



## **FUNCIONAMIENTO, REPRESENTACIÓN Y COMPORTAMIENTO DE DIFERENTES ALGORITMOS FRENTE AL CÁLCULO DE UN SMALL WORLD EN CIENCIA**

***María Peñaranda-Ortega (1), Julia Osca-Lluch (2), Mayte López Ferrer (3), Cristina Civera-Mollá (4), Francisco Tortosa-Gil (5)***

*(1) Universidad de Murcia, España, mariap@um.es (2) CSIC, España, m.julia.osca@uv.es  
(3) CSIC, España, mayte.lopez@uv.es (4) Universidad de Valencia, España, cristina.civera@uv.es  
(5) Universidad de Valencia, España, francisco.m.tortosa@uv.es*

### **RESUMEN**

Para el estudio de redes científicas, y basándonos en las propuestas metodológicas de Watts y Strogatz, y las teorías de grafos, partimos de una serie de algoritmos que nos permitirán obtener diversas composiciones de Small Worlds sobre colaboración científica. Los algoritmos estudiados (algoritmo de Dijkstra, algoritmo de Bellman-Ford y algoritmo de Floyd), crean configuraciones espaciales diversas y con diferentes cualidades sobre los nodos que conformarían las publicaciones de los autores en ciencia. El comportamiento de estos algoritmos respecto de la calidad del nodo, superficie de la red y su relación o aislamiento con dependencia de la red de nodos, y otras cuestiones y variables, se valora para ver sus bondades y deficiencias respecto del estudio en la coautoría científica. Finalmente, y tras el análisis de las características de cada uno de los algoritmos, vemos que el más adecuado para el análisis de la coautoría científica es el algoritmo de Floyd, el cual presenta unos resultados más realistas para el estudio de las características propias de la colaboración en ciencia.

### **ABSTRACT**

To study scientific networks and based on methodologic proposals of Watts and Strogatz and graphos theory, we start from a sequence of algorithms allowing us to obtain some Small Worlds compositions on scientific collaboration. Studied algorithms (Dijkstra algorithms, Bellman-Ford algorithm and Floyd algorithm) create several configuration on space and different qualities on nodes that will include publications of scientific authors. These algorithms' behavior regarding their quality, network surface and their relationship or isolation on their network of node, or even other issues, is assessed to discover their defaults and virtues respect of the study on scientific coauthor. Finally, after the analysis o their features , we consider that the most appropriate for scientific coauthor is Floyd algorithm, which has the most realistic results on the study of collaboration features.

# IX CONGRESS CONGRESO ISKO-SPAIN ISKO-ESPAÑA

Valencia 11th, 12th, 13th March 2009 11, 12 y 13 de Marzo de 2009

New Perspectives for the organisation  
and dissemination of knowledge

Nuevas perspectivas para la difusión  
y organización del conocimiento



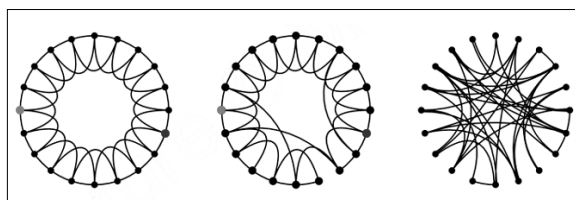
## **PALABRAS CLAVE:**

Colaboración científica, Teoría Small Worlds, Algoritmo de Floyd, Algoritmo de Bellman-Ford, Algoritmo de Dijkstra.

## INTRODUCCIÓN

Desde que Price (1973) pone en marcha la tradición bibliométrica para el estudio de la colaboración científica, hasta las propuestas sociométricas de análisis de Crane (1969, 1972), el estudio de la red de investigadores que publican en la Ciencia siempre se ha centrado en el establecimiento de la red total en la cual se hayan insertos estos autores, así como en el cálculo de la conectividad existente, determinando así los colegios invisibles que se conforman a lo largo del tiempo. Otra manera de estudiar la colaboración en base a colegios invisibles estaría basada en la técnica de estudio de los Small Worlds. La idea de conectar diferentes miembros de una comunidad que poseen alguna característica en común fue expuesta por Watts y Strogatz (1998) en un conocido artículo en la revista *Nature*, donde nos presentaron su modelo basado en los seis grados de separación. Desde este modelo, se propone que cualquier persona estaría conectada con el resto de seres humanos por un máximo de seis grados de separación. Por lo tanto, y siguiendo este modelo, un científico medianamente productivo estaría separado de cualquier otro por un máximo de seis intermediarios.

