

# Análisis de Redes Multiplexas en Escándalos de Corrupción

Luis Higuera<sup>1</sup>, Elizabeth Mesa<sup>1</sup>, Kevin Ramos<sup>1</sup>

---

## Abstract

Mediante el procesamiento de lenguaje natural se realizó una extracción de datos los cuales fueron utilizados para generar una red de corrupción de algunos de los escándalos públicos más conocidos en Colombia. Posteriormente, se aplicaron diferentes categorías a los vínculos haciendo la red multiplexa, después, se evaluaron ciertas medidas de centralidad para caracterizar la estructura de la red. Con el fin de obtener información de las características multiplexas de la red se comprobó la existencia de una interdependencia entre el grado de un nodo y sus tipos de relaciones, para esto se utilizaron métodos de Monte Carlo para simular una función de probabilidad conjunta utilizando las propiedades de las frecuencias ya conocidas de la red, con el fin de evaluar distintas probabilidades condicionales y así concluir que las probabilidades no son independientes. Esto tiene grandes implicaciones en el estudio de la evolución de la red en relación a sus propiedades estructurales.

**Keywords:** Corrupción, Redes, Complejidad, Redes multiplexas, Interdependencia estructural, Probabilidades conjuntas.

---

## 1. Introducción

Para Vito Tanzi (1998), la corrupción es el abuso del poder público para beneficio privado. De forma similar, para John Kramer (1977) la corrupción es el comportamiento de los funcionarios públicos que diverge de los deberes formales de un papel público para servir a fines privados. Incluso, Transparencia Gonzalez et al. (2018) siempre definió la corrupción como el abuso de cargos públicos para ganancia privada.

Estas definiciones coinciden en que la corrupción es una conducta que aprovecha o abusa de un puesto o cargo público con el fin de obtener una ganancia o un beneficio privado. Este fenómeno ha afectado en distintos factores a muchos países. Colombia no es ajena a la corrupción, pues con cierta frecuencia salen a la luz escándalos de corrupción de entidades, empresas u entes gubernamentales públicos o privados.

Según Newman and Ángel (2017), los indicadores de corrupción a nivel nacional e internacional muestran un panorama complejo para Colombia. En el Índice de Percepción de Corrupción elaborado en el año 2016, Colombia obtuvo una calificación de 37 puntos en una escala de 0 -mayor percepción de corrupción- a 100 -menor percepción de corrupción-. Entre 2009 y 2016 se registraron 3966 casos en el sistema penal oral acusatorio, relacionados a distintas formas de corrupción, esto es alrededor de 500 por año.

Existen diferentes organismos nacionales e internacionales que desempeñan un papel crucial en la lucha contra esta problemática social. Internacionalmente, la ONU promueve la lucha contra la corrupción a través de la Convención de las Naciones Unidas contra la Corrupción (UNCAC), que es el primer instrumento global jurídicamente vinculante para prevenir y combatir la corrupción, el Banco Mundial a través de programas y financiamiento, promueve la gobernanza transparente y la integridad en los países receptores de ayuda, y apoya iniciativas de fortalecimiento institucional para prevenir y combatir la cor-

rupción. World Bank (2021). También Transparencia Internacional, que es una organización que se dedica a combatir la corrupción a nivel global. Realiza investigaciones, promueve estándares de transparencia y ética, y aboga por la rendición de cuentas en todos los sectores de la sociedad. Transparencia Internacional también produce el Índice de Percepción de la Corrupción, que evalúa la percepción de la corrupción en diferentes países.

En Colombia, la Procuraduría General de la Nación es una entidad encargada de velar por la moralidad administrativa y la transparencia en el sector público en Colombia. Procuraduría de la Nación (2023). La Contraloría General de la República es responsable de ejercer el control fiscal en Colombia. Su objetivo es prevenir y detectar irregularidades en la gestión pública, incluyendo casos de corrupción, mediante la auditoría y fiscalización de los recursos del Estado. También el Departamento Administrativo de la Función Pública, el cual tiene como misión promover la eficiencia, la transparencia y la ética en el servicio público colombiano. Este organismo se encarga de formular y supervisar las políticas de gestión del talento humano en el sector público, con el fin de prevenir la corrupción y mejorar la calidad de los servicios. Función Pública (2023).

Estas y otras instituciones nacionales son clave en la lucha contra la corrupción en Colombia. Sin embargo, el fenómeno de la corrupción es complejo y afecta a diversas áreas de la sociedad. Por lo tanto, para abordar de manera integral este desafío, se requiere una perspectiva multidisciplinaria.

Diversas áreas del conocimiento como las ciencias políticas, la sociología, la antropología, la psicología, el derecho y la economía, han realizado algunas contribuciones sobre este fenómeno y han señalado los impactos económicos, políticos y sociales que genera.

Hoy en día, se suma el enfoque de la complejidad y la ciencia de redes, que permiten analizar la corrupción desde una perspectiva sistémica. Estas disciplinas nos brindan herramientas

para visualizar y comprender las interconexiones y dinámicas subyacentes en las redes de corrupción, revelando la complejidad de este fenómeno y abriendo nuevas oportunidades para su estudio y combate.

