

Análisis de Redes Sociales mediante Diagramas Estratégicos y Diagramas Estructurales

Fco. Fernando de la Rosa Troyano, Universidad de Sevilla¹

Rafael Martínez Gasca, Universidad de Sevilla²

Luis González Abril, Universidad de Sevilla³

Francisco Velasco Morente, Universidad de Sevilla⁴

Abstract

The aim of the present work is to define a theoretical framework which facilitates the analysis of the structure of a network at an individual level. To this end, we propose the integration of co-word analysis together with the analysis of social networks. As a result of this study, a series of diagrams called "structural diagrams" has been obtained. Thanks to these diagrams a pattern can be assigned to each node of the network, and a network can be classified into a set of typologies. The information provided by this theoretical framework will allow a deeper understanding of the dynamics of systems, modeled in the form of networks. In this context, structural diagrams technique improves strategies of the visual exploration of the networks as well as to orientate the definition of those procedures which enable the transformation of one typology of network into another. As a specific example of a real application of this theoretical framework, the social network of the Journal of Software Engineering and Databases (JISBD) scientific community has been analyzed, based on its co-authorship networks.

Key Words: Social Network Analysis, Structural Analysis, Structural Diagrams, Co-word Analysis, Patterns, Network Typology, Visualization, Co-Authorities, Dynamic Systems Dynamics.

Resumen

El objetivo del presente trabajo es definir un marco teórico que permita analizar la estructura de una red a nivel individual, para ello se ha integrado el análisis de palabras asociadas con el análisis de redes sociales. Como resultado de este estudio, se han obtenido una serie de diagramas que llamamos "diagramas estructurales". Con estos diagramas asignamos a cada nodo de la red un patrón, así como clasificamos una red dentro de un conjunto de tipologías. La información proporcionada por este marco teórico mejora la comprensión de la dinámica de los sistemas modelados en forma de redes. Dentro de este contexto podemos utilizar la técnica de los diagramas estructurales para mejorar las estrategias de exploración visual de las redes, así como para orientar la definición de actuaciones que permitan transformar la tipología de una red en otra. Como ejemplo concreto de una aplicación real de este marco teórico, se analiza la red social de la comunidad científica de las Jornadas de Ingeniería del Software y Bases de Datos (JISBD), obtenida a partir de su red de co-autorías.

Palabras clave: Análisis de Redes Sociales, Análisis Estructural, Diagramas Estructurales, Análisis de Palabras Asociadas, Patrones, Tipologías de Red, Visualización, Co-autorías, Dinámica de Sistemas.

¹ Dpto. de Lenguajes y Sistemas Informáticos. Enviar correspondencia a ffrosat@lsi.us.es

² Dpto. de Lenguajes y Sistemas Informáticos. Enviar correspondencia a gasca@lsi.us.es

³ Dpto. de Economía Aplicada I. Enviar correspondencia a luisgon@us.es

⁴ Dpto. de Economía Aplicada I. Enviar correspondencia a velasco@us.es

Introducción

El análisis estructural de redes (Callon, Law & Rip, 1986; Callon, Courter & Laville, 1991; Coulter, Monarch & Konda, 1998; Monarch, 2000) sobre diagramas estratégicos nos va a permitir presentar un nuevo marco teórico, al cual nos referiremos como análisis mediante *"diagramas estructurales"*. En estos trabajos, se presentan los diagramas estratégicos como una herramienta para analizar la estructura de las temáticas de un área, a partir de su producción documental. Por otro lado, los diagramas estructurales son una herramienta de análisis que permite la obtención sistemática de una serie de patrones que nos facilitan el análisis de las estructuras de las redes sociales, así como conocer su evolución.

La principal diferencia entre diagramas estratégicos y estructurales reside en que mientras los primeros exploran la posición estratégica de un conjunto de agregados definidos sobre una red (**mesoanálisis**), los segundos permiten explorar la posición estratégica de los nodos de una red (**microanálisis**). A pesar de analizar diferentes niveles de granularidad de una red, ambos diagramas son complementarios, de forma que la información obtenida por los primeros enriquecen la información obtenida por los segundos y a la inversa. Como se verá, la formalización de los diagramas estructurales, se realiza modificando la semántica de los diagramas estratégicos. Para ello, se sustituyen las medidas utilizadas en estos últimos (índice de cohesión interna y externa), por otro conjunto de medidas (centralidad, grado, redundancia, etc) utilizadas en el campo del análisis de redes sociales. De esta forma caracterizaremos la posición estratégica de los nodos.

La abundancia de artículos científicos y su disponibilidad en Internet, tanto de los documentos como de sus referencias, ha contribuido a que una de las mayores redes susceptibles de ser analizadas sean las redes de co-autorías. Por medio de las redes de co-autorías, es posible analizar una de las estructuras más importantes que componen una comunidad científica, su red social. Esto es posible gracias a que la red de co-autoría, en gran medida, refleja las relaciones sociales que existen entre los miembros de una comunidad. Como ejemplo descriptivo del análisis mediante diagramas

estructurales presentamos un estudio aplicado, sobre la comunidad perteneciente a las Jornadas de Ingeniería del Software y Bases de Datos (JISBD). Para ello se ha extraído de Internet, de forma automática con la ayuda de crawlers, la información bibliográfica de la comunidad, publicada en la base de datos DBLP¹ (Digital Bibliograph & Library Project).

Aunque pueda parecer que la técnica de análisis que se presenta sólo es aplicable a redes sociales, lo cierto es, que tiene gran aplicabilidad en cualquier sistema que pueda ser modelado como una red o grafo. Estas redes estarían compuestas por nodos, que representarían entidades (autores, palabras, páginas web, artículos, clientes, empleados, empresas, productos, etc) y por aristas, que informarían de las relaciones existentes entre las distintas entidades (publicación conjunta o colaboración, referencia, amistad, subordinado, compra, etc). De esta manera, esta técnica puede ser de utilidad para analizar las relaciones que se producen entre los miembros de un foro de discusión o para analizar la estructura de una comunidad de páginas web. Con objeto de generalizar los resultados obtenidos, a lo largo de este trabajo, hablaremos indistintamente de los actores y los autores, considerando los autores como actores de una red de co-autorías. También se considerará las redes de co-autorías como un subtipo de red.

Teniendo en cuenta los aspectos anteriores, la estructura de este trabajo se compone de las siguientes secciones:

- una formalización de los conceptos que define el contexto teórico en que se enmarca este trabajo,
- la presentación de una primera aportación original, donde se utilizan los diagramas estratégicos para caracterizar las redes sociales,
- la presentación de una segunda aportación novedosa, que describe el análisis mediante diagramas estructurales, así como la forma de aplicarlo e interpretarlo,

¹ DBLP es una base de datos cuyos servidores proporcionan información bibliográfica sobre las revistas y congresos más importantes sobre Computer Science.

- y finalizaremos con las conclusiones y el Anexo I. Éste refleja, en forma de diagrama conceptual, los conceptos teóricos y las propuestas realizadas en este trabajo.

Marco Teórico

En esta sección se introduce la notación, las diferentes medidas del análisis de redes sociales utilizadas a lo largo de este trabajo y se concluye describiendo el análisis de palabras asociadas que permite la construcción de los diagramas estratégicos de una red de términos.

Notación

Dado un conjunto de nodos (autores), V , y un conjunto de aristas (relaciones de co-autorías), E , se considera el grafo $G = \langle V, E \rangle$. Sea por tanto G , un grafo conexo, acíclico, sin pesos y no dirigido, que representa una red social. Se definen los siguientes conceptos:

La **matriz de adyacencia**, A , como:

$$A_{ij} = \begin{cases} 1 & \text{si } (i, j) \in E \\ 0 & \text{si } (i, j) \notin E \end{cases}$$

donde $(i, j) \in E$ significa que existe una arista que enlaza los nodos i y j de V . De la definición se sigue que A es una matriz cuadrada de orden $|V|$, siendo $|V|$ el tamaño o cardinalidad del conjunto V .

El conjunto de **nodos vecinos al nodo i** , N_i , como:

$$N_i = \{v \in V \mid A_{iv} = 1\}$$

La **matriz de intermediación**, g y la **matriz tridimensional de intermediación**, g^* , como:

$$g_{ij} = \text{Número de caminos mínimos entre los nodos } i \text{ y } j$$

$$g_{ikj}^* = \text{Número de caminos mínimos entre los nodos } i \text{ y } j \text{ que pasan por el nodo } k$$

La **matriz de distancias geodésicas**, D , como:

$$D_{ij} = \text{Número mínimo de aristas necesarias para llegar del nodo } i \text{ al nodo } j$$

Claramente como el grafo G es no dirigido se sigue que las matrices A , g , g^* y D son simétricas.

El **grafo egocéntrico del nodo i incluyendo el ego** como $G_i^+ = \langle V_i^+, E_i^+ \rangle$, donde:

$$V_i^+ = N_i \cup \{i\}$$

$$E_i^+ = \{(v, w) \mid v \in N_i \wedge w \in N_i\} \cup \{(i, k) \mid k \in N_i\}$$

y el **grafo egocéntrico del nodo i excluyendo el ego** como $G_i^- = \langle V_i^-, E_i^- \rangle$, donde:

$$V_i^- = N_i$$

$$E_i^- = \{(v, w) \mid v \in N_i \wedge w \in N_i\}$$

También se considera un **conjunto de agregados** $H = \{H_1, \dots, H_m\}$ como un conjunto de subgrafos conexos del grafo G , donde el **número de agregados** se representa como $|H|$, el **número de nodos que componen el agregado H_i** como $|H_i|$, el **conjunto de aristas del agregado H_i** como $In(i)$ y el **conjunto de aristas que unen al agregado H_i con el resto de agregados** como $Out(i)$.

Medidas de Análisis de Red Social

Para llevar a cabo el análisis de una red social, también necesitamos utilizar medidas que nos van a permitir caracterizar los nodos que la componen. Estas medidas se dividen en dos grupos o familias: 1) La **familia de medidas globales**, que informan de la posición que ocupa un nodo en relación al resto de nodos. 2) La **familia de medidas egocéntricas o locales**, que caracterizan la posición de un determinado nodo, i , respecto al conjunto de nodos vecinos, N_i .

