LA HEURÍSTICA EN LOS VIRUS

Francisco Eleazar Delgado Contreras
Jesús Humberto Rojas Rangel
José Luis Mares Monsiváis
Coautor: Julio César González Cervantes
FCFM-UANL
Facultad de Ciencias Físico Matemáticas
Universidad Autónoma de Nuevo León
San Nicolás de los Garza, Nuevo León, México

Resumen:

La heurística tiene dos objetivos: el uso de algoritmos en buen tiempo de ejecución y que sea óptimo para el trabajo, en caso de los virus su objetivo es penetrar al sistema con éxito sin ser detectado por algún sistema de protección, con la finalidad de infectar la computadora, archivos entre otras cosas. Para que los virus no sean detectados y que puedan sobrevivir más los *hackers*, o personas malintencionadas, utilizan métodos heurísticos para que los antivirus no los puedan detectar; tanto la base de datos del mismo programa, hasta el código malicioso con el que se generó. Generalmente, la programación heurística es considerada como una de las aplicaciones de inteligencia artificial y también como herramientas para la solución de problemas. Los antivirus utilizan diferentes técnicas para detectar los virus, *malware* entre otras cosas. Los antivirus pueden reconocer la firma de un virus, entre más instrucciones específicas contengan el virus o el código malicioso estas definen como comportarse y actuar.

Palabras claves:

heurística, antivirus, virus, detección, malware, malicioso

Introducción

Antes de saber cómo funciona la heurística en antivirus y virus, debemos saber lo que es una heurística como tal. Una heurística es un algoritmo que abandona uno o ambos objetivos fundamentales de la computación que son encontrar algoritmos con buenos tiempos de ejecución y que sean óptimos.

Ya hablando de antivirus, la heurística son los métodos que se emplean para reconocer códigos maliciosos (virus, gusanos, troyanos, etc.) que no se encuentran en la base de datos. En el caso de los virus, su heurística vendrían siendo los métodos que usán para evitar ser detectados por los antivirus; su importancia radica en el hecho de poder evadir la única defensa automática posible del antivirus frente a la aparición de nuevos códigos maliciosos de los que no se posean firmas en la base de datos del antivirus.

La heurística

La heurística, en la terminología de las tecnologías de la información, tiene dos objetivos importantes que son: el uso de algoritmos en buen tiempo de ejecución y que sea óptimo para el trabajo. En el caso del virus, su objetivo es penetrar al sistema con éxito sin ser detectado por algún sistema de protección, con la finalidad de infectar la computadora, archivos, entre otras cosas. Usando estos métodos en los virus, aumenta la probabilidad de que el sistema sea afectado por este programa y puede realizar su ejecución con facilidad. Con esto nos damos cuenta que la heurística de los virus también va de la mano con la heurística de los antivirus, que en el caso de los antivirus funcionan escaneando continuamente el sistema en busca de algún virus que contenga la base de datos. Los antivirus también están conformados de las instrucción de configuración donde se almacenan los patrones de comportamiento que sirven para identificar a este código malicioso. La vacuna que contiene este antivirus vendría siendo el método heurístico que en términos de informática es la detección de firmas genéricas, reconocimiento del código compilado, desensamblado, des empaquetamiento entre otros, más adelante hablaremos de estos temas.

Generalmente, la programación heurística es considerada como una de las aplicaciones de inteligencia artificial y también como herramientas para la solución de problemas. La programación heurística es utilizada en sistemas expertos, se construye bajo ciertas reglas extraídas de la experiencia y respuestas que fueron generadas por tal sistema, donde va adquiriendo la medicina que "aprende" mediante la experiencia y aumento en su base de conocimiento.

La programación heurística posee un doble rol en el desarrollo antivirus: velocidad y detección, por tal motivo de ahí surgió el término heurística.

Para que el virus no sea detectado y que pueda sobrevivir, mas los *hackers,* o personas malintencionadas, utilizan estos métodos para que los antivirus no los puedan detectar, tanto la base de datos del mismo programa, hasta el código malicioso con el que se generó. Los métodos de infección de los virus pueden ser de dos maneras:

1. Los antivirus pueden reconocer la firma de un virus, mientras más instrucciones específicas contenga el virus o el código malicioso que estas definen como comportarse y actuar. Por eso la firma digital es como la huella digital de los *hackers* o de las personas malintencionadas es única y distinta.
2. Es el comportamiento que contenga el virus, esto depende de cómo se realizó este virus porque pueden variar de infectar a varios archivos hasta corromper carpetas u otros archivos del sistema.

