



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 285 564**

51 Int. Cl.:  
**H04L 12/24** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **04808898 .3**

86 Fecha de presentación : **29.12.2004**

87 Número de publicación de la solicitud: **1700421**

87 Fecha de publicación de la solicitud: **13.09.2006**

54 Título: **Procedimiento para gestionar redes analizando su interconectividad.**

30 Prioridad: **30.12.2003 NO 20035852**

45 Fecha de publicación de la mención BOPI:  
**16.11.2007**

45 Fecha de la publicación del folleto de la patente:  
**16.11.2007**

73 Titular/es: **Telenor ASA**  
**Snaroyveien 30**  
**1331 Fornebu, NO**

72 Inventor/es: **Canright, Geoffrey;**  
**Engo-Monsen, Kenth y**  
**Weltzien, Asmund**

74 Agente: **Carpintero López, Francisco**

**ES 2 285 564 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

**DESCRIPCIÓN**

Procedimiento para gestionar redes analizando su interconectividad.

**5 Campo de la invención**

La presente invención se refiere a un conjunto de procedimientos para gestionar redes (tanto redes lógicas como físicas), en diversas áreas. Más particularmente la presente invención describe un procedimiento para analizar una red, donde la red está compuesta por diversos nodos de red conectados mediante enlaces.

**10 Antecedentes de la invención**

Casi cualquier estructura social o física donde las entidades individuales se enlazan juntas mediante cualquier clase de relación puede analizarse desde una perspectiva de red, sean grupos sociales, rutas aéreas, o grupos de ordenadores. Las redes son objetos interesantes. Tienen una gran cantidad de estructura, pero al mismo tiempo son sencillas: están compuestas, en su forma más sencilla, únicamente por nodos, conectados mediante enlaces. La idea abstracta de una red, o gráfico -el término se usa de forma intercambiable- también es muy útil en la modelización de las estructuras observadas en el mundo real. Los ejemplos incluyen: redes sociales, redes de comunicación, la red mundial (World Wide Web), redes metabólicas y genéticas en sistemas biológicos, redes alimentarias, redes de enfermedades, y redes de energía. Resumiendo, una red es una estructura abstracta sencilla, no trivial, fascinante por derecho propio, y también altamente trascendente para muchas ramas de la ciencia y la tecnología.

Dentro del área de las telecomunicaciones, durante mucho tiempo se han establecido teorías respecto a la gestión de la red y las estructuras de red. Es de crucial importancia entender una red. La eficacia de funcionamiento y mantenimiento de dicha red de telecomunicación dependerá en gran medida del conocimiento de la red en cuestión. Es importante tanto con respecto al tiempo medio entre fallos, así como con respecto a la propagación de la avería, tal como virus, gusanos o similares.

Para las redes de comunicación de datos la situación es muy similar. Son trascendentes consideraciones similares para el funcionamiento de redes de energía eléctrica, particularmente con respecto a seguridad. Dentro de la planificación y funcionamiento de la red eléctrica es importante tener una red robusta, evitando así, por ejemplo, situaciones en las que una gran parte de una población esté expuesta a un apagón eléctrico. El análisis de la conectividad de una red es importante para las consideraciones de robustez.

La administración del sistema implica invariablemente gestionar una red, que está compuesta por múltiples tipos de enlaces. Los ejemplos incluyen: los enlaces físicos entre las máquinas, los enlaces lógicos entre usuarios y archivos, y los enlaces sociales entre usuarios. Un aspecto importante de la administración del sistema es asegurar el flujo libre de información necesaria sobre la red, mientras que al mismo tiempo se inhibe el flujo de información dañina o perjudicial, sobre esta misma red.

La estructura de la red desempeña un papel crucial en la aplicación práctica de estos dos objetivos importantes, y parcialmente contradictorios, de administración del sistema. Ambos objetivos implican la propagación de información sobre los enlaces de la red; por lo tanto ambos problemas son muy sensibles a la estructura de la red. Debido a esta dependencia, la comprensión de la estructura de la red puede ser un componente valioso para la administración eficaz del sistema.

Adicionalmente, hay por supuesto redes que son sociales y tecnológicas. Los ejemplos incluyen redes de telefonía; redes de igual a igual [10] superpuestas sobre Internet; y la red combinada de ordenadores, archivos, y usuarios que es la preocupación diaria de cada administrador del sistema.

En esto, de nuevo, la seguridad parece una aplicación obvia para estas ideas: se desea identificar nodos a los que se debe dar la mayor prioridad de protección contra virus, por ejemplo.

Los estudios sobre redes han recibido una gran cantidad de atención en la última década más o menos. La mayor parte de las medidas de estructura de la red que se han estudiado hasta la fecha [8] toman la forma de propiedades de "gráfico completo", es decir, medidas o distribuciones escalares que se aplican al gráfico en su conjunto, y se calculan usando un promedio. Los ejemplos de dichas propiedades de gráfico completo incluyen el grado distribución del nodo, el diámetro o longitud media de la trayectoria, los coeficientes de agrupación, y la noción de "pequeños mundos", que se basa en los dos anteriores.

Las propiedades de gráfico completo son importantes y útiles; sin embargo no pueden dar una respuesta completa a la pregunta, "¿Cómo podemos entender la estructura de una red?"

