



OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA



11) Número de publicación: 2 703 861

61 Int. Cl.:

H04L 29/06 (2006.01) H04L 12/58 (2006.01) G06F 17/30 (2006.01) G06F 21/64 (2013.01) G06F 21/56 (2013.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

- Fecha de presentación y número de la solicitud europea: 08.01.2016 E 16150573 (0)
 Fecha y número de publicación de la concesión europea: 12.12.2018 EP 3190767
 - (54) Título: Técnica para detectar mensajes electrónicos maliciosos
 - (45) Fecha de publicación y mención en BOPI de la traducción de la patente: 12.03.2019

73) Titular/es:

RETARUS GMBH (100.0%) Aschauer Strasse 30 81549 München, DE

(72) Inventor/es:

HAGER, MARTIN y GRAUVOGL, MICHAEL

(74) Agente/Representante:

UNGRÍA LÓPEZ, Javier

DESCRIPCIÓN

Técnica para detectar mensajes electrónicos maliciosos

5 Campo técnico

La presente divulgación se refiere, en general, a los aspectos de seguridad en la tecnología de la información. En particular, la divulgación se refiere a una técnica para detectar mensajes electrónicos maliciosos que se transmiten desde el al menos un dispositivo de envío de mensajes a el al menos un dispositivo de recepción de mensajes.

Antecedentes

10

15

20

25

30

35

40

45

50

55

60

Los mensajes electrónicos, tales como los mensajes de correo electrónico (o e-mails), los mensajes instantáneos, los faxes etc., son el método de elección para el intercambio de información entre sí. Como el uso de mensajes electrónicos se ha vuelto muy popular, no es sorprendente que los mensajes electrónicos, tal como los correos electrónicos, se usen con frecuencia para la proliferación de software malicioso. En este contexto, el término "malware" o "software malicioso" se refiere a cualquier software o parte de software que se usa para alterar las operaciones informáticas, recopilar información confidencial u obtener acceso a sistemas informáticos privados o corporativos. El software malicioso incorporado o adjunto en los mensajes electrónicos, y distribuido a través de mensajes electrónicos puede incluir, entre otros, virus, gusanos, troyanos, ransomware, scareware, adware y otros programas maliciosos.

Con el fin de impedir la proliferación de software malicioso en una red de comunicaciones, existen en el mercado numerosas soluciones antisoftware malicioso que siguen diferentes estrategias de protección o combate. Por ejemplo, hay soluciones antisoftware malicioso disponibles que están diseñadas para proporcionar una protección en tiempo real contra la instalación de software malicioso en un dispositivo informático escaneando todos los datos entrantes de la red en busca de software malicioso y bloqueando inmediatamente cualquier amenaza detectada. También existen soluciones antisoftware malicioso disponibles que están diseñadas para detectar y eliminar el software malicioso que ya está instalado en un dispositivo informático.

Por otra parte, con el fin de impedir de manera eficaz la proliferación de software malicioso en las redes de comunicaciones, se proporcionan soluciones anti-software malicioso adecuadas para los servidores de mensajería remota de las redes de comunicaciones que están diseñadas para encaminar los mensajes procedentes de uno o más dispositivos de envío de mensajes a uno o más dispositivos de recepción de mensajes. Estas soluciones antisoftware malicioso se realizan en general en la forma de módulos de software y hardware implementados en los servidores de mensajería que se diseñan para realizar una comprobación antisoftware malicioso de cada mensaje. Es decir, los servidores de mensajería escanean los mensajes en busca de software malicioso y solo los mensajes limpios (es decir, los mensajes no maliciosos) se encaminan a los dispositivos de recepción de mensajes, mientras que los servidores de mensajería filtran los mensajes maliciosos, incluso antes de que el mensaje malicioso pueda llegar alcanzar un dispositivo de recepción.

Tales soluciones antisoftware malicioso funcionan en general sobre la base de una comparación del contenido del mensaje con firmas de virus conocidas. Una firma de virus es un algoritmo o un hash estático (es decir, un valor numérico de una parte del código exclusivo del virus) que puede usarse como huella digital para un virus específico. Una técnica de detección de antisoftware malicioso de este tipo es muy eficaz, pero tiene el inconveniente de que solo los virus ya conocidos por el software antisoftware malicioso pueden filtrarse de manera eficaz. Incluso en el caso de que las firmas de virus conocidas del módulo antisoftware malicioso se actualicen regularmente, existe cierto riesgo de que el software malicioso de última generación (denominado "software malicioso de día cero") permanezca sin detectarse. Por lo tanto, la detección de antisoftware malicioso conocida para los sistemas de comunicaciones o las redes de comunicaciones tiene el riesgo de que el software malicioso de última generación permanezca sin detectarse durante un período de tiempo más prolongado. En consecuencia, los mensajes maliciosos que se consideran erróneamente limpios pueden encaminarse a los dispositivos de recepción de mensajes en lugar de filtrarlos.

A partir del documento US 7.539.871 B1 se conoce una técnica para identificar la propagación de mensajes maliciosos. De acuerdo con esta técnica, las huellas digitales se generan para cada mensaje entrante y se almacenan en un servidor de correo como un conjunto de huellas digitales. El conjunto de huellas digitales almacenadas se compara con un conjunto de huellas digitales de comparación predeterminadas proporcionadas por un proveedor de AV. Es decir, se realiza una coincidencia entre el conjunto de huellas digitales almacenadas y el conjunto predeterminado de huellas digitales de comparación y, si se encuentra una coincidencia, los mensajes correspondientes a las huellas digitales coincidentes se identifican como maliciosos. Ya que las actualizaciones periódicas del conjunto de huellas digitales de comparación conducen a un refinamiento continuo de las huellas digitales de comparación, los mensajes que se han clasificado como no maliciosos en el pasado pueden identificarse como maliciosos más adelante basándose en el conjunto actualizado de huellas digitales de comparación.

Por consiguiente, hay una necesidad de una técnica de detección de software malicioso eficaz en las redes de comunicaciones que supere las desventajas técnicas mencionadas anteriormente.

