

UNIVERSIDAD EAN

FACULTAD DE INGENIERÍA

INGENIERÍA DE SISTEMAS

DETECCIÓN TEMPRANA DE ATAQUES CIBERNÉTICOS AVANZADOS (APT)
MEDIANTE MODELOS OCULTOS DE MARKOV APLICADOS AL ANÁLISIS DE
COMPORTAMIENTOS EN SISTEMAS DE SEGURIDAD

PROYECTO DE GRADO

AUTOR: ANDRÉS FELIPE CORTES ARANDIA

TUTORA: LINA MARÍA CHACÓN RIVERA

BOGOTÁ D.C.

2026

Tabla de contenido

Resumen	4
Abstract	5
Introducción	6
Marco teórico	10
Ataques cibernéticos avanzados y sistemas de detección	10
Modelos Ocultos de Markov	11
Estado del arte (metodología PRISMA)	12
Metodología	17
Resultados	20
Análisis de resultados.....	23
Discusión.....	25
Plan de divulgación	27
Análisis de viabilidad.....	28
Fuentes de inversión y participación de semilleros.....	29
Conclusiones	30
Referencias	32

Índice de tablas

Tabla 1. Estudios comparativos seleccionados (2019–2025)	15
Tabla 2. Caracterización de los datasets utilizados en el entrenamiento y la validación.....	20
Tabla 3. Desempeño global sobre el conjunto de prueba unificado.....	21
Tabla 4. Desempeño del HMM jerárquico desagregado por fase del ciclo de vida APT	22
Tabla 5. Resultados de la simulación operativa sobre Elastic Stack.....	23
Tabla 6. Comparación directa entre el modelo de referencia y el modelo propuesto	25
Tabla 7. Simulación financiera del despliegue por perfil organizacional	28

Resumen

Las amenazas persistentes avanzadas, conocidas por su sigla en inglés APT, pertenecen a los problemas de seguridad más difíciles de resolver para las organizaciones colombianas. A diferencia de los ataques convencionales, que buscan un efecto rápido y más fácil de detectar, las APT operan durante semanas o meses dentro de los sistemas comprometidos sin disparar ninguna alerta. Cuando finalmente son descubiertas, el daño acumulado suele ser irreversible. Los métodos de detección vigentes los cuales están basados en bases de firmas maliciosas conocidas o en reglas estáticas predefinidas, procesan cada evento de seguridad de manera independiente y son incapaces de leer el significado que emerge cuando esos eventos se encadenan en el tiempo. El presente proyecto responde a esa limitación con un sistema de detección temprana construido sobre modelos ocultos de Márkov (HMM) de arquitectura jerárquica, entrenado sobre secuencias de eventos reales provenientes de los datasets DARPA Transparent Computing, LANL y cierto grupo de logs extraídos de un ataque simulado usando máquinas virtuales Windows y Linux. El modelo infiere probabilísticamente la fase del ciclo de vida del ataque en que opera el adversario, produce alertas progresivas integradas en Elastic Stack y supera, en precisión, recall, F1-score, a métodos de referencia Random Forest y redes LSTM. Los hallazgos esperados confirman la viabilidad del enfoque para su aplicación en organizaciones que carecen de soluciones comerciales de alto costo (Centro Cibernético Policial, 2024, secc. “Cifras generales”; IBM Security, 2024, p. 8; Rabiner, 1989, p. 258; Ghafir et al., 2019, p. 99510).

Palabras clave: amenazas persistentes avanzadas, Modelos Ocultos de Markov, detección temprana de intrusiones, ciberseguridad, registros de seguridad, análisis secuencial de eventos.

Abstract

Advanced Persistent Threats (APTs) are among the most difficult cybersecurity challenges faced by organizations in Colombia. Unlike conventional cyberattacks, which are often designed to cause immediate and noticeable effects, APTs can remain hidden inside compromised systems for weeks or even months without being detected. During this time, attackers are able to carry out malicious activities while avoiding security controls. As a result, when the attack is finally discovered, the damage is often significant and difficult to undo. Most current detection methods rely on known malware signatures or predefined security rules. While these approaches can identify familiar threats, they usually examine security events separately and therefore have difficulty recognizing the patterns that emerge when multiple events are linked over time. This limitation makes it harder to detect sophisticated attacks that unfold gradually and in several stages. To address this problem, this project proposes an early detection system based on hierarchical Hidden Markov Models (HMMs), the model is trained using real-world event data from the DARPA Transparent Computing and Los Alamos National Laboratory (LANL) datasets, as well as logs generated from a controlled simulation of an APT attack carried out in virtualized Windows and Linux environments. By analyzing sequences of events, the system estimates the stage of the attack being carried out and generates progressive alerts through the Elastic Stack platform. Its performance is evaluated using precision, recall, and F1-score and compared with well-known approaches such as Random Forest and Long Short-Term Memory (LSTM) Networks, the expected results seek to demonstrate that this

approach can provide an effective and affordable alternative for organizations that do not have access to high-cost commercial cybersecurity solutions (Centro Cibernético Policial, 2024, secc. “Cifras generales”; IBM Security, 2024, p. 8; Rabiner, 1989, p. 258; Ghafir et al., 2019, p. 99510).

Keywords: advanced persistent threats, Hidden Markov Models, early intrusion detection, sequential event analysis, cybersecurity, security logs.

Introducción:

Colombia atraviesa una crisis silenciosa en materia de ciberseguridad. Entre enero y diciembre de 2024, el país registró más de 70.000 incidentes cibernéticos confirmados, lo que representa un incremento cercano al 19% respecto al año anterior (Centro Cibernético Policial, 2024, secc. “Cifras generales”). Sin embargo, la magnitud real del problema no se ve reflejada del todo en esos registros. Los incidentes que se cuentan son los que se logran detectar. Lo mas preocupante es lo contrario, es decir, los ataques que llevan meses o años activos dentro de los sistemas de sus victimas sin que nadie lo sepa.

América Latina y el Caribe registra el crecimiento más acelerado en incidentes cibernéticos reportados a nivel global, con una tasa promedio anual de 25% durante la última década, al tiempo que obtiene el puntaje mas bajo en preparación defensiva 10,2 sobre 20 de todas las regiones del mundo (Banco Mundial, 2024, párr. 3). Dentro de ese cuadro regional, los sectores más afectados son el público, que concentra el 21% de los ataques exitosos, y el financiero con el 13%, convirtiéndose ambos en el objetivo preferido

de los grupos que ejecutan amenazas persistentes avanzadas (Positive Technologies, 2024, secc. “Key findings”). Se estima que las organizaciones con equipos de seguridad establecidos tardan en promedio 204 días en detectar una brecha avanzada y otros 73 días adicionales en contenerla (IBM Security, 2024, p. 8), lo que significa que un atacante puede operar dentro de los sistemas durante casi 10 meses antes de que sus víctimas se den cuenta que están comprometidas. Para las pequeñas y medianas empresas colombianas las cuales constituyen más del 90% del tejido empresarial nacional (Confecámaras, 2022, p. 5) ese periodo es habitualmente mayor, y el impacto es proporcionalmente mas grave: el 68% de las organizaciones afectadas reportó pérdida directa de datos, el 78% sufrió periodos de inactividad operativa y el 34% enfrente un daño reputacional de carácter duradero (Infoblox, 2023, secc. “Key findings”).

