Sistemas de detección de intrusos: carencias actuales y nuevas tecnologías

El objetivo de la detección de intrusos es avisar en tiempo real de los ataques que se producen en las infraestructuras TIC. Se trata de un área reconocida de la seguridad informática que tradicionalmente ha sido basada en la detección de patrones de ataques sobre redes o sistemas finales. En este artículo se introducirán algunas de las limitaciones derivadas de este acercamiento, tanto directas (la incapacidad de detectar ataques desconocidos) como indirectas (ataques al detector de intrusos por el trabajo necesario de análisis de protocolos). Finalmente se presentarán algunas de las técnicas complementarias que están aún en fase de experimentación y estudio.

Javier Fernández-Sanguino Peña

Introducción a la detección de intrusos

La detección de intrusos es el área aplicada de la seguridad informática encargada de informar de eventos que puedan tener lugar en un sistema informático y pueda ser considerado, por unas u otras razones, como parte de un intento de intrusión. Como intrusión se entiende la realización de un acto no autorizado, como pueda ser el acceso a un sistema, la ejecución de programas no autorizados o el ataque a una red informática.

El concepto de intrusión está cercano al de un ataque dirigido, pero algunas de las tareas previas a un ataque, como pueda ser la recopilación de información o la búsqueda de servicios no está directamente tipificada como ataque. Actualmente existe una controversia sobre si dichas actividades constituyen o no una actividad ilegal como lo pueda ser el acceso sin autorización a un sistema informático. El hecho de que estemos ante un entorno complejo, formado por un sistema de información global interconectado a través de redes públicas de comunicación, hace difícil determinar si las actividades que realizan los sistemas por sí mismos pueden considerarse ataques cuando son realizados por personas con intención desconocida.

Un sistema de detección de intrusos ha de distinguir entre un acceso normal y habitual al sistema, que puede surgir de la puesta en marcha de servicios ofrecidos al exterior (entendiendo como exterior cualquier otro sistema ajeno al que ofrece los servicios), de un intento de vulnerar de algún modo dichos servicios, e incluso de aquellos que no debieran ser públicos, como parte del ataque a dicho sistema.

Es, por tanto, un sistema de detección de intrusos aquél capaz de advertir al administrador de todas aquellas situaciones que puedan ser consideradas como elementos o fases de una intrusión. El objetivo de dicho sistema es, en la medida de lo posible, proporcionar conocimiento de la puesta en marcha de un ataque sobre el sistema antes de que dicho ataque tenga éxito. Se ha de considerar, por tanto, como un mecanismo previo de alarma que está indisolublemente unido al mecanismo de respuesta. De esta forma se podrán poner en marcha las medidas necesarias para mitigar el impacto.

Los sistemas detección de intrusos quedan divididos en función, fundamentalmente, del lugar donde realizan la detección. Este lugar puede ser la red, basándose en el análisis del tráfico que pasa por ésta y su contenido, o puede ser el propio sistema operativo (basados en host) sobre el que se monitoriza el uso que las aplicaciones, procesos y usuarios hacen de él.

Aún con un desarrollo correcto en el tratamiento de los protocolos, los detectores de intrusos de red son muy susceptibles a ataques para inutilizarlos mediante la generación de tráfico falseado que consuma su capacidad

-

La técnica tradicionalmente aplicada a la detección de intrusos, en todos sus ámbitos, consisten generalmente en el uso de reconocimientos de patrones para determinar ataques conocidos. De igual forma que la tecnología aplicada a la detección de virus, basada en la introducción de firmas más algunos heurísticos para detectar ligeras desviaciones, la detección de intrusos habitualmente busca en patrones que permiten discriminar un ataque de algo que no lo es.

Sin embargo, los problemas de la aplicación de esta técnica se pueden resumir en:

* La incapacidad de detectar nuevos ataques, es decir, nuevos patrones. Esto deriva en que dichas herramientas han de ser actualizadas contínuamente y que, además, siempre existirán ataques aún no descubiertos que no podrán identificarse.

* La aparición de "falsos positivos". Esto es, la detección de ataques que realmente no lo son debidos a que algunos patrones pueden ser , en realidad, accesos legítimos a los servicios.

* Los "falsos negativos". Corresponden estos sucesos a ataques que pasan desapercibidos para el sistema de detección de intrusos.