A continuación exponemos las medidas globales:

- La **lejanía (farness, f_i) del nodo i** : Es la suma de las distancias del nodo i al resto de nodos de la red (suma por filas o por columnas de la matriz de distancias geodésicas).

$$f_i = \sum_{j=1}^{|V|} D_{ij}, \quad \forall i \in V$$

- La **cercanía o centralidad (closeness o centrality, c_i) del nodo i** (Freeman, 1979) Mide la proximidad del nodo i al resto de nodos de la red (inversa de la lejanía).

$$c_i = \frac{1}{f_i}, \quad \forall i \in V$$

- La **autoridad o importancia (authority, a_i) del nodo i** (Page & Brin, 1999; Bonacich, 1972) Proporciona una medida de la importancia del nodo calculada a partir del algoritmo **Pagerank**, el cual calcula la autoridad del nodo i en función de la autoridad de sus vecinos según la relación:

$$a_i^t = (1 - \gamma) + \gamma \sum_{j \in V_i^-} \frac{a_j^{t-1}}{|V_j^-|} \quad \text{siendo } 0 < \gamma < 1$$

donde a_i^t es la autoridad del nodo i en la iteración t , y γ es la **constante de amortiguamiento** que simula un modelo de selección aleatorio: para $\gamma = 0$ el modelo será totalmente aleatorio ya que todos los nodos tendrían la misma probabilidad de ser seleccionados y para $\gamma = 1$ la selección de los nodos sólo se haría en función de la importancia del conjunto de nodos vecinos.

- **El grado de intermediación (betweenness, g_k) del nodo k** (Freeman, 1979; Brandes, 2001): Índice que informa del número de caminos mínimos que pasan por el nodo k .

$$g_k = \sum_{i < k < j} \frac{g_{ikj}^*}{g_{ij}}, \quad \forall k \in V$$

Por otro lado dentro de las medidas egocéntricas o locales se tiene:

- El **grado (degree, d_i) del nodo i** (Freeman, 1979): número de aristas del nodo i .

$$d_i = \sum_{j \in V} A_{ij}, \quad \forall i \in V$$

- La **densidad (density, D_i^-) del nodo i** : proporción de aristas existentes en relación con las posibles aristas entre el conjunto de vecinos del nodo i .

$$D_i^- = \frac{2 * |E_i^-|}{|V_i^-| * (|V_i^-| - 1)}, \quad \forall i \in V$$

- La **redundancia (redundancy, r_i) del nodo i** (Burt, 1992; Borgatti, 1997): mide el grado de cohesión de los vecinos del nodo i .

$$r_i = \frac{2 * |E_i^-|}{|V_i^-|}, \quad \forall i \in V$$

Análisis de Palabras Asociadas o Análisis de Co-Palabras

El objetivo del análisis de las palabras asociadas (Callon, Law & Rip, 1986; Callon, Courter & Laville, 1991; Coulter, Monarch & Konda, 1998; Monarch, 2000) es la creación de mapas temáticos donde se identifiquen unos **focos o centros de interés**, así como las relaciones que existen entre ellos. Estos focos son asimilables a las **áreas temáticas o conceptuales** que componen un área y por ello, el análisis de las palabras asociadas se realiza a partir de un conjunto de documentos representativos de la producción de un área.

A partir de este análisis se obtienen dos productos: los **mapas** y los **supermapas (o diagramas estratégicos)**, los cuales se corresponden con los niveles de análisis **micro** y **meso**. La construcción de los diagramas estratégicos, presupone la reducción o resumen del espacio de términos que componen un conjunto de documentos, mediante un proceso de agregación, con objeto de hacer la información más comprensible.

El análisis de las palabras asociadas se realiza en tres etapas: cálculo de la red de términos, proceso de agregación y construcción de los diagramas estratégicos. Seguidamente se describen cada una de estas etapas:

Cálculo de la red de términos: A partir del conjunto de documentos, se construye una red donde los **nodos** son los términos o palabras que

componen los documentos y la fuerza de las **relaciones** entre cada pareja de términos (**co-ocurrencia** o **co-aparición**) se mide por el **índice de equivalencia** o **coeficiente de inclusión mutua** dado por:

$$0 \leq S(w_i, w_j) = S(k) = \frac{c_{ij}^2}{c_i c_j} \leq 1$$

donde w_i , w_j son dos términos, c_{ij} es el número de documentos en que aparecen conjuntamente los términos w_i y w_j , y c_i es el número de documentos en que aparece la palabra w_i (respectivamente con c_j). Como notación alternativa a $S(w_i, w_j)$ se utiliza la expresión $S(k)$, donde k es la arista que une las palabras w_i y w_j .

Proceso de Agregación (microanálisis): En esta etapa se construyen unas series de mapas cada uno de estos mapas se corresponden con uno de los centros de interés del área o agregados de la red temática. Para ello se asumen dos principios:

- 1) Debido a la redundancia de información en los documentos, se crean redes densas, donde es posible encontrar agregaciones de términos, asimilables a conceptos. Como identificadores de los agregados es posible escoger los términos más representativos.
- 2) Las relaciones entre conceptos se producen por el alto grado de co-ocurrencia que existe entre ellos.

El proceso de generación de los mapas se realiza mediante un proceso de agregación, realizado en dos fases:

Fase 1) Selección de **nodos internos**. En este proceso se tienen en cuenta sólo las aristas que sobrepasan un determinado umbral de co-ocurrencia. Inicialmente el algoritmo escoge la pareja de palabras con mayor co-ocurrencias y realiza una **búsqueda en anchura** para añadir nuevos nodos y aristas. En esta búsqueda tienen preferencia las aristas con mayor co-ocurrencia. También se limita la búsqueda a un número máximo de términos y de enlaces que pueden formar parte de un mapa (típicamente entre 15 y 20 nodos y entre 20 y 25 enlaces). Una vez definido el mapa, se elimina de la red los nodos asignados al mapa. Este

proceso se repite hasta que no existan más relaciones que sobrepasen el umbral inicial (*Figura 1 fase 1*).

Fase 2) Selección de **nodos externos**. En esta fase se añaden a cada mapa generado en la fase anterior, los nodos internos de otros mapas con fuerte relación. La selección de los nodos externos para un mapa, se realiza ordenando de forma decreciente y según su índice de equivalencia, las aristas que unen los nodos internos de ese mapa con los nodos internos del resto de mapas. Los motivos por los que se descarta un nodo externo son debidos a que la arista que lo une al mapa no sobrepasa un determinado umbral o porque el número máximo de nodos o aristas que puede tener un mapa se ha satisfecho. El algoritmo finalizará cuando no pueda seleccionar más aristas (*Figura 1 fase 2*).

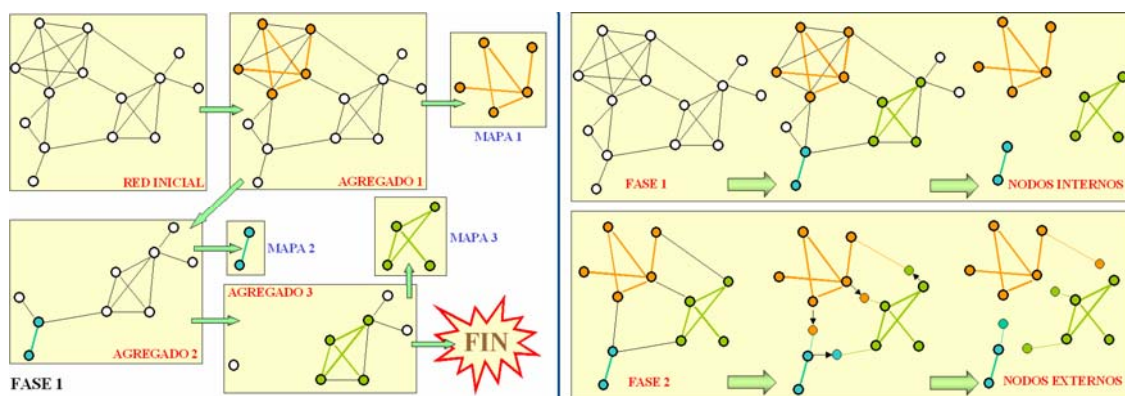


Figura 1: Fase 1 y 2 del proceso de agregación

Construcción de Diagramas Estratégicos (mesoanálisis): En esta etapa se estudia la evolución de la posición estratégica de los distintos agregados (mapas) obtenidos en la etapa anterior. Para ello se definen dos nuevas medidas sobre los enlaces seleccionados en la Fase 1 y Fase 2 de la etapa anterior.

- **Índice de cohesión externa o centralidad del agregado H_i , C_i :** Cuantifica las relaciones que une un agregado con el resto de agregados de la red. Cuanto mayor sea, más crucial será el tema de investigación para el área. El índice se calcula como:

$$C_i = \frac{\sum_{k \in Out(i)} S(k)}{|Out(i)|}$$

- **Índice de cohesión interna o densidad de un agregado H_i , D_i :**
 Cuantifica el grado de cohesión del agregado. Cuanto mayor es este índice, más capacidad tiene el agregado para desarrollarse y permanecer en el tiempo. Por otro lado, si el valor que toma es pequeño significa que el agregado está compuesto de relaciones débiles que no se cierran unas sobre otras, y por tanto dicho agregado se está descomponiendo o en vías de formación. Se calcula como:

$$D_i = \sqrt{\sum_{k \in In(i)} S^2(k)}$$

los autores del técnica de análisis de palabras asociadas, escogen esta definición de densidad con el objetivo de poder diferenciar entre los valores relativamente próximos.

A partir de estos índices¹ se construyen los diagramas estratégicos, colocando cada uno de los agregados, en un diagrama de dos dimensiones, donde el eje X y el eje Y representan la centralidad y la densidad respectivamente. De esta manera quedan definidas cuatro regiones, donde hemos tomado como origen, el punto medio (\bar{X}, \bar{Y}) de los valores de estos índices y como límites, el máximo y el mínimo de X y de Y . Cada región presenta una semántica que permite su adecuada interpretación.

Diagramas Estratégicos y Redes Sociales

Una de las principales aportaciones de este trabajo es establecer un paralelismo entre las redes de términos y las redes sociales, de forma que los nodos sean individuos en vez de términos y el índice de equivalencia indique la fuerza de las relaciones entre los individuos. Para el caso de las redes sociales el índice de equivalencia se redefine como:

$$0 \leq S(a_i, a_j) = S(k) = \frac{A_{ij}^2}{A_i A_j} \leq 1$$

donde a_i , a_j son dos actores distintos de la red; A_{ij} , es el peso de la relación entre a_i y a_j ; y A_i , la suma de los pesos de todas las relaciones del actor a_i .