Watts y Strogatz se basan en la idea de los "seis grados de separación", la teoría social que Stanley Milgram propuso en 1960, según la cual cualquier ser humano está "separado" de cualquier otro a una distancia de, como máximo, seis personas intermediarias. Las claves para diferenciar entre unos mundos (Small Worlds) y otros, se encuentran en la denominada teoría de los Grafos, especie de redes tanto del tipo ordenado como estructurados al azar. Nos referimos a grupos cuyos miembros están conectados internamente entre sí pero que no se relacionan con otros grupos como no sea por otros conductos. En el otro extremo de las posibilidades de organización está la red totalmente caótica: cada individuo se comunica con otros individuos de la organización, pero de manera totalmente aleatoria. Se pueden establecer qué tipos de grafos necesitan más o menos enlaces.



**Figura 1.-** Paso de un grafo regular a grafo aleatorio: Small Worlds. (Fuente: Watts y Strogatz, 1998).

La figura 1 ilustra un procedimiento aleatorio de interpolación de nodos entre un círculo regular y una red aleatoria, sin alterar los vértices del gráfico. Se muestran tres realizaciones de este proceso, para diferentes valores de  $p$ , siendo éste el valor aleatorio. Si  $p=0$ , el círculo original no cambia (figura izquierda), mientras que si el valor de  $p=1$ , el gráfico se convierte en un aumento desordenado (figura derecha). El nodo es elegido



al azar. Para valores intermedios de  $p$ , el gráfico es una red small world muy cohesionada (figura central).

Según Marchiori y Latora (2000), podemos estudiar las características de un grafo mediante el grado promedio, la distribución del grado, la longitud del atajo, y la estimación del Coeficiente de Homogeneidad.

El grado promedio, " $k$ ", es definido como la media de enlaces que poseería cada nodo. Si un nodo  $i$  contiene  $k$  enlaces, su grado promedio es  $K_i$ . Al ser un grafo bidireccional, cada enlace tiene dos principios conectados a un nodo, si existen  $n$  enlaces en el grafo, entonces tendremos  $2n$  en el promedio total.

La distribución del grado,  $P(k)$ , nos informa de la probabilidad de que un nodo seleccionado al azar posea  $k$  enlaces.

La longitud media del atajo, denotada como  $l$ , es la distancia media entre un par de nodos. Ésta describe la conectividad que mantiene el sistema como una red.

Por último, el coeficiente de homogeneidad denotado como  $C$ , describe lo cercana que está entre sí la comunidad por medio de sus enlaces. Vendría a ser una estimación de la "densidad" que la red de nodos y enlaces puede llegar a mostrar.

El Coeficiente de Homogeneidad de un nodo  $i$  es la relación entre el número de enlaces actuales y el número de enlaces posibles entre un nodo  $i$  y sus nodos adyacentes. La estimación del Coeficiente de Homogeneidad no depende de  $N$ , esto es, del número de nodos o individuos en una red social, por lo que el modelo presentado de Watts y Strogatz es el que mejor describe una red social, con la ventaja de poder ser analizable (Watts, Dodds y Newman, 2002) y estructuradas en diferentes tipos.

Si focalizásemos nuestro ejemplo en la producción científica, diríamos que como cualquier comunidad que se relaciona tanto formal como informalmente, se encuentra la conformada por la Ciencia y las personas que trabajan en ella: los científicos. Teniendo en consideración que todos los investigadores de una u otra rama del saber conocen a otros investigadores, cada uno de ellos puede ocupar el puesto de un eslabón en una "cadena", la cual no distaría más de cuatro eslabones para poder unir a todos los autores de la ciencia norteamericana. Si incluimos a todos los autores e investigadores que existen, la cadena se hace más larga, pero no tanto como podemos imaginar (Adamic y Adar, 2003). En esta investigación usaremos el término Small Worlds para delimitar las relaciones de algún miembro de un colegio invisible científico, es decir, un grupo que entre sus miembros colaboran en una misma área científica y publican juntos, con otros miembros de otros grupos científicos. No sólo conocemos el aumento de la zona de influencia del grupo, sino que ésta es otra forma de explicar la configuración de los grupos científicos y las relaciones entre sus miembros. Dentro de nuestra propuesta



tratamos de conectar a todos los autores más prolíficos en un área científica concreta mediante los nexos de unión más cortos y que más autores engloben.

De esta forma, se comprueba finalmente que la Ciencia no es más que la mera unión de ramas científicas más alejadas o separadas entre sí, pero englobadas al fin y al cabo en el mismo plano (Peñaranda, 2004). Desde este punto de vista, los autores investigan y firman los trabajos juntos, estableciendo una estrecha relación entre sí, y quedan vinculados con otros autores que, no habiendo publicado directamente con ellos, sí lo han hecho con sus colaboradores más directos. Esta red de relaciones sociales establece grupos coherentes y perennes de trabajo, con intereses y finalidades bien definidas, tanto teórica como metodológicamente.