Sumario

5

10

15

20

25

Para superar el problema técnico identificado anteriormente, de acuerdo con un primer aspecto, se proporciona un método de detección de mensajes electrónicos maliciosos transmitidos desde el al menos un dispositivo de envío de mensajes a el al menos un dispositivo de recepción mensajes. El método comprende las etapas de generar al menos una firma para un mensaje electrónico a transmitir desde el al menos un dispositivo de envío de mensajes a el al menos un dispositivo de recepción de mensajes; almacenar la al menos una firma generada en una unidad de almacenamiento de datos; determinar si el mensaje electrónico es malicioso; si se determina que el mensaje electrónico es malicioso, determinar sobre la base de la al menos una firma generada para el mensaje malicioso determinado, si los mensajes electrónicos comparables al mensaje malicioso determinado se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes en el pasado, donde la etapa de determinación comprende comparar la al menos una firma del mensaje malicioso determinado con las firmas ya almacenadas en la unidad de almacenamiento de datos que están asociadas con los mensajes transmitidos anteriormente; y si se determina que los mensajes electrónicos comparables con el mensaje malicioso determinado se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes en el pasado, notificar a el al menos un dispositivo de recepción de mensajes sobre una posible amenaza.

En la presente divulgación la expresión "mensaje electrónico" (o abreviado "mensaje"), puede tener que interpretarse en sentido amplio. Como "mensaje electrónico", o "mensaje", cualquier elemento de datos digitales o cualquier parte de datos digitales pueden significar que contiene un mensaje en la forma de símbolos, caracteres alfabéticos y/o numéricos, elementos gráficos, etc., y que puede usarse con el fin de embeber o adjuntar un software malicioso. Por ejemplo, como "mensaje electrónico", o "mensaje" puede significar un mensaje de correo electrónico, un mensaje instantáneo o cualquier otro tipo de mensaje electrónico.

El método puede realizarse en un servidor de mensajería. El servidor de mensajería puede estar dispuesto en una red de comunicaciones a través de la que pueden transmitirse mensajes desde el al menos un dispositivo de envío de mensajes a el al menos un dispositivo de recepción de mensajes. El servidor de mensajería puede estar diseñado para encaminar mensajes desde el al menos un dispositivo de envío de mensajes a el al menos un dispositivo de recepción de mensajes para el que están destinados los mensajes. Para este fin, el servidor de mensajería puede estar en comunicación (continua) con el al menos un dispositivo de envío de mensajes y el al menos un dispositivo de recepción de mensajes. El servidor puede implementarse como un único dispositivo informático o como un sistema informático que comprende dispositivos informáticos distribuidos que están configurados para realizar el método descrito anteriormente.

El al menos un dispositivo de envío de mensajes puede ser cualquier dispositivo configurado para enviar mensajes electrónicos, tales como un teléfono inteligente, una tableta, un ordenador personal, y/o cualquier otro dispositivo informático privado o corporativo. De manera similar, el al menos un dispositivo de recepción de mensajes puede ser cualquier dispositivo configurado para enviar mensajes electrónicos, tal como un teléfono inteligente, una tableta, un ordenador personal y/o cualquier otro dispositivo informático privado o corporativo.

Las etapas de generar al menos una firma, almacenar la al menos una firma en la unidad de almacenamiento de datos, y determinar si el mensaje es malicioso pueden repetirse para cada nuevo mensaje (es decir, para cada mensaje recibido actualmente en un flujo de mensajes continuo) a transmitir a el al menos un dispositivo de recepción de mensajes. Es decir, el servidor de mensajería puede repetir, para cada mensaje recibido actualmente desde el al menos un dispositivo de envío de mensajes, la generación de firmas y las etapas de almacenamiento, así como la etapa de determinación de antisoftware malicioso. Por lo tanto, con el tiempo puede recopilarse una gran cantidad de firmas, donde cada firma puede estar asociada con un mensaje específico que se ha recibido y distribuido. Por consiguiente, la unidad de almacenamiento de datos puede comprender un gran número de firmas de mensajes que pueden asociarse con los mensajes recibidos y distribuidos a el al menos un dispositivo de recepción de mensajes en el pasado.

55

60

Por otra parte, la etapa mencionada anteriormente de determinar si los mensajes comparables al mensaje malicioso determinado se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes en el pasado (es decir, la etapa d en la reivindicación 1), así como la etapa de notificar a el al menos un dispositivo de recepción de mensajes sobre una posible amenaza (es decir, la etapa e en la reivindicación 1) puede repetirse para los mensajes recibidos recientemente, para los que se ha identificado por primera vez un comportamiento malicioso o contenido malicioso específico. Para los mensajes limpios recibidos recientemente (es decir, los mensajes para los que no se ha determinado ningún un comportamiento malicioso o un contenido malicioso) estas etapas pueden omitirse.

La etapa de determinar si los mensajes comparables al mensaje malicioso determinado se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes en el pasado (es decir, la

etapa d en la reivindicación 1) puede comprender además la siguiente subetapa: seleccionar esas firmas de los mensajes transmitidos anteriormente desde la unidad de almacenamiento de datos, que se ha descubierto que son comparables a la al menos una firma del mensaje malicioso determinado. Las firmas almacenadas en la unidad de almacenamiento de datos pueden considerarse como comparables a la al menos una firma del mensaje malicioso determinado, si las firmas son idénticas o muy similares entre sí. Si las dos firmas son muy similares, puede determinarse estimando si un grado de similitud entre las dos firmas supera un valor de umbral predeterminado.

La etapa de generar al menos una firma puede comprender además generar un identificador de mensaje (ID) para cada mensaje electrónico a transmitir a el al menos un dispositivo de recepción. Además, la etapa de almacenar la al menos una firma generada en una unidad de almacenamiento de datos puede comprender además almacenar, para cada mensaje a transmitir, el ID de mensaje generado junto con la al menos una firma generada en la unidad de almacenamiento de datos. Al asociar las firmas generadas con los ID correspondientes, cada firma en la unidad de almacenamiento de datos puede asociarse con un mensaje específico transmitido a un dispositivo de recepción de mensajes específico más adelante.

