2025

Centro de Respuestas a Incidentes Informáticos
CSIRT Académico UNAD

Detección de intrusos con aprendizaje contrastivo y mezcla gaussiana bayesiana

UN ENFOQUE MÁS PRECISO Y ADAPTABLE PARA LA DETECCIÓN DE INTRUSIONES

VIEM
Vicerrectoría de Innovación y Emprendimiento

ECBTI
Escuela de Ciencias Básicas, Tecnología e Ingeniería

Semillero de Investigación
Ceros y Unos

E-boletín Informativo CSIRT Académico UNAD

Medio de divulgación: Correo Electrónico, Sitio Web

Incluye: Noticias, alertas o informes relacionados con la disciplina de la ciberseguridad

Número treinta y tres [33]
Junio de 2025

Universidad Nacional Abierta y a Distancia (UNAD)
Vicerrectoría de Innovación y Emprendimiento (VIEM)
Escuela de Ciencias Básicas Tecnología Ingeniería (ECBTI)
Maestría en Ciberseguridad
Especialización en Seguridad Informática
CSIRT Académico UNAD

Universidad Nacional Abierta y a Distancia
Calle 14 sur No. 14-23 | Bogotá D.C
Correo electrónico: csirt@unad.edu.co
Página web: https://csirt.unad.edu.co

Edición electrónica, financiada por la Universidad Nacional Abierta y a Distancia (UNAD)

Vicerrectoría de Innovación y Emprendimiento (VIEM)
Ing. Andrés Ernesto Salinas
Vicerrector

Escuela de Ciencias Básicas Tecnología e Ingeniería (ECBTI)
Ing. Claudio Camilo González Clavijo
Decano

Maestría en Ciberseguridad (ECBTI)
Ing. Sonia Ximena Moreno Molano
Líder Programa de Maestría en Ciberseguridad

Semillero de Investigación Ceros y Unos, adscrito al Grupo de Byte InDesign

Centro de Desarrollo Tecnológico CSIRT Académico UNAD
Ing. Luis Fernando Zambrano Hernández
Líder CSIRT Académico UNAD

Responsable de la Edición
Ing. Luis Fernando Zambrano Hernandez

Revisó
Adm. Libardo Cárdenas Corral
Analista CSIRT Académico UNAD

Estado legal:
Periodicidad: Mensual
ISSN: 2806-0164

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Contenido

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Glosario de Terminos

Aprendizaje contrastivo

Técnica de aprendizaje automático, comúnmente no supervisada, que entrena modelos a distinguir entre ejemplos similares y diferentes. Se usa para generar representaciones (embeddings) discriminativas útiles para tareas como clasificación o detección de anomalías.

Modelo bayesiano de mezcla gaussiana (BGMM)

Variante del modelo de mezcla gaussiana (GMM) que emplea inferencia bayesiana para estimar automáticamente el número de componentes y sus parámetros, facilitando el modelado de distribuciones complejas en los datos.

Distribución gaussiana

También conocida como distribución normal. Es una función de probabilidad simétrica que describe cómo se distribuyen los valores alrededor de la media. Es fundamental en estadística y en modelos probabilísticos como GMM y BGMM.

Embedding

Representación vectorial de datos (como tráfico de red o texto) en un espacio latente. Los embeddings permiten al modelo identificar patrones y relaciones subyacentes en los datos de forma más eficiente.

F1-Score

Métrica utilizada en clasificación que combina precisión y exhaustividad (recall). Es especialmente útil cuando hay un desequilibrio en las clases, como ocurre en muchos conjuntos de datos de ciberseguridad.

NetFlow

Protocolo de recolección de metadatos de tráfico IP. Utilizado para monitoreo y análisis de comportamiento de red sin necesidad de inspeccionar paquetes completos. Compatible con muchos sistemas NIDS.

Modelo de mezcla gaussiana (GMM)

Modelo estadístico que representa una distribución como la combinación de múltiples distribuciones gaussianas. Se utiliza para clustering, clasificación y detección de anomalías.

Sistema de detección de intrusiones en red (NIDS)

Solución de seguridad que monitorea el tráfico de red en busca de actividades maliciosas o violaciones de políticas. Puede basarse en firmas, comportamiento o inteligencia artificial.