En ese proyecto, se aprovecharon los recientes avances en el procesamiento del lenguaje natural y el aprendizaje automático para analizar de manera eficiente y a gran escala la gran cantidad de información textual relacionada con los escándalos. Los modelos de lenguaje, como GPT-3.5, brindaron una forma sofisticada de interpretar y analizar texto Wadhwa et al. (2023), lo que permitió extraer y categorizar entidades y relaciones a partir de la abundante cantidad de datos de noticias y artículos relacionados con dichos escándalos. A través de este enfoque, se logró construir una red que incluirá como nodos a los actores mencionados en las noticias de los escándalos de corrupción y sus relaciones los vínculos encontrados en forma de frases verbales.

## 1.1. Estado del Arte

### 1.1.1. Corrupción y Complejidad

Desde hace varias décadas se han implementado diferentes índices y medidas al estudio de la corrupción que permiten realizar diferentes análisis cuantitativos de esta problemática social. Rose-Ackerman (1999).

Posteriormente, autores como D'Orsogna and Perc (2015) empezaron a implementar herramientas de la mecánica estadística al estudio de la corrupción. Sin embargo, según Gómez-Gardeñes et al. (2020), los autores afirman que los enfoques tradicionales carecen de las herramientas analíticas para manejar los aspectos estructurales y dinámicos que caracterizan los sistemas sociales, políticos y tecnológicos modernos donde tiene lugar la corrupción. Los autores señalan que la corrupción es un fenómeno sistémico y adaptativo que requiere enfoques integrales y multidisciplinarios para su prevención y combate efectivo. En respuesta a esto, sugieren que los sistemas complejos han surgido como un marco integral para estudiar fenómenos altamente adaptativos en entornos naturales.

### 1.1.2. Redes de Corrupción

Cuando la ciencia de redes comienza a aplicarse al estudio de la corrupción se utilizan técnicas de análisis de redes sociales para investigar las conexiones y las interacciones entre actores corruptos en diversos contextos, como gobiernos y empresas, esto se puede evidenciar en Wang (2020), donde estudian la red social de la corrupción en China para encontrar patrones en su comportamiento.

Luego, los investigadores comienzan a utilizar modelos de simulación para estudiar cómo se propaga la corrupción en las redes. Estos modelos ayudan a comprender los factores que contribuyen a la persistencia de la corrupción y a identificar estrategias de prevención y detección. Mahler-Hutter (2009). En paralelo, estos modelos han ayudado a construir y estudiar la dinámica y propagación de estas redes. Ribeiro et al. (2018)

Poco a poco se produce un aumento en la colaboración multidisciplinaria para abordar el problema de la corrupción y se aplican métodos de análisis de redes para investigar redes

políticas y estructuras de poder en relación con la corrupción. Granados and Nicolás-Carlock (2021).

El estudio de la corrupción mediante la ciencia de redes continúa evolucionando y ampliándose. Cada vez se utilizan enfoques más sofisticados, como el análisis de múltiples capas y el análisis de flujos de dinero y recursos, para comprender la dinámica y los efectos de la corrupción en diferentes sistemas. Diviák et al. (2019).

### 1.1.3. Redes de Corrupción en Colombia

En Colombia, se ha llevado a cabo un estudio pionero que aborda el fenómeno de la corrupción política desde una perspectiva económica, partiendo del análisis del Public Choice e incorporando un análisis de redes para caracterizar la complejidad estructural que definen los procesos de corrupción política, clientelismo y favorecimientos judiciales en Colombia en el caso del Cartel de la Toga. Gonzalez et al. (2018).

## 2. Metodología y Datos

Debido a la posibilidad de representar sistemas compuestos por agentes y sus relaciones desde la teoría de redes, ha sido posible estudiar diferentes tipos de sistemas sociales, como lo serían las redes de corrupción. En este proyecto de investigación, se analizó la red de escándalos de corrupción en la política de Colombia.

Aquí presentamos una breve descripción de cada caso de corrupción:

- **AgroIngreso Seguro:** Es un escándalo de corrupción que surgió en Colombia en 2009, donde se reveló que los subsidios del gobierno destinados a apoyar a pequeños agricultores estaban siendo desviados de manera irregular hacia grandes terratenientes y figuras políticas.
- **Cartel de la toga:** Es un caso de corrupción que involucra a magistrados del sistema judicial colombiano, en el que se denunciaron ventas de fallos judiciales y tráfico de influencias a cambio de sobornos.
- **Caso Centros Poblados:** Este caso de corrupción surgió en 2021 en Colombia. Involucra irregularidades en un contrato de 1.07 billones de pesos para la instalación de internet gratuito en áreas rurales, y culminó con la captura de varios funcionarios públicos y contratistas.
- **Escándalo de Reficar:** Es uno de los escándalos de corrupción más grandes en la historia de Colombia. Reficar, una filial de la petrolera estatal colombiana, malversó miles de millones de dólares en la remodelación de una refinería de petróleo.
- **Caso Odebrecht:** Odebrecht, una constructora brasileña, pagó millones de dólares en sobornos a políticos y funcionarios colombianos para obtener contratos de infraestructura. El escándalo estalló en 2016 y llevó a la condena de numerosos políticos y funcionarios de alto nivel en Colombia.

- **Carrusel de la Contratación:** Este escándalo surgió en Bogotá en 2010, cuando se descubrió que políticos y contratistas estaban inflando de manera ilícita el precio de los contratos de obras públicas para enriquecerse. Las pérdidas para la ciudad se estiman en cientos de millones de dólares.

Una vez seleccionados los escándalos, se realiza una búsqueda minuciosa de noticias relacionadas con cada uno de ellos. Esta búsqueda no se limita a encontrar la mera mención de los hechos, sino que se busca una cobertura completa y detallada del escándalo. De entre las numerosas noticias que pueden surgir de esta búsqueda, se seleccionan las 15 más relevantes, con preferencia por aquellas publicadas por medios de comunicación reconocidos.

