(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 801 273**

(51) Int. CI.:

G06F 21/55 (2013.01)
G06F 21/57 (2013.01)
H04W 12/12 (2009.01)
H04L 29/06 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **24.02.2016 PCT/CN2016/074424**

(87) Fecha y número de publicación internacional: **09.09.2016 WO16138830**

(96) Fecha de presentación y número de la solicitud europea: **24.02.2016 E 16758446 (5)**

(97) Fecha y número de publicación de la concesión europea: **08.04.2020 EP 3267348**

(54) Título: **Método y aparato para reconocer el comportamiento de riesgo**

(30) Prioridad:

02.03.2015 CN 201510093725

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
08.01.2021

(73) Titular/es:

ALIBABA GROUP HOLDING LIMITED (100.0%)
Fourth Floor, One Capital Place P.O. Box 847
George Town, Grand Cayman, KY

(72) Inventor/es:

MAO, RENXIN;
SUN, CHAO;
LI, XINKAI y
HE, DIJUN

(74) Agente/Representante:

VIDAL GONZÁLEZ, Maria Ester

DESCRIPCIÓN

Método y aparato para reconocer el comportamiento de riesgo

Campo técnico

La presente solicitud se refiere al campo de las tecnologías informáticas y, en particular, a un método y un aparato para identificar un comportamiento de riesgo de la red.

Técnica relacionada

Con el desarrollo de Internet, los comportamientos de las personas en la red se entrelazan con mayor frecuencia. Conceptualmente, un comportamiento de la red se refiere a un proceso de adquisición, envío o transmisión de datos de red por cada individuo de red en la red, que generalmente incluye: consulta de información, descarga de archivos, envío de correo y similares. Además de los comportamientos normales de la red, los comportamientos anormales de la red realizados por individuos en la red de manera intencional o no, tal como la información de navegación irrelevante para el trabajo de un empleado de la empresa durante el trabajo o la consulta ilegal de un historial de gastos por parte del personal de servicio al cliente de la red, pueden causar pérdidas. Para tratar el problema anterior, surge un sistema de monitoreo de riesgos para monitorear un comportamiento de riesgo de la red.

En la actualidad, un sistema convencional de monitoreo de riesgos, mediante la construcción de un motor de reglas, extrae y analiza las características de los comportamientos de la red que se ajustan a las definiciones de las reglas, identificando así los riesgos de los comportamientos de la red. Sin embargo, las reglas empleadas por el motor de reglas generalmente tienen vulnerabilidades, y es necesario agregar reglas continuamente para remediar las vulnerabilidades de las reglas. Esto definitivamente puede aumentar la carga de trabajo de los desarrolladores, y la eficiencia es baja. Además, el propio motor de reglas necesita consumir recursos informáticos adicionales, lo que causa por lo tanto una carga para un sistema informático.

El documento US 2014/359777 A1 divulga un servidor de administración de dispositivos móviles y un método para determinar el riesgo de seguridad para los dispositivos móviles implementados. El servidor de administración de dispositivos móviles recibe mediciones de riesgo de dispositivos móviles que se utilizan para calcular un puntaje de riesgo basado en las reglas. El puntaje de riesgo también se puede ajustar correlacionando las mediciones de riesgo recibidas con violaciones anteriores de seguridad o medidas de uso típicas. El puntaje de riesgo calculado se compara con uno o más umbrales para determinar si se debe tomar una acción de protección que se asocie con que se excede un umbral.

El documento US 7 574 382 B1 divulga un motor de detección de anomalías que monitorea el tráfico de la red para detectar pedidos realizados por los usuarios desde un catálogo electrónico de artículos, agrega datos sobre los pedidos detectados por período de tiempo y analiza los datos agregados para detectar anomalías en niveles de actividad asociados con artículos específicos en el catálogo. Para detectar si existe una anomalía en los datos de actividad asociados con un elemento determinado, se utiliza un algoritmo de pronóstico, tal como un algoritmo de suavizado exponencial, para generar un volumen de pedido esperado para un período de tiempo actual, y el volumen de pedido esperado se compara con un volumen de pedido real.

Resumen de la invención

La presente invención está definida por las reivindicaciones. Las modalidades de la presente solicitud proporcionan un método y un aparato para identificar un comportamiento de riesgo para resolver el problema de baja eficiencia en la técnica anterior causado al remediar una vulnerabilidad de la regla durante la identificación de un riesgo de comportamiento de la red, y el problema que un motor de regla consume más recursos informáticos.

El método para identificar un comportamiento de riesgo proporcionado en las modalidades de la presente solicitud incluye:

adquirir datos de comportamiento de un usuario;

determinar un coeficiente de riesgo de un enlace de comportamiento específico en los datos de comportamiento; y

evaluar, de acuerdo con el coeficiente de riesgo, si el enlace de comportamiento específico es riesgoso.

El aparato para identificar un comportamiento de riesgo proporcionado en las modalidades de la presente solicitud incluye:

un módulo de adquisición configurado para adquirir datos de comportamiento de un usuario;

un módulo de determinación configurado para determinar un coeficiente de riesgo de un enlace de comportamiento específico en los datos de comportamiento; y

un módulo de evaluación configurado para evaluar, de acuerdo con el coeficiente de riesgo, si el enlace de comportamiento específico es riesgoso.

La al menos una solución técnica empleada en las modalidades de la presente solicitud puede lograr los siguientes efectos beneficiosos:

En las modalidades de la presente solicitud, se adquieren datos de comportamiento de un usuario, y se selecciona un enlace de comportamiento específico de los datos de comportamiento; el coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento se determina mediante cálculos y, finalmente, se determina, de acuerdo con el coeficiente de riesgo, si el enlace de comportamiento específico es riesgoso. En comparación con la manera del motor de reglas, en el proceso anterior, no es necesario remediar una vulnerabilidad de la regla manualmente, mejorando así la eficiencia de la identificación del riesgo de comportamiento. Además, el proceso anterior evita la desventaja de que el motor de reglas consuma recursos informáticos adicionales, aliviando así la carga de un sistema informático.

Breve descripción de los dibujos acompañantes

Los dibujos adjuntos descritos aquí se utilizan para proporcionar una mejor comprensión de la presente solicitud y constituyen una parte de la presente solicitud. Las modalidades esquemáticas de la presente solicitud y la descripción de la misma se usan para ilustrar la presente solicitud, pero no constituyen limitaciones inadecuadas para la presente solicitud. En los dibujos:

La Figura 1 muestra un proceso de un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud;

La Figura 2 muestra un proceso de selección de un enlace de comportamiento específico a partir de datos de comportamiento en un método para identificar un comportamiento riesgoso de acuerdo con una modalidad de la presente solicitud;

La Figura 3 muestra un proceso para determinar un coeficiente de riesgo a corto plazo en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud;

La Figura 4 muestra un proceso para determinar un coeficiente de riesgo histórico en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud;

La Figura 5 muestra un proceso para determinar un coeficiente de riesgo del equipo en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud;

La Figura 6 muestra un proceso de evaluar si un enlace de comportamiento específico es riesgoso en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud; y

La Figura 7 es un diagrama estructural esquemático de un aparato para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud.