## **PROPIEDADES, COMPORTAMIENTO Y CÁLCULO DE DIFERENTES ALGORITMOS FRENTE A LA COMUNIDAD DE COLABORACIÓN CIENTÍFICA**

Nuestra propuesta se basa en analizar qué algoritmos favorecerán el análisis de las coautorías científicas en un área determinada de la Ciencia, tratando de conectar a todos los autores más prolíficos mediante los nexos de unión más cortos y que más autores engloben. Además, no sólo querríamos obtener información global sobre la red de conexiones establecidas en una disciplina determinada, sino también información específica y más cualitativa sobre los nodos que componen esa red, es decir, sobre los propios autores. Para calcular los caminos más cortos entre uno y otro científico, y teniendo todos los nodos en cuenta, sabiendo a que nodo pertenece cada investigador, hemos seleccionado los algoritmos que calculan de manera eficaz caminos más cortos entre nodos que puedan o no tener cierta unión entre sí. Concretamente, estos son el algoritmo de Dijkstra (Dijkstra, 1982), el algoritmo de Bellman-Ford (Bellman, 1958) y el algoritmo de Floyd (Floyd, 1962). Hemos buscado entre las aplicaciones que permiten estos algoritmos todas las combinaciones posibles entre nodos y los atajos que mejor describan el Small World científico.

A continuación presentamos los algoritmos estudiados y su comportamiento ante la búsqueda de caminos entre nodos, y concretamente el camino más corto existente entre un par de nodos entre autores para analizar la colaboración científica, es decir un grafo con nodos limitados y con conexiones limitadas, lo que supone un verdadero “Small World”. Un grafo con  $n$  nodos puede simbolizarse por una matriz  $C$  ( $n \times n$ ), donde cada fila/columna representa un nodo. Cualquier combinación  $(i,j)$  es un par de nodos del grafo, excepto los pares  $(i,i)$ , o elementos de la diagonal principal de  $C$ . Por tanto,  $C$  tiene un total de  $n(n-1)$  pares de nodos. Por lo tanto, estamos interesados en conocer los caminos más cortos entre todos los  $n(n-1)$  pares de vértices de un grafo dirigido  $G$  formado por  $n$  vértices.



Considerando un grafo con  $n$  nodos, sabemos que posee  $n(n-1)$  pares de nodos debido a que tendríamos en total  $n^2$  nodos de los cuales restaremos  $n$  ya ésta es la cantidad de pares de nodos de la forma  $(i, i)$ , para todo  $i$ ; es decir, estos caminos no representan ningún par de nodos, sino que son el propio camino hacia el mismo nodo. Por ejemplo, el par de nodos  $(4, 4)$  no simboliza ningún par de nodos y por lo tanto no es válido para el cálculo de los caminos de dicho grafo.

## A) Algoritmo de Dijkstra

El algoritmo de Dijkstra, calcula los caminos más cortos entre un vértice y los demás. Para que este algoritmo pueda construir su cálculo se necesitan unas suposiciones tales como que las longitudes de todos los arcos deben de ser números enteros y positivos; además, la red ha de ser dirigida, es decir, los arcos que se crean en el cálculo de un nodo a otro solamente se pueden recorrer en una dirección.

La idea subyacente en este algoritmo consiste en ir explorando todos los caminos más cortos que parten del vértice origen y que llevan a todos los demás vértices. Cuando se obtiene el camino más corto desde el vértice origen, al resto de vértices que componen el grafo, el algoritmo se detiene. No funciona en grafos con aristas de coste negativo (al elegir siempre el nodo con distancia menor, pueden quedar excluidos de la búsqueda nodos que en próximas iteraciones bajarían el coste general del camino al pasar por una arista con coste negativo).

Su pseudocódigo es:

Dijkstra (N,s)

Inicializar

```
for cada v perteneciente a V[N]
  do d[v] = infinito
  p[v] = nulo
d[s] = 0
```

S = vacío

Q = V[N]

mientras Q no vacío

```
do u = nodo v con min d[v]
```

S = S unión u \se añade al conjunto de nodos finalizados

```
for cada v perteneciente Adyacente u
```



## Relajación

if  $d[v] > d[u] + w(u,v)$  then

$d[v] = d[u] + w(u,v)$

$p(v) = u$

Este algoritmo construye una sucesión de subconjuntos del conjunto de nodos totales (N) que posee un Small World, de manera que cumple dos condiciones. Primero, que las distancias que existen en el primer conjunto siempre serán mayores que las de los siguientes subconjuntos creados, ya que se reduce la distancia entre los nodos (número de arcos) en cada subconjunto calculado. Segundo, el funcionamiento de este algoritmo provoca el cálculo sistemático de todos los caminos posibles, hasta llegar a obtener el mínimo camino necesario para recorrer desde un nodo a otro.

