Fortalecimiento del trabajo Policial mediante Big Data y minería de datos para la identificación de criminales y redes delictivas

Enhancing police work through Big Data and data mining for the identification of criminals and criminal networks

• Fecha de recepción: 2025-10-05 • Fecha de aceptación: 2025-10-19 • Fecha de publicación: 2025-12-12

Guido Fabián Chamba Iza[1]

Resumen

El manejo de grandes volúmenes de datos, conocido como Big Data, se caracteriza por su variedad, velocidad y volumen, siendo la base tecnológica de múltiples herramientas disruptivas como la minería de datos, la inteligencia artificial (IA) y el aprendizaje automático (ML). En el ámbito de la seguridad pública, estas tecnologías permiten gestionar información estructurada y no estructurada proveniente de fuentes diversas: bases policiales, instituciones públicas y privadas, redes sociales, sensores urbanos y dispositivos móviles. La información procesada y transformada en inteligencia se convierte en un insumo esencial para que las actividades policiales sean más efectivas antes, durante y después de la comisión de un delito.

Por su parte, la minería de datos integra procedimientos estadísticos y computacionales orientados a identificar regularidades y asociaciones no evidentes en grandes bases de datos, posibilitando funciones de clasificación, segmentación, predicción, detección de anomalías y análisis de redes.

La combinación de ambas herramientas potencia la inteligencia policial predictiva, facilitando la identificación de individuos con comportamientos delictivos, la detección de vínculos con organizaciones criminales y la toma de decisiones basadas en evidencia. Este trabajo analiza un caso de integración de fuentes policiales y judiciales en Ecuador, demostrando cómo estas tecnologías fortalecen la identificación de redes criminales y visibilizan actores cuya detección sería compleja mediante métodos tradicionales.

Palabras clave: análisis de redes; Big Data; inteligencia policial; minería de datos; prevención del delito; seguridad pública

Abstract

Big Data refers to the management of massive volumes of data characterized by variety, velocity, and volume, forming the technological foundation for disruptive tools such as data mining, artificial intelligence, and machine learning. In public safety, these technologies enable the management of structured and unstructured information from diverse sources including police databases, public and private institutions, social media, urban sensors, and mobile devices. Processed and transformed into intelligence, this information becomes essential for making police operations more effective before, during, and after criminal events.

[1] Discente del Instituto Superior Tecnológico Policía Nacional, Quito-Ecuador, guido.chamba@policia.gob.ec, https://orcid.org/0009-0009-1608-584X

Data mining comprises analytical techniques aimed at discovering patterns, correlations, and hidden insights in large datasets, allowing functions such as classification, segmentation, prediction, anomaly detection, and network analysis.
The combination of both tools enhances predictive police intelligence by facilitating the identification of individuals with criminal behaviors, detecting links with criminal organizations, and enabling evidence-based decision-making. This study analyzes a case of integrating police and judicial sources in Ecuador to demonstrate how these technologies strengthen the identification of criminal networks and reveal actors that are difficult to detect through traditional means.
Keywords: Big Data; criminal networks; crime prevention; data mining; police intelligence; public safety

Introducción

El crecimiento exponencial de fuentes de datos digitales —como registros policiales, redes sociales, bases judiciales, cámaras de videovigilancia, sensores urbanos y dispositivos móviles— ha transformado las capacidades de investigación y prevención del delito. En este contexto, el uso de tecnologías emergentes de la cuarta revolución industrial, como Big Data y minería de datos, permite anticipar comportamientos delictivos mediante el análisis automatizado de información masiva (Chen et al. 2004).

Este artículo analiza cómo la integración de estas herramientas puede fortalecer el trabajo policial en Ecuador, especialmente en la identificación de antisociales vinculados a grupos delictivos organizados. El objetivo es demostrar que la correlación de múltiples fuentes de datos permite generar inteligencia criminal capaz de detectar relaciones y estructuras criminales ocultas, apoyando la toma de decisiones estratégicas y operativas en la lucha contra el crimen organizado.

Estado del Arte

La aplicación de Big Data y minería de datos en el ámbito policial ha ganado protagonismo en la literatura científica durante las dos últimas décadas. (Chen et al. 2004) propusieron un marco general para la minería de datos criminal, destacando su utilidad para la detección de patrones y correlaciones ocultas en información delictiva. (Xu y Chen 2005) demostraron que la visualización de vínculos facilita comprender estructuras criminales complejas.

Posteriormente, (Akoglu, Tong y Koutra 2015) revisaron enfoques de detección de anomalías en grafos aplicables al análisis de redes delictivas. (Silva y Rocha 2019) analizaron los desafíos institucionales en la adopción de Big Data en cuerpos policiales latinoamericanos, enfatizando la importancia de la infraestructura tecnológica, la interoperabilidad de sistemas y la ética en el uso de datos personales.