Existen muchos ejemplos en los que el conocimiento de redes, que toma una forma más abstracta que el de las redes de telecomunicaciones, de datos, o eléctricas, es importante. Por ejemplo, en el campo de la epidemiología, es importante tener una comprensión de las redes sociales y cómo estas redes facilitan la propagación de enfermedades. Dentro de la distribución de información es importante conocer los mecanismos que gobiernan la propagación de información dentro de una población, sea a nivel local o global.

Cuando se consideran las interrelaciones humanas o las redes sociales se presta atención a los enlaces entre los individuos en lugar de a sus categorías o a lo que les caracteriza. Una red social es por lo tanto cualquier grupo de personas en el que los individuos tienen cualquier clase de relación entre ellos. Las personas con un alto grado de influencia social en redes sociales a menudo se etiquetan como líderes de opinión. Son influyentes gracias a su experiencia o gracias a su posición social. En cualquier caso esta influencia a menudo se manifiesta de por sí dando a los líderes de opinión una gran cantidad de contactos sociales; están relacionados con un gran número de personas. Por supuesto, esto es lógico; tener influencia social significa que tienes la capacidad para llegar a un gran número de personas.

La utilidad de esta idea para redes sociales parece clara [4]. También es obviamente de interés identificar comunidades en una red social medida. Un ejemplo con un aire ligeramente diferente es la red de contactos sexuales. En esto también estas ideas pueden ser bastante útiles, en el trabajo dirigido a limitar la propagación de enfermedades de transmisión sexual: quizás podrían enfocarse los dos objetivos complementarios de (i) evitar la infección de los nodos centrales de cada comunidad, y (ii) evitar la transmisión de la enfermedad a través de los nodos puente.

Por estas razones, las redes merecen un estudio formal. Una red es una de las abstracciones más sencillas de estructura que pueden estudiarse; aún así, comprender la estructura de una red no es una tarea trivial.

### Técnica Anterior

En el campo científico del análisis de redes, hay diversas maneras para medir la centralidad de los nodos de red. Una de estas medidas se denomina centralidad del vector propio. La centralidad del vector propio (EVC) la definió Bonacich [2] a principios de los años setenta. La idea básica detrás de la EVC es que, no es sólo cuánta gente conoces, sino también cómo de importantes (centrales) son, lo que determina cómo de importante (central) eres tú. Esta es, de esta manera, realmente una definición recursiva: mi importancia (centralidad) depende de la de mis vecinos, que a su vez depende de la mía. La cuestión de dicha definición recursiva es permitirnos identificar aquellos distribuidores que son realmente influyentes desde la perspectiva de la red en su conjunto. De lo contrario una definición que considerara la importancia sólo por cuántos vecinos tienes correría el riesgo de nombrar los centros de agrupaciones aisladas como distribuidores de la red. Con respecto a redes sociales estos centros sólo son influyentes en un sentido limitado, ya que su influencia no se extiende más allá de sus vecinos inmediatos.

El trabajo de Kleinberg [7], aunque referido a redes con enlaces dirigidos, proporciona alguna perspectiva útil. Kleinberg consideraba que un gráfico dirigido, definía dos tipos distintos de papeles para los nodos en el gráfico, y daba una manera de calcular índices que cuantifican el grado en el que cada nodo desempeña los dos tipos de papel. Es decir, cada nodo en un gráfico dirigido puede asignarse a una puntuación de Autoridad y una puntuación de Distribuidor. Es importante observar que estas puntuaciones pueden basarse únicamente en la topología de cualquier pregunta de contenido o significado independiente del gráfico, o de cualquier “propiedad” de los nodos.

Los nombres de estos dos tipos de papel expresan su significado. Los nodos con alta Autoridad son nodos que son importantes porque están señalados por nodos importantes -de hecho, por nodos con altas puntuaciones de Distribuidor. Y estos últimos obtienen sus altas puntuaciones de Distribuidor señalando buenos nodos de Autoridad. Resumiendo: los Distribuidores señalan, y las Autoridades son señaladas. Estas ideas pueden ser muy útiles para analizar la estructura de la WWW: Es probable que las Autoridades sean puntos finales buenos de una búsqueda de información, mientras que los Distribuidores son útiles para conducir la búsqueda a las Autoridades. Parece claro que surgen papeles similares en redes sociales: a veces, se sabe quién tiene la información necesaria (la Autoridad); otras veces, es necesario preguntar a un buen Distribuidor.

El trabajo de Kleinberg nos da índices para cada nodo en la red. Estos índices nos dan información útil sobre el papel o papeles que desempeña el nodo en la red. Dicha información es bastante diferente de la información de gráfico completo; y eso que aún se deriva, al menos como se presenta originalmente, simplemente a partir de la estructura topológica del gráfico.

Otro aspecto de un gráfico, que de nuevo es diferente de las propiedades de gráfico completo, es la estructura de comunidad del gráfico. En el mismo documento, Kleinberg sugirió una manera para encontrar dichas comunidades en gráficos tales como el gráfico de la Web. Las herramientas matemáticas usadas son básicamente las mismas que las usadas para encontrar las puntuaciones de Distribuidor/Autoridad - lo que significa, entre otras cosas, que la descomposición del gráfico en comunidades se basaba también simplemente en la estructura del gráfico, sin recurrir a ningún conocimiento o propiedades de los nodos o enlaces. Adicionalmente, puede observarse que descomponer un gráfico en sub-comunidades proporciona nueva información sobre los papeles que desempeñan los nodos: pueden ser miembros de la comunidad X; puede que no se encuentren en una comunidad; pueden ser “líderes” en algún sentido de su comunidad, o pueden encontrarse en el “borde”; y pueden desempeñar un papel importante en la unión de múltiples comunidades.