La carencia de este algoritmo es que calcula caminos más cortos de un nodo a otro, cuando el comportamiento de un Small World de coautoría científica requiere que se calculen distancias mínimas de cualquier nodo a cualquier otro. Si implementásemos el pseudocódigo arriba indicado, tan sólo podríamos analizar un autor, en detrimento del resto.

## B) Algoritmo de Bellman-Ford

El algoritmo de Bellman-Ford genera los caminos mínimos desde un nodo origen de un grafo ponderado al resto de nodos del mismo. Soluciona el problema de la ruta más corta o camino mínimo desde un nodo origen, de un modo más general que el Algoritmo de Dijkstra, ya que permite valores negativos en los arcos. El algoritmo devuelve un valor booleano (dicotómico), si encuentra un circuito o lazo de peso negativo. En caso contrario calcula y devuelve el camino mínimo con su coste.

Para cada vértice  $v$  perteneciente a  $V$ , se mantiene el atributo  $d[v]$  como cota superior o coste del camino mínimo desde el origen  $s$  al vértice  $v$ .

A continuación se muestra el pseudocódigo del Algoritmo:

Bellman-Ford (N,s)

Inicializar

for cada  $v$  perteneciente a  $V[N]$

do  $d[v] = \text{infinito}$

$p[v] = \text{nulo}$

$p[s] = 0$

for  $i=1$  to  $V[N]-1$

do for cada arco  $(u,v)$  perteneciente a  $A[N]$

Relajación

if  $d[v] > d[u] + w(u,v)$  then



```
d[v] = d[u] + w(u,v)
p(v) = u
for cada arco (u,v) chequea lazo de peso negativo
do if d[v] > d[u] + w(u,v) then
    return FALSO 'el algoritmo no converge
return VERDADERO
```

El algoritmo de Dijkstra logra esta misma tarea con un coste de tiempo menor, pero requiere que los pesos de las aristas no sean negativos. Es por eso, que el algoritmo de Bellman-Ford se utiliza únicamente cuando hay aristas negativas presentes en el grafo. Por tanto, tampoco el algoritmo de Bellman-Ford se muestra cualificado para analizar un Small World con las características que requerimos. Este algoritmo es mejor que el algoritmo de Dijkstra porque admite longitudes negativas en los arcos, por lo que se podría utilizar para la incomunicación científica entre dos o varios autores. No obstante, posee el mismo problema que el anterior algoritmo presentado: únicamente calcula los caminos más cortos de un nodo al resto, obviando el conjunto.

### C) Algoritmo de Floyd

Este es un algoritmo que analiza el grafo de manera que encuentra el camino mínimo de forma dirigida y ponderada. El algoritmo encuentra el camino entre todos los pares de vértices en una única ejecución. Compara todos los posibles caminos a través del grafo entre cada par de vértices. El algoritmo es capaz de hacer esto con sólo  $V^3$  comparaciones. Lo hace mejorando paulatinamente una estimación del camino más corto entre dos vértices, hasta que se sabe que la estimación es óptima. Por tanto, es un buen calculador para una comunidad regulada por las características que dominan la publicación científica. Para este método se utiliza una matriz de costes. Además, puede ser utilizado para detectar todos los circuitos dirigidos de un grafo, y en particular circuitos dirigidos con peso total negativo.

Una carencia que hemos encontrado desde este tipo de estudio es que, para análisis cualitativos de intensidad en la productividad, como conocer el tipo de unión que afecta a los nodos y la potencia que tendría a nivel productivo, el algoritmo de Floyd no puede discriminar las posibles diferencias entre autores con mayor o menor relevancia productiva entre ellos, por lo que, en un nivel superior de análisis, y referido al análisis de científicos que publican sus obras y no a números de nodos que estima un computador, quedaría patente esta limitación cualitativa de la red (Adamic y Adar, 2003). Si se introdujera un cambio en el cual se especificara el tipo de enlace que une cada nodo respecto del resto conforme su tamaño en productividad, esta circunstancia sería obviada.

