Análisis Inteligente de Datos y Visualización aplicadas a la Investigación Criminal*

Sebastián P. WAHLER[1], Diego C. MARTÍNEZ[1], and Martín L. LARREA[1]

Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, Av. Alem 1253, B8000CPB Bahía Blanca, ARGENTINA.
spwahler@ing.unp.edu.ar, dcm@cs.uns.edu.ar, mll@cs.uns.edu.ar
https://cs.uns.edu.ar/

Abstract. Se presenta un estudio de las técnicas y metodologías actuales de análisis inteligente de datos y visualización para la asistencia en la investigación criminal, a partir de los registros de actividades delictivas, sus autores y las relaciones de datos que puedan derivarse a partir de ellas. Es de especial interés la identificación de redes ilegales, tales como bandas delictivas o criminales para propender a una persecución penal inteligente.

Keywords: Investigación Criminal · Análisis Inteligente de Datos · Redes Sociales · Visualización.

1 Introducción

En la actualidad las actividades criminales habituales en una ciudad o región van desde hurtos y robos de poca importancia, hasta otros de mayor gravedad como amenazas, cibercrimen, abusos sexuales y homicidios. Todos ellos son registrados de diferentes formas por las fuerzas de la ley, con datos de variada precisión que incluyen usualmente la tipificación del delito, los datos en tiempo y espacio, y en muchas ocasiones los autores correspondientes.

Toda esta información respalda los procesos de investigación judicial de cada caso, pero con el transcurso del tiempo constituyen una extensa base de conocimiento sobre la cual es posible extraer valiosa información para la prevención del delito y la búsqueda de la justicia. Por ejemplo, es posible identificar relaciones entre personas de acuerdo a un análisis transitivo de eventos criminales en tiempo y espacio que sugieren la conformación de bandas delictivas. Las relaciones de amistad o conveniencia entre diversos autores de actividades criminales también puede inferirse de los registros delictivos y es de extrema relevancia para la prevención del delito y la resolución de casos inconclusos.

En este trabajo es de especial interés la aplicación de estas técnicas y tecnologías utilizando los registros de actividades criminales de la Provincia de

* Universidad Nacional del Sur - Bahía Blanca - Argentina. Universidad Nacional de la Patagonia San Juan Bosco - Chubut - Argentina.

Chubut a través de la colaboración del Ministerio Público Fiscal de la provincia (parte del Poder Judicial con autonomía funcional para la investigación y persecución de conductas delictivas) y las instituciones que lo asisten.

1.1 Antecedentes

Las organizaciones criminales son grupos que operan fuera de la ley, realizando actividades ilegales en beneficio propio y en detrimento de otros individuos o grupos sociales [1]. Pueden ser de diverso tamaño y cubrir áreas geográficas variadas, en muchos casos en conflicto con otras organizaciones similares. Una de las características particulares de este tipo de organizaciones es que, al estar enfocadas en actividades ilegales perseguidas por los organismos de seguridad pública, el anonimato y/o la discreción de sus miembros es de vital importancia. Esto requiere estudios de la información existente con el fin de identificar los criminales y realizar acciones apropiadas para la prevención del delito.

Los miembros de las organizaciones criminales tienen a su vez diversos grados de compromiso con cada una de ellas. En muchos casos los hechos criminales que son evidentes en la sociedad ocurren por individuos de baja jerarquía y responsabilidad en el grupo, motivados por la recompensa inmediata, las aspiraciones de ascenso y la reputación en su propio círculo de contactos. Asimismo, existen otros individuos de mayor jerarquía y responsabilidad en la organización criminal, que ostentan cualidades de liderazgo, intereses a largo plazo, y la constante preocupación por la conservación del poder para el beneficio personal y de la organización. Con frecuencia, son los individuos del primer grupo los que cometen delitos percibidos y registrados por las fuerzas policiales, mientras que los miembros del segundo grupo se mantienen con mayor discreción. Adicionalmente, las estructuras jerárquicas, la forma de operar, y la cultura inherente de sus realidades socio-económicas imponen códigos propios que hacen difícil la identificación de la organización delictiva como un todo, con sus miembros y actividades relacionadas. Es aquí donde las Ciencias de la Computación puede aportar un rol significativo.

