Índice general

Índice de figuras

Índice de tablas

CAPÍTULO 1

Antecedentes

1.1. Introducción

La seguridad informática ha sido y actualmente es un sector en el cual, empresas importantes de gran prestigio gastan cientos de millones de dólares para protegerse al estar conectados a una red [1], tal es la preocupación que a nivel mundial se registró una inversión en la seguridad informática de 75 billones de dólares en 2015 [2], mientras que, tanto chicas como medianas empresas suelen gastan un mínimo relacionado a este tema. En la actualidad, con la era de la revolución tecnológica por la que se está pasando, las empresas se han visto obligadas a contratar nueva tecnología para su producción, publicidad y/o servicios conectada al mundo de la Internet. Las empresas al estar conectadas a la Internet, están conectadas a millones de usuarios con cientos de posibilidades de acceso a los servicios de las empresas.

Cuando los usuarios se conectan a la red de Internet, están conectados todos los usuarios de la misma simultáneamente, esto conlleva un alto riesgo de inseguridad. Para tratar de disminuir el impacto provocado por amenazas informáticas, existen programas de computadora enfocados a detectar y proteger a los usuarios de la red contra los impactos provocados por las amenazas [3].

Hoy en día, existen diversos tipos de amenazas en la red, algunas son muy conocidas como virus informáticos que son diseñados para infectar archivos, pero no sólo existen ese tipo de amenazas diseñadas cierta forma con una actuación "automatizada". También existen las amenazas humanas o conocidos como Piratas informáticos, los cuáles, por medio de diversas técnicas de vulneración, pueden infectar, tomar el control o incluso obtener información o privilegios de computadoras o servidores con el fin de aprender o poner en práctica nuevas técnicas de vulneración, vender la información obtenida en el mercado negro o inclusive realizar un daño directo a los archivos o computadora objetivo [4][5].

Con el incremento de nuevos servicios web en Internet, se han creado y desarrollado diversos tipos de ataques hacia los servidores que proporcionan estos servicios. En los últimos años han ido en incremento los ataques web, aunque los que han tenido un mayor crecimiento y un gran impacto a nivel global, son los ataques tipo Cross-Site Scripting (de ahora en adelante XSS) [6].

1.2. Descripción del problema

Un ataque XSS ocurre cuando un atacante es capaz de inyectar un script, normalmente JavaScript, en la salida de una aplicación web de forma que se ejecuta en el navegador del cliente. Los ataques se producen principalmente por validar incorrectamente datos de usuario, y se suelen inyectar mediante un formulario web o mediante un enlace alterado.

Existen tres tipos de ataques XSS:

- XSS persistente o directo: este tipo de ataque consiste en embeber código HTML peligroso en sitios que lo permitan por medio de etiquetas *¡script¿* o *¡iframe¿*. Es la más grave de todas ya que el código se queda implantado en la web de manera interna y es ejecutado al abrir la aplicación web.

- XSS reflejado: en este tipo de ataque el código malicioso no queda almacenado en el servidor sino que se pasa directamente a la víctima. Es la forma más habitual de XSS. El ataque se lanza desde una fuente externa como un correo aparentemente inofensivo, un mensaje de chat u otro sitio web [8].

- XSS basado en DOM: es una variable de XSS persistente y reflejado. En un ataque XSS basado en DOM, la cadena maligna no es realmente analizada por el navegador de la victima hasta que el JavaScript legítimo de la página web es ejecutado. Estos códigos son ejecutados del lado del cliente, por lo que los filtros utilizados en el servidor no funcionan para este tipo de vulnerabilidades.

A la hora de lanzar un ataque de este tipo, los atacantes pueden utilizar varios tipos de inyección de código distinto. Los más utilizados son:

- Inyección en un formulario: se trata del ataque más sencillo. Consiste en inyectar código en un formulario que después al enviarlo al servidor, será incluido en el código fuente de alguna página. Una vez insertado en el código fuente, cada vez que se cargue la página se ejecutará el código insertado en ella.
- Inyección por medio de elementos: en este tipo de sistema de inyección de código se utiliza cualquier elemento que viaje entre el navegador y la aplicación, como pueden ser los atributos usados en las etiquetas HTML utilizadas en el diseño de la página.
- Inyección por medio de recursos: Aparte de los elementos en la URL y los formularios, hay otras formas en la que se puede actuar como son las cabeceras HTTP. Estas cabeceras son mensajes con los que se comunican el navegador y el servidor. Aquí entran en juego las *cookies*[1] y las sesiones [9].

