Unión Internacional de Telecomunicaciones

UIT-T

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

X.1235

(01/2022)

SERIE X: REDES DE DATOS, COMUNICACIONES DE SISTEMAS ABIERTOS Y SEGURIDAD

Seguridad en el ciberespacio – Lucha contra el correo basura

Tecnologías contra la suplantación de sitios web para las organizaciones de telecomunicaciones

Recomendación UIT-T X.1235

RECOMENDACIONES UIT-T DE LA SERIE X

REDES DE DATOS, COMUNICACIONES DE SISTEMAS ABIERTOS Y SEGURIDAD

REDES PÚBLICAS DE DATOS	X.1–X.199
INTERCONEXIÓN DE SISTEMAS ABIERTOS	X.200–X.299
INTERFUNCIONAMIENTO ENTRE REDES	X.300–X.399
SISTEMAS DE TRATAMIENTO DE MENSAJES	X.400–X.499
DIRECTORIO	X.500–X.599
GESTIÓN DE REDES DE INTERCONEXIÓN DE SISTEMAS ABIERTOS Y ASPECTOS DE SISTEMAS	X.600–X.699
GESTIÓN DE INTERCONEXIÓN DE SISTEMAS ABIERTOS	X.700–X.799
SEGURIDAD	X.800–X.849
APLICACIONES DE INTERCONEXIÓN DE SISTEMAS ABIERTOS	X.850–X.899
PROCESAMIENTO DISTRIBUIDO ABIERTO	X.900–X.999
SEGURIDAD DE LA INFORMACIÓN Y DE LAS REDES	
Aspectos generales de la seguridad	X.1000–X.1029
Seguridad de las redes	X.1030–X.1049
Gestión de la seguridad	X.1050–X.1069
Telebiometría	X.1080–X.1099
APLICACIONES Y SERVICIOS CON SEGURIDAD	x.1150 –x.1159
Seguridad en la multidifusión	X.1100–X.1109
Seguridad en la red residencial	X.1110–X.1119
Seguridad en las redes móviles	X.1120–X.1139
Seguridad en la web (1)	X.1140–X.1149
Seguridad de las aplicaciones (1)	
Protocolos de seguridad	X.1150–X.1159
Seguridad en las comunicaciones punto a punto	X.1160–X.1169
Seguridad de la identidad en las redes	X.1170–X.1179
Seguridad en la TVIP	X.1180–X.1199
SEGURIDAD EN EL CIBERESPACIO	
Ciberseguridad	X.1200–X.1229
Lucha contra el correo basura	**X.1230–X.1249**
Gestión de identidades	X.1250–X.1279
APLICACIONES Y SERVICIOS CON SEGURIDAD	
Comunicaciones de emergencia	X.1300–X.1309
Seguridad en las redes de sensores ubicuos	X.1310–X.1339
Seguridad de las redes eléctricas inteligentes	X.1330–X.1339
Recomendaciones relacionadas con la PKI	X.1340–X.1349
Seguridad en la Internet de las cosas (IoT)	X.1350–X.1369
Seguridad en los sistemas de transporte inteligente (ITS)	X.1370–X.13979
Seguridad de tecnología de libro mayor distribuido (2)	X.1400–X.1429
Seguridad de las aplicaciones (2)	X.1450–X.1459
Seguridad de la web (2)	X.1470–X.1489
INTERCAMBIO DE INFORMACIÓN DE CIBERSEGURIDAD	
Aspectos generales de la ciberseguridad	X.1500–X.1519
Intercambio de estados/vulnerabilidad	X.1520–X.1539
Intercambio de eventos/incidentes/heurística	X.1540–X.1549
Intercambio de políticas	X.1550–X.1559
Petición de heurística e información	X.1560–X.1569
Identificación y descubrimiento	X.1570–X.1579
Intercambio asegurado	X.1580–X.1589
Ciberdefensa	X.1590– X.1599
SEGURIDAD DE LA COMPUTACIÓN EN NUBE	
Visión general de la seguridad de la computación en nube	X.1600 –X.1601
Diseño de la seguridad de la computación en nube	X.1602 –X.1639
Prácticas óptimas y directrices en materia de seguridad de la computación en nube	X.1640 –X.1659
Aplicación práctica de la seguridad de la computación en nube	X.1660 –X.1679
Otras cuestiones de seguridad de la computación en nube	X.1680 –X.1699
COMUNICCINOES CUÁNTICAS	
Terminología	X.1700 –X.1701
Generador cuántico de números aleatorios cuánticos	X.1702 –X.1709
Marco de seguridad para QKDN	X.1710 –X.1711
Seguridad de diseño para OKBN	X.1712 –X.1719
Técnicas de seguridad para OKDN	X.1720 –X.1729
SEGURIDAD DE LOS DATOS	
Seguridad de macrodatos	X.1750 –X.1759
Protección de datos	X.1770–X.1789
SEGURIDAD DE LAS IMT-2020	X.1800–X.1819

Para más información, véase la Lista de Recomendaciones del UIT-T.

Recomendación UIT-T X.1235

Tecnologías contra la suplantación de sitios web para las organizaciones de telecomunicaciones

Resumen

La suplantación de sitios web es una de las principales amenazas para las organizaciones de telecomunicaciones, especialmente para los operadores. Se recomienda que los operadores de telecomunicaciones adopten tecnologías contra la suplantación de sitios web a fin de proteger a sus clientes y cuidar su reputación y sus ingresos. En la Recomendación UIT-T X.1235, se analizan las principales actuaciones para la suplantación de un sitio web y se recomiendan tecnologías para detectar sitios web suplantados. Estas medidas se pueden considerar directrices para las organizaciones de telecomunicaciones a fin de proteger los sitios web contra la suplantación. Puede aplicarse un enfoque similar contra la suplantación de cualquier sitio web, incluidos los de bancos, compañías de seguros y tiendas en Internet, entre otros.

Historia

Edición	Recomendación	Aprobación	Comisión de Estudio	ID único*
1.0	ITU-T X.1235	2022-01-07	17	11.1002/1000/14797

Palabras clave

Contramedidas, suplantación de sitios web.