10

15

20

25

30

45

55

60

65

De acuerdo con una implementación de la etapa de generar al menos una firma puede comprender además generar, para cada mensaje a transmitir, la información de mensaje específico adecuada para clasificar el mensaje transmitido más adelante. La información de mensaje específico generada puede almacenarse junto con la al menos una firma generada (y el ID generado) en la unidad de almacenamiento de datos. La información de mensaje específica generada puede comprender al menos una de la siguiente información: tiempo de transmisión de mensaje, información de remitente de mensaje, información de destino de mensaje e información sobre los adjuntos de mensaje. Por consiguiente, la información de mensaje específica generada puede usarse para identificar cuándo se ha transmitido un mensaje específico, desde qué dispositivo de envío de mensajes se ha recibido el mensaje específico.

Para cada mensaje a transmitir puede generarse la al menos una firma sobre la base de al menos una de la siguiente información: al menos una propiedad de adjunto de mensaje; y una información de URL del dispositivo de envío de mensajes de transmisión. Como una propiedad de adjunto de mensaje puede usarse al menos uno de entre un nombre del adjunto, el tipo de archivo del adjunto (por ejemplo, es el adjunto de un archivo de texto, archivo de música, archivo de imagen, etc.), y la información de extensión de archivo (por ejemplo, el archivo es un archivo ejecutable). Las propiedades de adjuntos pueden proporcionarse como un valor hash. La información de URL puede comprender una URL completa o una parte de la URL que puede asociarse con un dispositivo de envío de mensajes específico.

La etapa de determinar si el mensaje electrónico es malicioso puede comprender comparar el mensaje electrónico con firmas de virus conocidos; y clasificar el mensaje electrónico como malicioso, si el mensaje coincide suficientemente con una de las firmas de virus conocidas. Para la comparación, puede usarse una lista negra (BL) que contiene firmas de virus conocidas. La BL con firmas de virus conocidas puede proporcionarse y actualizarse regularmente por un proveedor de software antivirus. Además, o como alternativa, pueden usarse métodos heurísticos para detectar los mensajes electrónicos maliciosos.

De acuerdo con una implementación, la etapa de notificar a el al menos un dispositivo de recepción de mensajes puede comprender proporcionar información a el al menos un dispositivo de recepción de mensajes que indica que un mensaje específico transmitido en el pasado es malicioso (es decir, puede comprender una amenaza).

El método puede comprender además las etapas de transmitir el mensaje (direccionar) a el al menos un dispositivo de recepción de mensajes, si el mensaje se ha clasificado como no malicioso; y bloquear el mensaje, si el mensaje se ha clasificado como malicioso.

De acuerdo con otro aspecto, se proporciona un producto de programa informático con partes de código de programa para realizar el método descrito anteriormente cuando el producto de programa informático se ejecuta en un dispositivo informático (por ejemplo un servidor de mensajería). El producto de programa informático puede almacenarse en un medio de grabación legible por ordenador (no transitorio).

De acuerdo con otro aspecto, se proporciona un servidor de mensajería, donde el servidor de mensajería está configurado para transmitir mensajes electrónicos recibidos desde al menos un dispositivo de envío de mensajes a al menos un dispositivo de recepción de mensajes. El servidor de mensajería está configurado además para detectar mensajes electrónicos maliciosos y comprende los siguientes componentes: una unidad de generación configurada para generar al menos una firma para un mensaje electrónico a transmitir desde el al menos un dispositivo de envío de mensajes a el al menos un dispositivo de recepción de mensajes; una unidad de almacenamiento de datos configurada para almacenar la al menos una firma generada; una unidad antivirus configurada para determinar si el mensaje electrónico es o no malicioso; una unidad de determinación configurada para determinar, en el caso de un mensaje malicioso determinado y basándose en al menos una firma generada para el mensaje malicioso determinado, si los mensajes electrónicos comparables al mensaje malicioso determinado se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes en el pasado, comparando la al menos una firma del mensaje malicioso determinado con firmas en la unidad de almacenamiento de datos que

están asociadas con los mensajes electrónicos transmitidos anteriormente; y una unidad de señalización configurada para señalizar una posible amenaza para el al menos un dispositivo de recepción de mensajes, si se determina que los mensajes electrónicos comparables con el mensaje malicioso determinado se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes en el pasado.

5

10

15

25

30

45

50

El servidor de mensajería puede implementarse en la forma de un solo dispositivo informático o en la forma de dispositivos informáticos distribuidos que están dispuestos en una red remota desde el al menos un dispositivo de envío de mensajes y el al menos un dispositivo de recepción de mensajes. El servidor de mensajería puede estar en comunicación (continua) con el al menos un dispositivo de envío de mensajes y el al menos un dispositivo de recepción de mensajes con el fin de recibir mensajes desde al menos un dispositivo de envío de mensajes y para transmitir los mensajes recibidos a el al menos un dispositivo de recepción de mensajes.

De acuerdo con otro aspecto más, se proporciona un sistema de mensajería, que comprende el servidor de mensajería descrito anteriormente, al menos un dispositivo de envío de mensajes, y al menos un dispositivo de recepción de mensajes. El servidor de mensajería está en comunicación con el al menos un dispositivo de envío de mensajes y con el al menos un dispositivo de recepción de mensajes.

Breve descripción de los dibujos

20 Otros detalles, aspectos y ventajas de la presente divulgación divulgadas en el presente documento serán evidentes a partir de los siguientes dibujos, en los que:

La figura 1: es un diagrama de bloques que ilustra un sistema de mensajería configurado para detectar

mensajes electrónicos maliciosos de acuerdo con una realización a modo de ejemplo de la

presente invención;

La figura 2: es un diagrama de bloques que ilustra un servidor de mensajería configurado para detectar

mensajes electrónicos maliciosos de acuerdo con una realización a modo de ejemplo de la

presente invención;

Las figuras 3a/3b: son diagramas de flujo que ilustran un método para detectar mensajes electrónicos maliciosos

de acuerdo con una realización a modo de ejemplo de la presente invención; y

La figura 4: es una representación esquemática adicional del método de la figura 3.

