Lógica aplicada en IDS's

IDS = Intrusion Detection System

Carlos Ledesma

Gestión Inteligente de la Información

Índice

Índice

Lógica cuántica

La primera aplicación de la lógica que se me viene a la cabeza, es la lógica cuántica:

- **Problema:** Suena interesante, pero mi conocimiento sobre la computación cuántica no pasa del nivel divulgativo, y lo veo un futuro muy lejano y abstracto. Me pongo a leer lo más básico y me pierdo. Pierdo el interés...

Análisis de malware

Hace un mes, me pongo a buscar aplicaciones de la lógica en el análisis de malware:

- **Problema:** Es lo que más me interesa actualmente, pero tras una búsqueda superficial, no encuentro aplicación existente interesante. Supongo que el análisis de malware es demasiado caótico para formalizarlo...
- **Investigación:** ¿Puede ser útil la lógica aplicada en la automatización del análisis de malware? Idea abierta...

Sistemas de detección de intrusos

Leyendo sobre seguridad informática, me entero de la existencia de los sistemas de detección de intrusos:

- **Definición:** Su nombre dice mucho. La idea es identificar una intrusión según patrones de eventos en el host o en la red. Ciertos eventos por sí solos no son síntoma de intrusión, pero juntos y dispuestos en el *tiempo* de cierta forma, sí...
- **Hmmm...** Eso suena a lógica temporal. Una serie de eventos relacionados de cierta forma en el tiempo, vamos a ver...

Índice

Primera aproximación

Primera parada: Google (Lo siento... Por algún lado había que empezar, ¿no?). Términos de la busqueda: *intrusion detection logic*

- Me sale de todo menos lo que busco. Así que busco buscadores académicos. Me quedo principalmente con dos, Google Scholar y BASE

Segunda parada: Google Scholar y BASE: Mismos terminos de búsqueda. Tengo la opción de limitar por años. De 2012 a la actualidad, la seguridad informática cambia en pocos años...

- En Google Scholar, me salen demasiados resultados irrelevantes. En BASE, me salen unas 100 publicaciones...

Investigar el estado del arte es duro...

- Reviso a mano los títulos de las 100 publicaciones y leo los resúmenes de las que tienen títulos relevantes
- Descargo unas 7 publicaciones (más 8 en una revisión posterior) para leerlas
- Me doy cuenta de que no sé leer publicaciones (pierdo mucho tiempo leyendo y comprendiendo partes innecesarias)
- Algunas publicaciones se basan en otras y tienes que leerte esas otras para enterarte de algo

Tipos de publicaciones encontradas (I)

- **Lógica sin razonador:** Como ejemplo, uso de un subconjunto bastante limitado de una exótica lógica llamada MSFOMTL para especificar los patrones. Pero no usa razonador, sino que traduce las fórmulas que representan patrones a queries en StreamSQL [1].

- **Publicaciones incompletas** Ya sea por demasiado abstractas [2] o porque dicen estar construyendo prototipos que a día de hoy siguen sin aparecer [3].

Tipos de publicaciones encontradas (II)

- **Lógica difusa:** Muchas publicaciones que usan clasificadores basados en la teoría de conjuntos difusos (con reglas difusas tipo if-then) y te hablan de la lógica difusa muy alegremente [4].

- **Lógica temporal:** Las publicaciones que finalmente tratan sobre la aplicación de la lógica temporal, con razonadores automáticos y mención de su complejidad computacional [5].

Índice

Índice

Una serie de eventos como evidencia de la intrusión

De forma genérica, la evidencia de la intrusión en un sistema está compuesta por una serie de eventos relacionados en el tiempo de una forma característica:

- A nivel de host, un evento puede ser una llamada al sistema, un inicio de sesión...
- Y a nivel de red, un paquete IP con ciertas características, o a más alto nivel, una conversación TCP...

Aplicación de la lógica temporal

Esta serie de eventos se puede traducir a un modelo de una lógica temporal, llamémoslo **M**

Y una intrusión se puede traducir a una fórmula en esa lógica, llamémosla **p**

Diremos que existe una intrusión representada por **p** si esta se puede inferir de la serie de eventos representada por **M** (es decir, $\mathbf{M} \models \mathbf{p}$)

Esto es lo que se conoce como **verificación de modelos**.

Índice

Algunas lógicas temporales

A la hora de elegir una lógica concreta, tenemos que saber qué queremos expresar con ella (el dominio lo dirá), para elegir una lógica con la expresividad adecuada. Las siguientes lógicas han sido usadas en IDS's anteriores:

- **Linear Temporal Logic (LTL):** Lógica temporal modal, también llamada Propositional Temporal Logic, no es muy expresiva. **No puede expresar eventos concurrentes**.