* Para acceder a la Recomendación, sírvase digitar el URL http://handle.itu.int/ en el campo de dirección del navegador, seguido por el identificador único de la Recomendación. Por ejemplo, http://handle.itu.int/11.1002/1000/11830-en.

PREFACIO

La Unión Internacional de Telecomunicaciones (UIT) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones y de las tecnologías de la información y la comunicación. El Sector de Normalización de las Telecomunicaciones de la UIT (UIT-T) es un órgano permanente de la UIT. Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Asamblea Mundial de Normalización de las Telecomunicaciones (AMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución 1 de la AMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

NOTA

En esta Recomendación, la expresión "Administración" se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

La observancia de esta Recomendación es voluntaria. Ahora bien, la Recomendación puede contener ciertas disposiciones obligatorias (para asegurar, por ejemplo, la aplicabilidad o la interoperabilidad), por lo que la observancia se consigue con el cumplimiento exacto y puntual de todas las disposiciones obligatorias. La obligatoriedad de un elemento preceptivo o requisito se expresa mediante las frases "tener que, haber de, hay que + infinitivo" o el verbo principal en tiempo futuro simple de mandato, en modo afirmativo o negativo. El hecho de que se utilice esta formulación no entraña que la observancia se imponga a ninguna de las partes.

PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no adopta ninguna posición en cuanto a la demostración, validez o aplicabilidad de los derechos de propiedad intelectual reivindicados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT no ha recibido notificación de la existencia de propiedad intelectual, protegida por patente o derecho de autor, que puede ser necesaria para implementar esta Recomendación. Sin embargo, debe señalarse a los implementadores que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar las correspondientes bases de datos del UIT T disponibles en el sitio web del UIT T en http://www.itu.int/ITU-T/ipr/.

ÍNDICE

Introducción

La suplantación de identidad en la web ha desempeñado un papel fundamental en el fraude en Internet en los últimos años. Los estafadores suelen elegir sitios web de organizaciones o empresas conocidas para recoger las credenciales de los visitantes o para difundir software malicioso. El resultado es una pérdida financiera tanto para los visitantes como para los operadores de telecomunicaciones, y una pérdida de reputación para los operadores de telecomunicaciones

Dado que los sitios web de los operadores de telecomunicaciones se han convertido en uno de los portales más importantes para que sus clientes se informen y suscriban todo tipo de servicios, se ha observado en todo el mundo que los estafadores intentan continuamente falsificar los sitios web de los operadores de telecomunicaciones con el objetivo de engañar a los clientes. Esta Recomendación hace un análisis exhaustivo de la falsificación de sitios web y recomienda una serie de contramedidas organizadas.

Recomendación UIT-T X.1235

Tecnologías contra la suplantación de sitios web para las organizaciones de telecomunicaciones

1 Alcance

En la presente Recomendación se recomiendan tecnologías que permiten a las organizaciones de telecomunicaciones detectar a tiempo la suplantación de sitios web y proteger sus propios sitios contra la suplantación. Tras realizar un análisis sistemático de las medidas y características de la suplantación, se recomiendan prácticas idóneas para el uso de tecnologías del lado de la red, combinadas con el apoyo de tecnologías del lado del usuario, para luchar contra la suplantación de sitios web.

El cumplimiento de la presente Recomendación no se considerará, ni podrá ser utilizado, como prueba del cumplimiento de ningún reglamento, ley o política nacional o regional. Los medios técnicos, organizativos y de procedimiento descritos en la presente Recomendación no garantizan en modo alguno el nivel de seguridad que un reglamento, ley o política en concreto, nacional o regional, pudiera exigir para una determinada correspondencia.

Opcionalmente puede aplicarse un enfoque similar contra la suplantación de cualquier sitio web, incluidos los de bancos, compañías de seguros, tiendas de Internet, entre otros.

2 Referencias

Las siguientes Recomendaciones UIT-T y demás referencias contienen disposiciones que, por referencia a las mismas en este texto, constituyen disposiciones de esta Recomendación. En la fecha de publicación, las ediciones citadas estaban en vigor. Todas las Recomendaciones y demás referencias están sujetas a revisión, por lo que se alienta a los usuarios de esta Recomendación a que consideren la posibilidad de aplicar la edición más reciente de las Recomendaciones y demás referencias que se indican a continuación. Se publica periódicamente una lista de las Recomendaciones UIT-T vigentes. La referencia a un documento en el marco de esta Recomendación no confiere al mismo, como documento autónomo, el rango de Recomendación.

Ninguna.

3 Definiciones

3.1 Términos definidos en otros documentos

Ninguno.

3.2 Términos definidos en esta Recomendación

En la presente Recomendación se definen los términos siguientes:

3.2.1 red neuronal convolucional: algoritmo de aprendizaje profundo que toma una imagen de entrada, asigna importancia a diversos aspectos/objetos de la imagen y es capaz de diferenciar unos de otros.

3.2.2 redes de memoria a largo-corto plazo: tipo de red neuronal recurrente capaz de aprender la dependencia del orden en problemas de predicción de secuencias.

3.2.3 red neuronal recurrente: clase de red neuronal que permite utilizar elementos de salida anteriores como elementos de entrada, aunque tengan estados ocultos.

3.2.4 transformación de características invariante a la escala (SIFT): algoritmo de detección de características en visión por ordenador, que permite detectar y describir características locales en imágenes y compone un identificador descriptivo invariante a las traslaciones, rotaciones y modificaciones de escala en el ámbito de la imagen y resistente a las transformaciones moderadas de la perspectiva y las variaciones en términos de iluminación. A título experimental, el identificador descriptivo SIFT se ha revelado sumamente útil para la correspondencia de imágenes y el reconocimiento de objetos en condiciones reales.

3.2.5 características robustas aceleradas: detector de características locales e identificador descriptivo en visión por ordenador que permite detectar y describir características locales en imágenes, es varias veces más rápido que la transformación de características invariantes (SIFT) y resiste mejor a las transformaciones de las imágenes que la SIFT.

3.2.6 sitio web suplantado: sitio web creado con técnicas de suplantación de sitios web (véase la cláusula 3.2.8).

3.2.7 máquina de vectores de soporte: modelo de aprendizaje automático supervisado que resuelve problemas de clasificación entre dos grupos con los conjuntos dados de datos de entrenamiento etiquetados para cada categoría y es capaz de categorizar nuevos conjuntos.

