

Red de coautorías del 2^{sd} International Conference on Industrial Engineering and Industrial Management – CIO2008

José I. Santos¹, José M. Galán¹, Luis R. Izquierdo¹, Ricardo del Olmo¹

¹ Área de Organización de Empresas. Dpto. de Ingeniería Civil. Escuela Politécnica Superior. Universidad de Burgos. C/Villadiego s/n 09001, Burgos. jmgalan@ubu.es, lrizquierdo@ubu.es, jisantos@ubu.es, rdelolmo@ubu.es

Palabras clave: redes sociales, redes científicas, innovación

Resumen

En este artículo ofrecemos los principales resultados del estudio de la red de coautorías del 2^{sd} International Conference on Industrial Engineering and Industrial Management – CIO2008 celebrado en Burgos. Formalizamos la red a partir de los datos del libro de actas y analizamos sus principales características estadísticas y estructurales que nos permiten inferir con rigor algunas propiedades de la red social de investigadores del área de Ingeniería de Organización en España. Este trabajo forma parte de una línea de investigación sobre redes sociales que pretende estudiar y formalizar las redes españolas de I+D+I.*

1. Introducción

El desarrollo económico en un mercado cada vez más internacionalizado y competitivo requiere de una apuesta inequívoca por la innovación. Las Universidades juegan un papel muy importante en los sistemas de innovación a través de los diferentes institutos, centros y grupos de investigación que alberga y tutela, y que desarrollan una amplia gama de proyectos y líneas de investigación. Cualquier estudio sobre los sistemas de innovación españoles no puede, por tanto, olvidar al sistema universitario. Dentro de la diversidad académica universitaria, la Ingeniería de Organización representa a nuestro juicio un ámbito de estudio especialmente interesante, puesto que sus líneas de investigación han estado siempre muy relacionadas con las innovaciones organizacionales, metodológicas e instrumentales surgidas en las empresas en los últimos años.

Una dimensión muy importante de cualquier grupo humano es el tejido de relaciones interpersonales en el que se desarrolla su actividad, lo que formalmente llamamos su red social. En el ámbito de la investigación y la innovación la red social es si cabe más importante, en la medida en que los procesos de creación de conocimiento se ven significativamente influenciados por cómo se intercambia y difunde el mismo. En consecuencia, querer comprender cómo se relacionan los miembros de una comunidad de investigadores resulta incuestionable. Hoy en día disponemos de un amplio y creciente conjunto de herramientas matemáticas y estadísticas que permiten formalizar una red social para tratar de inferir los patrones relacionales que mejor la describen, y que muy probablemente pueden explicar una parte significativa de su comportamiento.

* Este trabajo se deriva de la participación de sus autores en un proyecto de investigación financiado por el Ministerio de Educación y Ciencia con referencia TIN2008-06464-C03-02 y en un proyecto de investigación financiado por la Junta de Castilla y León con referencia BU034A08

Con esta motivación, hemos llevado a cabo un primer análisis de la red de profesores e investigadores de Ingeniería de Organización, quienes a través de sus colaboraciones en las comunicaciones presentadas en el congreso CIO2008 muestran una instantánea particular de su red social que puede ser estudiada formalmente. Además de realizar un estudio descriptivo de los parámetros que caracterizan esta red, hemos aprovechado los últimos avances en teoría de redes para la identificación de estructuras locales (Girvan y Newman 2002, Newman 2006).

2. Formalización de la red de coautorías del CIO2008

La comunidad de profesores e investigadores del ámbito de la Ingeniería de Organización en España definen una red social, que manifiesta su actividad de muy diversas formas (Santos *et al* 2006), particularmente en sus diferentes congresos anuales. Un aspecto importante en el estudio de cualquier red son el tipo y naturaleza de las relaciones que se producen entre sus miembros. Como en cualquier otra red social, estas relaciones son muy diversas y en muchas ocasiones difíciles de medir y estudiar. Sin embargo, la celebración de un congreso anual permite identificar algunas de estas relaciones, por ejemplo, a través del conjunto de comunicaciones presentadas en el congreso. Cada comunicación es el resultado de un trabajo de investigación y colaboración más o menos intenso entre los diferentes autores que la firman, consecuentemente las comunicaciones nos permiten formalizar el conjunto de relaciones entre todos los participantes del congreso.