Asimismo, (Maimon y Rokach 2010) y (Hand, Mannila y Smyth 2001) proporcionaron los fundamentos algorítmicos de la minería de datos, mientras que (Bhattacharyya y Jha 2011) exploraron su aplicación en la detección de fraudes, ofreciendo analogías metodológicas útiles para el análisis criminal.

Estos estudios demuestran que la integración de fuentes heterogéneas, el análisis de redes sociales y la visualización de datos son factores clave para fortalecer la inteligencia policial. No obstante, existe una brecha en la aplicación contextualizada de estas tecnologías en países latinoamericanos,

particularmente en Ecuador, donde la investigación empírica en seguridad basada en datos aún es incipiente.

Big Data en el contexto policial

El concepto de Big Data indica al manejo de información masiva que, por su diversidad y rapidez de generación, requiere tecnologías capaces de procesarla casi en tiempo real (Gandomi y Haider 2015). En el ámbito policial, su uso posibilita integrar registros estructurados, como antecedentes judiciales, con contenidos no estructurados —videos o publicaciones digitales— para construir una visión integral del fenómeno criminal (Gandomi y Haider 2015).

El análisis en tiempo casi real de estos datos posibilita detectar patrones de comportamiento, correlaciones entre individuos y conexiones con grupos delictivos. Las plataformas de análisis, como IBM i2 Analyst's Notebook (IBM 2020), o soluciones basadas en ML, ofrecen capacidades de integración, filtrado y visualización relacional que facilitan la construcción de mapas criminales y modelos predictivos de riesgo.

Minería de datos para la identificación de patrones delictivos

La minería de datos permite extraer conocimiento útil mediante algoritmos de clasificación, asociación, agrupamiento y análisis de redes sociales. Estas técnicas se aplican en investigaciones policiales para descubrir comunidades delictivas, identificar jerarquías internas y detectar relaciones ocultas entre actores y eventos (Xu y Chen 2005).

Entre las funciones principales destacan:

Clasificación: permite etiquetar individuos como sospechosos o no sospechosos según patrones históricos.

Asociación: identifica relaciones frecuentes entre delitos, personas o ubicaciones.

Clusterización: detecta grupos o células delictivas con características comunes.

Análisis de redes sociales: mapea relaciones jerárquicas y roles dentro de organizaciones criminales.

Su aplicación fortalece la inteligencia operativa y estratégica, optimizando recursos y permitiendo intervenciones más focalizadas (Maimon y Rokach 2010).

Metodología

El presente estudio empleó un enfoque analítico basado en la integración de tres fuentes de información institucionales:

Fuente A: Base de datos de antisociales con vínculos comprobados o presuntos a grupos delictivos organizados, elaborada por unidades policiales.

Fuente B: Base de datos de partes de detención policial, que incluye información de personas aprehendidas, fechas, lugares y motivos.

Fuente C: Base de datos de noticias del delito de la Fiscalía General del Estado (FGE), que registra los casos investigativos y la tipificación delictiva.

Se aplicó una metodología de cuatro etapas:

1. **Recolección de datos**: Captura de registros estructurados y validación cruzada de identidades a través de cédulas, alias y grupos delictivos.
2. **Preprocesamiento**: Limpieza y normalización de datos para eliminar duplicados, corregir inconsistencias y estandarizar entidades.
3. **Análisis de datos**: Aplicación de minería descriptiva, segmentación y análisis de vínculos para identificar relaciones directas e indirectas entre individuos.
4. **Visualización**: Construcción de diagramas relacionales y mapas de calor georreferenciados para representar la concentración espacial de eventos y asociaciones.

El enfoque metodológico propuesto se enmarca en estudios previos sobre análisis criminal basados en minería de datos, asegurando rigor científico y replicabilidad.

Además, durante el proceso de análisis se aplicaron procedimientos manuales de correlación de entidades mediante coincidencia aproximada de nombres, alias y números de cédula. Para ello se utilizó una hoja de cálculo en Microsoft Excel, usando herramientas y fórmulas que ayudan la identificación de duplicados, lo que permitió comparar registros entre las tres bases de datos y detectar coincidencias parciales que no eran evidentes bajo una búsqueda exacta.

Este enfoque manual de vinculación fortaleció la detección de relaciones interpersonales entre individuos asociados a grupos delictivos organizados, incrementando la precisión en la construcción de redes y perfiles delictivos.