3.2.8 suplantación de sitios web: conjunto de comportamientos maliciosos cuyo objetivo es recrear un sitio web bien conocido por el público o por un grupo de personas; el sitio web falso se utiliza entonces para abusar de la confianza de los visitantes y lograr objetivos maliciosos/ilegales, como el fraude, la invasión de la privacidad, etc.

4 Abreviaturas y acrónimos

En esta Recomendación se utilizan las siguientes abreviaturas y acrónimos:

DNS Servicio de nombre de dominios (*domain name service*)

GUI Interfaz gráfica de usuario (*graphical user interface*)

IA Inteligencia artificial

LSTM Memoria a largo-corto plazo (*long short-term memory*)

OCR Reconocimiento óptico de caracteres (*optical character recognition*)

PII Información de identificación personal (*personally identifiable information*)

QR Respuesta rápida (*quick response*)

SIFT Transformación de característica invariante (*scale-invariant feature transform*)

SURF Características robustas aceleradas (*speeded up robust features*)

SVM Máquina de vectores de soporte (*support vector machine*)

URL Localizador uniforme de recursos (*uniform resource locator*)

5 Convenios

En la presente Recomendación se utilizan los siguientes convenios:

La expresión "**se recomienda**" indica que se trata de un requisito recomendado pero que no es absolutamente obligatorio. Su cumplimiento no es indispensable para poder declarar la conformidad.

La expresión "**se tiene la opción de**" u "**opcionalmente**" indica que el requisito se permite, sin que ello signifique que se recomiende.

En el cuerpo de la presente Recomndación, ocasionalmente puede aparece la palabra "**puede**", en cuyo caso deben interpretarse como "**es capaz de**".

6 Análisis de la suplantación de sitios web

La suplantación de sitios web comprende un conjunto de comportamientos maliciosos cuyo objetivo es recrear un sitio web conocido por el público o por un grupo de personas. Los sitios web suplantados recopilan las credenciales de los 'visitantes' o difunden programas maliciosos para lograr objetivos maliciosos/ilegales, como el fraude, la invasión de la privacidad, etc.

A través de la suplantación de sitios web, los atacantes pueden causar daños graves de varias maneras. Por ejemplo, pueden provocar pérdidas económicas tanto para los clientes como para los operadores, dañar la reputación de los 'operadores'. Además, la información obtenida sobre los usuarios puede ser utilizada de forma indebida durante mucho tiempo, incluso aunque el sitio web suplantado desaparezca. Por tanto, es importante que las organizaciones de telecomunicaciones, y especialmente los operadores, adopten tecnologías útiles para luchar contra la suplantación de sitios web.

6.1 Caso de suplantación de sitios web

Normalmente, el proceso de creación de un sitio web suplantado para lograr un objetivo indebido se articula en torno a tres pasos, según se indica en la Figura 6-1.

– **Primer paso**: crear el sitio web
  - Primer paso: desarrollar y explotar un sitio web suplantado que sea idéntico o similar a algunos sitios web conocidos.

– **Segundo paso**: difundir el localizador de recursos uniforme (URL)
  - Segundo paso: difundir el sitio web suplantado utilizando diversos métodos, entre ellos el correo electrónico, el servicio de mensajes breves y los servicios de mensajería instantánea.

– **Tercer paso**: obtener credenciales o propagar programas maliciosos
  - Tercer paso: engañar a los visitantes para que introduzcan sus credenciales o descarguen programas maliciosos. Una vez que las víctimas han accedido al sitio web suplantado, pueden creer erróneamente que el sitio web en cuestión es el original e iniciar sesión con sus credenciales o descargar programas maliciosos.
  - Cuarto paso: extraer las credenciales de los 'visitantes' y preparar la siguiente actividad fraudulenta o maliciosa.

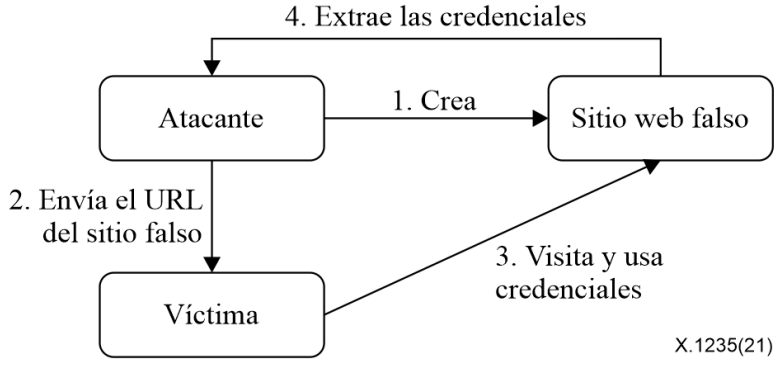

Figura 6-1 – Cuatro pasos para utilizar un sitio web suplantado a fin de obtener las credenciales de los usuarios

6.2 Características de la suplantación de sitios web

6.2.1 Características de los sitios web suplantados

Los sitios web suplantados utilizan técnicas similares para engañar a los visitantes y hacerles creer que están visitando los sitios originales, y presentan varias características comunes:

– Son visualmente similares: un sitio web suplantado puede ser una copia íntegra del sitio original, incluidos su contenido visual y su lógica de interacción.

– Toman prestados elementos visuales: un sitio web suplantado puede tomar prestados algunos elementos visuales importantes del sitio web original.

– Enlazan sitios web conocidos: un sitio web suplantado puede incluir enlaces a otros sitios web conocidos para recrear un sitio web asociado autorizado, un sitio web relacionado o una versión diferente del sitio web.

– Tienen un nombre de dominio similar: el nombre de dominio de un sitio web suplantado podría ser muy similar al del sitio original.

– Tienen un nombre de dominio oculto: a veces los atacantes utilizan servicios de acortamiento de URL o códigos de respuesta rápida (QR) para ocultar el nombre de dominio real de un sitio web al compartir un enlace.

– Tienen un nombre de dominio duplicado: cabe la posibilidad de reservar un conjunto de nombres de dominio para un sitio web suplantado para prolongar la vida del sitio web en cuestión. Varios de ellos se publicarán a la vez.