¹ Mustchke, 2004 propone como medidas alternativas para el cálculo de la centralidad y la densidad las siguientes: $C_i = \frac{|Out(i)|}{(|H|-1)}$ $D_i = \frac{2 * |In(i)|}{|H_i| * (|H_i|-1)}$

(de forma análoga para a_j). Para el caso particular de una red de co-autorías, estos parámetros coincidirían con el número de publicaciones conjuntas de dos individuos y el número total de publicaciones que realiza un individuo.

Basándonos en este paralelismo, en esta sección, se hace uso de los diagramas estratégicos para analizar la estructura organizativa de las JISBD. El diagrama estratégico de las JISBD se ha construido a partir de la red de co-autorías de dichas jornadas. Por tanto, esta red, se compone de los autores que publicaron en dichas jornadas desde 1996 hasta 2003 y de las relaciones de co-autorías que los unen. Debido al proceso seguido para representar el diagrama estratégico, se ha sustituido el índice de equivalencia por el siguiente **índice de co-autorías**:

$$ca_{ij} = \begin{cases} 0 & \text{si no existe publicación conjunta entre los autores } a_i \text{ y } a_j \\ 1 & \text{si existe publicación conjunta entre los autores } a_i \text{ y } a_j \end{cases}$$

El proceso seguido, para construir el diagrama estratégico, ha consistido en:

1. la descomposición de la red en agregados (Givan & Newman, 2002),
2. el cálculo de los índices de cohesión (Mustchke, 2004),
3. y la representación del diagrama estratégico (Callon, Law & Rip, 1986; Callon, Courter & Laville, 1991).

En la *Figura 2* se observa la descomposición de la red en agregados, cada agregado equivale a un colectivo o subcomunidad dentro de la red. Y en la *Figura 3* se presenta el diagrama estratégico de la red. En este diagrama cada agregado se simboliza mediante un círculo. El radio del círculo es representativo del número de autores, cuanto mayor sea, más autores tiene el agregado; y su etiqueta informa sobre los autores más relevantes del agregado, estos autores se seleccionan a partir de su autoridad.

Como se observa en la *Figura 3* la densidad de los agregados y su tamaño son inversamente proporcionales, cuanto mayor es el número de actores que componen un agregado menos densidad tiene. Se ha demostrado en varios estudios que esta característica se presenta en varios tipos de redes, entre ellas destacan las redes sociales (Friendkin, 1981; Niemeijer, 1973;

Snijders, 1981). Debido a este sesgo es preferible utilizar el grado medio de un agregado como medida de cohesión interna, los resultados obtenidos utilizando este nuevo indicador se presentan en la *Figura 4*.

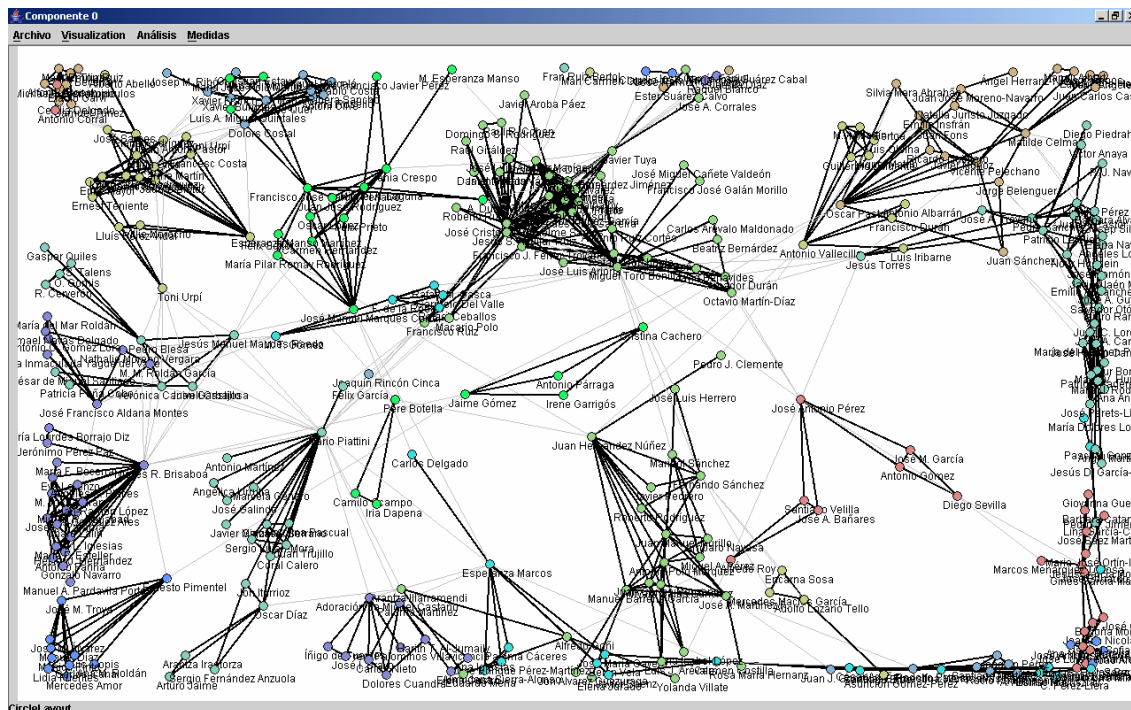


Figura 2: Sociograma de las JISBD. Los actores asociados a un agregado son del mismo color

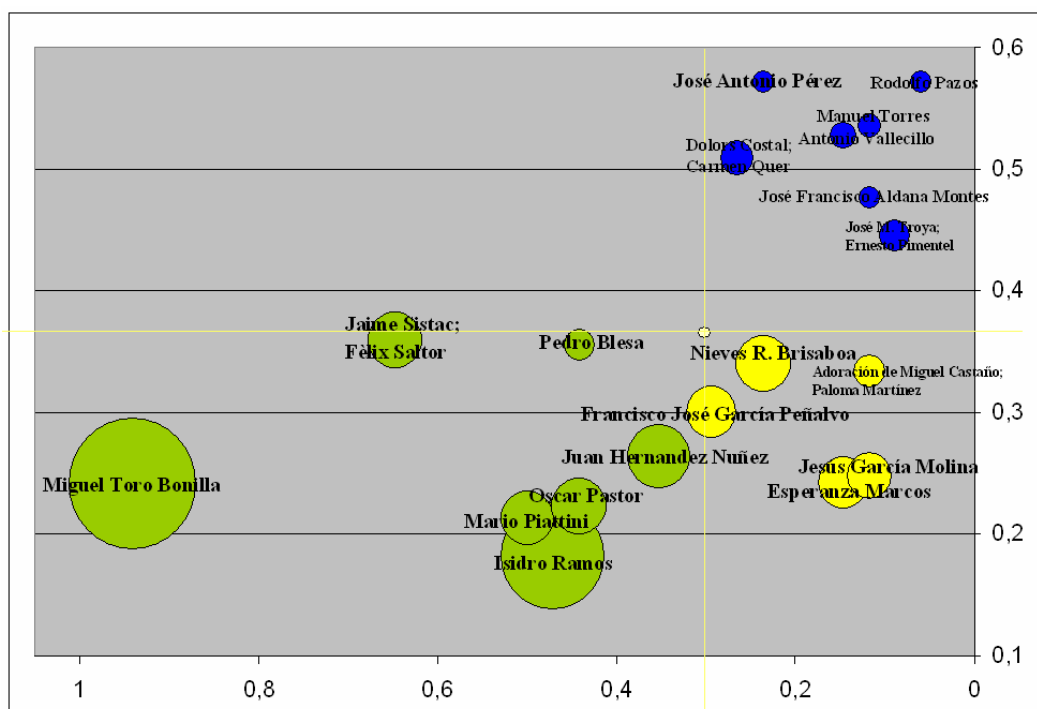


Figura 3: Diagrama estratégico de la comunidad de las JISBD. El eje X, representa la centralidad y el eje Y, la densidad

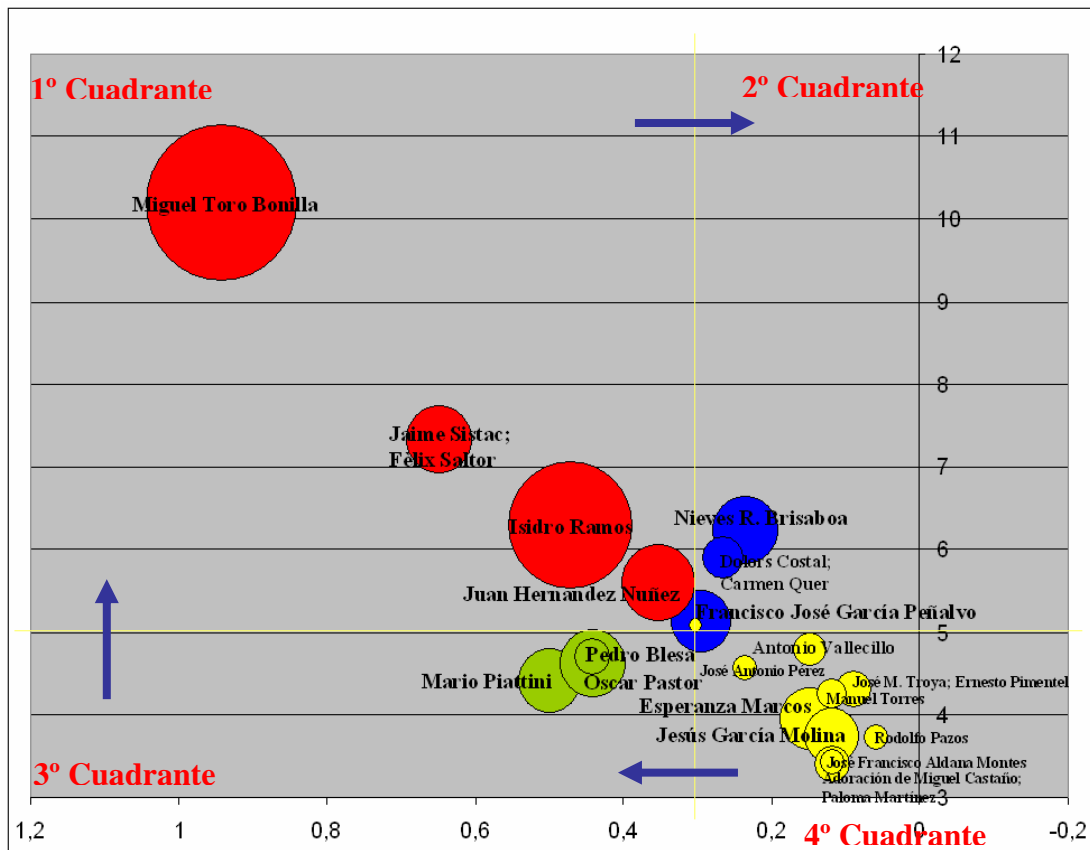


Figura 4: Diagrama estratégico de las JISBD. El eje X, representa la centralidad y el eje Y, el grado medio del agregado