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Aprendizaje no supervisado

Tipo de aprendizaje automático que extrae patrones de los datos sin etiquetas predefinidas. Se utiliza ampliamente para clustering, reducción de dimensionalidad y detección de anomalías.

Remuestreo (resampling)

Técnica utilizada para balancear clases desiguales en un conjunto de datos, como sobremuestreo de la clase minoritaria o submuestreo de la mayoritaria. El modelo propuesto por Liu y Xu evita esta etapa, gracias al aprendizaje contrastivo.

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Introducción

En los últimos años, la detección de intrusiones en redes ha evolucionado hacia enfoques avanzados que integran aprendizaje profundo y modelos probabilísticos adaptativos. En este contexto, el trabajo *A Newark intrusion detection method based on contrastive learning and Bayesian Gaussian Mixture Model* (Liu & Xu, 2025) propone un sistema en dos etapas: aprendizaje contrastivo y modelo bayesiano de mezcla gaussiana, logrando alta precisión sin necesidad de remuestreo ni ingeniería manual de características.

Este enfoque complementa estudios recientes que exploran arquitecturas modernas para NIDS. Por ejemplo, Xi, Wang y Wang (2024) presentan un modelo multi-escala basado en Transformer (IDS-MTran) que alcanza >99 % de precisión en datasets como UNSW-NB15, mejorando la captación de patrones intrusivos en distintos niveles de tráfico. Asimismo, Singh y Jang-Jaccard (2022) desarrollaron una red autoencoder multiescala (MSCNN-LSTM-AE) que aprovecha correlaciones espaciales y temporales para detección no supervisada, logrando un F1-score alto en NSL-KDD, UNSW-NB15 y CIC-DDoS2019.

Estas propuestas reflejan una tendencia sólida en los proyectos de I+D+i en la comunidad académica, que buscan combinar métodos de aprendizaje profundo con modelado probabilístico o atencional, optimizando tanto rendimiento como adaptabilidad. La metodología de Liu y Xu (2025) se inscribe en esta corriente, ofreciendo una solución práctica y escalable para la detección de intrusiones, incluso en escenarios de tráfico desbalanceado.

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Detección de intrusos con aprendizaje contrastivo y mezcla gaussiana bayesiana

UN ENFOQUE MÁS PRECISO Y ADAPTABLE PARA LA DETECCIÓN DE INTRUSIONES

Autores

Hernando José Peña Hidalgo
Docente Investigador
CSIRT Académico UNAD
Universidad Nacional Abierta y a Distancia
ORCID: 0000-0002-3477-2645

Luis Fernando Zambrano Hernandez
Docente Investigador
Líder CSIRT Académico UNAD
Universidad Nacional Abierta y a Distancia
ORCID: 0000-0002-4690-3526

Sonia Ximena Moreno Molano
Líder Maestría en Ciberseguridad
Universidad Nacional Abierta y a Distancia
ORCID: 0000-0003-0392-1983

Néstor Raúl Cárdenas Corral
Analista CSIRT Académico UNAD
Universidad Nacional Abierta y a Distancia
ORCID: orcid.org/0000-0003-3691-0148

Fundamento Metodológico del Modelo Propuesto

Dos etapas clave: representación y detección inteligente

El sistema se estructura en dos etapas diseñadas para trabajar de forma sinérgica en la detección de intrusiones en redes:

1. Aprendizaje contrastivo para representación automática

El método comienza con un entrenamiento de tipo aprendizaje contrastivo, una técnica no supervisada que genera representaciones (embeddings) robustas a partir de pares de ejemplos similares y disímiles (Liu & Xu, 2025). Este enfoque permite al modelo discernir patrones de tráfico normal frente a tráfico malicioso sin necesidad de etiquetas manuales. Estudios recientes como Tan y Cheng (2025) demuestran que técnicas contrastivas pueden alcanzar hasta un 98% de precisión en datasets CIC-IDS2017 y KDDCUP99, validando su idoneidad en escenarios no equilibrados.

2. Clasificación adaptativa con modelo bayesiano de mezcla gaussiana

Sobre los embeddings generados, se aplica un Modelo de Mezcla Gaussiana Bayesiano (BGMM). A diferencia de un GMM tradicional, el BGMM emplea inferencia bayesiana para estimar tanto los parámetros como el número

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

óptimo de componentes de forma adaptativa (Liu & Xu, 2025). Esto significa que puede ajustarse automáticamente a nuevas distribuciones de tráfico sin intervención humana. Métodos similares, como Yu et al. (2023), han utilizado GMMs combinados con técnicas estadísticas para detectar anomalías de tráfico con alta precisión.

