modo de realización hace intervenir un método de estimación de cuantílicos, que permiten determinar $p_{0,N}$ a partir de la estimación de la probabilidad global *P(X)*, es decir de la relación (4) y los $\hat{F}_{i,N}$ y $\hat{\theta}_{i,N}$.

[0066] Como variante, el número No de eventos etiquetados como sospechosos que se van a extraer del fichero NETFLOW fuente es predeterminado. Así, en la etapa 120, cuando el valor *p(X(N))* forma parte de las N_0 probabilidades más reducidas del fichero NETFLOW fuente, el evento *X(N)* se etiqueta como sospechoso.

[0067] En la etapa 122, cuando el evento *X(N)* verifica la regla de alerta, la línea correspondiente del fichero NETFLOW se aumenta en un campo Alerta que toma el valor unidad.

[0068] Después, el número N se aumenta en una unidad, y las diferentes etapas 110 a 122 se ejecutan en el evento siguiente del fichero NETFLOW fuente.

[0069] El experto en la materia comprenderá que el presente procedimiento no realiza ninguna hipótesis a priori sobre los eventos sospechosos. En particular, el presente procedimiento no prevé ninguna fase previa de aprendizaje de lo que podría ser un evento sospechoso. La única hipótesis tomada es la debilidad de la probabilidad de ocurrencia de los eventos considerados como unos ataques.

[0070] El experto en la materia, especialista de las probabilidades, constatará que el presente procedimiento está particularmente bien adaptado a un evento descrito por la mezcla de variables cualitativas y de variables cuantitativas.

[0071] Durante unas primeras iteraciones del procedimiento, las diferentes estimaciones evolucionan rápidamente. No obstante, cuanto más aumenta la profundidad del histórico tenido en cuenta, más estables son los estimadores y más pertinente es el etiquetado de un evento como sospechoso. En la práctica, hace falta más de un fichero NETFLOW para alcanzar una convergencia, de modo que ventajosamente varios ficheros NETFLOW sean analizados sucesivamente, sin reinicialización de los estimadores entre cada fichero.

[0072] En la etapa 122, cuando el evento *X(N)* verifica la regla de alerta, la línea correspondiente del fichero NETFLOW se aumenta en un campo Alerta que toma el valor unidad.

[0073] Después, el número entero N se aumenta en una unidad, y las diferentes etapas 110 a 122 se ejecutan en el evento siguiente del fichero NETFLOW fuente.

[0074] El experto en la materia comprenderá que el presente procedimiento no realiza ninguna hipótesis a priori sobre los eventos sospechosos. En particular, el presente procedimiento no prevé ninguna fase previa de aprendizaje de lo que podría ser un evento sospechoso. La única hipótesis tomada es la debilidad de la probabilidad de ocurrencia de los eventos considerados como unos ataques.

[0075] El experto en la materia, especialista de las probabilidades, constatará que el presente procedimiento está particularmente bien adaptado a un evento descrito por la mezcla de variables cualitativas y de variables cuantitativas.

[0076] Durante unas primeras iteraciones del procedimiento, las diferentes estimaciones evolucionan rápidamente. No obstante, cuanto más aumenta la profundidad del histórico tenido en cuenta, más estables son los estimadores y más pertinente es el etiquetado de un evento como sospechoso. En la práctica, hace falta más de un fichero NETFLOW para alcanzar una convergencia, de modo que ventajosamente varios ficheros NETFLOW sean analizados sucesivamente, sin reinicialización de los estimadores entre cada fichero.

REIVINDICACIONES

1. Procedimiento de detección de eventos sospechosos en un fichero de eventos generado por un componente informático de recopilación de informaciones relativas a un flujo de datos que circulan en una red de comunicaciones IP, estando un evento **caracterizado por** una pluralidad de parámetros,
constando el procedimiento de una etapa de estimación de una probabilidad (P) de ocurrencia de un evento (X) a partir de una muestra constituida por unos eventos (X(n)) del fichero, y una etapa de asociación de una alerta a un evento particular del fichero cuando dicho evento particular respeta una regla de alerta basada en el valor de la probabilidad de ocurrencia de dicho evento particular calculada utilizando la probabilidad de ocurrencia estimada, **caracterizado porque** la estimación de una probabilidad (P) de ocurrencia de un evento consta, para cada parámetro (Xi) de un evento (X), una etapa (112) de estimación de una probabilidad individual (Pi) de ocurrencia de dicho parámetro, considerada como independiente de la estimación de los otros parámetros, a partir de una muestra constituida por los valores (Xi(n)) de dicho parámetro para unos eventos del fichero y una etapa (114) de estimación de una dependencia (θ_{ij}) entre los parámetros de un evento.

2. Procedimiento según la reivindicación 1, en el que, cuando dicho parámetro (Xi) es equiparable a una variable aleatoria cuantitativa, la estimación de la probabilidad individual de dicho parámetro aplica un estimador con núcleo.

3. Procedimiento según la reivindicación 2, en el que el estimador con núcleo aplicado es un estimador de Parzen-Rozenblatt.

4. Procedimiento según cualquiera de las reivindicaciones 1 a 3, en el que la estimación de una dependencia entre los parámetros de un evento aplica un par (C).

5. Procedimiento según la reivindicación 4, en la que el par aplicado es un par de Farlie-Gumbel-Morgenstern, que hace intervenir, para cada par de parámetros de un evento, un parámetro de dependencia entre los dos parámetros de dicho par.