Con las noticias en mano, se procede a descargar su contenido textual por medio de la librería BeautifulSoup de Python. Este texto, en su forma más cruda y sin modificaciones ni formatos adicionales, será el material que se someterá al procesamiento de un modelo de lenguaje llamado GPT-3.5.

Este modelo de lenguaje procesa los textos de las noticias y realiza dos tareas cruciales. En primer lugar, identifica y clasifica las entidades presentes en el texto, ya sean organizaciones o individuos. En segundo lugar, el modelo busca relaciones entre estas entidades. Así, al final de este proceso, se obtiene una lista de entidades y una serie de relaciones entre ellas, todas extraídas directamente de las noticias seleccionadas.

Sin embargo, antes de proceder con el análisis de los datos, es necesario llevar a cabo una fase de limpieza y depuración. En esta etapa, recurrimos a la librería spaCy de Python, para eliminar todas las entidades que no correspondan a individuos. SpaCy emplea un componente de reconocimiento de entidades nombradas que se basa en transiciones y que identifica rangos de tokens etiquetados sin superposiciones. Aunque este algoritmo podría no ser idóneo para todas las tareas, debido a que presupone que la información más crucial acerca de las entidades se hallará cerca de sus primeros tokens y optimiza la precisión total de la entidad, lo que puede ser ineficaz si las entidades son extensas o si existe desacuerdo en cuanto a sus límites. Posteriormente, se eliminan las relaciones duplicadas. Este paso garantiza que los datos que se emplearán para el análisis sean únicos y relevantes.

Las relaciones seleccionadas se clasifican en cinco categorías distintas: relacionesDeInfluencia, relacionesDeCorrupción, relacionesDeInvestigación, relacionesFamiliares, y relacionesProfesionales. Cada categoría representa un tipo de conexión específica entre los individuos o entidades dentro del conjunto de datos.

1. Relaciones de Influencia: Corresponden a la transferencia de recursos y a la colaboración. Se sustenta en la idea de que "La transferencia o intercambio de recursos (si es recíproco) es el componente principal de numerosas formas de actividad delictiva organizada" Bright et al. (2015) y en que "Los vínculos en esta dimensión pueden definirse ampliamente como interacciones intencionales, que involucran comunicación que puede ser tanto directa (cara a cara) como indirecta (por ejemplo, llamadas telefónicas, correos electrónicos)" Krebs (2002).

2. Relaciones de Corrupción: Se asocian a la transferencia de recursos, basándose en la afirmación de que "En nuestro caso, los recursos principales transferidos son dinero en forma de sobornos o coimas para políticos y contratos manipulados para empresarios" Bright et al. (2015).

3. Relaciones de Investigación: Aunque no se mencionan explícitamente en la literatura citada, podrían ser consideradas como parte de la colaboración, ya que este tipo de relación requiere un intercambio de información y cooperación para llevar a cabo una investigación. Esto se basa en la descripción de que "Los vínculos en esta dimensión pueden definirse ampliamente como interacciones intencionales, que involucran comunicación que puede ser tanto directa (cara a cara) como indirecta (por ejemplo, llamadas telefónicas, correos electrónicos)" Krebs (2002).

4. Relaciones Familiares: Se encuentran dentro de la categoría de vínculos preexistentes, como se evidencia en la declaración de que "Los vínculos preexistentes, es decir, los vínculos establecidos antes del propio acto delictivo, pueden ser fuentes cruciales de confianza... los vínculos preexistentes pueden tomar la forma de relaciones de parentesco" Morselli and Roy (2008).

5. Relaciones Profesionales: También están categorizadas como vínculos preexistentes, de acuerdo con la afirmación de que "Los vínculos preexistentes incluyen la membresía mutua en un consejo de administración de una empresa" Morselli and Roy (2008).

Cada uno de estos pasos se repite para cada uno de los escándalos de corrupción seleccionados. Así, al final de esta metodología, se obtiene una vista la red de escándalos de corrupción en Colombia, lista para ser analizada y estudiada.

## 2.1. Herramientas utilizadas

### 2.1.1. Medidas de red

Para analizar las estructuras y patrones de las redes de corrupción que surgen de los datos, se emplean diversas métricas de red. Dichas métricas facilitan la comprensión de la estructura, la cohesión y la distribución de la red, además de identificar los nodos más influyentes.

1. Grado: Esta métrica de centralidad cuenta el número de enlaces de un nodo con otros nodos en la red. Se puede expresar como:

$$k_i = \sum_j A_{ij}$$

donde  $A_{ij}$  es la entrada de la matriz de adyacencia que conecta los nodos  $i$  y  $j$ . Newman (2010).

2. Intermediación (Betweenness): Cuantifica la cantidad de veces que un nodo actúa como un "puente" en el camino más corto entre dos otros nodos. Se calcula como:

$$C_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

donde  $\sigma(s,t)$  es el número total de caminos más cortos de  $s$  a  $t$  y  $\sigma(s,t|v)$  es el número de esos caminos que pasan por  $v$ . Brandes (2001).

3. Clustering: Indica la tendencia de los nodos a agruparse. Un coeficiente de agrupamiento elevado puede señalar subredes densamente conectadas.

$$C_i = \frac{2E_i}{k_i(k_i - 1)}$$

donde  $E_i$  es el número de enlaces entre los vecinos del nodo  $i$ . Watts and Strogatz (1998).