Descripción detallada

Para aclarar los objetivos, las soluciones técnicas y las ventajas de la presente solicitud, las soluciones técnicas de la presente solicitud se describen clara y completamente a continuación con referencia a las modalidades específicas de la presente solicitud y los dibujos correspondientes. Aparentemente, las modalidades divulgadas son simplemente algunas de las modalidades de la presente solicitud, en lugar de todas las modalidades. En base a las modalidades de la presente solicitud, todas las demás modalidades obtenidas por los expertos en la técnica sin pagar esfuerzos creativos pertenecen al alcance de protección de la presente solicitud.

La Figura 1 muestra un proceso de un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud, que incluye las siguientes etapas: S11: Se adquieren los datos de comportamiento de un usuario.

En la modalidad de la presente solicitud, los datos de comportamiento se obtienen a través de un sistema de monitoreo de red. El sistema de monitoreo de red monitorea y controla los ordenadores en una red para registrar las actividades de Internet (comportamiento de la red) realizadas por los usuarios en la red en una dimensión de tiempo. El sistema de monitoreo de red incluye hardware de monitoreo o software de monitoreo, y la red incluye una red de área local, una red de área metropolitana o una red de área amplia. Los datos de comportamiento anteriores se

almacenan en un medio de almacenamiento particular, y de acuerdo con un requisito de análisis real, los datos de comportamiento correspondientes se extraen del medio de almacenamiento para su análisis.

En esta descripción, se toma como ejemplo un sitio web de comercio electrónico para presentar las soluciones técnicas de la presente solicitud. Por lo tanto, el método para identificar un comportamiento de riesgo se utiliza para controlar si un comportamiento de la red del personal de servicio al cliente de un sitio web de comercio electrónico es riesgoso.

S12: Se selecciona un enlace de comportamiento específico de los datos de comportamiento. Un enlace de comportamiento se refiere a una combinación obtenida al ordenar secuencialmente múltiples comportamientos de acuerdo con los tiempos de ocurrencia. Como un enlace de comportamiento está más cerca de una intención de comportamiento real del usuario, se mejora la credibilidad de la identificación del riesgo de comportamiento de la red.

La Figura 2 muestra un proceso de selección de un enlace de comportamiento específico a partir de datos de comportamiento en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud. En la modalidad de la presente solicitud, la etapa S12 incluye específicamente las siguientes etapas:
S121: Los datos de fragmentos en un período de tiempo específico se seleccionan a partir de los datos de comportamiento.

Todavía tomando el ejemplo en esta descripción, se supone que un individuo de la red que necesita un análisis de riesgo de comportamiento es un usuario M. En este caso, los datos fragmentados del usuario M en un período de tiempo específico en un día particular D se extraen del medio de almacenamiento. Si el período de tiempo específico es de 15 minutos, por ejemplo, desde las 13:10 a las 13:25, los datos del fragmento se refieren a datos sobre comportamientos realizados por el usuario M en el período de 13:10 a 13:25 en ese día.

S122: Se adquieren los comportamientos incluidos en los datos del fragmento.

En el ejemplo anterior, se supone que en el período de 13:10 a 13:25 de ese día, los comportamientos realizados por el usuario M incluyen un comportamiento X, un comportamiento Y y un comportamiento Z.

S123: Los comportamientos se ordenan en orden cronológico de acuerdo con los tiempos de ocurrencia para obtener un enlace de comportamiento.

En el ejemplo anterior, la clasificación se lleva a cabo en orden cronológico de acuerdo con los tiempos de ocurrencia del comportamiento X, el comportamiento Y y el comportamiento Z, y un enlace de comportamiento específico obtenido G es: comportamiento X → comportamiento γ → comportamiento Z.

S13: Se determina un coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento.

En la modalidad de la presente solicitud, el coeficiente de riesgo es un valor numérico para expresar el grado de rareza de un enlace de comportamiento específico G. Generalmente, si el comportamiento de la red tiene una probabilidad relativamente alta de ocurrencia, es decir, el comportamiento de la red es relativamente común, indica que el comportamiento de la red es un comportamiento normal, por ejemplo, un comportamiento de visualización de información de la tienda por parte del personal de servicio al cliente. Si un comportamiento de la red tiene una probabilidad relativamente baja de ocurrencia, es decir, el comportamiento de la red solo ocurre en condiciones extremadamente raras, indica que el comportamiento de la red es un comportamiento de riesgo, por ejemplo, un comportamiento de consulta a los historiales de gastos de familiares y amigos por parte del personal de servicio al cliente. La presente solicitud evalúa, de acuerdo con el coeficiente de riesgo, si un comportamiento de la red es riesgoso.

En la modalidad de la presente solicitud, el coeficiente de riesgo anterior incluye uno o más de un coeficiente de riesgo a corto plazo a, un coeficiente de riesgo histórico b, y un coeficiente de riesgo del equipo c. Ciertamente, en otras modalidades de la presente solicitud, el coeficiente de riesgo analizado puede no estar limitado a los tres tipos anteriores. El coeficiente de riesgo a corto plazo a se refiere al grado de rareza al operar el enlace de comportamiento específico G por el usuario M en un primer período de tiempo t_1 (como por ejemplo un día). El coeficiente de riesgo histórico b se refiere a un grado de rareza al operar el enlace de comportamiento específico G por el usuario M en un período de tiempo total t2 de registro del usuario (un intervalo desde un tiempo de registro hasta un tiempo actual). Si se define que una población de usuarios a la que pertenece el usuario M es un grupo de usuarios y el grupo de usuarios incluye múltiples usuarios, el coeficiente de riesgo del equipo c se refiere a un grado de rareza al operar el enlace de comportamiento específico G por el grupo de usuarios al que el usuario M pertenece.

Los procesos para determinar los coeficientes de riesgo anteriores se describirán en detalle a continuación:

La Figura 3 muestra un proceso para determinar un coeficiente de riesgo a corto plazo en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud, que incluye específicamente las siguientes etapas:

S131: Se adquiere un número total de operaciones s_1 en las que el usuario M opera todos los enlaces de comportamiento en un primer período de tiempo t_1.

Todavía tomando el ejemplo en esta descripción, suponiendo que el primer período de tiempo t_1 es un día, el número de todos los enlaces de comportamiento (es decir, el número total de operaciones s_1) realizadas por el usuario M en ese día puede contarse en función de datos de comportamiento del usuario M en ese día. En la modalidad de la presente solicitud, el número total de operaciones s_1 del usuario M en ese día se cuenta usando un intervalo de tiempo para que un único enlace de comportamiento específico G dure como referencia. Específicamente, si t_G es de 15 minutos, el número total de operaciones $s_1 = 24 * 60/15 = 96$.