El caso con mas evidencias documentales de este tipo de amenaza es el grupo Blind Eagle, también conocido como APT-C-36, activo desde 2018 y especializado en comprometer agencias gubernamentales con acceso a datos nacionales sensibles, instituciones financieras y sistemas de pago en línea (Check Point Research, 2025, párr. 4). Su efectividad no depende de herramientas altamente complejas, sino en una estrategia de ocultamiento sostenida, distribuye sus cargas maliciosas a través de plataformas legítimas como Google Drive, Dropbox y Github, evade los motores de detección por firma y diseña sus rutinas de infección para que la activación del malware pueda producirse con acciones tan comunes como hacer click derecho sobre un archivo (Darktrace, 2025, secc. “Attack overview”). Entre noviembre de 2024 y marzo de 2025, una sola campaña de este grupo comprometió más de 1.600 dispositivos colombianos en un solo día, una operación paralela generó cerca de 9000 infecciones en una semana, y la mayoría de los archivos empleados

no fueron detectados por ningún motor antivirus disponible (Check Point Research, 2025, párr. 6).

Los sistemas de detección convencionales están diseñados para reconocer eventos discretos tales como una firma maliciosa conocida, un valor que supera un límite establecido o un comportamiento que coincide con reglas predefinidas. No están diseñados para leer secuencias. Un inicio de sesión fuera del horario habitual no activa ninguna alarma. Un proceso poco frecuente ejecutado en segundo plano tampoco. Una transferencia de datos de volumen moderado hacia una dirección externa pasa desapercibida. Pero cuando esos tres eventos ocurren en ese orden, sobre el mismo sistema, en el transcurso de días o semanas, están describiendo las fases sucesivas de un ataque en progreso que ninguna regla fija es capaz de identificar. Los Modelos Ocultos de Markov (HMM) ofrecen una respuesta matemática directa a esa limitación, en lugar de evaluar cada evento por separado, modelan la secuencia completa de observaciones y, a partir de ella, infieren probabilísticamente el estado interno que las genera, en este caso, la fase del ciclo de vida del ataque en que opera el adversario Rabiner (1989, p. 258), cuyo trabajo constituye referencia canónica del campo, demostró que los HMM son especialmente poderosos para analizar sistemas donde el estado interno no es directamente observable y debe deducirse de sus efectos sobre señales medibles. Investigaciones recientes han confirmado que los APT exhiben patrones de transición entre fases lo suficientemente estables como para que un HMM entrenado sobre ellos pueda anticipar la siguiente acción del atacante antes de que el daño sea irreversible (Ghafir et al., 2019, p. 99512; Ou et al., 2025, secc. “Conclusions”).

El proyecto esta enfocado en desarrollar un método de detección temprana de ataques cibernéticos avanzados con el propósito de identificar de manera anticipada las fases críticas de un ataque antes de que lleguen a presentar consecuencias negativas de forma significativa. Para esto se analiza la secuencia de logs de seguridad buscando patrones que permitan revelar el estado y el progreso de la amenaza. Este objetivo es el que establece el alcance de la investigación en la evaluación de como un análisis centrado en la periodicidad y orden de los eventos de logs de seguridad puede mejorar la detección frente a los métodos tradicionales que consideran los incidentes de seguridad de manera aislada.

De esta manera, El propósito del proyecto no es solamente definir una solución tecnológica, sino también proporcionar resultados sobre la efectividad de enfoques basados en secuencias de eventos con el objetivo de identificar de forma temprana ataques cibernéticos avanzados.

La investigación se enfoca en identificar en que medida desarrollar un sistema basado en Modelos ocultos de Markov jerárquicos entrenados a partir de secuencias de logs de seguridad, permite la detección temprana de las fases de un ataque cibernético avanzado en un entorno controlado, en comparación con los métodos convencionales de detección. La hipótesis planteada en el proyecto establece que la observación a detalle de los logs de seguridad considerando el orden y concurrencia de estos, puede revelar indicios tempranos de ataques que los métodos convencionales de detección pasan por alto, ofreciendo la posibilidad de intervenir antes de que el daño sea significativo

La propuesta de solución consiste en un sistema de detección temprana construido sobre modelos ocultos de Markov jerárquicos de dos niveles, entrenado sobre los datasets DARPA Transparent Computing y LANL. El nivel 1 modela el comportamiento individual

de cada host a partir de los símbolos observacionales obtenidos por discretización K-means de los logs de seguridad, mientras que el nivel 2 correlaciona las secuencias inferidas en múltiples hosts para reconocer movimiento lateral y fases coordinadas del ciclo de vida de la amenaza. El sistema integra sus salidas en Elastic Stack mediante alertas progresivas etiquetadas por fase, lo cual permite a un analista priorizar la respuesta sin esperar a que el ataque alcance una fase irreversible. La validación compara el rendimiento del HMM jerárquico con redes LSTM utilizando precisión, recall, F1-score, tiempo medio de detección, FPR (Falsos positivos) y un indicador si la detección fue por fases del ataque o no. Además, se obtendrán métricas clave sobre el uso del modelo HMM alojado en una maquina virtual Ubuntu sobre una simulación de ataque realizada sobre logs generados con la herramienta Sysmon en una máquina virtual Windows y atacados desde una agente aislado y controlado en Caldera desde una máquina virtual Kali Linux

Marco teórico

Ataques cibernéticos avanzados y sistemas de detección:

Las amenazas persistentes avanzadas (APT) consisten en operaciones que persisten en el tiempo y a su vez son sigilosas, están diseñadas para infiltrarse en sistemas críticos sin ser detectadas, obtener acceso continuo y acceder a información sensible o comprometer funciones esenciales prolongadamente (Kaspersky, 2024, párr. 1). A diferencia de un ataque convencional cuyo objetivo es un daño inmediato, los APT buscan operar con la mayor discreción posible, camuflando cada acción para evitar ser detectados por mecanismos de detección basados en firmas o anomalías individuales (Darktrace, 2025, secc. “Attack overview”; Check Point Research, 2025, párr. 4). En Colombia y America

Latina, estos ataques han crecido de manera acelerada, afectando principalmente a los sectores público y financiero y a las pequeñas y medianas empresas, que constituyen la mayor parte del tejido empresarial colombiano y en su mayoría tienen defensas limitadas (Banco Mundial, 2024, párr. 3; Infoblox, 2023, secc. “Key findings”; Confecámaras, 2022, p. 5).

Operativamente, los APT siguen un ciclo de vida que incluye fases de acceso inicial, persistencia, escalamiento de privilegios, movimiento lateral y exfiltración hacia sistemas de mayor valor (Fortinet, 2024, párr. 2; Kaspersky, 2024, párr. 3). Cada fase deja evidencia en los logs de seguridad, como modificaciones sutiles del sistema, patrones de autenticación atípicos o conexiones a infraestructura de comando y control, que los sistemas de detección de intrusiones (IDS) pueden registrar y analizar (Ortega Candel, 2021, p. 45; Cano Martínez, 2022, p. 112). Sin embargo, los sistemas de detección tradicionales, basados en firmas o en detección de anomalías aisladas, presentan limitaciones frente a ataques diseñados para evadirlos. Las acciones de un atacante avanzado no siempre superan los umbrales de alerta ni coinciden con patrones definidos, lo que permite que el ataque permanezca oculto durante meses (Scarfone & Mell, 2007, p. 3-1; Galicia, 2024, p. 87). Por esta razón, surge la necesidad de métodos que analicen secuencias de eventos para anticipar las fases del ataque y generar alertas tempranas, sin depender de la detección de eventos aislados.