La formalización de la red de coautorías del CIO2008 es semejante a la ya empleada en otros estudios de redes de coautorías científicas (Newman 2001a, 2001b). La información necesaria ha sido obtenida del libro de actas del congreso (Saíz *et al* 2008) que contiene datos suficientes sobre las características intrínsecas de los agentes, p.ej. su filiación, así como de sus colaboraciones en las diferentes comunicaciones presentadas en el congreso (todos los datos utilizados en este estudio se encuentran a publicados http://ingor.ubu.es/cio2008/red_coautorias/).

Nuestra red de coautorías no es más que una abstracción matemática compuesta por (1) un conjunto de nodos, cada uno representativo de un autor, y (2) un conjunto de vínculos entre pares de nodos, cada vínculo representativo de una comunicación en la que dos nodos son coautores de la misma. La red así definida es una red recíproca y ponderada puesto que cada enlace contiene un peso que cuantifica el número de comunicaciones en las que dos autores coinciden (la Figura 1 muestra la representación gráfica de una parte de la red).

En los siguientes apartados caracterizamos esta red a través de sus principales parámetros globales: tamaño, densidad, número de componentes, distancia geodésica media, coeficiente de clustering; así como otros representativos de la centralidad de los nodos: distribución de grado y de intermediación de los nodos. Además, realizamos un análisis clúster con el objeto de identificar estructuras locales. Para llevar a cabo este trabajo hemos empleado la aplicación de redes Pajek (Nooy et al, 2005), así como Mathematica para ejecutar algunos de los algoritmos de análisis clúster.

3. Características de la red

3.1. Tamaño y componentes

En el congreso CIO2008 se presentaron un total de 187 artículos firmados por 244 autores. Hemos considerado conveniente simplificar estos datos iniciales eliminando aquellas comunicaciones de un solo autor que evidentemente no aportan información relacional alguna. La red de coautorías resultante está compuesta por 226 autores y 157 artículos, que muestran 333 vínculos. La Tabla 1 recoge las principales características de la red.

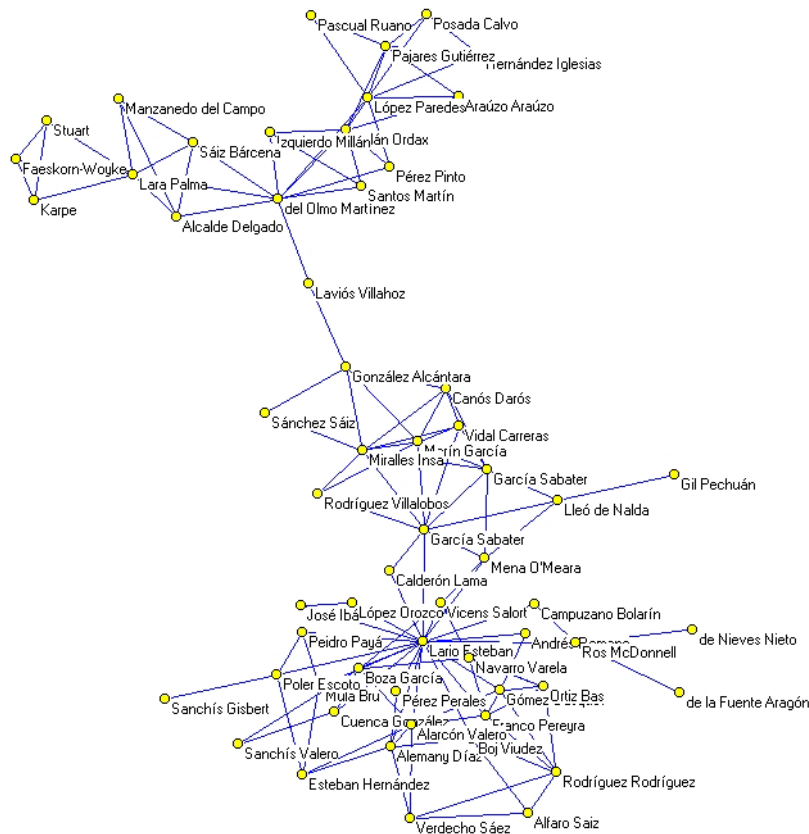


Figura 1. Representación gráfica del componente principal de la red de coautorías CIO2008.

