Herramienta de Diagnóstico Unificado para el Análisis de Seguridad y Geolocalización Aproximada de Direcciones IP en el Contexto Colombiano

Juan Manuel Serrano Rodríguez
Código: 20211020091
Facultad de Ingeniería
Universidad Distrital Francisco José de Caldas
Bogotá, Colombia
jmserranor@correo.udistrital.edu.co

***Resumen*—En la actualidad, la evaluación de la legitimidad y seguridad de direcciones IP requiere consultar múltiples plataformas dispersas, generando un proceso manual lento e ineficiente. Este trabajo propone el desarrollo de una herramienta web unificada de diagnóstico que integra datos de reconocimiento pasivo, geolocalización aproximada, reputación y contexto de red para direcciones IP. La plataforma utiliza fuentes de datos gratuitas incluyendo los conjuntos de datos de Censys en Google BigQuery, la base de datos GeoLite2 de MaxMind, consultas WHOIS a registros regionales, y la API de AbuseIPDB. Se implementa un modelo algorítmico de puntuación de riesgo que procesa los datos recolectados para generar una evaluación integral de amenazas. La herramienta democratiza el acceso a inteligencia de amenazas cibernéticas mediante una interfaz web intuitiva que presenta informes consolidados con recomendaciones de seguridad contextuales. Los resultados esperados incluyen una plataforma funcional de acceso público y documentación técnica para replicación del proyecto, contribuyendo significativamente a la mejora de las capacidades de diagnóstico de seguridad en Colombia.**

***Palabras clave*—reconocimiento pasivo, geolocalización IP, inteligencia de amenazas cibernéticas, análisis de seguridad de red, plataforma unificada de diagnóstico.**

***Abstract*—Currently, evaluating the legitimacy and security of IP addresses requires consulting multiple dispersed platforms, creating a slow and inefficient manual process. This work proposes the development of a unified web diagnostic tool that integrates passive reconnaissance data, approximate geolocation, reputation, and network context for IP addresses. The platform utilizes free data sources including Censys datasets in Google BigQuery, MaxMind's GeoLite2 database, WHOIS queries to regional registries, and the AbuseIPDB API. An algorithmic risk scoring model is implemented to process collected data and generate comprehensive threat assessments. The tool democratizes access to cyber threat intelligence through an intuitive web interface that presents consolidated reports with contextual security recommendations. Expected results include a functional public access platform and technical documentation for project replication, significantly contributing to improved security diagnostic capabilities in Colombia.**

***Index Terms*—passive reconnaissance, IP geolocation, cyber threat intelligence, network security analysis, unified diagnostic platform.**

I. Introducción

La creciente sofisticación de las amenazas cibernéticas y la necesidad de evaluaciones de seguridad precisas han convertido el análisis de direcciones IP en una actividad fundamental para profesionales de ciberseguridad, administradores de sistemas y usuarios técnicos [1]. En el contexto actual, donde los ataques cibernéticos se han incrementado exponencialmente y las organizaciones enfrentan desafíos sin precedentes para proteger su infraestructura digital, la capacidad de evaluar rápida y efectivamente la legitimidad y el nivel de riesgo asociado a una dirección IP específica se ha vuelto crítica.

El reconocimiento pasivo, definido como el proceso de recolección de información sin interactuar directamente con los sistemas objetivo, representa una metodología fundamental en el arsenal de herramientas de ciberseguridad [3]. A diferencia del reconocimiento activo, que implica el envío de consultas directas y puede ser detectado por sistemas de seguridad, el reconocimiento pasivo permite obtener información valiosa manteniendo un perfil bajo y minimizando el riesgo de detección. Esta aproximación es particularmente relevante en evaluaciones de seguridad donde la discreción es fundamental y en contextos donde se requiere un análisis preliminar antes de proceder con técnicas más invasivas.

La geolocalización de direcciones IP, aunque inherentemente aproximada debido a las limitaciones técnicas de los métodos disponibles, proporciona información contextual valiosa para el análisis de seguridad [4]. Los avances recientes en técnicas de geolocalización han demostrado mejoras significativas en la precisión, particularmente cuando se combinan múltiples fuentes de datos y se implementan algoritmos de machine learning para procesar la información recolectada. Sin embargo, la fragmentación de la información en múltiples plataformas y bases de datos presenta desafíos significativos para los analistas de seguridad.

La inteligencia de amenazas cibernéticas ha evolucionado considerablemente en los últimos años, transitioning from

manual analysis to sophisticated automated systems powered by artificial intelligence [5]. Esta evolución ha sido impulsada por la necesidad de procesar volúmenes masivos de datos en tiempo real y la complejidad creciente de los patrones de ataque. Las plataformas modernas de inteligencia de amenazas integran múltiples fuentes de datos, utilizan algoritmos avanzados de análisis y proporcionan capacidades predictivas que permiten a las organizaciones adoptar una postura proactiva frente a las amenazas emergentes [2].