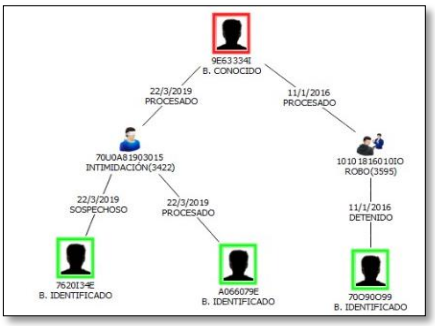
Resultados

Los resultados evidenciaron que la integración de datos provenientes de fuentes policiales y judiciales permitió detectar individuos relacionados con grupos delictivos organizados (GDO). Se identificaron cabecillas y miembros recurrentes en distintos eventos delictivos, cuyas relaciones interpersonales y territoriales fueron mapeadas a través de diagramas de red. Para lo cual se describen dos casos:

En el primer ejemplo (véase gráfico 1), el antisocial (B. CONOCIDO) de siglas C.H.J.G. con número de identificación 9E63334I quien es cabecilla de un grupo delictual que opera en Durán. Al verificar en la base de datos de la Fiscalía registra 7 procesos (NDD) como procesado o sospechoso. En el NDD 10101816010IO aparece relacionado con el antisocial de siglas H.V.J.A. con número de identificación 70O90O99 como detenido, quien al verificar el sistema del parte policial registra 8 aprehensiones entre el 2018 y 2024. Mientras que, en el NDD 70U0A81903015 aparece relacionado con los antisociales de siglas: L.N.V.V. con número de identificación 7620I34E como sospechoso y P.C.C.P. con número de identificación A066079E como procesado, quien al verificar el sistema del parte policial registra 3 aprehensiones en el 2019.

Gráfico 1

Diagrama de relaciones entre individuos asociados a GDO

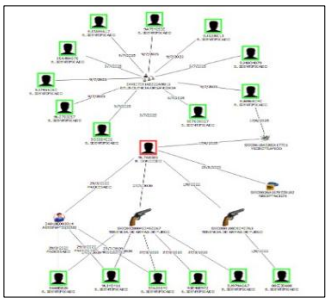


Fuente: Elaboración propia, 2025

En el segundo ejemplo (véase gráfico 2), el antisocial (B. CONOCIDO) de siglas J.V.J.I. con número de identificación 9U7A8585 quien es integrante de un grupo delictual que opera en la Zona 8. Al verificar en la base de datos de la Fiscalía registra 3 NDD como procesado. En el NDD 0A018E003314 aparece relacionado con los antisociales de siglas: B.G.T.S. con número de identificación 9EA36640 como procesado quien al verificar el sistema del parte policial registra 1 aprehensión en el 2020; C.C.R.H. con número de identificación 9A68E826 como procesado quien al verificar el sistema del parte policial registra 3 aprehensiones entre 2020 y 2024; y, U.R.B.A. con número de identificación 9UI48411 como procesado quien al verificar el sistema del parte policial registra 3 aprehensiones entre 2020 y 2024. El cabecilla en tres partes policiales aparece detenido con trece individuos por delincuencia organizada y tenencia de armas de fuego.

Gráfico 2

Diagrama de relaciones entre individuos asociados a GDO

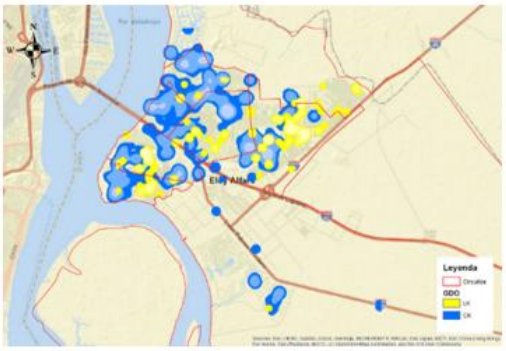


Fuente: Elaboración propia, 2025

El análisis de correlaciones reveló vínculos entre detenidos que, de manera aislada, no presentaban relación aparente, pero compartían conexiones indirectas mediante terceros. Asimismo, los mapas de calor georreferenciados de detenciones (véase gráfico 3) mostraron zonas de alta concentración criminal en sectores específicos del cantón Durán, coincidiendo con territorios de influencia de dos organizaciones delictivas antagónicas.

Gráfico 3

Mapa de calor de detenciones vinculadas a GDO en Durán



Fuente: Elaboración propia, 2025

Estos hallazgos confirman que la minería de datos aplicada al contexto policial es capaz de descubrir estructuras criminales ocultas y de proporcionar una visión integral del fenómeno delictivo (Xu y Chen 2005; Silva y Rocha 2019).