Lo primero que debemos destacar es que la red tiene un número importante de componentes (38) que constituyen islas de autores sin relación entre ellos (la Tabla 2 recoge la distribución de tamaño de los principales componentes). Esta fragmentación de la red se explica en la fuente de datos utilizada, el congreso CIO2008 solo es una instantánea puntual que no recoge todo el espacio de relaciones entre los profesores e investigadores de la Ingeniería de organización (esta fragmentación se podría reducir completando los datos con las comunicaciones de congresos pasados).

Tabla 1. Principales propiedades de la red de coautorías CIO2008

| | |
|--------------------------------------|-------|
| Número de autores (nodos) | 226 |
| Número de comunicaciones | 187 |
| Número de vínculos | 333 |
| Promedio de comunicaciones por autor | 2,022 |
| Promedio de autores por comunicación | 2,6 |
| Densidad | 0,013 |
| Número de componentes | 38 |
| Tamaño componente mayor | 60 |
| Tamaño segundo componente | 15 |
| Coefficiente de Clustering | 0,581 |
| Distancia geodésica media | 3,912 |

Aun así, existe una significativa asimetría en los tamaños de los componentes. Por ejemplo, el componente mayor contiene a 60 autores mientras que el siguiente más grande agrupa solo a 15, lo que evidencia una significativa estructura relacional en los trabajos presentados en un único congreso.

Tabla 2. Tamaño y distancias geodésicas medias de los principales componentes (60% de la red)

| | | | | | | | | |
|---------------------|------|------|------|------|------|------|------|------|
| Tamaño | 60 | 15 | 14 | 14 | 12 | 7 | 7 | 6 |
| Distancia geodésica | 4,46 | 2,61 | 2,54 | 2,25 | 1,91 | 1,67 | 1,67 | 1,53 |

3.2. Proximidad (efecto “small world”) y clustering

Un parámetro frecuentemente utilizado para medir el grado de proximidad entre individuos de una red es el promedio de las distancias geodésicas entre todos los agentes. La distancia geodésica entre dos nodos se define como el número de vínculos de la secuencia de nodos (camino) más corta que los une. La red del CIO2008 presenta un valor de 3.91, si bien, al existir una significativa fragmentación, resulta más interesante observar las distancias geodésicas de cada componente (la Tabla 2 muestra las distancias geodésicas de los componentes principales). Así, la distancia media en el componente principal es de 4,46, es decir que en promedio menos de 5 personas separan la actividad de cualquier par de autores de este grupo.

Se suele hablar del efecto “small world” cuando en una red, aun cuando pudiera contener muchos nodos, la distancia geodésica media es sensiblemente pequeña, y próxima a la que cabría esperar en una red con los mismos nodos pero relacionados unos con otros de forma aleatoria. El conocido estudio de Milgram (1967) sugirió que la distancia geodésica media de la red social mundial de contactos no era mayor de 6, lo que posteriormente se ha popularizado con el nombre de los seis grados de separación. Podemos concluir diciendo que la red de coautorías del CIO2008 también presenta un efecto “small world”. Esta propiedad es importante desde el punto de vista de difusión y transferencia de conocimiento, pensemos en la red de coautorías como si de un sistema de información mediante contactos personales se tratase, la velocidad de difusión será tanto mayor cuanto menor distancia geodésica media presente la red.

Otra característica a tener en cuenta una red social es el fenómeno de transitividad, por el que frecuentemente en un grupo de personas los amigos de una de ellas son a su vez amigos. El coeficiente de clustering mide esta propiedad, y se expresa como la probabilidad de que dos personas que mantienen un enlace en la red con un tercero tengan a su vez un enlace entre ellos mismos.

En redes puramente aleatorias tiende a cero conforme el tamaño de la red se hace mayor. Sin embargo, en muchas redes sociales no ocurre lo mismo, y aunque la red tenga un número grande de nodos el valor del coeficiente de clustering es diferente de cero. Nuestra red de coautorías muestra un coeficiente de 0,581, lo que evidencia un grado de transitividad muy importante entre los autores (por ejemplo, sólo una de las redes de colaboraciones científicas analizadas por Newman (2001a) muestra un coeficiente de clustering mayor). El hecho de que muchas comunicaciones tengan tres a más coautores contribuye a que la transitividad de la red sea tan importante; razones de proximidad disciplinar, cercanía geográfica o afinidad personal pueden explicar este tipo de triadas en las colaboraciones científicas de los investigadores.

