

Análisis de redes organizacionales con herramientas multicriterio y análisis de componentes principales

Enrique Álvarez Beltrán¹

Centro de Investigación en Matemáticas (CIMAT), Unidad Aguascalientes

Humberto Martínez Bautista

Centro de Investigación en Matemáticas (CIMAT), Unidad Aguascalientes

Bey Jamelyd López Torres

Universidad Autónoma de Zacatecas

RESUMEN

En la actualidad muchas organizaciones han adoptado el análisis de datos para respaldar su toma de decisiones, buscando con ello mejorar su competitividad y operación. El análisis de redes sociales (ARS) constituye una opción eficiente para analizar la estructura de una red organizacional usando indicadores representativos de las relaciones internas de una empresa. La complejidad de integración e interpretación de los indicadores de ARS ha limitado su utilización en las organizaciones empresariales, por tanto, el objetivo en este trabajo es construir una medida integral que combine los indicadores de ARS considerando dos enfoques: análisis multicriterio (TOPSIS) y análisis de componentes principales (PCA). Analizamos la estructura organizacional del Centro de Investigación en Matemáticas (unidad Aguascalientes) utilizando el enfoque de ARS. Identificamos los actores con mayor influencia en dicha organización usando TOPSIS y PCA. Los resultados mostraron que las dos técnicas de integración son adecuadas para el análisis de redes organizacionales (ARO). Sin embargo, TOPSIS mostró mayor eficiencia computacional y facilidad de ordenación comparada con PCA. Se concluye que los indicadores de ARS fueron suficientes para caracterizar la estructura de la red organizacional. El trabajo desarrollado facilita la adopción de métricas de análisis estructural en organizaciones de diversa índole.

Palabras clave: Análisis de Redes Organizacionales (ARO), Análisis de Redes Sociales (ARS), PCA, TOPSIS.

ABSTRACT

Currently many organizations have adopted data analysis to support their decision-making seeking to improve their competitiveness and operation. Social network analysis (SNA) is an efficient option to analyze the structure of an organizational network using representative indicators of a company's internal relationships. The complexity of integration and interpretation of SNA indicators has limited their use in business organizations, therefore, the objective in this work is to build a comprehensive measure that combines SNA indicators, considering two approaches: multicriteria analysis (TOPSIS) and principal components analysis (PCA). We analyzed the structure of the Mathematics Research Center (Aguascalientes) using the SNA approach. We identified the actors with the greatest influence in the said organization using TOPSIS and PCA. The results showed that the two integration techniques were suitable for organizational network analysis (ONA). However, TOPSIS showed higher computational efficiency and order facility compared to PCA. It is concluded that the SNA indicators are sufficient to characterize the structure of the organizational network. The work developed in this research facilitates the adoption of structural analysis metrics in organizations of various types.

Key words: Organizational Network Analysis (ONA), PCA, Social Network Analysis (SNA), TOPSIS.

¹ Contacto con los autores: Enrique Álvarez Beltrán (enrique.alvarez@cimat.mx), Humberto Martínez Bautista (humberto.martinez@cimat.mx), Bey Jamelyd López Torres (bey.jamelyd@uaz.edu.mx).

INTRODUCCIÓN

El análisis de redes sociales (ARS) es un paradigma de investigación utilizado para estudiar a los sujetos y sus relaciones dentro de una red. En años recientes, el ARS se ha popularizado debido a su amplio campo de aplicación, que se ha extendido más allá de las ciencias sociales y ha incursionado en disciplinas como la física, epidemiología y biología (Borgatti y Halgin, 2011; McHale y Relton, 2018; Rovere, 1999), incluyendo su versatilidad para analizar fenómenos que van desde la agricultura hasta la política (López y Ibarra, 2021).

Con la ayuda del ARS es posible comprender fenómenos complejos usando un enfoque de red. Una red está conformada por actores y sus relaciones. Los actores pueden representar individuos, instituciones, empresas, etc; las relaciones permiten conocer la circunstancia por la que dos actores están conectados, que puede ser un lazo de tipo familiar, de amistad, de colaboración o laboral, por mencionar algunos ejemplos.

Estos patrones de relacionamiento dan forma a lo que se conoce como estructuras, a partir de las cuales es posible obtener indicadores fundamentales de la red y con ellos obtener conocimiento del fenómeno bajo estudio. Los fundamentos de la teoría de redes sociales y su análisis tienen su base en las matemáticas, específicamente en la teoría de grafos (Hansen et al., 2020).

Representando un sistema complejo a través de una red, podemos visualizarlo mejor y observar las interconexiones entre los distintos actores. En una red social, es posible establecer un orden jerárquico entre los nodos o actores usando sus atributos, que son características mediante las cuales se puede explicar el comportamiento de un conjunto de actores, o bien, de la red en su totalidad (Bonacich, 1987). De la misma forma que cada actor posee atributos, los vínculos entre ellos también los tienen, estos representan la fuerza e intensidad de su interacción, o bien, el tipo de relación entre dos actores (Hanneman y Riddle, 2005).

Borgatti (2013) en su investigación "El problema del actor clave" aborda dos problemas que denominó PAC-1 y PAC-2. PAC-1 consiste en identificar un conjunto de actores cuya eliminación perturbaría al máximo la comunicación en la red y, PAC-2 consiste en encontrar los actores que pueden llegar a la mayor cantidad de nodos en la red con el objetivo de dispersar de manera óptima una información. En dichas circunstancias propone identificar actores clave utilizando los algoritmos conocidos como harvest, diffuse y disrupt en Key Player;

que abordan específicamente el alcance y fragmentación en la red, sin embargo, no consideran otros atributos de los actores como su centralidad y popularidad, por ejemplo.

En ocasiones las organizaciones tienen poca comprensión de cómo los equipos y las personas se asocian dentro de la organización, este conocimiento es relevante ya que puede ser utilizado por ejemplo para comprender las estructuras de colaboración que generan mejores resultados comerciales (Krackhardt y Hanson, 1993). Actualmente, los tomadores de decisiones tienen la oportunidad de conocer mejor la estructura de sus organizaciones, cómo colaboran y de qué manera pueden ser más eficientes por medio del ARS.

