

Desmantelamiento óptimo de redes delincuenciales. Una perspectiva desde el modelado matemático y computacional

Tomas Angel Sarmiento Bahoque ¹, John Fredys Cantillo Palacio ²,
John Eduardo Realpe Gómez ³ y Javier Antonio Montoya Martínez ⁴

Recepción: 18-01-2016 | Aceptación: 20-07-2016 | En línea: 15-11-2016

MSC:91A06,91A12,91A20,91C99,91F99 | PACS:87.23.Ge

doi:10.17230/ingciencia.12.24.4

Resumen

El objetivo de este trabajo es estudiar y comparar diferentes estrategias para el desmantelamiento óptimo de redes delincuenciales, las cuales están representadas en algoritmos que permiten la identificación óptima de los individuos claves en la red. La estrategia de mayor complejidad se basa en la métrica de centralidad de Katz-Bonacich como medida de influencia en la red, y da lugar a un problema NP-difícil por lo que se debe recurrir a métodos heurísticos para encontrar soluciones aproximadas. Aquí se desarrolla un algoritmo basado en el método Monte Carlo y se compara con un método basado en algoritmos voraces introducido recientemente en la literatura. En este trabajo se compara además el desempeño de éstos con estrategias menos sofisticadas y se proporciona evidencia que dichos algoritmos se desempeñan relativamente bien, contribuyendo así a proporcionar

¹ Universidad de Cartagena, tsarmientob@unicartagena.edu.co, ORCID:<http://orcid.org/0000.0002-6413-7204> Cartagena, Colombia.

² Universidad de Cartagena, jcantillop@unicartagena.edu.co, Cartagena, Colombia.

³ Universidad de Cartagena, jrealpeg@unicartagena.edu.co, ORCID:<http://orcid.org/0000.0002-5473-2464>, Cartagena, Colombia.

⁴ Universidad de Cartagena, jmontoyam@unicartagena.edu.co, ORCID:<http://orcid.org/0000.0003-0507-45XX>, Cartagena, Colombia.

un mejor entendimiento de éstos. Se discute además un modelo introducido recientemente que justifica el uso de la centralidad de Katz-Bonacich desde el punto de vista de la teoría de juegos sobre redes.

Palabras clave: Sistemas complejos; modelado computacional; modelos en redes; teoría de juegos; redes delincuenciales; mecánica estadística; TICs.

Optimal Dismantling of Criminal Networks. A Perspective from the Mathematical and Computational Modeling.

Abstract

This work deals with the study and comparison of different strategies for the optimal dismantling of delinquent networks, which aim to optimally identify the most relevant individuals in the network. The strategy of greater complexity that we have studied here, is based on the Katz-Bonacich centrality criteria as a measure of influence of the individuals in the network. This results in an NP-hard type of problem, therefore, in order to apply that criteria, we must use heuristic methods which allow us to find approximate solutions. In particular, the methods used in this work are the Monte Carlo and greedy algorithms. We compared their performance against less sophisticated strategies and we were able to find that these algorithms perform relatively better, which contributes to improve our understanding of these approaches. In addition, we discuss a model that was recently introduced, which justifies the use of Katz-Bonacich centrality from the point of view of game theory on networks.

Key words: Complex systems; computational modeling; network models; game theory; delinquent networks; statistical mechanics; ICT.

1 Introducción

Avances recientes en las tecnologías de la información y las comunicaciones han permitido el reconocimiento del importante rol que desempeñan las interacciones sociales que emergen de diversos fenómenos socio-económicos [1],[2],[3]; esta idea en gran medida se ha inspirado en conceptos y métodos empleados en el estudio de sistemas físicos constituidos por muchas partículas [4],[5], incentivando la investigación de problemas formulados en la intersección entre las matemáticas, la ciencia de la computación y las ciencias socio-económicas [6]; comprendiendo así que en esencia, una sociedad o una economía, pueden conceptualizarse como un sistema de muchos

agentes que interactúan entre sí, de modo que cada agente actúa buscando obtener la mejor utilidad posible para sí, condicionado al comportamiento de los demás agentes y condicionando a su vez el comportamiento de éstos. La formulación matemática de estas ideas toma lugar dentro de la llamada teoría de juegos [3],[7].

Esto ha hecho además que las ciencias socio-económicas adquieran un carácter más cuantitativo, abriendo así un espacio donde las llamadas ciencias duras, como la matemática y la física, pueden contribuir con toda una artillería analítica desarrollada a lo largo de los siglos. Esto constituye una excelente oportunidad para contribuir desde una perspectiva más cuantitativa al progreso social de los países en vía de desarrollo, los cuales se enfrentan usualmente a fenómenos tan complejos como la pobreza, la inequidad, y en particular, a redes delincuenciales que constituyen un ejemplo persistente de este tipo de fenómenos, el cual ha tenido un creciente interés para muchos investigadores en los últimos años desde esta perspectiva [8],[9],[10],[11],[12],[13],[14].

Una de las hipótesis que se han explorado, conocida como la teoría de las ventanas rotas [15], asume que el comportamiento criminal se contagia como una epidemia cuya difusión genera patrones espacio-temporales de actividad criminal en la región geográfica en la cual se encuentra inmersa [12].

Otro enfoque consiste en estudiar la influencia que puede tener la estructura de la red social que conforman un grupo de individuos en su comportamiento criminal [8],[16],[17]. La motivación detrás de esta manera de atacar el problema es la posibilidad de que las políticas de reducción del crimen se puedan hacer más efectivas al incorporar técnicas de análisis que permitan dilucidar las propiedades estructurales de la red que se busca dismantelar.