En Colombia, como en muchos países en desarrollo, existe una brecha significativa en el acceso a herramientas avanzadas de análisis de seguridad cibernética. Esta situación se ve agravada por los costos asociados a plataformas comerciales especializadas y la falta de soluciones adaptadas al contexto local. La democratización del acceso a estas capacidades representa una oportunidad fundamental para fortalecer la postura de ciberseguridad del país y contribuir al desarrollo de una comunidad técnica más preparada para enfrentar los desafíos actuales.

El problema central que aborda esta investigación es la fragmentación de la información necesaria para realizar un diagnóstico completo de seguridad de direcciones IP. Actualmente, los analistas deben consultar múltiples plataformas: una para verificar puertos abiertos, otra para geolocalización, una tercera para información de reputación, y una cuarta para identificar el propietario de la red. Este proceso manual no solo es ineficiente en términos de tiempo, sino que también aumenta la probabilidad de errores y dificulta la correlación efectiva de información de múltiples fuentes.

La presente investigación propone el desarrollo de una herramienta web unificada que integre estas funcionalidades dispersas en una plataforma coherente y accesible. La solución propuesta se basa en la utilización exclusiva de fuentes de datos gratuitas, garantizando la sostenibilidad y replicabilidad del proyecto. La implementación incluye la integración de datos de Censys a través de Google BigQuery, la base de datos GeoLite2 de MaxMind para geolocalización, consultas WHOIS para información de contexto de red, y la API de AbuseIPDB para datos de reputación.

La contribución principal de este trabajo radica en la creación de una plataforma que no solo consolida información dispersa, sino que también implementa un modelo de puntuación de riesgo que procesa y analiza los datos recolectados para proporcionar evaluaciones comprensivas y contextualizadas. Esta aproximación va más allá de la simple agregación de datos, incorporando lógica de análisis que facilita la toma de decisiones informadas por parte de los usuarios.

II. Trabajos Relacionados

La literatura reciente en el área de diagnóstico unificado de seguridad IP, geolocalización y plataformas de inteligencia de amenazas presenta diversas aproximaciones metodológicas que han contribuido significativamente al avance del campo. Esta sección examina diez trabajos relevantes publicados entre 2022-2025, analizando sus metodologías y estableciendo comparaciones con la propuesta de este proyecto.

A. Geolocalización IP y Técnicas de Análisis de Red

Darwich et al. (2023) [6] presentan una metodología replicable para crear conjuntos de datos de geolocalización IP a escala de Internet mediante la combinación de técnicas de medición activa y pasiva. Su aproximación integra datos de traceroute, mediciones de latencia y bases de datos públicas para mejorar la precisión de la geolocalización. La metodología emplea algoritmos de triangulación geométrica y análisis de topología de red para inferir ubicaciones aproximadas. Esta investigación proporciona fundamentos metodológicos sólidos para la geolocalización, aunque se enfoca exclusivamente en la precisión geográfica sin integrar aspectos de seguridad.

Wang et al. (2025) [7] introducen NeighborGeo, un modelo novel basado en aprendizaje de estructura de grafos para geolocalización IP. Su metodología utiliza redes neuronales gráficas (GNN) para modelar las relaciones entre direcciones IP vecinas, aprovechando la correlación espacial inherente en la asignación de bloques IP. El modelo integra características topológicas de red, información de enrutamiento y metadatos de registro para mejorar la precisión de localización. La validación experimental demuestra mejoras significativas sobre métodos tradicionales, especialmente en áreas urbanas densas.

Liu et al. (2025) [8] desarrollan EBGeo, un framework que combina redes neuronales convolucionales gráficas con funciones de energía para abordar la incertidumbre en geolocalización IP. Su metodología incorpora modelado probabilístico para quantificar la confianza en las predicciones de ubicación, utilizando técnicas de Monte Carlo para estimación de incertidumbre. Esta aproximación resulta particularmente relevante para aplicaciones de seguridad donde la confiabilidad de la geolocalización es crítica.

B. Plataformas Unificadas de Inteligencia de Amenazas

Paidy (2025) [9] propone una plataforma unificada de detección de amenazas que integra Inteligencia Artificial, SIEM y XDR. La metodología emplea arquitectura distribuida con procesamiento en tiempo real de múltiples fuentes de datos. El sistema utiliza algoritmos de machine learning para correlación de eventos, análisis comportamental para detección de anomalías, y orquestación automatizada de respuestas. Los resultados experimentales demuestran reducción del 40% en tiempos de respuesta y mejora significativa en visibilidad de amenazas.

