Seguridad en Servicios Web

Edgardo Bernardis, Hernán Bernardis, Mario Berón, Germán Montejano

Departamento de Informática
Facultad Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis

Ejército de los Andes 950 – San Luis – Argentina
{ebernardis, hbernardis, mberon, gmonte}@unsl.edu.ar

Resumen

Con el auge de internet y las distintas dinámicas de la sociedad actual, ha cambiado en gran medida la forma de interactuar entre las personas y las empresas. Este cambio notable se observa en la forma de intercambiar información entre los distintos actores. Este intercambio se vuelve de particular interés siendo blanco de ataque por parte de todos aquellos actores que quieren obtener información útil y valiosa a sus propios intereses o de terceros. Es aquí donde cobra particular relevancia implementar todo tipo de medidas y acciones tendientes a evitar estos ataques, por tal motivo surge lo que se denomina Seguridad Informática.

En este artículo se describe una línea de investigación cuyo principal objetivo es el desarrollo de métodos, técnicas y estrategias orientadas a incrementar el nivel de seguridad de Servicios Web.

Palabras clave: Aplicaciones Web, Servicios Web, Seguridad Informática, Seguridad de la Información.

Introducción

Con los avances de la tecnología, sobre todo en el ámbito de internet, se vuelve sumamente importante y necesario la protección de todo tipo de información. En la actualidad, es realmente alta la cantidad de delitos que se llevan adelante en contra de información personal o de empresas. Todo tipo de información es valiosa, ya sea desde simples datos personales hasta sistemas y bases de datos empresariales.

Con el auge de internet, el intercambio de archivos se ha vuelto un punto esencial en la sociedad actual. Los consumidores intercambian información no sólo entre ellos sino también con los vendedores. Para el intercambio de información se utilizan diferentes medios entre los que se pueden mencionar redes sociales, correo electrónico, sistemas punto a punto, pagos online, juegos. Todo esto se fundamenta en la confianza y el correcto funcionamiento del software y del hardware subyacente a dicho proceso.

Por lo antes mencionado es que surge lo que se conoce como *Seguridad Informática* (SI). Existen diversas definiciones de SI, algunas más extensas que otras, pero todas enfocadas en los mismos aspectos comunes. A los fines de este trabajo, se adhiere a la siguiente definición de SI: *Preservación de la confidencialidad, integridad y disponibilidad de la información en el Ciberespacio. A su vez, el Ciberespacio se define como el entorno complejo que resulta de la interacción de las personas, software y servicios en Internet por medio de redes y dispositivos tecnológicos conectados a el, y que no existe en ninguna forma física* [1].

La información es un conjunto organizado de datos, que cambia su enfoque y su estado de conocimiento dependiendo del ámbito en

la que se la utilice. Por ejemplo, si la información se conceptualiza bajo el punto de vista de la ingeniería: *Estudio de las características y estadísticas del lenguaje que permite su análisis desde un enfoque matemático, científico y técnico*.
Desde el punto de vista de una empresa: *Conjunto de datos propios que se gestionan y mensajes que se intercambian personas y/o máquinas dentro de una organización* [2].

La información se ve afectada por muchos factores, motivo por el cual se vuelve importante su seguridad. De aquí que *Seguridad de la Información* es: *una disciplina, cuyo principal objetivo es mantener el conocimiento, datos y sus significados libres de eventos indeseables, tales como el robo, espionaje, daños, amenazas y otros peligros. La Seguridad de la Información incluye todas las acciones tomadas con anticipación, para evitar eventos no deseados* [3].

El objetivo de la SI es obtener un nivel aceptable de seguridad, entendiéndose por aceptable un nivel de protección suficiente para que la mayor parte de potenciales intrusos, interesados en los equipos con información de una organización o persona, fracasen en cualquier intento de ataque contra los mismos. Asimismo, se encarga de establecer los mecanismos para registrar cualquier evento fuera del comportamiento normal y tomar las medidas necesarias para re-establecer las operaciones críticas a la normalidad [4].