S132: Se adquiere el número de operaciones s_2 en el que el usuario M opera el enlace de comportamiento específico G en el primer período de tiempo t_1.

En el ejemplo anterior, el primer período de tiempo establecido t_1 es un día y, por lo tanto, se cuenta el número de veces (es decir, el número de operaciones s_2) que el usuario M opera el enlace de comportamiento específico G en ese día. Específicamente, si t_G es de 15 minutos, el día se divide en varios segmentos de tiempo de 15 minutos, y se evalúa secuencialmente si el enlace de comportamiento específico G ocurre en cada segmento de tiempo de 15 minutos; en caso afirmativo, el número de operaciones s_2 se incrementa en 1, y si no, el número de operaciones s_2 se incrementa en 0, hasta que se obtenga el número de operaciones s_2 en ese día.

S133: Se determina una relación entre el número total de operaciones s_1 y el número de operaciones s_2 para obtener el coeficiente de riesgo a corto plazo a.

En la modalidad de la presente solicitud, una fórmula para calcular el coeficiente de riesgo a corto plazo a es la siguiente:

$$a = s_1/s_2.$$

La Figura 4 muestra un proceso para determinar un coeficiente de riesgo histórico en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud, que incluye específicamente las siguientes etapas:

S134: Se adquiere una duración de tiempo total t_2 del usuario M desde un tiempo de registro hasta un tiempo actual t_a.

Todavía tomando el ejemplo en esta descripción, se supone que el tiempo de registro t_0 del usuario M en un sistema de servicio al cliente de un sitio web de comercio electrónico es 1ro de enero del 2014, y un tiempo actual t_a es 1ro de enero del 2015; en este caso, la duración total del tiempo t_2 es de 365 días.

S135: Se adquiere una duración de tiempo real t_3 en la que el usuario M opera el enlace de comportamiento específico G.

En la modalidad de la presente solicitud, en la etapa de adquirir una longitud de tiempo real t_3 en la que el usuario M opera el enlace de comportamiento específico G, el cálculo se lleva a cabo diariamente. En este caso, los datos de comportamiento del usuario M en los 365 días se dividen en 365 fragmentos de datos a diario, y se evalúa secuencialmente si el enlace de comportamiento específico G ocurre en los datos de fragmentos de cada día; en caso afirmativo, la duración de tiempo real t_3 se incrementa en 1; y si no, la duración de tiempo real t_3 se incrementa en 0, hasta que se obtiene el número real de días (es decir, la duración de tiempo real t_3) en que el usuario M opera el enlace de comportamiento específico G.

S136: El coeficiente de riesgo histórico b se determina de acuerdo con el tiempo total t_2 y el tiempo real t_3.

En la modalidad de la presente solicitud, para un usuario antiguo, como el usuario se registra en un momento anterior, el tiempo total de duración t_2 es relativamente largo (tal como 3 años). Suponiendo que la duración de tiempo real t_3 en el que el usuario antiguo opera el enlace de comportamiento específico G es de 2 días, finalmente se concluye que la probabilidad de operar el enlace de comportamiento específico G por el usuario antiguo en el período de tiempo total t_2 es relativamente baja. Sin embargo, para un nuevo usuario, como el usuario se registró recientemente, el tiempo total t_2 es relativamente corto (como 5 días). Suponiendo que la duración de tiempo real t_3 en el que el nuevo usuario opera el enlace de comportamiento específico G es de 2 días, finalmente se concluye que la probabilidad de que el nuevo usuario opere el enlace de comportamiento específico G en el período de tiempo total t_2 es relativamente alta. Como puede verse, la diferencia entre usuarios nuevos y antiguos puede afectar la

credibilidad del coeficiente de riesgo histórico b, y para suavizar la diferencia entre usuarios nuevos y antiguos, la etapa S136 incluye específicamente:

En primer lugar, la duración de tiempo total t_2 y la duración de tiempo real t_3 se suavizan para obtener una duración de tiempo total suavizado $t_{2\,k}$ y una duración de tiempo real suavizado t_{3k}. En la modalidad de la presente solicitud, el suavizado puede ser por procesamiento logarítmico, por procesamiento de módulo, por procesamiento de extracción de raíz o similares. Tomando la forma de procesamiento logarítmico como ejemplo, $t_{2k} = \lg t_2$; y $t_{3k} = \lg t_3$. Ciertamente, la base del procesamiento logarítmico no está limitada.

Luego, los cálculos se llevan a cabo en la longitud de tiempo total de suavizado t_{2k} y la longitud de tiempo total de suavizado t_{3k} para obtener el coeficiente de riesgo histórico b. En la modalidad de la presente solicitud, una fórmula para calcular el coeficiente de riesgo histórico b es la siguiente:

$$b=(1+ t_{3k})/(1+t_{2k})=(1+\lg t_3)/(1+\lg t_2).$$

La Figura 5 muestra un proceso para determinar un coeficiente de riesgo del equipo en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud, que incluye específicamente las siguientes etapas:

S137: Se determina un número total de usuarios n incluidos en un grupo de usuarios al que pertenece el usuario M.

Todavía tomando el ejemplo en esta descripción, se supone que el usuario M es personal de servicio al cliente de un sitio web de comercio electrónico. En este caso, un departamento al que pertenece el usuario M es el grupo de usuarios. Se supone que el número total de usuarios n incluidos en este departamento es 20.

S138: Un número real de usuarios m que han operado el enlace de comportamiento específico G en un segundo período de tiempo t_4 se adquiere en el grupo de usuarios.

En el ejemplo anterior, si el segundo período de tiempo t_4 es un día, la etapa S138 se utiliza para contar el número de personas que han operado el enlace de comportamiento específico G (es decir, el número real de usuarios m) en un día particular entre Las 20 personas del departamento al que pertenece el usuario M. Específicamente, los datos de comportamiento de las 20 personas en el departamento en ese día se obtienen por separado por adelantado, y luego se ve secuencialmente si los 20 usuarios han operado el enlace de comportamiento específico G ese día; en caso afirmativo, el número real de usuarios m se incrementa en 1; y si no, el número real de usuarios m se incrementa en 0, hasta que se obtenga el número real de usuarios m que han operado el enlace de comportamiento específico G ese día.

S139: El coeficiente de riesgo del equipo c se determina de acuerdo con el número total de usuarios n y el número real de usuarios m.