4. Closeness: Se refiere a la proximidad promedio de un nodo a todos los demás nodos.

$$C(x) = \frac{1}{\sum_y d(y, x)}$$

donde  $d(y, x)$  es la distancia más corta de  $y$  a  $x$ . Bavelas (1950)

5. Distancia Geodésica Promedio: Es la longitud promedio del camino más corto entre cada par de nodos.

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}$$

donde  $d_{ij}$  es la distancia más corta entre los nodos  $i$  y  $j$ . Newman (2003).

6. Modularidad: Evalúa la estructura de la red dividiéndola en módulos o comunidades.

$$Q = \frac{1}{2m} \sum_{vw} [A_{vw} - \frac{k_v k_w}{2m}] \delta(c_v, c_w)$$

donde  $A_{vw}$  son los elementos de la matriz de adyacencia,  $k_v$  y  $k_w$  son los grados de los nodos  $v$  y  $w$ ,  $m$  es el total de los enlaces,  $c_v$  y  $c_w$  son las comunidades de los nodos y  $\delta(c_v, c_w)$  es 1 si  $v$  y  $w$  son de la misma comunidad y 0 si no. Newman (2006).

### 2.1.2. Conteo de frecuencias

Las frecuencias y las probabilidades están relacionadas de manera directa. La probabilidad se puede calcular a partir de las frecuencias de los eventos. Si se tiene un conjunto de eventos, cada uno con su propia frecuencia, se puede convertir esas frecuencias en probabilidades. La probabilidad de un evento específico es simplemente la frecuencia de ese evento dividida por la suma de las frecuencias de todos los eventos.

Matemáticamente, si se tiene un evento  $E$  con frecuencia de  $f(E)$  en  $N$  experimentos totales, entonces la probabilidad  $P(E)$  del evento  $E$  se calcula de la siguiente manera:

$$P(E) = \frac{f(E)}{N}$$

### 2.1.3. Teoría de la información y factor sorpresa

La entropía de Shannon es una medida cuantitativa de la incertidumbre en una distribución de probabilidades, se utiliza para cuantificar la incertidumbre asociada a una variable aleatoria:

$$H(x) = - \sum P(x) \log P(x)$$

El factor de sorpresa es un concepto relacionado que cuantifica cuán sorprendente o inesperado es un evento particular, dado lo que se sabe. Varley (2023). Un evento que es muy probable de acuerdo con el modelo o conocimiento previo tiene un factor de sorpresa bajo, mientras que un evento que es muy poco probable tiene un factor de sorpresa alto. Este factor sorpresa se define como:

$$h(x) = \log \left( \frac{1}{P(x)} \right) \quad (1)$$

### 2.1.4. Asignación de probabilidades a la distribución de grado y la distribución de betweenness

Una vez obtenida una distribución de grado en donde según los grados de los nodos se contabilizan cuántos nodos con este valor en la red están presentes, es posible transformar esta distribución en una función de probabilidades; esta función de probabilidades dice qué probabilidad tendría un nodo de presentar un grado. ??

$$P(k) = \frac{\alpha}{f_{max} k^\beta} \quad (2)$$

### 2.1.5. Probabilidades conjuntas, condicionales y Teorema de Bayes

Las probabilidades conjuntas hacen referencia a la probabilidad de que dos o más eventos ocurran simultáneamente. Si hay dos eventos  $A$  y  $B$  independientes, la probabilidad conjunta de que ocurran ambos eventos se denota como  $P(A \cap B)$ , y se calcula multiplicando las probabilidades individuales de cada evento:

$$P(A \cap B) = P(A)P(B)$$

Si los eventos son dependientes, hace falta definir la probabilidad condicional para encontrar la probabilidad conjunta.

Las probabilidades condicionales se utilizan para calcular la probabilidad de que ocurra un evento dado que otro evento ya ha ocurrido. Se denotan como  $P(A|B)$ , que es la probabilidad de que ocurra el evento  $A$  dado que el evento  $B$  ha ocurrido. La expresión para calcular la probabilidad condicional es:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Esta ecuación indica que la probabilidad condicional de  $A$  dado  $B$  es igual a la probabilidad conjunta de  $A$  y  $B$  dividida por la probabilidad de  $B$ . Canavos (1994).

El teorema de Bayes es una herramienta que permite actualizar los conocimientos previos sobre un evento dado cuando se obtiene nueva información relevante. Se utiliza para calcular las probabilidades condicionales en sentido inverso. La ecuación del teorema de Bayes es:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3)$$

### 2.1.6. Redes multiplexas

Las redes multiplexas son aquellas en las cuales sus nodos interactúan a través de distintas capas de relaciones. Hay una multitud de sistemas que es posible modelarles como redes multiplexas, por ejemplo las neuronas se conectan entre ellas por medio de sinapsis química y sinapsis eléctrica CITA. Para modelar este tipo de redes existen varios enfoques, en este trabajo se utilizará el enfoque multilíneal.