Si bien el estudio de fenómenos mediante el ARS se ha extendido, una de las desventajas es que las herramientas actuales para su aplicación se enfocan principalmente en el cálculo de indicadores, y, el análisis de éstos puede ser problemático para los tomadores de decisiones que se encuentran con indicadores aislados, por lo que su interpretación suele parecerles complicada. Debido a esto, dichas herramientas son menos eficientes cuando se trata del apoyo para la toma de decisiones dentro de una organización.

Dada la dificultad de los tomadores de decisiones para identificar a los actores clave de una red en el proceso de análisis, el objetivo en este trabajo es investigar la oportunidad de dos métodos: uno multicriterio y el otro estadístico, para construir una medida integral que combine los indicadores de ARS con la finalidad de identificar actores influyentes de la red y simplificar la toma de decisiones en las organizaciones.

Análisis de Redes Organizacionales (ARO)

Una organización se define como una unidad social o agrupación humana deliberadamente construida para alcanzar objetivos específicos (Krackhardt y Hanson, 1993). Las personas que integran una organización desarrollan actividades comunes con el objetivo de alcanzar el éxito de la organización. Como en todo fenómeno social los individuos interactúan entre sí buscando un beneficio laboral o personal. Cuando los individuos dentro de una organización se representan mediante nodos y sus interacciones mediante conexiones con otros nodos, entonces se genera una red organizacional.

El ARO es un método a través del cual es posible visualizar y analizar las conexiones formales e informales que se generan de manera natural y espontánea entre los integrantes de una organización (Nurek y Michalski, 2020). Dicha metodología de análisis permite: 1. Que las

organizaciones identifiquen el tipo de relaciones de cada colaborador y cómo estos en conjunto forman conexiones dentro de la red. 2. Comprender las relaciones entre los actores de una red, esto se vuelve fundamental para entender cómo se distribuye, viaja y retiene el conocimiento dentro de las organizaciones y, 3. Conocer aquellos factores que impactan positiva o negativamente el negocio.

Al representar a una organización como una red, donde los agentes están conectados entre sí por relaciones interpersonales, se puede analizar cómo circula el flujo de información dentro del organismo de la empresa y de esta forma descubrir propiedades invisibles a simple vista (Liu y Moskvina, 2015). Además, las medidas de centralidad son capaces de detectar las posiciones con mayor influencia en la red y, por lo tanto, proporcionan un análisis cuantitativo del poder en las organizaciones.

El ARO cuenta con dos enfoques distintos para la recolección de información acerca de la red a estudiar (Adams, 2020). El primero es el análisis activo, en este el investigador utiliza herramientas como sondeos o encuestas a través de las cuales los colaboradores identifican y reconocen a aquellos individuos con los que interactúan de manera recurrente ya sea por una razón de tipo laboral o personal. El segundo enfoque para la recolección de información es el análisis pasivo el cual consiste en realizar un sondeo más profundo de la organización y sus interacciones a través de herramientas como el correo electrónico, calendarios, plataformas de la empresa, eventos de la compañía, medios sociales internos y externos, así como bases de datos existentes en internet.

METODOLOGÍA

Para recolectar la información sobre las relaciones de los actores en la red dentro de la organización, se optó por aplicar el método de *red completa* (Hanneman y Riddle, 2005). Es decir, se consideraron en el estudio a todos los sujetos de la población.

La etapa de recolección de datos del estudio siguió un enfoque de análisis activo de Adams (2020), es decir, a los sujetos de estudio se les proporcionó una encuesta en la cual indicaron con quienes habían tenido algún tipo de interacción tanto de tipo profesional dentro de la organización como de tipo personal fuera de esta.

Para construir la propuesta, fue empleado un caso de estudio en el Centro de Investigación en Matemáticas A.C. (CIMAT) Unidad Aguascalientes, éste es un centro público del Consejo Nacional de Humanidades, Ciencias y

Tecnologías (Conahcyt) en México, que desarrolla actividades de formación de recursos humanos, investigación, vinculación y divulgación relacionadas con las áreas de la matemática. Se utilizó un diseño no experimental, con temporalidad transversal (año 2022) y con alcance descriptivo.

Este estudio incluyó a profesores y personal administrativo del programa de posgrado de la Maestría en Modelación y Optimización de Procesos. La población de dicha institución al momento de la investigación estaba conformada por 19 integrantes (profesores y personal administrativo). Se diseñó una encuesta *exprofeso* para su aplicación dentro del CIMAT Unidad Aguascalientes.

Las redes analizadas en esta investigación se consideraron de tipo dirigidas. Es decir, la interacción del actor v con el actor u no supone la interacción del actor u con el actor v , a menos que esta segunda interacción esté especificada. Otra consideración es que aquellos nodos aislados, es decir, sin conexiones de entrada o salida no fueron considerados para este análisis.

A continuación, se describen los indicadores que fueron empleados para el estudio de la red.

Centralidad de grado: Es el número de actores adyacentes a un nodo, es decir, a cuántos actores está conectado. En el ARS dos nodos son adyacentes cuando su interacción se representa mediante un vínculo (Freeman, 1978), debido a esto se considera un indicador de primer orden. Para el caso de redes no dirigidas este parámetro es únicamente un valor, sin embargo, en redes dirigidas es necesario distinguir entre grados de entrada y grados de salida (Hanneman y Riddle, 2011). Este parámetro representa la exposición que tiene un nodo o actor al interior de la red, es decir, un actor con un grado alto puede tener mayor probabilidad de ser alcanzado por otros actores en la red.

Cercanía: Es un indicador de la independencia de un nodo con respecto a los demás nodos que conforman una red. Un nodo es distante cuando necesita de varios nodos para alcanzar a otros más, por otro lado, un nodo es cercano cuando está en contacto con otros y necesita de pocos nodos para llegar a otros puntos de la red. Un nodo cercano tiene una mayor centralidad de cercanía que un nodo distante (Freeman, 1978), se interpreta como el tiempo mínimo que tomaría a un nodo alcanzar el recurso que está fluyendo dentro de la red (Aguilar, et al., 2017). Así un nodo con un valor alto de cercanía tardaría menos tiempo ya que estaría más cercano a la red, caso contrario al de un nodo con un valor bajo de este indicador, el cual necesita de otros nodos ya que se encuentra lejano al recurso y, por lo tanto, tardará más tiempo en alcanzarlo.