Ventajas del enfoque mixto:

Sin necesidad de remuestreo ni ingeniería manual, ambos bloques aprenden directamente de los datos originales, reduciendo la complejidad operativa.

Robustez frente a desequilibrios: el aprendizaje contrastivo maneja la escasez de ejemplos maliciosos generando representaciones balanceadas; el BGMM aprende automáticamente la densidad de datos.

Escalabilidad y adaptabilidad: al ajustarse a nuevas distribuciones, el sistema se adapta a cambios en el tráfico, similar a esfuerzos recientes como el método EG-ConMix de gráfica contrastiva.

Tabla 1.
Comparativa metodológica

Comparativa metodológica		
Componente	**Ventajas del modelo Liu & Xu (2025)**	**Referencias académicas similares**
Aprendizaje contrastivo	Representaciones discriminativas sin etiquetas ni remuestreo	Tan & Cheng (2025): < 98 % precisión en NIDS via contrastive learning (DOI: 10.1145/3731867.3731894)
BGMM adaptativo	Aprendizaje bayesiano que ajusta número de componentes dinámicamente	Yu et al. (2023): alto rendimiento en detección de anomalías (DOI: 10.3390/electronics12061397)

Esta arquitectura combinada representa una solución práctica y escalable, alineada con la dirección actual de los proyectos de I+D+i en la comunidad académica, donde los métodos de aprendizaje profundo y los modelos probabilísticos se integran para ofrecer seguridad de red automatizada, evaluada en benchmarks globales.

Evaluación y Resultados

El rendimiento del modelo propuesto por Liu y Xu (2025) se evaluó exhaustivamente en dos reconocidos conjuntos de datos:

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

UNSW-NB15[1] y CIC-IDS2017[2]. Estos benchmarks permiten una comparación objetiva frente a enfoques previos.

Resultados en UNSW-NB15

- Precisión: 91,27 %
- F1-score: 92,30 %

Estas cifras representan mejoras de +1,85 % (precisión) y +2,35 % (F1) respecto al estado del arte (Liu & Xu, 2025).

Este avance es significativo en un campo donde cada décima de punto porcentual puede marcar la diferencia entre detección efectiva y riesgo no identificado.

Resultados en CIC-IDS2017

- Precisión: 99,66 %
- F1-score: 99,12 %

Aunque el aumento respecto a SOTA es más modesto (+0,05 % y +0,12 %), alcanza niveles casi óptimos, reflejando una excelente eficiencia en entornos con mayor complejidad de tráfico (Liu & Xu, 2025).

Comparativa con otros enfoques recientes

Para contextualizar estos resultados, se contrasta con otros modelos relevantes: Li et al. (2024) implementaron un sistema contrastivo de extremo a extremo (CNN+GRU), logrando una precisión del 99,9 % en ataques conocidos y un recall del 95 % para ataques desconocidos en CIC-IDS2017, lo que demuestra la competitividad del enfoque de Liu & Xu (Li et al., 2024)

Luo et al. (2023) propusieron un método multicanal con aprendizaje contrastivo y autoencoder, logrando 98,43 % de precisión en CIC-IDS2017. Aunque robusto, su rendimiento es ligeramente inferior al modelo híbrido con BGMM de Liu & Xu (Luo et al., 2023).

Estos ejemplos evidencian que el enfoque mixto de Liu & Xu combina lo mejor de lo probabilístico y lo contrastivo, posicionándolo muy cerca del límite del rendimiento en la detección de intrusiones.

Impacto del aprendizaje contrastivo y BGMM

Los componentes clave del modelo influyen de manera directa en los resultados:

Aprendizaje contrastivo:

Mejora la representación de patrones sutiles en tráfico malicioso y raro.

Estudios como Tan & Cheng (2025) evidencian que esta técnica puede superar el 98 % de precisión en detección no supervisada.

Modelo Bayesiano de Mezcla Gaussiana (BGMM):

Ajuste automático del número de componentes, permitiendo una clasificación más precisa en

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset

[2] https://www.unb.ca/cic/datasets/ids-2017.html

presencia de datos nuevos o cambiantes (Yu et al., 2023).
La combinación de ambos resulta en un sistema robusto que supera las limitaciones comunes de métodos basados únicamente en redes neuronales o técnicas estadísticamente puras.