Cabe destacar que el segundo de estos problemas es, sin embargo, difícil de resolver completamente debido a que, independientemente de la técnica empleada, siempre existirá un margen de error no nulo a la hora de clasificar un determinado evento como ataque o no.

Una limitación: aplicaciones desarrolladas a medida

El hecho de que los sistemas de detección de intrusos dependan intrínsecamente de las “firmas” que incluyen para detectar ataques es un grave problema a la hora de detectar ataques a aplicaciones que se acceden de forma supuestamente legítima. Debido al auge de Internet se ha utilizado ésta, y más particularmente el protocolo http, como la forma de generar aplicaciones “a medida” independientes de plataforma al hacer uso de un elemento común, el navegador. Así, han surgido multitud de sistemas de gestión, de comercio electrónico, de información, etc. implementados sobre este esta plataforma. Las aplicaciones que se utilizan, no son generalmente productos adquiridos e instalados, sino que existe una gran heterogeneidad al tratarse, fundamentalmente, de utilidades realizadas a medida para las organizaciones, o de éstas para sus clientes.

El diseño de aplicaciones introduce nuevas vulnerabilidades dentro de los sistemas informáticos. Estas vulnerabilidades están habitualmente relacionadas con el tratamiento de los parámetros de entrada (valores de formularios, datos de sesión) que permitirán, en función de la implementación de la aplicación, posibilidades de ejecución de código arbitrario, bien dentro del sistema operativo del sistema, bien en la aplicación mediante la ejecución de sentencias embebidas en el lenguaje de desarrollo (java para los servidores de aplicaciones, visual Basic para páginas ASP, php...).

Evidentemente, los fabricantes de las herramientas de detección de intrusos no pueden proporcionar firmas para aplicaciones que desconocen, por esto habitualmente fracasan (entendiendo como fallos los falsos negativos) en la detección de estos ataques. Es posible, sin embargo, que los detectores de intrusos basados en host puedan detectar en cierta medida el comportamiento anormal de un sistema bajo ataque, pero esto dependerá sustancialmente de las acciones tomadas una vez el sistema ha sido comprometido.

De igual manera, un tipo determinado de ataques basados en la sobrecarga de búfer pueden ser en algunos casos detectados por las sondas de en red, aunque estos ataques pueden ser, también, modificados para que pasen inadvertidos para éstas (debido a la gran variedad de instrucciones en ensamblador que generalmente se pueden llevar a cabo en un sistema remoto).

Ventana de vulnerabilidad en sistemas de detección de intrusos

Las herramientas de detección de intrusos, al igual que los sistemas que pueden estar bajo ataque, sufren del efecto conocido como "ventana de vulnerabilidad". Es decir, existe un tiempo de exposición a un ataque en el que el riesgo es mayor. Dicho tiempo es el que pasa desde que el el impacto, o materialización de una amenaza, es posible, hasta que se desarrolla una protección (parche) para la misma. Cuando sale a la luz pública una nueva vulnerabilidad el sistema de detección de intrusos no podrá reconocerla de forma inmediata hasta que se inserte en su sistema de reconocimiento la firma que la identifica unívocamente.

La disponibilidad de las firmas dependerá de la celeridad del fabricante en crear éstas una vez conocido el funcionamiento del ataque. Sin embargo, puede llegar a darse incluso la paradoja de que la generación de una firma sea superior al tiempo de aparición del arreglo de dicha vulnerabilidad (o parche). En este caso, si se sigue una política adecuada de actualización es más que probable que el riesgo desaprezca porque el sistema atacado ya no es vulnerable a éstos.

Más problemas: sobrecarga de análisis en la red

Además, en el caso de los detectores de intrusos de red hay que añadir unas limitaciones inherentes a su funcionamiento. Debido a que deben tratar las conexiones en el nivel superior de la pila de protocolos para aplicar las firmas, han de ser capaces de “entender” las conexiones realizadas. Esto significa que cada paquete recibido por el detector de intrusos debe ser analizado como parte de una conexión, decodificado y tratado para analizar su contenido. Esta sobrecarga de trabajo ha llevado a muchas limitaciones. Algunas de éstas se han ido superando por parte de los fabricantes. Las ya superadas son:

- La capacidad de trabajar en redes de tránsito elevado (velocidades de Gigabit)