Se consideran a las redes multiplexas compuestas por múltiples capas como muchas redes simples, las cuales se denotan red de la capa  $\alpha$  y se representan por medio de  $G_\alpha$ , cada una de estas redes está compuesta por una cantidad  $N$  de nodos y  $M_\alpha$  enlaces. Descrita la red multiplexa de esta forma es posible definir una generalización multilíneal de la matriz de adyacencia Cozzo et al. (2018), a la cual se denotan sus componentes como  $a_{ij}^\alpha$  y en el caso de redes multiplexas no pesadas tomará valor de 1 si hay un vínculo entre los nodos  $i, j$  en la capa  $\alpha$ . El modelo de red multiplexa utilizado en este trabajo se ejemplifica en la siguiente imagen:

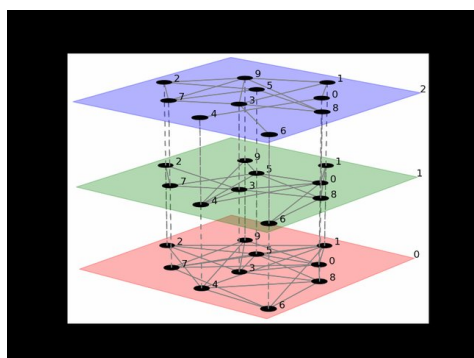


Figure 1: Ejemplo de red multiplexa

Las redes multiplex son útiles en el análisis de redes de corrupción porque proporcionan una visión más realista y detallada de la realidad social. Permiten diferenciar entre diferentes tipos de relaciones e intercambios entre actores, que es esencial para entender las actividades delictivas y cómo las redes desempeñan un papel en su organización. La multiplicidad de relaciones puede ser un factor importante en entornos ilegales donde compensa la falta de instituciones formales. Un análisis simplex de redes agregadas puede llevar a una sobre o subestimación de la prominencia de actores individuales. Un actor puede ser central en una dimensión pero periférico en otras, lo que incorrectamente lo hace parecer poco importante en la red simplex agregada general. Diviák et al. (2019).

## 3. Resultados y Discusión

Una vez completada la red mediante la filtración de los artículos proporcionados y la limpieza de algunos datos repetidos 2, se observó que en el grafo generado en networkx, los vínculos se encontraban diferenciados por el nombre que se les asignaba, sin embargo, estos nombres no representaban categorías que tuviesen implicaciones prácticas, sino que solo se tenían diferenciaciones cualitativas de las conexiones, por lo

que mediante el uso de los nombres se crearon diferentes tipos de categorías para de esta forma hacer el grafo multiplexo.



Figure 2: Grafo de la red de los seis escándalos de corrupción estudiados

Este grafo multiplexo se encuentra conformado por 191 nodos y 312 enlaces, lo cual corresponde a una densidad muy baja ( $\rho = 0.0085$ ), también, su forma a simple vista parece tener características muy propias de las redes de corrupción para otros países reportadas en Martins et al. (2022), además, el promedio de los agentes de la red tienen entre 3 y 4 conexiones ( $K_{average} = 3.26$ ), por otro lado el grafo multiplexo tiene 6 tipos de enlaces, los cuales son: **Relación de Influencia**, **Relación de Corrupción**, **Relación Familiar**, **Relación de Investigación**, **Relación familiar** y **Relación Desconocida**.

Asumiendo parcialmente que los tipos de relación no influyen en las medidas clásicas de centralidad, es posible por un momento asumir la red con todas sus conexiones en una sola categoría, de esta manera se obtuvieron varias medidas de centralidad las cuales dan luz acerca de la estructura de la red y la importancia de los actores, como medidas tenemos el grado 6, la intermediación 3, la cercanía 4 entre otras:

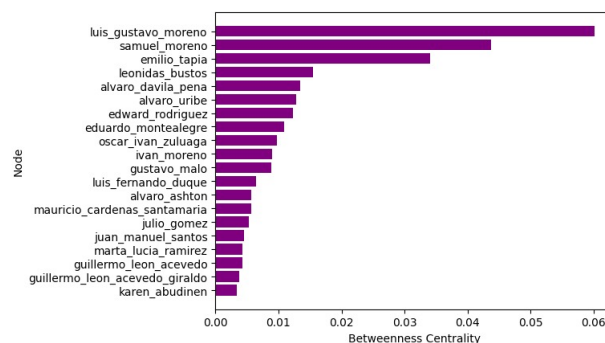


Figure 3: Niveles de intermediación para diferentes agentes de la red

Con el fin de extraer información de las redes multiplexas se realizó un conteo de la cantidad de enlaces entre nodos según su tipo de relación, para posteriormente hablar sobre qué relaciones son más frecuentes que otras y también para asignar probabilidades en caso de que se quiera construir una red con

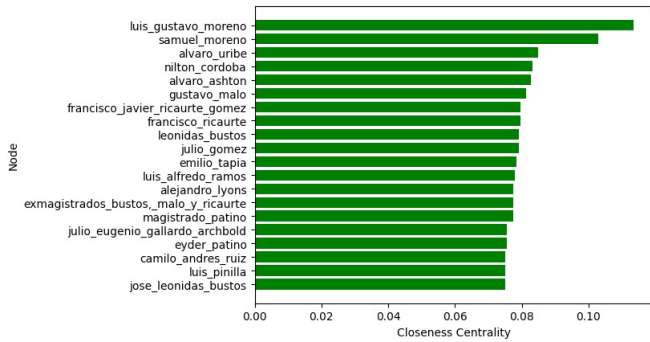


Figure 4: Niveles de cercanía para diferentes agentes de la red

esta misma frecuencia de aparición de enlaces multiplexos, es decir se creó una distribución que muestra con qué probabilidad se encontraría un tipo de relación si se toma un enlace al azar en la red.