Muchos otros trabajos han abordado el mismo problema de cómo encontrar comunidades “naturales” en un gráfico dirigido tal como la Web. En contraste, Girvan y Newman [5] han considerado esta pregunta para gráficos no dirigidos. Su procedimiento básico es definir comunidades encontrando en primer lugar sus “límites” - específicamente, encontrando enlaces con alta “propiedad de estar en medio” que, cuando se retiran, descomponen el gráfico en sub-comunidades.

## ES 2 285 564 T3

La publicación “Archipelago: A Nodos Security Analysis Tool” de GEOFFREY CANRIGHT *et al.* describe un procedimiento para determinar la capacidad de una red para propagar información o tráfico físico con las siguientes etapas:

5 mapear la topología de una red, computar un valor para la fuerza de enlace entre los nodos, computar un índice de Centralidad del Vector Propio para todos los nodos, agrupar los nodos en regiones que rodean cada nodo central identificado, medir la susceptibilidad de la red para propagar información, basándose en el número de regiones, su tamaño, y cómo están conectadas. El procedimiento en este documento, sin embargo, no es suficientemente preciso puesto que pone un gran número de nodos en el conjunto del borde (es decir no los asigna a una región particular).

### Sumario de la invención

15 Un objetivo de la presente invención es proporcionar un procedimiento para el análisis de red, a aplicar a redes físicas, o a redes lógicas que existen como redes superpuestas encima de la red física. El aspecto común importante es la identificación de enlaces (físicos o lógicos), sobre qué información puede fluir.

20 Otro objetivo de la presente invención es proporcionar un medio “natural” -es decir, uno basado únicamente en la topología del gráfico- para decomponer un gráfico no dirigido en comunidades. Se definirá un conjunto de papeles para los nodos del gráfico, de manera que a cada nodo se le asigna uno, y solo un, papel. Es decir, a diferencia del procedimiento de Kleinberg, para la presente solicitud es deseable que los papeles sean propiedades binarias (Sí/No) de los nodos -y también exclusivas.

25 La técnica anterior [13, 3] ha mostrado con más detalle cómo aplicar el análisis presentado aquí para ordenadores conectados en red con muchos usuarios. La presente invención proporciona una manera natural de decomponer una red en agrupaciones bien conectadas, y de asignar papeles significativos en el flujo de información a cada nodo en la red.

30 Estos objetivos se consiguen en un procedimiento para el análisis de red como se describe en la reivindicación adjunta 1. En particular, la presente invención proporciona un procedimiento para analizar la capacidad de una red para propagar información o tráfico físico, incluyendo dicha red numerosos nodos de red interconectados mediante enlaces, incluyendo dicho procedimiento las etapas de

- 35 • mapear la topología de una red,
- computar un valor para la fuerza de enlace entre los nodos,
- computar un índice de Centralidad del Vector Propio para todos los nodos, basado en dichos valores de la fuerza de enlace
- 40 • identificar nodos que son máximos locales del índice de Centralidad del Vector Propio como nodos centrales,
- agrupar los nodos en regiones que rodean cada nodo central identificado,
- 45 • asignar un papel a cada nodo a partir de su posición en una región, como nodos centrales, nodos miembro de una región, nodos de borde, nodos puente, nodos oscilantes,
- medir la susceptibilidad de la red a la propagación, basándose en el número de regiones, su tamaño, y cómo están conectadas.

50 Aparecen realizaciones ventajosas de la invención a partir de las siguientes reivindicaciones dependientes.

### Breve descripción de las figuras

55 Para hacer a la invención más fácil de entender, la invención se analizará ahora en detalle con referencia a las figuras adjuntas, en las que:

60 La Figura 1 es un diagrama de flujo esquemático que muestra un Nodo Intermedio (izquierda) y un Nodo Intermedio y Oscilantes (derecha).

La Figura 2 muestra un gráfico sencillo con dos regiones.

65 La Figura 3 muestra el mismo gráfico que en la Figura 2, pero con las regiones definidas usando otra regla. Se muestran también los valores de EVC para los nodos.

Las Figuras 4, 5, y 6 muestran los gráficos resultantes del proyecto MANA [4] usando tres medidas diferentes para la fuerza de enlace.

La Figura 7 es un flujo diagrama de flujo que ilustra el procedimiento usado para calcular el índice de Centralidad del Vector Propio.

### Descripción detallada de la invención

5

La presente invención describe aplicaciones útiles e interesantes de ideas de análisis de red. El único prerrequisito es que los enlaces de la redes sean no dirigidos. Por enlaces no dirigidos se entienden enlaces que no señalan en una dirección específica. En la World Wide Web una página web puede señalar a otra, pero esta página no necesariamente tiene que señalar de vuelta. En este caso las páginas se conectarán mediante un enlace dirigido. Si ambas páginas estuvieran hiperenlazadas entre sí, un enlace en cada dirección, estos enlaces podrían colapsar en un solo enlace no dirigido. La presente invención trata a todas las redes como compuestas por enlaces no dirigidos.