Después de ver cómo actúan los virus, veremos qué métodos o que tipos de infección realizan de diferentes maneras que se explicarán a continuación:

- Infección Directa: Se enfoca a atacar primordialmente en infectar el disco o los discos duras que tenga la computadora; ataca a uno o más archivos del sistema. Esto se realiza mediante un programa o un archivo infectado que nos pasaron o descargamos por Internet; la infección se propaga tanto en una computadora o servidor; varía si es que están en la misma red o los archivos compartidos del sistema, este tipo de infección es el más común que es muy fácil de detectar por un antivirus.
- Infección Rápida: este tipo de infección es un poco peligrosa porque puede infectar cualquier archivo e incluso el mismo antivirus.
- Infección Lenta: método tiene el fin de ocultarse o enmascararse dentro del sistema, haciendo difíciles de detectar por el antivirus.
- Infección Minimalista: puede que a veces los infecte o a veces no lo hará, hace que la infección sea lenta y aumenta la probabilidad de no ser detectado por el sistema.
- Infección RAM Residente: puede infectar los programas de arranque del sistema. Estos virus no se propagan por la red o por el Internet, solo insertando

físicamente un disquete, USB u otro medio de almacenamiento; puede infectar a los demás equipos que estén en la misma red.

Análisis Heurístico

Utiliza un acercamiento basándose en reglas para diagnosticar un archivo potencialmente ofensivo o malicioso. Debido a que el motor de análisis trabaja a través de su base de reglas, chequeando el archivo mediante criterios que indican que podría ser un archivo sospechoso, cuando hay archivos semejantes de posible *malware*, realiza un puntaje igual o superior o semejante a él. Si el puntaje es igual o superior al umbral estimado, ese archivo es señalado como sospechoso (malicioso o *spam*) y será procesado de acuerdo a los criterios e instrucciones del sistema o antivirus.

Lo opuesto a la Heurística

Lo que en realidad es lo opuesto a la heurística o lo opuesto al análisis heurístico en los antivirus no es la exploración por firmas sino la exploración algorítmica, en la exploración por firmas es un caso especial de la misma. La exploración algorítmica es la decodificación algorítmica, que está basada en procedimientos matemáticos. Lo que es referido en la industria como la exploración algorítmica, normalmente es entendido como algo que se basa en un algoritmo (buscar una cadena estática o una secuencia estructurada de *bytes*) que es especificada de acuerdo al virus que se intenta detectar.

Método heurístico en los antivirus

Como se había mencionado antes, los antivirus ,utilizando este método, sirven para detectar los patrones o comportamientos de los virus. Tiene como finalidad detectar este virus mediante el uso de firmas digitales o genéricas, el reconociendo del código malicioso, etc. No es un método exacto, pero se aproxima mucho a la realidad; como los virus cambian día con día sus patrones, es más difícil detectar por dicha razón este método; se aproxima, y cuando ya está seguro dependiendo de qué haya analizado, lo detectará como virus o no.

Métodos de detección heurística en los antivirus

Los antivirus utilizan diferentes técnicas para detectar los virus, *malware*, entre otras cosas. Cada método es para analizar las características y el comportamiento para determinar si es un virus o un archivo sospechoso. Ahora mencionaremos algunas de estas técnicas:

- Emulación del archivo: También conocido como la caja de arenas de prueba (exploración dinámica) donde se realiza la emulación de archivos en un ambiente virtual controlado (caja de arena) para ver qué es lo que realiza. Tiene como fin ver si este archivo se comporta como virus, si es el caso, se trata como virus.
- Análisis del archivo: El antivirus analizará el archivo sospechoso con la finalidad de checar cuál es su intención, destino y propósito; pero si tienen funciones que pongan en riesgo al sistema, como eliminación de archivos del sistema, se reconocerá automáticamente que es un virus.
- Detección de Firma Genérica (Firma Digital): Esta técnica es especial, debido a que es muy común para detectar los virus. Cuando se crea un virus o se da a conocer por su nombre; la mayoría de la gente cree que es un nuevo virus que amenaza el sistema, pero es todo lo contrario debido a que provienen de la misma familia o clasificación. La ventaja que tienen los antivirus en esto es que utilizan las mismas definiciones que utilizaron anteriormente para localizar a estos nuevos virus o "primos" similares e incluso si utilizan otro nombre diferente o que contenga otro tipo de función [6].