2 Análisis de Redes Sociales (SNA)

Desde hace algunos pocos años, el Análisis de Redes Sociales (o SNA por sus siglas en inglés de Social Network Analysis) ha contribuido a las investigaciones criminales y a las actividades de inteligencia relacionadas.

Una red social modela individuos como nodos, vinculados entre sí por arcos o aristas que representan las relaciones entre esos individuos. El estudio de estas redes es importante porque se enfoca en la abstracción de las relaciones humanas sobre uno o más aspectos particulares [2] [3]. De esta manera, las redes conforman estructuras de grafos en las cuales es posible identificar diversas propiedades, tales como la relevancia o la importancia relativa de los nodos individuales en función de las conexiones existentes o el flujo de información. De acuerdo a Sage [4] , existen cuatro pilares fundamentales del análisis de redes: el

reconocimiento de la importancia de las relaciones sociales entre los individuos, la recolección y análisis de datos sobre estas relaciones entre los individuos, la importancia de la representación visual de estos datos y la necesidad de modelos matemáticos y computacionales que expliquen los patrones de conexión entre los individuos.

En particular, la vinculación entre el estudio de las redes sociales y la investigación criminal ha sido encarada por varios autores. A mediados de los 70 se utilizaban modelos básicos para establecer y cualificar las relaciones entre individuos o actores de un escenario particular, definiendo grafos de acuerdo a la información recolectada [5], pero el procesamiento era mayoritariamente manual y con varias etapas de refinamiento y valoración de datos. Esta es la que según Klerk en [6] sería la primera generación de análisis de redes en criminalística. La segunda generación involucra el uso de herramientas computacionales que automatiza parte de la tarea de registro y estructuración de datos. Estas herramientas además aumentaron notoriamente la cantidad de datos que se pueden analizar, haciendo mucho más ágil su registro y consulta. La tercera y actual generación establece la definición de modelos y técnicas matemáticas para la generación de nuevo conocimiento, como la identificación de posiciones de poder e influencia o la calidad de potenciales testigos o informantes. Métricas como la centralidad de un nodo en un grafo son especialmente útiles en este escenario.

Uno de los trabajos más importantes al respecto es el de Krebs [7], en donde se identifica una parte de la red de terroristas que fué responsable de los atentados del 11 de septiembre de 2001 en Nueva York. Aquí identifica agrupaciones de individuos que se conectan entre sí por los pilotos responsables del secuestro de las aeronaves. Otros estudios similares han sido efectivos en consecuencia [8] [9] [10]. Por otro lado, el análisis de redes sociales ha cobrado también interés en la investigación criminal tradicional como las estructuras de la mafia o el narcotráfico [11] [12] [13] [14] [15]. Estudios como el de Malm [16] han permitido identificar roles en la cadena de suministros para la fabricación de drogas ilícitas, lo que acarrea diferentes riesgos penales para cada uno de los colaboradores. Otros estudios se enfocan en el uso del análisis de las redes sociales para otras actividades criminales, como el tráfico ilícito de arte [17], el lavado de dinero [18] [19], corrupción policial [20] y bandas juveniles [21] [22]. Existen también líneas de investigación en la disciplina referente al cibercrimen [23] [24] [25]. Es claro entonces que el análisis de redes sociales puede ser aplicado a un amplio rango de actividades criminales y ha demostrado modelar apropiadamente características propias de las organizaciones ilegales, asistiendo a la prevención del delito y al diseño de políticas adecuadas para enfrentar estas actividades.

Existen sin embargo algunas dificultades que requieren aún estudios intensivos. La cantidad de información que debe manejarse es enorme, en muchos casos con información incompleta, contradictoria y no menos frecuentemente incorrecta. Además, las relaciones humanas tradicionales se mezclan naturalmente con las interacciones ilícitas entre los individuos por lo que es necesario identificar apropiadamente su naturaleza y consecuencias y determinar los límites sensatos de la red social analizada.

Actualmente los organismos estatales encargados de la Justicia y la prevención del delito cuentan con registros informatizados de las actividades criminales detectadas, así como de las etapas y eventos del subsecuente proceso penal. En particular, para este Plan de Trabajo es de especial interés la información producida a tal efecto por las fuerzas policiales de la Provincia del Chubut y su Poder Judicial de la mano del Ministerio Público Fiscal (MPF). Existen decenas de miles de registros que son utilizados principalmente para la acción penal, pero que pueden ser empleados para modelar diferentes redes sociales sobre las cuales aplicar un análisis matemático y computacional en la búsqueda de nueva información. Esto permitirá conocer más sobre las actividades criminales y sus autores en la jurisdicción de esa provincia, con las particularidades propias de la información registrada digitalmente.