En conclusión, mientras más cercano esté un nodo a la red, más fácil éste puede alcanzar información u otro recurso y, además, su valor de centralidad es mayor. El cálculo de este indicador está dado por:

$$C(u) = \frac{n-1}{N-1} \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)}$$

La cercanía se normaliza mediante el factor $\frac{n-1}{N-1}$ donde n es el número de nodos en el componente conectado que incluye a u y N es el número total de nodos en la red. Mientras que $d(v,u)$ es el camino más corto entre los nodos v y u . Finalmente, $n-1$ representa el número de nodos que pueden llegar a u . La cercanía de un nodo va de 0 a 1. Los valores cercanos a 0 indican que un nodo dado está lejos de otros nodos en la red (se deben atravesar muchos enlaces para llegar de ese nodo a otros nodos en la red). Para el caso contrario, un valor cercano a 1 indica que un nodo dado está cerca de otros nodos en la red (se deben atravesar pocos enlaces para llegar de ese nodo a otros nodos en la red) (Goldstein y Vitevitch, 2017).

Eigenvector: Este indicador fue introducido por Bonacich (1987). El eigenvector es una medida de la popularidad de un actor dentro de una red. Si bien la popularidad de un actor puede ser medida simplemente mediante el número de vínculos que tiene, el eigenvector representa un índice de segundo orden al considerar y ponderar los vínculos de los actores a los que el nodo original está conectado. Por lo tanto, el eigenvector se puede interpretar como una medida de la popularidad de un actor en el sentido de que un nodo con alto valor del indicador se encuentra relacionado con nodos que a su vez están bien conectados (Borgatti, et al., 2013). Este indicador, al tomar en cuenta relaciones indirectas, ofrece más información con respecto a la calidad de los vínculos de un nodo en comparación con el parámetro de centralidad de grado. Puede que existan nodos con centralidad de grado bajas, pero que estén conectados a nodos con grados altos y, por tanto, su eigenvector será alto; al contrario, puede ser que haya actores altamente conectados (grados elevados) pero que éstos no están conectados a nadie, por tanto, su eigenvector será bajo (Aguilar et al., 2017). Este indicador se calcula obteniendo la centralidad de un nodo en base a la centralidad de sus vecinos.

El eigenvector del nodo i es el i -ésimo elemento del vector x definido por la ecuación:

$$Ax = \lambda x$$

donde A es la matriz de adyacencia del grafo G con valor propio λ , cuya solución única es dada

por el teorema de Perron-Frobenius (Bapat, 2014).

Diversidad: Una red consiste en un conjunto de individuos y sus interacciones. Las interacciones entre estos pueden representar un gran número de relaciones, entre estas se encuentran la amistad o comunicación informal entre dos actores, o bien, que existe un intercambio de algún tipo, ya sea de información o material. Los lazos de amistad pueden clasificarse como internos o externos (López y Ibarra, 2021). El grupo de actores internos es aquel integrado por actores con características similares a las del actor v , por otro lado, en el grupo de actores externos se encuentran todos aquellos actores con características distintas a las del actor v . La diversidad es un índice que se calcula restando el número de lazos internos del grupo al número de sus lazos externos y dividiéndolo entre el número total de lazos (Krackhardt y Stern, 1988). La fórmula de cálculo del índice de diversidad es la siguiente, donde D_i : diversidad del nodo i , L_E : lazos externos, L_I : lazos internos y T_L : Total de lazos (López y Ibarra, 2021).

$$D_i = \frac{L_E - L_I}{L_T}$$

Como se observa en la fórmula, este indicador solamente puede tomar valores entre -1 y $+1$. Cuando la diversidad toma valores cercanos a $+1$ entonces el nodo i presenta mayor diversidad en sus relaciones dentro de la red, al presentar más relaciones con el grupo externo en comparación al grupo interno. Mientras que si la diversidad toma valores cercanos a -1 entonces el nodo analizado presenta más relaciones con actores similares a él que con el grupo externo. Es importante mencionar que de acuerdo con Krackhardt y Stern (1988) en tiempos de crisis, las organizaciones más efectivas serán aquellas con lazos de amistad entre los grupos, a diferencia de lazos de amistad dentro de los grupos.

A continuación, se describen las dos metodologías de análisis. La primera fue el Análisis de Componente Principales (PCA) y la segunda es una combinación de dos técnicas de análisis multicriterio: el Proceso de Análisis Jerárquico (AHP) y la Técnica de Orden de Preferencia por Semejanza a la Solución Ideal (TOPSIS). Ambos enfoques han sido utilizados en la literatura como auxiliares del ARS en la determinación de los actores más importantes de una red (Du, et al., 2014; Fox y Everton, 2013; Sam, et al., 2021).

Análisis de Componentes Principales (PCA)

El PCA es uno de los métodos multivariados más comúnmente empleados en la investigación de

datos multidimensionales y altamente correlacionadas, cuyo fin es tomar p variables X_1, X_2, \dots, X_p , y encontrar combinaciones de estos para producir índices Z_1, Z_2, \dots, Z_p que no están correlacionados y, en orden de importancia en términos de la variación en los datos. La falta de correlación significa que los índices están midiendo diferentes dimensiones del fenómeno, y el orden es tal que $Var(Z_1) \geq \dots \geq Var(Z_p)$, donde $Var(Z_i)$ denota la varianza de Z_i . Los índices Z_i son entonces los componentes principales (Manly y Navarro, 2017).

Los pasos en un análisis de componentes principales son:

1. Se estandarizan las variables X_1, X_2, \dots, X_p , para que tengan medias cero y varianzas unitarias.
2. Calcular la matriz de covarianza o correlación.
3. Obtener los valores propios $\lambda_1, \lambda_2, \dots, \lambda_p$ y los vectores propios correspondientes a_1, a_2, \dots, a_p .
4. Descartar cualquier componente que solo represente una pequeña proporción de la variación en los datos. Por ejemplo, comenzando con 20 variables, se podría encontrar que los primeros tres componentes explican el 90% de la varianza total. Sobre esta base, los otros 17 componentes pueden ignorarse.
5. Realizar la combinación lineal usando los vectores propios y trabajar con la o las componentes seleccionadas.

Canali, et al. (2010) proponen una metodología cuantitativa con PCA que busca apoyar el ARS para identificar actores relevantes aprovechando el potencial de PCA para visualización y, por ende, comprensión del problema. En PCA las primeras componentes, suponiendo que Z_1 y Z_2 , contienen la mayor cantidad de información acerca del fenómeno analizado, como una consecuencia de esta situación es posible emplear únicamente las primeras dos dimensiones. Dada esta ventaja de usar PCA es factible analizar las relaciones entre las características de distintos actores de la red mediante gráficas de dispersión de los primeros dos componentes principales Z_1 y Z_2 . Finalmente, mediante el análisis de estas gráficas es posible obtener los nodos más influyentes de la red.