A partir de estas probabilidades fue posible calcular la entropía para cada tipo de relación y en general la entropía para la distribución de probabilidades. (2.109 bits) Luego, a partir de esta distribución de probabilidades también fue posible calcular el factor sorpresa, 1 ; el cual se puede ver en la gráfica 5, como se puede evidenciar la relación que más se tiende a repetir en la red es la relación familiar, seguida de la relación de corrupción y la profesional. A partir de la observación de esta gráfica es posible conocer qué tipo de relación es la que más se da a esperar, también sugeriría a qué tipo de relaciones en la política se le deberían de prestar mucha más atención para evitar la creación de una red robusta que haga más difícil el desmantelamiento de agentes que incurran en actividades ilícitas.

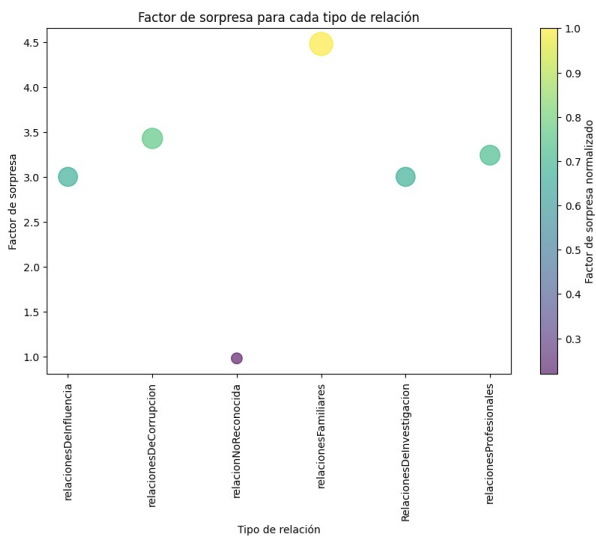


Figure 5: Factor sorpresa para la distribución de probabilidad de los tipos de relaciones en la red

Para obtener la distribución de grado se contó la cantidad de vecinos para cada nodo de la red para después crear un histograma que diera información acerca cuántas veces se repetía el escenario en el que un nodo de la red tuviese cierto grado,

y como era de esperar, este comportamiento parecía dirigido por una ley de potencias, en donde muchos nodos tienen pocos vecinos y muy pocos nodos tienen muchos vecinos. 6

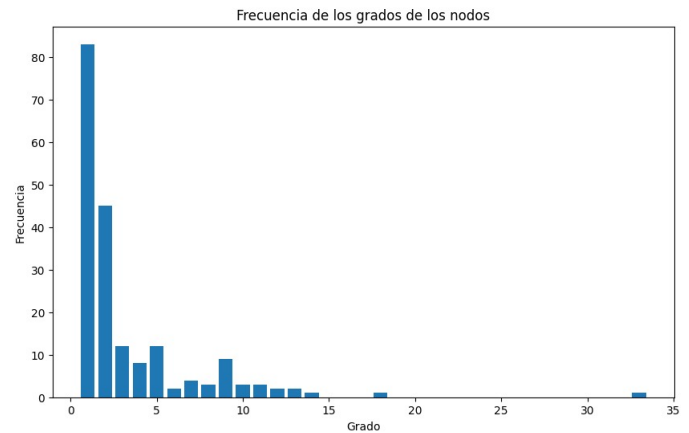


Figure 6: Histograma de cantidad de nodos con cierto grado

Aproximadamente, a partir de los valores de las frecuencias en relación a los grados es posible tomar estos puntos para ajustar una ley de potencias y encontrar una función definida entre  $[1, \infty)$  (1 - grado mínimo que puede exhibir un nodo) que nos hable de la probabilidad de encontrar un nodo con cierto grado en la red  $P(K)$ . 7 Aunque claramente, si se quiere una función de probabilidad exacta, lo primero es obtener una función de densidad de probabilidad, no normalizando el ajuste obtenido por medio de los datos de las frecuencias sino dividiendo cada frecuencia sobre el total de veces en las que se presentaba un nodo con cierto grado en la red (suma de todas las frecuencias), de esta forma se obtiene una función de densidad de probabilidad discreta que al integrarla toda nos daría 1, y qué también bajo ciertas consideraciones y un ajuste haríamos continua para posteriormente ser integrada sobre un intervalo para obtener una función de probabilidad exacta; sin embargo, por simplicidad y falta de tiempo, se toma una forma aproximada la cual no difiere mucho en los resultados y conserva la idea de lo que se quiere hacer para extraer la información de las redes multiplexas.

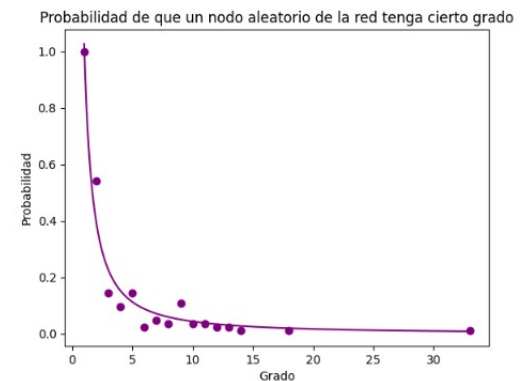


Figure 7: Función de probabilidad aproximada  $P(k)$

Debido a que se tiene una función de probabilidad y una dis-

tribución de probabilidades, es pertinente hacerse la pregunta de qué probabilidad hay de que un nodo aleatorio tenga cierto grado o que su grado se encuentre sobre una región debido a que este está conectado a algún nodo con un tipo de relación en específico ( $P(K = 5|r = 'Familiar')$ ); o incluso seguir haciéndose preguntas que se basen en factores condicionales y que de esta forma sea posible conocer cómo parte de la estructura de grados o de intermediaciones de la red se relaciona con la estructura de sus tipos de relaciones multiplexas.

