Aprendizaje por Refuerzo Aplicado al Etiquetado de Tráfico de Red.

Carlos A. Catania, Elina Pacini, Rodrigo Gonzalez, Jorge Guerra, Tatiana Parlanti, Luciano Robino y Eduardo Pavez

Universidad Nacional de Cuyo, Facultad de Ingeniería, LABSIN
Campus Universitario, Mendoza, Argentina.

{harpo,elina.pacini,rodrigo.gonzalez,jorge.guerra,tatiana.parlanti,luciano.robino,eduardo.pavez}@ingenieria.uncuyo.edu.ar

RESUMEN

Durante los últimos años, se han aplicado varios enfoques basados en el aprendizaje automático (AA) para aligerar el análisis y el reconocimiento de comportamientos maliciosos en las redes informáticas. Estos enfoques se han centrado en facilitar la tarea del personal de seguridad de la red, en mejorar su capacidad de detección y en intentar aumentar el nivel de automatización en el reconocimiento y el análisis de los comportamientos maliciosos. Sin embargo, todos estos enfoques requieren un conjunto de datos anotados con etiquetas. Al utilizar conjuntos de datos etiquetados, estos sistemas pueden generalizar comportamientos no observados previamente. Por desgracia, los conjuntos de datos etiquetados no suelen tener la representatividad y la precisión necesarias. Esta última carencia está asociada, entre otros factores, a la falta de recursos humanos formados y a la dificultad en el proceso de creación de estos conjuntos de datos etiquetados.

En consecuencia, el objetivo general de este proyecto es desarrollar estrategias para reducir el esfuerzo humano durante el proceso de etiquetado y análisis de conjuntos de datos con tráfico de red. Estas estrategias se centran en la aplicación de diferentes técnicas de Aprendizaje por Refuerzo (RL) para establecer políticas adecuadas que minimicen el grado de interacción del usuario durante el proceso de etiquetado. Al utilizar RL es posible aprender nuevas políticas que tengan en consideración otras recompensas como ser la experiencia o capacidad del usuario junto a aspectos distintivos del proceso de etiquetado como ser la calidad, oportunidad y relevancia, entre otros. De esta manera se evita depender de una única política a la hora de tomar la decisión de consultar al usuario. Esto último resulta fundamental para el desarrollo de sistemas de detección basados en técnicas de AA.

La principal contribución que se espera obtener de la investigación propuesta en el corto plazo es disponer de un conjunto de técnicas que faciliten el proceso de etiquetado en trazas de tráfico de red para su utilización en sistemas de detección de comportamiento malicioso basados en algoritmos de AA, mejorando de manera continua su capacidad de detección.

CONTEXTO

El presente proyecto se desarrolla en el marco de Facultad de Ingeniería dentro Laboratorio de sistemas inteligentes (LABSIN). Este trabajo es parte del proyecto de investigación que dio inicio en septiembre de 2022 en el marco de los proyectos bienales de secretaria de Investigación, Internacionales y Posgrados (SIIP) de la Universidad Nacional de Cuyo.

1 INTRODUCCIÓN

El proceso para la construcción de un modelo basado en aprendizaje automático (AA) asociado a un malware consiste de tres etapas [1,2]. En una primera etapa, se genera un modelo de comportamiento considerando diferentes características de la red (por ej. el tamaño, periodicidad y duración, de las trazas de tráfico). Este modelo permite visualizar de manera sencilla un gran número de conexiones de red a lo largo del tiempo y constituye además una herramienta de análisis. Luego, en una segunda etapa de reconocimiento, se realiza un proceso de asociación entre el modelo de comportamiento y la aplicación maliciosa, tarea normalmente denominada **etiquetado** y que requiere del apoyo de un especialista humano. Finalmente, en una tercera etapa, mediante el uso de algoritmos con base estadística o de aprendizaje automático se generaliza el modelo de comportamiento a fin de mejorar la capacidad del modelo para reconocer de forma automática otras aplicaciones maliciosas que presenten un comportamiento similar.