Descripción detallada

En la siguiente descripción, para fines de explicación y no de limitación, se exponen los detalles específicos con el fin de proporcionar un conocimiento profundo de la técnica presentada en el presente documento. Será evidente para un experto en la materia que la técnica divulgada puede practicarse en otras realizaciones que se apartan de estos detalles específicos.

40 La figura 1 ilustra, en la forma de un diagrama de bloques, una realización a modo de ejemplo de un sistema de mensajería 1 que puede implementar la técnica descrita a continuación para detectar mensajes electrónicos maliciosos.

El sistema de mensajería 1 comprende al menos un servidor de mensajería 1000, uno o más dispositivos de envío de mensajes 110, 120, 130, y uno o más dispositivos de recepción de mensajes 210, 220, 230. El uno o más dispositivos de envío de mensajes 110, 120, 130 y el uno o más dispositivos de recepción de mensajes 210, 220, 230 pueden implementarse cada uno en la forma de terminales de usuario portátiles (tales como PDA, teléfonos móviles, teléfonos inteligentes, ordenadores portátiles) o dispositivos informáticos fijos. El servidor de mensajería 1000 puede realizarse en la forma de un solo dispositivo informático o en la forma de dos o más dispositivos informáticos distribuidos a través de una red 2. La invención no depende de los detalles de implementación de hardware del servidor de mensajería 1000 descritos anteriormente, de los dispositivos de envío de mensajes 110, 120, 130 y de los dispositivos de recepción de mensajes 210, 220, 230.

Como se ilustra en la figura 1, el servidor de mensajería 1000 está dispuesto alejado de los dispositivos de envío de mensajes 110, 120, 130 y los dispositivos de recepción de mensajes 210, 220, 230 correspondientes. El servidor de 55 mensajería 1000 está configurado para encaminar los mensajes 100 recibidos desde el uno o más dispositivos de envío de mensajes 110, 120, 130 al uno o más dispositivos de recepción de mensajes específicos 210, para los que están destinados los mensajes. Para este fin, el servidor de mensajería 1000 está dispuesto para estar en comunicación (continua) con cada uno de los dispositivos de envío de mensajes 110, 120, 130 y cada uno de los 60 dispositivos de recepción de mensajes 210, 220, 230. La comunicación entre el servidor de mensajería 1000 por un lado, y los dispositivos de envío de mensajes 110, 120, 130 y los dispositivos de recepción de mensajes 210, 220, 230 correspondientes por otro lado, puede establecerse a través de canales de comunicación inalámbricos y/o cableados correspondientes. Además, puede usarse cualquier estándar de comunicación conocido con el fin de comunicar uno o más mensajes 100 entre los dispositivos de envío de mensajes 110, 120, 130 y los dispositivos de 65 recepción de mensajes 210, 220, 230. Por ejemplo, puede usarse el protocolo TCP/IP para la comunicación de mensajes.

ES 2 703 861 T3

El servidor de mensajería 1000 está diseñado para recibir uno o más mensajes 100 desde los dispositivos de envío de mensajes de comunicación 110, 120, 130, y para encaminar los mensajes recibidos 100 hacia los dispositivos de recepción de mensajes correspondientes 210, 220, 230, para los que están destinados los mensajes recibidos 100. De acuerdo con la ilustración a modo de ejemplo de la figura 1, el servidor de mensajería 1000 está diseñado para encaminar un mensaje 100 recibido desde el dispositivo de envío de mensajes 110 al dispositivo de recepción de mensajes de destinado 230.

El servidor 1000 está configurado además para realizar una comprobación de antisoftware malicioso (por ejemplo, un análisis de antivirus) para cada mensaje recibido 100 antes de que el mensaje 100 se encamine a los dispositivos de recepción de mensajes correspondientes 210, 220, 230. El servidor 1000 bloqueará todos aquellos mensajes 100 que se identifiquen como mensajes maliciosos (es decir, que comprenden contenidos maliciosos), mientras que los mensajes 100 que se identifican como no maliciosos (es decir, mensajes limpios) se encaminan directamente a los dispositivos de recepción de mensajes correspondientes 210, 220, 230.

10

20

45

50

55

60

65

15 Haciendo referencia a la figura 2, se describe adicionalmente la estructura y la funcionalidad del servidor de mensajería 1000.

El servidor 1000 comprende una unidad de generación 1010, una unidad de antivirus 1020 (en la figura 2 se hace referencia como unidad de AV), una unidad de comparación 1030, una unidad de señalización 1040 y una unidad de almacenamiento de datos 1050. Además, el servidor 1000 comprende una primera interfaz de comunicación 1070 y una segunda interfaz de comunicación 1080. Las unidades 1010, 1020, 1030, 1040, 1050, así como las interfaces 1070, 1080 están en comunicación entre sí.

Cada una de entre la unidad de generación 1010, la unidad de comparación 1030, la unidad de señalización 1040 y la unidad de AV 1020 puede implementarse como un módulo de software separado, un módulo de hardware o un módulo de software/hardware combinado. Como alternativa, la unidad de generación 1010, la unidad de comparación 1030, la unidad de señalización 1040 y la unidad de AV 1020 también pueden implementarse como submódulos de un módulo de software y/o hardware comúnmente diseñados, como se ilustra mediante la caja de puntos en la figura 1. Un experto en la materia apreciará que las unidades 1010, 1020, 1030, 1040 mencionadas anteriormente pueden implementarse usando un software que funcione junto con un microprocesador programado, usando un circuito integrado de aplicación específica (ASIC), un procesador de señal digital (DSP) o un ordenador de fin general.

Independientemente de los detalles de implementación mencionados anteriormente, la unidad de generación 1010 está en comunicación con la unidad de entrada 1070 y está configurada para generar, sobre la base de los mensajes recibidos 100 desde la unidad de entrada 1070, al menos una firma para cada mensaje electrónico 100. De acuerdo con una implementación, la unidad de generación 1010 también está configurada para generar un identificador de mensaje (ID) para cada mensaje 100 para el que se genera al menos una firma.

40 La unidad de almacenamiento de datos 1050 está configurada para almacenar, para cada mensaje 100, la al menos una firma y el ID asociado generado y entregado por la unidad de generación 1010.