En la modalidad de la presente solicitud, si el grupo de usuarios que necesita ser analizado incluye un gran número de usuarios (por ejemplo, n = 1000), y si se obtiene el número real de usuarios m que han operado el comportamiento específico el enlace G en un día en particular es igual a 5, en este punto, indica que la probabilidad de que el enlace de comportamiento específico G haya sido operado en el grupo de usuarios es relativamente baja. Sin embargo, si el grupo de usuarios que necesita ser analizado incluye un pequeño número de usuarios (por ejemplo, n = 10), y si se obtiene que el número real de usuarios m que han operado el comportamiento específico vinculan a G en un día en particular es igual a 5, en este punto, indica que la probabilidad de que el enlace de comportamiento específico G haya sido operado en el grupo de usuarios es relativamente alta. Como puede verse, diferentes números de usuarios en diferentes grupos de usuarios pueden afectar la credibilidad del coeficiente de riesgo del equipo c, y para suavizar los diferentes números de usuarios en diferentes grupos de usuarios, la etapa S139 incluye específicamente:

En primer lugar, el número total de usuarios n y el número real de usuarios m se suavizan para obtener un número total suavizado de usuarios p y un número real suavizado de usuarios q. En la modalidad de la presente solicitud, el suavizado puede ser por procesamiento logarítmico, por procesamiento de módulo, por procesamiento de extracción de raíz o similares. Tomando la forma de procesamiento logarítmico como ejemplo, $p = \lg n$; y $q = \lg m$. Ciertamente, la base del procesamiento logarítmico no está limitada.

Luego, los cálculos se llevan a cabo sobre el número total suavizado de usuarios p y el número real suavizado de usuarios q para obtener el coeficiente de riesgo del equipo c. En la modalidad de la presente solicitud, una fórmula para calcular el coeficiente de riesgo del equipo c es la siguiente:

$$c=(1+p)/(1+q)=(1+\lg n)/(1+\lg m).$$

S14: Se evalúa, de acuerdo con el coeficiente de riesgo r, si el enlace de comportamiento específico G es riesgoso.

En la modalidad de la presente solicitud, una fórmula para calcular el coeficiente de riesgo r es la siguiente:

$$r = a \times b \times c.$$

Ciertamente, en otras modalidades de la presente solicitud, el coeficiente de riesgo r= a + b + c.

La Figura 6 muestra un proceso para evaluar si un enlace de comportamiento específico es riesgoso en un método para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud. En la modalidad de la presente solicitud, la etapa S14 incluye específicamente:

S141: Los coeficientes de riesgo r de los enlaces de comportamiento se ordenan en orden descendente.

Todavía tomando el ejemplo en esta descripción, se supone que los datos de comportamiento extraídos son todos los enlaces de comportamiento del usuario M en un día particular D. En los datos de comportamiento, hay 100 piezas de enlaces de comportamiento monitoreados; en este caso, los coeficientes de riesgo de r_1 a r_{100} de los 100 enlaces de comportamiento se determinan por separado de acuerdo con el método anterior, y luego los coeficientes de riesgo de r_1 a r_{100} se ordenan en orden descendente.

S142: Se evalúa si el coeficiente de riesgo r_G correspondiente al enlace de comportamiento específico G está en los rangos de riesgo.

En la modalidad de la presente solicitud, un rango más alto de un coeficiente de riesgo indica un mayor grado de rareza del enlace de comportamiento y un coeficiente de riesgo más alto del mismo. Suponiendo que un rango de riesgo preestablecido es el 3ro más alto, se evalúa si el coeficiente de riesgo r_G correspondiente al enlace de comportamiento específico G se clasifica en el 3ro más alto.

S143: En caso afirmativo, se determina que el enlace de comportamiento específico G es riesgoso.

Si el coeficiente de riesgo r_G correspondiente al enlace de comportamiento específico G se clasifica en el 3ro más alto, indica que el enlace de comportamiento específico G es riesgoso y, posteriormente, el enlace de comportamiento específico G puede publicarse como un comportamiento de riesgo para informar al personal de servicio al cliente de un sitio web de comercio electrónico que no opere el enlace de comportamiento.

S144: Si no, se considera que el enlace de comportamiento específico G no es riesgoso.

Si el coeficiente de riesgo r_G correspondiente al enlace de comportamiento específico G no se clasifica en el 3ro más alto, indica que el enlace de comportamiento específico G no es riesgoso.

La Figura 7 es un diagrama estructural esquemático de un aparato para identificar un comportamiento de riesgo de acuerdo con una modalidad de la presente solicitud. Basado en la misma idea, el aparato incluye:

un módulo de adquisición 10 configurado para adquirir datos de comportamiento de un usuario;

un módulo de selección 20 configurado para seleccionar un enlace de comportamiento específico de los datos de comportamiento;

un módulo de determinación 30 configurado para determinar un coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento; y

un módulo de evaluación 40 configurado para evaluar, de acuerdo con el coeficiente de riesgo, si el enlace de comportamiento específico es riesgoso.

En la modalidad de la presente solicitud, el módulo de selección 20 está configurado específicamente para:

seleccionar, a partir de los datos de comportamiento, datos del fragmento en un período de tiempo específico;

adquirir comportamientos incluidos en los datos del fragmento; y

ordenar los comportamientos en orden cronológico de acuerdo con los tiempos de ocurrencia para obtener el enlace de comportamiento específico.

En la modalidad de la presente solicitud, el coeficiente de riesgo incluye uno o más de un coeficiente de riesgo a corto plazo, un coeficiente de riesgo histórico y un coeficiente de riesgo del equipo.

En la modalidad de la presente solicitud, el módulo de determinación 30 incluye un módulo de determinación de riesgo a corto plazo 31 configurado para:

adquirir un número total de operaciones en las que el usuario opera todos los enlaces de comportamiento en un primer período de tiempo;

adquirir el número de operaciones que el usuario opera el enlace de comportamiento específico en el primer período de tiempo; y

determinar una relación entre el número total de operaciones y el número de operaciones para obtener el coeficiente de riesgo a corto plazo.

En la modalidad de la presente solicitud, el módulo de determinación 30 incluye un módulo de determinación de riesgo histórico 32 configurado para:

adquirir una duración de tiempo total para el usuario desde un tiempo de registro hasta un tiempo actual;

adquirir una duración de tiempo real para que el usuario opere el enlace de comportamiento específico; y

determinar el coeficiente de riesgo histórico de acuerdo con el tiempo total y el tiempo real.

En la modalidad de la presente solicitud, el módulo de determinación 30 incluye un módulo de determinación de riesgo de equipo 33 configurado para:

determinar un número total de usuarios incluidos en un grupo de usuarios al que pertenece el usuario;

adquirir, en el grupo de usuarios, un número real de usuarios que han operado el enlace de comportamiento específico en un segundo período de tiempo; y

determinar el coeficiente de riesgo del equipo de acuerdo con el número total de usuarios y el número real de usuarios.

En la modalidad de la presente solicitud, el módulo de determinación de riesgo histórico 32 incluye una primera unidad de suavizado configurada para:

suavizar el tiempo total y el tiempo real para obtener un tiempo total suavizado y un tiempo real suavizado; y

realizar cálculos sobre la duración del tiempo real suavizado y la duración del tiempo total suavizado para obtener el coeficiente de riesgo histórico.