Ya hablamos de los tipos de infección y cómo la heurística aplicada en los virus y antivirus; a continuación, veremos los tipos de heurística que aplican en el análisis para la detección, estos temas son impartidos por el antivirus ESET NOD 32 que utiliza estos temas para la detección proactiva de *malware* [4].

Tipos de heurística

Los antivirus utilizan tres tipos de variantes más comunes para el análisis, que son normalmente utilizados para la detección heurística. Son los siguientes:

- Heurística Genérica: Realiza un análisis de la similitud del objeto a otro, ya sea código malicioso. Si al analizar un archivo es similar a un código malicioso, este será detectado y será considerado como amenaza al sistema.
- Heurística Pasiva: Se analiza el archivo tratando de determinar qué es lo que realiza el programa. Si realiza actividades sospechosas, será considerado como malicioso.
- Heurística activa: Como objetivo crear un entorno seguro y ejecutar el código de tal forma que se pueda

reconocer cuál es el comportamiento del código. Esta técnica es conocida por los siguientes nombres: "Caja de Arena (*Sandbox*)", "Virtualización" o "Emulación".

Las ventajas y desventajas de la heurística en los antivirus

Utilizando la heurística es una manera eficaz para identificar las amenazar desconocidas para la protección de nuestro sistema actualizada día a día en tiempo real, pero tiene sus desventajas. Este tipo de análisis y exploración puede llevar tiempo debido a que hace un chequeo a todos los archivos para verificar si existe una amenaza, por lo tanto afecta el rendimiento de nuestro sistema haciéndolo lento.

La preocupación principal con la detección heurística es que a menudo aumentan los falsos positivos. Los falsos positivos son que el antivirus determina un archivo es malicioso (lo pone en cuarentena o elimina la misma), cuando en perfectas condiciones. Esto se debe a que algunos archivos pueden parecerse a los virus, pero en realidad no lo son, son restringidos o que dejaron de trabajar en el sistema.

Referencias

[1] Virus y Antivirus. Recuperado de: http://www.monografias.com/trabajos18/virus-antivirus/virus-antivirus2.shtml

[2] Estudio sobre virus informáticos Recuperado de; http://www.monografias.com/trabajos/estudiovirus/estudiovirus.shtml

[3] La Heurística y los Antivirus. Recuperado de: http://www.expresionbinaria.com/la-heuristica-y-los-antivirus/

[4] Análisis Heurístico: detectando malware desconocido. Recuperado de: http://www.eset-la.com/pdf/prensa/informe/analisis_heuristico_detectando_malware_desconocido.pdf

[5] How viruses avoid detection. Recuperado de: http://etutorials.org/Misc/computer+book/Part+2+Dangerous+Threats+on+the+Internet/Chapter+7+Viruses+and+Worms/HOW+VIRUSES+AVOID+DETECTION/

[6] What is Heuristic Antivirus Detection? Recuperado de: http://internet-security-suite-review.toptenreviews.com/premium-security-suites/what-is-heuristic-antivirus-detection-.html

Datos de los autores:

Francisco Eleazar Delgado Contreras

Escuela: Facultad de Ciencias Físico-Matemáticas

Dirección: Plan de Galeana No. 2804, La republica, Monterrey, Nuevo León, C.P: 64900

Email: francisco.fcfm92@gmail.com

Jesús Humberto Rojas Rangel

Escuela: Facultad de Ciencias Físico-Matemáticas

Dirección: Abril en Portugal No. 319, Residencial San Nicolás, San Nicolás de los Garza, Nuevo León, C. P: 66414

Email: rojaslcc@msn.com

José Luis Mares Monsiváis

Escuela: Facultad de Ciencias Físico-Matemáticas

Dirección: Bambú No. 118, La enramada 3er Sector, Apodaca, Nuevo León, C. P: 66635

Email: huero_xtm@hotmail.com

Coautor:

Ing. Julio César González Cervantes

Dirección: Lazaro Cardenas 1212 col. Las Puentes, San Nicolas de los Garza

Email: julioglezc@gmail.com