Discusión

Los resultados obtenidos demuestran la pertinencia de aplicar tecnologías de Big Data y minería de datos en el fortalecimiento de la inteligencia policial. De acuerdo con el National Institute of Justice (2013), el uso de analítica predictiva favorece la transición de estrategias reactivas a políticas preventivas sustentadas en datos.

Entre los principales beneficios destacan:

- Mayor precisión en la detección de sospechosos y patrones delictivos.
- Reducción de tiempos de investigación mediante análisis automatizado.
- Identificación de líderes y redes criminales ocultas.
- Optimización de recursos humanos y logísticos en operativos policiales.

No obstante, la implementación de estas tecnologías enfrenta desafíos éticos y técnicos, entre ellos la necesidad de garantizar la privacidad de los datos personales, la interoperabilidad entre instituciones y la capacitación del personal en análisis de datos. La ausencia de marcos regulatorios claros sobre el uso de información sensible podría generar riesgos de uso indebido o sesgos algorítmicos. Por tanto, se requiere una estrategia nacional de datos en seguridad ciudadana que defina políticas de gobernanza, auditoría y estándares éticos de aplicación (Silva y Rocha 2019).

Conclusiones

El uso de Big Data y minería de datos en el ámbito policial constituye un cambio de paradigma en la gestión de la seguridad pública. La capacidad de integrar fuentes heterogéneas y analizar correlaciones complejas permite identificar actores clave, mapear estructuras criminales y anticipar posibles eventos delictivos (Chen et al. 2004).

El estudio demostró que la correlación de bases policiales y judiciales revela vínculos relevantes entre individuos, fortaleciendo la inteligencia operativa y estratégica. Estos resultados refuerzan la necesidad de institucionalizar sistemas analíticos interconectados que faciliten la prevención del delito y la lucha contra el crimen organizado. Asimismo, la adopción de estas herramientas debe

acompañarse de políticas públicas que regulen su aplicación ética, garantizando transparencia, rendición de cuentas y respeto a los derechos humanos.

Recomendaciones

Institucionalización tecnológica: Implementar plataformas de análisis de *Big Data* interinstitucionales, que permitan compartir información policial, judicial y social bajo protocolos de seguridad y confidencialidad.

Capacitación especializada: Desarrollar programas de formación en analítica de datos, IA y visualización para agentes policiales y analistas criminales.

Gobernanza de datos: Establecer marcos legales que definan la propiedad, tratamiento y protección de la información utilizada en análisis policial.

Líneas de investigación futura: Profundizar en modelos predictivos basados en ML y redes neuronales que permitan anticipar comportamientos delictivos y patrones espaciales de criminalidad.

Colaboración internacional: Promover alianzas con universidades y centros de investigación para el desarrollo de sistemas de inteligencia policial basados en datos abiertos y tecnologías emergentes.

Estas acciones contribuirían a consolidar un modelo de seguridad sustentado en la evidencia, eficiente y éticamente responsable.

Bibliografía

Akoglu, Leman, Hanghang Tong, y Danai Koutra. 2015. “Graph-Based Anomaly Detection and Description: A Survey.” *Data Mining and Knowledge Discovery* 29 (3): 626–688.

Bhattacharyya, Siddhartha, y Sanjay Jha. 2011. “Data Mining for Credit Card Fraud: A Comparative Study.” *Decision Support Systems* 50 (3): 602–613.

Chen, Hsinchun, Wingyan Chung, Jennifer J. Xu, Gang Wang, Yi Qin, y Michael Chau. 2004. “Crime Data Mining: A General Framework and Some Examples.” *Computer* 37 (4): 50–56.

Gandomi, Amir, y Murtaza Haider. 2015. *Beyond the Hype: Big Data Concepts, Methods, and Analytics. International Journal of Information Management* 35 (2): 137–144.

Hand, David J., Heikki Mannila, y Padhraic Smyth. 2001. *Principles of Data Mining*. Cambridge, MA: MIT Press.

IBM. 2020. *IBM i2 Analyst’s Notebook*: *Turning Data into Intelligence*. White Paper. https://www.ibm.com

Maimon, Oded, y Lior Rokach, eds. 2010. *Data Mining and Knowledge Discovery Handbook*. 2ª ed. New York: Springer.

National Institute of Justice (NIJ). 2013. *Using Predictive Analytics to Prevent Crime*. U.S. Department of Justice. https://nij.ojp.gov/library/publications

Silva, Mariana, y Helder Rocha. 2019. “Big Data in Law Enforcement: Applications and Challenges.” *Journal of Big Data* 6 (1): 12–23.

Xu, Jennifer J., y Hsinchun Chen. 2005. “Criminal Network Analysis and Visualization.” *Communications of the ACM* 48 (6): 100–107.