Introducción

Para nuestros tiempos las tecnologías de la información han alcanzado niveles importantes de capacidad y funcionalidad; se han aplicado en numerosas ocaciones y al mismo tiempo han existido aciertos y errores en sus implementaciones. Los programadores y diseñadores de sistemas digitales y lógicos van cada día más adelante y no paran de desarrollar cosas nuevas e innovadoras. Es interesante e indispensable pensar por un momento si todos estos avances han traído únicamente beneficios. Con seguridad no se puede contar con alguna respuesta totalmente acertada. Sin embargo, la realidad se sustenta en la cantidad de problemas que enfrentan los usuarios de computadoras alrededor del mundo por causa de una ineficiente calidad de programación en los productos de cómputo y por la carencia de conocimientos de estos usuarios en materia de contenidos electrónicos.

Es entonces cuando, probablemente, surge un término muy utilizado en el ámbito computacional "virus informáticos", los cuales resultan equivalentes a las enfermedades o malestares de millones de equipos de cómputo en el planeta. Actualmente a los virus se les ha clasificado en general con un nombre estándar, Malware, proveniente de la contracción de “Malicious Software” y definido como todo aquel código malicioso que lleva a cabo una actividad del mismo carácter y con fines de obtener información, controlar los recursos de cómputo e incluso llegar a paralizar infraestructuras completas.

Es muy común escuchar en las noticias sobre un nuevo gusano informático, sobre un nuevo virus de computadora o acerca de una nueva técnica de engaño donde se ve involucrado un código malicioso. Por la cantidad de nuevos códigos maliciosos que aparecen día a día, resulta inaceptable tener a grupos numerosos de investigadores analizando cada una de esas muestras. Por lo tanto surge la idea y la necesidad de crear una herramienta que ayude a los analistas, proporcionándoles de manera automatizada y confiable, información relevante sobre la infección, cómo pueden ser su comportamiento, actividad maliciosa y la razón de llevarla a cabo. Como

resultado se complementa su labor de informar a los usuarios de computadoras para ayudarlos a prevenir la infección de sus equipos.

Este informe se compone de tres principales secciones. La primera “Antecedentes Teóricos”, detalla todos los conocimientos teóricos que previamente debe poseer el lector para entender el resto del mismo. La segunda “Análisis y metodología”, describe las herramientas y técnicas utilizadas para alcanzar el objetivo. Finalmente la sección de “Diseño y Desarrollo”, explica como se construyó cada una de las herramientas utilizadas y desarrolladas.

Objetivo

Desarrollar los programas, técnicas forenses y de análisis de tráfico que logren entregar un reporte sustancial y condensado acerca del comportamiento de códigos maliciosos. Lo anterior mediante el desarrollo, instalación e implementación de una red de arena para lograr un análisis integral, completo y automatizado de estos. Finalmente demostrar que con una implementación básica se obtiene un resultado suficiente para observar el comportamiento del software malicioso.

Justificación

Desde hace algunos años en diferentes instituciones dedicadas a la seguridad informática se han ocupado en la tarea de recolectar muestras de malware y posteriormente analizarlas en ambientes controlados con el objetivo de conocer la carga maliciosa que implican (en inglés, payload). El escenario de análisis utilizado en muchas ocasiones ha sido mediante máquinas virtuales; debido a características especiales descritas más adelante, resultaron ser una solución muy cómoda y funcional por algún tiempo. Sin embargo, los atacantes comenzaron a perfeccionar sus técnicas para evadir los ambientes virtuales, es decir, comenzaron a implementar tácticas para detectar cuando la ejecución de su binario o archivo ejecutable se realizaba sobre una máquina virtual. Poniendo así a prueba, la efectividad de los laboratorios construidos con base en ambientes virtuales para el análisis efectivo de malware.

De acuerdo con el párrafo anterior, la alternativa más aplicable que resta para el análisis de códigos maliciosos actuales y sofisticados se reduce a infectar un equipo real (no virtual), para así obtener la información de sus procesos, archivos y conexiones con los que trabaja. Lo que se propone en este informe es exactamente lo anterior; con la finalidad de que el agente malicioso

no se inhiba y nos muestre toda la información posible y de utilidad para conocer su comportamiento.