Y es precisamente el concepto de red, o grafo, uno de los que ha generado mayor interés debido a su ubicuidad y generalidad: un conjunto de puntos (nodos) y un conjunto de líneas (vínculos) que une algunos pares de nodos. Un problema importante en las aplicaciones socio-económicas es la caracterización de los nodos más relevantes en una red y el diseño de algoritmos para identificarlos. Esto da lugar a un problema computacional que resulta ser NP-difícil (NP-hard), lo cual significa que no se conoce un

algoritmo eficiente para resolverlo y, más aún, que no se espera que exista ya que esto equivaldría a resolver de manera eficiente toda una vasta clase de problemas de gran relevancia en diversas áreas de la ciencia y la ingeniería [18]. De manera que uno de los retos en esta área es el desarrollo de algoritmos eficientes, si bien aproximados, que permitan realizar los análisis necesarios en un tiempo razonable. Tener la capacidad de encontrar los nodos más relevantes permitiría, por ejemplo, concentrarse en los integrantes más influyentes en una red social, ya sea para maximizar el impacto de una campaña publicitaria o para bloquear de manera eficiente la propagación de una epidemia [19],[20]. Con este fin se han introducido diferentes versiones de un concepto denominado centralidad, el cual cuantifica precisamente el grado de influencia de un nodo dado en una red. Estas variaciones van desde las que tienen en cuenta propiedades netamente locales, como el número de vínculos que emergen de un nodo, hasta propiedades globales más complejas como el número de caminos que lo atraviesan [21].

En este trabajo se estudia el problema de redes delincuenciales desde el punto de vista del modelado matemático y computacional. El modelo de redes delincuenciales aquí estudiado fue propuesto recientemente [8] y se basa en la teoría de juegos [22],[23], la cual provee un marco conceptual bastante general para estudiar interacciones sociales de toda índole. Este modelo sugiere una estrategia de desmantelamiento basada en la medida de centralidad de Katz-Bonacich [24],[25], la cual cuantifica el número de caminos que atraviesan un nodo dado; la relevancia de dichos caminos disminuye exponencialmente con su longitud. Aquí se estudian varios métodos algorítmicos para implementar dicha estrategia y se hace un análisis comparativo con otro tipo de estrategias más tradicionales. En particular, se adapta un algoritmo basado en el método Monte Carlo y se compara con un método basado en algoritmos voraces introducido recientemente en la literatura [8]. En este trabajo se compara además el desempeño de éstos con estrategias menos sofisticadas y se proporciona evidencia que dichos algoritmos se desempeñan relativamente bien, contribuyendo así a proporcionar un mejor entendimiento de éstos.

La topología de las redes delincuenciales aquí estudiadas corresponde a la de un grafo aleatorio, o de Erdős-Rényi, el cual se obtiene conectando pares de nodos al azar con una cierta probabilidad [26],[27]. Esto se hace con el objeto de facilitar este primer estudio sobre el tema; se debe tener

en cuenta, sin embargo, que es de esperar que la topología de las redes delincuenciales se asemeje más a la de una red compleja [28] y por lo tanto el análisis aquí propuesto debe ser extendido en futuros trabajos.

2 Definición del problema

2.1 Descripción de redes delincuenciales

El modelo aquí investigado hace uso de la hipótesis de racionalidad de la teoría de juegos [22],[23]. Esta se debe tomar como una primera aproximación al problema que debe ser luego comparada con modelos más realistas que incluyen dinámicas de aprendizaje basadas en teorías cognitivas [29], tales como el modelo investigado por algunos de los autores en el contexto de cooperación [30]. También se asume que el proceso de desmantelamiento de la red sucede suficientemente rápido de modo que la red no tiene tiempo de reorganizarse manteniéndose por lo tanto constante, excepto por los cambios debidos a la remoción de individuos producto del desmantelamiento de ésta.

Considere un conjunto de n delincuentes y un conjunto de vínculos entre ellos. Esto define una red delincencial que se denotará por g . La *matriz de adyacencia* que describe la red se denotará por G y sus componentes para todo par de delincuentes i y j se denotará por g_{ij} . Por definición $g_{ij} = 1$ si y solo si los delincuentes i y j han cometido un delito juntos y $g_{ij} = 0$ de otro modo. Se denotara por x_i el esfuerzo delincencial del delincuente i y por $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ el perfil de esfuerzos de la población delincencial con $0 \leq x_i \leq 1$ para cada i . Siguiendo a Ballester *et al.* [8] la *ganancia delincencial esperada* para el delincuente i está dada por

$$u_i(\mathbf{x}) = y_i(\mathbf{x}) - p_i(\mathbf{x}, g)f. \quad (1)$$

Aquí $y_i(\mathbf{x})$ corresponde a la ganancia bruta del delincuente i , mientras que $p_i(\mathbf{x}, g)$ y f se refieren a la probabilidad que el delincuente i sea capturado y a la intensidad de la sanción recibida (valor de la multa o tiempo de cárcel), respectivamente. Dicha ganancia o utilidad de un delincuente depende no solo del esfuerzo del delincuente en consideración, sino también del esfuerzo de todos los delincuentes y de la estructura de la red delincencial. Esto se

debe a la *competencia global* por los recursos limitados y a la *colaboración local* entre delincuentes.

De acuerdo con la hipótesis de racionalidad de la teoría de juegos, los delincuentes escogerán el nivel de su esfuerzo delincencial x_i^* que maximice su ganancia $u_i(\mathbf{x}, g)$ dados los niveles de esfuerzo de los demás delincuentes $\mathbf{x}_{-i} = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$. El perfil de esfuerzos dado por $\mathbf{x}^* = \{x_1^*, \dots, x_n^*\}$ constituye lo que se conoce en la teoría de juegos como un *equilibrio de Nash* del sistema. Se asumirá por simplicidad que [8]

$$y_i(\mathbf{x}) = x_i \text{ máx} \left\{ 1 - \delta \sum_{j=1}^n x_j, 0 \right\}, \quad (2)$$

$$p_i(\mathbf{x}, g) = p_0 x_i \text{ máx} \left\{ 1 - \phi \sum_{j=1}^n g_{ij} x_j, 0 \right\}. \quad (3)$$

La primera expresión modela la competencia global entre delincuentes, dependiendo por tanto de la suma de las actividades delincuenciales de *todos* los individuos en la red. La segunda expresión, por otra parte, modela la colaboración local entre individuos cercanos en la red delincencial, dependiendo por tanto de la suma de las actividades delincuenciales de los individuos vecinos en dicha red. Ambas expresiones incluyen la función máx con el fin de modelar procesos de saturación; por ejemplo, una vez que hay una alta actividad delincencial a nivel global, cualquier actividad delincencial adicional no cambiará mucho los incentivos de los individuos. Los parámetros $\delta > 0$ y $\phi \geq 0$ caracterizan la intensidad de dicha competencia global en la ganancia bruta y la intensidad de dicha colaboración y complementariedad local entre los delincuentes, respectivamente.