En Colombia, la protección frente a APT se encuentra respaldada por normativas y estrategias nacionales, como la Estrategia Nacional de Seguridad Digital 2025-2027 y las buenas practicas de MITRE ATT&CK, que guían la gestión de riesgos, la prevención de intrusiones y la planificación de monitoreo de sistemas críticos (MITRE Corporation, 2024,

secc. “Enterprise matrix”; Ministerio de Tecnologías de la Información y las Comunicaciones, 2024, p. 12).

Modelos Ocultos de Markov:

Los Modelos Ocultos de Markov (HMM) consisten en modelos matemáticos los cuales permiten inferir estados internos de un sistema a partir de observaciones, como los logs de seguridad generados a partir de una actividad APT. Parten de las cadenas de Markov, que modelan sistemas estocásticos donde el estado futuro depende únicamente del estado presente, simplificando la representación de la dinámica del sistema (Norris, 1997, p. 1; Lay, Lay & McDonald, 2016, p. 268)

Un HMM se define mediante tres componentes: la matriz de transición, que indica la probabilidad de pasar de un estado a otro, la matriz de emisión la cual relaciona los estados ocultos con las observaciones visibles, y la distribución inicial de los estados (Rabiner, 1989, p. 260). Mediante algoritmos como Forward-Backward, Viterbi y Baum-Welch, es posible calcular la probabilidad de secuencias de eventos, determinar la trayectoria más probable de los estados ocultos y ajustar los parámetros del modelo a partir de datos de entrenamiento. Este enfoque permite identificar patrones temporales en los logs que los métodos tradicionales no capturan, anticipando fases críticas de los ataques antes de que el daño se concrete (Ortega Candel, 2022, p. 203; Ghafir et al., 2019, p. 99512; Ou, Tsai & Wu, 2025, secc. “Conclusions”).

Estudios recientes han demostrado que los HMM jerárquicos, al considerar múltiples niveles de estados ocultos, mejoran la detección temprana de ataques de larga duración y sigilosos, superando en precisión y tiempo de detección a técnicas basadas en Random Forest o redes LSTM (Du, Zhang & Liu, 2023, p. 1047; Shen, Mariconti, Vervier

& Stringhini, 2022, p. 594; Sakthivelu & Vinoth Kumar, 2024, secc. “Results”). De esta manera, los HMM consolidan un enfoque que combina teoría matemática, análisis de secuencias y conocimiento del comportamiento de APT para ofrecer soluciones mas predictivas frente a amenazas avanzadas.

Estado del arte (metodología PRISMA)

Para la construcción del estado del arte se siguió con una estrategia de revisión de literatura orientada por los principios del protocolo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), el cual proporciona lineamientos para la identificación y selección sistemática de estudios relevantes. La búsqueda se realizó en bases de datos especializadas en ingeniería y ciencias de la computación: Scopus, Web Of Science (WOS), IEEE Xplore y ACM Digital Library.

La ecuación de búsqueda empleada fue la siguiente: (“Advanced Persistent Threat” OR “APT detection”) AND (“Hidden Markov Model” OR “HMM”) AND (“intrusion detection” OR “cybersecurity”). Como criterios de inclusión se consideraron artículos publicados entre 2019 y 2025, escritos en ingles o español, con validación empírica mediante datasets de seguridad o experimentos controlados. Se excluyeron artículos de opinión, revisiones narrativas sin metodología explicita y trabajos sin evaluación experimental reproducible.

La búsqueda arrojó diez resultados de documentos con relación directa al enfoque de esta investigación. No todos son artículos empíricos, Rabiner (1989) y Russell y Norvig (2020) son referencias teóricas, mientras que Scarfone y Mell (2007) corresponde a una guía normativa del NIST, los siete restantes si incluyen una validación experimental. El punto de partida sigue siendo Rabiner (1989), cuyo tutorial sobre HMM formalizo los

algoritmos Baum-Welch, Viterbi y Forward-Backward que hoy siguen siendo el núcleo computacional de cualquier implementación del modelo. Sin ese marco, los trabajos posteriores simplemente no existirían tal como lo conocemos. Entre los estudios con evaluación empírica, los mas cercanos al problema que Aquí se aborda son los de Ghafir et al. (2019) y Chadza et al. (2020), ambos con un enfoque en HMM aplicados a secuencias de alertas IDS. El primero reporta una precisión del 91,3% en la predicción de la siguiente acción del atacante, aunque trabaja con un dataset privado que no puede reproducirse. El segundo, en cambio, no propone un sistema nuevo, sino que compara seis variantes de inicialización del algoritmo Baum-Welch. Su hallazgo más útil para este proyecto es que la inicialización con K-means reduce de manera importante el numero de iteraciones necesarias para converger, algo que se tienen en cuenta directamente en el diseño metodológico. Milajerdi (2019) van por un camino distinto, su sistema HOLMES no usa HMM, sino grafos de procedencia sobre logs del kernel, y consigue detectar los nueve escenarios de ataque evaluados con baja tasa de falsas alarmas. Es un referente válido, pero su dependencia de instrumentación a nivel del sistema operativo lo hace difícil de desplegar en entornos como las organizaciones colombianas de tamaño medio. Chadza et al. (2020) comparan distintas variantes del algoritmo de entrenamiento Baum-Welch para HMM sobre el dataset DARPA 2000 que tiene ya más de veinte años, y ese es precisamente uno de sus limites, los patrones de ataque que captura ese conjunto de datos no representan las campañas APT actuales.

Los tres estudios más recientes del conjunto Sakthivelu y Vinoth Kumar (2024), Ou et al. (2025) y Du, Zhang y Liu (2023) coinciden en que los enfoques híbridos o jerárquicos superan a los HMM Planos cuando el ataque involucra movimiento lateral entre hosts.

Sakthivelu y Vinoth Kumar combinan HMM con grafos de procedencia y técnicas de ciberengaño, y logran reducir el tiempo de contención en un 34% respecto a un HMM convencional, el problema es que su validación se hizo sobre datos sintéticos, lo que deja la pregunta de si ese rendimiento se sostendría ante tráfico real. Du et al., en cambio, si trabajan con datasets de referencia y demuestran que un HMM jerárquico de dos niveles (host y red) mejora la detección del movimiento lateral frente a Random Forest y redes LSTM. Ese resultado es posiblemente el mas relevante para el presente proyecto, ya que valida exactamente la arquitectura propuesta. Ou et al (2025) aportan una variante diferente ya que combinan selección automática de variables con Elastic Net antes de entrenar el HMM y reportan un F1-score de 0,91 sobre un dataset gubernamental el cual es de acceso privado. En el extremo opuesto en cuanto a enfoque, Shen et al. (2022) con Tiresias muestran hasta donde llegan las redes recurrentes en este problema precisión de 0,93 en predicción del siguiente evento, lo que fija un umbral de referencia que cualquier propuesta basada en HMM tiene que discutir de manera honesta. En cuanto a Scarfone y Mell (2007), su guía del NIST no aporta metodología experimental, pero si el vocabulario y las métricas como precisión, recall, F1, tasas de falsos positivos que se usan de manera estándar para comparar sistemas de detección de intrusiones, y eso la convierte en una referencia esencial cuando se evalúa un producto de estas características.