10

La idea perseguida por la presente solicitud es que puede observarse una “buena conexión” como una función de la altura sobre el espacio discreto (el gráfico). Si la función de la altura de la presente invención es suficientemente regular, pueden emplearse ideas apropiadas para superficies regulares sobre un espacio continuo. Es decir, la presente invención usará una imagen topográfica para definir regiones en una red. Las regiones se corresponderán a “montañas”, siendo el centro de cada región la cima de la montaña correspondiente. Los límites entre regiones se definirán entonces como aquellos puntos que fallan a la hora de asociarse exclusivamente con una región de montaña.

15

20

Los papeles definidos son: “líder” de una comunidad (región); miembro de una comunidad; y dos tipos de papeles para nodos en el “conjunto del borde”, es decir, nodos que no pertenecen a ninguna comunidad.

25

El enfoque seguido es prácticamente doble según el de Girvan y Newman [5]. La presente invención comienza, no con los “bordes”, sino con los “centros” de las comunidades. Desde este punto de partida, se trabaja “hacia fuera” para encontrar los miembros, y finalmente los nodos de borde. El conjunto de papeles presentado es completo y consistente, en el sentido de que las definiciones permiten una asociación única y no ambigua de un único papel a cada nodo en el gráfico.

### Realización de la presente invención

30

La gente que se comunica entre sí forma una red social, donde los enlaces se basan en su comunicación. Estos enlaces pueden distinguirse de acuerdo con el tipo de medio que se está usando, sea telefonía, comunicaciones cara a cara, o correo. De esta manera, la red social puede describirse como *multiplex*: es una red en la que los nodos están relacionados entre sí por diferentes tipos de enlaces. Aunque las relaciones sociales que enlazan juntas a diferentes personas pueden existir independientemente del tipo de medio usado, el tipo de medio desempeña un papel importante en la definición de los enlaces, puesto que cada medio es un canal diferente para el flujo de información. Los diferentes medios de comunicación son en este sentido análogos a idiomas. Por ejemplo, una persona que quiere alcanzar muchos nodos de la red tiene que poder comunicarse con múltiples tipos de medios -tiene que hablar el “idioma” preferido de los otros nodos. Esta idea de enlaces diferenciados por medios es válida para la mayoría de clases de redes: una enfermedad puede propagarse, por ejemplo, a través de numerosos vehículos de infección diferentes, y los enlaces en las redes de transporte pueden constar de muchos medios diferentes para transporte, por ejemplo coches, aviones, o trenes.

35

40

### Fuerza del enlace y medidas de EVC

45

La fuerza de los enlaces en este tipo de red *multiplex* puede determinarse de diferentes maneras. A continuación se mencionan cuatro:

50

1) Puede establecerse simplemente si un enlace (de cualquier tipo) existe o no. Numéricamente, se asigna 0 cuando “no hay enlace” y 1 cuando hay “algún enlace”.

55

2) Puede contarse el número de medios diferentes que conectan cualquier par de nodos, es decir, el número de medios diferentes que tiene cualquier flujo de sustancias o de información entre dos nodos cualesquiera en la red.

3) Puede medirse el flujo total entre dos nodos cualesquiera en la red. Para hacer esto deben convertirse los datos que están disponibles a una medida común. De esta manera, esta medida da la cantidad neta de flujo (por ejemplo minutos o palabras para los medios de comunicación) entre dos nodos cualquiera en la red.

60

4) Una cuarta alternativa es determinar la fuerza de los enlaces en una mezcla de 2) y 3). Es decir, contra cada medio [como en 2)], pero ponderado [como en 3)] por la fracción de flujo por este medio, que usa un par dado.

65

La manera tradicional para determinar las fuerzas de enlace se indica con el número 1). El procedimiento 3) también es conocido. Los procedimientos 2) y 4) son procedimientos nuevos e innovadores para determinar las fuerzas de enlace.

El índice de centralidad del vector propio (EVC) se define matemáticamente como el vector propio principal de una matriz. El procedimiento más sencillo y habitual para encontrar el vector propio principal de una matriz es el “Procedimiento de las Potencias” [14]. Este procedimiento implica la multiplicación repetida sobre un vector de

ponderaciones por la matriz. La multiplicación sobre el vector de ponderación por la matriz es equivalente a lo que se denomina “propagación de la ponderación”: redistribuye un conjunto de ponderaciones de acuerdo con una regla. La redistribución repetida de las ponderaciones (con la normalización global de la ponderación total) produce una distribución estacionaria, que es el vector propio dominante o principal. Estas son las puntuaciones, que se usan como índice de centralidad en la presente invención.

Por claridad, se ilustra la aplicación del Procedimiento de las Potencias, en la Fig. 7. En este documento, usando las ecuaciones explicadas anteriormente, el proceso comienza y se elige un vector inicial  $w_0$  (S401). En cada iteración, se calcula una nueva ponderación  $p_{nueva}$  (S403) redistribuyendo las ponderaciones de acuerdo con la acción del operador de la matriz. Esta nueva ponderación se normaliza después (S405). Después se realiza un ensayo de convergencia (S407). Si la ponderación ha convergido, el proceso finaliza. De lo contrario, se calcula una nueva ponderación y se repite el proceso hasta que la ponderación converge.