El análisis y exploración de estos grandes conjuntos de datos y sus relaciones debe ser asistido por técnicas y herramientas que faciliten este proceso y reduzcan la carga cognitiva que recae sobre los usuarios. En tal sentido, el área de Visualización de Información, en particular la Visualización de Grandes Conjuntos de Datos, busca asistir a los usuarios de tal manera. La aplicación de técnicas visuales para la representación de este tipo de información no es nueva [26] [27] [28]. Sin embargo, se requiere analizar los trabajos desarrollados para identificar aquellos aspectos que sean más favorables para la representación de la información propia que se manejara en esta tesis. También es importante el estudio de las tareas e interacciones que la visualización debe soportar [29], ya que son estas interacciones las que facilitan la exploración de la visualización de información.

3 Marco de Trabajo - Ministerio Público Fiscal del Chubut

En el marco de trabajo se ha optado por investigar en base al análisis criminal de redes delictivas en el Ministerio Público Fiscal del Chubut [30], perteneciente al Poder Judicial de la provincia.

Coirón, es el sistema informático que colabora con la administración del flujo de casos ingresados al Ministerio Público Fiscal del Chubut.

Es una herramienta que permite registrar, comunicar y gestionar las actividades, trámites y actuaciones que se realizan para un caso penal, desde la denuncia hasta su finalización.

Como herramienta de registro construye una base de datos con el historial de cada caso, así como de las personas involucradas y de los responsables de la gestión en cada oficina. Como herramienta de comunicación, agrupa la información, entrecruza las relaciones, identifica pertenencias y vinculaciones entre casos, personas, sus antecedentes y sus lazos. Como herramienta de gestión administra el flujo de casos y el trabajo de los integrantes de las oficinas responsables de los mismos. Permite planificar, organizar, coordinar y controlar el flujo de trabajo afín a cada caso y la sumatoria de ellos. Ha sido desarrollado a medida de las necesidades del Ministerio Público Fiscal del Chubut, tomando como base

el Código Procesal Penal vigente en el Chubut y adaptado a los lineamientos estratégicos de diseño y gestión de Oficinas Fiscales definidos por la Procuración General.

Su progreso, mantenimiento y mejora continua está a cargo de un Equipo de Desarrollo del Departamento de Informática del Área de Planificación y Control de Gestión de la Procuración General, del cual formo parte.

Una vez registrada toda la información relacionada a un hecho, y con ayuda de herramientas y vinculaciones con otros sistemas, se pueden obtener salidas que permiten llevar adelante la investigación de un caso o de un conjunto de hechos con características comunes.

Actualmente me encuentro trabajando en la incorporación de herramientas de visualización de información que permitirán ver en modo gráfico lo que hoy se muestra en grillas, y listados, potenciando el análisis que realizarán luego los especialistas.

Un medio de enfrentar, "a través del análisis criminal, la persecución penal de sujetos prolíficos es el perfilamiento relacional entre ellos, a través de la realización de vinculación de compañeros de delitos y de redes sociales; a fin de identificar si forman parte de una banda o alguna organización criminal mayor o de un fenómeno delictual más extenso" [31].

Se denomina "Grupo de Pertenencia" en el Sistema Coirón a la relación directa que existe entre un individuo dentro del universo de personas cargadas como actores de delitos (rol: denunciado, sospechoso o imputado) y otros individuos del mismo universo, con los cuales existan uno o más casos penales en común.

Crear una aplicación de software "Red de Grupos de Pertenencia" donde se muestre gráficamente las relaciones entre las personas involucradas en los casos penales es el objetivo principal de esta investigación. No sólo enfocarse en el grupo de pertenencia de una persona en particular, sino que mediante una visualización y con diversos filtros de búsqueda se logre mostrar gráficamente las relaciones entre un determinado grupo de personas y de esta manera poder inferir la conformación de posibles bandas delictivas.