El punto o centro de ataque a la seguridad informática se da en una *Vulnerabilidad*: *debilidad de un activo o control que puede ser explotado por una o más amenazas* [5]. La presencia de una vulnerabilidad no puede causar daño en sí misma, ya que es necesario que exista una amenaza que la aproveche. Una vulnerabilidad que no tiene una amenaza, puede no requerir la aplicación de un control, pero debe ser reconocida, supervisada por si tiene cambios y, en lo posible, eliminada.

Las amenazas surgen a partir de la existencia de vulnerabilidades, es decir que una amenaza sólo puede existir si existe una vulnerabilidad que pueda ser aprovechada, independientemente de que se comprometa o no la seguridad de un sistema de información. Una amenaza se puede definir como: *Cualquier elemento o acción que es capaz de aprovechar una vulnerabilidad y comprometer la seguridad de un sistema de información* [6]. Las amenazas se pueden clasificar o dividir en dos tipos; las intencionales, en caso de que deliberadamente se intente producir un daño (por ejemplo el robo de información). Las no intencionales, en donde se producen acciones u omisiones de acciones que si bien no buscan explotar una vulnerabilidad, ponen en riesgo los activos de información y pueden producir un daño (por ejemplo las amenazas relacionadas con fenómenos naturales).

Dentro de las diferentes áreas y enfoques que se pueden llevar adelante, en lo que respecta a seguridad informática, existe uno particularmente interesante y de gran crecimiento como lo son las aplicaciones que ejecutan en la web. Estas aplicaciones son de lo más utilizado en la web por parte de los usuarios, debido a su versatilidad, beneficios, fácil utilización, comodidad, etc. Existe una gran variedad y cantidad de las mismas; desarrolladas en diferentes tecnologías y lenguajes de programación. Todas contienen información formal (código fuente) e informal (identificadores, comentarios, documentación, etc.). Es a partir de estos tipos particulares de información que se puede detectar y medir el nivel de comprensibilidad de la misma y a través de esto, aumentar o disminuir su seguridad.

La organización de este artículo se expone a continuación. La sección 2 describe la línea de investigación y desarrollo abordada. La sección 3 presenta los resultados obtenidos hasta el momento, junto con todos aquellos esperados a corto plazo. Finalmente, la sección 4 describe las tareas realizadas por los recursos humanos en formación.

Líneas de Investigación y Desarrollo

En las sub-secciones siguientes se describen de manera concisa la línea de investigación presentada en este trabajo.

Técnicas para Medir el Nivel de Comprensibilidad de las Aplicaciones Web

Una de las formas de analizar si un software es o no seguro, se puede llevar a cabo a través de la estimación de su nivel de comprensibilidad. Entendiendo tal concepto como una medida que establece a priori la facilidad de entender la tarea que un sistema de software lleva a cabo. Es de suponer que mientras más fácil sea de comprender más fácil será de vulnerar. Por lo antes mencionado, se estudian métodos y estrategias que permitan medir la comprensibilidad con alto nivel de precisión.

En este contexto, el grupo de investigación está centrado en la elaboración de estrategias que permitan medir a priori o a posteriori el nivel de comprensibilidad de Servicios Web (SW). Dicho enfoque se lleva a cabo debido al crecimiento, interoperabilidad y ventajas que ofrecen los SW, lo que los vuelve un blanco atractivo por aquellos actores interesados en robar o manipular la información valiosa que pueden contener.

Visualización de Vulnerabilidades y partes Comprensibles

No solo basta con detectar los puntos vulnerables y comprensibles de un sistema de software sino que se deben visualizar de la manera adecuada. Esta tarea es necesaria ya que cuando se analizan sistemas de gran envergadura, la información extraída puede ser muy grande y compleja de entender. Por consiguiente, si no se plantea apropiadamente la forma de visualizarla, es decir, presentada de forma sintetizada y con extrema facilidad de análisis será sumamente complejo examinarla y por lo tanto, aplicar las estrategias tendientes a mejorar e incrementar el nivel de seguridad de los puntos vulnerables será todo un desafío. En este contexto, los estudios se centran en la generación de visualizaciones innovadoras que permitan disminuir la brecha existente entre la información extraída de las aplicaciones y la estructura de conocimiento del programador.