En los últimos años, se han desarrollado varios **métodos asistidos** para abordar el problema del etiquetado aplicado a las trazas de tráfico de red. En estos métodos, la inclusión de un usuario con cierto nivel de experiencia en el proceso de etiquetado mejora la calidad de los conjuntos de datos resultantes. Dado que los expertos son un recurso escaso, el tiempo de etiquetado debe utilizarse de forma eficiente. Es por esto que los métodos de etiquetado asistidos se apoyan en diferentes técnicas como el Active Learning (AL), las técnicas de visualización, o una combinación de ambas, para tratar con un tráfico de red más realista y acelerar el proceso de etiquetado [3, 4, 5, 6, 7].

Frente a los enormes desafíos que presenta el tener que lidiar con los volúmenes de tráfico de las redes actuales como ser almacenamiento, capacidad de cómputo y el gran esfuerzo humano requerido para el análisis, el proceso de etiquetado podría verse como un problema que presenta pocas diferencias respecto del proceso de detección de comportamiento malicioso. Sin embargo, en un trabajo reciente presentado por el director del presente proyecto [8], se observó que existen diferencias significativas en el proceso de etiquetado de tráfico frente a un problema de detección de tráfico convencional, como el que lleva a cabo un SDI. Estas diferencias pueden enmarcarse bajo los siguientes aspectos:

Oportunidad: en el proceso de etiquetado no existe la necesidad de realizar la detección en un margen de tiempo relativamente corto. Durante el proceso de etiquetado es posible contar con el tiempo necesario para confirmar la potencial presencia de malware.

Relevancia: un error de etiquetado no presenta la misma importancia para un sistema de detección en tiempo real que para un sistema de creación de conjuntos de datos etiquetados. Como ejemplo, puede mencionarse el hecho de que en el contexto de un SDI, los falsos positivos pueden resultar una molestia para el usuario. Sin embargo, durante un proceso de etiquetado simplemente representa parte del ruido que pudiera ocurrir en el conjunto de datos resultante.

Calidad: durante el proceso de etiquetado se busca fundamentalmente obtener etiquetas de calidad que representen las características de la red. El objetivo es contar con datos de calidad, a fin de alimentar los distintos algoritmos de AA. Como consecuencia, no resulta necesario etiquetar

todas las trazas de tráfico que pertenezcan a un mismo comportamiento malicioso, sino que bastará con etiquetar adecuadamente un subconjunto lo suficientemente representativo.

Economía: el proceso de etiquetado no presenta consecuencias económicas en lo inmediato para el sistema vulnerado. Es decir, frente a un comportamiento malicioso no detectado, en principios no existe una consecuencia más allá de datos espurios para la construcción de un modelo estadístico de predicción. Mientras que, en el caso de un sistema de detección, el no reconocimiento de un comportamiento malicioso puede ocasionar pérdidas importantes a la organización donde se encuentre la red.

Curiosamente, estas diferencias no son tenidas en cuenta en la mayoría de los enfoques orientados al etiquetado de trazas de tráfico, y generalmente se aplica un enfoque que no se diferencia mucho del problema de reconocimiento de comportamiento malicioso en el tráfico de red. La realidad es que hoy en día el reconocimiento de comportamiento malicioso en el tráfico red es una tarea que demanda una gran cantidad de recursos tanto de índole computacionales como humanos. Es por ello que en la actualidad estos sistemas basados en AA solo han sido desplegados en ambientes controlados donde el volumen de tráfico evaluado dista mucho de los volúmenes actuales reales, lo que resta peso práctico a los enfoques. De esta manera, muchas de las tareas de análisis y reconocimiento se realizan fuera de línea y el proceso de etiquetado se realiza de manera manual o asistida por herramientas de visualización.