Para el análisis de redes sociales *multiplex*, se ha generalizado la medida de EVC para que incorpore otras tres medidas de fuerza de enlace (2-4), como se ha mencionado anteriormente. La modificación de la idea general de EVC, como se aplica en los nuevos procedimientos 2) y 4), es de la siguiente manera: un nodo es central si tiene muchos vecinos con alta centralidad -y usa muchos tipos diferentes de medios. A continuación se describe cómo poner en práctica esta idea general para cada uno de los cuatro enfoques para la fuerza de enlace analizados anteriormente:

1) El enfoque tradicional, en el que la matriz de contigüidad  $\underline{A}$  está compuesta únicamente por ceros y unos, podría usarse con redes *multiplex*; pero es totalmente insensible al número de medios usados por cada par de nodos.

2) En este documento simplemente se ha sustituido la matriz  $\underline{A}$ , cuyas entradas son todas 0 ó 1, por la matriz  $\underline{A}_{color}$  definida de la siguiente manera: la entrada  $(\underline{A}_{color})_{ij}$  es igual al número de “colores” (medios diferentes) que conectan los nodos  $i$  y  $j$ .

3) En este documento los unos en la matriz  $\underline{A}$  tradicional se sustituyen por un número real positivo, dando el volumen total de flujo (sumado para todos los medios, y medido en una unidad común de medida) durante un intervalo de tiempo dado. Es decir:  $(\underline{A}_{volumen})_{ij} = \sum V_{c,ij}$ , en la que  $c$  es un índice que varía sobre los “colores” (medios), y  $V_{c,ij}$  es el volumen total de comunicaciones en el medio  $c$  entre los nodos  $i$  y  $j$ .

4) Finalmente la presente invención propone una mezcla de los enfoques 2) y 3), para dar importancia tanto al volumen de flujo como a la existencia de múltiples medios. Por lo tanto, para cada medio  $c$  y par de nodos  $ij$ , se da una “puntuación” que es la fracción (aportada por el par  $ij$ ) de la comunicación total que usa el medio  $c$  en la red. Si  $V_{T,c}$  denota el volumen total (por toda la red) de comunicación que usa el medio  $c$ . Entonces nuestra medida “mixta” de fuerza de enlace puede escribirse como

$$(\underline{A}_{mixta})_{ij} = \sum (V_{c,ij} / V_{T,c}).$$

El procedimiento de acuerdo con la presente invención convierte los datos de flujo en una matriz de contigüidad, usando uno de los cuatro procedimientos descritos anteriormente para dar a cada entrada de la matriz una medida de fuerza de enlace. Después calcula el vector propio principal de la matriz de contigüidad modificada resultante. Esto nos permite asignar un índice (un número positivo) a todos los nodos en la red, dando su centralidad de acuerdo con una de nuestras cuatro medidas. Aquellos nodos con los valores de centralidad más altos se contemplan como los nodos más centrales en la red. Esto permite al procedimiento producir una lista de los distribuidores de la red y sus vecinos inmediatos. El índice de centralidad hace posible también producir un mapa topográfico de la estructura de la red, es decir, una visualización gráfica de la red que muestra los nodos más centrales como “picos” locales.

#### Papeles en las redes

El objetivo final de la presente invención es asignar un papel natural y único a cada nodo en la red, basándose únicamente en la topología del gráfico. Como se ha indicado anteriormente, Kleinberg encontró dos de dichos papeles para gráficos dirigidos: Distribuidores y Autoridades. Los Distribuidores son naturalmente buenos para señalar buenas Autoridades; y las Autoridades son naturalmente buenas para ser señaladas por buenos Distribuidores. A partir de estas afirmaciones gramaticales sencillas puede verse ya que la distinción entre Distribuidores y Autoridades se desvanece cuando los arcos del gráfico se convierten en no dirigidos (de manera que “señalar a” = “ser señalado por”). Las matemáticas dan el mismo resultado: para el caso de no dirigidos, la matriz de contigüidad es simétrica,  $A = A^T$  y, por lo tanto, las matrices que definen Distribuidores y Autoridades se hacen iguales.

Resumiendo, para gráficos no dirigidos, los dos tipos de papeles colapsan en uno solo. Este papel único (más precisamente, un índice que cuantifica el grado en el que el nodo desempeña el papel) es la centralidad del vector propio.

El operador del Distribuidor  $AA^T$  y el operador de Autoridad  $A^T A$  simplemente se convierten en  $A^2$ , cuyo vector propio principal es el mismo que para  $A$ .

## ES 2 285 564 T3

Por lo tanto se ha descubierto que dos de los papeles identificados en el trabajo de Kleinberg con gráficos dirigidos se convierte en un único (tipo de) papel para un gráfico no dirigido. Este tipo de papel se denomina buena conectividad en las siguientes secciones, o centralidad del vector propio. Se buscan además distinciones entre los nodos de un gráfico no dirigido -en otras palabras, múltiples papeles diferentes, a los que pueden asignarse cualquier nodo dado. Estos papeles se definirán en la siguiente sección. La centralidad del vector propio (EVC) será la función de la altura, y por lo tanto el punto de partida.