Proceso de Análisis Jerárquico (AHP)

Es uno de los métodos de decisión multicriterio más utilizados debido a que mediante su

aplicación es posible seleccionar una opción a partir de un conjunto de alternativas en función de una combinación de criterios o variables de selección, cuya característica principal es que son conflictivos entre sí (Zoran et al., 2011). Se dice que es un método jerárquico ya que para la aplicación del método se requiere el planteamiento de un árbol dividido en niveles. Este árbol presenta una configuración vertical con orientación de arriba a abajo empezando por el objetivo final, seguido de los criterios con sus subcriterios y finalmente las alternativas a comparar. Uno de los pasos cruciales para el funcionamiento adecuado de este método es elegir y definir correctamente los criterios y subcriterios de selección, asegurándose que sean mutuamente excluyentes.

En la metodología AHP se requiere realizar comparaciones pareadas entre los criterios y los subcriterios considerados en el problema a analizar (Kulakowski, 2020). Estas comparaciones se realizan mediante la Escala Fundamental de Saaty (Saaty, 1987). Esta escala es una de las razones principales del éxito de este método de análisis multicriterio, dado que a través de esta es posible transformar aspectos cualitativos en cuantitativos, lo que permite la comparación entre las distintas alternativas. Una vez obtenidas las matrices de comparaciones pareadas el siguiente paso es obtener los pesos relativos para cada uno de los criterios y subcriterios.

El método AHP pertenece al grupo de métodos populares por su análisis de la consistencia en el proceso de comparación de los elementos de la jerarquía. Considerando que las comparaciones pareadas se basan en una estimación subjetiva del tomador de decisiones, es necesario que este parámetro sea monitoreado constantemente para asegurar la precisión requerida. Si la razón de consistencia (C.R.) es menor a 0.10, se dice que mediante la matriz de comparaciones pareadas se pueden calcular los pesos relativos sin mayores complicaciones (Zoran et al., 2011), para el caso contrario cuando $C.R. > 0.10$, se procede a revisar la matriz de comparaciones pareadas antes de seguir con el procedimiento.

Técnica de Orden de Preferencia por Semejanza a la Solución Ideal (TOPSIS)

TOPSIS es un método de análisis multicriterio, es utilizado para encontrar la mejor alternativa, para lo que este método primero identifica la mejor y la peor solución de acuerdo con el contexto del problema, posteriormente busca aquella alternativa que presente la distancia geométrica más corta desde la solución ideal positiva y la distancia geométrica más larga desde la solución ideal negativa (Hwang y Yoon, 1981). Para lograr esto el método requiere que

el tomador de decisiones establezca un peso para cada criterio para posteriormente normalizar las puntuaciones por criterio y calcular la distancia geométrica entre cada alternativa y la alternativa ideal, que es la mejor puntuación en cada criterio (Hamurcu y Eren, 2018). Una suposición de TOPSIS es que los criterios aumentan o disminuyen monótonamente. Por lo general, se recomienda la normalización ya que los parámetros o criterios a menudo son de dimensiones incompatibles en problemas multicriterio. Los métodos compensatorios, como es el caso de TOPSIS, permiten hacer concesiones entre criterios. Esto quiere decir que un mal resultado en un criterio de alguna de las alternativas analizadas puede ser solventado por un buen resultado en algún otro criterio. Esta característica es una de las principales ventajas y diferencias al comparar estos métodos con los métodos no compensatorios, en los cuales existen cortes estrictos para incluir o excluir alternativas.

El proceso TOPSIS se lleva a cabo de la siguiente manera:

1. Crear una matriz de evaluación, digamos D , que consta de n alternativas y p criterios, con la intersección de cada alternativa y criterio dada como x_{ij} , dando como resultado una matriz $(X_{ij})_{n \times p}$.

$$D = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

2. La matriz D es normalizada para formar la matriz $R = (R_{ij})_{n \times p}$, utilizando el método de normalización distributiva siguiente:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}}$$

para $i = 1, 2, 3, \dots, n$ y $j = 1, 2, 3, \dots, p$.

3. Calcular la matriz normalizada ponderada de decisión. Primero se requiere establecer los pesos. Estos pueden venir ya sea del tomador de decisión o por cómputo.
4. Utilizar los pesos del tomador de decisiones para los atributos x_1, x_2, \dots, x_p o calcular los pesos mediante el uso de comparaciones pareadas mediante la escala fundamental de Saaty (Saaty, 1987) a través de la metodología AHP para obtener los pesos como el vector propio de la matriz de

comparación por pares de atributos contra atributos.

5. Multiplicar los pesos de cada una de las entradas de la columna en la matriz del Paso 2 para obtener la matriz T .

$$T = (t_{ij})_{n \times p} = (w_i r_{ij})_{n \times p}, i = 1, 2, \dots, n$$

6. Determinar la peor alternativa (A_w) y la mejor alternativa (A_b): Examinar la columna de cada atributo y seleccionar el mejor valor y el peor. Si el objetivo es maximizar (beneficio) entonces las mejores alternativas son los valores más grandes, por otro lado, si el objetivo es minimizar (el costo) entonces la mejor alternativa es el valor más pequeño.
7. Calcular la distancia entre la alternativa objetivo i y la peor condición A_w :

$$d_{iw} = \sqrt{\sum_{j=1}^n (t_{ij} - t_{wj})^2}, i = 1, 2, \dots, n$$

Calcular la distancia entre la alternativa objetivo i y la mejor condición A_b :

$$d_{ib} = \sqrt{\sum_{j=1}^n (t_{ij} - t_{bj})^2}, i = 1, 2, \dots, n$$

donde d_{iw} y d_{ib} son las distancias euclidianas desde la alternativa objetivo i hacia la mejor y peor condición, respectivamente.

8. Calcular el coeficiente de cercanía de cada alternativa:

$$s_{iw} = \frac{d_{ib}}{d_{iw} + d_{ib}}, 0 \leq s_{iw} \leq 1, i = 1, 2, \dots, n$$

El indicador s_{iw} tiene un valor próximo a 1 cuando la alternativa se acerca a la solución ideal, si está más cercana a la solución no ideal se aproxima a 0.