Se trata de un inconveniente de gran impacto si se tiene en cuenta el hecho de que en muchos casos resulta necesario realizar un reentrenamiento periódico de los modelos de AA a causa de cambios observados en la distribución del tráfico de la red, es decir, que el etiquetado de un conjunto de datos estático en el tiempo resulta inadecuado para hacer frente a las características del tráfico de red actual. Además, es importante destacar que estos inconvenientes no sólo se aplican al problema de etiquetado en el contexto de detección de intrusiones, sino que también se observan en muchos otros problemas donde se requiere la aplicación de algoritmos de AA. Como consecuencia, la comunidad científica en general y de seguridad en particular, ha comenzado a trabajar en el desarrollo de enfoques orientados a datos (data-centric) en los cuales se pone el foco en el desarrollo de herramientas para la construcción, mantenimiento y evaluación de conjuntos de datos de manera eficiente y fácilmente repetible. Más aún, se ha observado que la comunidad se ha tenido que enfrentar al hecho de que si bien el desarrollo y despliegue de sistemas basados en AA es relativamente simple y rápido, su mantenimiento a lo largo del tiempo puede volverse difícil y costoso. Recientemente, para hacer frente a estos inconvenientes, han surgido distintos roles y procedimientos entre los que se destacan MLE (Machine Learning Engineer) y MLOps (Machine Learning Operations). El rol de MLE combina aspectos de la ingeniería de software tradicional aplicado a los procesos de generación y mantenimiento de aplicaciones basadas en AA. De manera similar, MLOps se focaliza en el desarrollo de procedimientos para simplificar y automatizar el proceso de despliegue y mantenimiento a gran escala de los modelos de AA en ambientes de producción. Entre los aspectos más importantes que involucran el concepto de MLOps se destaca el etiquetado de calidad de manera periódica requeridos para el ajuste y mantenimiento de los modelos de AA. La realidad es que muchos de estos roles todavía no han sido considerados en las propuestas para el desarrollo de Sistemas de detección de intrusos basados en AA.

El presente proyecto se enfoca en la aplicación de técnicas de aprendizaje por refuerzo (RL) dentro de un ciclo de AL de tal manera que permita desarrollar herramientas para facilitar el proceso de etiquetado de tráfico de red. El aprendizaje por refuerzo (RL) es un área del aprendizaje automático

que se focaliza en el desarrollo de agentes inteligentes capaces de realizar acciones en un entorno a fin de maximizar una recompensa acumulada.

Normalmente, en un flujo de trabajo de AL, se aplica una política fija para decidir cuándo preguntar al usuario por el valor para la etiqueta de una instancia en particular. La política más común está basada en el concepto de incertidumbre que consiste en seleccionar aquellas trazas de tráfico que se encuentran cerca de la frontera de decisión del modelo de AA, es decir, para las cuales la política no puede decidir de manera autónoma cómo etiquetar. Sin embargo, esta política puede no resultar adecuada cuando se trabaja con grandes volúmenes de datos. Las técnicas de RL son un enfoque de reciente aplicación en el aprendizaje de una política dinámica de AL a partir de los datos. Al utilizar RL es posible aprender nuevas políticas que tengan en consideración otras recompensas como ser la experiencia o capacidad del usuario junto a aspectos distintivos del proceso de etiquetado como ser la calidad, oportunidad y relevancia, entre otros. De esta manera se evita depender de una única heurística a la hora de tomar la decisión de consultar al usuario. Esto último resulta fundamental para el desarrollo de sistemas de detección basados en técnicas de AA.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO.

El proyecto se enmarca en dos de las áreas de investigación del instituto de investigaciones del LABSIN, en particular la captura y procesamiento de datos a gran escala y el aprendizaje estadístico.

Para el desarrollo del presente proyecto pueden diferenciarse 4 etapas principales:

A) Análisis preliminar del problema. En esta etapa se incluye la recopilación de información bibliográfica sobre el tema poniendo especial énfasis en la aplicación de técnicas de aprendizaje por refuerzo a problemas relacionados con la temática etiquetado de conjuntos de datos. La intención es evaluar no solo las estrategias de etiquetado de tráfico sino otras que podrían ser fácilmente aplicables al problema formulado en el presente proyecto.

B) Generación de conjuntos de datos. Esta etapa se focaliza en la construcción o adaptación de uno (o varios) conjuntos de datos adecuados para algún problema de seguridad informática. En principio se considerará la utilización de los conjuntos previamente producidos en el marco del proyecto SIIP B06/036 bajo la dirección del Dr. Carlos Catania.