3.3. Distribución de los grados y de la intermediación (betweenness) de los nodos

Ningún estudio sobre la actividad investigadora de un grupo olvida una clasificación de los investigadores, de acuerdo a algún criterio de importancia. El criterio más sencillo, aunque evidentemente no muy adecuado a las características de un congreso, puede ser el número de trabajos presentados. Así, la Figura 2 muestra el histograma del número de comunicaciones por autor presentadas en el congreso. Destaca como un grupo pequeño de

investigadores (ver Tabla 3) firma un número destacado de comunicaciones, frente a la mayoría de participantes que firman 1 o 2 trabajos.

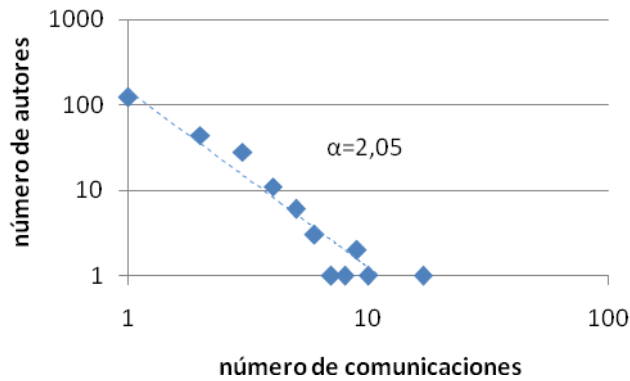


Figura 2. Histograma del número de comunicaciones por autor (los ejes están en escala logarítmica); se muestra además (línea punteada) la función de potencia ajustada a los datos junto con el valor del exponente estimado.

Si bien, un congreso como el CIO2008 no es más que una muestra puntual de la actividad investigadora en la Ingeniería de Organización, resulta interesante comprobar cómo el número de comunicaciones por autor se ajusta bastante bien a una ley de potencia, al igual que muchas distribuciones de publicaciones científicas estudiadas (Newman 2001a). Por tanto, la probabilidad de que un autor presente k comunicaciones puede expresarse mediante la ecuación (1), donde c y α son constantes.

$$p(x = k) = ck^{-\alpha} \quad (1)$$

Sin embargo, la información de un congreso puede darnos una información más útil desde el punto de vista relacional. Una tarea muy interesante en el análisis de redes es la de identificar aquellos individuos que desempeñan un papel central en el espacio de relaciones, lo que en teoría de redes se llaman medidas de centralidad. La forma más sencilla de discriminar entre los diferentes actores de una red es a través del grado de cada nodo, que se define para redes recíprocas como el número total de vínculos que comparten extremo con él.

No debemos confundir el grado de un nodo con el número de comunicaciones firmadas por un autor. En nuestra red de coautorías cada vínculo puede provenir de una o más comunicaciones, de hecho, una comunicación de n autores aporta $\binom{n}{2}$ nuevos vínculos a la red e incrementa el grado de cada nodo en $n-1$. La Figura 3 muestra la distribución del grado de la red de coautorías. En la medida en que las comunicaciones y los vínculos de la red están relacionados, no nos sorprende que la distribución del grado también siga una ley de potencia, al igual que el ranking de los principales autores por número de comunicaciones y grado sea muy parecido (ver Tabla 3).

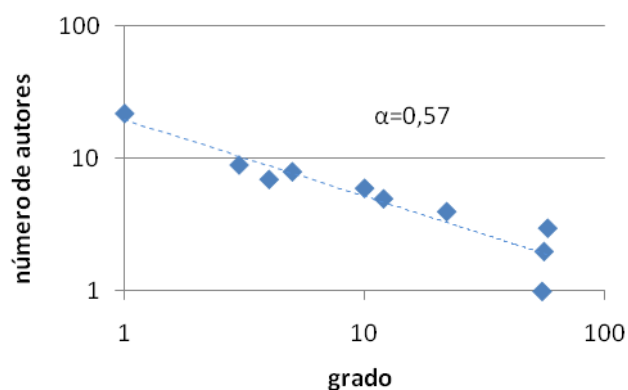


Figura 3. Histograma del grado de los nodos (los ejes están en escala logarítmica); se muestra además (línea punteada) la función de potencia ajustada a los datos junto con el valor del exponente estimado.