En la modalidad de la presente solicitud, el módulo de determinación de riesgo de equipo 33 incluye una segunda unidad de suavizado configurada para:

suavizar el número total de usuarios y el número real de usuarios para obtener un número total suavizado de usuarios y un número real suavizado de usuarios; y

realizar cálculos sobre el número total suavizado de usuarios y el número real suavizado de usuarios para obtener el coeficiente de riesgo del equipo.

En la modalidad de la presente solicitud, el módulo de determinación 30 está configurado específicamente para: multiplicar o sumar el coeficiente de riesgo a corto plazo, el coeficiente de riesgo histórico y el coeficiente de riesgo del equipo para obtener el coeficiente de riesgo.

En la modalidad de la presente solicitud, el módulo de evaluación 40 está configurado específicamente para:

ordenar los coeficientes de riesgo de los enlaces de comportamiento en orden descendente;

evaluar si el coeficiente de riesgo correspondiente al enlace de comportamiento específico está en las filas de riesgo; y

en caso afirmativo, evaluar que el enlace de comportamiento específico es riesgoso; y si no, evaluar que el enlace de comportamiento específico no es riesgoso.

El método y el aparato proporcionados en las modalidades de la presente solicitud adquieren datos de comportamiento de un usuario, seleccionan un enlace de comportamiento específico a partir de los datos de comportamiento, determinan un coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento mediante cálculos, y finalmente, determina, de acuerdo con el coeficiente de riesgo, si el enlace de comportamiento específico es riesgoso. En comparación con la manera del motor de reglas, en el proceso anterior,

no es necesario remediar una vulnerabilidad de la regla manualmente, mejorando así la eficiencia de la identificación del riesgo de comportamiento. Además, el proceso anterior evita la desventaja de que el motor de reglas consuma recursos informáticos adicionales, aliviando así la carga de un sistema informático.

En la modalidad de la presente solicitud, tres factores: a corto plazo (tal como un día en particular), historial (desde una hora de registro hasta la hora actual) y equipo (un grupo de usuarios al que pertenece el usuario), se consideran de manera integral para analizar si el comportamiento de un usuario es riesgoso, lo que reduce el impacto de algunas transiciones repentinas de factores (tal como el ajuste de la orientación del servicio del equipo o la transferencia de trabajo del usuario) en el enlace de comportamiento del usuario, mejorando así la precisión y la credibilidad de identificación de conductas de riesgo.

Vale la pena mencionar que, el aparato para identificar un comportamiento de riesgo descrito en esta descripción se genera de acuerdo con la misma idea basada en el método para identificar un comportamiento de riesgo. Por lo tanto, el método para identificar un comportamiento de riesgo puede continuar utilizando todas las características técnicas del aparato anterior para identificar un comportamiento de riesgo. Los detalles no se describen aquí nuevamente.

Debe observarse adicionalmente que, las fórmulas para calcular los coeficientes de riesgo en la presente solicitud no se limitan a las modalidades divulgadas. Por ejemplo, en otras modalidades, el coeficiente de riesgo a corto plazo a $= s_2/s_1$; el coeficiente de riesgo histórico $b=(1 + lgt_2)/(1 + lgt_3)$; y el coeficiente de riesgo del equipo $c=(1 + lgm)/(1 + lgn)$. En consecuencia, durante la evaluación posterior de si el enlace de comportamiento es riesgoso, los coeficientes de riesgo de los enlaces de comportamiento se ordenan en orden ascendente para evaluar si el coeficiente de riesgo correspondiente al enlace de comportamiento específico está en los rangos de riesgo.

Los expertos en la técnica deben comprender que las modalidades de la presente invención pueden proporcionarse como un método, un sistema o un producto de programa informático. Por lo tanto, la presente invención puede implementarse en forma de una modalidad de hardware completa, una modalidad de software completa o una modalidad que combina software y hardware. Además, la presente invención puede emplear la forma de un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por ordenador (que incluyen, pero no se limitan a, una memoria de disco magnético, un CD-ROM, una memoria óptica y similares) que incluyen código de programa utilizable por ordenador.

La presente invención se describe con referencia a diagramas de flujo y/o diagramas de bloques del método, dispositivo (sistema) y producto de programa informático de acuerdo con las modalidades de la presente invención. Debe entenderse que las instrucciones del programa informático pueden usarse para implementar cada proceso y/o bloque en los diagramas de flujo y/o diagramas de bloque y una combinación de un proceso y/o un bloque en los diagramas de flujo y/o los diagramas de bloque. Estas instrucciones de programa informático pueden proporcionarse para un ordenador de propósito general, un ordenador de propósito especial, un procesador incorporado o un procesador de otro dispositivo de procesamiento de datos programable para generar una máquina, de modo que las instrucciones ejecutadas por un ordenador o un procesador de otro dispositivo de procesamiento de datos programable genera un aparato para implementar una función específica en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

Estas instrucciones del programa informático también pueden almacenarse en una memoria legible por ordenador que puede indicarle al ordenador u otro dispositivo de procesamiento de datos programable que trabaje de una manera particular, de modo que las instrucciones almacenadas en la memoria legible por ordenador generen un artículo de fabricación que incluya un aparato de instrucción. El aparato de instrucción implementa una función especificada en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

Estas instrucciones del programa informático también pueden cargarse en un ordenador u otro dispositivo de procesamiento de datos programable, de modo que se realicen una serie de etapas operativas en el ordenador u otro dispositivo programable, generando así el procesamiento implementado por ordenador. Por lo tanto, las instrucciones ejecutadas en el ordenador u otro dispositivo programable proporcionan etapas para implementar una función específica en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

En una configuración típica, el dispositivo informático incluye uno o más procesadores (CPU), una interfaz de entrada/salida, una interfaz de red y una memoria.

La memoria puede incluir una memoria volátil, una memoria de acceso aleatorio (RAM) y/o una memoria no volátil o similar en un medio legible por ordenador, por ejemplo, una memoria de solo lectura (ROM) o una RAM flash. La memoria es un ejemplo del medio legible por ordenador.

El medio legible por ordenador incluye medios no volátiles o volátiles, y medios móviles o no móviles, y puede implementar el almacenamiento de información mediante cualquier método o tecnología. La información puede ser

una instrucción legible por ordenador, una estructura de datos y un módulo de un programa u otros datos. Un medio de almacenamiento de un ordenador incluye, por ejemplo, pero no se limita a, una memoria de cambio de fase (PRAM), una memoria estática de acceso aleatorio (SRAM), una memoria dinámica de acceso aleatorio (DRAM), otros tipos de memorias de acceso aleatorio (RAM), una memoria de solo lectura (ROM), una memoria de solo lectura programable y borrable eléctricamente (EEPROM), una memoria flash u otras tecnologías de memoria, una memoria de solo lectura de disco compacto (CD-ROM), un disco versátil digital (DVD) u otros almacenamientos ópticos, una cinta de casete, una cinta magnética/almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio sin transmisión, y pueden usarse para almacenar información accesible al dispositivo informático. Según la definición en este texto, el medio legible por ordenador no incluye medios transitorios, como una señal de datos modulada y una portadora.