2.2 Caracterización del equilibrio de Nash

Ballester *et al.* [8] demostraron que el juego definido por las ecuaciones (1),(2), y (3) tiene un único equilibrio de Nash, el cual puede expresarse en términos de una noción de centralidad. Se entiende por centralidad a una caracterización matemática del grado de influencia que un nodo tiene en

una red; existen diferentes tipos de centralidades, tales como la *centralidad de grado (degree centrality)* y la *centralidad de Katz-Bonacich*, entre otras [31],[32],[33]. La centralidad de Katz-Bonacich es la noción relevante en este caso como se discute a continuación.

Para comenzar, se asume que en el equilibrio de Nash \mathbf{x}^* se tiene

$$1 - \delta \sum_{j=1}^n x_j^* \geq 0 \quad \text{y} \quad 1 - \phi \sum_{j=1}^n g_{ij} x_j^* \geq 0; \quad (4)$$

la consistencia de esta hipótesis y su rango de validez deben ser chequeados al final del razonamiento que sigue. Teniendo en cuenta esta hipótesis en las ecuaciones (1),(2), y (3) que definen la función de utilidad de cada delincuente, y después de un poco de álgebra se tiene que

$$u_i(\mathbf{x}, g) = (1 - \pi)x_i - \delta x_i^2 - \delta \sum_{\substack{j=1 \\ j \neq i}}^n x_i x_j + x_i \pi \phi \sum_{j=1}^n g_{ij} x_j, \quad (5)$$

donde $\pi = p_0 f$ es el costo marginal esperado de la pena para un delincuente aislado.

La condición de equilibrio está dada por la expresión

$$0 = \frac{\partial u_i(\mathbf{x})}{\partial x_i} = (1 - \pi) - \delta x_i - \delta \sum_{j=1}^n x_j + \pi \phi \sum_{j=1}^n g_{ij} x_j. \quad (6)$$

Introduciendo el escalar $x = \sum_{j=1}^n x_j$, el vector $\mathbf{1} \in \mathbb{R}^n$ cuyas componentes son todas iguales a uno, y la matriz identidad $n \times n$, denotada por \mathbf{I} , la Ecuación (6) puede escribirse en forma vectorial como

$$\mathbf{0} = (1 - \pi - \delta x)\mathbf{1} - \delta \mathbf{x} - \phi \pi G \mathbf{x},$$

cuya solución está dada por

$$\mathbf{x}^* = \mathbf{b}(g, \alpha) \left(\frac{1 - \pi - \delta x^*}{\delta} \right). \quad (7)$$

Aquí $\alpha = \pi \phi / \delta$, y

$$\mathbf{b}(g, \alpha) = (\mathbf{I} - \alpha G)^{-1} \mathbf{1}, \quad (8)$$

asumiendo que $\mathbf{I} - \alpha G$ es invertible, es conocido como el *vector de centralidades de Katz-Bonacich*. Finalmente, se debe tener en cuenta que

$$x^* = \sum x_j^* = \left(\frac{1 - \pi - \delta x^*}{\delta} \right) b(g, \alpha), \quad (9)$$

donde $b(g, \alpha) = \sum_{j=1}^n b_j(g, \alpha)$ es el nivel de centralidad global de la red.

Despejando x^* en (9) y reemplazando en (7) se tiene finalmente que

$$\mathbf{x}^* = \mathbf{b}(g, \alpha) \frac{1 - \pi}{\delta[1 + b(g, \alpha)]}. \quad (10)$$

El rango de validez de esta solución está determinado por la condición $\alpha \rho(g) < 1$, donde $\rho(g) = \max_k |\lambda_k|$, con λ_k , los autovalores de la matriz de adyacencia G , es el radio espectral de G . Esta condición garantiza que la matriz $\mathbf{I} - \alpha G$ en la Ecuación(8) es invertible y además que cada $x_i^* \geq 0$.

El vector de centralidades de Katz-Bonacich $\mathbf{b}(g, \alpha)$, se puede expresar en términos de sus componentes como

$$b_i(g, \alpha) = 1 + \alpha \sum_{j \in \partial i} b_j(g, \alpha), \quad (11)$$

donde ∂i denota el conjunto de vecinos del delincuente i .

Esta ecuación se puede resolver escogiendo una centralidad inicial para cada nodo, introduciéndola en el lado derecho para calcular nuevos valores de centralidad, e iterando este proceso hasta su convergencia, es decir hasta que no haya más variación en los valores de centralidad obtenidos mediante una iteración dentro de una tolerancia dada.

Este proceso iterativo se puede considerar como un proceso difusivo no conservativo en una red. Un *proceso difusivo* en una red es un proceso estocástico que distribuye alguna cantidad entre cada uno de sus nodos. Si el total de la cantidad en consideración distribuida en la red se mantiene constante el proceso difusivo es *conservativo*, en caso contrario, el proceso difusivo es *no conservativo* [34].

En este caso la cantidad que se distribuye es la centralidad de Katz-Bonacich estimada, y el peso total de la red corresponde a la suma de cada

una de las centralidades estimadas. Dado que esta última cantidad no se conserva durante el método iterativo descrito anteriormente, éste es un proceso difusivo no conservativo. Un ejemplo importante de proceso difusivo conservativo es el celebrado algoritmo *PageRank* utilizado por Google para clasificar páginas web.