En nuestros días, y ya desde hace algunos años, el malware presenta un patrón de comportamiento independiente del tipo de código malicioso del que se trate. La gran mayoría tiene una técnica de infección que puede variar entre ingeniería social hasta la explotación de vulnerabilidades de software que se encuentran en las computadoras. También se valen de una manera para mantener el acceso como pueden ser puertas traseras o conexiones a servidores IRC. Finalmente, desde luego, llevan a cabo una actividad maliciosa y por lo regular consiste en la recolección de datos sensibles como contraseñas, cuentas de correo, cuentas de banco, datos personales, etcétera; y probablemente concluir en la utilización de esos recursos de cómputo para causar estragos en las redes de datos, como negaciones de servicio generalizadas, es decir, las infecciones de actualidad no sólo quieren introducirse en los equipos y hacer algo molesto para complicarle la existencia al usuario, como el típico virus de la pelotita en la pantalla; sino que buscan llevar a cabo algo que les deje beneficios (muchas veces económicos como el caso del phishing, el pharming y las botnets) y se valen de todo lo que encuentren a su alcance como archivos, conexiones, dispositivos, servicios, vulnerabilidades, etcétera.

A continuación se citan y describen dos soluciones que satisfacen la tarea del análisis dinámico de malware (estos dos conceptos serán más detallados en partes posteriores de este informe).

Caja de arena es un ambiente real o virtual controlado donde el malware se puede ejecutar y posteriormente se obtiene información acerca de su actividad maliciosa en el sistema de archivos, procesos, peticiones de red, etcétera. Limitándose a una sola entidad.

Una red de arena es un ambiente real o virtual controlado donde existe más de un equipo conectado a una red de datos aislada. En una configuración mínima se cuenta con el equipo a infectarse, un servidor que levanta servicios simulados y captura todo el tráfico de red generado por el equipo infectado. Con ello se pretende obtener información que va más allá de los cambios en el sistema de archivos o llaves de registro, como pueden ser los datos de una sesión IRC (servidor, puerto de conexión, usuario, contraseña, canal, etcétera) la cual, dicho agente malicioso utiliza para descargar más software malicioso o esperar instrucciones de algún controlador.

Acorde con lo anterior, el concepto de realizar una caja de arena - lo cual sería una idea correcta para objetivos menores- no es suficiente para obtener un análisis completo del malware en la

actualidad. Como una solución más adecuada, aunque compleja, se propone la implementación de una red arena que permita obtener mucho más información que con una simple caja de arena. Lo anterior permitiría conocer información sobre las conexiones que se establecen con el exterior, las peticiones que se realizan y a dónde las hace; sin mencionar que también analizaría el sistema de archivos y el registro. Reuniendo todos esos argumentos se puede garantizar que entregará un reporte con información muy relevante y que seguramente será utilizada en el futuro para la detección, prevención y mitigación de infecciones y ataques.

Sede Laboral

El Departamento de Seguridad en Cómputo DSC/UNAM-CERT pertenece a la Dirección General de Servicios de Cómputo Académico (DGSCA, fig. I.1). La cual se ubica físicamente en el Circuito Exterior S/N Frente a la Facultad de Contaduría y Administración, Delegación Coyoacán, C.P. 04510, México D.F.

Fig. I.1. Fachada exterior de la DGSCA.

El Departamento de Seguridad en Cómputo (DSC) nace en el año 1999, y sus antecedentes fueron el Equipo de Seguridad en Cómputo (ESC) y el Área de Seguridad en Cómputo (ASC), los cuales datan desde 1994. Surge bajo la necesidad de contar con un equipo de expertos en seguridad informática que pudiera satisfacer el requerimiento en esta materia que la DGSCA y la misma Universidad tenían. En el año de 2001 obtiene el reconocimiento ante FIRST (Forum of Incident Response and Security Teams) de CERT (Computing Emmergency Response Team, o bien Equipo de Respuesta a Incidentes en Cómputo). Este título lo sigue manteniendo hasta la fecha y es distintivo de la labor de atender las emergencias en cómputo a lo largo y ancho de la red de cómputo universitaria.

En la actualidad esta organización desarrolla, mantiene y desempeña proyectos muy importantes en el campo de la seguridad sobre tecnologías de la información. Por mencionar algunos se cuenta con los proyectos honeynet, usuario casero, malware, entre otros.

A partir de noviembre de 2008 comencé a trabajar para el DSC/UNAM-CERT. He participado en varios proyectos que el departamento ha venido realizando. Sin embargo, he tenido especial presencia en el desarrollo de un analizador dinámico y automático de códigos maliciosos. Este proyecto lo realicé en un trayecto de seis meses. En el proceso me encontré con una gran cantidad de problemas, así como funcionalidades nuevas por agregarle. En el desarrollo de este informe describiré el fundamento teórico y técnico de esta herramienta, el proceso seguido para llegar al resultado final obtenido, la metodología que justifica su funcionamiento y propondré una liberación final para la misma.

Los sitios oficiales de la DGSCA y el DSC/UNAM-CERT son los siguientes:

- http://www.dgsca.unam.mx/
- https://www.seguridad.unam.mx/