Debe observarse además que, los términos "incluir", "comprender" o cualquier variante de los mismos están destinados a cubrir una inclusión no exclusiva, de modo que un proceso, un método, un producto o un dispositivo que incluya una serie de elementos no solo incluye dichos elementos, sino que también incluye otros elementos no especificados expresamente, o puede incluir elementos inherentes del proceso, método, producto o dispositivo. Sin más restricciones, un elemento limitado por la frase "incluir un/unos..." no excluye otros mismos elementos existentes en el proceso, método, producto o dispositivo que incluye el elemento.

Los expertos en la técnica deben comprender que las modalidades de la presente solicitud pueden proporcionarse como un método, un sistema o un producto de programa informático. Por lo tanto, la presente solicitud puede implementarse en forma de una modalidad de hardware completa, una modalidad de software completa o una modalidad que combina software y hardware. Además, la presente solicitud puede emplear la forma de un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por ordenador (que incluyen, pero no se limitan a, una memoria de disco magnético, un CD-ROM, una memoria óptica y similares) que incluyen código de programa utilizable por ordenador.

Lo anterior son meramente las modalidades de la presente solicitud, que no se usan para limitar la presente solicitud. Para los expertos en la técnica, la presente solicitud puede tener varios cambios y alteraciones. Las modificaciones, reemplazos equivalentes y mejoras realizadas a la presente solicitud deben incluirse en el alcance de las reivindicaciones de la presente solicitud.

REIVINDICACIONES

1. Un método para identificar comportamientos de riesgo dentro de una red informática, el método que comprende:

adquirir datos de comportamiento del usuario, los datos de comportamiento del usuario que comprenden un registro de actividades de red realizadas por el usuario en una red en un período de tiempo (S11);
seleccionar un enlace de comportamiento específico de los datos de comportamiento organizando secuencialmente las actividades de la red dentro de la red informática de acuerdo con los tiempos de ocurrencia (S12);
determinar un coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento, el coeficiente de riesgo comprende un valor numérico para expresar un grado de rareza del enlace de comportamiento específico (S13), que comprende determinar un coeficiente de riesgo histórico mediante:

la adquisición de una duración de tiempo total para el usuario desde un tiempo de registro hasta un tiempo actual;
la adquisición de una duración de tiempo real para que el usuario opere el enlace de comportamiento específico; y
el suavizado de la duración del tiempo total y la duración del tiempo real utilizando el procesamiento logarítmico para obtener una duración del tiempo total suavizado y una duración del tiempo real suavizado, en donde la duración del tiempo total suavizado comprende un logaritmo de la duración del tiempo total y la duración del tiempo real suavizado comprende un logaritmo de la duración del tiempo real; y
el cálculo del coeficiente de riesgo histórico dividiendo i) la duración del tiempo real suavizado más uno, por ii) la duración del tiempo total suavizado más uno; y

la determinación, de acuerdo con el coeficiente de riesgo, si el enlace de comportamiento específico es riesgoso (S14).

2. El método de acuerdo con la reivindicación 1, en donde la selección del enlace de comportamiento específico de los datos de comportamiento comprende específicamente:

seleccionar, a partir de los datos de comportamiento, datos del fragmento en un período de tiempo específico (S121);
adquirir comportamientos comprendidos en los datos del fragmento (S122); y
ordenar los comportamientos en orden cronológico de acuerdo con los tiempos de ocurrencia para obtener el enlace de comportamiento específico (S123).

3. El método de acuerdo con la reivindicación 1, en donde el coeficiente de riesgo comprende uno o más de un coeficiente de riesgo a corto plazo, un coeficiente de riesgo histórico, y un coeficiente de riesgo del equipo.

4. El método de acuerdo con la reivindicación 3, en donde la determinación del coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento comprende específicamente determinar un coeficiente de riesgo a corto plazo mediante:

la adquisición de un número total de operaciones en las que el usuario opera todos los enlaces de comportamiento en un primer período de tiempo (S131);
la adquisición del número de operaciones en las que el usuario opera el enlace de comportamiento específico en el primer período de tiempo (S132); y
la determinación de una relación entre el número total de operaciones y el número de operaciones para obtener el coeficiente de riesgo a corto plazo (S133).

5. El método de acuerdo con la reivindicación 3, en donde la determinación de un coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento comprende específicamente la determinación de un coeficiente de riesgo del equipo mediante:

la determinación de un número total de usuarios comprendidos en un grupo de usuarios al que pertenece el usuario (S137);
la adquisición, en el grupo de usuarios, de un número real de usuarios que han operado el enlace de comportamiento específico en un segundo período de tiempo (S138); y
la determinación del coeficiente de riesgo del equipo de acuerdo con el número total de usuarios y el número real de usuarios (S139).

6. El método de acuerdo con la reivindicación 5, en donde la determinación del coeficiente de riesgo del equipo de acuerdo con el número total de usuarios y el número real de usuarios comprende específicamente:

suavizar el número total de usuarios y el número real de usuarios para obtener un número total suavizado de usuarios y un número real suavizado de usuarios; y
realizar cálculos sobre el número total suavizado de usuarios y el número real suavizado de usuarios para obtener el coeficiente de riesgo del equipo.

7. El método de acuerdo con la reivindicación 6, en donde el suavizado comprende procesamiento logarítmico, procesamiento de módulo o procesamiento de extracción de raíz.

8. El método de acuerdo con la reivindicación 3, en donde la determinación del coeficiente de riesgo del enlace de comportamiento específico en los datos de comportamiento comprende específicamente:

multiplicar o sumar el coeficiente de riesgo a corto plazo, el coeficiente de riesgo histórico y el coeficiente de riesgo del equipo para obtener el coeficiente de riesgo.

9. El método de acuerdo con la reivindicación 1, en donde la determinación, de acuerdo con el coeficiente de riesgo, de si el comportamiento objetivo es riesgoso comprende específicamente:

ordenar los coeficientes de riesgo de los enlaces de comportamiento en orden descendente (S141);
determinar si el coeficiente de riesgo correspondiente al enlace de comportamiento específico está en los rangos de riesgo (S142); y
en caso afirmativo, determinar que el enlace de comportamiento específico es riesgoso; y
en caso negativo, determinar que el enlace de comportamiento específico no es riesgoso (S143).

10. El método de acuerdo con la reivindicación 1, en donde el logaritmo de la duración del tiempo total y el logaritmo de la duración del tiempo real comprenden logaritmos de base arbitraria.

11. Un aparato para identificar un comportamiento de riesgo, el aparato que comprende una pluralidad de módulos configurados para llevar a cabo el método de una cualquiera de las reivindicaciones de la 1 a la 10.