– Tienen el código fuente ofuscado: el código fuente de un sitio web suplantado puede ofuscarse para confundir a los programas de análisis de seguridad.

– Inducen a la introducción de credenciales: los sitios web suplantados pueden engañar a los visitantes para que introduzcan sus credenciales.

– Redirigen a sitios web conocidos: se tiene la opción de que un sitio web suplantado se redirija a diversos sitios web conocidos (por ejemplo, google.com) durante su tiempo de inactividad para eludir el escaneo automático de fraudes en línea.

– Tienen una barra de direcciones falsa: algunos sitios web suplantados pueden estar específicamente diseñados para navegadores móviles. El sitio web puede insertar una copia falsa de la barra de direcciones de un navegador móvil en la parte superior de la pantalla del teléfono móvil y mantener la copia bloqueada abusando de la función de la interfaz gráfica de usuario (GUI) del navegador móvil.

6.2.2 Características de actividades conexas

Para atraer o engañar a las personas y lograr que visiten los sitios web suplantados e interactúen con ellos, se puede recurrir a diversos métodos. Estos métodos pueden incluir una o varias de las actividades o tácticas que se enumeran a continuación:

– Difundir mensajes para dar a conocer la existencia del sitio web suplantado al mayor número de personas posible.

– Piratear o envenenar un servicio de nombres de dominios (DNS). De esa forma, todas las visitas al sitio web genuino se redirigirán al suplantado.

– Prometer beneficios a los usuarios (por ejemplo, cupones y regalos, etc.) para inducirles a hacer clic en enlaces a sitios web suplantados.

– Instalar un programa malicioso en el teléfono de la víctima para sustituir la solicitud del sitio web original por la del suplantado.

– Contactar con los usuarios imitando un servicio de atención al cliente y proporcionarles enlaces a sitios web suplantados.

6.3 Consecuencias

Un sitio web suplantado puede perjudicar a los usuarios y los propietarios de los sitios web de formas muy diversas. Entre sus posibles efectos perjudiciales figuran los siguientes:

– Pérdida de elementos propiedad de los usuarios: las credenciales y la información de identificación personal (PII) de los usuarios podrían filtrarse. Los estafadores pueden utilizar esa información para seguir cometiendo fraudes. En ese sentido, pueden engañar a los usuarios para que descarguen troyanos, códigos de minería de criptomonedas y otros programas maliciosos. También pueden tratar de controlar los terminales de los usuarios para robar cuentas financieras, difundir correo basura e infectar a otros usuarios. Estas actividades podrían dañar la batería o el funcionamiento de los terminales y entrañar pérdidas financieras directas e ingentes.

– Pérdida de elementos propiedad de los operadores: las transacciones normales entre los usuarios y los operadores podrían perderse e incluso verse restringidas durante un largo plazo.

– Pérdida de reputación: la reputación de los sitios web originales e incluso la confianza pública en el operador podrían verse afectadas, habría gran número de quejas de clientes y evaluaciones negativas a largo plazo.

7 Contramedidas

Las medidas para contrarrestar la suplantación de sitios web pueden dividirse en dos: las de detección y las de protección. Las primeras detectan los sitios web suplantados; véase la cláusula 7.1. Las segundas protegen a los usuarios para que no accedan a sitios web suplantados conocidos, o sean engañados por los mismos; véase la cláusula 7.2.

No es necesario que las organizaciones de telecomunicaciones utilicen simultáneamente todas las tecnologías recomendadas en esta cláusula. Dichas organizaciones deberían adoptar de forma flexible las tecnologías adecuadas tal como se describen en esta cláusula, en función de los datos de que dispongan, de su entorno legislativo y de las necesidades de los abonados, entre otros factores.

7.1 Detección

Partiendo de las características de los sitios web suplantados que se describen en la cláusula 6.2.1, se recomienda considerar las siguientes contramedidas:

7.1.1 Comparación de nombres de dominio similares

Se recomienda que los operadores mantengan una lista con los nombres de dominio de sitios web conocidos susceptibles de suplantación. Si un nombre de dominio es muy similar (pero no idéntico) a uno de los nombres de dominio de la lista, podría tratarse del nombre de dominio de un sitio web suplantado. Si el nombre de dominio ha sido acortado con ayuda de un servicio de acortamiento de URL, se recomienda que este se convierta en su forma URL original antes de efectuar la comparación.

Existen numerosos métodos para calcular la similitud de dos nombres de dominio, entre ellos: la distancia de edición, la similitud de Jaccard, la subsecuencia común más larga, la conversión de similitud visual, etc.

- **Método de la distancia de edición**: partiendo de dos nombres de dominio A y B, este método permite calcular el número mínimo de ediciones necesarias para convertir A en B. Cuanto menor sea la distancia de edición, mayor será la similitud.
- **Método de similitud de Jaccard**: este método consiste en dividir la cardinalidad de la intersección de dos conjuntos de caracteres de nombres de dominio por la cardinalidad de la intersección de su unión. Cuanto mayor sea la relación, mayor será la similitud.

- **Método de la subsecuencia común más larga:** este método consiste en calcular la longitud de la subsecuencia más larga común a dos nombres de dominio. Cuanto mayor sea la longitud, mayor será la similitud.

  NOTA – La subsecuencia común más larga hace referencia a la subsecuencia más larga común a dos secuencias dadas, cuyos elementos no tienen que ocupar posiciones consecutivas dentro de las secuencias originales. La subsecuencia común es una secuencia estrictamente ascendente de los índices de las dos secuencias dadas. Por ejemplo, si "abcde" y "akcve" son las dos secuencias dadas, entonces "ace" es la subsecuencia más larga común a las dos secuencias.

- **Método de conversión de la similitud visual:** este método consiste en sustituir los caracteres que son visualmente similares antes de efectuar la comparación. Por ejemplo, "0" y "o" pueden sustituirse entre sí; "1" e "i" pueden sustituirse entre sí, etc. Así, el nombre de dominio "z00.com" podría convertirse en "zoo.com". Este método podría mejorar el rendimiento de la comparación. No obstante, en ciertos casos, la conversión de la similitud visual también podría disminuir el rendimiento. Por lo tanto, sería prudente comparar la similitud antes y después de la conversión.