La teoría de redes nos proporciona una medida muy interesante de la importancia que juegan los individuos en una red, el llamado índice de intermediación (betweenness) de un nodo, que se calcula como la proporción de todos los caminos más cortos (distancia geodésica) entre pares de nodos de la red que lo incluyen. Si estamos interesados en los fenómenos de transferencia y difusión de información, este índice puede ser especialmente útil, puesto que aquellos individuos con índices más elevados, desempeñan un papel central en el proceso de difusión al ser intermediarios en un mayor número de caminos.

La Tabla 3 muestra el ranking de autores por índice de intermediación (último par de columnas). Podemos destacar cómo algunos autores que desde la perspectiva del número de comunicaciones habían pasado desapercibidos en el estudio, son importantes en la dimensión de la transferencia de información.

Tabla 3. Top-10 de autores por número de comunicaciones, grado e intermediación del CIO2008

| Nºcom. | Autor | Grado | Autor | Interm. | Autor |
|--------|--------------------|-------|---------------------|---------|--------------------|
| 17 | Lario Esteban | 22 | Lario Esteban | 0,0428 | Lario Esteban |
| 17 | García Sabater | 16 | García Sabater | 0,0350 | García Sabater |
| 10 | Marín García | 9 | del Olmo Martínez | 0,0313 | del Olmo Martínez |
| 9 | Rodríguez Monroy | 9 | Alarcón Valero | 0,0302 | González Alcántara |
| 7 | Vidal Carreras | 8 | Martínez Costa | 0,0293 | Laviós Villahoz |
| 6 | Carrasco Arias | 8 | de la Fuente García | 0,0155 | Miralles Insa |
| 6 | Miralles Insa | 8 | Miralles Insa | 0,0140 | Marín García |
| 6 | Gómez Gasquet | 8 | Rodríguez Monroy | 0,0084 | López Paredes |
| 5 | López Paredes | 8 | López Paredes | 0,0074 | Lara Palma |
| 5 | Palacios Fernández | 7 | Marín García | 0,0046 | Ros McDonnell |

4. Análisis clúster

Finalmente, vamos a centrar nuestra atención en la detección de estructuras locales en la red. Recientes estudios ponen de manifiesto el papel que juegan las estructuras locales -regularidades en la asociación de nodos entre el nivel micro y macro- en el comportamiento de la red en su conjunto (Newman 2006). Sin embargo, a diferencia de los análisis clusters tradicionales, la tarea de identificación de particiones dentro de una red no resulta trivial, y constituye hoy en día una de las áreas de estudio que más interés está despertando en la teoría y aplicación de las redes sociales (Porter *et al*, 2009).

Teniendo en cuenta el grado de fragmentación de la red CIO208, decidimos llevar a cabo el análisis clúster en el componente principal de la red (60 nodos).

4.1. K-núcleos

Una técnica muy sencilla para identificar grupos homogéneos son los k -núcleos. Bajo la hipótesis de que aquellos individuos que comparten alguna característica (similitud) deberían mostrar un grado de relación entre ellos mayor que con el resto de la red, podemos emplear el grado de los nodos para visualizar estas estructuras. Un subconjunto de nodos se dice que forman un k -núcleo si todos sus nodos están conectados al menos a k -nodos de dicho grupo. La Figura 4 muestra la distribución de los k -núcleos del componente principal (cada nodo está asociado al núcleo de mayor grado posible), junto con la representación gráfica del núcleo de mayor grado (4-núcleo).