La unidad de AV 1020 está configurada para proporcionar un análisis de antivirus (análisis AV) para cada mensaje electrónico 100 recibido a través de la unidad de entrada 1070. El análisis de AV implica un análisis de comparación realizado sobre la base de una lista negra actualizable (BL) proporcionada por un proveedor de software de antivirus. La BL comprende una lista completa de firmas de software malicioso de todos los softwares maliciosos conocidos contra la que se compara el contenido del mensaje (incluidos los archivos adjuntos). Además, ya que la BL puede actualizarse, la BL puede complementarse regularmente con las firmas de software malicioso más recientes. Se observa que la presente invención no depende de la técnica de análisis de AV específica. También es concebible que la unidad de AV 1020 use técnicas heurísticas y/o técnicas de emulación además de la comparación de BL mencionada anteriormente.

La unidad de AV 1020 está configurada además para filtrar y bloquear esos mensajes 100 que se encuentran que comprenden contenido malicioso. La unidad de AV 1020 clasifica tales mensajes como mensajes maliciosos. Por lo tanto, la unidad de AV 1020 permite una transmisión de solo aquellos mensajes 100 que se encuentran limpios a los dispositivos de recepción de mensajes 210, 220, 230.

La unidad de comparación 1030 está configurada para comparar las firmas generadas de los mensajes maliciosos recibidos actual o recientemente 100 revelando nuevos tipos de amenazas (es decir firmas generadas por la unidad de generación 1010 para los mensajes que comprenden software malicioso de día cero) con las firmas de los mensajes recibidos anteriormente 100, que se almacenan en la unidad de almacenamiento de datos 1050. La unidad de comparación 1030 está configurada además para seleccionar aquellas firmas asociadas con los mensajes transmitidos anteriormente 100 desde la unidad de almacenamiento de datos 1050, que se consideran comparables con las firmas generadas de los mensajes maliciosos recibidos actual o recientemente 100. Aún más, la unidad de comparación 1030 está configurada para identificar, sobre la base de las firmas seleccionadas y los ID de mensajes asociados, los mensajes correspondientes que se han transmitido en el pasado y que se han considerado mensajes

limpios (es decir, los mensajes no maliciosos).

La unidad de señalización 1040 está configurada para generar, sobre la base de los mensajes identificados, al menos una notificación 105 que indica que los mensajes específicos transmitidos en el pasado pueden ser maliciosos. La notificación(es) 105 puede comprender una lista de aquellos mensajes que se consideraron como mensajes limpios en el pasado, pero que han revelado firmas comparables a la firma(s) del mensaje actual, que se ha clasificado como maliciosa durante la primera vez.

La primera interfaz de comunicación 1070 está configurada para recibir mensajes electrónicos 100 desde los dispositivos de envío de mensajes correspondientes 110, 120, 130 y para proporcionar los mensajes electrónicos recibidos 100 a la unidad de comparación 1010 y a la unidad de AV 1020. Además, la segunda interfaz de comunicación 1080 está configurada para transmitir esos mensajes 100, que se determina que están limpios por la unidad de AV 1020 (y, por lo tanto, no bloqueados por la unidad de AV 1020), a los dispositivos de recepción de mensajes 210, 220, 230. Además, la segunda interfaz de comunicación 1080 también puede configurarse para transmitir la notificación(es) 105 generada a los respectivos dispositivos de recepción de mensajes 210, 220, 230. Ambas interfaces de comunicación 1070, 1080 pueden implementarse en la forma de una interfaz de comunicación inalámbrica (por ejemplo, una interfaz de transmisión de radio) y/o una interfaz de comunicación cableada, en función de cómo se implemente la comunicación entre los respectivos dispositivos de envío de mensajes 110, 120,130 y los dispositivos de recepción de mensajes 210, 220, 230.

20

25

10

15

Las funcionalidades de las unidades 1010, 1020, 1030, 1040 mostradas en la figura 2, se explican adicionalmente junto con los dos diagramas de flujo 300, 350 de las figuras 3a y 3b. Los diagramas de flujo 300, 350 ilustran un método para detectar mensajes electrónicos maliciosos 100 transmitidos desde el uno o más dispositivos de envío de mensajes 110, 120, 130 a el uno o más dispositivos de recepción de mensajes 210, 220, 230. El método se realiza en el servidor de mensajería 1000 como ya se ha explicado anteriormente junto con la figura 2 y comprende las siguientes etapas para cada mensaje 100 recibido de al menos uno de los al menos un dispositivo de envío de mensajes 210, 220, 230.

El método comienza con la etapa 310 en la figura 3a, de acuerdo con la cual para cada mensaje recibido 100 se genera al menos una firma indicativa de las propiedades del mensaje por el módulo de generación 110. La al menos una firma generada puede mapear las propiedades del mensaje, tales como las propiedades de adjuntos de mensaje (por ejemplo, al menos una de entre un nombre, tipo de archivo y extensión de archivo de un adjunto de mensaje), una información de URL asociada con el dispositivo de recepción de mensajes, desde la que proviene el mensaje. Se observa que las firmas de mensajes se generan por la unidad de generación 1010 en tiempo real para cada mensaje entrante 100. Además, las firmas de mensajes generadas pueden ser diferentes de las firmas de software malicioso usadas por la unidad de AV 1020 y proporcionadas por un proveedor de software de antivirus. Además, para cada mensaje recibido 100, se genera un ID de mensaje y se asigna a la al menos una firma generada de tal manera que las firmas puedan asociarse más adelante con el mensaje 100.

40 Por otra parte, de acuerdo con un mensaje específico de implementación puede generarse adicionalmente para cada mensaje 100 una información adecuada para clasificar e identificar los mensajes transmitidos. Tal información puede comprender al menos uno de entre una hora de mensaje, una información de mensaje remitente, una información de destino de mensaje y una información de adjunto de mensaje disponibles para el servidor de mensajería 1000 en el momento en que se recibe el mensaje desde el al menos un dispositivo de envío de mensajes 45 110, 120, 130. La hora del mensaje puede indicar un instante en el tiempo en el que se recibió el mensaje 100 desde un dispositivo de envío de mensajes 110, 120, 130 o se transmitió a un dispositivo de recepción de mensajes 210, 220, 230. La información del remitente puede comprender un nombre, dirección IP u otra información para identificar el dispositivo de envío de mensajes 110, 120, 130 desde el que se ha recibido el mensaje 100. De manera similar, la información de destino puede comprender un nombre, dirección IP u otra información para identificar el dispositivo 50 de recepción de mensajes 210, 220, 230 al que se ha enviado el mensaje 100. La información de adjunto de mensaje puede, por ejemplo, comprender al menos un nombre de adjunto o información similar para identificar el adjunto.