RESULTADOS Y DISCUSIÓN

En nuestro trabajo empleamos un conjunto suficiente de características o indicadores ARS que componen a un actor clave en el análisis de redes organizacionales (ARO): en primer lugar, ser cercanos a la red (por lo que se utilizan los indicadores de centralidad de grado y cercanía); en segundo lugar, se prioriza a actores que tengan buen posicionamiento por tener relaciones con actores a su vez bien posicionados, lo que se conoce como una medida

de popularidad (indicador de *eigenvector*); y por último, la diversidad de las relaciones que el actor posee, lo que le da influencia en grupos con características y recursos distintos al que pertenece, lo que representa una ventaja en comparación con trabajos como el de Borgatti (2013). Este enfoque considera que los actores clave tienen roles y tipos de relaciones diversas con docentes, personal administrativo y alumnos, lo que les proporciona distinta influencia.

Con la finalidad de ilustrar la comparación entre el análisis PCA y TOPSIS fue realizada una encuesta a los docentes y personal administrativo del Centro de Investigación en Matemáticas, Unidad Aguascalientes, con la cual, se generó una red a partir de la pregunta: *En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?* Dicha red representa el flujo del intercambio de información entre el equipo de docentes y personal administrativo. Un total de 19 personas forman esta red y todas fueron encuestadas, es decir se aplicó un método de red completa para la recolección de datos

Tabla 1

Indicadores por nodo de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

	Salida	Entrada	Diversidad	Eigenvector	Cercanía
A-1	18	15	0.091	0.282	1.000
A-2	13	9	0.000	0.188	0.783
A-3	10	17	0.111	0.315	0.692
A-4	13	8	0.048	0.155	0.783
A-5	11	15	0.154	0.289	0.720
A-6	9	6	-0.467	0.115	0.667
A-7	5	13	-0.111	0.251	0.563
A-8	12	16	0.071	0.301	0.750
A-9	4	7	0.455	0.135	0.563
B-1	13	14	0.704	0.281	0.783
B-2	7	9	0.500	0.187	0.621
B-3	11	12	0.652	0.241	0.720
C-1	17	13	1.000	0.242	0.947
D-1	16	9	1.000	0.186	0.900
E-1	17	14	0.677	0.273	0.947
E-2	7	10	0.529	0.207	0.621
E-3	13	9	0.636	0.186	0.783
E-4	1	6	0.714	0.122	0.419
F-1	18	13	1.000	0.247	1.000

usando un cuestionario estructurado y exprofeso para esta investigación. Las encuestas se aplicaron a través del correo institucional de los empleados usando la herramienta *Google Forms*. Es importante mencionar que se garantiza la confidencialidad de los datos y por lo tanto en ningún momento se exponen los nombres de los actores involucrados. Para garantizar la confidencialidad de los datos a cada encuestado se le asignó un identificador único de la forma X-00, donde X representa su puesto o rol dentro de la red seguido de un número consecutivo único.

Cálculo de indicadores de la red

Para cada nodo en la red se obtuvieron 5 indicadores (Tabla 1): centralidad de grado de salida (X_1) y de entrada (X_2), diversidad (X_3), *eigenvector* (X_4) y cercanía (X_5). A partir de estos 5 indicadores o dimensiones se llevó a cabo tanto un análisis mediante la metodología PCA y TOPSIS. Los resultados obtenidos mediante ambos enfoques, así como sus ventajas y desventajas se discuten a continuación.

Esta matriz de 5 dimensiones se analizó usando PCA con la intención de reducir las dimensiones de los datos y conocer si las interacciones dentro de la red podían ser explicadas mediante una o dos dimensiones con el objetivo de determinar los nodos más influyentes de la red. Se aplicó un PCA a través de la paquetería *scikit-learn de Python*.

La Tabla 2 muestra los datos estandarizados. Este proceso es importante ya que en lugar de trabajar con datos heterogéneos se trabaja con un conjunto de datos con media cero y varianza unitaria.

Tabla 2

Indicadores por nodo estandarizados de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

	Salida	Entrada	Diversidad	Eigenvector	Cercanía
A-1	1.402	1.098	-0.782	1.001	1.612
A-2	0.353	-0.690	-1.005	-0.545	0.208
A-3	-0.276	1.693	-0.732	1.543	-0.376
A-4	0.353	-0.988	-0.888	-1.089	0.208
A-5	-0.066	1.098	-0.627	1.123	-0.197
A-6	-0.486	-1.584	-2.153	-1.756	-0.541
A-7	-1.324	0.502	-1.279	0.484	-1.214
A-8	0.143	1.395	-0.830	1.321	-0.003
A-9	-1.534	-1.286	0.113	-1.427	-1.214
B-1	0.353	0.800	0.726	0.977	0.208
B-2	-0.905	-0.690	0.225	-0.565	-0.838
B-3	-0.066	0.204	0.599	0.323	-0.197
C-1	1.192	0.502	1.455	0.336	1.272
D-1	0.982	-0.690	1.455	-0.579	0.966
E-1	1.192	0.800	0.661	0.855	1.272
E-2	-0.905	-0.392	0.297	-0.232	-0.838
E-3	0.353	-0.690	0.560	-0.574	0.208
E-4	-2.163	-1.584	0.752	-1.627	-2.144
F-1	1.402	0.502	1.455	0.428	1.612

Una vez estandarizados los datos se aplicó la metodología PCA. Los resultados del análisis se muestran en la Tabla 3. El componente 1 explica el 58.9% de la variación total, mientras que el componente 2 el 25.8% y el componente 3 el 14.9%. Un criterio aceptado es utilizar aquellos componentes que sumen más del 80% de la variación total. Para nuestro caso los primeros dos componentes explican el 84.7%

de la variación total, y por tanto, fueron elegidos sólo dos componentes para construir la medida. Cabe resaltar que la idea central de la aplicación del análisis PCA en esta investigación no fue la reducción de dimensiones, si no la visualización de los datos en dos dimensiones mediante un diagrama de dispersión como se muestra en la Figura 2.

