Informe del Artículo “Machine learning-based PortScan attacks detection using OneR classifier”

Resumen:

Objetivos: El objetivo principal del estudio era desarrollar un método de alta precisión y eficiencia para detectar ataques de escaneo de puertos (PortScan). Los autores buscaron abordar las limitaciones de las soluciones tradicionales de detección de intrusiones mediante el uso de aprendizaje automático, proponiendo un sistema que pudiera distinguir eficazmente entre el tráfico de red benigno y el malicioso asociado con los escaneos de puertos. Un objetivo secundario clave fue optimizar el modelo mediante un método de selección de características para mejorar la velocidad y el rendimiento computacional.

Metodología: La investigación utilizó el conjunto de datos CICIDS2017, que contiene tráfico de red etiquetado como benigno o malicioso, incluyendo varios tipos de ataques. Se propuso una arquitectura que comienza con el preprocesamiento de los datos, seguido de una división del 50% para entrenamiento y 50% para pruebas. El paso más importante fue la aplicación de un método híbrido de selección de características, que combinó el Filtrado por Varianza y la Relación de Ganancia de Información (IGR) para reducir el número de características de 84 a las 13 más relevantes. Finalmente, se entrenaron y compararon dos algoritmos de clasificación basados en reglas:

JRip y OneR, para clasificar el tráfico. La evaluación del rendimiento se basó en métricas como exactitud (accuracy), precisión, sensibilidad (recall) y F-Measure.

Resultados Principales: El modelo propuesto que utiliza el clasificador JRip demostró un rendimiento superior en comparación con el clasificador OneR. El modelo JRip alcanzó una exactitud del 99.84%, una precisión del 99.80%, una sensibilidad del 99.80% y una F-Measure del 99.80%. Estos resultados indicaron que la combinación del método de selección de características híbrido y el algoritmo JRip fue altamente efectiva para identificar ataques PortScan dentro del conjunto de datos CICIDS2017.

Discusión de Fortalezas, Limitaciones y Posibles Mejoras

Fortalezas:

- Alta Precisión y Eficiencia: El estudio logra una exactitud excepcionalmente alta (99.84%), lo que demuestra la efectividad de la metodología en el contexto del dataset utilizado.

- Selección de Características Innovadora: El enfoque híbrido para la selección de características es una fortaleza clave. Reducir la dimensionalidad de 84 a 13 características no solo optimiza la eficiencia computacional, sino que también minimiza el riesgo de sobreajuste al eliminar el ruido.

- Interpretabilidad del Modelo: El uso de clasificadores basados en reglas como JRip y OneR es ventajoso, ya que sus modelos son más fáciles de interpretar por analistas de seguridad en comparación con modelos de "caja negra" como las redes neuronales profundas.

Limitaciones:

- Dependencia de un Dataset Antiguo: La principal limitación es el uso exclusivo del conjunto de datos CICIDS2017, recopilado en 2017. Las tácticas de ciberataque, incluidas las técnicas de escaneo de puertos sigilosos, evolucionan constantemente. Por lo tanto, la eficacia del modelo contra amenazas contemporáneas no está garantizada.

- Riesgo de Sobreajuste (Overfitting): Una exactitud tan elevada, obtenida a partir de un único conjunto de datos, sugiere un posible sobreajuste. El modelo podría estar perfectamente ajustado a los patrones específicos de CICIDS2017, pero podría no generalizar bien a otros entornos de red.

- Métricas de Evaluación Incompletas: El análisis omite métricas cruciales para los sistemas de detección de intrusiones, como la Tasa de Falsos Positivos (FPR). En un entorno real, una alta tasa de falsos positivos puede generar "fatiga de alertas" y hacer que el sistema sea impracticable, incluso con una alta exactitud general.

- Comparativa de Algoritmos Limitada: El estudio se centra únicamente en JRip y OneR. No se compara el rendimiento del modelo JRip con otros algoritmos de

aprendizaje automático más robustos y comúnmente utilizados en ciberseguridad, como Random Forest, SVM o XGBoost.

Posibles Mejoras:

- Validación con Datasets Modernos: Para demostrar la relevancia actual del modelo, sería fundamental validarlo con conjuntos de datos más recientes que reflejen las tácticas de ataque actuales (el propio estudio sugiere usar CICDDoS2019).

- Implementar Validación Cruzada: En lugar de una única división 50/50, emplear técnicas como la validación cruzada k-fold proporcionaría una evaluación más robusta y fiable de la capacidad de generalización del modelo.

- Ampliar la Evaluación de Métricas: Incluir y analizar la Tasa de Falsos Positivos (FPR) y las curvas ROC/AUC para ofrecer una visión completa del rendimiento del modelo en un escenario práctico.

- Benchmarking Extensivo: Realizar un análisis comparativo del modelo JRip contra un espectro más amplio de algoritmos de aprendizaje automático y aprendizaje profundo para contextualizar mejor sus resultados.

Aportes al Campo de la Ciberseguridad

A pesar de sus limitaciones, el artículo realiza contribuciones valiosas al campo de la ciberseguridad:

- Refuerza el valor del Aprendizaje Automático: Demuestra con éxito cómo el aprendizaje automático puede automatizar y mejorar significativamente la precisión en la detección de actividades de reconocimiento como los escaneos de puertos.

- Destaca la Importancia de la Ingeniería de Características: Subraya que una selección de características inteligente y eficiente es tan crucial como la elección del algoritmo clasificador. El método híbrido propuesto es un ejemplo práctico de cómo reducir la complejidad del modelo y aumentar su eficiencia.

- Proporciona un Baseline de Alto Rendimiento: Establece un punto de referencia (benchmark) de rendimiento muy alto para la detección de PortScan en el dataset CICIDS2017 utilizando un modelo interpretable. Esto puede servir de base para que futuras investigaciones comparen y mejoren sus propios modelos.

Reflexión Personal

Este estudio presenta una solución metodológicamente sólida y bien ejecutada para un problema fundamental en la ciberseguridad defensiva. El enfoque en la reducción de características es particularmente relevante, ya que la eficiencia computacional es clave para la detección en tiempo real. Los resultados obtenidos son, en papel, impresionantes.

Sin embargo, desde una perspectiva práctica, el análisis crítico revela una brecha entre el éxito en un entorno de laboratorio controlado y la viabilidad en un entorno de producción real. La dependencia de un dataset obsoleto y la ausencia de métricas críticas como la Tasa de Falsos Positivos limitan la confianza en su aplicabilidad inmediata. Un analista de seguridad necesita un sistema que no solo sea preciso, sino también fiable y que no lo abrume con falsas alarmas.

En conclusión, el artículo es un excelente ejercicio académico que demuestra el potencial de los clasificadores basados en reglas y la selección inteligente de características. No obstante, para considerarlo una solución robusta y lista para el despliegue, sería indispensable una validación más rigurosa frente a amenazas modernas y en condiciones que simulen de cerca un entorno de red real. Es un paso prometedor, pero que requiere un desarrollo y pruebas adicionales para madurar completamente.

Referencias

Kareem, M. I., Abood, M. J. K., & Ibrahim, K. (2023). Machine learning-based PortScan attacks detection using OneR classifier. Bulletin of Electrical Engineering and Informatics, 12(6), 3690–3696. https://doi.org/10.11591/eei.v12i6.4142