Los daños potenciales que pueden causar un ataque XSS, pueden afectar tanto a los servidores en donde está contenida la aplicación web o pueden provocar serios problemas para el usuario final, éstos pueden varían en el grado de impacto, pueden ir desde una molestia para el usuario hasta un compromiso completo de la cuenta del mismo. Uno de los efectos más graves de los ataques XSS implica la divulgación de cookies de sesión del usuario, lo que permite a un atacante secuestrar la sesión del usuario y tomar control total de la cuenta. Otros ataques dañinos incluyen la divulgación de los archivos de los usuarios finales, la instalación de programas dañinos para el equipo del usuario final, redirigir al usuario a otra página o sitio web con fines malicioso, o modificar la presentación de los contenidos [10]. Los ataques XSS explotan vulnerabilidades no en el navegador

[1] Una cookie es un pequeño elemento de información que un servidor Web envía al navegador al visitar ciertas páginas web y que ambos comparten cada que este navegador vuelve a visitar [7].

del usuario, sino en las aplicaciones Web de terceros a las que accede el usuario. En este tipo de ataque el navegador no puede distinguir entre el contenido que un usuario haya podido incluir en una petición Web, y el contenido inyectado a través de un ataque XSS [11].

1.3. Solución propuesta

El presente Trabajo Terminal implementa un sistema que mediante reconocimiento de patrones, aprendizaje máquina y minería de datos; sea capaz de analizar y clasificar texto en el idioma español a partir de un conjunto de datos estadísticos. Aplicado al caso de estudio *online grooming*.

1.4. Estado del arte

El interés por desarrollar sistemas que apoyan el análisis de textos se ha incrementado. En la tabla 1 se mencionan algunos sistemas que hacen uso de estos aplicado a diferentes problemáticas sociales.

Nombre	Descripción	Tipo	Año	Lugar de desarrollo
ChildDefence	La aplicación ChildDefence permite recolectar conversaciones (escritas en inglés) que un niño ha tenido con una persona, provenientes de mensajes de texto o chats en línea, y luego hacer que se analicen en el propio teléfono. La aplicación, ha sido entrenada para identificar si la persona es un niño o un adulto.	Gratuita	2011	Isis Forensics
LIWC	Es un programa de computadora analizador de textos, que determina el porcentaje de emociones expresadas en un texto, con base en diversas categorías.	Comercial	2001	Austin Texas University, EUA
KidsWatch	Es un programa de control parental. Funciona de tal manera que cuando en la conversación de un chat encuentra palabras referentes al sexo, drogadicción, armas, suicidio entre otros se les avisa a los padres acerca del comportamiento sospechoso.	Comercial	2002	Computer Business Solutions

Tabla 1.1: Comparativa de Sistemas que realizan análisis de textos

Dentro de la Escuela Superior de Cómputo, también se encontró un Trabajo Terminal que implementa procesamiento de lenguaje natural y minería de datos para la toma de decisiones. La descripción se encuentra en la tabla 1.

1.5. Objetivo

1.5.1. Objetivo general

Nombre	Descripción	Tipo	Año	Lugar de desarrollo
Prototipo de sistema de información con minería de datos para la toma de decisiones. (20060106)	Sistema con una adaptación de minería de datos dirigida hacia las PyMES, obteniendo información relevante de sus bases de datos mostrando interpretaciones de los resultados de la minería en enunciados de lenguaje natural.	Trabajo Terminal	2006	ESCOM

Tabla 1.2: Trabajos Teminales relacionados con el procesamiento de lenguaje natural

Desarrollar un sistema que sea capaz de analizar y clasificar texto en diferentes categorías a partir de un conjunto de datos estadísticos mediante: reconocimiento de patrones, aprendizaje máquina y minería de datos. A fin de poder clasificar conversaciones como peligrosas o no peligrosas.

1.5.2. Objetivos específicos

Desarrollar:

- Aplicación para el intercambio de conversaciones.
- API para el análisis de conversaciones.
- Aplicación para el procesamiento de texto.
- Aplicación que funcione, como conexión entre los módulos del sistema.
- Un set de pruebas de desempeño del sistema para verificar su eficiencia.

1.6. Justificación

La llegada de internet abrió las puertas a grandes posibilidades de comunicación: redes sociales, foros, chats, etcétera. El término *grooming* hace referencia a las acciones que lleva acabo un adulto para establecer amistad con un menor por medio de internet, con el objetivo de obtener una satisfacción sexual mediante la obtención de imágenes con contenido erótico o sexual del menor. A pesar de que estas situaciones tienen su origen dentro de la red, muy frecuentemente terminan en abuso físico a menores o tráfico de pornografía infantil.

Los problemas que se presentan para atacar el grooming son principalmente: La inocencia de los menores, el anonimato en el que se mantienen los adultos implicados y la facilidad con la que se puede acceder a internet hoy en día.

La detección de una posible amenaza hacia un infante podría realizarse de manera manual, donde un padre de familia tenga acceso a todos los mensajes que se comparten en un canal. Sin embargo, esta detección es poco factible ya que el análisis de los mensajes podría ser tardado dependiendo del volumen de la información.

Es por ello que este sistema pretende actuar como una herramienta capaz de automatizar el análisis de dichas conversaciones y con ello hacer uso del sistema implementado.