Los diagramas estratégicos se proponen como un instrumento para estudiar la forma en que se estructuran los temas que componen un campo de investigación. En este trabajo se analiza la estructura de una red social a partir de su diagrama estratégico. De esta forma, se define una semántica que caracteriza los agregados que componen la red dependiendo de la posición que ocupen en el diagrama. Utilizando el valor medio de los índices de cohesión se definen cuatro regiones (*Figura 4*):

- **Región o Cuadrante 1:** Se corresponde con los agregados que componen el *núcleo de la comunidad*. Estos agregados tienen un alto grado de desarrollo e integración y disponen de una buena posición estratégica dentro de la red.
- **Región o Cuadrante 2:** Se corresponde con los agregados con alta densidad y baja centralidad. Estos agregados están muy desarrollados. Probablemente fueron comunidades que pertenecieron al núcleo y que han ido perdiendo interés, quedando aisladas.

- **Región o Cuadrante 3:** Se corresponde con los *colectivos emergentes*. Estos agregados están bien conectados, pero poco desarrollados. Son susceptibles de pertenecer al núcleo de la red en el futuro.
- **Región o Cuadrante 4:** Se corresponden con los *colectivos periféricos*. Son agregados poco desarrollados, que definen los límites de la red. Estas comunidades son susceptibles de convertirse en emergentes.

Atendiendo a la clasificación anterior, se definen tres categorías que describen la estructura organizativa de una red (Figura 5):

- **Categoría 1:** Primera bisectriz. Es una red organizada en torno a unos colectivos bien estructurados y desarrollados, con los cuales se relacionan otros colectivos periféricos y poco desarrollados. Corresponde a una comunidad altamente especializada.
- **Categoría 2:** Segunda bisectriz. Representa una red en vía de estructuración.
- **Categoría 3:** Es una red muy compleja y rica. En ella encontramos todos los posibles tipos de colectivos. Sugiere una dinámica importante de la red.

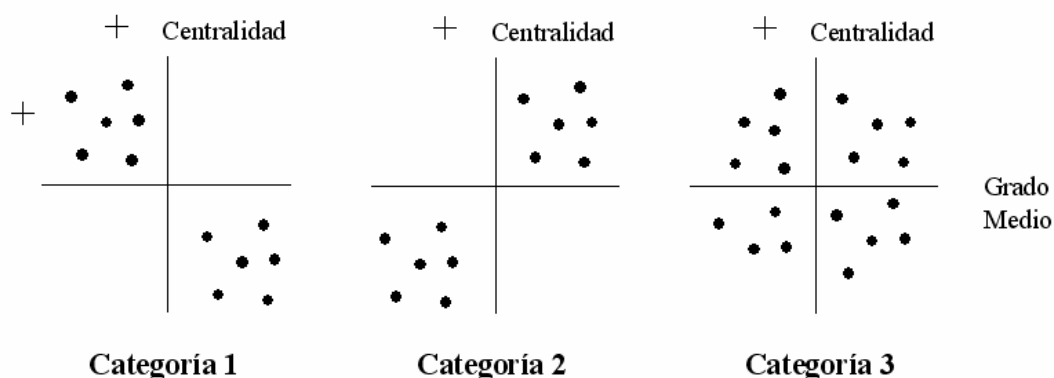


Figura 5: Categorización de los diagramas estratégicos en tres estructuras organizativas de una red social

Basándonos en las tres categorías definidas, la estructura organizativa de la comunidad de autores que publican en las JISBD presenta una dinámica compleja y rica, con un grado de especialización alto. También es posible, gracias a trabajos previos (Callon, Law & Rip, 1986; Callon, Courter &

Laville, 1991; Coulter, Monarch & Konda, 1998; Monarch, 2000), predecir cual será la dinámica futura de los distintos agregados (el flujo de agregados entre regiones está representado en la *Figura 4* con flechas azules).

Diagramas Estructurales y Redes Sociales

Aunque los diagramas estratégicos proporcionan una forma simplificada y sintética de analizar la estructura de una red, hay que tener en cuenta que no todos los individuos que componen un agregado tienen las mismas funciones. Basándonos en los trabajos citados arriba y en varias medidas utilizadas en el análisis de redes sociales (Scott, 2000) (autoridad, intermediación, grado, redundancia y densidad), se propone como elemento original en este trabajo que los diagramas estructurales sea una forma complementaria de analizar la estructura de las redes sociales. Mientras los diagramas estratégicos analizan la posición estratégica de los agregados dentro de una red (mesoanálisis), los diagramas estructurales analizan la posición estratégica de los nodos o actores de una red (microanálisis), permitiendo:

- Asociar a cada nodo de la red un patrón distintivo
- y estudiar la dinámica de la red en el futuro.

Para ilustrar este marco de trabajo, en esta sección se lleva a cabo el análisis mediante diagramas estructurales de la comunidad de las JISBD.

Al igual que los diagramas estratégicos, los diagramas estructurales posicionan los nodos en una representación 2D. Para ello se definen dos índices, que representarán cada uno de los ejes del diagrama y que son equivalentes a los índices utilizados en el análisis de diagramas estratégicos:

- **Índice de cohesión externa de un nodo (eje X):** Con este índice se persigue representar la posición estratégica de un nodo respecto al resto de nodos de la red, para ello es posible elegir cualquiera de las medidas globales. En este estudio se han utilizado la centralidad, la autoridad y la intermediación.

- **Índice de cohesión interna de un nodo (eje Y):** Con este índice se persigue representar la posición del nodo respecto a su entorno más cercano, sus vecinos, para ello es posible elegir cualquiera de las medidas locales. En concreto se han utilizado el grado, la redundancia y la densidad.

A diferencia de los diagramas estratégicos, los diagramas estructurales proponen utilizar una familia de mapas para analizar la estructura de una red. De esta manera, eligiendo diferentes medidas se consigue que los diagramas proporcionen distintas perspectivas de la red social, aunque no todas admitan una interpretación útil y clara. Como ejemplo de aplicación se ha elegido el diagrama estructural grado-autoridad y redundancia-intermediación, los cuales han permitido definir una serie de **patrones de comportamiento** de gran interés para la clasificación de los nodos.

Respecto al diagrama grado-autoridad, una vez representados los nodos y tomando como eje de referencia el nodo (\bar{X}, \bar{Y}) (nodo promedio), se ha dividido el plano en cuatro regiones, las cuales admiten la siguiente interpretación (*Figura 6*):

- **Región o Cuadrante 1:** Se corresponde al **núcleo de la comunidad**, compuesto por los individuos con mayor autoridad y grado. Estos individuos tienen un alto grado de desarrollo e integración y disponen de una buena posición estratégica dentro de la red.
- **Región o Cuadrante 2:** Se corresponde con los **individuos distanciados o aislados** de la comunidad, individuos con alto grado y baja autoridad. Estos individuos pertenecen a un grupo bien estructurado, pero que no se relacionan con individuos ajenos al grupo. Tiene una posición estratégica aislada.
- **Región o Cuadrante 3:** Se corresponde con los **individuos emergentes**, son individuos con una gran autoridad y con poco grado. Estos individuos se relacionan con un grupo pequeño de vecinos entre los cuales realiza la función de intermediador. Son susceptibles de pertenecer al núcleo de la red en el futuro.

- **Región o Cuadrante 4:** Se corresponden con individuos *periféricos*, son individuos con poca autoridad y poco grado. Es una comunidad poco desarrollada, que definen los límites de la red social. Los individuos que componen esta región son susceptibles de convertirse en emergentes.

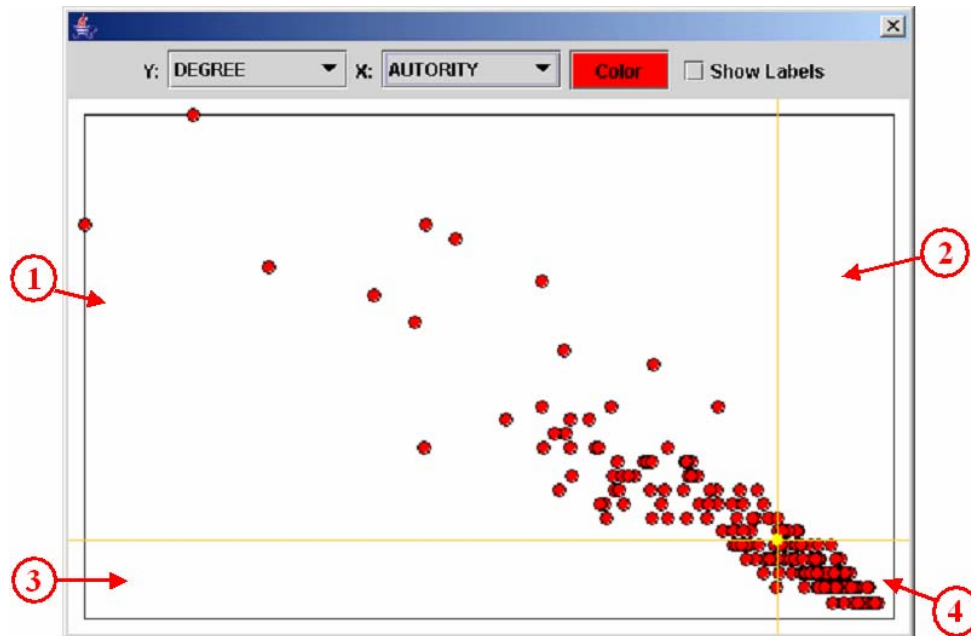


Figura 6: Diagrama estructural grado-autoridad de las JSIBD. Enumeración de las cuatro regiones en que se divide un diagrama estructural. El nodo de color amarillo es la media

La asignación de un color a cada comunidad permite utilizar éste como un indicador del posicionamiento estratégico del individuo dentro de la comunidad, facilitando la comprensión del **sociograma** (Moreno, 1934) así como la exploración de las distintas familias que componen la comunidad. De esta forma se ha coloreado los individuos como sigue: rojo=núcleo, azul=distanciado, verde=emergente y amarillo=periférico (Figuras 7 y 8).