7.1.2 Detección de logotipos oficiales

Se recomienda comparar las imágenes u otros elementos visuales de un sitio web desconocido con los logotipos oficiales. Son logotipos oficiales las marcas de empresas u organizaciones, los diseños publicitarios y otros elementos icónicos. El método para detectar los logotipos oficiales comprende los siguientes pasos:

- Antes de proceder a la detección, se recomienda recopilar los logotipos oficiales y almacenarlos en una base de datos de logotipos.
- A continuación, los elementos visuales de la página web desconocida se descargan, o se hace una captura de pantalla de estos, y se comparan con los logotipos de la base de datos de logotipos. La similitud entre dos imágenes puede calcularse con ayuda de dos descriptores populares, que son la transformación de características invariante a la escala (SIFT) y las características robustas aceleradas (SURF).
- Si el código fuente de la página web detectada está ofuscado, la descarga de los elementos visuales podría resultar compleja. En ese caso, cabe la posibilidad de utilizar algún tipo de marco de pruebas de automatización de aplicaciones web basado en el navegador, para hacer una captura de pantalla de la página web. Antes de proceder a la comparación de logotipos, estos se pueden recortar de la captura de pantalla de la página web utilizando un modelo de detección de logotipos basado en IA. Este modelo de detección de logotipos basado en IA puede entrenarse adecuadamente utilizando bases de datos de logos de código abierto.

7.1.3 Detección de ofuscación de código

Los sitios web suplantados pueden ofuscar el código malicioso o de suplantación para dificultar su análisis. Un código ofuscado es muy diferente de un código normal. Numerosos métodos permiten detectar un código ofuscado, por ejemplo, los clasificadores basados en redes neuronales convolucionales y recurrentes, entre otros.

- Los clasificadores basados en redes neuronales convolucionales pueden extraer automáticamente las características de los *n*-gramas del código fuente normal y ofuscado. Un *n*-grama es una secuencia contigua de *n* palabras de un código fuente dado. El número de palabras puede establecerse en cinco o más, a fin de capturar íntegramente las características discriminativas para la clasificación.
- Los clasificadores basados en redes neuronales recurrentes se basan principalmente en redes de memoria a largo-corto plazo (LSTM) o en redes de transformación. Estos clasificadores tratan el código fuente como una secuencia de caracteres y detectan automáticamente los patrones de secuencia entre el código normal y el código ofuscado, para luego determinar si se trata o no de un código ofuscado.

7.1.4 Detección de introducción de credenciales

El código fuente de las páginas web suele incluir "formularios de entrada" para la introducción de credenciales. A su vez, el código fuente del formulario de entrada suele incluir un atributo de contraseña o "password" (que garantiza que no se muestre el contenido de la contraseña cuando el usuario la introduce). La información del atributo puede localizarse rápidamente utilizando expresiones regulares o herramientas de análisis de páginas web. Si el código fuente del sitio web está ofuscado, cabe la posibilidad de utilizar el método de reconocimiento óptico de caracteres (OCR) en capturas de pantalla del sitio web, para detectar la introducción de credenciales. El método OCR permite localizar las zonas de texto en una captura de pantalla de una página web y, a continuación, convertir las zonas de texto en textos. Si los textos convertidos contienen palabras clave como "contraseña", se detecta la introducción de credenciales en la página web.

7.1.5 Servicio de seguridad de terceros

Se recomienda que los operadores utilicen servicios de seguridad proporcionados por terceros proveedores de seguridad para obtener los atributos y las estadísticas de los sitios web, incluidos el tráfico, la reputación, las funcionalidades, el registro del nombre de dominio, la certificación y otra información de seguridad de los sitios en cuestión. Estos datos podrían utilizarse como información adicional a efectos de la detección de sitios web suplantados.

7.1.6 Análisis de enlaces

El sitio web suplantado puede copiar los hipervínculos del sitio web oficial para reutilizar los elementos visuales de dicho sitio en sus páginas web suplantadas. Este comportamiento puede detectarse mediante el análisis de enlaces. El análisis de enlaces consiste en recopilar los enlaces de los elementos visuales del sitio web oficial correspondiente y compararlos con los enlaces del sitio objeto de análisis. Si ambos sitios web tienen los mismos vínculos o comparten muchos de ellos, es posible que se trate de un sitio web suplantado.

7.1.7 Detección final basada en la puntuación de la reputación

Se recomienda que los operadores utilicen un modelo de decisión basado en inteligencia artificial (IA) para combinar todos los resultados de los análisis o procesos de detección incluidos en el marco de las contramedidas *supra* y decidir en términos generales si están ante un sitio web suplantado o no. En el modelo de decisión basado en IA, el peso de cada contramedida queda determinado automáticamente tras la fase de entrenamiento.