Conclusión del apartado

Los resultados obtenidos por Liu & Xu (2025) no solo muestran mejoras estadísticas; también reflejan un enfoque viable para escenarios reales con necesidades dinámicas. El sistema presenta:

- Alta precisión y balance (F1) en tráficos complejos.
- Adaptabilidad sin necesidad de reajustes manuales.
- Material suficiente para evaluación en proyectos de I+D+i en la comunidad académica.

Aplicabilidad práctica en entornos reales

El enfoque híbrido propuesto por Liu & Xu (2025) no solo demuestra eficacia en datasets, sino que también presenta atributos clave para su implementación real en entornos productivos:

Implantación en plataformas de ciberseguridad activas

- **SOC, NDR y EDR**: El modelo puede integrarse en Sistemas de Respuesta y Detección de Redes (NDR) y en entornos de Respuesta a Endpoint (EDR), aprovechando los embeddings logrados con aprendizaje contrastivo para monitoreo inmediato (Liu & Xu, 2025).
- **Tráfico NetFlow estándar:** El uso de representaciones basadas en NetFlow lo hace compatible con infraestructuras existentes, como demuestra Sarhan, Layeghy y Portmann (2021), quienes estandarizaron datasets NetFlow para facilitar la implementación real.

Adaptabilidad a tráfico dinámico y nuevas amenazas

- **Aprendizaje online/autónomo:** Métodos recientes como AOC-IDS (Zhang et al., 2024) emplean aprendizaje contrastivo en línea para adaptarse a variaciones de tráfico y etiquetas débiles, una capacidad que se alinea con la flexibilidad del modelo de Liu & Xu (2025).
- **Adaptación sin intervención humana**: El uso de BGMM permite ajustar automáticamente el número de componentes, manteniendo la precisión en contextos cambiantes, similar a sistemas federados con GMM y autoencoders (Wang et al., 2025) (DOI: 10.1007/s10207-025-01000-8).

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Aplicaciones en IoT e infraestructuras críticas

- **Detección en IoT y sistemas industriales:** El modelo se adapta a entornos con recursos limitados, donde no es viable implementar remuestreo o etiquetado extensivo. En el caso de IoT, sistemas como EG-ConMix han usado aprendizaje contrastivo para tráfico en grafos de IoT con buena velocidad y precisión (Wu et al., 2024) (DOI: 10.48550/arXiv.2403.17980).
- **Escalabilidad para alto rendimiento:** Modelos comerciales que combinan autoencoders con GMM, como ThunderSecure, han sido aplicados con éxito en redes de alta capacidad (100 Gb/s), validando la escalabilidad de aproximaciones similares (DOE, 2022).

Tabla 2.
Ventajas prácticas del modelo híbrido propuesto para detección de intrusiones

Resumen de ventajas reales	
Característica	**Impacto práctico**
Integración con NetFlow	Compatibilidad con infraestructuras de monitoreo existentes
Adaptación dinámica	Ajuste automático a nuevos patrones sin intervención humana
Aplicación en IoT y 5G	Viable en dispositivos con recursos limitados
Validado en alta velocidad	Escalable a redes de gran volumen y rendimiento

En conjunto, este enfoque se convierte en una solución robusta para SOCs, NDRs, IoT y entornos industriales, y representa un avance real para los proyectos de I+D+i en la comunidad académica y su transferencia al sector productivo.

Conclusiones

El modelo híbrido propuesto por Liu y Xu (2025), basado en aprendizaje contrastivo y un modelo bayesiano de mezcla gaussiana, representa una contribución significativa al fortalecimiento de los sistemas de detección de intrusiones en entornos complejos y con limitaciones operativas.
Sus resultados, validados en datasets ampliamente utilizados como UNSW-NB15 y CIC-IDS2017, evidencian un alto rendimiento tanto en precisión como en F1-score, sin recurrir a técnicas de remuestreo ni etiquetado manual. Esto lo convierte en una solución particularmente atractiva para la implementación en centros de monitoreo, redes distribuidas y dispositivos de borde (IoT).