### *Definiciones de los papeles*

Tiene que aclararse la diferencia entre el “tipo de papel” y el “papel”. Los índices de valor real o “puntuaciones” pueden asociarse con cada nodo: las puntuaciones de Distribuidor y Autoridad para el caso dirigido, y la puntuación de EVC para los casos no dirigidos. Estos son los tipos de papeles; de hecho, es justo decir que las tres puntuaciones representan algún tipo de centralidad. Todos los nodos tienen algún grado de centralidad; y “ser central” es ciertamente un tipo de papel. Por papel, sin embargo en este documento, se entiende una distinción binaria (sí/no) aplicada a cada nodo, de manera que cada nodo recibe un único Sí y por lo tanto se le asigna un papel único y no ambiguo. La centralidad (un tipo de papel) dará una función de altura regular sobre el gráfico, permitiendo el uso de criterios topográficos para asignar un papel (Sí o No) a cada nodo.

### *Centros*

Por simplicidad y capacidad de lectura se mantiene la imagen de montañas, valles, collados etc. para la función de la altura. Cada montaña puede definirse por su pico. El pico es un máximo local de la función de la altura. El primer papel es entonces el pico de la montaña.

**Centro:** cualquier nodo que sea un máximo local de la centralidad del vector propio es un Centro.

### *Regiones*

Cada cima de montaña define una montaña. Por lo tanto el número de Regiones en el gráfico es igual al número de centros. (En lo sucesivo, excepto cuando se definen los papeles, se suprimen las letras mayúsculas; el significado debería estar claro a partir del contexto). Las regiones están compuestas normalmente por más de un nodo; por lo tanto el papel para un nodo no puede ser una región, sino en lugar de ello un Miembro de una Región.

**Miembro de una Región:** cada nodo que puede asociarse exclusivamente con un único Centro, de acuerdo con una regla no ambigua, es un miembro de esta Región del Centro, y por lo tanto un Miembro de una Región.

Queda por especificar la “regla no ambigua”. De acuerdo con la presente invención, se dan dos posibles elecciones para la “regla no ambigua”.

*Regla 1 (distancia).* Un nodo está asociado con el Centro C si está más próximo (en número de saltos de trayectoria más cortos) a C que cualquier otro Centro  $C_0$ .

*Regla 2 (ascenso pronunciado).* Para cada nodo  $i$ , una trayectoria de ascenso pronunciado que comienza en  $i$  terminará en uno (o más) Centros. Si termina en un único Centro, entonces el nodo  $i$  está asociado con ese Centro.

Estas reglas son simplemente la versión de dominio discreto del proceso de asociar una parte del dominio (espacio base) con cada cima de montaña- que define, por lo tanto, cada montaña. Hay que tener cuidado aquí para dividir la definición de región en dos partes: la propia definición, que se refiere a una regla pero que no la especifica; y la regla. Esto se hace porque es posible más de una regla para el caso discreto; y se establece la definición de la región de una manera que captura la idea de “montaña”, pero deja la regla sin especificar.

Ambas reglas establecidas anteriormente satisfacen el criterio intuitivamente razonable de que los vecinos cercanos a un centro deben pertenecer (en general) a su región. (Es, después de todo, el número y conectividad de los vecinos de un centro lo que da a este centro su alta EVC). Ambas reglas son también fáciles de poner en práctica de una sencilla forma iterativa - partiendo de los centros, y trabajando hacia fuera desde ellos, “coloreando” nodos de acuerdo con las regiones (centros) a las que pertenecen. La regla de ascenso pronunciado es, sin embargo, la regla que es la más fiel a la imagen topográfica.

### *Bordes - entre regiones*

Sobre una superficie topográfica continua hay algunos puntos que se encuentran entre montañas, y no pertenecen a una única montaña. Puede suceder que también existan puntos análogos para el caso discreto.

Los nodos que no pueden asociarse con ninguna montaña se asignan al conjunto del Borde.

**Nodos de Borde:** cualquier nodo para el que la regla no ambigua para miembros de una Región da más de una respuesta es un Nodo de Borde.

Intuitivamente, se cree que los nodos de borde son “regiones de conexión”. Y aún, pensándolo un poco más, se pone de manifiesto que no todos los nodos de borde son iguales en este aspecto. Algunos nodos de borde de hecho desempeñan un papel importante en la conexión de dos o más regiones: se encuentran en las trayectorias que conectan los centros respectivos (por lo tanto regiones). Véase el panel izquierdo de la Figura 1. Otros nodos pueden retirarse, sin ninguna pérdida en el grado de conexión entre las regiones. Véase el panel derecho de la Figura 1. Por lo tanto es natural definir dos papeles diferentes para ajustar los nodos de borde.

**Nodo Intermedio:** un Nodo de Borde que se encuentra en al menos una trayectoria no auto-retrazable que conecta dos Centros es un Nodo Intermedio.

**Oscilante:** cualquier Nodo de Borde que no sea un Nodo Intermedio es un Oscilante.

Los Oscilantes, por supuesto, pueden introducir nueva información en la red; pero no desempeñan un papel significativo en el transporte de información entre regiones.

Finalmente, es deseable distinguir una clase de enlaces que desempeñan un papel importante en regiones de conexión. La razón para hacer esto aquí es que el conjunto del borde para la regla de ascenso pronunciado es en general muy pequeño o cero. En este caso es aún útil destacar aquellos elementos rojos que conectan las regiones. Por lo tanto se define:

**Enlaces de Conexión:** cualquier enlace cuyos puntos finales se encuentran en dos Regiones diferentes es un Enlace de Conexión.