A continuación se muestra el pseudocódigo del Algoritmo de Floyd:



Inicializar

$C^0 = A$  ' matriz de distancias = matriz de arcos

si  $i=j$  o  $c_{ij} = \infty$  entonces  $z_{ij} = \text{nulo}$  sino  $z_{ij} = j$  'matriz de caminos  
for  $k = 1$  to  $V$

for  $i = 1$  to  $V$

for  $j = 1$  to  $V$

$c_{ij} = \min(c_{ij}, c_{ik} + c_{kj})$

si  $\min = c_{ik} + c_{kj}$  entonces

$z_{ij} = z_{kj}$

$c_{ij} = c_{ik} + c_{kj}$

fin

Realizamos un ejemplo de cálculo de Algoritmo de Floyd, al ser el que mejor se adapta a nuestro propósito, o sea, es el que más información nos puede facilitar para el cálculo de un Small World de colaboración científica. Mostramos un ejemplo de grafo para calcular la distancia más corta entre cualquier par de nodos mediante la aplicación del Algoritmo de Floyd (Ver Figura 2: Ejemplo de grafo). Se muestra un grafo que contiene cinco nodos:  $N = \{N_1, N_2, N_3, N_4, N_5\}$  y que contiene un total de ocho enlaces:  $V = \{V_{12}, V_{24}, V_{25}, V_{32}, V_{35}, V_{41}, V_{42}, V_{54}\}$

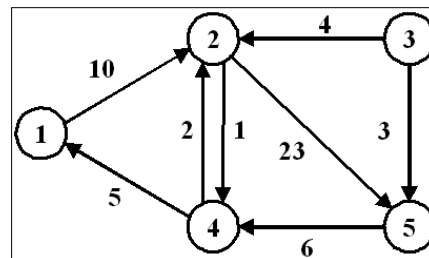


Figura 2.- Ejemplo de grafo

Creamos la matriz  $C^0$  en la que guardamos las distancias entre todos los pares de nodos, en caso de que dos nodos no estén conectados por un enlace, la distancia es  $\infty$  para esa matriz.

$$C^0 = \begin{pmatrix} 0 & 10 & \infty & \infty & \infty \\ \infty & 0 & \infty & 1 & 23 \\ \infty & 4 & 0 & \infty & 3 \\ 5 & 2 & \infty & 0 & \infty \\ \infty & \infty & \infty & 6 & 0 \end{pmatrix}$$



La matriz  $Z^0$  queda configurada de la siguiente manera: Si la posición  $c_{ij}^0$  es igual a  $\infty$ , lo cual indica la no conexión directa entre los nodos  $i$  y  $j$ ,  $z_{ij}^0$  será igual a 0. Mientras que si  $c_{ij}^0$  es distinto de  $\infty$ ,  $z_{ij}^0$  será igual a  $j$ , que es la posición por columnas en la matriz. Este es el paso 1 del algoritmo anteriormente definido.

$$Z^0 = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 4 & 5 \\ 0 & 2 & 3 & 0 & 5 \\ 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix}$$

Hacemos  $k = 0$ , donde  $k$  la haremos variar para todos los nodos. Este es el paso 2 del algoritmo de Floyd. Hacemos  $k = k+1 = 0+1 = 1$ , y para todo  $i \neq 1$  tal que  $c_{i1} \neq \infty$ , y para todo  $j \neq 1$  tal que  $c_{1j} \neq \infty$  (este es el paso 3 del algoritmo): calculamos  $M$  como el mínimo entre  $c_{ij}$  y  $c_{i1} + c_{1j}$  y si  $M$  es menor que  $c_{ij}$  realizamos  $z_{ij} = z_{i1}$  y  $c_{ij} = M$ . Como resultado de este paso obtenemos las siguientes matrices. Estos son los pasos 3.1 y 3.2. del algoritmo que seguimos.

$$C^1 = \begin{pmatrix} 0 & 10 & \infty & \infty & \infty \\ \infty & 0 & \infty & 1 & 23 \\ \infty & 4 & 0 & \infty & 3 \\ 5 & 2 & \infty & 0 & \infty \\ \infty & \infty & \infty & 6 & 0 \end{pmatrix} \quad Z^1 = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 4 & 5 \\ 0 & 2 & 3 & 0 & 5 \\ 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix}$$

Como los elementos de la diagonal principal de la matriz  $C^1$  son todos mayores o iguales que 0, continuamos al siguiente paso. Este es el paso 4.1 del algoritmo. Como  $k$  es distinto de  $n$ , siendo  $n$  el número de vértices o nodos de nuestro grafo, que es 5, volvemos al paso 3 como indica el paso 4.3. Así hacemos  $k = k + 1 = 1 + 1 = 2$ , y para todo  $i \neq 2$  tal que  $c_{i2} \neq \infty$ , y para todo  $j \neq 2$  tal que  $c_{2j} \neq \infty$  (este es el paso 3 del algoritmo): calculamos  $M$  como el mínimo entre  $c_{ij}$  y  $c_{i2} + c_{2j}$  y si  $M$  es menor que  $c_{ij}$  hacemos  $z_{ij} = z_{i2}$  y  $c_{ij} = M$ . Como resultado de este paso obtenemos las siguientes matrices. Estos son los pasos 3.1 y 3.2. del algoritmo que seguimos.