2.3 Estrategias de desmantelamiento

El objetivo principal de esta investigación es estudiar algunas estrategias que permitan propinar la máxima reducción posible en el nivel agregado de actividad delincriminal dado por

$$x^*(g) = \frac{1 - \pi}{\delta[1 + b(g, \alpha)]} b(g, \alpha). \quad (12)$$

En general, estas se pueden clasificar en estrategias globales que intervienen la red de manera homogénea, o locales que tienen en cuenta la heterogeneidad del sistema inducida por la topología de la red. Algunos ejemplos son

Global: Esta consiste en incrementar a $\pi = p_0 f$ mediante, por ejemplo, un incremento en la intensidad de las sanciones f o un incremento de la probabilidad de atrapar a un delincuente p_0 . Estas no tienen en cuenta de manera explícita la topología de la red.

Local: Asumiendo que sólo se tienen recursos para la captura o remoción de k delincuentes de la red, esta estrategia consiste en concentrar dichos recursos en los k individuos que mayor impacto tengan en la reducción del nivel agregado de actividad delincriminal dado por la expresión (12). Si se denota con $K = \{i_1, \dots, i_k\}$ el conjunto de delincuentes a ser removidos de la red, y la red resultante como $g - K$, el objetivo es entonces encontrar el conjunto K^* tal que

$$K^* = \arg \min_{\{K: |K|=k\}} \{x^*(g - K, \alpha)\}.$$

Dado que $\partial x^* / \partial b(g, \alpha) > 0$ esto equivale a

$$K^* = \arg \min_{\{K: |K|=k\}} \{b(g - K, \alpha)\}. \quad (13)$$

Nótese que bajo la hipótesis que usamos en este trabajo de que la red *original* a desmantelar g es constante, este problema de minimización es equivalente a la maximización de la llamada intercentralidad $d_K(g, \alpha) = b(g, \alpha) - b(g - K, \alpha)$ introducida en la definición 3 de la referencia [8], como se puede observar en la Ecuación (11) de la misma referencia. Esta hipótesis es válida siempre y cuando el desmantelamiento se haga suficientemente rápido en relación con el tiempo que le toma a la red re-estructurarse.

La estrategia global es tal vez la que tradicionalmente más se ha utilizado. Sin embargo, con el advenimiento del análisis matemático y computacional de redes de toda índole, estrategias del segundo tipo pasaron a ser el centro de atención. La estrategia global es tal vez la que tradicionalmente más se ha utilizado. Sin embargo, con el advenimiento del análisis matemático y computacional de redes de toda índole, estrategias del segundo tipo pasaron a ser el centro de atención.

Este trabajo se enfoca en el estudio de algunos métodos que permitan implementar la estrategia local descrita anteriormente. Note que ésta estrategia da lugar a un problema de optimización combinatoria donde la función objetivo no está dada de manera explícita. En efecto, el vector de centralidades de Katz-Bonacich depende de la topología de la red mediante la Ecuación (12), y ésta varía al remover k nodos. Dicho problema de optimización es NP-difícil o ‘NP-hard’ [8]. El método más directo de resolver este problema de optimización sería resolviendo dicha ecuación para todos los posibles conjuntos de k nodos a remover, calculando la suma de las centralidades de cada nodo, y escogiendo el conjunto que corresponda al valor mínimo de ésta. Sin embargo, dicho método no es práctico pues escala exponencialmente con el número de nodos a remover. Aquí exploraremos algunas heurísticas que permiten obtener respuestas aproximadas de una manera más eficiente.

2.4 Aproximación analítica de la centralidad

El resultado que se discute a continuación no se debe tomar como una demostración rigurosa sino más bien como una aproximación heurística con el objetivo de ganar un poco de intuición sobre el problema. Que tan buena resulta esta aproximación analítica se discutirá en la Sección 4 donde se comparará con métodos computacionales más precisos.

Introduciendo un vector de variables binarias $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ tal que para cada i se tiene

$$\sigma_i = \begin{cases} 0, & \text{si el nodo ha sido removido,} \\ 1, & \text{en otro caso,} \end{cases} \quad (14)$$

la centralidad de Katz-Bonacich de un delincuente i , cuando la configuración de delincuentes removidos está dada por el vector σ , puede escribirse como

$$b_i(\sigma) = \sigma_i \left(1 + \alpha \sum_{j \in \partial i} b_j(\sigma) \right). \quad (15)$$

La centralidad global b de la red está dada por

$$b = \sum_{i=1}^N b_i(\sigma) = \sum_{i=1}^N \sigma_i + \alpha \sum_{i=1}^N \sigma_i \sum_{j \in \partial i} b_j(\sigma) = \sigma + \alpha \sum_{i=1}^N b_i(\sigma) \sum_{j \in \partial i} \sigma_j,$$

donde $\sigma = \sum_{i=1}^N \sigma_i$. Para poder proseguir con el cálculo se asumirá que

$$\frac{1}{|\partial i|} \sum_{j \in \partial i} \sigma_j \approx \frac{1}{N} \sum_{j=1}^N \sigma_j, \quad (16)$$

donde $|\partial i|$ es el número de vecinos de i . Este tipo de aproximación se conoce en física usualmente como la aproximación de campo medio, pues equivale a aproximar el promedio local con respecto a los vecinos de i con el promedio global sobre toda la red. Finalmente, aproximando el número de vecinos de i con el número de vecinos promedio en la red C , es decir $|\partial i| \approx C$, se obtiene

$$b \approx \sigma + \alpha b \frac{C}{N} \sigma;$$

despejando b de esta ecuación se llega a

$$\langle b \rangle \approx \frac{\langle \sigma \rangle}{1 - \alpha C \langle \sigma \rangle} = \frac{1 - f}{1 - \alpha C(1 - f)}, \quad (17)$$

donde se introdujo la fracción de nodos removidos f . Este resultado se espera que sea válido cuando las correlaciones en el sistema sean suficientemente débiles.