En una etapa posterior 315, para cada mensaje recibido 100 la al menos una firma generada (y, opcionalmente, la información de mensaje específico) se almacena junto con el ID generado en la unidad de almacenamiento de datos 1050. Por lo tanto, dentro de un flujo de mensajes continuo desde el al menos un dispositivo de envío de mensajes 110, 120, 130 a el al menos un dispositivo de recepción de mensajes 210, 220, 230, las firmas de mensaje (y, opcionalmente, la información de mensaje específico) se registran para cada mensaje 100 y se ponen a disposición para su análisis en retrospectiva si los mensajes 100 transmitidos en el pasado contenían o no software malicioso.

60

En una etapa posterior 320, cada mensaje recibido 100 se proporciona a la unidad de AV 1020. La unidad de AV 1020 realiza inmediatamente un análisis de AV con el fin de determinar si el mensaje recibido 100 es o no malicioso. Con el fin de determinar si el mensaje 100 recibido es o no malicioso, la unidad de AV 1020 puede aplicar una técnica de detección de AV como se ha descrito anteriormente en la figura 2.

Si se determina, mediante la unidad de AV 1020, que el mensaje recibido 100 no es malicioso (véase la etapa de decisión 325 y la rama "NO" en la figura 3a), el mensaje 100 se transmite a los dispositivos de recepción de mensajes correspondientes 210, 220, 230 (etapa 330 en la figura 3a). En tal caso, el método termina para el mensaje considerado 100 y se reinicia con un mensaje recibido recientemente 100 (es decir, el siguiente mensaje recibido 100 en el flujo de mensajes).

Sin embargo, si se determina, por la unidad de AV 1020, que el mensaje recibido 100 es malicioso (véase la rama "SÍ" en la figura 3a), el mensaje 100 a transmitir a el al menos un dispositivo de recepción de mensajes 210, 220, 230 se bloquea. Opcionalmente, puede transmitirse una notificación al dispositivo de recepción de mensajes destinado 302, 304, 306 que indica que el mensaje 100 se ha bloqueado debido a que el mensaje 100 se ha clasificado como mensaje malicioso (no mostrado en la figura 3a).

10

15

20

25

30

35

Además, si se determina, por la unidad de AV 1020, que el mensaje recibido 100 es malicioso (véase la rama "SÍ" en la figura 3a), se comprueba adicionalmente en la etapa de método posterior 345 si el mensaje malicioso 100 incluye un nuevo software malicioso (es decir, una nueva amenaza) que hasta ahora no se había detectado por la unidad de AV 1020. Más específicamente, se verifica si el software malicioso determinado constituye un software malicioso de día cero que solo podría detectarse por la unidad de AV 1020 teniendo en cuenta las últimas actualizaciones de firmas de software malicioso disponibles por un proveedor de antivirus. Si se descubre que el software malicioso o la amenaza determinada ya se había detectado en el pasado, se considera que el software malicioso no es nuevo y el algoritmo se detiene para el mensaje actualmente considerado (véase la rama "NO" en la figura 3a). Si se descubre que el módulo de AV 1020 detecta el software malicioso por primera vez, se considera que es un software malicioso de día cero y el método continúa con las etapas 350 y 360 (véase la rama "SÍ" en la figura 3a).

En la etapa 350 se determina, sobre la base de la al menos una firma generada, si los mensajes 100 comparables al mensaje malicioso determinado 100 se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes 210, 220, 230 en el pasado. La etapa de determinación 350 se realiza mediante la unidad de comparación 1030 y comprende las siguientes subetapas ilustradas en el diagrama de flujo de la figura 3b. En la subetapa 352, la al menos una firma generada del mensaje malicioso determinado 100 se compara con las firmas de la unidad de almacenamiento de datos 1050 que se han generado y almacenado en la unidad de almacenamiento de datos 1050 para los mensajes electrónicos transmitidos anteriormente 100. La comparación se realiza sobre la base de una verificación de similitud. Es decir, las firmas comparadas entre sí se consideran comparables si se encuentra que las firmas son idénticas o muy similares. Se observa que las firmas comparadas se consideran idénticas si su información de firma corresponde al 100 %. Las firmas que se comparan entre sí se consideran altamente similares si la correspondencia entre la información de firma respectiva supera un valor de umbral predeterminado. Preferentemente, se requiere una correspondencia de información de firma del 80 % o más para considerar las firmas como altamente similares. Más preferentemente, se requiere una correspondencia de información de firma del 90 % o más para considerar las firmas como altamente similares.

Si ninguna de las firmas almacenadas en la unidad de almacenamiento 1050 es comparable (es decir, idéntica o altamente similar) a la al menos una firma generada (véase la etapa de decisión 353, rama "NO"), el algoritmo se detiene (etapa 354) para el mensaje recibido actualmente 100 y se reinicia en la etapa 310 para un mensaje recibido recientemente 100. Sin embargo, si algunas de las firmas almacenadas en la unidad de almacenamiento 1050 son comparables a la al menos una firma generada (etapa de decisión 353, rama "Sí"), el método continúa con la subetapa 355 identificando aquellos mensajes electrónicos transmitidos en el pasado que se ha encontrado que tienen firmas comparables. La identificación de mensaje se realiza sobre la base de los ID de mensaje (y la información de mensaje adicional) almacenados en la unidad de almacenamiento 1050 junto con cada firma almacenada.