C) Algoritmos para aprendizaje de políticas. Se analizarán e implementarán algoritmos de etiquetado basados en técnicas clásicas de RL, como políticas del gradiente (Policy Gradient) o Q-learning adaptados para el problema de etiquetado de tráfico de red. Durante esta primera actividad se tendrá en cuenta una única acción posible que consiste en preguntar (o no) al usuario por el valor de una etiqueta. Para las diferentes implementaciones de algoritmos de RL, se considerará en primer lugar un mecanismo de recompensa enfocado principalmente al aspecto de calidad necesario para el etiquetado de trazas de tráfico. Esta actividad tiene como objeto determinar cómo interactúan los elementos básicos de un algoritmo de RL, y sus ventajas y limitaciones para su aplicación dentro de un flujo de trabajo de AL para el etiquetado de trazas de tráfico.

D) Experimentación. Diseño de los experimentos computacionales destinados a medir el rendimiento de los algoritmos desarrollados durante la actividad (C). Para este punto resulta necesario considerar métricas que vayan más allá de la precisión de las etiquetas (i.e. Exactitud, Sensibilidad, Sensitividad), sino otros aspectos necesarios para medir el rendimiento de los enfoques de

AL como por ejemplo la tasa de aprendizaje. Esta última métrica apunta a medir el número mínimo de trazas etiquetadas necesario para que el algoritmo de AA comience a funcionar con niveles aceptables de exactitud. Otra métrica de gran importancia que se considerará durante la experimentación es la tolerancia al ruido, la cual permite conocer el grado de conexiones incorrectamente etiquetadas por el usuario que permite el algoritmo de AA subyacente sin perder exactitud.

3 RESULTADOS ESPERADOS

Al término de los dos años de duración del plan de trabajo se pretende que se haya logrado:

1. Fortalecer la línea de investigación en la aplicación de modelos de aprendizaje automático relacionados con el tráfico de red.
2. Contar con el prototipo de una herramienta computacional que permita facilitar el proceso de etiquetado de tráfico de red dentro de un ciclo de AL. Dicha herramienta deberá ser capaz lidiar no solo con la precisión de las etiquetas sino también con su representatividad.
3. Se espera además contar con un marco de trabajo que facilite la evaluación de variantes aprendizaje por refuerzo a otros de problemas de clasificación más allá de la seguridad informática

4 FORMACIÓN DE RECURSOS HUMANOS

Se espera capacitar en el ámbito de la investigación a profesores y alumnos interesados en participar en un entorno académico y tecnológico innovador y a todos aquellos actores interesados en los resultados del proyecto.

Sobre la temática de este proyecto se está desarrollando el doctorado en ciencias de la computación de la Lic Tatiana Parlanti, en la Universidad Nacional del centro de la provincia de Buenos Aires.

5 BIBLIOGRAFÍA

[1] J. L. Guerra, C. A. Catania, and E. Veas. Active learning approach to label network traffic datasets. Journal of Information Security and Applications, 49:102388, 2019. [Indexada SCI/SCI-E].

[2] P. Torres, C. Catania, S. Garcia, and C. G. Garino. An analysis of recurrent neural networks for botnet detection behavior. In 2016 IEEE Biennial Congress of Argentina (ARGENCON), pages 1–6, 2016.

[3] N. Görnitz, M. Kloft, K. Rieck, and U. Brefeld. Toward Supervised Anomaly Detection.

[4] A. Beaugnon, P. Chifflier, and F. Bach. Ilab: An interactive labelling strategy for intrusion detection. In M. Dacier, M. Bailey, M. Polychro nakis, and M. Antonakakis, editors, Research in Attacks, Intrusions, and Defenses, pages 120–140, Cham, 2017. Springer International Publishing

[5] X. Fan, C. Li, X. Yuan, X. Dong, and J. Liang. An interactive visual analytics approach for network anomaly detection through smart labeling. Journal of Visualization, 22(5):955–971, 2019.

[6] Y. Yang, Z. Ma, F. Nie, X. Chang, and A. G. Hauptmann. Multi-Class Active Learning by Uncertainty Sampling with Diversity Maximization. International Journal of Computer Vision, 113(2):113–127, 2015.

[7] S. McElwee. Active learning intrusion detection using k-means clustering selection. In SoutheastCon 2017, pages 1–7, 2017.

[8] J. Guerra, C. Catania, and E. Veas. Datasets are not enough: Challenges in labeling network traffic, 2021. arxiv.org/abs/2110.05977