- **Interval Temporal Logic (ITL):** Ampliación de la LTL, esta sí permite expresar eventos concurrentes. **No puede expresar restricciones temporales entre eventos**.

RASL, la lógica adecuada

Ya que la serie de eventos tratada por un IDS contiene eventos concurrentes y restricciones temporales entre eventos, es necesaria una nueva lógica que tenga la expresividad suficiente. Debido a esa necesidad surge **RASL**, con las siguientes características:

- Parece lo suficientemente expresiva para expresar cualquier tipo de intrusión en un sistema (hipótesis personal)
- Es decidible, y tiene complejidad exponencial.
- Para verificación de modelos, existe un algoritmo óptimo.

Sintaxis y semántica de RASL

RASL es esencialmente **ITL** más el operador **prj** y el operador **;I**

Diapositivas sobre sintaxis y semántica de ITL (página 22 del PDF

Explicación del operador *prj* (página 83 del PDF)

El operador **;I** es como el operador **;** pero con un intervalo de tiempo entre ambos eventos definido por **I**

Ejemplos de intrusiones expresadas en RASL

Publicación principal (página 5 del PDF)

Índice

Índice

Estado del arte en la actualidad

Creo que el estudio del estado del arte realizado por los autores de la publicación principal es bastante bueno, y dada la antigüedad de la publicación (1 año), considero que en su mayor parte sigue vigente, y que esta publicación representa bastante bien el estado del arte actual, respecto a la aplicación de la lógica temporal a IDS's para detección de intrusiones según las series de eventos que provocan.

Y al ser la lógica temporal más expresiva posible (sin contar su versión híbrida, que para esta aplicación no parece necesaria), un IDS no puede necesitar más expresividad (sin salirse de la temporal). Sería interesante, eso sí, optimizar la ejecución real de algoritmo de verificación de modelos, todo lo que sea posible.

Índice

Cosas que he aprendido

- El proceso de investigación del estado del arte (buscar publicaciones, seguir la pista a sus autores...)
- Leer publicaciones de una forma eficiente
- El estado del arte actual en lógica aplicada a IDS's y lógicas temporales con más profundidad
- El mundo de las publicaciones no es tan elitista, hay bastantes publicaciones de calidad dudosa

Índice

Cosas que tendría que aprender

- Aprender más sobre la lógica temporal, en general.
- Entender perfectamente la semántica formal de la lógica RASL.
- Estudiar los algoritmos actuales de verificación de modelos, para entender y optimizar el propuesto para RASL.
- Aprender más sobre los sistemas de detección de intrusos.

Bibliografía (I)


Abdulbasit M. Ahmed; *Online Network Intrusion Detection System Using Temporal Logic and Stream Data Processing; University of Liverpool PhD Thesis; 2013*.
http://cgi.csc.liv.ac.uk/~alexei/TeSTid/thesisfinal.pdf


Ahmad Salahi, Morteza Ansarinia; *Predicting Network Attacks Using Ontology-Driven Inference; International Journal of Information and Communication Technology (IJICT), Volume 4, Issue 1; 2012*.
http://arxiv.org/ftp/arxiv/papers/1304/1304.0913.pdf

Bibliografía (II)

M. Couture, B. Ktari, M. Mejri, F. Massicotte *A Declarative Approach to Stateful Intrusion Detection and Network Monitoring; 2nd Annual Conference on Privacy, Security and Trust (PST), Fredericton, New Brunswick, Canada, pages 175-179; 2004.*
http://cg.scs.carleton.ca/~mathieu/MCouture_PSTFredericton2004.pdf

Richard A. Wasniowski; *Agent Based Intrusion Detection with Fuzzy Logic; WSEAS Transactions on Systems, Issue 10, Volume 3, pages 3169-3173; 2004.*
http://www.wseas.us/e-library/conferences/athens2004-b/papers/474-297.pdf

Bibliografía (III)

Weijun Zhu, Qinglei Zhou, Weidong Yang, Haibin Zhang; *A Novel Algorithm for Intrusion Detection Based on RASL Model Checking; Hindawi, Mathematical Problems in Engineering, Article ID 621203; 2013.*
`http://downloads.hindawi.com/journals/mpe/2013/621203.pdf`

Nota: No aparecen reflejadas las publicaciones leídas que son referenciadas en alguna de las ya reflejadas. Ni tampoco las usadas para aprender conceptos generales puntuales.