3 Métodos computacionales

3.1 Método de Ballester *et al.*

Este método, propuesto en la referencia [8], se basa en una heurística *greedy* o avara. Esta heurística se utiliza para resolver problemas de optimización y consiste en tomar la elección óptima en cada iteración. Dado que esta heurística no utiliza información global, ésta puede quedar atrapada en un óptimo local.

Para resolver aproximadamente la Ecuación (13) mediante esta heurística se siguen los siguientes pasos:

1. Se calcula la centralidad de cada uno de los nodos de la red resolviendo la Ecuación (11) mediante el proceso difusivo no conservativo discutido anteriormente.
2. Se remueve de la red un nodo que tenga el mayor valor de centralidad y se calcula nuevamente cada una de las centralidades tomando como condición inicial del proceso difusivo no conservativo las centralidades obtenidas en el paso anterior.
3. Se itera este proceso hasta que hayan sido removidos k nodos.

3.2 Método de Monte Carlo

El método de Monte Carlo es un método no determinista muy versátil desarrollado por el físico Enrico Fermi y el matemático Stanislaw Ulam con el objetivo de estudiar proceso nucleares. Con el advenimiento de la era digital, este método ha encontrado importantes aplicaciones en diversas áreas de la ciencia y la ingeniería. A continuación se describe como se puede utilizar el método de Monte Carlo para encontrar una solución aproximada de la Ecuación (13).

1. Se escogen k nodos al azar para ser removidos y se calcula la centralidad de cada uno de los nodos de la nueva red así obtenida, resolviendo la Ecuación (11) mediante el proceso difusivo no conservativo descrito anteriormente.

2. Se intercambian dos nodos escogidos al azar, uno de los k nodos removidos en el paso anterior y uno de los $N - k$ nodos restantes, y se calcula nuevamente cada una de las centralidades tomando como condición inicial del proceso difusivo no conservativo las centralidades obtenidas en el paso anterior. Si la centralidad global disminuye se acepta el cambio, de lo contrario se acepta con probabilidad $e^{-\beta\Delta b}$, donde Δb es la diferencia de la nueva centralidad global menos la anterior, y β es un parámetro a fijar.
3. Se itera el proceso anterior un número determinado de veces, guardando en cada paso los valores mínimos obtenidos y el conjunto de nodos removidos correspondiente.

4 Resultados numéricos

Inicialmente este estudio se concentró en el estudio de redes delincuenciales que tienen la topología de un grafo aleatorio o de Erdős-Rényi. Para generar estos grafos se tuvo en cuenta el siguiente procedimiento.

1. Se genera un arreglo de N nodos.
2. Se escoge un par de nodos al azar que no estén conectados y se adiciona un vínculo entre ellos.
3. Se itera el paso anterior hasta que se hayan adicionado M vínculos.

La conectividad, o número de vínculos, promedio por nodo en el grafo está dado por $C = 2M/N$ [35].

Luego de generar estos grafos, se buscó tener una idea de cómo se distribuye la influencia en redes delincuenciales que tienen la topología de un grafo aleatorio. Para tal fin, en la Figura 1 se presentan la fracción de nodos que tienen un valor de centralidad de Katz-Bonacich dado promediado sobre 100 grafos de 100 nodos, para diferentes valores de los parámetros del modelo: número M de vínculos en el grafo, el cual fija la conectividad promedio en $C = 2M/N$, y parámetro de centralidad α .

Los histogramas obtenidos se asemejan mucho a una distribución de Poisson, lo cual puede ser una consecuencia de la topología del grafo ya

que la fracción de nodos con una conectividad dada también sigue una distribución de Poisson. Para una demostración de este hecho se puede ver por ejemplo [35]. Es probable que siguiendo una estrategia similar se pueda extender dicha demostración al caso de la centralidad de Katz-Bonacich, la cual es una noción global de centralidad en contraste con la conectividad que es una noción netamente local. Sin embargo esto es sólo una hipótesis que encontramos plausible.

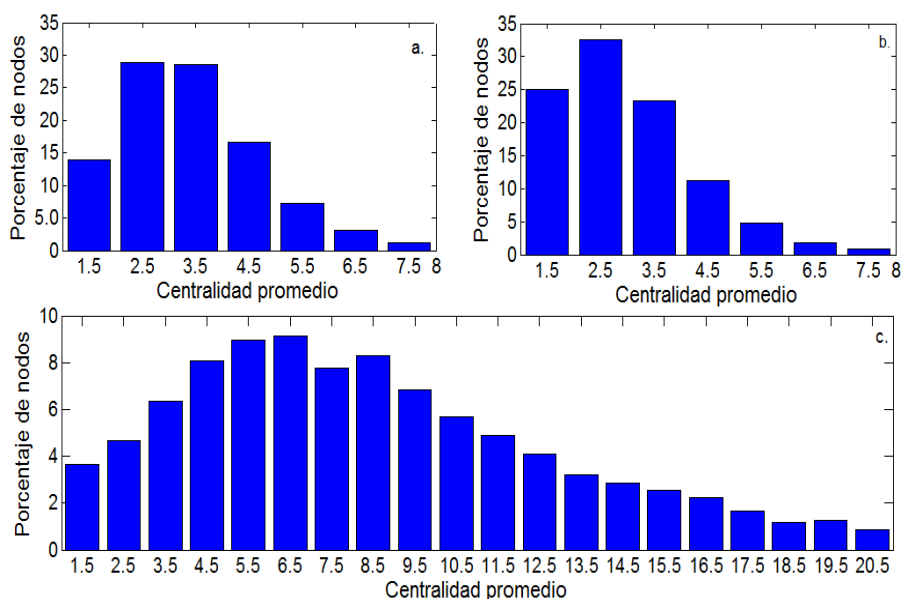


Figura 1: *Distribución de la centralidad.* Fracción de nodos que tienen un valor de centralidad dado promediado sobre 100 grafos aleatorios de 100 nodos, para diferentes valores de los parámetros del modelo: a) $M = 200$ y $\alpha = 0.15$ b) $M = 200$ y $\alpha = 0.18$, y c) $M = 150$ y $\alpha = 0.18$.