Basándose en los mensajes identificados, la unidad de señalización 1040 notifica en una etapa posterior 360 (véase de nuevo la figura 3a) a el al menos un dispositivo de recepción de mensajes 210, 220, 230 sobre una posible amenaza(s) en el mensaje(s) transmitido 100. De acuerdo con una implementación, la notificación comprende generar y comunicar una notificación 105 a el al menos un dispositivo de recepción de mensajes 210, 220, 230, donde la notificación generada 105 comprende una lista de los mensajes identificados, y opcionalmente, el tipo de amenaza identificada para cada mensaje. Por lo tanto, al recibir la notificación 105, se puede advertir a el al menos un dispositivo de envío de mensajes 210, 220 230 sobre un nuevo software malicioso infiltrado en el dispositivo de envío de mensajes 210, 220, 230. Basándose en esta información, el usuario del dispositivo de envío de mensajes 210, 220, 230 puede tomar medidas de seguridad adecuadas para limitar los posibles daños provocados por el software malicioso infiltrado.

Como no todos los mensajes identificados 100 pueden transmitirse a cada uno de los dispositivos de recepción de mensajes 210, 220, 230, la lista de mensajes puede generarse individualmente para cada dispositivo de recepción de mensajes 210, 220, 230 comprendiendo solo aquellos mensajes identificados que se han transmitido al dispositivo de recepción de mensajes específico 210, 220, 230. Es decir, para cada dispositivo de recepción de mensajes específico 210, 220, 230 se genera una lista específica que contiene solo los mensajes transmitidos a los dispositivos de recepción de mensajes específicos 210, 220, 230.

ES 2 703 861 T3

Se observa que para cada mensaje recibido 100, el método termina con la etapa 356 y se reinicia en la etapa 310 para un nuevo mensaje posterior.

El beneficio de la técnica de detección de software malicioso descrito anteriormente se trata más adelante junto con la figura 4. La figura 4 ilustra un diagrama que comprende una línea de tiempo que discurre verticalmente y una línea de evento que discurre horizontalmente. La línea de tiempo comprende diferentes puntos en el tiempo t0, t1 y t2, donde t0 representa un punto actual en el tiempo, que en la figura 4 también se conoce como "tiempo de observación". Los puntos en el tiempo t1 y t2 representan puntos anteriores en el tiempo. Es decir, t2 representa un punto en el tiempo, en el que puede aparecer un nuevo software malicioso o amenaza en un mensaje 100, pero la unidad de AV 1020 no lo detecta porque la unidad de AV 1020 aún no es sensible al nuevo software malicioso. Además, t1 representa un punto en el tiempo entre t0 y t2, en el que están disponibles las nuevas firmas de virus para la unidad de AV 1020 con el fin de detectar y filtrar los mensajes que comprenden el nuevo software malicioso.

10

Por lo tanto, tal como se ilustra por las flechas verticales en la figura 3b, solamente para los puntos en el tiempo posteriores a t1 puede detectarse el nuevo software malicioso y los mensajes que comprenden el nuevo software malicioso pueden filtrarse. Además, existe un intervalo de tiempo desde t1 a t2 en el que el nuevo software malicioso ya existe, pero aún no puede detectarse debido a la falta de firmas de software malicioso adecuadas para la unidad de AV 1020. En los servidores de mensajería convencionales no se implementa la técnica de detección divulgada, el presente intervalo de tiempo no puede cerrarse y el nuevo software malicioso puede propagarse sin obstáculos en una pluralidad de dispositivos de mensajería. Además, ya que los usuarios de los dispositivos de mensajería suelen confiar en las capacidades de filtrado de software malicioso del servidor de mensajería, es probable que el nuevo software malicioso permanezca sin detectarse durante largos períodos de tiempo en los dispositivos de recepción de mensajes.

25 La técnica de detección de software malicioso divulgada cierra este intervalo de tiempo mediante la recopilación y el almacenamiento de firmas para todos los mensajes transmitidos, es decir, también para los mensajes enviados entre t1 y t2. Las firmas recopiladas se comparan con las firmas de los mensajes maliciosos que comprenden el nuevo software malicioso (es decir, los mensajes de software malicioso de día cero) detectados en instancias de tiempo posteriores a t1. Si existen mensajes transmitidos entre t1 y t2 que tengan firmas comparables a las firmas de los 30 mensajes de software malicioso de día cero detectados, se generan notificaciones y se transmiten a los dispositivos de recepción de mensajes correspondientes, lo que indica que los mensajes de software malicioso de día cero se han transmitido en el pasado. Por lo tanto, con la técnica de detección de software malicioso reivindicada, los usuarios de los dispositivos de recepción de mensajes pueden ser advertidos en una etapa temprana contra los mensajes de software malicioso de día cero de tal manera que los usuarios puedan tomar las contramedidas 35 apropiadas. Además, la técnica de detección de software malicioso reivindicada es compatible con las técnicas de detección de AV convencionales. Por lo tanto, la técnica de detección de software malicioso reivindicada puede implementarse fácilmente en servidores de mensajería ya existentes.

Mientras que la técnica presentada en el presente documento se ha descrito con respecto a realizaciones específicas, los expertos en la materia reconocerán que la presente invención no se limita a las realizaciones específicas descritas e ilustradas en el presente documento. Debería entenderse que la divulgación es solamente ilustrativa. La invención se define en las reivindicaciones adjuntas.