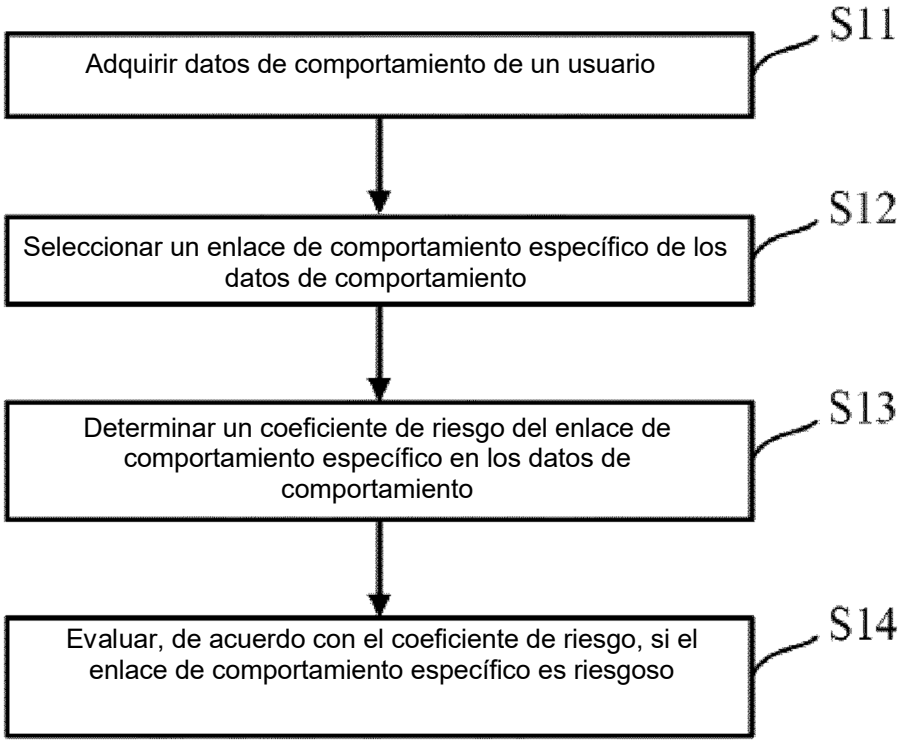

Figura 1

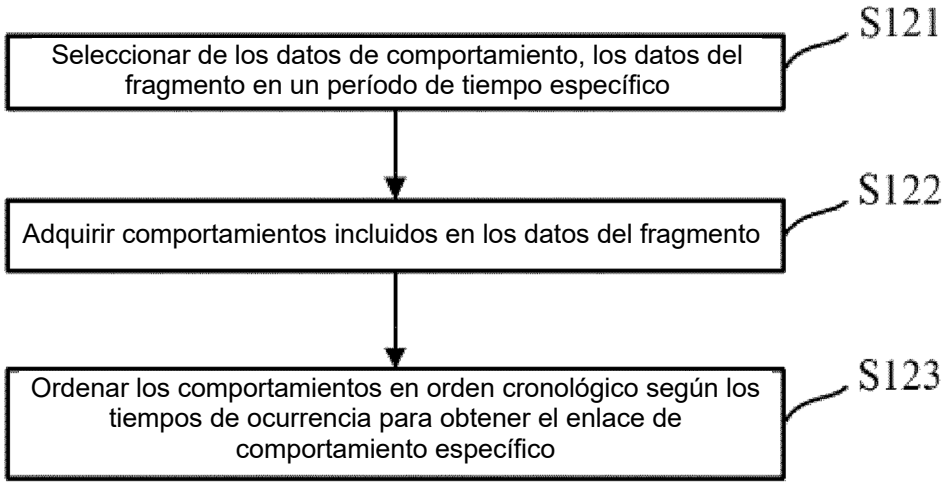

Figura 2

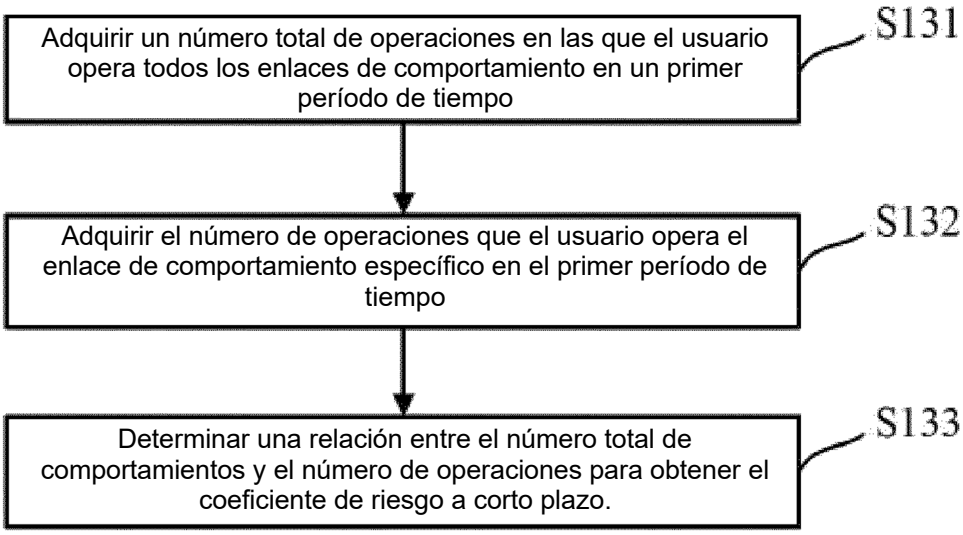

Figura 3

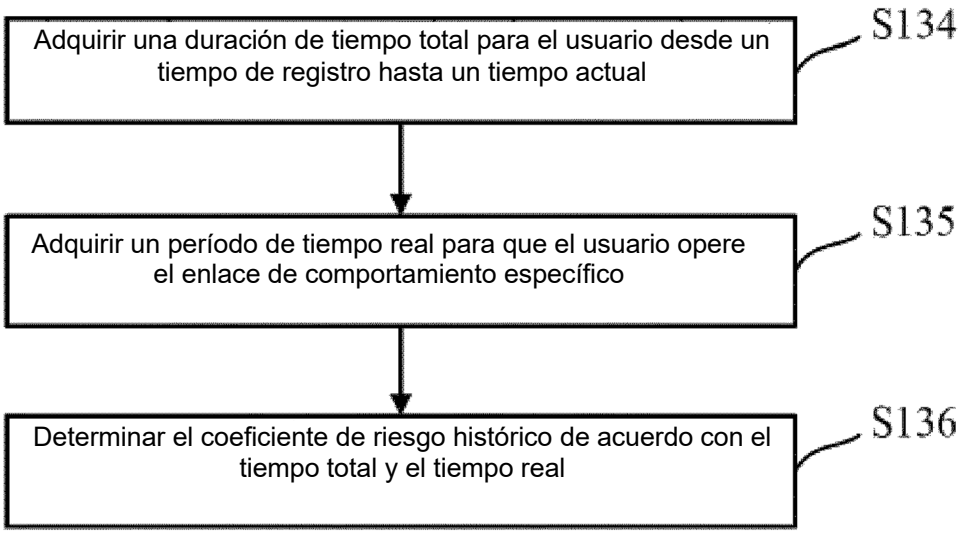

Figura 4

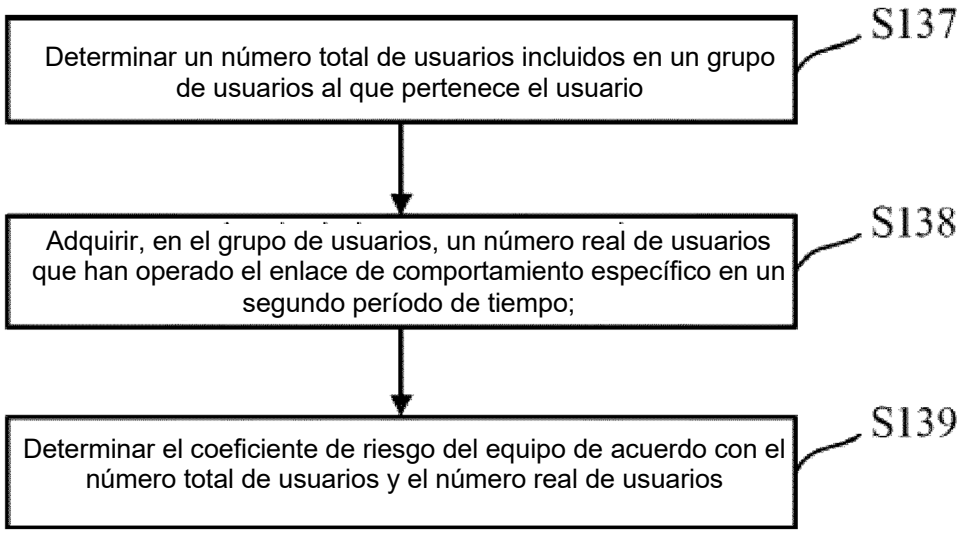

Figura 5

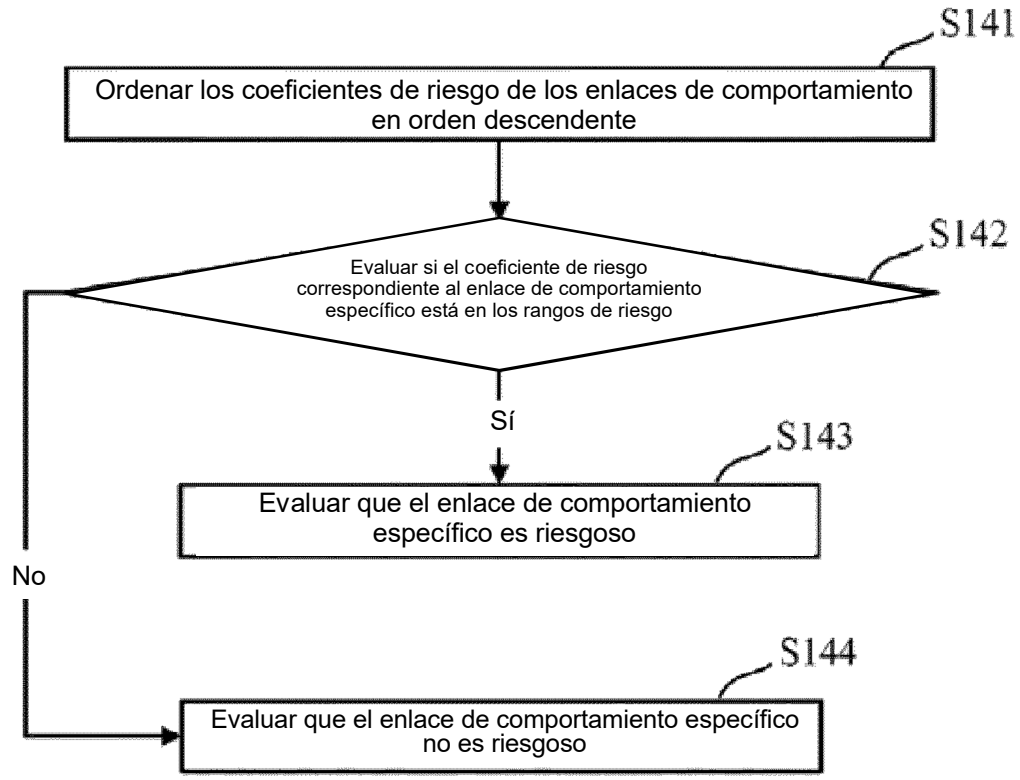