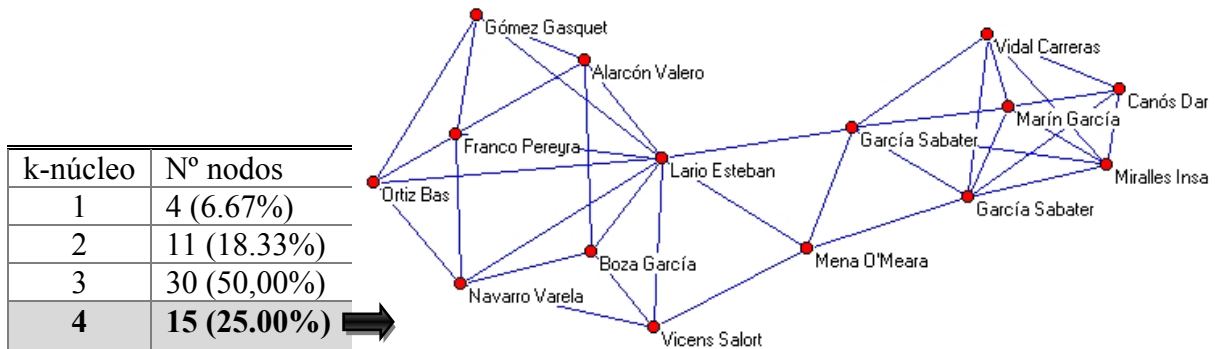
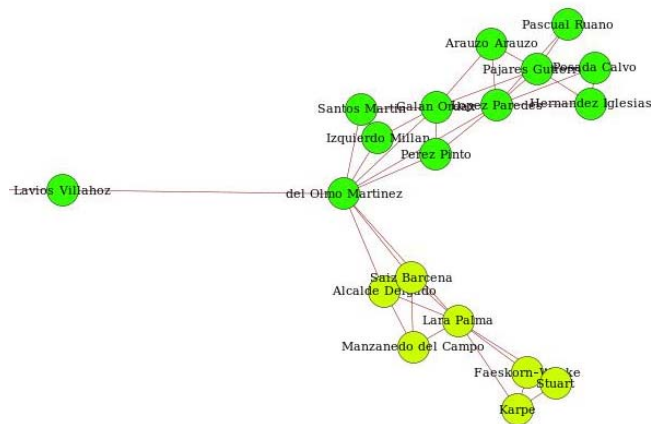


Figura 4. Histograma de la distribución de los k -núcleos de la red de participaciones.

Sin embargo, aunque el análisis de k -núcleos nos ha permitido detectar un grupo significativamente homogéneo, no nos ofrece una auténtica partición de la red en grupos disjuntos (pensemos que un k -núcleo siempre engloba a los núcleos de grado igual o mayor que k).

5. Comunidades

El principio empleado en la detección de los k -núcleos nos sirve para definir el concepto de comunidad. Una comunidad es un conjunto de nodos que comparten comparativamente un mayor número de vínculos entre ellos que con el resto de la red. Existe una gran diversidad de algoritmos de detección de comunidades, según cómo se defina lo que se entiende por “mayor número de vínculos” (Porter *et al* 2009); nosotros empleamos uno de los más habituales basado en el concepto de modularidad (Newman 2006). La Figura 5 muestra la partición en comunidades del componente principal. A diferencia de los k -núcleos, obtenemos una partición competa de la red, así como una mayor coherencia de los grupos identificados.



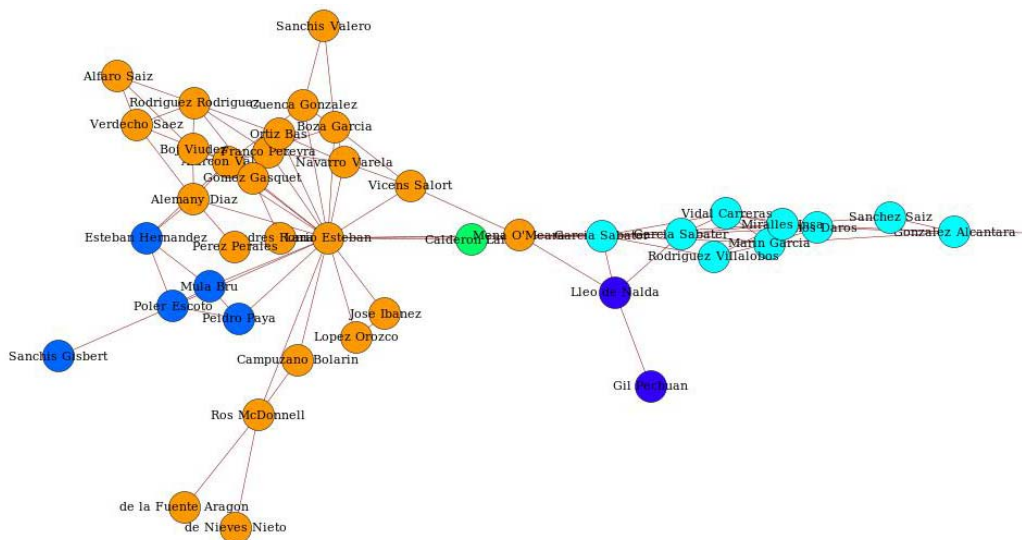


Figura 5. Partición del componente principal por comunidades (teniendo en cuenta el peso de las relaciones entre autores).