$$C^2 = \begin{pmatrix} 0 & 10 & \infty & 11 & 33 \\ \infty & 0 & \infty & 1 & 23 \\ \infty & 4 & 0 & 5 & 3 \\ 5 & 2 & \infty & 0 & 25 \\ \infty & \infty & \infty & 6 & 0 \end{pmatrix} \quad Z^2 = \begin{pmatrix} 1 & 2 & 0 & 2 & 2 \\ 0 & 2 & 0 & 4 & 5 \\ 0 & 2 & 3 & 2 & 5 \\ 1 & 2 & 0 & 4 & 2 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix}$$



Como los elementos de la diagonal principal de la matriz  $C^2$  son todos mayores o iguales que 0, continuamos al siguiente paso. Este es el paso 4.1 del algoritmo. Como  $k$  es distinto de  $n$ , siendo  $n$  el número de vértices de nuestro grafo, volvemos al paso 3 como indica el paso 4.3. Así hacemos  $k = k + 1 = 2 + 1 = 3$ , y para todo  $i \neq 3$  tal que  $c_{i3} \neq \infty$ , y para todo  $j \neq 3$  tal que  $c_{3j} \neq \infty$  (este es el paso 3 del algoritmo): calculamos  $M$  como el mínimo entre  $c_{ij}$  y  $c_{i3} + c_{3j}$  y si  $M$  es menor que  $c_{ij}$  hacemos  $z_{ij} = z_{i3}$  y  $c_{ij} = M$ . Como resultado de este paso obtenemos las siguientes matrices. Estos son los pasos 3.1 y 3.2. del algoritmo que seguimos.

$$C^3 = \begin{pmatrix} 0 & 10 & \infty & 11 & 33 \\ 6 & 0 & \infty & 1 & 23 \\ 10 & 4 & 0 & 5 & 3 \\ 5 & 2 & \infty & 0 & 25 \\ 11 & 8 & \infty & 6 & 0 \end{pmatrix} \quad Z^3 = \begin{pmatrix} 1 & 2 & 0 & 2 & 2 \\ 4 & 2 & 0 & 4 & 5 \\ 2 & 2 & 3 & 2 & 5 \\ 1 & 2 & 0 & 4 & 2 \\ 4 & 4 & 0 & 4 & 5 \end{pmatrix}$$

Como los elementos de la diagonal principal de la matriz  $C^3$  son todos mayores o iguales que 0, continuamos al siguiente paso. Este es el paso 4.1 del algoritmo. Como  $k$  es distinto de  $n$ , siendo  $n$  el número de vértices de nuestro grafo, volvemos al paso 3 como indica el paso 4.3. Así hacemos  $k = k + 1 = 3 + 1 = 4$ , y para todo  $i \neq 4$  tal que  $c_{i4} \neq \infty$ , y para todo  $j \neq 4$  tal que  $c_{4j} \neq \infty$  (este es el paso 3 del algoritmo): calculamos  $M$  como el mínimo entre  $c_{ij}$  y  $c_{i4} + c_{4j}$  y si  $M$  es menor que  $c_{ij}$  hacemos  $z_{ij} = z_{i4}$  y  $c_{ij} = M$ . Como resultado de este paso obtenemos las siguientes matrices. Estos son los pasos 3.1 y 3.2. del algoritmo que seguimos.

$$C^4 = \begin{pmatrix} 0 & 10 & \infty & 11 & 33 \\ 6 & 0 & \infty & 1 & 23 \\ 10 & 4 & 0 & 5 & 3 \\ 5 & 2 & \infty & 0 & 25 \\ 11 & 8 & \infty & 6 & 0 \end{pmatrix} \quad Z^4 = \begin{pmatrix} 1 & 2 & 0 & 2 & 2 \\ 4 & 2 & 0 & 4 & 5 \\ 2 & 2 & 3 & 2 & 5 \\ 1 & 2 & 0 & 4 & 2 \\ 4 & 4 & 0 & 4 & 5 \end{pmatrix}$$

Como los elementos de la diagonal principal de la matriz  $C^4$  son todos mayores o iguales que 0, continuamos al siguiente paso. Este es el paso 4.1 del algoritmo. Como  $k$  es distinto de  $n$ , siendo  $n$  el número de vértices de nuestro grafo, volvemos al paso 3 como indica el paso 4.3. Así hacemos  $k = k + 1 = 4 + 1 = 5$ , y para todo  $i \neq 5$  tal que  $c_{i5} \neq \infty$ , y para todo  $j \neq 5$  tal que  $c_{5j} \neq \infty$  (este es el paso 3 del algoritmo): calculamos  $M$  como el mínimo entre  $c_{ij}$  y  $c_{i5} + c_{5j}$  y si  $M$  es menor que  $c_{ij}$  hacemos  $z_{ij} = z_{i5}$  y  $c_{ij} = M$ .