La tabla 1 reúne los diez estudios con sus metodologías, variables, resultados y limitaciones. La mayoría de los trabajos validan sobre datos propietarios o sintéticos, pocos consideran infraestructuras heterogéneas de tamaño medio, y ninguno ha sido probado en el contexto organizacional colombiano. Esas brechas con las que justifican el desarrollo del sistema que se describe en las secciones siguientes.

Tabla 1

Estudios comparativos seleccionados (2019–2025)

Autor(es) / Año	Título del estudio	Objetivo principal	Metodología empleada	Variables analizadas	Principales hallazgos	Limitaciones reportadas	Relación con el proyecto
Rabiner (1989)	A tutorial on hidden Markov models and selected applications in speech recognition	Establecer los fundamentos matemáticos de los HMM y sus algoritmos de entrenamiento e inferencia.	Teórico-formal. Derivación matemática de los algoritmos Forward-Backward, Viterbi y Baum-Welch.	Matrices de transición y emisión, distribución inicial, secuencias de observaciones.	Formalización completa del marco HMM. Demostración de convergencia del algoritmo Baum-Welch.	No incluye validación empírica sobre datos de seguridad ni aplicaciones a detección de intrusiones.	Referencia canónica del marco matemático adoptado. Los algoritmos se implementan directamente en el entrenamiento del modelo.
Milajerdi et al. (2019)	HOLMES: Real-time APT detection through correlation of suspicious information flows	Detectar campañas APT en tiempo real correlacionando flujos de información sospechosos mediante grafos de procedencia.	Cuantitativo experimental. Sistema HOLMES sobre logs de auditoría del kernel; evaluado en 9 escenarios de ataque reales.	Flujos de información entre procesos, archivos y red; grafos de procedencia; tasa de falsos positivos.	Detección de todos los escenarios APT con alta precisión y baja tasa de falsas alarmas. Reducción de 1875:1 en tamaño del grafo.	Requiere instrumentación del kernel. Evaluado en entorno de laboratorio, no en producción a gran escala.	Referente para la correlación de eventos multi-etapa. El presente proyecto adopta la correlación secuencial, pero mediante HMM probabilístico en lugar de grafos.
Ghafir et al. (2019)	Hidden Markov models and alert correlations for the prediction of APTs	Construir un sistema de predicción de la siguiente acción del atacante mediante correlación de alertas IDS con HMMs.	Cuantitativo experimental. HMM entrenado con Baum-Welch y decodificación con Viterbi sobre alertas Snort de dataset propio.	Secuencias de alertas IDS, tácticas MITRE ATT&CK, tasa de falsos positivos.	Precisión del 91,3 % en predicción de la siguiente fase. Reducción significativa de falsos positivos.	Dataset privado. Sin evaluación de movimiento lateral entre hosts ni integración con SIEM.	Valida el uso de HMMs para correlación de alertas APT. El presente trabajo extiende el enfoque con arquitectura jerárquica e integración en Elastic Stack.
Chadza et al. (2020)	Analysis of hidden Markov model learning algorithms for multi-stage network attack detection	Comparar seis variantes del algoritmo de entrenamiento de HMMs para determinar cuál maximiza la precisión en ataques multi-etapa.	Cuantitativo comparativo. Seis variantes de Baum-Welch evaluadas sobre DARPA 2000 con alertas Snort, incluyendo inicialización K-means.	Precisión de detección del estado actual, predicción del siguiente estado y de la observación futura.	Baum-Welch con inicialización K-means supera todas las variantes; la inicialización tiene impacto crítico en la convergencia.	Dataset DARPA 2000 considerado desactualizado. Solo evalúa ataques DDoS de cinco pasos, no APTs completos.	Fundamenta la elección de Baum-Welch con K-means en el presente proyecto y justifica el uso de DARPA TC como benchmark más representativo.
Shen et al. (2022)	Tiresias: Predicting security events through deep learning	Predecir el siguiente evento de seguridad en una máquina a partir de observaciones previas usando redes neuronales recurrentes.	Cuantitativo experimental. RNN/LSTM entrenada sobre 3.400 millones de eventos de seguridad de un sistema comercial de prevención de intrusiones.	Secuencias de eventos de seguridad, tipo de evento siguiente, precisión, recall y F1-score.	Precisión de hasta 0,93 en predicción del siguiente evento. La memoria a largo plazo de las RNN es clave para el rendimiento.	Dataset comercial no reproducible públicamente. No modela explícitamente fases de ciclo de vida del ataque.	Referente de comparación para métodos de aprendizaje profundo. Su precisión es el umbral de referencia que el HMM jerárquico propuesto debe superar.

Du, Zhang y Liu (2023)	Hierarchical hidden Markov model for APT lateral movement detection	Detectar el movimiento lateral en ataques APT prolongados mediante un HMM de arquitectura jerárquica de dos niveles.	Cuantitativo experimental. HMM jerárquico (host-level y network-level) entrenado y evaluado sobre datasets IEEE TIFS con validación cruzada.	Trayectoria de movimiento lateral, patrones de autenticación entre hosts, precisión, recall y F1-score.	El HMM jerárquico supera a Random Forest y LSTM en F1-score y tiempo de detección para ataques de larga duración.	Evaluated sobre datasets cerrados del laboratorio. No incluye integración con plataformas SIEM de producción.	Valida directamente la arquitectura de dos niveles adoptada en el presente proyecto. Sus resultados sirven de línea base para la comparación experimental.
Sakthivelu y Vinoth Kumar (2024)	A multi-step APT attack detection using hidden Markov models by molecular magnetic sensors	Proponer detección reactiva y proactiva combinando HMMs con grafos de procedencia y técnicas de ciberengaño para contener el movimiento lateral.	Cuantitativo experimental. HMM combinado con análisis de grafos de procedencia y cyber deception. Validación en dataset sintético.	Trayectoria del ataque, movimiento lateral entre hosts, tiempo de detección y tasa de detección por fase APT.	Detección de movimiento lateral del 87 %. Reducción del tiempo de contención en 34 % frente a HMMs planos.	Dataset sintético no reproducible en entornos reales. No evalúa el rendimiento del sistema bajo carga elevada de eventos de log.	Valida la viabilidad del enfoque jerárquico para capturar movimiento lateral, incorporado en el nivel 2 del HMM de este proyecto.
Ou et al. (2025)	Detecting trajectory of targeted attack by Hidden Markov Model	Identificar trayectorias de ataques dirigidos en secuencias de eventos de red mediante un marco híbrido Elastic Net-HMM.	Cuantitativo experimental. Elastic Net para selección automática de features + HMM para modelado de transiciones latentes. Validación cruzada k-fold.	Features de logs de red, fases del ataque como estados ocultos, precisión, recall, F1-score y tiempo medio de detección.	Precisión del 93,7 % y F1-score de 0,91. Supera a LSTM y Random Forest en escenarios de ataque sigiloso de larga duración.	Un único dataset gubernamental privado. Sin consideración de entornos multi-atacante ni infraestructura heterogénea.	Referente directo del sistema propuesto. Este proyecto adopta el esquema HMM jerárquico y lo valida sobre datasets públicos (DARPA TC, LANL).
Scarfone y Mell (2007)	Guide to intrusion detection and prevention systems (IDPS) (NIST SP 800-94)	Proveer un marco de referencia para la clasificación, evaluación y despliegue de sistemas de detección y prevención de intrusiones.	Revisión técnica normativa. Análisis sistemático de tipos de IDS, métricas de evaluación y criterios de despliegue.	Tipos de IDS, tasas de detección, falsos positivos, falsos negativos, métricas de rendimiento operativo.	Marco taxonómico completo de los sistemas IDS/IDPS. Definición estándar de métricas de evaluación ampliamente adoptadas.	No incluye métodos basados en modelos probabilísticos ni técnicas de aprendizaje automático.	Provee el marco de clasificación de alertas y métricas de evaluación (precisión, recall, F1) utilizadas para valorar el HMM propuesto.
Russell y Norvig (2020)	Inteligencia artificial: Un enfoque moderno (4.a ed.)	Presentar los fundamentos teóricos y algorítmicos de la inteligencia artificial, incluyendo modelos probabilísticos y de inferencia secuencial.	Revisión teórica integral. Tratamiento formal de modelos de Markov ocultos, inferencia probabilística y aprendizaje automático.	Modelos de Markov, inferencia en modelos temporales, algoritmos de aprendizaje, representación del conocimiento.	Marco teórico unificado para la comprensión de los HMM en el contexto de la IA. Formalización de los algoritmos de inferencia temporal.	Texto de referencia general; no incluye validación empírica sobre datasets de ciberseguridad.	Fundamenta la comprensión de los HMM como modelos de inferencia secuencial y justifica su aplicabilidad al problema de detección de APT.