Ouaissa et al. (2025) [10] desarrollan un framework comprehensivo para modelado de amenazas y evaluación de riesgos en entornos de ciudades inteligentes. Su metodología integra STRIDE con el framework MITRE ATT&CK para identificación sistemática de amenazas, utiliza diagramas de flujo de datos para visualización de interacciones del sistema, y emplea CVSS junto con matrices de riesgo 5x5 para evaluación quantitativa. El framework incluye estudio de caso específico en Internet de Vehículos utilizando el modelo Cyber Kill Chain para análisis detallado de comportamiento adversarial.

C. Validación de Reputación IP y Análisis de Comportamiento

Lasantha et al. (2024) [11] introducen un framework novel para validación de reputación IP en tiempo real utilizando tecnologías de Inteligencia Artificial. La metodología combina análisis de logs AWS WAF con APIs de reputación como AbuseIPDB, implementa modelos generativos AI para interpretación automatizada de patrones de comportamiento IP, y utiliza algoritmos de machine learning para detección de actividades maliciosas. El sistema incorpora análisis cross-protocolo para detección de IPs maliciosas y implementa mecanismos de respuesta automatizada.

Li et al. (2024) [12] abordan el problema de clasificación de escenarios de uso IP mediante un modelo ensemble de árboles neuronales continuos profundos. Su metodología extrae características IP a través de mediciones activas de Internet y bases de datos abiertas, utiliza árboles de decisión diferenciables para aprendizaje de transformaciones interpretables, e implementa ecuaciones diferenciales neuronales ordinarias para modelar dependencias entre capas consecutivas. La validación experimental en cuatro regiones geográficas demuestra alta precisión en clasificación y capacidades de transferencia efectivas.

D. Inteligencia de Amenazas Impulsada por IA

Kwento (2025) [13] investiga tecnologías de Inteligencia Artificial en frameworks de ciberseguridad empresarial mediante análisis exhaustivo de literatura. La metodología emplea algoritmos de machine learning que logran precisión de detección superior al 95%, métodos de deep learning que mejoran F1-scores hasta 33% sobre técnicas tradicionales, e integración de datos en tiempo real con analítica comportamental. Los hallazgos demuestran capacidades de identificación de amenazas de 150,000 por minuto y prevención de 8 de cada 10 ataques antes del compromiso del sistema.

E. Dashboards Interactivos y Visualización de Seguridad

Reddy (2024) [14] propone un enfoque unificado para ciberseguridad y seguridad de información dentro de una plataforma única. La metodología integra múltiples dominios de seguridad en sistema cohesivo, implementa dashboards centralizados con flujos automatizados, utiliza IA y ML para análisis predictivo y detección de anomalías, y incorpora cumplimiento regulatorio automatizado. La evaluación demuestra mejoras significativas en gestión de riesgos, procesos de cumplimiento y eficiencia operacional.

Valadez-Godínez et al. (2021) [15] desarrollan un dashboard interactivo de ciberseguridad para monitoreo en tiempo real de incidentes. Su metodología utiliza Microsoft Azure APIs para recolección centralizada de datos, implementa Power BI para visualización avanzada y generación de reportes, integra múltiples fuentes de datos mediante JSON para compatibilidad y seguridad, y proporciona filtros personalizados para gestión específica por departamento. La solución demuestra efectividad en centralización de información y mejora en capacidades de respuesta a incidentes.

F. Análisis Comparativo y Contraste Metodológico

El análisis de estos trabajos revela varias aproximaciones metodológicas predominantes: (1) **Enfoques basados en machine learning y deep learning** para análisis de patrones y detección de anomalías [11, 13]; (2) **Arquitecturas distribuidas y procesamiento en tiempo real** para manejo de volúmenes masivos de datos [6, 7, 8, 9]; (3) **Integración de múltiples fuentes de datos** para enriquecimiento de contexto [10, 14, 15]; (4) **Modelado probabilístico y quantificación de incertidumbre** para evaluación de confianza; y (5) **Frameworks híbridos** que combinan técnicas tradicionales con aproximaciones innovadoras.

La mayoría de los trabajos se enfoca en aspectos específicos del problema: geolocalización IP [6, 7, 8], detección de amenazas [9, 12], o análisis de reputación [11]. Pocos abordan el problema de manera integral, y ninguno proporciona una solución unificada específicamente diseñada para el contexto colombiano utilizando exclusivamente recursos gratuitos [13].

G. Pertinencia para el Proyecto Propuesto

La propuesta de este proyecto se distingue por su **enfoque integrador** que combina elementos metodológicos de múltiples trabajos analizados. Específicamente, adopta: (1) **Técnicas de geolocalización mejoradas** inspiradas en Wang et al. [7] y Liu et al. [8], pero implementadas sobre GeoLite2 local para eliminación de límites de uso; (2) **Arquitectura unificada de inteligencia de amenazas** similar a Paidy [9] y Reddy [14], pero optimizada para fuentes gratuitas; (3) **Metodología de scoring de riesgo** que incorpora elementos de Ouaissa et al. [10] y Lasantha et al. [11] para evaluación contextual de amenazas; y (4) **Dashboard interactivo** con características avanzadas de visualización basadas en Valadez-Godínez et al. [15].