Como resultado de este paso obtenemos las siguientes matrices. Estos son los pasos 3.1 y 3.2. del algoritmo que seguimos.

$$C^5 = \begin{pmatrix} \infty & 10 & \infty & 11 & 33 \\ 6 & 0 & \infty & 1 & 23 \\ 10 & 4 & 0 & 5 & 3 \\ 5 & 2 & \infty & 0 & 25 \\ 11 & 8 & \infty & 6 & 0 \end{pmatrix} \quad Z^5 = \begin{pmatrix} 1 & 2 & 0 & 2 & 2 \\ 4 & 2 & 0 & 4 & 5 \\ 2 & 2 & 3 & 2 & 5 \\ 1 & 2 & 0 & 4 & 2 \\ 4 & 4 & 0 & 4 & 5 \end{pmatrix}$$

Como los elementos de la diagonal principal de la matriz  $C^5$  son todos mayores o iguales que 0, continuamos al siguiente paso. Este es el paso 4.1 del algoritmo. Ahora  $k$  es igual a  $n$ , y en este caso, como manifiesta el paso 4.2 del algoritmo, hemos obtenido las matrices óptimas que contienen los caminos más cortos entre todos los pares de nodos (matriz  $C^5$ , que a partir de ahora llamaremos  $C'$ ) y la matriz que da el primer índice posterior al camino (matriz  $Z^5$ , que a partir de ahora llamaremos  $Z'$ ).

Para mayor comprensión del uso de estas matrices, vamos a realizar un ejemplo calculando el camino más corto entre los nodos 3 y 1. Este camino comienza en el nodo 3, para saber cual es el siguiente nodo miraremos en la matriz  $Z'$  la posición (3,1), que será el elemento  $z_{31} = 2$ , con lo que para llegar del nodo 3 al nodo 1 tendremos que pasar por el 2, miramos el elemento  $z_{21} = 4$ , con lo que para llegar del nodo 2 al nodo 1 hay que pasar por el 4, miramos ahora la posición  $z_{41} = 1$ , con lo que hemos llegado al final del camino. Entonces el camino más corto para ir del nodo 3 al nodo 1 es el siguiente, el nodo inicial es el 3, seguido del nodo 2, seguido del nodo 4 y por último el nodo 1, y la longitud de ese camino es la posición  $c_{31} = 10$ , que es la longitud del camino más corto para llegar al nodo 1 desde el nodo 3.

## CONCLUSIONES

Hemos comprobado la implementación de varios algoritmos en el estudio de la colaboración científica. Concretamente, el algoritmo de Floyd simplifica enormemente el cálculo de longitud entre rutas de nodos, lo cual es muy beneficioso en el análisis de la colaboración científica y la visibilidad de sus autores. El algoritmo de Floyd es una herramienta precisa en el cálculo de los caminos más cortos a trazar para el análisis de las publicaciones entre diferentes autores, incluso desde diferentes materias o áreas de conocimiento, ya que como se ha mostrado, permite analizar una enorme red de nodos (Peñaranda, López, Quiñones y López, 2006). La entrada en un mundo "Small Worlds" de coautoría científica vendría a ser relativamente fácil si se cumplen ciertos criterios de publicación y firmas conjuntas con otros autores, aunque la cadena sería relativamente



extensa si analizásemos autores que han publicado en reducidas ocasiones. Una desventaja, no obstante, es que cuando se detecta una colaboración, permanece a lo largo del tiempo, tanto si esa relación científica tuvo su fin o continúa, ya que el algoritmo de Floyd mantiene esa información, por lo que para el tipo de crecimiento que tuvo esa comunidad se debería de parcelar la observación del Small World tomando los datos en cortes temporales, dotándolo así de cierta dinamicidad.

Otra carencia del modelo es que no podemos saber la “calidad de los nodos”, o dicho de otra manera, el tipo de unión que afecta a los autores y la potencia que tendría a nivel productivo. Un autor “A” puede tener distar un nexo de unión con un autor “B” que publica una vez con él, pero mantener una relación mucho más intensa y prolífica con un colaborador directo “C” y con su colaborador cercano “D” que sería un nodo con valor dos. En este sentido, los Small Worlds no pueden diferenciar entre las posibles diferencias de los nodos “B” y “C” respecto de “A”. Y en el caso de la colaboración científica, estas diferencias son reales, por lo que, en un nivel superior de análisis, y referido al análisis de científicos que publican sus obras y no a números de nodos que estima un computador, quedaría patente esta limitación cualitativa de la red (Adamic y Adar, 2003).