No cabe duda de que las comunidades detectadas en nuestra red son el reflejo de algún tipo de afinidad entre sus miembros. Una manera de estudiar esta afinidad subyacente es observar las características socio-académicas que comparten los autores de una comunidad. Por ejemplo, la información del libro de actas nos ha permitido conocer la filiación de los autores. Hemos comprobado el grado de relación entre las comunidades detectadas y la partición del componente principal de acuerdo a la institución de origen del autor, y mediante una sencilla prueba χ^2 concluimos rechazando la independencia de estos dos factores. Por tanto, los datos de la red de coautorías parecen apoyar la hipótesis intuitiva de que la proximidad geográfica condiciona significativamente la actividad investigadora en las universidades.

Para terminar, nos ha parecido ilustrativo mostrar las relaciones del componente principal desde el punto de vista de la filiación de los autores. La Figura 6 muestra una visión global de las universidades que integran este componente, así como una visión parcial de los miembros de cada una de ellas y sus relaciones con el resto de instituciones académicas. Destaca cómo los investigadores de la Universidad de Burgos juegan un papel central en el espacio de relaciones entre universidades, aunque aportan al componente menos autores y comunicaciones que otras universidades.

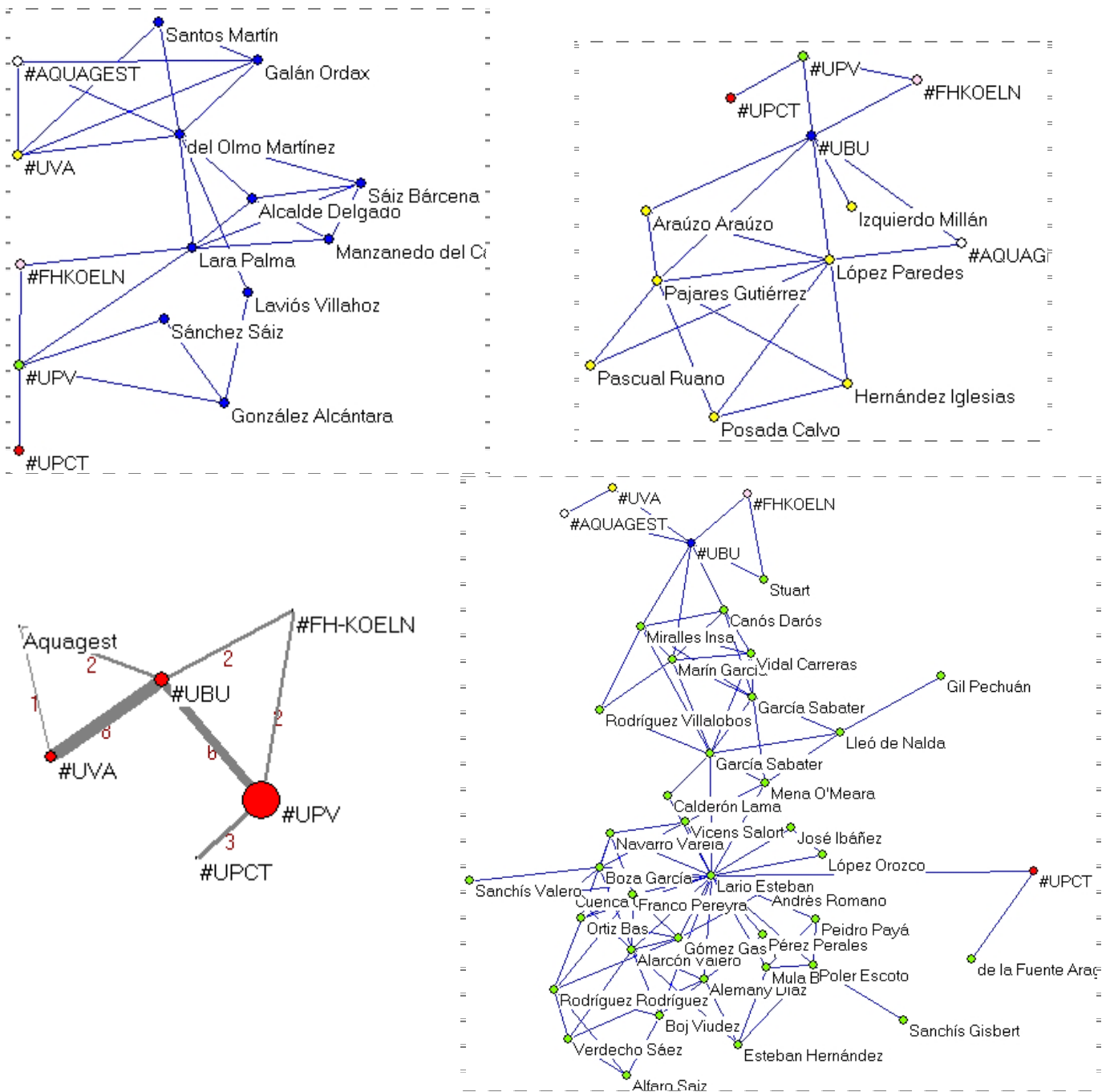


Figura 6. Representación global del componente principal por filiación de los autores (abajo izquierda): Universidad Politécnica de Valencia (UPV), Universidad Politécnica de Cartagena (UPCT), Universidad de Valladolid, Universidad de Burgos (UBU), Universidad de FH de Colonia (FH-KOELN) y empresa Aquagest. El tamaño de los nodos y vínculos son proporcionales a los pesos de las relaciones intragrupos e intergrupos respectivamente. Además, la figura contiene la representación contextual de los autores de la UBU (arriba izquierda), UVA (arriba derecha) y UPV (abajo derecha), donde se muestra simultáneamente la relaciones entre sus miembros y con el resto de instituciones.

6. Conclusiones