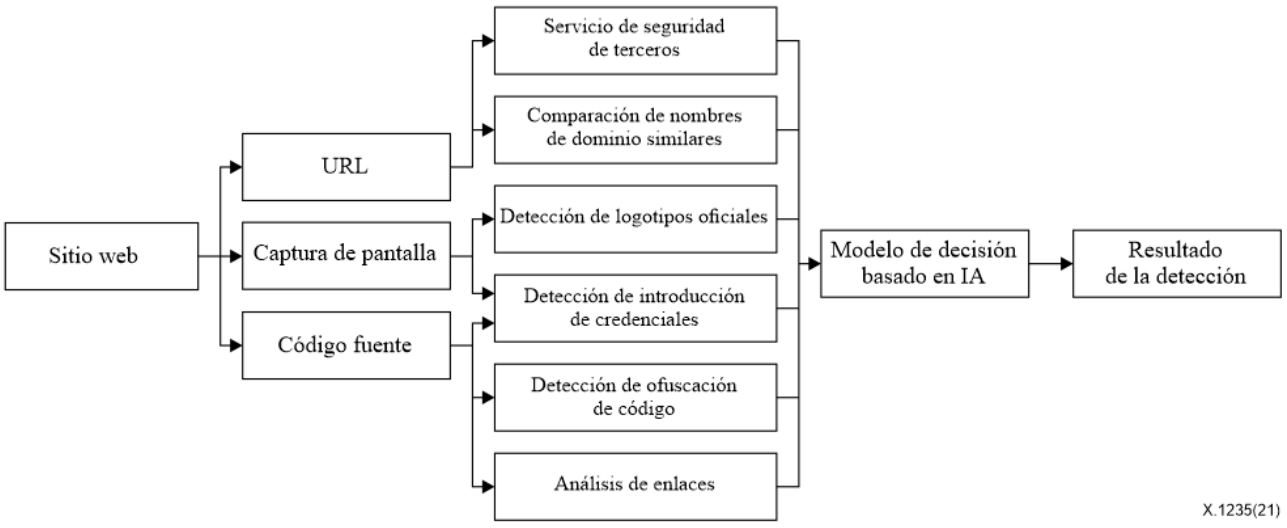

Figura 7-1 – Detección final basada en la puntuación de la reputación

Según se indica en la Figura 7-1, el sitio web se analiza a partir del URL, las capturas de pantalla y el código fuente. Para el URL de la página web, se calcula la puntuación relativa a la similitud del nombre de dominio y se comprueba el resultado de los servicios de seguridad de terceros. Para las capturas de pantalla, se detectan y analizan los logotipos oficiales y la introducción de credenciales. Para el código fuente, se detecta la introducción de credenciales y se realiza un análisis de enlaces. Para determinar si se trata de un sitio web suplantado, se recomienda que el modelo de decisión basado en IA tenga en cuenta todos esos aspectos.

A fin de entrenar dicho modelo de decisión basado en IA, es recomendable recopilar como muestras todos los sitios web oficiales que cabe proteger, los sitios web suplantados conocidos y otros sitios web normales. Todas las muestras (es decir, todos los sitios web recopilados) reciben una puntuación en función de las contramedidas recomendadas *supra*, tal y como ilustra la Figura 7-1, y cada muestra posee su propio vector de puntuación. Todos esos vectores de puntuación y los tipos (suplantados o no suplantados) de todas las muestras pueden utilizarse como datos para entrenar un clasificador que conforme el modelo de decisión basado en IA. Los clasificadores pueden estar basados en máquinas de vectores de soporte (SVM) o en redes neuronales profundas.

Sería de utilidad que revisores humanos verificasen los sitios web suplantados que se hayan detectado con ayuda del modelo de decisión basado en IA, para evitar errores.

7.2 Protección

Se recomienda que el operador proteja a los usuarios adoptando las siguientes contramedidas, previa autorización de dichos usuarios:

- Advertir a los usuarios: cuando un usuario trate de visitar un sitio web suplantado conocido, se recomienda que el operador retenga la petición, redirija al usuario a una página con un mensaje de advertencia del riesgo de suplantación y le pida que confirme la petición.
- Crear una lista de bloqueo: se recomienda al operador que cree una lista de bloqueo de sitios web suplantados, a fin de interrumpir todas las solicitudes relacionadas con los sitios incluidos en la lista. La lista de bloqueo podría resultar útil en las pasarelas de las redes del operador o en los servidores DNS del operador u otras entidades de red adecuadas.
- Evitar la propagación: se recomienda al operador que bloquee los mensajes interferentes que contengan enlaces a sitios web suplantados.
- Orientar a los usuarios para que sepan protegerse contra sitios web suplantados sospechosos: conviene recordar periódicamente a los usuarios los riesgos y las características de los sitios web suplantados e indicarles las mejores prácticas para evitarlos. Entre esas mejores prácticas figuran las siguientes:
  - No abrir enlaces extraños en correos electrónicos o mensajes.
  - Analizar cuidadosamente la información del nombre de dominio en la barra de direcciones y compararla con la del sitio web oficial.
  - Utilizar la función de seguridad del navegador web para comprobar la autenticidad del sitio web.
- Proteger a los usuarios descubriendo los distintos nombres de dominio activos de los mismos sitios web suplantados: un conjunto de nombres de dominio de un sitio web suplantado puede utilizarse para predecir otros nombres de dominio activos del mismo sitio web. Si el contenido de los sitios web a los que corresponden los nuevos nombres de dominio activos coincide con el de los sitios web suplantados conocidos, los primeros podrían tratarse como los mismos sitios web suplantados directamente. En estos casos, convendría tomar las medidas de protección anteriormente recomendadas.

8 Mecanismo

Se recomienda adoptar un mecanismo sistemático para contrarrestar la suplantación de sitios web, que combine todas las contramedidas descritas en la cláusula 7.

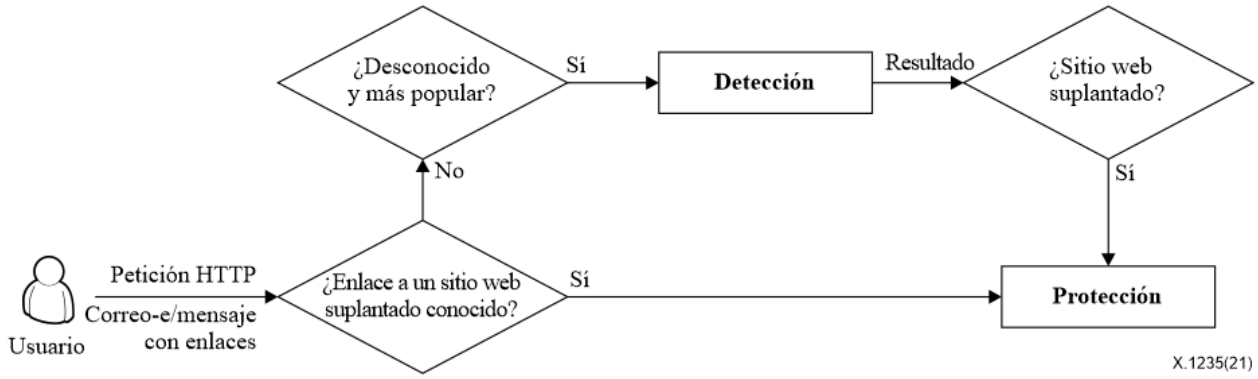

Figura 8-1 – Mecanismo para contrarrestar la suplantación de sitios web

1) Tal y como muestra la Figura 8-1, los URL pueden provenir de peticiones HTTP o de correos electrónicos o mensajes instantáneos de los usuarios (compartidos por terceros o autorizados por los usuarios).

2) Si el URL dirige a un sitio web conocido, conviene adoptar medidas de protección para gestionar la petición o el mensaje.