- La de ser susceptibles de técnicas “stealth", que evitan al propio detector de intrusos bien por fallos en el tratamiento de los protocolos (por ejemplo por la utilización de fragmentación de paquetes), bien porque la implementación del tratamiento de los datos no es totalmente correcta (se puede evitar, por ejemplo, las firmas de ataques http codificación los URLs de formas distintas)

Sin embargo hay limitaciones aún no superadas. Obviamente, los detectores de intrusos de red son incapaces de tratar el tráfico cifrado en la red, por ejemplo, en el caso de que se utilice HTTPS o Redes Privadas Virtuales. Aunque algunos fabricantes empiecen a añadir capacidades de descifrado (introduciendo las claves privadas en el detector) para este tipo de tráfico, este análisis supone una sobrecarga adicional (y elevada) al detector de intrusos que puede llevar a que sea incapaz de analizar las conexiones en tránsito sino dispone de un sistema de aceleración de cifrado adicional (encareciendo el producto).

Pocos productos de detección de intrusos son capaces de integrarse con consolas de gestión de red (Patrol, Tivoli, HP Openview) para ofrecer un entorno de gestión y operación único

Igualmente, este tipo de análisis hace que los propios detectores de intrusos sean susceptibles de ataque. Una mala programación en el tratamiento de los datos de entrada (esto es, los paquetes) puede llevar a que se introduzca una vulnerabilidad en el propio detector de intrusos y éste pueda ser atacado de forma indirecta. Un ejemplo, es el envío de paquetes maliciosos a otros destinos sabiendo que serán tratados por el propio detector. Aún con un desarrollo correcto en el tratamiento de los protocolos, los detectores de intrusos de red son susceptibles a ataques para inutilizarlos mediante la generación de tráfico falseado que consuma su capacidad. Estos ataques podrían ser de denegación de servicio al propio sensor.

Interoperabilidad e integración

Otro de los problemas actuales de los sistemas de detección de intrusos es la falta de interoperabilidad entre éstos así como con herramientas de gestión de la seguridad independientes. La mayor parte de los fabricantes diseñan los detectores de intrusos pensando en “su consola”. Este hecho lleva a que los usuarios, a la hora de abordar la introducción de un sistema de detección de intrusos para un entorno heterogéneo (redes de área local, redes gigabit, sistemas host), no pueden decantarse por el sistema de detección de intrusos mejor para dicho entorno, sino que deben elegir un fabricante en concreto, que seguramente no ofrezca la mejor solución en todos los entornos. La única alternativa es desplegar distintos sensores de detección de intrusos y disponer de más de una consola, con los consiguientes problemas de operación. Cabe destacar también que pocos productos de detección de intrusos son capaces de integrarse con consolas de gestión de red (Patrol, Tivoli, HP Openview) para ofrecer un entorno de gestión y operación único.

Existe, sin embargo, un esfuerzo de estandarización que los fabricantes tendrán que asumir como suyo, conocido como CIDF (Common Intrusion Detection Framework) que ha tenido como resultado el Grupo de Trabajo Formato de Intercambio de Detección de Intrusos (idwg) de la IETF. Dicho grupo de trabajo ha publicado ya cinco borradores en los que se define un modelo de formato de datos y mensajes relacionados con elementos de detección de intrusos.

Si los fabricantes introducen este estándar (actualmente en fase de borrador) podrá ser posible, en un futuro, la integración de elementos de detección de intrusos a consolas de distintos fabricantes, así como dentro de consolas generales de gestión de la seguridad.

Necesidad de contínuos ajustes

A nivel operativo, los detectores de intrusos también tienen unas carencias importantes que hace que su despliegue sea dificultoso. Aquellos que hayan tenido la experiencia de instalar un producto de detección de intrusos en una red en funcionamiento habrán observado que, de partida, el sistema de detección de intrusos va a emitir, practicamente al momento, alarmas de ataques aún cuando la red no esté sufriendo ninguno. Esta generación de falsos positivos se debe a que determinados usos de la infraestructura de red o de los sistemas finales tienen características similares a los ataques definidos en su base de datos. La generación de estos falsos positivos puede incluso llegar a abrumar al operador por un exceso de información.