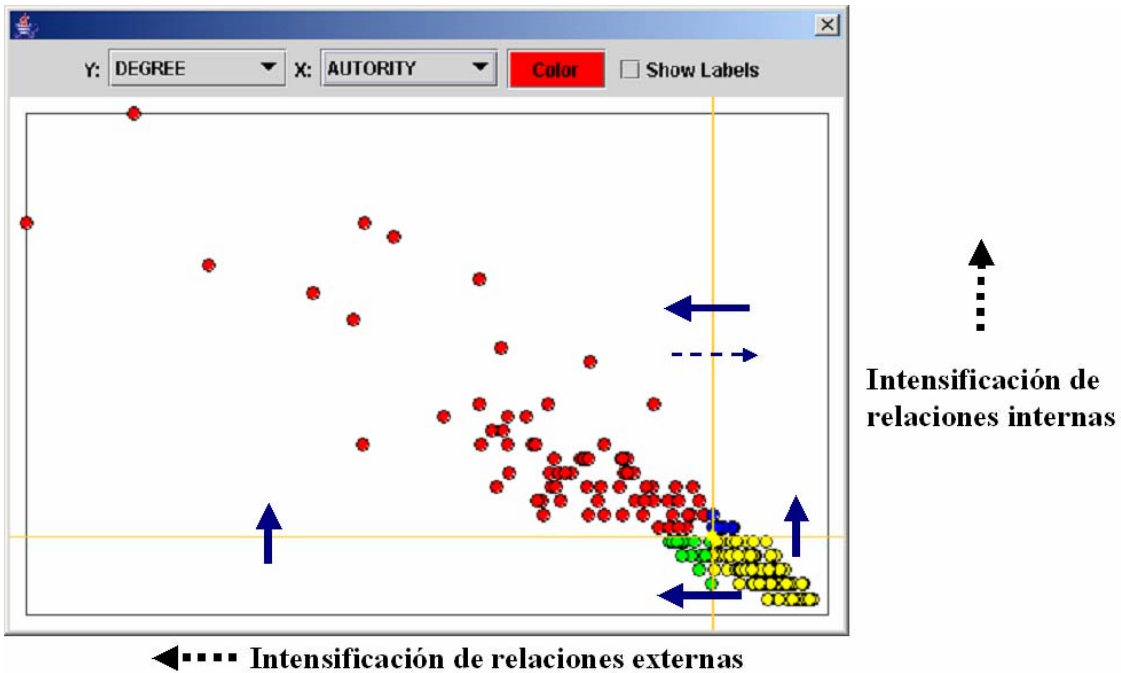


Figura 7: Diagrama estructural grado-autoridad de las JISBD. El eje X, representa la autoridad, y eje Y, el grado. El coloreado de los nodos es: rojo=núcleo, azul=aislados verde=emergentes y amarillo=periféricos. Las flechas azules indican el flujo de nodos entre regiones y las flechas negras punteadas indican el tipo de relación a intensificar para determinada dirección de flujo

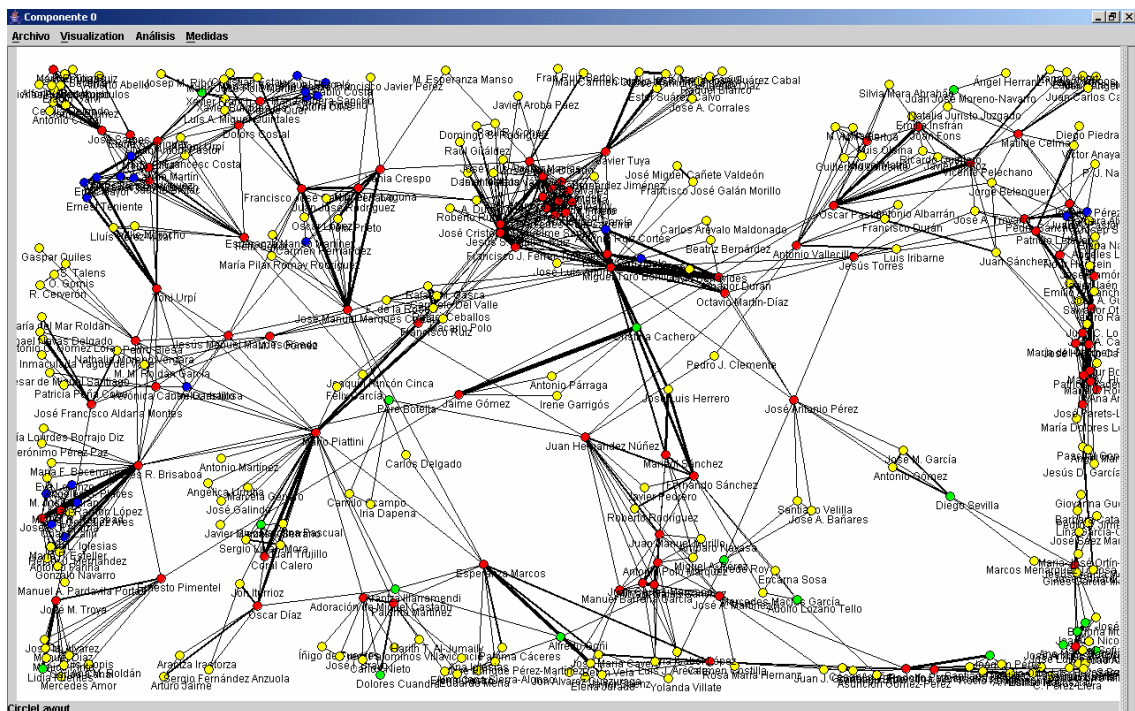


Figura 8: Sociograma de las JISBD. Los nodos se han coloreado según los patrones definidos en el diagrama estructural grado-autoridad de la Figura 7

Para llevar a cabo un estudio de esta red, desde el punto de vista de la dinámica del sistema, supongamos que esta se caracteriza por las leyes que rigen en las redes complejas de tipo *scale-free*. Según los trabajos de

(Barabási & Albert, 1999; Newman, 2000; Barabási, 2002; Barabási & Bonabeau, 2003), estas redes se generan a partir de dos leyes fundamentales:

- **Crecimiento lineal:** a medida que pasa el tiempo, nuevos nodos se van agregando linealmente a la red.
- **Enlace preferencial:** Cada nuevo enlace de la red se enlaza preferentemente a los nodos con mayor número de enlaces. Por tanto, sería más probable que un nodo emergente pase a formar parte del núcleo que a la periferia o que directamente se transformase en un nodo distanciado. Al igual que un nodo que pertenezca al núcleo es prácticamente imposible que pasase a formar parte de la zona que define la periferia.

Es por ello por lo que se estudian los posibles flujos de los nodos (éstos están representados en la *Figura 7* con flechas azules):

- **Desde periféricos a emergentes:** Actores periféricos que han sabido aprovechar los agujeros estructurales de la red, relacionándose con Actores de otros grupos, para mejorar la calidad de sus trabajos.
- **Desde periféricos a distanciados:** Actores periféricos que han intensificado sus relaciones con otros miembros periféricos dentro de su mismo grupo.
- **Desde emergente a núcleo:** Actores cuyos trabajos se han ido consolidando dentro de la red, realizando proyectos con otros miembros de la comunidad y construyendo un grupo afín.
- **Desde distanciados a núcleo:** Actores distanciados que han intensificado sus relaciones con otros miembros de la red ajenos a su grupo.
- **Desde núcleo a distanciados:** En menor medida y a largo plazo, puede ocurrir que actores que dejan de publicar queden distanciados, alejándose cada vez más del núcleo de la red.

Partiendo de estas definiciones, se observa, como para transformar nodos desde periférico a distanciado y desde emergente a núcleo es necesario

intensificar sus relaciones internas y para el caso de las transformaciones de flujo de periférico a emergente y de distanciado a núcleo se debe intensificar las relaciones externas (*Figura 7* flechas negras punteadas).

Por último destacar, la posible existencia de colaboraciones con miembros pertenecientes a otras redes, en este caso estaríamos ante una posible revolución conceptual, debido a la colisión entre dos áreas. Esto podría provocar una reestructuración de la red.

Al igual que se ha utilizado el grado y la autoridad para definir las distintas familias que componen una comunidad, podemos utilizar otro diagrama estructural para definir los patrones de nodos intermediadores. En este caso utilizaríamos un diagrama estructural representado la redundancia, r_i , y la intermediación, g_k , como índices de estructuración interna y externa respectivamente (*Figura 9*). En este diagrama lo más interesante es resaltar dos regiones:

- **Región o Cuadrante 1:** Esta compuesta por actores con un alto grado de intermediación y una alta redundancia de sus vecinos. Son los actores del núcleo con pocas posibilidades de explotar sus posiciones estratégicas (en la *Figura 10* podemos comprobar como la mayoría de los actores que pertenecen a este cuadrante pertenecen al núcleo, pero no es cierto a la inversa). Esta región nos permite definir el patrón **intermediador débil**.
- **Región o Cuadrante 3:** Esta compuesta por actores con un alto grado de intermediación y una baja redundancia entre sus vecinos. Serían actores que están en posición de aprovechar los agujeros estructurales de la red. Esta región nos permite definir el patrón **intermediador fuerte**.

En la *Figura 11* podemos observar el sociograma donde el coloreado de las coronas de los nodos permite distinguir entre los dos tipos de nodos intermediadores.