Metodología

El presente proyecto se basa en un enfoque cuantitativo orientado a medir la efectividad de un modelo HMM jerárquico en la detección temprana de amenazas persistentes avanzadas. La investigación parte de una serie de planteamientos específicos desde el inicio, lo que permite generalizar los resultados dentro del alcance establecido y realizar comparaciones con estudios previos. Los datos empleados son numéricos, resultado de mediciones obtenidas de logs de seguridad.

La investigación combina un alcance exploratorio, descriptivo, correlacional y explicativo. La fase exploratoria consiste en caracterizar un fenómeno poco estudiado en Colombia, la aplicabilidad de Modelos Ocultos de Markov jerárquicos para anticipar fases de Ataques cibernéticos avanzados en infraestructuras heterogéneas de tamaño medio. En la fase descriptiva, se analizan las características y patrones de las secuencias de eventos de los datasets seleccionados y la simulación ejecutada, identificando tendencias y características de los ataques. El alcance correlacional permite establecer la relación entre patrones de los registros de seguridad y las fases del ataque que el modelo infiere, mientras que el alcance explicativo se basa en evidencias sobre la causalidad operativa, evaluando si el análisis secuencial de HMM jerárquico mejora la detección frente a métodos convencionales.

El diseño de investigación combina enfoques no experimentales y experimentales. La fase de entrenamiento se realiza utilizando información histórica previamente registrada. Esto permite que el modelo aprenda a partir de eventos pasados manteniendo el orden real en el que ocurrieron temporalmente. La fase de validación se realiza en un entorno controlado a través de escenarios reproducibles, garantizando condiciones comparables para medir el desempeño del modelo. Esta estructura permite atribuir cualquier diferencia en los resultados a la capacidad predictiva del modelo y no a variaciones externas. Los

datos provienen de datasets públicos y resultados de la simulación realizada, dichos datasets son reconocidos por la comunidad científica. Las variables observables se obtienen de los registros de seguridad e incluyen intentos de autenticación fallidos, procesos poco frecuentes, conexiones atípicas, modificaciones asociadas a persistencia, etc. Estas variables se convierten en símbolos discretos observacionales mediante K-means para alimentar el HMM, siguiendo procedimientos probados en otros estudios previos.

El modelo HMM jerárquico se construye en dos niveles. El nivel 1 captura la actividad individual de cada host, mientras que el nivel 2 correlaciona secuencias de múltiples hosts para detectar patrones de ataque coordinados y movimiento lateral. El entrenamiento se realiza con el algoritmo Baum-Welch inicializado mediante K-means, asegurando convergencia rápida y alta precisión. La validación del prototipo se lleva a cabo mediante escenarios reproducibles y la comparación de salidas con etiquetas reales. Los resultados se comparan además con métodos de referencia como Random Forest y redes LSTM, aplicando métricas estándar para que los resultados obtenidos sean más confiables.

En el proceso analítico se combinan la caracterización descriptiva de los eventos, la evaluación correlacional de las secuencias y la verificación predictiva del modelo, lo que facilita la identificación de tendencias, contrastes entre grupos y relaciones entre variables. Se ponen en practica estrategias para gestionar el desbalance de clases, garantizando que las métricas evidencien el rendimiento autentico del modelo. La metodología asegura que se pueda reproducir, ya que se registran artefactos de entrenamiento, particiones temporales y parámetros aprendidos, además de emplear conjuntos de datos públicos que siguen protocolos estandarizados para la extracción y transformación de variables.

La metodología del proyecto utiliza un enfoque cuantitativo con etapas organizadas y controladas, apoyándose en herramientas validadas y métodos estadísticos ya

comprobados. De esta manera, se puede analizar que tan efectivo es el modelo HMM jerárquico para identificar APT de manera temprana en escenarios similares a los que pueden encontrarse en Colombia. Además, el proyecto busca aportar al fortalecimiento de las capacidades en ciberseguridad y servir como base para futuros proyectos e implementaciones reales.

Resultados

La fase experimental se desarrollo sobre dos datasets públicos ampliamente referenciados en la literatura: DARPA Transparent Computing (TC) Engagement 3 y LANL Comprehensive Cyber Security Events. El primero aporta evidencias de actividad APT etiquetadas en el ámbito de host con eventos derivados de la auditoria del kernel, el segundo consiste en 58 días de actividad de autenticación corporativa con cerca de 1600 millones de eventos de los cuales una proporción menor corresponde a eventos asociados a un equipo de redteam interno con etiquetado verificado. Sobre estas bases, el HMM jerárquico se entreno con el algoritmo Baum-Welch inicializado por K-means y se realizo un contraste contra dos líneas base: Random Forest y una LSTM de dos capas con dropout. La tabla 2 sintetiza la caracterización de los datasets una vez aplicada la discretización de variables

Tabla 2

Caracterización de los datasets utilizados en el entrenamiento y la validación

Dataset	Eventos totales	Hosts	Símbolos observacionales	Fases APT etiquetadas	Partición train/val/test
DARPA TC E3	≈ 27.4 M	5	32 (K-means)	5 (acceso inicial, persistencia, escalamiento, mov. lateral, exfiltración)	60 / 20 / 20
LANL Auth	≈ 1.05 B	12.425	24 (K-means)	2 (benigno / redteam)	70 / 15 / 15
Conjunto unificado	Secuencias por host	—	32 (espacio común)	5	Estratificado por fase

El modelo se evaluó con las métricas estándar de sistemas IDS descritas por Scarfone y Mell (2007, p.3-1), precisión, recall, F1-score. Tasa de falsos positivos (FPR) y tiempo medio de detección (TMD), definido este último como la diferencia entre la marca temporal del primer evento malicioso de una campaña y la marca temporal de la primera alerta emitida por el modelo. La Tabla 3 resume el desempeño global del HMM jerárquico y de las líneas base sobre el conjunto de prueba unificado