Entonces, surge la pregunta, ¿de alguna manera el grado de un nodo guarda relación con el tipo de conexiones que pueda tener?, y para responder esta pregunta es posible recurrir al teorema de Bayes 3 en donde dado el caso, si ambas probabilidades son independientes, el resultado de la probabilidad condicional no se verá influenciada por la relación que existe y sería igual a la probabilidad de solo una parte:

$$\begin{aligned} P(K = 5|r = 'Familiar') &= \frac{P(K = 5, r = 'Familiar')}{P(r = 'Familiar')} \quad (4) \\ &= \frac{P(K = 5)P(r = 'Familiar')}{P(r = 'Familiar')} \\ &= P(K = 5) \end{aligned}$$

Para verificar si esto es así o no, y poder concluir si existe una interdependencia entre los valores de centralidad y los tipos de relaciones que se dan en esta red, se creó una distribución de probabilidad conjunta mediante un método de Monte Carlo, de esta manera, siguiendo la distribución de probabilidad de los tipos de relaciones multiplexas y las frecuencias a las que aparecían los nodos según su grado, fue posible construir un generador de valores aleatorios para de esta forma tener un espacio conjunto con puntos generados según las respectivas distribuciones para posteriormente asignarle probabilidades. 8

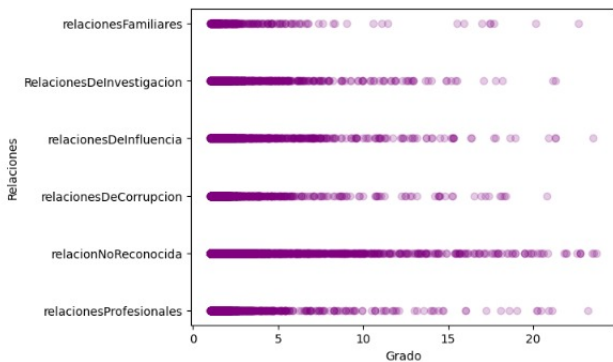


Figure 8: Espacio conjunto de variables aleatorias para generar probabilidades conjuntas

Una vez obtenido el espacio conjunto fue posible calcular las probabilidades conjuntas mediante el conteo de puntos en relación a la cantidad de puntos totales generados en el espacio, de esta forma, para probar que existe relación entre el grado y los tipos de vínculos se decidió calcular  $P(K = 2|r = 'Influencia')$ , lo cual dio como resultado 60.29%, un valor muy

diferente de lo obtenido si se hubiese asumido que estas probabilidades eran independientes 39.63%. Este hallazgo es muy importante, ya que permite conocer más a fondo la estructura de esta red.

Esto último tiene grandes implicaciones en el caso de poseer una red de corrupción mucho más extensa que abarque escándalos políticos más allá de los seleccionados para este trabajo, ya que si se piensa en que de alguna forma se puede predecir la evolución de la red basándose en el tipo de conexiones que podrían surgir mediante el conocimiento de estas frecuencias propias de la estructura, sería posible también conocer información acerca de cómo se podrían generar alternativas de combate y prevención en la corrupción, como lo sería el estudiar antes de tiempo una estructura de relaciones que puedan llegar a maximizar el tipo de vínculos que generen un escenario en donde se esté dando la corrupción; sin embargo, más estudios se tienen que hacer en torno a las características estructurales que tienen los escenarios políticos, en especial la corrupción, y de qué otras formas se pueden extraer de maneras más precisas y dicientes conclusiones que nos provean de luces en este aspecto.

#### 4. Conclusiones

En este estudio, se ha explorado la estructura de una red de corrupción política, utilizando técnicas de análisis de redes y estadísticas para extraer información significativa sobre las relaciones entre los actores y la naturaleza de sus interacciones. A través de la investigación, se ha descubierto que la red es altamente compleja y multiplexa, con diferentes tipos de relaciones que se entrelazan y se superponen de maneras que no son inmediatamente evidentes a simple vista.

Se ha encontrado que las relaciones familiares y de corrupción son las más prevalentes en la red, lo que sugiere que los lazos familiares y las actividades corruptas están profundamente arraigados en la estructura de la red. Sin embargo, también se ha observado que la red tiene una densidad muy baja, lo que indica que la mayoría de los actores en la red tienen solo unas pocas conexiones. Esto sugiere que, aunque la corrupción puede ser generalizada, está concentrada en un pequeño número de actores clave.

Al analizar las medidas de centralidad de la red, se ha descubierto que hay una pequeña cantidad de nodos que tienen un alto grado de intermediación y cercanía, lo que indica que estos actores pueden tener un papel importante en la facilitación de la corrupción en la red. Estos actores pueden ser objetivos clave para las intervenciones destinadas a dismantelar la red de corrupción.

Además, se han utilizado técnicas de simulación de Monte Carlo para explorar la relación entre el grado de un nodo y el tipo de relaciones que tiene. Los resultados sugieren que existe una interdependencia entre estas dos características, lo que indica que la estructura de la red y la naturaleza de las relaciones entre los actores no son independientes. Este hallazgo tiene implicaciones importantes para la comprensión de cómo se forma y evoluciona la red de corrupción.