Tabla 3

Resultados del análisis PCA de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

	PC1	PC2	PC3	PC4	PC5
A-1	-2.404	0.405	1.185	0.170	-0.074
A-2	0.476	-0.146	1.295	-0.146	-0.069
A-3	-1.111	2.132	-0.462	0.022	0.100
A-4	0.862	-0.597	1.417	-0.055	0.081
A-5	-0.835	1.475	-0.154	-0.072	0.011
A-6	2.479	0.122	2.144	0.028	0.049
A-7	1.015	2.069	-0.167	0.088	-0.069
A-8	-1.244	1.687	0.030	-0.039	0.066
A-9	2.676	-0.377	-0.415	0.234	-0.020
B-1	-1.258	0.184	-0.806	-0.150	-0.024
B-2	1.447	-0.119	-0.501	-0.014	-0.080
B-3	-0.219	-0.030	-0.696	-0.123	0.005
C-1	-1.884	-1.321	-0.449	0.105	0.108
D-1	-0.609	-2.106	-0.118	-0.070	-0.012
E-1	-2.148	-0.476	-0.035	0.027	-0.054
E-2	1.132	0.124	-0.715	-0.025	-0.099
E-3	0.234	-1.098	0.078	-0.142	0.008
E-4	3.597	-0.459	-1.356	0.016	0.074
F-1	-2.206	-1.467	-0.276	0.145	0.000
%Var exp	0.589	0.259	0.149	0.002	0.001

En la Tabla 4 se muestran los coeficientes obtenidos para cada uno de los cinco componentes principales usando las variables X_1, X_2, X_3, X_4 y X_5 estandarizadas. Por otro lado, en la Figura 1 se muestra un mapa de calor ilustrativo de la influencia de cada coeficiente por componente principal. A partir de este se puede concluir, del color de cada variable, que los coeficientes del primer componente principal son en su mayoría negativos. Situación similar para el segundo componente principal.

El primer componente principal (PC_1), ver Tabla 4, está dado por:

$$PC_1 = -0.506X_1 - 0.481X_2 - 0.163X_3 - 0.483X_4 - 0.503X_5$$

Los coeficientes de las variables X_1, X_2, X_4 y X_5 son prácticamente iguales y representan una ponderación negativa de los indicadores

centralidad de entrada, salida, eigenvector y cercanía de un actor, por otro lado, el indicador diversidad tiene una menor influencia en este componente principal.

El segundo componente principal (PC_2), ver Tabla 4, está dado por:

$$PC_2 = -0.324X_1 + 0.463X_2 - 0.600X_3 + 0.446X_4 - 0.350X_5$$

Este representa un contraste entre las variables X_2 y X_4 con las variables X_1, X_3 y X_5 , como se puede observar por sus signos. Es decir, PC_2 tomará un valor grande cuando X_2 y X_4 son valores positivos grandes, por otro lado, este componente tendrá un valor bajo cuando X_1, X_3 y X_5 sean valores grandes en combinación con valores pequeños de X_2 y X_4 . En colusión PC_2 representa las diferencias en los indicadores de los actores de la red.

Tabla 4

Coefficientes de cada componente obtenidos mediante el PCA, a partir de la matriz de datos normalizada, de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

		PC ₁	PC ₂	PC ₃	PC ₄	PC ₅
Salida	X ₁	-0.506	-0.324	0.376	-0.591	0.385
Entrada	X ₂	-0.481	0.463	-0.228	0.396	0.588
Diversidad	X ₃	-0.163	-0.600	-0.782	-0.004	0.039
Eigenvector	X ₄	-0.483	0.446	-0.268	-0.386	-0.590
Cercanía	X ₅	-0.503	-0.350	0.350	0.587	-0.396

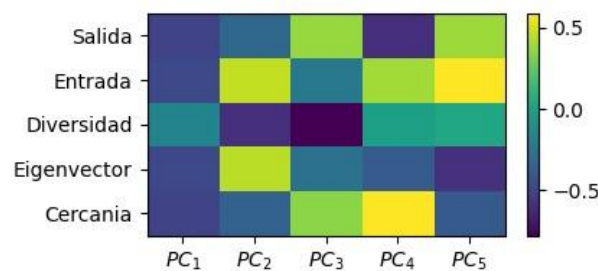


Figura 1

Mapa de calor representativo de los coeficientes de los componentes principales obtenidos de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

Una vez analizados los datos de los indicadores de cada nodo mediante PCA se procedió a determinar los nodos más influyentes en la red a partir de los resultados obtenidos. El primer paso fue graficar el componente principal (PC1) contra componente principal (PC2) como se observa en la Figura 2. Los nodos F-1, C-1, E-

1, B-1 y A-1 son los más influyentes de la red. Para identificación de dichos nodos nos guiamos del análisis PCA (Tabla 4 y Figura 1): actores con valores negativos del componente 1 y valores negativos o cercanos a cero del componente 2 (Figura 2).

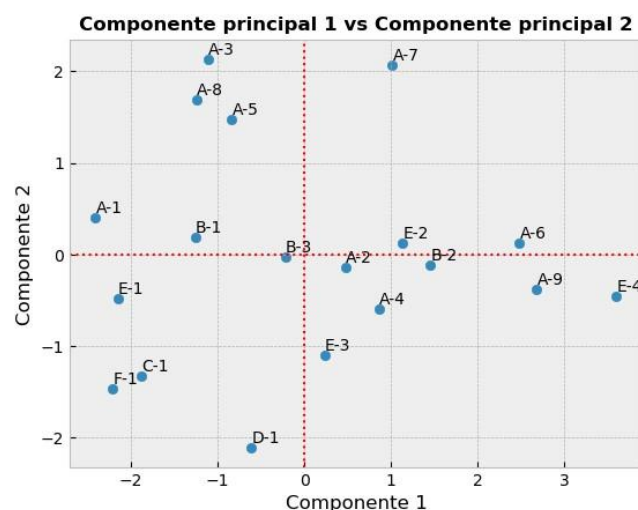


Figura 2

Componente principal 1 contra componente principal 2 de los resultados obtenidos del análisis PCA a partir de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

De la misma manera, se llevó a cabo un análisis de los nodos de la red a través de sus indicadores mediante la metodología TOPSIS. Esta metodología requiere un peso por criterio, en este caso por indicador, por lo que se propuso una matriz de comparaciones pareadas de los criterios mediante la escala fundamental de Saaty (Saaty, 1987) para posteriormente obtener los pesos mediante la metodología AHP. La matriz de preferencias y los pesos por criterio se muestran en la Tabla 5. Para este caso el indicador C.R de la matriz

de preferencias fue 0.0981, menor a 0.10, es decir, a partir de la matriz de comparaciones pareadas se pueden calcular los pesos relativos.