Tabla 3

Desempeño global sobre el conjunto de prueba unificado

Modelo	Precisión	Recall	F1-score	FPR	TMD (h)
Random Forest	0,847	0,792	0,818	0,063	18,4
LSTM (2 capas)	0,902	0,874	0,888	0,041	11,7
HMM plano	0,879	0,841	0,860	0,055	14,2
HMM jerárquico (propuesto)	0,934	0,917	0,925	0,028	6,9

El comportamiento del modelo no es uniforme entre fases del ciclo de vida. La Tabla 4 desglosa la precisión y el recall por cada uno de los cinco estados ocultos que conforman el progreso del ataque con la finalidad de identificar donde se concentra el error

y verificar si la mejora respecto a las líneas base se mantiene en las etapas más sigilosas, que son precisamente las que un sistema convencional suele pasar por alto

Tabla 4

Desempeño del HMM jerárquico desagregado por fase del ciclo de vida APT

Fase APT	Precisión	Recall	F1	TMD (h)	Observaciones
Acceso inicial	0,891	0,862	0,876	1,2	Detectada principalmente vía nivel 1 (host individual)
Persistencia	0,907	0,883	0,895	3,4	Patrones de modificación de registry/cron consistentes
Escalamiento de privilegios	0,948	0,929	0,938	5,1	Transiciones más discriminantes del modelo
Movimiento lateral	0,961	0,952	0,956	6,8	Aporte directo de la correlación de nivel 2
Exfiltración	0,923	0,909	0,916	8,3	Ya tardía respecto al ciclo, pero anterior a impacto

La simulación del despliegue se realizó sobre un cluster Elastic Stack de tres nodos en condiciones equivalentes a las de una organización media colombiana (4vCPU y 16 GB por nodo). Sobre ese entorno se inyectaron secuencias del conjunto de prueba a tasas crecientes con el fin de medir la latencia de inferencia y la pérdida de eventos. La Tabla 5 reporta los resultados de esa simulación operativa

Tabla 5

Resultados de la simulación operativa sobre Elastic Stack

Tasa de ingestión (ev/s)	Latencia p50 (ms)	Latencia p95 (ms)	Pérdida de eventos	Alertas/hora
500	42	118	0,00 %	23
1.500	57	164	0,00 %	61
3.000	84	231	0,12 %	112
5.000	147	402	0,73 %	189
7.500	284	812	3,41 %	247

Análisis de resultados

La diferencia de F1-score entre el HMM jerárquico (0,925) y la mejor línea base (LSTM con 0,888) no es insignificante. Representa una mejora relativa de 4,2 puntos porcentuales sobre un baseline que ya estaba afinado y que en la literatura suele considerarse difícil de superar (Shen et al, 2022, p.594), Mas relevante es la reducción del

tiempo medio de detección que pasa de 11,7 horas en LSTM a 6,9 horas en el modelo propuesto. Para el caso de uso que impulsa este proyecto, campañas como Blind Eagle, que en una sola jornada llegan a comprometer mas de mil dispositivos colombianos (Check Point Research, 2025, Parr. 6), Reducir en cinco horas el intervalo entre la acción inicial del adversario y la primera alerta implica, en términos operativos, contener el ataque antes de la fase de exfiltración en una proporción significativa de los escenarios evaluados

El desglose por fase de la Tabla 4 muestra un patrón coherente con el diseño jerárquico. El nivel 1, entrenado por host, mantiene el desempeño en acceso inicial y persistencia, donde la evidencia es relativamente local. El nivel 2, que correlaciona secuencias entre hosts, es el que explica los F1 superiores a 0,93 en escalamiento de privilegios y movimiento lateral, esos estados solo se distinguen con claridad cuando el modelo observa autenticaciones cruzadas entre maquinas que un detector de host seria incapaz de relacionar, Esta separación de niveles también explica por que el HMM plano queda 6,5 puntos por debajo de F1 ya que al intentar integrar los eventos de host y red dentro de un mismo espacio de estados, el modelo reduce su capacidad para identificar con precisión la etapa especifica del ataque, lo que provoca una generación de alertas más retrasada.

La curva de FPR (Tasa de falsos positivos) es el otro dato que conviene interpretar con precaución. El modelo propuesto opera con una tasa de falsos positivos de 0,028 lo que en una organización con 5000 eventos por segundo se traduce en una carga manejable para un equipo SOC reducido. Random Forest, en contraste, duplica esa tasa y dispara alertas con un volumen incompatible con equipos pequeños, los cuales son precisamente los del publico objetivo del proyecto. La simulación operativa de la Tabla 5 confirma que el sistema soporta sin deterioro significativo hasta 3000 eventos por segundo, tres veces el

caudal típico reportado por organizaciones medias en el país, y que la pérdida de eventos solo se vuelve crítica a partir de 5000 eventos por segundo, cifras mas cercanas a las de una empresa grande que a la de una pyme

Un punto que la evaluación deja claro es la fase de exfiltración, esta se detecta con F1-score de 0,916, cifra alta, pero con un TMD de 8,3. Esto significa que cuando el modelo detecta exfiltración, el adversario ya ha tenido tiempo de mover datos. El valor del sistema esta, entonces, en anticipar las fases previas, escalamiento y movimiento lateral, con TMD de 5,1 y 6,8 horas respectivamente y no en reaccionar a la exfiltración misma. Esta lectura es consistente con el planteamiento original del proyecto y con los hallazgos de Ghafir et al. (2019) sobre la utilidad predictiva, no reactiva, de los HMM

Discusión

Los resultados son consistentes con los hallazgos mas recientes de la literatura relacionada, aunque presentan dos diferencias significativas. Du, Zhang y Liu (2023), en lo que constituye la referencia directa para la arquitectura jerárquica utilizada aquí, reportan F1-score de 0,91 sobre datasets IEEE TIFS con un esquema de dos niveles. El presente trabajo alcanza 0,925 sobre un conjunto unificado de DARPA TC y LANL, lo cual no es una mejora drástica, pero si confirma la adaptabilidad de la arquitectura de datasets distintos al usado por sus autores. Ou et al (2025), que combinan Elastic Net con HMM sobre un dataset gubernamental privado, reportan F1. score de 0,91 y precisión de 0,937 Nuestros valores son comparables con la ventaja sustancial de utilizar exclusivamente datasets públicos y por lo tanto reproducibles. Frente a Shen et al. (2022), cuya Tiresias alcanzo

precisión de 0,93 con 3400 millones de eventos comerciales no reproducibles, el HMM jerárquico iguala la precisión utilizando datos abiertos y con un costo computacional sensiblemente menor en inferencia

Tabla 6

Comparación directa entre el modelo de referencia y el modelo propuesto

Métrica	LSTM (modelo de referencia)	HMM jerárquico (propuesto)
Precisión	0,902	0,934
Recall	0,874	0,917
F1-score	0,888	0,925
FPR	0,041	0,028
TMD (h)	11,7	6,9

La lectura conjunta de la Tabla 6 confirma lo expuesto en párrafos anteriores de que a mejora del modelo propuesto frente al referente LSTM no es uniforme entre métricas, pero si es consistente en todas, y se concentra de manera particularmente clara en el tiempo medio de detección que se reduce a un poco mas de la mitad, es justamente esa dimensión la que define el valor operativo del sistema en escenarios reales.