3) Si el URL dirige a un sitio web popular desconocido, conviene adoptar medidas de detección para identificar el tipo de sitio web se trata.

4) Si, una vez aplicadas las medidas de detección, resulta que se trata de un sitio web suplantado, conviene adoptar medidas de protección para evitar que el usuario acceda al mismo.

Apéndice I

Ejemplo de mecanismo para contrarrestar la suplantación de sitios web

(Este Apéndice no forma parte integrante de la presente Recomendación.)

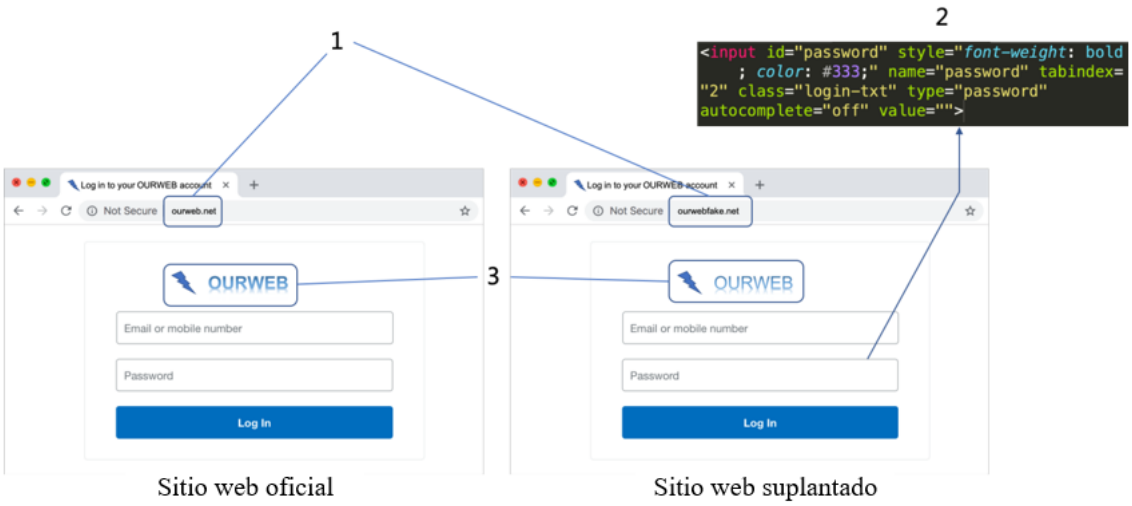

Figura I.1 – Ejemplo de un sitio web original y un sitio web suplantado

Supongamos que existe un sitio web muy conocido llamado OURWEB, cuyo nombre de dominio es "ourweb.net", y que un pirata informático intenta crear un sitio web suplantado, cuyo nombre de dominio es "ourwebfake.net". Supongamos también que el operador no conoce el nombre de dominio del sitio web suplantado. Entonces, cuando se producen un gran número de peticiones para visitar el sitio web suplantado, dicho sitio se convierte en un sitio web popular desconocido. En ese caso, el sitio web en cuestión debe someterse a las contramedidas de detección. Tal y como ilustra la Figura I.1, existen tres tipos de contramedidas que permiten detectar un sitio web suplantado.

1) Los nombres de dominio de los dos sitios web son similares. Supongamos que la puntuación relativa a la similitud, es decir la proporción de caracteres comunes, es 0,8.

2) Existe un mecanismo de introducción de credenciales en el sitio web suplantado, que puede detectarse a partir del código fuente. En consecuencia, la puntuación relativa a la introducción de credenciales es 1.

3) El logotipo del sitio web suplantado es similar al del sitio web oficial. Supongamos que la puntuación relativa a la similitud, es decir la proporción obtenida por el identificador visual descriptivo común (por ejemplo, un identificador descriptivo SIFT), es 0,9.

4) Supongamos que las puntuaciones obtenidas a raíz de las otras contramedidas equivalen a cero.

De esta forma, es posible componer un vector de puntuación [0,8; 0,9; 0; 1; 0; 0]. El primer valor del vector es la puntuación relativa a la similitud de los nombres de dominio; el segundo es la puntuación relativa a la similitud de los logotipos; y el cuarto es la puntuación relativa a la introducción de credenciales. Este vector puede introducirse en un clasificador SVM para obtener el resultado correspondiente. Si el resultado indica que se trata de un sitio web suplantado, este se someterá a una serie de contramedidas de protección para evitar que los usuarios accedan a él. Por ejemplo, convendría añadir el nombre de dominio del sitio web a una lista de bloqueo en las pasarelas de las redes de los operadores.

Apéndice II

Ejemplos de medidas técnicas

(Este Apéndice no forma parte integrante de la presente Recomendación.)

Ejemplo II.1: Comparación de nombres de dominio similares

Supongamos que el nombre de dominio oficial es "abc123.cn" y que el nombre de dominio del sitio web suplantado es "abc123cn.org". Al calcular la similitud de Jaccard (que consiste en dividir la cardinalidad de la intersección de dos conjuntos de caracteres de nombres de dominio por la cardinalidad de su unión) de los dos nombres de dominio, obtenemos una puntuación de 8/11=0,73.

Ejemplo II.2: Detección de logotipo oficial

La Figura II.1 muestra el sitio web oficial de OURWEB a la izquierda y el correspondiente sitio web suplantado a la derecha. Supongamos que el logotipo oficial de OURWEB se añade a la base de datos de logotipos. El logotipo del sitio web suplantado de la derecha es muy similar al logotipo oficial de OURWEB. En este caso, el algoritmo de comparación de similitudes permite calcular la puntuación relativa a la similitud de dichos logotipos (por ejemplo, 0,9).