Aproximaciones para Incrementar la Seguridad de las Aplicaciones Web

Como se mencionó anteriormente, toda aplicación web está conformada por distintos tipos de información, tanto formal como informal. Al analizar detalladamente y estimar su nivel de comprensibilidad, permitirá detectar las partes más entendibles, lo que las vuelven más vulnerables para su manipulación y/o robo de información. En este punto, es posible definir estrategias que permitan subsanar las vulnerabilidades y proteger las partes del software que sean susceptibles de ataques.

En este contexto, se toma como base aproximaciones basadas en el análisis del código de las aplicaciones y las transformaciones correspondientes que permitan incrementar el nivel de seguridad.

Existen diversas formas para poder analizar y extraer información de un código fuente. En el caso de esta línea de investigación, se utilizan las técnicas de compilación tradicionales [16] que usan la representación de Árbol de Sintaxis Abstracta decorado para representar el código fuente. De manera simplificada, como primer paso se utiliza un Analizador Lexicográfico (Lexer), que toma el código fuente y lo divide en fragmentos o partes denominadas tokens [7]. Estas partes o tokens son la entrada del Analizador Sintáctico (Parser), el cual tiene dos funcionalidades principales: Verificar si la especificación del Programa no tiene errores sintácticos y realizar acciones semánticas para llevar a cabo actividades tales como: recolección de información específica, control, transformación de código, etc. La salida del parser es un Árbol de Sintaxis Abstracta sobre el cual se pueden aplicar diferentes recorridos para recuperar, efectivamente, la información compleja del sistema. Una vez obtenida la información requerida y de acuerdo al grado de comprensibilidad de las mismas, se realizan las modificaciones y/o transformaciones necesarias, según los métodos elegidos, para mejorar la seguridad de la aplicación al volver menos comprensible dicha información.

Resultados Obtenidos/Esperados

Hasta el momento se han llevado a cabo las siguientes tareas:

- Se desarrolló una métrica que posibilita medir a priori la comprensibilidad de aplicaciones. Esta métrica utiliza un método multicriterio que permite que el ingeniero de software pueda volcar su experiencia en el dominio de la aplicación. Experiencia que beneficiará en la precisión del cálculo del nivel de comprensibilidad.
- También, como parte del proceso de validación, se desarrolló un prototipo que implementa la métrica antes mencionada y en la cual se llevaron a cabo pruebas para validar los resultados. Además, como parte de la herramienta, se agregaron diagramas de barras e iluminación de código fuente para visualizar el nivel de comprensibilidad del mismo. Todas las pruebas se aplicaron sobre Servicios Web especificados en WSDLs (Web Services Description Language), esto es debido a la simplicidad que proveen para la extracción de la información. Es importante notar que todas las ideas probadas para un WSDL [11] pueden ser extendidas y utilizadas con relativa facilidad en aplicaciones que usan otros lenguajes de programación.
- Se generaron, como parte de las tareas mencionadas previamente, diferentes publicaciones en congresos nacionales, internacionales, capítulos de libros y revistas indexadas.
- Se desarrolló un prototipo con diferentes recorridos en el árbol de sintaxis abstracta que permite extraer la información de los identificadores. Dicho proceso se aplicó a WDSLs mediante la utilización de un parser. El lenguaje utilizado por los WDSL es el XML [12] y para este lenguaje existen varios parsers. En el caso particular de este trabajo se utiliza DOM (Document Object Model), que facilita las distintas estrategias de inspección de información y manipulación de la misma [13]. Así mismo, la herramienta es parte de un proyecto de mayor envergadura, cuyas operaciones sobre el código fuente son más complejas y en las cuales utilizar DOM facilita el trabajo.
- Se estuvo trabajando con métodos simples de ofuscación y encriptación de código para mejorar el nivel de seguridad de la información contenida en los identificadores [14, 15]. Las Técnicas de Comprensibilidad permiten, a través de métricas, medir el nivel o grado de entendimiento de un WSDL [8, 9, 10]. Mientras mayor sea el nivel de comprensibilidad o entendimiento de la información, mayor será su vulnerabilidad. Si las métricas indican que el nivel de comprensibilidad de un WDSL es alto, se lo puede manipular y transformar con métodos adecuados de ofuscación y/o encriptación. Al aplicar dichos procesos la información se volverá menos entendible, por consiguiente su nivel de comprensibilidad será mucho menor y por lo tanto más segura.