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

En el marco de los proyectos de I+D+i en la comunidad académica, este enfoque puede servir como base para:

- Desarrollar pruebas piloto de detección automática en ambientes controlados, tales como laboratorios de simulación de tráfico o escenarios de entrenamiento en seguridad ofensiva/defensiva.
- Fortalecer la formación técnica avanzada, al incluir esta arquitectura en módulos de aprendizaje automático aplicado a la ciberseguridad.
- Promover la colaboración entre grupos de investigación y CSIRTs universitarios, con el fin de adaptar el modelo a contextos locales de tráfico y amenazas.

Desde el CSIRT Académico UNAD, recomendamos:

1. Explorar la implementación experimental del modelo en entornos de pruebas virtualizados (por ejemplo, con tráfico NetFlow sintético o capturado).
2. Evaluar la escalabilidad del enfoque en infraestructura institucional, considerando dispositivos de monitoreo ya existentes.
3. Difundir este tipo de soluciones dentro de la red de colaboración entre CSIRTs académicos, con miras a construir una base de conocimiento compartida para detección avanzada de amenazas.

La consolidación de métodos como el propuesto no solo fortalece las capacidades técnicas, sino que refuerza la autonomía tecnológica de las instituciones en materia de seguridad digital.

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html

Bibliografía

Liu, L., & Xu, M. (2025). A network intrusion detection method based on contrastive learning and Bayesian Gaussian Mixture Model. Cybersecurity. Advance online publication. https://doi.org/10.1186/s42400-025-00364-7

Tan, X., & Cheng, J. (2025). Contrastive learning for network intrusion detection: A comprehensive survey. CIoTSC 2024. https://doi.org/10.1145/3731867.3731894

Yu, B., Zhang, Y., Xie, W., Zuo, W., Zhao, Y., & Wei, Y. (2023). A network traffic anomaly detection method based on Gaussian mixture model. Electronics, 12(6), Article 1397. https://doi.org/10.3390/electronics12061397

Li, L., Lu, Y., Yang, G., & Yan, X. (2024). End-to-End network intrusion detection based on contrastive learning. Sensors, 24(7), 2122. https://doi.org/10.3390/s24072122

Luo, J., Zhang, Y., Wu, Y., Xu, Y., Guo, X., & Shang, B. (2023). A multi-channel contrastive learning network based intrusion detection method. Electronics, 12(4), 949. https://doi.org/10.3390/electronics12040949

Sarhan, M., Layeghy, S., & Portmann, M. (2021). NetFlow Datasets for Machine Learning-Based Network Intrusion Detection Systems. In ICCCN 2021: Proceedings. https://link.springer.com/chapter/10.1007/978-3-030-72802-1_9

Zhang, X., Zhao, R., Jiang, Z., Sun, Z., Ding, Y., Ngai, E. C. H., & Yang, S.-H. (2024). AOC-IDS: Autonomous Online Framework with Contrastive Learning for Intrusion Detection. arXiv. https://doi.org/10.48550/arXiv.2402.01807

Wang, X., et al. (2025). Federated learning for misbehaviour detection with variational Gaussian Mixture Models. International Journal of Information Security. https://doi.org/10.1007/s10207-025-01000-8

Wu, L., Lei, S., Liao, F., Zheng, Y., Liu, Y., Fu, W., Song, H., & Zhou, J. (2024). EG-ConMix: An Intrusion Detection Method based on Graph Contrastive Learning. arXiv. https://doi.org/10.48550/arXiv.2403.17980

Department of Energy. (2022). ThunderSecure: Deploying Real-time Intrusion Detection for 100G networks. OSTI. https://www.osti.gov/servlets/purl/1867680

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset

[2] https://www.unb.ca/cic/datasets/ids-2017.html

Contactenos

📧 **Correo electrónico:** csirt@unad.edu.co
🌐 **Página web:** https://csirt.unad.edu.co

🛡️ El CSIRT Académico UNAD está siempre disponible para apoyarte ante consultas o inquietudes relacionadas con la protección de la información en la universidad. No dudes en ponerte en contacto con nuestro equipo para recibir asesoría, reportar incidentes o recibir orientación en temas de seguridad digital. ¡Tu seguridad es nuestra prioridad!

[1] https://research.unsw.edu.au/projects/unsw-nb15-dataset
[2] https://www.unb.ca/cic/datasets/ids-2017.html