Figura 6

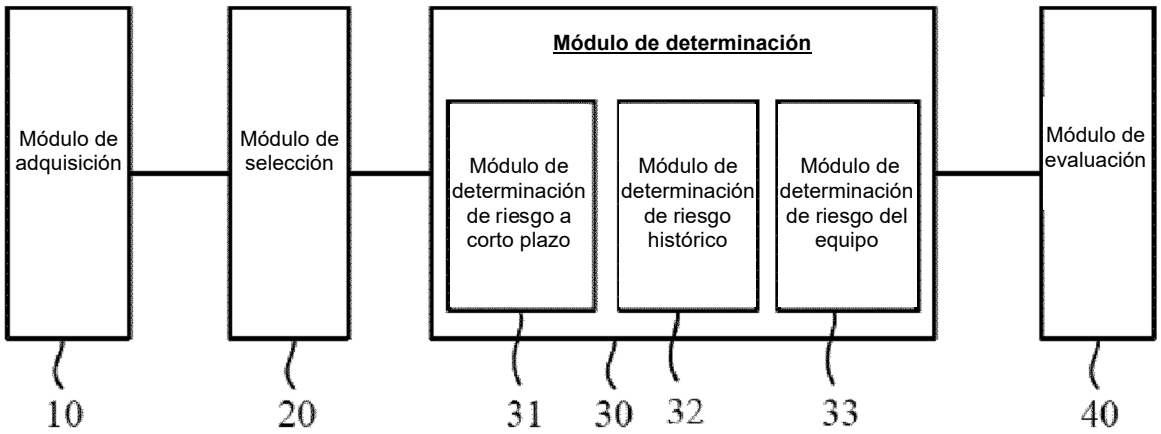
Módulo de adquisición
Módulo de selección
Módulo de determinación
Módulo de determinación de riesgo a corto plazo
Módulo de determinación de riesgo histórico
Módulo de determinación de riesgo del equipo
Módulo de evaluación
10
20
31
30
32
33
40

Figura 7