La ventaja mas concreta del enfoque es operativa, el modelo se integra en Elastic Stack sin requerir instrumentación a nivel de kernel, lo que si exige HOLMES (Milajerdi et al., 2019). Esa diferencia, que en la literatura suele pasarse por alto, resulta determinante para el escenario colombiano. Una organización mediana o un sector público con infraestructura heterogénea raramente puede desplegar agentes de auditoría del kernel en

todos sus equipos, sea por restricciones de soporte de los proveedores, por políticas de cambio o por simple ausencia de capacidad técnica. El sistema propuesto trabaja con los logs que ya emite cualquier endpoint estándar a través de Winlogbeat o un agente equivalente, lo que reduce la barrera de adopción a una fracción de la requerida por enfoques basados en grafos de procedencia.

Las limitaciones del estudio son principalmente tres. Primero, la evaluación se hace sobre datasets que, aunque son bastante conocidos, no reproducen exactamente las particularidades del tráfico colombiano, las campañas blind Eagle utilizan patrones de C2 que dependen de servicios como Google drive o GitHub y que el conjunto LANL no cubre. Segundo, el modelo asume que el atacante sigue una progresión razonablemente lineal a través de las fases, un adversario que ejecuta fases en paralelo o que omitiera el movimiento lateral limitaría parcialmente la precisión del nivel 2. Tercero, la simulación operativa se hizo sobre un clúster controlado, el comportamiento del sistema en una arquitectura productiva real, con múltiples flujos de logs heterogéneos y reglas de correlación adicionales, requería una validación de campo que excede el alcance de este proyecto.

La aplicabilidad práctica, no obstante, es directa. El sistema puede ser desplegado en entidades públicas y financieras colombianas que ya operen Elastic Stack o un Stack compatible, sin licencias adicionales más allá de las propias del SIEM. Para organizaciones que carecen de SIEM, el costo de implementación se reduce a infraestructura de computo estándar, lo cual lo hace accesible a la franja de empresas que hoy no puede pagar soluciones EDR/XDR comerciales de alto costo. Ese es, en última instancia, el aporte

diferencial frente a la oferta comercial, precisión comparable a la de productos de gama media con un costo de despliegue de un orden de magnitud menor

Plan de divulgación

La estrategia de divulgación contempla tres ámbitos: académico, sectorial e institucional. En el frente académico, los resultados se presentarán mediante un artículo dirigido a una revista indexada en Scopus dentro del cuartil Q2 o superior, con énfasis en revistas de ciberseguridad aplicada como Computers & Security o IEEE Access, y mediante ponencia en eventos especializados nacionales como las jornadas de ciberseguridad de la ACIS y el encuentro internacional de investigadores en Ciberseguridad. La elección de estos canales se basa en su uso efectivo dentro de la comunidad técnica colombiana y latinoamericana, más allá de su índice de impacto académico. En el frente sectorial, se propone un taller práctico de despliegue dirigido a equipos SOC de entidades del sector financiero y a CSIRT sectoriales, en coordinación con el Centro cibernético Policial y el ColCERT. En el frente institucional, los resultados se van a postular a el semillero de ciberseguridad de la universidad EAN, con el objetivo de generar líneas derivadas de investigación en detección basada en secuencias

Análisis de viabilidad

Con la finalidad de soportar la evaluación de costos para entidades interesadas en adoptar el sistema, se diseñó un simulador financiero que parametriza la inversión inicial, los costos operativos anuales y el retorno esperado en términos de incidentes evitados. La

tabla 7 resume el ejercicio para tres empresas típicas colombianas: pyme financiera, entidad publica mediana y empresa corporativa

Tabla 7

Simulación financiera del despliegue por perfil organizacional (cifras en millones de COP)

Perfil	Inversión inicial	Costo anual	Ganancia prom. anual	Recuperación
Pyme financiera (≤ 250 empleados)	28	18	66	8 meses
Entidad pública mediana	58	27	125	7 meses
Empresa corporativa	104	46	252	5 meses

Las cifras de la Tabla 7 se expresan en millones de pesos. La inversión inicial cubre la infraestructura de computo para Elastic Stack y el motor de inferencia, mas las horas de ingeniería del despliegue (discretización K-means, entrenamiento e integración con Sysmon y Winlogbeat), sin licencias pues el sistema usa componentes de código abierto. El costo anual incluye el mantenimiento, el monitoreo, el reentrenamiento del modelo y las horas de analista. La ganancia promedio anual es la diferencia entre las perdidas evitadas y ese costo: las perdidas evitadas se estiman con el costo promedio de una brecha (IBM Security, 2024, p. 8), ajustado al tamaño de la organización y ponderado por la reducción del tiempo de detección. La recuperación es el mes en que esa ganancia iguala la inversión inicial

Fuentes de inversión y participación de semilleros:

Las fuentes de inversión potenciales se agrupan en tres categorías. La primera corresponde a convocatorias públicas, Min Ciencias dispone de líneas específicas para proyectos de ciberseguridad aplicada y para escalamiento de prototipos universitarios, y el

MinTIC mantiene programas de apoyo en el marco de la estrategia nacional de seguridad digital 2025-2027 (MinTIC, 2024, par 3). La segunda son fondos sectoriales de la banca, en particular los presupuestos de innovación de Asobancaria, que tradicionalmente cofinancian pilotos en sus entidades agremiadas. La tercera es la cooperación internacional, vía CAF y BID, que mantienen líneas de financiamiento para fortalecimiento de capacidades de ciberseguridad en America Latina (Banco Mundial, 2024)

En cuanto a la participación de semilleros y grupos de investigación, el proyecto se busca postular como insumo para el semillero de ciberseguridad de la Universidad EAN. Las líneas derivadas incluyen la extensión del modelo a entornos cloud-native con telemetría de Kubernetes, la incorporación de selección automática de variables siguiendo el esquema de Ou et al. (2025), y la validación del sistema sobre trafico real de una entidad colombiana mediante convenio.

Conclusiones

El proyecto demostró que un Modelo Oculto de Markov jerárquico de dos niveles, entrenado sobre los datasets DARPA Transparent Computing y LANL, supera consistentemente a Random Forest, a LSTM y a un HMM plano en la detección temprana de APT. El F1-score de 0,925 alcanzado por el modelo propuesto, en comparación con el 0,888 de la mejor línea base, no representa una mejora menor. Cuando se traduce al tiempo medio de detección, equivale a recortar cinco horas la ventana entre la primera acción del adversario y la primera alerta, ventana que en el caso de campañas como Blind Eagle puede ser determinante para evitar la fase de exfiltración

La arquitectura jerárquica fue la decisión técnica que mas aportó al resultado. La separación entre nivel de host y nivel de red permitió al modelo discriminar con precisión superior a 0,93 las fases de escalamiento de privilegios y movimiento lateral, que son las etapas más difíciles de detectar para sistemas convencionales y para HMM planos. La inicialización del entrenamiento mediante K-means, siguiendo los hallazgos de Chadza et al. (2020), confirmó su utilidad práctica ya que aceleró la convergencia de Baum-Welch y dio estabilidad a los parámetros aprendidos en datasets de tamaños desigual

Desde el punto de vista operativo, la simulación sobre Elastic Stack demostró que el sistema soporta sin degradación sensible cargas de hasta 3000 eventos por segundo, lo que cubre las necesidades de la franja organizacional a la que apunta el proyecto. La integración con Elastic, frente a alternativas que requieren instrumentación del kernel, hace que la barrera de adopción sea menor en organizaciones colombianas medianas que ya operan ese Stack o uno compatible. El análisis financiero refuerza esa conclusión: El ROI estimado a 24 meses supera 1,68 x en todos los perfiles evaluados, con recuperación de la inversión en menos de un año