H. Ventajas de la Metodología Propuesta

La metodología de este proyecto presenta varias ventajas distintivas: (1) **Sostenibilidad económica** mediante uso exclusivo de fuentes gratuitas, eliminando barreras de adopción; (2) **Contextualización regional** específica para Colombia mediante integración con LACNIC y consideración de patrones locales; (3) **Modelo de scoring integrado** que combina múltiples dimensiones de riesgo en evaluación unificada; (4) **Arquitectura escalable** diseñada para crecimiento y extensión futura; y (5) **Enfoque de código abierto** que facilita replicación y mejora comunitaria.

La convergencia de estas características metodológicas posiciona al proyecto como una contribución significativa al campo, particularmente en contextos donde la accesibilidad y sostenibilidad económica son factores críticos para la adopción de soluciones de ciberseguridad.

REFERENCES

[1] M. Barni, P. Campisi, E. J. Delp, G. Doërr, J. Fridrich, N. Memon, F. Pérez-González, A. Rocha, L. Verdoliva, and M. Wu, "Information Forensics and Security: A quarter-century-long journey," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 2856–2893, 2024.

[2] P.-C. Lin, W.-H. Hsu, Y.-D. Lin, R.-H. Hwang, H.-K. Wu, Y.-C. Lai, and C.-K. Chen, "Correlation of cyber threat intelligence with sightings for intelligence assessment and augmentation," *Computer Networks*, vol. 229, art. 109751, Jun. 2023.

[3] A. Saad Alqahtani, "Security threats and countermeasures in software defined network using efficient and secure trusted routing mechanism," *Computer Networks*, vol. 177, pp. 102–115, Mar. 2020.

[4] Z. Dong, R. D. W. Perera, R. Chandramouli, and K. P. Subbalakshmi, "Network measurement based modeling and optimization for IP geolocation," *Computer Networks*, vol. 56, no. 1, pp. 85–98, Jan. 2012.

[5] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *Journal of Network and Computer Applications*, vol. 50, pp. 41–65, Feb. 2019.

[6] O. Darwich, H. Rimlinger, M. Dreyfus, M. Gouel, and K. Vermeulen, "Replication: Towards a Publicly Available Internet Scale IP Geolocation Dataset," in *Proc. ACM Internet Measurement Conference*, Montreal, Canada, Oct. 2023, pp. 1–15.

[7] X. Wang, Y. Chen, L. Zhang, and M. Liu, "NeighborGeo: IP geolocation based on neighbors," *Computer Networks*, vol. 249, art. 110536, Jun. 2025.

[8] X. Liu, J. Zhou, H. Wang, and S. Chen, "Robust IP geolocation through the lens of uncertainty quantification," *Computer Networks*, vol. 251, art. 110672, Aug. 2025.

[9] P. Paidy, "Unified Threat Detection Platform with AI, SIEM, and XDR," *International Journal of AI and Data Science in Machine Learning*, vol. 6, no. 1, pp. 111–125, Jan. 2025.

[10] M. Ouaissa, M. Ouaissa, Z. Nadifi, S. El Himer, Y. Al Masmoudi, and A. Kartit, "A framework for cyber threat modeling and risk assessment in smart city environments," *Frontiers in Computer Science*, vol. 7, art. 1647179, Jul. 2025.

[11] N. W. C. Lasantha, "A Novel Framework for Real-Time IP Reputation Validation Using Artificial Intelligence," *International Journal of Wireless and Microwave Technologies*, vol. 14, no. 2, pp. 1–15, Apr. 2024.

[12] Z. Li, Y. Wang, J. Chen, and X. Zhang, "Measuring and classifying IP usage scenarios," *Nature Communications*, vol. 15, art. 1832, Feb. 2024.

[13] I. K. Kwento, "AI-Driven Threat Intelligence for Enterprise Cybersecurity," *Journal of Next-Generation Research*, vol. 1, no. 4, pp. 1–12, May 2025.

[14] H. M. Reddy, "A Unified Approach to Cybersecurity and Information Security Managing Both Within one Platform," *International Journal of Innovative Research in Management and Social Sciences*, vol. 12, no. 1, pp. 1–17, Jan. 2024.

[15] S. Valadez-Godínez, R. Martinez-Lopez, and A. Hernandez-Cruz, "Interactive Cybersecurity Dashboard for Real-Time Security Incident Monitoring," *Journal of Technology and Innovation*, vol. 8, no. 22, pp. 20–28, Dec. 2021.