Posteriormente se utilizaron diferentes métodos para estudiar la variación de la centralidad total de Katz-Bonacich b en función de la fracción de nodos removidos f .

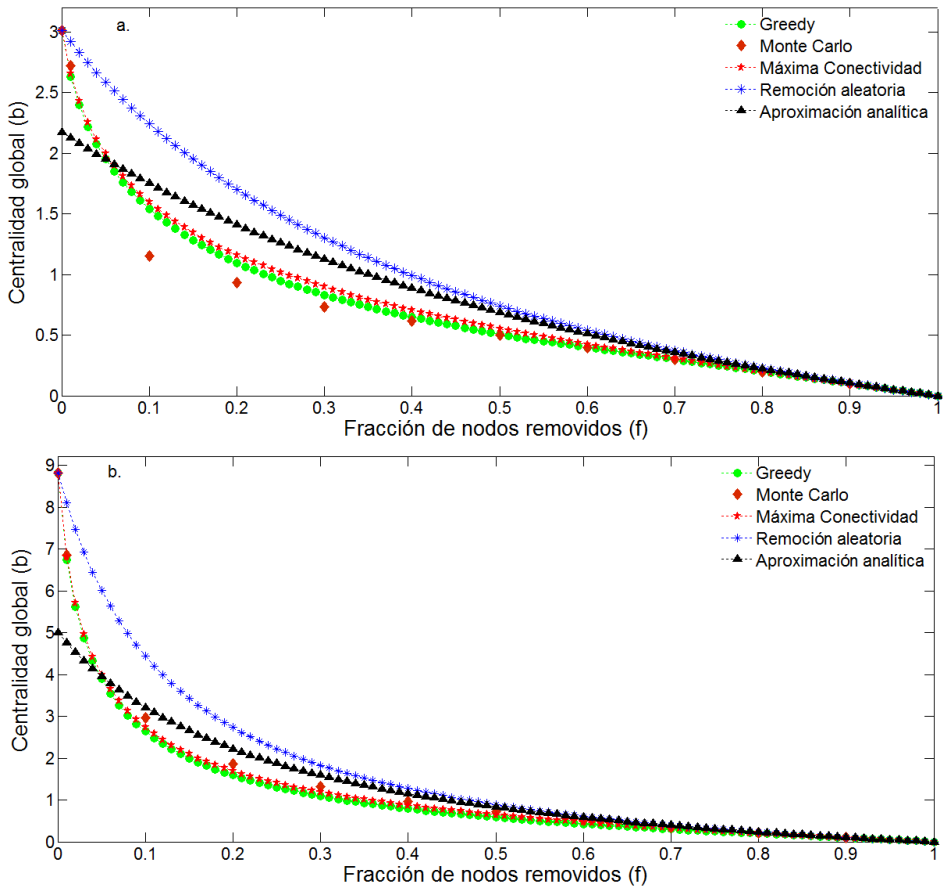
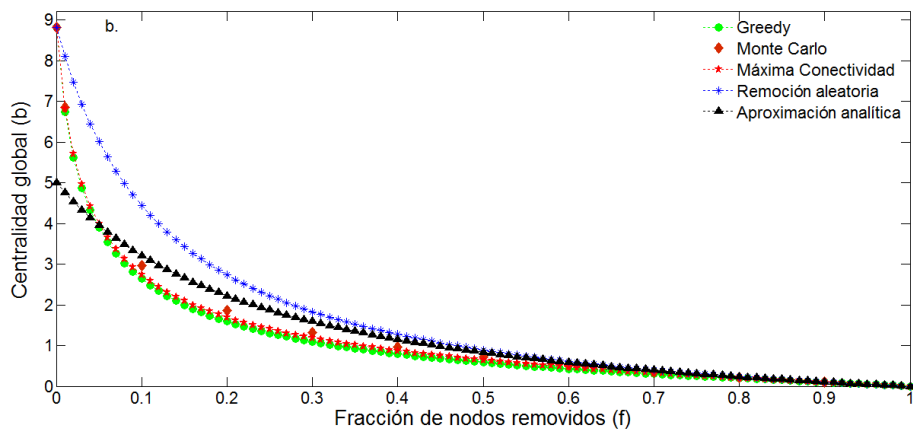


Figura 2: *Contraste de los diferentes métodos.* Centralidad global b como función de la fracción de nodos removidos f en la red obtenida con los diferentes métodos discutidos en el texto, a saber: método Monte Carlo, método de Ballester *et al.*, o greedy, remoción iterativa de los nodos con máxima conectividad, remoción de nodos aleatoria, y aproximación analítica. Los resultados numéricos corresponden a promedios sobre 2000 grafos aleatorios de 100 nodos y un parámetro de centralidad de $\alpha = 0,18$. En la implementación del método de Monte Carlo se utilizaron 10^4 repeticiones y un valor de $\beta = 10^4$; este valor de β fue el que produjo los mejores resultados. Los dos casos mostrados corresponden a grafos con diferente número de vínculos: a) $M = 150$, y b) $M = 400$.



La Figura 2 muestra los resultados obtenidos para diferentes valores de los parámetros del modelo. Entre los métodos utilizados están el método de Ballester *et al.*, o método *greedy*, y el método Monte Carlo, discutidos en la sección anterior. Como referencia se utilizó el caso donde los nodos se remueven al azar sin considerar información alguna sobre el sistema, lo cual corresponde al peor desempeño que se podría lograr. Un método adicional que se utilizó fue el de remover iterativamente el nodo con mayor conectividad (no centralidad) hasta remover un número especificado de nodos. Finalmente, se incluyó también la aproximación analítica dada por la Ecuación (17).