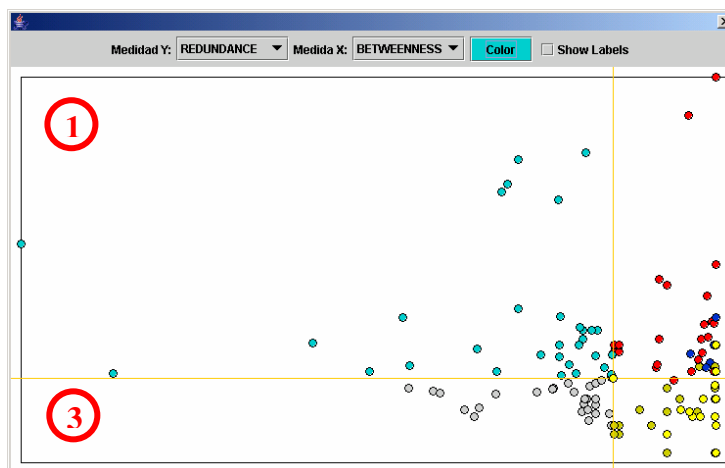


Figura 9: Diagrama estructural redundancia-intermediación de las JISBD. El eje X representa la intermediación, y el Y, la redundancia. Coloreado de patrones: cyan=intermediador_débil y gris=intermediador_fuerte

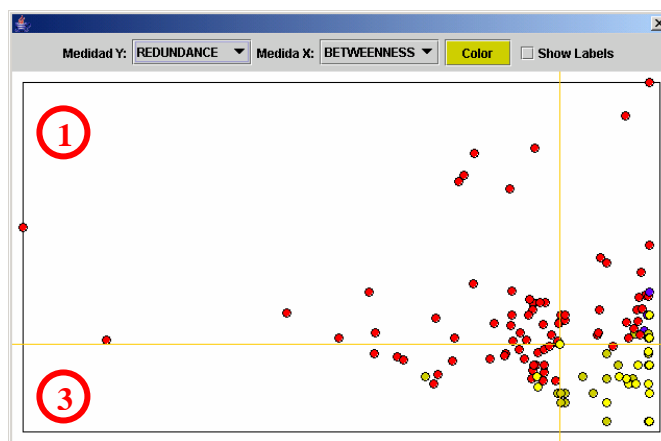


Figura 10: Diagrama estructural redundancia-intermediación de las JISBD. El eje X, representa la intermediación y el Y, la redundancia. Los nodos se han coloreado según los patrones definidos en el diagrama grado-autoridad de la Figura 7

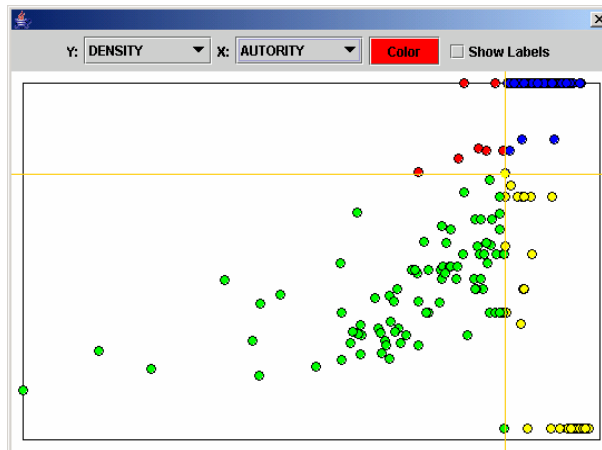


Figura 12: Diagrama estructural densidad-autoridad de las JISBD. El eje X, representa la autoridad, y el Y, la densidad. Coloreado de los nodos según las distintas regiones: rojo=cuadrante1, azul=cuadrante2, verde=cuadrante3 y amarillo=cuadrante4

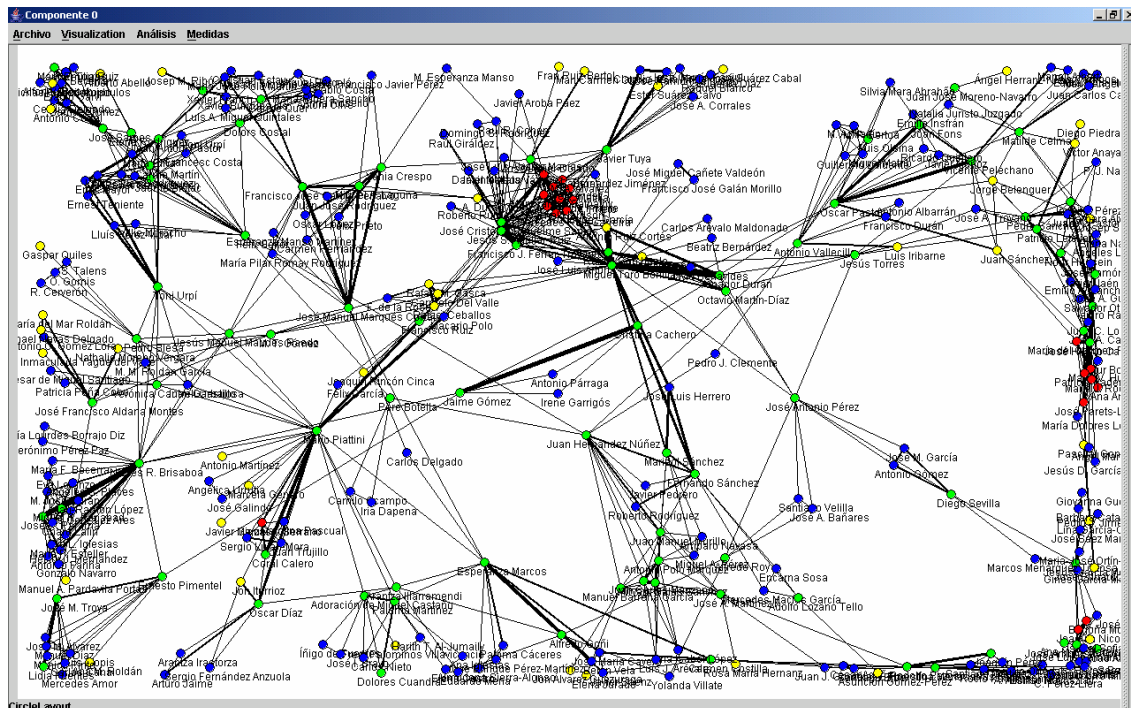


Figura 13: Visualización del sociograma de las JISBD. Coloreado de los nodos según las distintas regiones definidas en diagrama estructural densidad-autoridad de la Figura 9

En diversos estudios (Friendkin, 1981; Niemeijer, 1973; Snijders, 1981) se aconseja cautela cuando se comparan las densidades de redes con tamaños distintos. En el contexto de este trabajo no se aconseja utilizar la densidad D_i , ya que favorece a los individuos con pocos vecinos (en el ejemplo de la Figura 14 podemos observar el impacto que sufre la densidad al añadir un nuevo nodo a una red). Debido a este sesgo es preferible utilizar el grado en vez de la densidad local como medida de estructuración interna.

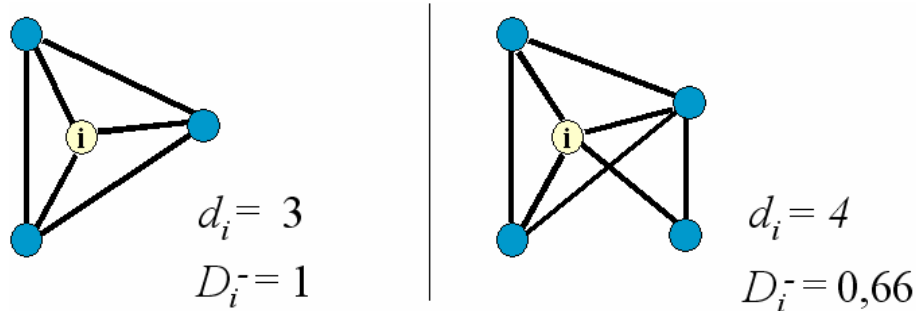


Figura 14: Ejemplo del sesgo de la densidad a favor de las redes con menos nodos

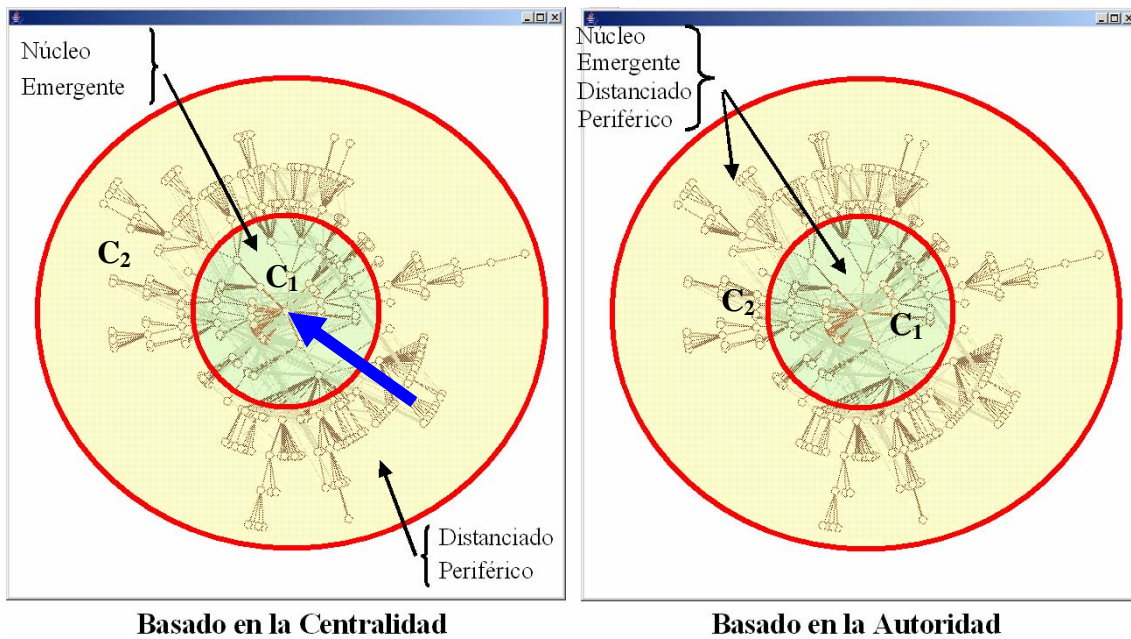
Otro aspecto interesante en esta discusión es decidir si utilizar como indicador de cohesión externa la centralidad o la autoridad. Se ha estudiado este problema desde dos perspectivas, el **significado de los patrones** y la **estabilidad** de los diagramas.