Una vez obtenidos los pesos por cada criterio el siguiente paso fue emplear TOPSIS a la matriz de datos de la Tabla 1. Fue aplicada una normalización distributiva en donde cada valor es dividido por la raíz cuadrada de la suma de los cuadrados de los elementos del criterio evaluado.

Tabla 5

Matriz de preferencias mediante la escala fundamental de Saaty y pesos por criterio.

	Salida	Entrada	Diversidad	Eigenvector	Cercanía	Peso por criterio
Salida	1	1/3	1/3	1/5	1/4	0.055
Entrada	3	1	1/4	1/5	1/4	0.088
Diversidad	3	4	1	1/4	1/3	0.156
Eigenvector	5	5	4	1	3	0.450
Cercanía	4	4	3	1/3	1	0.251

Los resultados obtenidos a partir del análisis de los indicadores de la red mediante TOPSIS se muestran en la Tabla 6. En esta los nodos están ordenados del más influyente al menos influyente de acuerdo con los criterios y pesos por indicador de la matriz de la Tabla 5. Los 5 nodos mas influyentes de la red fueron: E-1, F-1, B-1, C-1 y B-3.

La red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo? (Figura 3) se analizó mediante PCA y TOPSIS. Para la primera metodología los nodos más influyentes fueron F-1, C-1, E-1, B-1 y A-1, mientras que para TOPSIS fueron E-1, F-1, B-1, C-1 y B-3. La coincidencia de identificación de actores clave fue del 80% (4/5) lo cual muestra la consistencia de los resultados obtenidos con ambos métodos. Es importante mencionar que una ventaja de TOPSIS es que los resultados son una clasificación ordenada de los nodos más influyentes de la red, mientras que a través de PCA solamente se determinó un grupo de actores importantes. Dicha ventaja en ordenación hace más oportuno el uso de TOPSIS.

Las desventajas de identificar los nodos más influyentes mediante PCA es que se requiere interpretar los coeficientes de las ecuaciones de los componentes principales y, que para visualización de los datos se requiere trabajar con uno o dos componentes principales, siempre que éstos expliquen un alto porcentaje (superior al 80%) de la varianza total de los datos. TOPSIS presenta como desventajas la dependencia de ponderaciones y que no permite la reducción de dimensionalidad del problema.

Ambas metodologías empleadas fueron adecuadas y consistentes, dependiendo del objetivo de análisis y la naturaleza de los datos podría preferirse alguna. Por ejemplo, en nuestro caso TOPSIS proporciona una clasificación ordenada de los actores más influyentes, sin necesidad de interpretar ecuaciones o gráficas como fue el caso de PCA. Este es un punto importante ya que cabe recordar que en esta investigación se busca proponer una metodología de ARS orientada a las organizaciones, por lo que evitar el paso de interpretación de coeficientes resulta de gran ayuda para el tomador de decisiones dentro de la organización. Nuestra propuesta proporciona una perspectiva robusta e integral en la identificación de actores clave en redes organizacionales.

Tabla 6

Nodos más influyentes obtenidos mediante el análisis TOPSIS de los indicadores de la red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo?

Rank	Actor	Cercanía	Diversidad	Eigenvector	Entrada	Salida
1	E-1	0.947	0.677	0.273	14	17
2	F-1	1.000	1.000	0.247	13	18
3	B-1	0.783	0.704	0.281	14	13
4	C-1	0.947	1.000	0.242	13	17
5	B-3	0.720	0.652	0.241	12	11
6	D-1	0.900	1.000	0.186	9	16
7	A-3	0.692	0.111	0.315	17	10
8	A-1	1.000	0.091	0.282	15	18
9	A-5	0.720	0.154	0.289	15	11
10	A-8	0.750	0.071	0.301	16	12
11	E-3	0.783	0.636	0.186	9	13
12	E-2	0.621	0.529	0.207	10	7
13	B-2	0.621	0.500	0.187	9	7
14	A-7	0.563	-0.111	0.251	13	5
15	E-4	0.419	0.714	0.122	6	1
16	A-2	0.783	0.000	0.188	9	13
17	A-9	0.563	0.455	0.135	7	4
18	A-4	0.783	0.048	0.155	8	13
19	A-6	0.667	-0.467	0.115	6	9

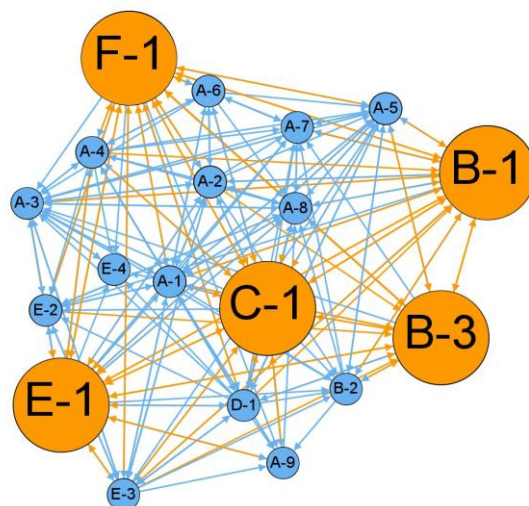


Figura 3

Red generada para la planta docente y personal administrativo a partir de la pregunta: En el último año, ¿Con quienes ha intercambiado información de temas relacionados a su trabajo? (nodos más influyentes de acuerdo con TOPSIS en color naranja).

CONCLUSIÓN

Usando PCA y TOPSIS en conjunto con los indicadores clásicos de nodos obtenidos mediante ARS logramos tener el alcance suficiente para analizar estos indicadores y a partir de ellos proponer un conjunto de nodos influyentes en la red. En esta investigación TOPSIS presentó como ventaja en comparación con el análisis PCA la ordenación de los nodos y con ello facilidad para identificar los más influyentes, en tanto, con el análisis PCA se obtiene un grupo de nodos influyentes de la red. Debido a esto la interpretación de resultados dentro de una organización se vuelve más asequible mediante el análisis TOPSIS. El análisis de los resultados obtenidos mediante el PCA se puede volver complejo cuando el tomador de decisiones no es experto en esta metodología.

Otra ventaja de la aplicación de TOPSIS con el objetivo de obtener los nodos más importantes de una red es que el tomador de decisiones puede asignar la importancia de cada variable a través una Matriz de preferencias mediante la escala fundamental de Saaty según se requiera en cada caso de aplicación.