En general, se puede observar que el método *greedy* alcanza la mayor reducción posible en la centralidad global, seguido de cerca por el método de máxima conectividad. Esto indica que hay una estrecha correlación entre conectividad y centralidad de Katz-Bonacich en grafos aleatorios, lo cual es consistente con lo que se observó en la Figura 1. El método de Monte Carlo, con los parámetros utilizados, también se desempeña bastante bien aunque no siempre tan bien como el método *greedy*. Todos estos métodos alcanzan valores de centralidad global mucho menores que el método de remoción aleatoria, especialmente cuando se han removido menos del 60 % de los nodos. Note que aquí se está graficando la centralidad global normalizada por el número de nodos de la red inicial, por lo que en la práctica, dichas diferencias deben multiplicarse por el número de nodos de la red a desmantelar y por lo tanto es mucho mayor que lo que indica la figura.

Finalmente, la estimación analítica de la centralidad dada por la Ecuación (17) como función de la fracción de nodos removidos funciona relativamente bien cuando se han removido poco más del 40% de los nodos en la red. Esto es comprensible dado que en dicha aproximación se ignoran las correlaciones entre variables correspondientes a distintos nodos en el grafo y como tal se espera que su calidad se deteriore con el aumento de la conectividad del grafo.

5 Conclusiones

En este trabajo se estudiaron diferentes métodos para el desmantelamiento de redes delincuenciales basados en el modelo propuesto por Ballester *et al.* [8]. Este problema es NP-difícil (NP-hard) por lo cual se requieren métodos heurísticos que permitan obtener soluciones aproximadas en un tiempo razonable. Uno de los métodos estudiados fue propuesto en la misma referencia citada anteriormente, el cual garantiza una solución dentro del 33% del valor óptimo. Otro método estudiado fue una adaptación del método de Monte Carlo de la física estadística. También se consideró la remoción iterativa del nodo con máxima conectividad, y la remoción aleatoria de nodos. Este último sirvió como referencia ya que no utiliza información alguna sobre el sistema, y por lo tanto corresponde al peor desempeño que se puede obtener.

En general, salvo por el método de remoción aleatoria, todos los métodos presentaron un desempeño similar y mucho mejor que el de referencia anteriormente mencionado. En particular, se observó una fuerte correlación entre la medida de influencia netamente local dada por la conectividad, o número de vecinos, de un nodo con la medida de influencia más global dada por la centralidad de Katz-Bonacich, que tiene en cuenta todos los posibles caminos que atraviesan un nodo. Sin embargo, la topología de las redes estudiadas corresponde a la de un grafo aleatorio o de Erdős-Rényi. Se espera que la topología de redes delincuenciales reales se asemeje más a la de una red compleja [27],[28] y por lo tanto el análisis aquí desarrollado debe ser extendido a este tipo de topología en futuros estudios.

Agradecimientos

J.F. Cantillo agradece al programa Jóvenes Investigadores de Colciencias. J.A. Montoya y J.E. Realpe agradecen a la Vicerrectoría de Investigación de la Universidad de Cartagena su apoyo a través de convocatorias para el fortalecimiento de grupos de investigación.

Referencias

- [1] L. E. Blume, W. A. Brock, S. N. Durlauf, and Y. M. Ioannides, “Identification of social interactions,” pp. 853–964, 2011. [Online]. Available: <http://dx.doi.org/10.1016/B978-0-444-53707-2.00001-3> 84
- [2] S. N. Durlauf and Y. M. Ioannides, “Social interactions,” *Annu. Rev. Econ.*, vol. 2, no. 1, pp. 451–478, 2010. [Online]. Available: <http://www.annualreviews.org/doi/abs/10.1146/annurev.economics.050708.143312?journalCode=economics> 84
- [3] M. O. Jackson and Others, *Social and economic networks*. Princeton University Press, 2008, vol. 3. [Online]. Available: http://www.nber.org/econometrics_minicourse_2014/Jackson-NBER-slides2014_lecture1.pdf 84, 85
- [4] S. N. Durlauf, “How can statistical mechanics contribute to social science?” *Proceedings of the National Academy of Sciences*, vol. 96, no. 19, pp. 10 582–10 584, 1999. [Online]. Available: <http://www.pnas.org/content/96/19/10582.full> 84
- [5] T. C. Schelling, “Dynamic models of segregation,” *Journal of mathematical sociology*, vol. 1, no. 2, pp. 143–186, 1971. [Online]. Available: <http://dx.doi.org/10.1080/0022250X.1971.9989794> 84
- [6] S. Fortunato, M. Macy, and S. Redner, “Editorial,” *Journal of Statistical Physics*, vol. 151, no. 1-2, pp. 1–8, 2013. [Online]. Available: <http://dx.doi.org/10.1007/s10955-013-0703-2> 84
- [7] M. O. Jackson and Y. Zenou, “Games on networks,” *Handbook of game theory*, vol. 4, p. 1, 2014. [Online]. Available: <http://voxeu.org/sites/default/files/file/DP9127.pdf> 85