Dependiendo de la medida utilizada (centralidad o autoridad) el **significado de los patrones** sufre alteraciones:

- En el caso de seleccionar la centralidad (*Figuras 16 y 17*), un mismo patrón se agrupa en regiones donde los nodos tienen centralidad parecida. Por ejemplo, si dividimos los nodos en dos conjuntos, los que tienen centralidad mayor y menor que la media, conjuntos c_1 y c_2 respectivamente, podemos observar que los patrones núcleo y emergente solo aparecen en el conjunto c_1 y los patrones distanciado y periférico solo aparecen en el conjunto c_2 (*Figura 15*).
- El seleccionar la autoridad, nos permite determinar dentro de una misma región con centralidad parecida, patrones distintos en los individuos que la componen (*Figura 15*). Por tanto esta segunda opción se ajusta mejor a los objetivos planteados en este trabajo, al clasificar cada nodo con un patrón, con independencia de la región en que se ubique.

Respecto a la **estabilidad** se observa, para el diagrama estructural grado-centralidad, que si se tienden enlaces entre los conjuntos c_1 y c_2 , los individuos distanciados y periféricos, afectados por estos enlaces, transforman sus patrones a núcleo y a periférico respectivamente, con una probabilidad alta (flecha azul de la *Figura 15*). Sin embargo el diagrama centralidad-autoridad es más robusto a la inserción de estos enlaces, siendo necesario un mayor número de enlaces para modificar la asignación de

patrones a los nodos (*Figuras 18 y 19*). Teniendo en cuenta estos datos se recomienda utilizar la autoridad antes que la centralidad.



Basado en la Centralidad **Basado en la Autoridad**
Figura 15: Zonas de disposición de los patrones dependiendo de la medida utilizada en el diagrama estructural: centralidad o autoridad. Los círculos delimitan dos áreas de la comunidad, definidas por el centralidad media de los nodos (círculo interior)

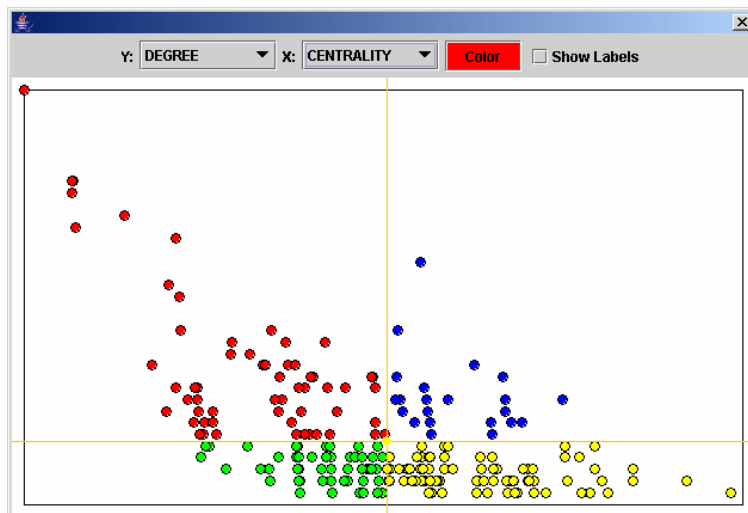


Figura 16: Diagrama estructural grado-centralidad de las JISBD. Coloreado de los nodos según los distintos patrones: rojo=núcleo, azul=distanciado, verde=emergentes y amarillo=periféricos

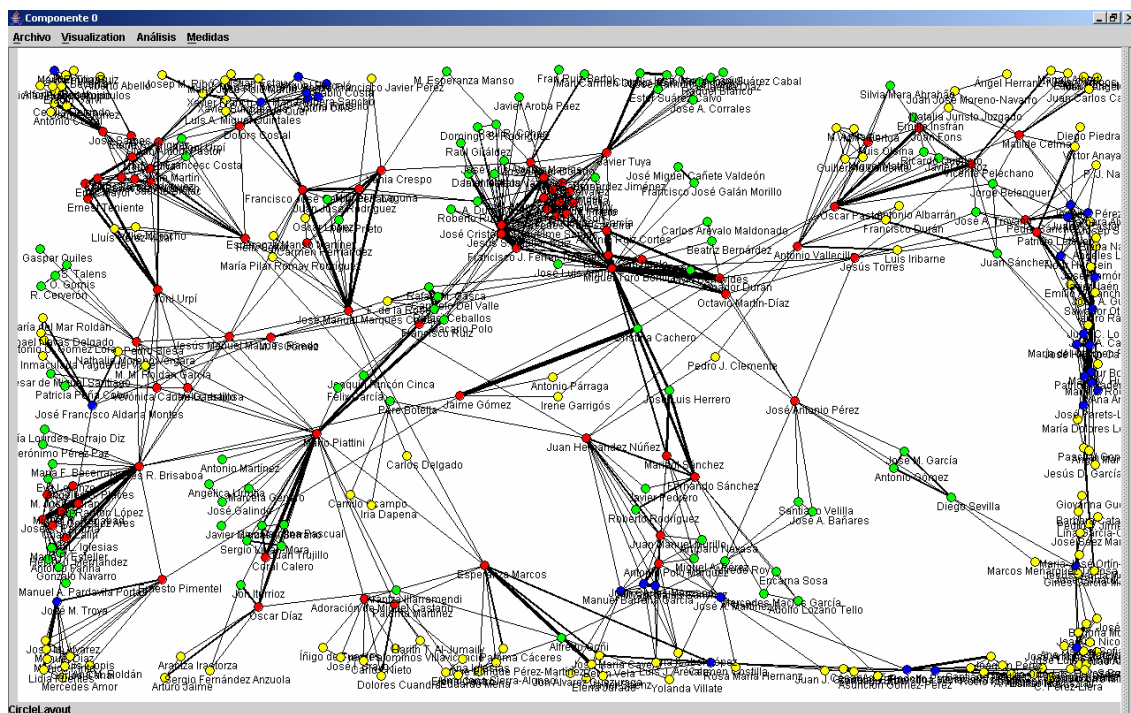


Figura 17: Sociograma de las JISBD. Coloreado de los nodos según los distintos patrones definidos en el diagrama estructural grado-centralidad de la Figura 16

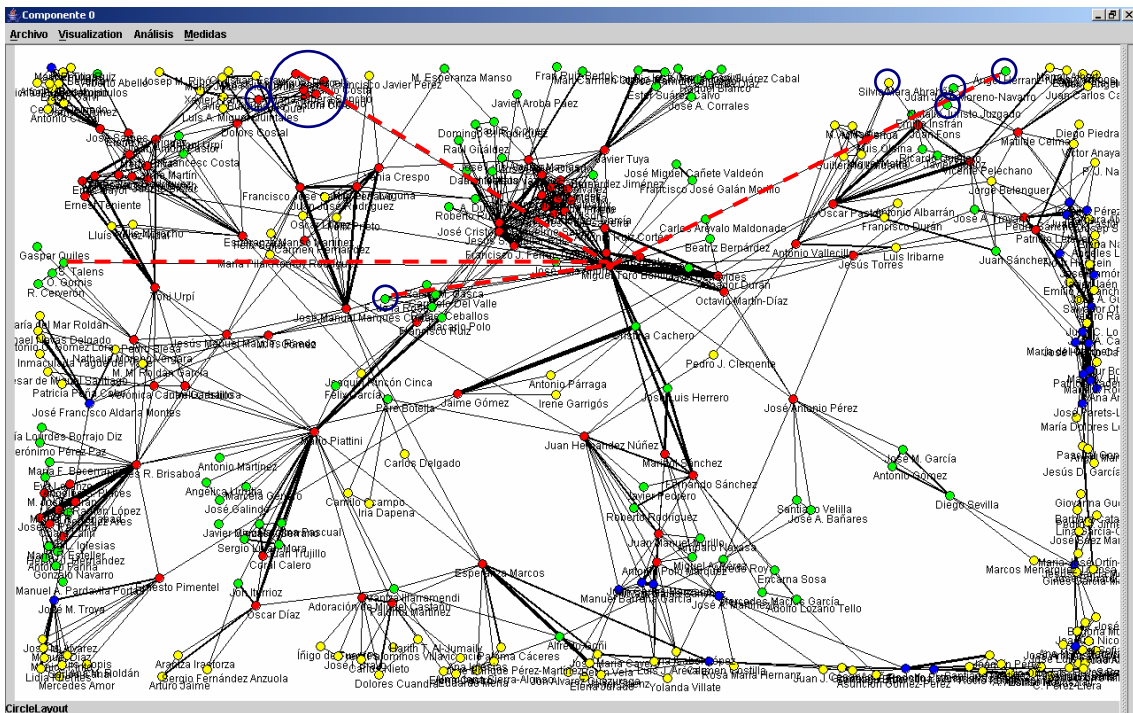


Figura 18: Sociograma de las JISBD. Coloreado de los nodos según los distintos patrones definidos en el diagrama estructural grado-centralidad de la Figura 16. Las aristas punteadas son los enlaces añadidos y los círculos grandes los patrones modificados