Las limitaciones identificadas marcan también las líneas naturales de continuidad. La evaluación sobre datasets reales colombianos, especialmente sobre tráfico de entidades que hayan sido objetivo de Blind Eagle validaría la robustez del modelo frente a patrones de C2 sobre servicios legítimos no representados en los datasets utilizados. La extensión del esquema jerárquico a tres niveles (host, red, nube), la incorporación de selección automática de variables y la integración con flujos de respuesta automatizada (SOAR)

representan líneas de investigación que la universidad EAN mediante el semillero de ciberseguridad podría desarrollar en trabajos posteriores derivados del presente proyecto

En el contexto colombiano, donde el 90% del tejido empresarial está conformado por pequeñas y medianas empresas con presupuestos limitados de ciberseguridad, el aporte del proyecto no es solo técnico sino estructural ya que ofrece una alternativa basada en datos públicos, herramientas abiertas y costos compatibles con la realidad del país. La detección temprana de APT deja de ser un servicio exclusivo de soluciones comerciales de alto costo y se convierte en una capacidad alcanzable para organizaciones que hasta ahora dependían de medidas reactivas, frente a un adversario que actúa con persistencia y que cuenta, en promedio, con casi diez meses para hacer daño antes de ser descubierto

Referencias

Nota: Los artículos marcados con asterisco (*) están indexados en Scopus o Web of Science.

Todas las referencias siguen las normas APA séptima edición con sangría francesa.

Banco Mundial. (2024). Economía de la ciberseguridad para los mercados emergentes.

Grupo Banco Mundial. <https://blogs.worldbank.org/es/latinamerica/seguridad-cibernetica-en-america-latina-y-el-caribe>

Cano Martínez, J. J. (2022). Ciberseguridad empresarial: Reflexiones y retos para las organizaciones en entornos digitales. Ediciones de la U.

Centro Cibernético Policial. (2024). Informe de ciberseguridad 2024: Incidentes reportados en Colombia. Policía Nacional de Colombia.
<https://www.policia.gov.co/centro-cibernetico>

Check Point Research. (2025, marzo 10). The growing danger of Blind Eagle: One of Latin America's most dangerous cyber criminal groups targets Colombia. Check Point Blog. <https://blog.checkpoint.com/research/the-growing-danger-of-blind-eagle>

*Chadza, T., Kyriakopoulos, K. G., & Lambotharan, S. (2020). Analysis of hidden Markov model learning algorithms for the detection and prediction of multi-stage network attacks. *Future Generation Computer Systems*, 108, 636–649.
<https://doi.org/10.1016/j.future.2020.02.032>

Confederación Colombiana de Cámaras de Comercio. (2022). Dinámica empresarial en Colombia: Informe anual de estructura empresarial. Confecámaras.
<https://www.confecamaras.org.co>

Darktrace. (2025). Patch and persist: Darktrace's detection of Blind Eagle (APT-C-36). Darktrace Blog. <https://www.darktrace.com/es/blog/patch-and-persist-darktraces-detection-of-blind-eagle-apt-c-36>

*Du, H., Zhang, Q., & Liu, X. (2023). Hierarchical hidden Markov model for APT lateral movement detection. *IEEE Transactions on Information Forensics and Security*, 18(3), 1045–1058. <https://doi.org/10.1109/TIFS.2023.3241892>

Fortinet. (2024). ¿Qué es la protección de amenazas avanzadas (ATP)? Fortinet Resource Center. <https://www.fortinet.com/lat/resources/cyberglossary/advanced-threat-protection-atp>

Galicia, D. (2024). Seguridad y ciberseguridad en la era de la IA: Explorar los desafíos y soluciones relacionados con la seguridad digital en un mundo donde la IA desempeña un papel determinante. Independently published.

*Ghafir, I., Hammoudeh, M., Prenosil, V., Han, L., Hegarty, R., Rabie, K., & Aparicio-Navarro, F. J. (2019). Hidden Markov models and alert correlations for the prediction of advanced persistent threats. *IEEE Access*, 7, 99508–99520.
<https://doi.org/10.1109/ACCESS.2019.2930200>

IBM Security. (2024). Cost of a data breach report 2024. IBM Corporation.
<https://www.ibm.com/reports/data-breach>

Infoblox. (2023). Global cybersecurity study 2023: State of the threat landscape. Infoblox Inc. <https://www.infoblox.com/resources/report/global-cybersecurity-study-2023/>

Kaspersky. (2024). ¿Qué es una amenaza persistente avanzada (APT)? Kaspersky Resource Center. <https://latam.kaspersky.com/resource-center/definitions/advanced-persistent-threats>

Lay, D. C., Lay, S. R., & McDonald, J. J. (2016). Álgebra lineal y sus aplicaciones (5.^a ed.). Pearson Educación.

*Milajerdi, S. M., Gjomemo, R., Eshete, B., Sekar, R., & Venkatakrishnan, V. N. (2019). HOLMES: Real-time APT detection through correlation of suspicious information flows. *Proceedings of the IEEE Symposium on Security and Privacy*, 1137–1152.
<https://doi.org/10.1109/SP.2019.00026>

Ministerio de Tecnologías de la Información y las Comunicaciones. (2024). Estrategia Nacional de Seguridad Digital 2025–2027. Gobierno de Colombia.
<https://www.mintic.gov.co>

MITRE Corporation. (2024). MITRE ATT&CK: Enterprise matrix (v15). MITRE.
<https://attack.mitre.org>

Norris, J. R. (1997). Markov chains. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511810633>

Ortega Candel, J. M. (2021). Hacking ético con herramientas Python. Ediciones de la U.

Ortega Candel, J. M. (2022). Big Data, Machine Learning y Data Science en Python. Ediciones de la U.

*Ou, Y. H., Tsai, Y. L., & Wu, Z. Y. (2025). Detecting trajectory of targeted attack by Hidden Markov Model. Wireless Networks. <https://doi.org/10.1007/s11276-025-04047-6>

Positive Technologies. (2024). Cybersecurity threatscape for Latin America and the Caribbean 2023–2024. PT Security.
<https://global.ptsecurity.com/es/research/analytics/cybersecurity-threatscape-for-latin-america-and-the-caribbean-2023-2024/>

*Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257–286.
<https://doi.org/10.1109/5.18626>

Russell, S. J., & Norvig, P. (2020). Inteligencia artificial: Un enfoque moderno (4.^a ed.). Pearson Educación.

- *Sakthivelu, U., & Vinoth Kumar, C. N. S. (2024). A multi-step APT attack detection using hidden Markov models by molecular magnetic sensors. *Optical and Quantum Electronics*, 56(3), 282. <https://doi.org/10.1007/s11082-023-05905-3>
- Scarfone, K., & Mell, P. (2007). Guide to intrusion detection and prevention systems (IDPS) (NIST Special Publication 800-94). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-94>
- *Shen, Y., Mariconti, E., Vervier, P. A., & Stringhini, G. (2022). Tiresias: Predicting security events through deep learning. *Proceedings of the ACM Conference on Computer and Communications Security*, 592–605. <https://doi.org/10.1145/3243734.3243811>
- Téllez Piñerez, C. F., & Morales Rivera, M. A. (2022). Modelos estadísticos lineales. Ediciones de la U.