- [8] C. Ballester, Y. Zenou, and A. Calvó-Armengol, “Delinquent networks,” *Journal of the European Economic Association*, vol. 8, no. 1, pp. 34–61, 2010. [Online]. Available: <https://dx.doi.org/10.1111/j.1542-4774.2010.tb00494.x> 85, 86, 87, 88, 92, 94, 99
- [9] G. S. Becker, “Crime and punishment: An economic approach,” in *Journal of Political Economy*. Springer, 1968, vol. 76, no. 2, pp. 169–217. [Online]. Available: <http://dx.doi.org/10.1086/259394> 85
- [10] S. N. Durlauf and D. S. Nagin, “The deterrent effect of imprisonment,” in *Controlling Crime: Strategies and Tradeoffs*, J. M. Philip Cook; Jens Ludwig; Ed. University of Chicago Press, 2010, pp. 43–94. [Online]. Available: <http://www.nber.org/chapters/c12078.pdf> 85
- [11] C. Ballester, A. Calvó-Armengol, and Y. Zenou, “Who’s who in networks. wanted: the key player,” *Econometrica*, vol. 74, no. 5, pp. 1403–1417, 2006. 85
- [12] M. B. Short, P. J. Brantingham, A. L. Bertozzi, and G. E. Tita, “Dissipation and displacement of hotspots in reaction-diffusion models of crime,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 9, pp. 3961–3965, 2010. [Online]. Available: <http://www.pnas.org/content/107/9/3961.full> 85
- [13] M. B. Gordon, “A random walk in the literature on criminality: A partial and critical view on some statistical analyses and modelling approaches,” *European Journal of Applied Mathematics*, vol. 21, no. 4-5, pp. 283–306, 2010. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00404405/file/CrimeModeling_Gordonv2.pdf 85
- [14] E. L. Glaeser, B. Sacerdote, and J. A. Scheinkman, “Crime and Social Interactions,” *The Quarterly Journal of Economics*, vol. 111, no. 2, pp. 507–548, 1996. [Online]. Available: <http://dx.doi.org/10.2307/2946686> 85
- [15] K. Keizer, S. Lindenberg, and L. Steg, “The spreading of disorder,” *Science*, vol. 322, no. 5908, pp. 1681–1685, 2008. [Online]. Available: <http://dx.doi.org/10.1126/science.1161405> 85
- [16] A. Calvó-Armengoi and Y. Zenou, “Social networks and crime decisions: The role of social structure in facilitating delinquent behavior,” *International Economic Review*, vol. 45, no. 3, pp. 939–958, 2004. [Online]. Available: <http://www.jstor.org/stable/3663642> 85
- [17] C. Morselli, *Inside criminal networks*. Springer, 2009. [Online]. Available: <http://dx.doi.org/10.1007/978-0-387-09526-4> 85

- [18] C. Moore and S. Mertens, *The nature of computation*. OUP Oxford, 2011. [Online]. Available: http://dannyreviews.com/h/Nature_Computation.html?iframe=true&width=90%& 86
- [19] F. Altarelli, A. Braunstein, L. Dall’Asta, J. R. Wakeling, and R. Zecchina, “Containing Epidemic Outbreaks by Message-Passing Techniques,” *Phys. Rev. X*, vol. 4, no. 2, p. 21024, May 2014. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevX.4.021024> 86
- [20] C. E. A. Cabrera, E. J. A. Lotero, and V. G. Umaña, “Modelos epidemiológicos en redes: una presentación introductoria,” *Boletín de Matemáticas*, vol. 22, no. 1, pp. 21–37, 2015. [Online]. Available: <http://www.revistas.unal.edu.co/index.php/bolma/article/viewFile/51844/51641> 86
- [21] E. Estrada, *The Structure of Complex Networks: Theory and Applications*. OUP Oxford, 2011. [Online]. Available: <http://dx.doi.org/10.1093/acprof:oso/9780199591756.001.0001> 86
- [22] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1947. 86, 87
- [23] A. T. Villalón and A. M. Caraballo, “Un paseo por la historia de la Teoría de Juegos,” *Boletín de Matemáticas*, vol. 22, no. 1, pp. 77–95, 2015. [Online]. Available: <http://www.revistas.unal.edu.co/index.php/bolma/article/view/51847> 86, 87
- [24] L. Katz, “A new status index derived from sociometric analysis,” *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953. [Online]. Available: <http://dx.doi.org/10.1007/BF02289026> 86
- [25] P. Bonacich, “Power and Centrality: A Family of Measures,” *American Journal of Sociology*, vol. 92, no. 5, pp. pp. 1170–1182, 1987. [Online]. Available: <http://www.jstor.org/stable/2780000> 86
- [26] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, 2002. [Online]. Available: <https://doi.org/10.1103/RevModPhys.74.47> 86
- [27] M. E. J. Newman, “Resource Letter CS-1: Complex Systems,” *American Journal of Physics*, vol. 79, no. 8, pp. 800–810, 2011. [Online]. Available: <http://dx.doi.org/10.1119/1.3590372> 86, 99
- [28] J. Xu and H. Chen, “The Topology of Dark Networks,” *Commun. ACM*, vol. 51, no. 10, pp. 58–65, 2008. [Online]. Available: <https://doi.org/10.1145/1400181.1400198> 87, 99

- [29] P. W. Glimcher and E. Fehr, *Neuroeconomics: Decision Making and the Brain*. Elsevier Science, 2013. 87
- [30] J. Realpe-Gomez, B. Szczesny, L. Dall’Asta, and T. Galla, “Fixation and escape times in stochastic game learning,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2012, no. 10, p. P10022, 2012. [Online]. Available: <http://stacks.iop.org/1742-5468/2012/i=10/a=P10022> 87
- [31] L. C. Freeman, “A set of measures of centrality based on betweenness,” *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977. [Online]. Available: <http://dx.doi.org/10.2307/3033543> 89
- [32] —, “Centrality in social networks conceptual clarification,” *Social networks*, vol. 1, no. 3, pp. 215–239, 1978. [Online]. Available: <http://leonidzhukov.ru/hse/2013/socialnetworks/papers/freeman79-centrality.pdf> 89
- [33] P. Bonacich and P. Lloyd, “Eigenvector-like measures of centrality for asymmetric relations,” *Social Networks*, vol. 23, no. 3, pp. 191–201, 2001. [Online]. Available: [http://dx.doi.org/10.1016/S0378-8733\(01\)00038-7](http://dx.doi.org/10.1016/S0378-8733(01)00038-7) 89
- [34] R. Ghosh, K. Lerman, T. Surachawala, K. Voevodski, and S.-h. Teng, “Non-conservative Diffusion and its Application to Social Network Analysis,” *arXiv:1102.4639*, p. 11, 2011. [Online]. Available: <http://arxiv.org/abs/1102.4639> 90
- [35] A. K. Hartmann and M. Weigt, *Phase Transitions in Combinatorial Optimization Problems: Basics, Algorithms and Statistical Mechanics*. Wiley, 2006. 95, 96