Al basarnos en la Teoría Small Worlds, no intentamos reflejar todos los autores que publican con los más importantes, sino la vía más corta en el tiempo y en número de académicos, de enlazar a los autores más eminentes de la revista. Esta es una forma de entender un colegio invisible como una entidad con vida propia, tal y como observamos, en forma de redes neuronales, de “webrings” o páginas de Internet, en la que la práctica totalidad de los científicos, aunque pertenezcan a ciencias no afines, estarían comunicados entre sí. No obstante, una de las limitaciones que encontramos al aplicar los “Small Worlds” a la colaboración científica, es que nunca podrán estar representados todos los científicos. En especial, aquellos que no tengan colaboraciones representativas en los documentos científicos que hayan publicado, y aquellos que únicamente lo hayan hecho en revistas con un índice de impacto y una visibilidad poco representativas. Así, la entrada en un mundo “Small Worlds” vendría a ser relativamente fácil si se cumplen ciertos criterios de publicación y firmas conjuntas con otros autores, e imposible si no se ha publicado o se ha hecho de forma individual.

Finalmente, hemos comprobado que utilizando el algoritmo de Floyd para el cálculo de nodos particulares, no sólo conseguiremos obtener la globalidad de un Small World, sino también información mucho más detallada respecto del mismo. Este tipo de análisis se muestra idóneo para el análisis de colaboraciones científicas que requieran ser descritas y analizadas en profundidad, tal como el desarrollo de teorías en el tiempo por un grupo de científicos, el análisis de un autor y sus discípulos, o el estudio minucioso de la literatura de una revista o campo científico determinado.

Desde este punto de vista, la aplicación del algoritmo de Floyd en el examen de un colegio invisible, garantiza un análisis minucioso de todos sus componentes,



complementando de esta manera los datos resultantes de cálculos más globales, como por ejemplo el grado promedio o la distribución del atajo entre nodos. Los autores que conforman los colegios invisibles pasan de ser meros “nexos de unión” entre literatura científica para tener una representación más cualitativa en el estudio de los grafos, algo de lo que se carecía anteriormente.

## REFERENCIAS BIBLIOGRÁFICAS

ADAMIC, Lada A. y ADAR, Eytan. Friends and neighbors on the web. *Social Networks*, 25, 211-230, 2003.

BELLMAN, Richard. On a routin problem. *Quarterly of Applied Mathematics*, 16, 87-90, 1958.

CRANE, Diane. *Invisible colleges: Diffusion of knowledge in scientific communities*. University of Chicago Press: Chicago, 1972. ISBN: 0-226-11857-6.

CRANE, Diane. Social structure in a group of scientist: a test of the “invisible college” hypothesis. *American Sociology Review*, 34, 335-352, 1969.

DIJKSTRA, Edsger. *Selected writings on computing: A personal perspective, texts and monographs in computer science*. Springer-Verlag: Amsterdam, 1982. ISBN: 0-387-90652-5.

FLOYD, Robert W. Algorithm 97: Shortest path. *Communications of the ACM*, 5(6), 345.

MARCHIORI, Massimo y LATORA, Vito. Harmony in the Small World. *Physica A*, 285, 539-546, 2000.

PEÑARANDA ORTEGA, María. *La colaboración científica en la psicología social y de la personalidad: Análisis bibliométrico del Journal of Personality and Social Psychology*. Servicio de publicaciones de la Universidad de Murcia: Murcia, 2004. ISSN 84-8371-491-4.

PEÑARANDA ORTEGA, María; LÓPEZ SERRANO, Rafael; QUIÑONES VIDAL, Elena y LÓPEZ GARCÍA, Juan José. Los Small Worlds y el algoritmo de Floyd: Una manera de estudiar la colaboración científica. *Psicothema*, 18, 78-83, 2006.

PRICE, Derek J. Solla. *Hacia una ciencia de la ciencia*. Barcelona: Ariel, 1973. ISBN: 84-344-0739-6.

# IX CONGRESS CONGRESO ISKO-SPAIN ISKO-ESPAÑA

Valencia 11th, 12th, 13th March 2009 11, 12 y 13 de Marzo de 2009

New Perspectives for the organisation and dissemination of knowledge Nuevas perspectivas para la difusión y organización del conocimiento



WATTS, Duncan J. y STROGATZ, Steve H. Collective dynamics of small-world networks. *Nature*, 393, 440-442, 1998.

WATTS, Duncan J.; DODDS, Peter S. y NEWMAN, M. Elizabeth. Identity and search in social networks. *Science*, 296, 1302-1305, 2002.