La idea central es reflejar de manera gráfica, mediante un Grafo, los grupos de pertenencia. Se le llama nodo o node a cada círculo, y representa a una persona (con rol de imputado, sospechoso o denunciado) involucrada en dos o más casos penales. El tamaño del nodo posee una relación directa con la cantidad de casos penales en los que se encuentre involucrada la persona. Cuanto mayor sea el tamaño del nodo en más cantidad de casos penales estará involucrado.

Las líneas vinculan a las personas entre sí y representan el o los casos que tienen en común. Cuanto mayor sea el grosor de la misma, más casos hay entre ellas. Hay nodos que se encontrarán aislados en el grafo, esto no significa que no estén involucrados en casos, sino que quizás no existan relaciones para el filtro de búsqueda que se utilice en esa vista en particular.

Sample Heading (Third Level) Only two levels of headings should be numbered. Lower level headings remain unnumbered; they are formatted as run-in headings.

Sample Heading (Fourth Level) The contribution should contain no more than four levels of headings. Table 1 gives a summary of all heading levels.

Table 1. Table captions should be placed above the tables.

Heading level	Example	Font size and style
Title (centered)	**Lecture Notes**	14 point, bold
1st-level heading	**1 Introduction**	12 point, bold
2nd-level heading	**2.1 Printing Area**	10 point, bold
3rd-level heading	**Run-in Heading in Bold.** Text follows	10 point, bold
4th-level heading	*Lowest Level Heading.* Text follows	10 point, italic

Displayed equations are centered and set on a separate line.

$$x + y = z \tag{1}$$

Please try to avoid rasterized images for line-art diagrams and schemas. Whenever possible, use vector graphics instead (see Fig. 1).

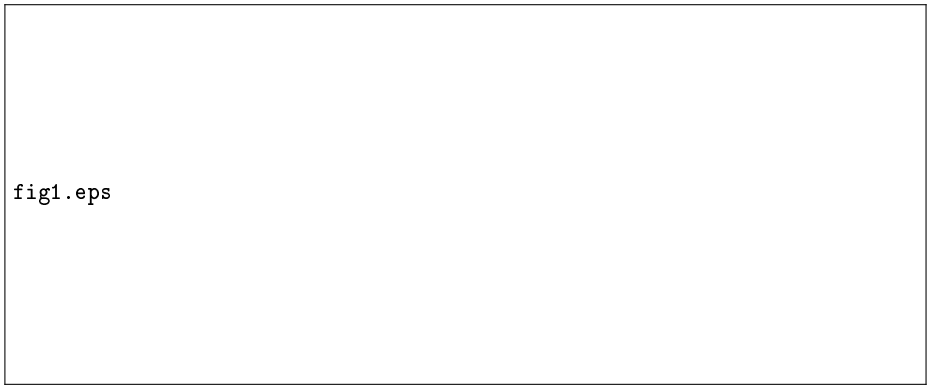

Fig. 1. A figure caption is always placed below the illustration. Please note that short captions are centered, while long ones are justified by the macro package automatically.

Theorem 1. *This is a sample theorem. The run-in heading is set in bold, while the following text appears in italics. Definitions, lemmas, propositions, and corollaries are styled the same way.*

Proof. Proofs, examples, and remarks have the initial word in italics, while the following text appears in normal font.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable. The following bibliography provides a sample reference list with entries for journal articles [1], an LNCS chapter [?], a book [?], proceedings without editors [?], and a homepage [?]. Multiple citations are grouped [1, ?,?], [1, ?,?,?].

Acknowledgements Please place your acknowledgments at the end of the paper, preceded by an unnumbered run-in heading (i.e. 3rd-level heading).