6. Procedimiento según la reivindicación 5, en el que el parámetro de dependencia θ_{ij} entre el parámetro X_i y el parámetro X_j del par de Farlie-Gumbel-Morgenstern es dado por la relación:

$$\theta_{ij} = 3\frac{V(F_i(X_i)+F_j(X_j))-V(F_i(X_i))-V(F_j(X_j))}{2\sqrt{V(F_i(X_i))V(F_j(X_j))}}$$

En la que *F(Y)* es la función de distribución de la variable aleatoria *Y*.

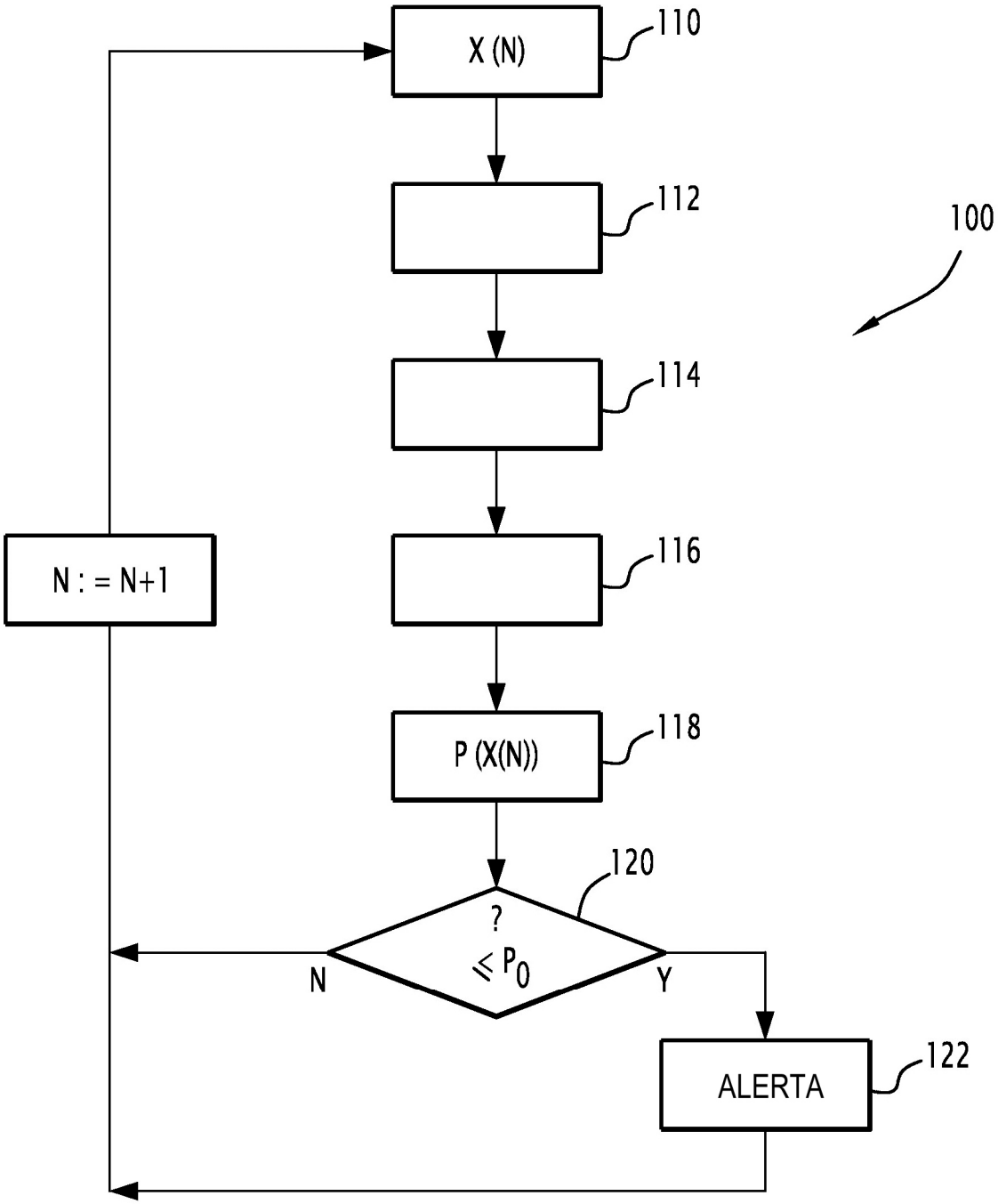
7. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la estimación de una probabilidad de ocurrencia de un evento se realiza de manera incremental, actualizando una estimación corriente de dicha probabilidad teniendo en cuenta un nuevo evento (X(N)) del fichero.

8. Procedimiento según cualquiera de las reivindicaciones 1 a 7, en el que la etapa (124) de asociación de una alerta a un evento particular del fichero consiste en calcular la probabilidad de ocurrencia de este evento particular utilizando la probabilidad estimada y en comparar esta probabilidad calculada con una probabilidad de referencia (p_0), determinada dinámicamente a partir del fichero.

9. Procedimiento según cualquiera de las reivindicaciones 1 a 8, en el que el fichero de eventos es un fichero NETFLOW.

10. Soporte de registro de informaciones, **caracterizado porque** consta de unas instrucciones para la ejecución de un procedimiento conforme al procedimiento según cualquiera de las reivindicaciones 1 a 9, cuando estas instrucciones son ejecutadas por un calculador electrónico.

11. Sistema que consta de una unidad de control, **caracterizado porque** dicha unidad de control está programada para ejecutar un procedimiento conforme al procedimiento según cualquiera de las reivindicaciones 1 a 9.


X (N)
110
112
100
114
N : = N+1
116
P (X(N))
118
120
?
≤ P0
N
Y
ALERTA
122