Los enlaces de conexión se encontrarán para cualquiera de las reglas de región anteriores. Pueden imaginarse reglas para definir regiones que dan bordes “gruesos”. Por ejemplo, podrían asociarse nodos con centros de acuerdo con:

**Regla 1’ (distancia con el corte).** Un nodo se asocia con el Centro  $C$  si está más cerca (en número de saltos) a  $C$  que cualquier otro Centro  $C_0$ , y si su distancia desde  $C$  no es mayor de  $h$  saltos.

Los bordes “Gruesos” surgen para dicha regla ya que puede haber muchos nodos que están más allá de  $h$  saltos desde cualquier centro. En general, surgen límites “gruesos” si se elige una regla diseñada para evitar que “crezcan juntas” regiones a partir de sus centros respectivos. La distancia a la que se permite el crecimiento puede medirse entonces en saltos (como en la Regla 1’), o en disminuciones de EVC.

Los límites de acuerdo con la Regla 1 son “finos”: esencialmente de un nodo de anchura. Los límites de acuerdo con la Regla 2 son aún más finos: en general, son nodos de una anchura 0, ya que es raro que un nodo tenga dos o más trayectorias de ascenso pronunciado, que conducen a diferentes máximos locales.

### Las matemáticas

Los problemas matemáticos según se resuelven mediante la presente invención se resuelven centrándose en funciones “regulares” sobre un espacio discreto.

Supongamos que el espacio del dominio es continuo. Entonces las funciones armónicas son las funciones más regulares disponibles. Estas funciones son soluciones a la ecuación de Laplace,

$$\nabla^2 \phi = 0 \tag{1}$$

Para un espacio dado, se obtienen diferentes soluciones para (1) a partir de diferentes condiciones límite sobre  $\phi$ .

Pueden identificarse inmediatamente algunos problemas con la imagen continua. Un problema es que no hay máximos, o mínimos, más allá del límite. Por lo tanto la imagen topográfica de acuerdo con la presente invención no puede trabajar con dichas funciones regulares: cada cima de la montaña se encontrará en el límite. Adicionalmente, la presente invención describe una manera natural de definir regiones. En este punto “natural” significa guiado tanto como sea posible por la topología del gráfico. Por lo tanto es indeseable tener que asignar valores para la función  $\phi$  en el límite - sería preferible que la topología resolviera este problema.

Por supuesto este último problema puede resolverse fijando  $\phi = \text{constante}$ , por ejemplo, cero, en el límite. Es decir, al límite se le da únicamente un valor de referencia nominal. Esto es suficientemente “natural”; sin embargo puede conseguirse que  $\phi = \text{constante}$  en todo el espacio, debido a la propiedad de promedio de la ecuación de Laplace.

## ES 2 285 564 T3

La versión discreta de la ecuación de Laplace es

$$L\phi = 0 \quad (2)$$

5

donde  $L = K - A$  es la matriz laplaciana,  $K = \text{Diag}(k_1, k_2, \dots)$  es una matriz diagonal cuya entrada es el grado del nodo  $k_i$ , donde  $k_i$  es el número de vecinos conectados del nodo  $i$ , y  $A$  es la matriz de contigüidad, con  $A_{ij} = 1$  si hay un enlace de  $i$  a  $j$ , y 0 en caso contrario.

10 Es fácil observar que la propiedad de promedio también se mantiene aquí: las soluciones para (2) cumplen

$$\phi_i = \frac{1}{k_i} \sum_{j=\text{nn}(i)} \phi_j \quad (3)$$

15

Aquí “nn” significa “vecino cercano”. La ecuación de Laplace discreta de esta manera ofrece funciones ‘más regulares’ para el caso discreto; pero tiene todos los problemas observados para las funciones armónicas continuas, más uno más. El problema adicional es el resultado del hecho crucial de que la especificación del límite de un espacio discreto no es única -de hecho, no hay una manera natural para definir dicho límite. Por supuesto puede suponerse, quizás menos arbitrariamente, que ninguno de los puntos son puntos límite -todos tienen que tener su altura determinada por la estructura del gráfico- pero después recuperar la constante  $\phi_i = \text{constante}$ .

### 25 *Centralidad del Vector Propio*

Siguiendo el análisis a partir de la expresión (3). Un pequeño cambio en la imagen según se da en (3) resuelve todos los problemas de una vez. El pequeño cambio es el siguiente: se busca una función de la altura que cumpla, en lugar de la propiedad de promedio (3), lo siguiente:

30

$$\phi_i = \frac{1}{\lambda} \sum_{j=\text{nn}(i)} \phi_j \quad (4)$$

35 Es decir, en lugar de tomar la media estricta sobre todos los vecinos, se divide la suma de vecinos por una constante  $\lambda$ , que es la misma para todos los nodos. Esta ecuación puede escribirse como

$$A\phi = \lambda\phi \quad (5)$$

40

donde  $A$  es de nuevo la matriz de contigüidad. Ahora tenemos una ecuación de eigenvalor, y la función de la altura  $\phi$  es un vector propio de la matriz de contigüidad. La presente invención quiere de hecho que el vector propio que es la solución iterativa estable de (4), porque se supone que la altura significa “buena conectividad”. Es decir, (4) codifica la idea de que se determina una buena conectividad del nodo  $i$ , dentro de una escala constante  $\lambda$ , por la de todos los vecinos de  $i$ . Iterando este requisito, desde cualquier punto de partida, dará el vector propio principal de la matriz de contigüidad. Este vector propio da la solución estable, auto-consistente de (4); tiene también la propiedad de que es positivo semi definido, ya que  $A$  lo es.