En estas condiciones, un detector de intrusos no tiene una utilidad práctica si no es con un refinado posterior. Trabajo arduo de ingeniería que es estrictamente dependiente del tráfico que tiene lugar en la red o de las características del host donde se instala el sensor y que, en muchos casos, puede llegar a alargarse en el tiempo, entorpeciendo el despliegue de estos sistemas. Además, estos ajustes no se realizan “una vez y ya está” sino que si se modifican el uso de los recursos que analiza el sensor puede ser necesario realizar contínuamente. Existe un riesgo implícito dentro de los ajustes y es que se pase de un lado a otro de la balanza, es decir, que de tener “falsos positivos” y tras un ajuste de descarte se puede llegar a introducir condiciones que generan “falsos negativos”. Estas condiciones, además, pueden no llegar a ser observadas hasta que el impacto de un ataque no detectado por el sensor se hace evidente.

Inteligencia artificial y detección de intrusos

Como una herramienta adicional que pueda ayudar a resolver los problemas y limitaciones de los sistemas de detección de intrusos existen acercamientos que hacen uso de técnicas de inteligencia artificial, en concreto de la rama de la minería de datos.

La aplicación de técnicas de inteligencia

artificial a la detección de intrusos, e incluso de la minería de datos en sí, no es un hecho novedoso, como se puede observar en **[lee01], [eskin00] o [ghosh99]**, existen trabajos de desarrollo de estas tecnologías desde 1992. Sin embargo, los analistas de los sistemas de detección de intrusos, coinciden en que las investigaciones en tecnologías de detección de intrusos se dirigen, hoy en día, hacia la aplicación de la minería de datos. La intención es evitar la aparición de falsos positivos/negativos así como mejorar las capacidades de los sistemas de detección para hacer frente a nuevos ataques.

La detección de intrusos puede ser enriquecida con la aplicación del campo conocido como KDD (Knowledge Discovery in Databases). Su objetivo es la extracción de información implícita, no trivial, previamente desconocida y potencialmente útil a partir de los datos.

Quizás la alternativa óptima, y la más aceptada en la aplicación de la minería de datos a la detección de intrusos, es la de "detección de anomalías" **[portnoy00]**. Esta detección puede realizarse de forma supervisada o no. Para llegar a esta detección de anomalías se pueden aplicar una serie de técnicas **[lee00]**: Clasificación, Episodios Frecuentes, Asociación de Valores (correlaciones), y Análisis adaptativos. Este tipo de aplicación lleva a la utilización del detector de intrusos en dos modos. Inicialmente se instala el detector de intrusos en modo “aprendizaje” en el que analiza la información que recibe para determinar los patrones de comportamiento habituales y, por tanto, permitidos. Tras este modo de aprendizaje, y después de un análisis de los criterios obtenidos, el detector de intrusos se dispone en modo “análisis”.

A diferencia de los detectores de intrusos basados en firmas conocidas, en los que el ajuste para eliminar los falsos positivos debe realizarse con un conocimiento en profundidad del entorno donde éste se ha desplegado, aquí el conocimiento del entorno es adquirido, directamente, por el sensor que va a realizar la detección.

Este tipo de tecnologías puede aplicarse tanto a detección de intrusos en red como detección de intrusos en host. La ventaja adicional en el campo de la detección de intrusos en red es que un detector de intrusos basado en comportamientos no tiene por qué analizar el trafico en detalle y su contenido. Es, por tanto, menos susceptible de ataques directos al detector de intrusos. La sobrecarga de análisis de paquetes es también consecuentemente menor.

Sin embargo, esta tecnología no está aún lo suficientemente madura. Existen algunos productos que incorporan ya este tipo de análisis basado en comportamiento estadístico. En el caso de detección de intrusos basada en red se pueden destacar: **Snort** (a través del módulo ACID), **Anzen Flight Jacked** y **Centrax** (ambos adquirido el año pasado por **NFR**).

En el caso de la detección basada en host, se aplica muy orientada a las características y comportamiento de los usuarios, las características de los procesos que se ejecutan en el sistema, e incluso del comportamiento de los propios proceso en sí (a través de sus llamadas al sistema y a las librerías). Algunas de las herramientas que se pueden destacar son las descritas en **[lee98], [esk01]** o **[esklee01]**, así como la herramienta **HostSentry** (de Psionic).

Conclusiones

En el artículo se han expuesto las limitaciones conocidas en la actualidad de la tecnología utilizada para la detección de intrusos. Estas limitaciones incluyen la incapacidad de tratar aplicaciones a medida, la falta de interoperabilidad entre fabricantes, la sobrecarga de análisis que lleva a la posibilidad de ataques contra el propio detector de intrusos, así como, finalmente, la necesidad de actualizaciones y ajustes contínuos del cortafuegos.