El análisis de la red de corrupción ha revelado una serie de características clave que podrían ser útiles para informar las estrategias de prevención y combate a la corrupción. Sin embargo, también se han identificado varias áreas que requieren más investigación, incluyendo la necesidad de estudios más detallados sobre las características estructurales de las redes de corrupción y la búsqueda de métodos más precisos para extraer información de estas redes. A pesar de estos desafíos, se cree que este trabajo proporciona un punto de partida valioso para futuras investigaciones en este campo.

## References

- Bank, W., 2021. Combating corruption. Disponible en Internet a través de. URL: <https://n9.c1/5d4fu>. accessed on August 12, 2022.
- Bavelas, A., 1950. Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America* 22, 725–730.
- Brandes, U., 2001. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology* 25, 163–177.
- Bright, D., Hughes, C., Chalmers, J., 2015. Illuminating dark networks: A social network analysis of an Australian drug trafficking syndicate. *Crime, Law and Social Change* 63, 203–223.
- Canavos, 1994. Probabilidad y Estadística - Aplicaciones y Metodos (Spanish Edition). McGraw-Hill Companies. URL: <https://n9.c1/1inw8>.
- Cozzo, E., De Arruda, G.F., Rodrigues, F.A., Moreno, Y., 2018. Multiplex networks: basic formalism and structural properties. Springer.
- Diviák, T., Dijkstra, J.K., Snijders, T.A., 2019. Structure, multiplexity, and centrality in a corruption network: the czech rath affair. *Trends in Organized Crime* 22, 274–297.
- Diviák, T., Dijkstra, J.K., Snijders, T.A.B., 2019. Structure, multiplexity, and centrality in a corruption network: the czech rath affair. *Trends in Organized Crime* 22, 274–297. URL: <https://doi.org/10.1007/s12117-018-9334-y>, doi:10.1007/s12117-018-9334-y.
- D’Orsogna, M.R., Perc, M., 2015. Statistical physics of crime: A review. *Physics of Life Reviews* 12, 1–21. URL: <https://n9.c1/3yik5>, doi:<https://doi.org/10.1016/j.plrev.2014.11.001>.
- Gómez-Gardeñes, J., Villarreal, S., Romance, M., Criado, R., 2020. Corruption and complexity: a scientific framework for the analysis of corruption networks. *Applied Network Science* 5, 1–22.
- Gonzalez, M.A., et al., 2018. The impact of corruption on healthcare outcomes in latin america. *Revista de Economía e Investigación* 15, 180–197. URL: [http://www.scielo.org.co/scielo.php?script=sci\\_rtttextpid=S0121-47052018000100180](http://www.scielo.org.co/scielo.php?script=sci_rtttextpid=S0121-47052018000100180).Granados, O.M., Nicolás Carlock, J.R., 2021. *Corruption networks : Concepts and applications*.
- Internacional, T., 2017. Índice de percepción de corrupción 2016. Disponible en Internet a través de. URL: <https://n9.c1/wy8mb>.
- Kramer, J.M., 1977. Political corruption in the u. s. s. r. *The Western Political Quarterly* 30, 213–224.
- Krebs, V., 2002. Mapping networks of terrorist cells. *Connections* 24, 43–52.
- Mahler-Hutter, M., 2009. Modeling corruption. na.
- Martins, A.F., da Cunha, B.R., Hanley, Q.S., Gonçalves, S., Perc, M., Ribeiro, H.V., 2022. Universality of political corruption networks. *Scientific Reports* 12, 6858.
- Morselli, C., Roy, J., 2008. Brokerage qualifications in ringing operations. *Criminology* 46, 71–98.
- de la Nación, P.G., 2023. Transparencia. Disponible en Internet a través de. URL: <https://n9.c1/e371u>. accessed on June 22, 2023.
- Newman, M., 2010. *Networks: An Introduction*. Oxford University Press.
- Newman, M.E.J., 2003. The structure and function of complex networks. *SIAM Review* 45, 167–256.
- Newman, M.E.J., 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 103, 8577–8582.
- Newman, V., Ángel, M.P., 2017. Sobre la corrupción en Colombia: marco conceptual, diagnóstico y propuestas de política. Number 56 in Cuadernos de Fedesarrollo, Fedesarrollo, Bogotá.
- Pública, F., 2023. Funciones generales. Disponible en Internet a través de. URL: <https://n9.c1/1n6gk>. accessed on June 22, 2023.
- Ribeiro, H.V., Alves, L.G.A., Martins, A.F., Lenzi, E.K., Perc, M., 2018. The dynamical structure of political corruption networks. *Journal of Complex Networks* 6, 989–1003. URL: <https://doi.org/10.1093/comnet/cny002>, doi:10.1093/comnet/cny002.
- Rose-Ackerman, S., 1999. Close but no cigar: The measurement of corruption. *Journal of Public Policy* 19, 139–167.
- Tanzi, V., 1998. Corruption Around the World: Causes, Consequences, Scope and Cures. Working Paper WP/98/63. International Monetary Fund.
- Varley, T.F., 2023. Information theory for complex systems scientists. *arXiv:2304.12482*.
- Wadhwa, S., Amir, S., Wallace, B.C., 2023. Revisiting relation extraction in the era of large language models *arXiv:2305.05003*.
- Wang, Y., 2020. A network-exchange approach to corruption: Brokers and institution spanning in a chinese corruption network. *Deviant Behavior* 41, 1636–1649.
- Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of ‘small-world’ networks. *Nature* 393, 440–442.