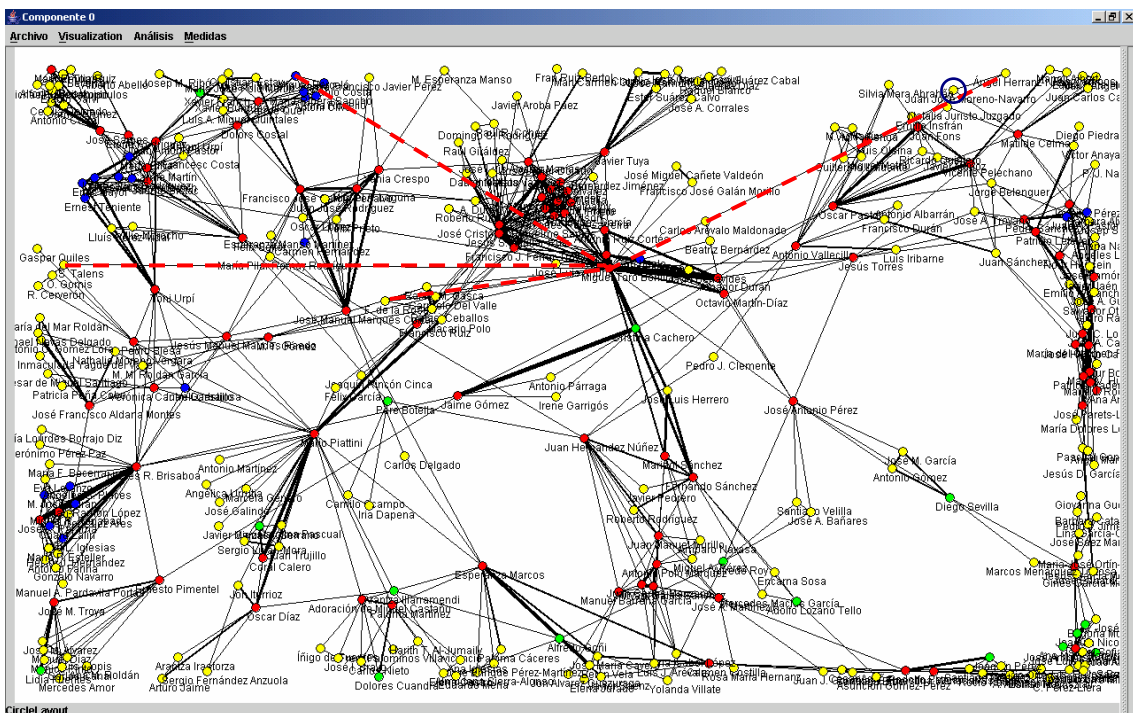


Figura 19: Sociograma de las JISBD. Coloreado de los nodos según los distintos patrones definidos en el diagrama estructural grado-autoridad de la Figura 7. Las aristas punteadas son los enlaces añadidos y los círculos grandes los patrones modificados

Conclusiones

Dentro de una comunidad científica los congresos juegan un papel muy importante como herramienta para intercambio de ideas, creación de redes de trabajo, proyectos de investigación, etc. En general es uno de los motores del proceso creativo científico¹. A pesar de todo ello, es la producción científica expresada en forma de artículos el único producto tangible que perdura en el tiempo. Este producto se obtiene a partir de la evaluación de los trabajos propuestos, como resultado de la crítica colectiva, y muchas veces representan la vanguardia del conocimiento. Desde el punto de vista del investigador lo importante no son los trabajos en sí, sino tener la capacidad de sintetizar las estructuras que se generan en la comunidad a la cual pertenece, como pueden ser su red social ó sus áreas temáticas. Esto le ayudará a satisfacer las necesidades informativas que necesita para orientar sus investigaciones y desempeñar correctamente su trabajo. Con este fin, en este trabajo, se han desarrollado dos nuevos instrumentos que permiten observar como se organiza la red social de una comunidad. Permitiendo observar la posición estratégica de los agregados que la componen. Estos agregados representarían los "colegios invisibles" que se forman en estas comunidades en su etapa de madurez, y que se han definido (Solla Price, 1963) como "círculos de investigadores influenciados por unos pocos investigadores de alta productividad".

En principio, la aplicabilidad de los instrumentos o técnicas de análisis propuestas en este trabajo, no se reducen al estudio de las comunidades científicas, también pueden ser utilizados por las empresas con la idea de obtener alguna ventaja competitiva. Por ejemplo una empresa podría estar interesada en buscar expertos que colaborasen en algún proyecto conjunto o conocer los intereses de los investigadores de empresas competidoras. También serviría para explorar las relaciones que existen entre los miembros de un foro de discusión o para explorar las estructuras de un conjunto de páginas Web. En general serviría para explorar cualquier información estructurable en forma de red.

¹ En otras actividades, que se desarrollan en sociedad, existen herramientas que realizan una función similar a los congresos y que se conocen con otros nombres como ferias, conferencias, asociaciones, etc.

Las técnicas propuestas permiten analizar como se organizan las estructuras que componen una red, cada técnica realiza este análisis a diferentes niveles de granularidad, pero a su vez la información obtenida se complementa: los diagramas estratégicos, estudian la organización de las comunidades o agregados de una red (mesoanálisis), mientras que los diagramas estructurales estudian la posiciones estratégicas de los actores que la componen (microanálisis). Tras analizar la aplicabilidad de estas técnicas se concluye que:

- Los diagramas estratégicos son un instrumento útil para representar de forma simplificada la organización de los agregados o comunidades de una red social, permitiendo **clasificar la estructura organizativa de una red social** en una serie de patrones.
- Los diagramas estructurales grado-autoridad y redundancia-intermediación añaden una nueva dimensión a la información proporcionada por los diagramas estratégicos. La **asignación de patrones** a los actores de un red, permiten diferenciar la posición estratégica que ocupa cada individuo tanto a nivel global (la red) como a nivel local (el agregado). Así mismo la combinación de estos patrones con técnicas de visualización potencian el análisis y la exploración visual de las estructuras de una red.

La información aportada por estas dos técnicas permite utilizarlas como un instrumento para comprender la dinámica que subyace en una red y aprovechar este conocimiento para modelar la red. Por ejemplo para enriquecer la actividad desarrollada en una red clasificada como categoría 1, nuestro objetivo sería transformarla en una red clasificada dentro de la categoría 3. Para conseguir este cambio, las actuaciones deberían perseguir crear comunidades emergentes, a partir de comunidades periféricas, para ello se deberían aprovechar los agujeros estructurales de la red, fomentando las colaboraciones entre los actores periféricos de las distintas comunidades periféricas. Durante el periodo de actuación, se pueden planificar actividades de monitorización de la red y aplicar medidas correctoras en caso de observar una evolución no esperada (Hackthorn, 2003).

Estas técnicas no solamente presentan ventajas, sino que desde el punto de vista del análisis de redes sociales podrían tener las siguientes limitaciones:

- La teoría de los lazos débiles (Granovetter, 1973) advierte que las estructuras sociales se componen de núcleos fuertes y de relaciones débiles o esporádicas que se producen entre estos núcleos y que pueden resultar ventajosas para los individuos que las disfrutan. En este trabajo no se han tenido en cuenta todos los lazos débiles que pueden surgir en una comunidad científica, ya que esta información no está reflejada en la producción científica disponible.
- También es necesario, para poder aplicar estas técnicas, que la comunidad se encuentre en su fase de madurez, momento en el cual emergen los colegios invisibles. En este trabajo se considera que esta fase se produce cuando es posible observar en la red una o varias componentes conexas importantes.
- Por último apuntar que las técnicas propuestas se encuentran limitadas por las técnicas de visualización disponibles, ya que existe un número de nodos a partir del cual se hace difícil una representación visual comprensible de la red.

Finalmente es importante considerar que aunque hemos obtenido buenos resultados al utilizar el grado medio en los diagramas estratégicos como medida de cohesión interna de los agregados (no existe una relación directa entre el grado medio y el número de investigadores que componen un agregado) resulta interesante proponer como trabajo futuro el refinamiento de esta medida, ya que por su definición, existe la tendencia en los agregados a concentrarse alrededor del punto medio que separa las distintas regiones del diagrama.

Bibliografía

- Barabási, A. L. (2002). *Linked*, Perseus Publishing.
- Barabási, A. L. and Albert, R., (1999) "Emergence of scaling in random networks", *Science*, Vol 286, Issue 5439, pp. 509-512.
- Barabási, A. L. and Bonabeau, E. (2003) "Scale-Free Networks". *Scientific America*. nº 288, pp. 60-69.
- Bonacich, P. (1972) "Factoring and weighting approaches to status scores and clique identification". *Journal of Mathematical Sociology*. nº 2, pp. 113-120.
- Borgatti S. P. (1997) "Structural Holes: Unpacking Burt's Redundancy Measures", *Connections*, 20(1), pp. 35-38.
- Brandes, U. (2001). "A Faster Algorithm for Betweenness Centrality". *Journal of Mathematical Sociology*, 25(2): pp. 163-177.
- Burt, R. S. (1992). *Structural Holes: the Social Structure of Competition*, Cambridge: Harvard University Press.
- Callon, M., Courtial, J.P. and Laville, F. (1991). "Co-Word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry". *Scientometrics*, vol. 22, nº 1, pp. 155-205.
- Callon, M., Law, J., and Rip, A. (1986). *Mapping the dynamics of science and technology: Sociology of science in the real world*. London: Macmillan.
- Coulter, N., Monarch, I. and Konda, S. (1998). "Software engineering as seen through its research literature: A study in co-word analysis". *Journal of the American Society for Information Science*, 49(13), pp. 1206-1223.
- Derek J. de Solla Price. (1963) "Little Science, Big Science." Columbia Univ. Press, New York.
- Freeman, L. C. (1979) "Centrality in social networks. Conceptual Classification". *Social Network*, nº 1.
- Friendkin, N. E. (1981). "The Development of Structure in Random Networks". *Social Networks*, nº 3.
- Givan, M. and Newman, M. E. J. (2002). "M. Community structure in social and biological networks", *Proceedings of the National Academy of Sciences of United States of America*, June 11, Vol 99, nº12.

- Granovetter, Mark (1973). "The Strength of Weak Ties" *American Journal of Sociology* 78 (6), pp. 1360-1380.
- Hackthorn, R. (2003). "The Link is the Thing". *DM Review*, August-October.
- Monarch, I. (2000). "Information Science and Information Systems: Converging or Diverging?" <http://www.slis.ualberta.ca/cais2000/monarch.htm>
- Moreno, J. L. (1934). *Who shall survive?* New York: Beacon Press.
- Mustchke, P. and Quaan Haase, A. (2004) "Collaboration and cognitive structures in social science research field," *Scientometric*, 4(2), pp. 47-66.
- Newman, M. E. J. (2000). "Models of the Small World: A Review", *Journal of Statistical Physics*, n°101, pp. 819-841.
- Niemeijer, R. (1973). *Some Applications of Notion of Density*, in Boissevain and Mitchell (eds).
- Page, L. and Brin, S. (1999). "The Anatomy of a Large-Scale Hypertextual Web Search Engine".
- Scott, J. P. (2000) *Social Network Analysis: A Handbook. Second edition*, Sage Publications.
- Snijders, T. A. B. (1981). "The Degree Variance", *Social Networks*, n° 3.

Anexo I. Resumen Conceptual