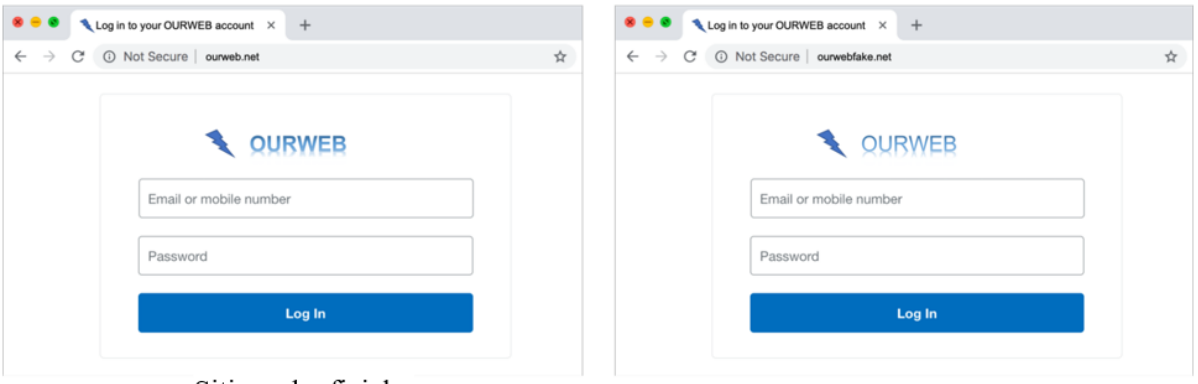

Figura II.1 – Un sitio web suplantado con un logotipo oficial

Ejemplo II.3: Detección de ofuscación de código

Algunos sitios web recurren a JavaScript para generar contenido web de forma dinámica y utilizan un servicio de ofuscación en línea (por ejemplo, https://obfuscator.io/) para codificar el código JavaScript.

Como se muestra en la Figura II.2, el código ofuscado difiere significativamente del código normal. En este caso, puede utilizarse un modelo de clasificación de texto basado en IA para predecir si el código está ofuscado. El algoritmo otorga al código fuente un valor de probabilidad de predicción (por ejemplo, 0,7).

Figura II.2 – Ejemplos de ofuscación de código JavaScript

Ejemplo II.4: Detección de introducción de credenciales

Como se muestra en la Figura II.1, el sitio web suplantado engaña a los usuarios para que introduzcan sus credenciales. La puntuación del sitio web puede depender de la existencia de un formulario para la introducción de credenciales en la página. De contar con dicho formulario, el sitio recibiría una puntuación de 1; en caso contrario, recibiría una puntuación de 0.

Ejemplo II.5: Servicio de seguridad de terceros

Los servicios de consulta de URL en línea de distintos proveedores de seguridad (por ejemplo, https://www.urlvoid.com/) pueden ayudarnos a determinar si estamos ante un sitio web suplantado. Si el sitio web detectado se clasifica como suplantado, la puntuación es 1, de lo contrario es 0.

Ejemplo II.6: Análisis de enlaces

En la Figura I.1, el logotipo puede hacer referencia a los recursos del sitio web oficial de OURWEB. Mediante el análisis de enlaces, puede determinarse si el recurso de imagen procede de ourweb.net (no de ourwebfake.net). Si se detectan enlaces a recursos de imágenes oficiales en el sitio web, la puntuación puede ser 1, de lo contrario puede ser 0.

Ejemplo II.7: Combinación de resultados

Las puntuaciones obtenidas a partir de los diferentes métodos de suplantación de sitios web se combinan en un vector. El Cuadro II.1 ilustra un ejemplo de puntuación de un sitio web. La segunda columna del cuadro conforma un vector de puntuación. Este vector puede introducirse en un clasificador para determinar si el sitio web es un sitio web suplantado.

Cuadro II.1 – Ejemplo de puntuación de un sitio web

Métodos de identificación	Puntuación
Comparación de nombres de dominio similares	0,8
Detección de logotipos oficiales	0,9
Detección de ofuscación de código	0,7
Detección de introducción de credenciales	1
Servicio de seguridad de terceros	0
Análisis de enlaces	1

Cuadro II.2 – Correspondencias entre los métodos de identificación y las características de los sitios web suplantados

Métodos de identificación	Características
Comparación de nombres de dominio similares	El nombre de dominio es similar
Detección de logotipos oficiales	Es visualmente similar Toma prestados elementos visuales
Detección de ofuscación de código	El código fuente ha sido ofuscado
Detección de introducción de credenciales	Introducción de credenciales
Servicio de seguridad de terceros	
Análisis de enlaces	Enlaza sitios web conocidos

Bibliografía

[b-UIT-T X.1126] Recomendación UIT-T X.1126 (2017), *Directrices para la mitigación de los efectos negativos de los terminales infectados en las redes móviles*.

[b-UIT-T X.1244] Recomendación UIT-T X.1244 (2008), *Características generales de la lucha contra el correo basura (spam) en aplicaciones multimedios basadas en IP*.

[b-SVM] Nature biotechnology, 2006, 24(12) 1565-1567: *What is a support vector machine?*

SERIES DE RECOMENDACIONES DEL UIT-T

Serie A	Organización del trabajo del UIT-T
Serie D	Principios de tarificación y contabilidad y cuestiones económicas y políticas de las telecomunicaciones/TIC internacionales
Serie E	Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos
Serie F	Servicios de telecomunicación no telefónicos
Serie G	Sistemas y medios de transmisión, sistemas y redes digitales
Serie H	Sistemas audiovisuales y multimedia
Serie I	Red digital de servicios integrados
Serie J	Redes de cable y transmisión de programas radiofónicos y televisivos, y de otras señales multimedia
Serie K	Protección contra las interferencias
Serie L	Medio ambiente y TIC, cambio climático, ciberdesechos, eficiencia energética, construcción, instalación y protección de los cables y demás elementos de planta exterior
Serie M	Gestión de las telecomunicaciones, incluida la RGT y el mantenimiento de redes
Serie N	Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión
Serie O	Especificaciones de los aparatos de medida
Serie P	Calidad de la transmisión telefónica, instalaciones telefónicas y redes de líneas locales
Serie Q	Conmutación y señalización, y mediciones y pruebas asociadas
Serie R	Transmisión telegráfica
Serie S	Equipos terminales para servicios de telegrafía
Serie T	Terminales para servicios de telemática
Serie U	Conmutación telegráfica
Serie V	Comunicación de datos por la red telefónica
Serie X	**Redes de datos, comunicaciones de sistemas abiertos y seguridad**
Serie Y	Infraestructura mundial de la información, aspectos del protocolo Internet, redes de próxima generación, Internet de las cosas y ciudades inteligentes
Serie Z	Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación

Impreso en Suiza
Ginebra, 2022