En este artículo hemos resumido los principales resultados del análisis formal de la red de coautorías del congreso CIO2008. Pensamos que este tipo de trabajos son necesarios, pues nos permiten conocer la estructura de relaciones de los agentes que integran un sistema complejo, en este caso la comunidad de investigadores de Ingeniería de Organización, como

paso previo para comprender su funcionamiento y proponer mecanismos de intervención eficaces. Este artículo se integra en una línea de investigación más general y ambiciosa que pretende analizar las redes españolas de investigación, principalmente vinculadas al mundo académico universitario, mediante la formalización y análisis de diferentes bases de datos sobre la actividad investigadora, como por ejemplo las comunicaciones de un congreso como el que hemos estudiado.

Referencias

Girvan, M. & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America* 99(12), pp. 7821-7826.

Newman, M. E. J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E* 64, 016131.

Newman, M. E. J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E* 64, 016132.

Newman, M. E. J. (2003). The structure and function of complex networks. *Siam Review* 45(2), pp. 167-256.

Newman, M. E. J. (2006). "Modularity and community structure in networks". *Proceedings of the National Academy of Sciences of the United States of America* 103(23), pp. 8577-8582.

Nooy, W., Mrvar, A., Batagelj, V. (2005). *Exploratory Social Network Analysis with Pajek*. Cambridge University Press.

Milgram, S. (1967). The small-world problem. *Psychology Today*, No. 2, pp. 60-67.

Porter, M.A., Onnela, J.P., Mucha, P.J. (2009). *Communities in Networks*. <http://www.citebase.org/abstract?id=oai:arXiv.org:0902.3788>

Saíz, L., Izquierdo, L.R. & Santos, J.I. (editores) (2008). *Insights on Current Organization Engineering*. *Proceedings of the Second International Conference on Industrial Engineering and Industrial Management - XI Congreso de Ingeniería y de Organización*. Burgos: University of Burgos.

Santos, J.I., del Olmo, R. & Pajares, J. (2006). Estudio de la red de participaciones en tribunales de tesis doctorales de organización y gestión de empresas en España. En *Actas del X Congreso de Ingeniería y Organización*, pp.183-184, Vicens, E. & Lario, F.C. Valencia: Universidad Politécnica de Valencia.