Referencias Bibliográficas

1. Finckenauer JO. Problems of definition: What is organized crime? Trends in Organized Crime. 2005; 8(3):63–83. https://doi.org/10.1007/s12117-005-1038-4
2. Krishna Raj P.M.Ankith MohanK.G. Srinivasa. Practical Social Network Analysis with Python. Springer, 2018. ISBN: 978-3-319-96746-2.
3. Burcher, Morgan. Social Network Analysis and Law Enforcement: Applications for Intelligence Analysis. Serie Crime Prevention and Security Management, 2020. ISBN 978-3-030-47770-7
4. L.C. Freeman, The SAGE handbook of social network analysis, eds. J. Scott, P.J. Carrington, SAGE Publications Ltd., 2011.
5. W.R. Harper, D.H. Harris, The application of link analysis to police intelligence. Hum. Factors 17(2), 157–164 (1975).
6. P. Klerks, The network paradigm applied to criminal organisations: theoretical nitpicking or relevant doctrine for investigators? Recent developments in the Netherlands. Connections 24(3), 53–65 (1999)
7. V. Krebs, Mapping networks of terrorist cells. Connections 24(3), 43–52 (2002)
8. R.M. Medina, Social network analysis: a case study of the Islamist terrorist network. Secur. J. 27(1), 97–121 (2014)
9. J. Qin, J.J. Xu, D. Hu, M. Sageman, H. Chen, Analyzing terrorist networks: a case study of the global jihad. Lect. Notes Comput. Sci. 3495, 287–304 (2005)
10. E. Stollenwerk, T. Dörfler, J. Schibberges, Taking a new perspective: mapping the Al Qaeda network through the eyes of the UN Security Council. Terror. Political Violence 28(5), 950–970 (2016)
11. M. Bouchard, J. Amirault, Advances in research on illicit networks. Glob. Crime 14(2–3), 119–122 (2013)
12. D.A. Bright, C. Greenhill, M. Reynolds, A. Ritter, C. Morselli, The use of actor-level attributes and centrality measures to identify key actors: a case study of an Australian drug trafficking network. J. Contemp. Crim. Justice 31(3), 262–278 (2015a)
13. L. Giommoni, A. Aziani, G. Berlusconi, How do illicit drugs move across countries? a network analysis of the heroin supply to Europe. J. Drug Issues 47(2), 217–240 (2016)
14. C. Morselli, Hells Angels in springtime. Trends Org. Crime 12(2), 145–158 (2009)
15. C. Morselli, Assessing vulnerable and strategic positions in a criminal network. J. Contemp. Crim. Justice 26(4), 382–392 (2010)

16. A. Malm, G. Bichler, Networks of collaborating criminals: assessing the structural vulnerability of drug markets. J. Res. Crime Delinq. 48(2), 271–297 (2011)
17. G. Bichler, A. Malm, Small arms, big guns: a dynamic model of illicit market opportunity. Glob. Crime 14(2–3), 261–286 (2013)
18. A.F. Colladon, E. Remondi, Using social network analysis to prevent money laundering. Expert Syst. Appl. 67, 49–58 (2017)
19. M.R.J. Soudijn, Using strangers for money: a discussion on money-launderers in organized crime. Trends Org. Crime 17(3), 199–217 (2014)
20. M. Lauchs, R. Keast, N. Yousefpour, Corrupt police networks: uncovering hidden relationship patterns, functions and roles. Polic. Soc. 21(1), 110–127 (2011)
21. J.M. McGloin, Policy and intervention considerations of a network analysis of street gangs. Criminol. Public Policy 4(3), 607–635 (2005)
22. G. Bichler, A. Malm, J. Enriquez, Magnetic facilities: identifying key juvenile convergence places with social network analysis. Crime Delinq. 60(7), 971–998 (2014)
23. D. Décary-Hétu, Information exchange paths in IRC hacking chat rooms, in Crime and networks, ed. by C. Morselli, (Routledge, New York, 2014)
24. D. Décary-Hétu, B. Dupont, The social network of hackers. Glob. Crime 13(3), 160–175 (2012)
25. D. Décary-Hétu, B. Dupont, Reputation in a dark network of online criminals. Glob. Crime 14(2–3), 175–196 (2013)
26. M1. Xu, Jennifer, and Hsinchun Chen. "Criminal network analysis and visualization." Communications of the ACM 48.6 (2005): 100-107.
27. M2. Feng, Mingchen, et al. "Big data analytics and mining for effective visualization and trends forecasting of crime data." IEEE Access 7 (2019): 106111-106123.
28. M3. Mathew, Ammu, et al. "Criminal Networks Mining and Visualization for Crime Investigation." Caniya and P, Mufeed and Raj, Asha, Criminal Networks Mining and Visualization for Crime Investigation (July 8, 2021) (2021).
29. M4. Chen, Hsinchun, et al. "Visualization in law enforcement." CHI'05 extended abstracts on Human factors in computing systems. 2005.
30. https://www.mpfchubut.gov.ar/
31. L. González, G. Rua. Sistemas Judiciales: Una perspectiva integral sobre la administración de justicia - Análisi Criminal, Publicación anual del CEJA e INECIP. N°23. Año 2019.