REIVINDICACIONES

- 1. Un método para detectar mensajes electrónicos maliciosos (100) transmitidos desde el al menos un dispositivo de envío de mensajes (110, 120, 130) a el al menos un dispositivo de recepción de mensajes (210, 220, 230), realizándose el método mediante un servidor de mensajería (1000) y que comprende las etapas de:
 - a) generar, mediante el servidor de mensajería (1000), al menos una firma para un mensaje electrónico (100) a transmitir desde el al menos un dispositivo de envío de mensajes (110, 120, 130) a el al menos un dispositivo de recepción de mensajes (210, 220, 230);
- b) almacenar, mediante el servidor de mensajería (1000), la al menos una firma generada en una unidad de almacenamiento de datos (1050);
 - c) determinar, mediante el servidor de mensajería (1000), si el mensaje electrónico (100) es malicioso;
 - d) si se determina que el mensaje electrónico (100) es malicioso, determinar, mediante el servidor de mensajería (1000) sobre la base de la al menos una firma generada para el mensaje malicioso determinado, si los mensajes electrónicos (100) comparables al mensaje malicioso determinado (100) se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes (210, 220, 230) en el pasado, comprendiendo la etapa de determinación comparar la al menos una firma del mensaje malicioso determinado con las firmas en la unidad de almacenamiento de datos (1050) que están asociadas con los mensajes electrónicos transmitidos anteriormente (100); y
- e) si se determina que los mensajes electrónicos (100) comparables con el mensaje malicioso determinado (100) se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes (210, 220, 230) en el pasado, notificar, mediante el servidor de mensajería (1000), a el al menos un dispositivo de recepción de mensajes (210, 220, 230) sobre una posible amenaza.
- 25 2. El método de acuerdo con la reivindicación 1, donde se repiten al menos las etapas (a) a (c), mediante el servidor de mensajería (1000), para cada nuevo mensaje a transmitir desde el al menos un dispositivo de envío de mensajes (110, 120, 130) a el al menos un dispositivo de recepción de mensajes (210, 220, 230).
- 3. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, donde la etapa (a) y la etapa (b) comprenden además: generar, mediante el servidor de mensajería (1000), un identificador de mensaje (ID) para el mensaje electrónico (100) y almacenar, mediante el servidor de mensajería (1000), el ID de mensaje generado junto con la al menos una firma generada en la unidad de almacenamiento de datos (1050).
- 4. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, donde la etapa (a) comprende además: generar, mediante el servidor de mensajería (1000), para el mensaje electrónico (100) a transmitir, una información de mensaje específica adecuada para clasificar el mensaje transmitido más tarde.
- 5. El método de acuerdo con la reivindicación 4, donde la etapa (b) comprende además: almacenar, mediante el servidor de mensajería (1000), la información de mensaje específica generada junto con la al menos una firma 40 generada en la unidad de almacenamiento de datos (1050).
 - 6. El método de acuerdo con la reivindicación 4 o 5, donde la información de mensaje específica comprende al menos una de la siguiente información:
- 45 tiempo de transmisión de mensaje;

15

- información de remitente de mensaie:
- información de destino de mensaje; e
- información sobre los adjuntos de mensaje.
- 7. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, donde la firma se genera, mediante el servidor de mensajería (1000), sobre la base de al menos uno de los siguientes elementos:
 - al menos una propiedad de adjunto de mensaje; y
 - la información de URL asociada con el dispositivo de envío de mensajes (110, 120, 130) del que proviene el mensaje.
 - 8. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, donde la etapa (c) comprende además:
- comparar, mediante el servidor de mensajería (1000), el mensaje electrónico (100) con firmas de virus conocidas; y
 - clasificar, mediante el servidor de mensajería (1000), el mensaje electrónico (100) como malicioso, si el mensaje (100) coincide suficientemente con una de las firmas de virus conocidas.
- 9. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, que comprende además al menos una de las siguientes etapas:

ES 2 703 861 T3

- transmitir, mediante el servidor de mensajería (1000), el mensaje electrónico (100) a el al menos un dispositivo de recepción de mensajes (210, 220, 230), si el mensaje (100) se ha clasificado como no malicioso; y
- bloquear, mediante el servidor de mensajería (1000), el mensaje electrónico (100), si el mensaje (100) se ha clasificado como malicioso.

5

10. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, donde la etapa (e) comprende además: proporcionar, mediante el servidor de mensajería (1000), información a el al menos un dispositivo de recepción de mensajes (110, 120, 130) que indica que un mensaje específico transmitido en el pasado puede ser malicioso.

10

25

- 11. Un producto de programa informático con partes de código de programa para realizar el método de acuerdo con al menos una de las reivindicaciones anteriores 1-10, cuando el producto de programa informático se ejecuta en un dispositivo informático.
- 15 12. El producto de programa informático de acuerdo con la reivindicación 11, que se almacena en un medio de grabación legible por ordenador.
- 13. Un servidor de mensajería (1000) que está configurado para transmitir mensajes electrónicos (100) recibidos desde el al menos un dispositivo de envío de mensajes (110, 120, 130) a el al menos un dispositivo de recepción de mensajes (210, 220, 230), estando el servidor de mensajería (1000) diseñado para detectar mensajes electrónicos maliciosos (100) y comprendiendo:
 - una unidad de generación (1010) configurada para generar al menos una firma para un mensaje electrónico (100) a transmitir desde el al menos un dispositivo de envío de mensajes (110, 120, 130) a el al menos un dispositivo de recepción de mensajes (210, 220, 230);
 - una unidad de almacenamiento de datos (1050) configurada para almacenar la al menos una firma generada;
 - una unidad antivirus (1020) configurada para determinar si el mensaje electrónico (100) es o no malicioso;
 - una unidad de determinación (1030) configurada para determinar en el caso de un mensaje malicioso determinado y basándose en al menos una firma generada para el mensaje malicioso determinado, si los mensajes electrónicos (100) comparables al mensaje malicioso determinado (100) se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes (210, 220, 230) en el pasado, comparando la al menos una firma del mensaje malicioso determinado con las firmas en la unidad de almacenamiento de datos (1050) que están asociadas con los mensajes electrónicos transmitidos anteriormente (100): v
- una unidad de señalización (1040) configurada para señalar una posible amenaza para el al menos un dispositivo de recepción de mensajes (210, 220, 230), si se determina que los mensajes electrónicos (100) comparables con el mensaje malicioso determinado (100) se han clasificado como no maliciosos y se han transmitido a el al menos un dispositivo de recepción de mensajes (210, 220, 230) en el pasado.
- 40 14. Un sistema de mensajería (1) que comprende:
 - al menos un dispositivo de envío de mensajes (110, 120, 130);
 - al menos un dispositivo de recepción de mensajes (210, 220, 230); y
- el servidor de mensajería (1000) de acuerdo con la reivindicación 13, estando el servidor de mensajería (1000)
 en comunicación con el al menos un dispositivo de envío de mensajes (110, 120, 130) y con el al menos un dispositivo de recepción de mensajes (210, 220, 230).