Entre los objetivos planteados a corto y largo plazo relacionados a este trabajo se pueden mencionar:

- Extender el nivel de seguridad no solo a la información contenida en los identificadores, sino a todos los componentes que forman una especificación de un WDSL.
- Ampliar y aplicar el prototipo a especificaciones escritas en BPEL debido a que este lenguaje es muy utilizado para la ejecución de procesos de negocios. Afortunadamente dicho lenguaje también utiliza el lenguaje XML, lo cual facilita su procesamiento.

- Estudiar, comprender y ampliar el número de métodos de encriptación y ofuscación de código utilizados.
- Crear un entorno que permita aplicar y utilizar métodos de terceros tendientes a mejorar la seguridad; ajenos a los implementados en la herramienta.

Formación de Recursos Humanos

Las tareas realizadas en el contexto de la presente línea de investigación están siendo desarrolladas como parte de trabajos para optar al grado de Magister en Ingeniería de Software. En el futuro se piensa generar diferentes tesis de licenciatura, maestría y doctorado a partir de los resultados obtenidos en la presente línea de investigación. Todas las actividades están enmarcadas en el proyecto de investigación "Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de software de Calidad".

Bibliografía

[1] ISO/IEC. Iso/iec 27032:2012 information technology - security techniques - guidelines for cybersecurity.

[2] Jorge Ramió Aguirre. Libro Electrónico de Seguridad Informática y Criptografía. Universidad Politécnica de Madrid, 2006.

[3] Jeremy Hilton Yulia Cherdantseva. Understanding information assurance and security. 2013.

[4] Alejandra Stolk. Técnicas de seguridad informática con software libre, 2013. Parque Tecnológico de Mérida. ESLARED.

[5] ISO/IEC. Iso/iec 27000:2016 information technology - security techniques - information security management systems - overview and vocabulary, 2016.

[6] http://www.seguridadinformatica.unlu.edu.ar/. UNLU. 2016.

[7] Mario Berón, Germán Montejano, Daniel Riesco, Pedro Rangel Henriques, Narayan Debnath. SIP: A Simple Tool for Inspecting and Evaluating WSDL Specifications. 10th International Conference on Information Technology: New Generations. 2013.

[8] Hernán Bernardis, Edgardo Bernardis, Mario Berón, Daniel Riesco, Pedro Rangel Henriques, Maria Joao Pereira. Cálculo de Métricas para Medir el Grado de Entendimiento de una Descripción WSDL. WICC. 2016.

[9] Mario M. Berón, Hernán Bernardis, Enrique A. Miranda, Daniel E. Riesco, Maria João Pereira, Pedro Rangel Henriques. "WSDLUD: a Metric to Measure the Understanding Degree of WSDL Descriptions". Proceedings of the 2015 Symposium on Languages, Applications and Technologies, SLATE'15. Madrid, España 2015.

[10] Bernardis, Hernán; Berón Mario; Bernardis, Edgardo; Riesco, Daniel; Henriques, Pedro. "Extracción de información y cálculo de métricas en WSDL 1.1 y 2.0". II Congreso Nacional de Ingeniería Informática / Sistemas de información (CoNaIISI). Argentina. 2014.

[11] WSDL Specification for W3C https://www.w3.org/TR/wsdl.

[12] Extensible Markup Language (XML) 1.0 (Fifth Edition). https://www.w3.org/TR/REC-xml/.

[13] Parser DOM specification for W3C. https://www.w3.org/DOM.

[14] Cappaert, J. Code obfuscation techniques for software protection. Katholieke Universiteit Leuven. 2012.

[15] Stallings W. Cryptography and Network Security Principles and Practice. Fifth Edition.

[16] A. V. Aho, R. Sethi, and J. D. Ullman. "Compilers Principles, Techniques and Tools". Addison-Wesley, 1986.