En un mundo donde cada vez existen y se generan más datos el ARS se vuelve fundamental para la toma de decisiones dentro de una organización, primero porque elimina la posible subjetividad del tomador de decisiones y segundo porque es capaz de describir las relaciones sociales mediante indicadores a partir de los cuales se pueden hacer inferencias sobre una red, por ejemplo, que tan bien conectada está, identifica la existencia de grupos dentro de la organización, así como los actores más influyentes. El trabajo desarrollado en esta investigación facilitará la adopción de métricas adecuadas de análisis estructural en organizaciones de diversa índole.

Para que los tomadores de decisiones en las organizaciones puedan utilizar la propuesta aquí presentada se identifica un área de oportunidad para el desarrollo una interfaz gráfica a través de la cual sea posible obtener un análisis estadístico y gráfico de la red. Dicha herramienta deberá estar enfocada en ofrecer un análisis de tipo cuantitativo para los tomadores de decisiones en una organización de tipo empresarial, académico o social.

Reconocimientos. La fuente de financiación de EAB fue el Consejo Nacional de Humanidades, Ciencias y Tecnologías de México (Conahcyt). EAB y HMB realizaron la investigación de los métodos estadísticos y los cálculos. BJLT realizó la conceptualización teórica. EAB aplicó la encuesta. Todos participaron en el análisis de los resultados y las conclusiones, así como en la revisión y edición final del artículo. Humberto Martínez Bautista es autor de correspondencia.

REFERENCIAS

- Adams, J. (2020).** *Gathering Social Network Data*. SAGE Publications, Inc.
- Aguilar G., N., Martínez, E. G. y Aguilar-Ávila, J. (2017).** *Análisis de redes sociales: Conceptos clave y cálculo de indicadores*. México. Universidad Autónoma de Chapingo.
- Bapat, R. B. (2014).** *Graphs and matrices* (2.^a ed.). Springer-Verlag.
- Bonacich, P. (1987).** Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5), 1170–1182. <http://www.jstor.org/stable/2780000>
- Borgatti, S. P. (2013).** El problema del actor clave. *REDES Revista Hispana para el Análisis de Redes Sociales*. 24(2), 1-20. <https://doi.org/10.5565/rev/redes.482>
- Borgatti, S. P., Everett, M. G. y Johnson, J. C. (2013).** *Analyzing Social Networks*. 85. SAGE Publications Limited.
- Borgatti, S. P. y Halgin, D. S. (2011).** On network theory. *Organization Science*, 22(5), 1168–1181.
- Canali, C., Casolari, S. y Lancellotti, R. (2010).** A quantitative methodology to identify relevant users in social networks. En *2010 IEEE international workshop on: Business applications of social network analysis (BASNA)* (pp. 1–8). IEEE.
- Du, Y., Gao, C., Hu, Y., Mahadevan, S. y Deng, Y. (2014).** A new method of identifying influential nodes in complex networks based on TOPSIS. *Physica A: Statistical Mechanics and its Applications*, 399, 57–69. <https://doi.org/10.1016/j.physa.2013.12.031>
- Fox, W. y Everton, S. (2013).** Mathematical modeling in social network analysis: using TOPSIS to find node influences in a social network. *Journal of Mathematics and System Science*, 3(10), 531-541. <http://hdl.handle.net/10945/53057>
- Freeman, L. C. (1978).** Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.
- Goldstein, R. y Vitevitch, M. S. (2017).** The influence of closeness centrality on lexical processing. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01683>
- Hamurcu, M. y Eren, T. (2018).** An application of multicriteria decision-making for the evaluation of alternative monorail routes. *Mathematics*, 7(1), 16.
- Hanneman, R. A. y Riddle, M. (2005).** *Introduction to social network methods*.

University of California, Riverside, California
<http://faculty.ucr.edu/hanneman/>

Hanneman, R. A. y Riddle, M. (2011). Concepts and measures for basic network analysis. *The SAGE handbook of social network analysis*, 340-369.

Hansen, D. L., Shneiderman, B., Smith, M. A. y Himmelboim, I. (2020). *Analyzing social media networks with NodeXL: Insights from a connected world* (2.ª ed.). Morgan Kaufmann.

Hwang, C.-L. y Yoon, K. (1981). *Multiple Attribute Decision Making*. Springer.

Krackhardt, D. y Hanson, J. R. (1993). Informal networks: the company behind the charts. *Harvard Business Review*, 104-111.

Krackhardt, D. y Stern, R. N. (1988). Informal Networks and Organizational Crises: An Experimental Simulation. *Social Psychology Quarterly*, 51(2), 123.

Kulakowski, K. (2020). *Understanding the analytic hierarchy process*. CRC Press.

Liu, J. y Moskvina, A. (2015, August). Hierarchies, ties and power in organizational networks: model and analysis. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (pp. 202-209). <https://doi.org/10.1145/2808797.2808874>

López T., B. J. y Ibarra E., R. (2021). *Análisis de Redes Sociales: Aplicaciones en las Ciencias Sociales*. Taberna Librería Editores.

Manly, B. F. y Navarro Alberto, J. A. (2017). *Multivariate statistical methods: A primer* (4.ª ed.). CRC Press.

McHale, I. G. y Relton, S. D. (2018). Identifying key players in soccer teams using network analysis and pass difficulty. *European Journal of Operational Research*, 268(1), 339-347.

Nurek, M. y Michalski, R. (2020). Combining machine learning and social network analysis to reveal the organizational structures. *Applied Sciences*, 10(5), 1699. <https://doi.org/10.3390/app10051699>

Rovere, M. R. (1999). *Redes en salud: Un Nuevo Paradigma Para El Abordaje De Las Organizaciones Y La Comunidad*. Secretaría de Salud Pública Municipalidad de Rosario.

Saaty, R. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling*, 9(3-5), 161-176. [https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)

Sam, C., Naicker, N. y Adebisi, M. (2021). Dimensionality reduction of social media application attributes for ubiquitous learning using principal component analysis. *Mobile Information Systems*, 2021, 1-10. <https://doi.org/10.1155/2021/6633223>

Zoran, D., Saša, M., y Dragi, P. (2011). Application of the AHP Method for Selection of a Transportation System in Mine Planning. *Underground mining engineering*, 19, 93-99.

Remitido: 16-03-2024

Corregido: 25-06-2024

Aceptado: 11-10-2024



© Los autores