El área de la minería de datos es perfectamente aplicable a la detección de intrusos en sistemas informáticos y ofrece, además una cobertura de ataques que no puede conseguirse mediante la utilización de técnicas de firma de ataques conocidos. Como se demuestra en **[schultz01]** el acercamiento de minería de datos está demostrando ser un buen acercamiento para tareas que anteriormente se hacían con detección de firmas.

En cualquier caso, en la detección de intrusos no se puede hablar de una única herramienta que cubra todo el espectro y avise de todas las intrusiones. En el caso del análisis de logs, no se puede esperar detectar determinados ataques de denegación de servicio que fuercen un comportamiento anómalo del servidor y que no queden registrado (por ejemplo, ataques de desbordamiento de búfer). Así, las herramientas de minería de datos para detección de intrusos suponen un complemento más al bagaje de herramientas de este tipo que permitirá, en un futuro, reducir el número de falsos positivos y falsos negativos de forma global al integrarse con otras técnicas de detección de intrusos. Para esto es necesario desplegar agentes de distintas tecnologías y métodos que informe a un sólo sistema que agregue el trabajo de todos. Algo que sólo podrá llegar a ser posible si los fabricantes empiezan a hacer uso de tipos de intercambio de datos como CIDF. De otra forma será necesario esperar a que éstos apliquen todas las técnicas y las integren en sus consolas únicas, sin que sea posible integrar distintos detectores de intrusos (y de distinto fabricante) en una misma consola.

Javier Fernández-Sanguino Peña
Jefe de Proyecto
División de Seguridad
de Germinus Solutions
jfernandez@germinus.com

Bibliografía

[mit99] 1999 DARPA Intrusion Detection Evaluation: http://www.ll.mit.edu/IST/ideval/index.html, 1999.

[lee01] Real Data Mining-based Intrusion Detection, Wenke Lee, Salvatore Stolfo, Phillip Chan, Eleazar Eskin, Wei Fan, Mathew Miller, Shlomo Herkshop, and Junxin Zhang, Department of Computer Science, Universidad de Columbia, USA, 2001.

[esk01] Data Mining Methods for Detection of New Malicious Executables, Eleazar Eskin, Matthew Schultz, Erez Kadok, and Salvator Stolzo, Department of Computer Science, Universidad de Columbia, USA, 2001.

[esklee01] Modeling System Calls for Intrusion Detection with Dynamic Window Sizes, Eleazar Eskin, Wenke Lee, and Salvator Stolzo, 2001.

[schultz01] MEF: Malicious Email Filter. A UNIX Mail Filter that Detects Malicious Window Executable, Matthew Schultz, Eleazar Eskin, Wenke Lee, Salvator Stolzo, and Manasi Bhattacharyya, 2001.

[eskin00] Adaptative Model Generation for Intrusion Detection Systems, Eleazar Eskin, Zhi-da Zhong, George Yi, Wei-Ang Lee, Mathew Miller, and Salvatore Stolfo, Department of Computer Science, Universidad de Columbia, USA, 2000.

[lee00] A Data Mining Framework for Building Intrusion Detection Models, Wenze Lee, Salvatore Stolfo, and Kui Mok, Department of Computer Science, Universidad de Columbia, USA, 2000.

[axelsson00] Intrusion Detection Systems: A Survey and Taxonomy, Stefan Axelsson, Department of Computer Engineering, Chalmers University of Technology, Göteborg, Suecia, 14 marzo 2000.

[portnoy00] Intrusion detection with unlabeled data using clustering, Leonid Portnoy, Data Mining Lab, Department of Computer Science, Universidad de Columbia, 2000.

[ghosh99] A study in Using Neural Networks for Anomaly and Miuse Detection, Anap Ghosh and Aaron Schwartzbad, 1999.

[cann98] Artificial Neural Networks for misuse detection, J Cannady, Proceedings of the 1998 National Information Security Conferences, Octubre 5-8 1998.

[lee98] Data Mining Approaches for Intrusion Detection, Wenke Lee and Salvatore Stolfo, Department of Computer Science, Universidad de Columbia, USA, Proceedings of the Seventh USENIX Security Symposium, Octubre 5-8 1998, 1998.