1.7. Caso de estudio

1.7.1. Grooming

Child grooming o internet grooming es un término que hace referencia a una serie de acciones o conductas deliberadamente emprendidas por un adulto con el objetivo de ganarse la confianza de un menor de edad haciendo

uso de internet. El adulto intenta crear una conexión emocional con el infante con el propósito de disminuir las inhibiciones del niño y posiblemente llegar a un abuso sexual. En casos severos se puede inducir al menor al mundo de la prostitución infantil o a la producción de material pornográfico. [**?**]

5

Bibliografía

[1] Kaplan J., Sharma S. & Weinberg A. (2011). "Meeting the cybersecurity challenge". McKinsey & Company. Recuperado 12 Marzo 2017, de http://www.mckinsey.com/business-functions/business-technology/our-insights/meeting-the-cybersecurity-challenge

[2] THE EDITORS AT CYBERSECURITY VENTURES. (2014). "The Cybersecurity Market Report covers the business of cybersecurity, including market sizing and industry forecasts, spending, notable M&A and IPO activity, and more..Çybersecurity Ventures. Recuperado 26 Septiembre 2016, de http://cybersecurityventures.com/cybersecurity-market-report/

[3] King, S. (2016). "Assessing the real risk of being online". ComputerWeekly. Recuperado 26 Septiembre 2016, de http://www.computerweekly.com/feature/Assessing-the-real-risk-of-being-online

[4] Computer Hope (2016). "Why do people hack computers?". Computerhope.com. Recuperado 26 Septiembre 2016, de http://www.computerhope.com/issues/ch001530.htm

[5] Cloudbric. (2016). "6 Reasons Why Hackers Want to Hack Your Website". Recuperado 26 Septiembre 2016, de https://www.cloudbric.com/blog/2015/10/6-reasons-why-hackers-want-to-hack-your-website/

[6] Imperva Inc. (2016). "2015 Web Application Attack Report (WAAR)". WAAR 2015. Recuperado de https://www.imperva.com/docs/HII_Web_Application_Attack_Report_Ed6.pdf

[7] Gutierrez, E. (2009). "JavaScript". 1st ed. Barcelona: Ed. ENI, p.233.

[8] Assis, R. (2016). "Primero post de la serie sobre vulnerabilidades XSS". Sucuri Español. Recuperado 19 Diciembre 2016, de https://blog.sucuri.net/espanol/2016/04/pregunte-sucuri-que-es-una-vulnerabilidad-xss.html

[9] (2016). "Qué es y cómo funciona un ataque Cross - Site Scripting". Hostalia. Recuperado 19 Diciembre 2016, de http://pressroom.hostalia.com/wp-content/themes/hostalia_pressroom/images/cross-site-scripting-wp-hostalia.pdf

[10] Ramos Pereira, K. (2016). Çross-Site Scripting". Revistasbolivianas.org.bo. Recuperado 20 Noviembre 2016, de http://www.revistasbolivianas.org.bo/scielo.php?pid=S1997-40442013000100023&script=sci_arttext

[11] (2014). "Su navegador esta desnudo: por qué los navegadores protegidos siguen siendo vulnerables.". Panda Security. Recuperado 27 Noviembre 2016, de http://resources.pandasecurity.com/enterprise/solutions/7.%20WP%20PCIP%20ESP%20Su%20Navegador%20esta%20desnudo.pdf

[12] Greene, T. (2016). "8 cyber security technologies DHS is trying to commercialize". Network World. Recuperado 26 Septiembre 2016, de http://www.networkworld.com/article/3056624/security/8-cyber-security-technologies-dhs-is-trying-to-commercialize.html

[13] (2016). "School of Computer Science and Information Technology University of Nottingham". Firewalls, Intrusion Detection Systems and Anti-Virus Scanners (p. 57). NOTTINGHAM NG8 1BB, UK. Recuperado de http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.2262&rep=rep1&type=pdf

[14] Mookhey , K. K., Nilesh, B. (2011). "Detection of SQL Injection and Cross-site Scripting Attacks — Symantec Connect". Symantec.com. Recuperado 26 Septiembre 2016, de http://www.symantec.com/connect/articles/detection-sql-injection-and-cross-site-scripting-attacks10

[15] Tim, K. (2016). "Strategies to Reduce False Positives and False Negatives in NIDS — Symantec Connect". Symantec.com. Recuperado 26 Septiembre 2016, de http://www.symantec.com/connect/articles/strategies-reduce-false-positives-and-false-negatives-nids

[16] Garcia-Alfaro, J. & Navarro-Arribas, G. (2005). "Prevención de ataques de Cross-Site Scripting en aplicaciones Web".Recuperado 25 Noviembre 2016, de http://www-public.tem-tsp.eu/~garcia_a/web/papers/recsi08-xss.pdf

[17] González Márquez, V. (2009). "Sistema de detección de intrusos basado en sistema experto (Tesis de maestría)". Centro de Investigación en Computación. México.

7