50 Con esta única modificación, los problemas como se han observado anteriormente con la ecuación de Laplace (discreta o de otro tipo) ya no están presentes nunca más. La EVC puede tener máximos locales lejos del límite. De hecho, como mide una buena conectividad, los máximos locales de la EVC tienden a encontrarse bien lejos de cualquier nodo al que se pretenda llamar “nodos del límite”. Adicionalmente, no hay necesidad de definir un límite para el caso discreto: todos los nodos pueden tener valores de EVC determinados por la Ecuación (4), sin introducir valores como “condiciones límite”.

55

Específicamente, las contribuciones aquí son:

1) Las dos nuevas formas modificadas para la matriz de contigüidad, que dan dos nuevas medidas de centralidad que permiten elegir los centros rojos.

60

2) La definición y procedimiento para identificar regiones rojas.

3) Las definiciones y procedimientos para asignar papeles de red discretos a cada nodo en la red.

65

4) Aplicar las nuevas medidas de centralidad, regiones, y papeles a una amplia variedad de aplicaciones.

## Ejemplos

A continuación se dan ejemplos de realización de la presente invención así como comparaciones entre las dos reglas para definir regiones.

Las Figuras 4, 5, y 6 muestran los resultados del proyecto de investigación MANA como se presenta en [4]. Los gráficos representan una pequeña red social, un grupo de trabajo de 11 personas. Usando el procedimiento presentado de diferentes medidas por la fuerza de enlace, se crearon los índices de centralidad basados en EVC para la red. Las visualizaciones topográficas muestran la centralidad de los nodos como diferencias de altura. En la figura 4, la fuerza de enlace se mide basándose en el número de medios diferentes usados por cada nodo (procedimiento 2). La Figura 5 muestra el gráfico cuando la fuerza de enlace se basa en la cantidad neta de flujo entre los nodos (procedimiento 3). Finalmente, la figura 6 muestra el gráfico que se basa en una mezcla de los procedimientos anteriores para determinar la fuerza de enlace, es decir, tanto el número de medios usados como la cantidad neta de flujo (procedimiento 4).

La Figura 2 muestra un gráfico sencillo con dos centros. El Borde consiste en tres nodos. Uno (nodo 11) es un nodo puente que claramente desempeña un papel esencial en la conexión de dos regiones, los otros dos son oscilantes.

Aplicando la Regla 2 al mismo gráfico nos da la Figura 3. En él, puede observarse que todo el borde ha sido “tragado” por el centro dominante (nodo 9). El papel más periférico de los nodos 10 y 12 -clasificados anteriormente como oscilantes- se refleja ahora en su distancia (2 saltos) desde su centro (y por supuesto en su baja EVC).

Comparando estas dos figuras de esta manera se confirman las expectativas sobre las diferencias entre las dos reglas: un conjunto del borde, con o sin pendientes, está presente típicamente con la Regla 1, pero ausente con la Regla 2.

Para ilustrar la aplicación de estas ideas, se supone que los nodos en las Figuras 2 y 3 son usuarios en una red de ordenadores, mientras que los enlaces son conexiones eficaces entre usuarios que permiten el flujo de información. Aquí se usa el término conexiones “eficaces”, porque los enlaces no pueden ser directos: tienen que estar mediados por archivos a los que los usuarios tienen acceso tanto leído como escrito [3]. Puede concluirse inmediatamente a partir del análisis que el sistema de usuarios está compuesto de forma natural por dos grupos principales. Adicionalmente, el nodo 9 es el más central para el grupo amarillo, mientras que el nodo 13 es el más central para el grupo azul. Finalmente, el nodo 11 es un nodo puente que es crucial para el flujo de información entre los dos grupos.

Supongamos además que la seguridad para este pequeño sistema es de interés. Entonces pueden identificarse inmediatamente los nodos 9, 13, y 11 como los que necesitan protección más urgentemente de cualquier amenaza a la que se enfrente el sistema. Los nodos 9 y 13 deben protegerse porque son centros de sus regiones: si se infectan, entonces hay una alta probabilidad que se infecte también toda la región.

Adicionalmente, puede darse al nodo 9 una mayor prioridad de protección que al nodo 13, ya que su región es más grande. Finalmente, el nodo 11 merece protección extra, ya que si se le hace inmune a las amenazas, entonces estas amenazas no tienen un canal preparado para propagarse de una región a otra.

Obsérvese que el uso de la Regla 2 no resalta ningún nodo del borde para una protección especial - aunque el nodo 11 claramente desempeña un papel importante en la conexión de dos regiones. Sin embargo, la Regla 2 identificará el enlace entre 11 y 13 como un enlace intermedio. La consecuencia obvia de esto es que los nodos en cada extremo de cada enlace intermedio merecen medidas protectoras especiales.

Este problema puede plantearse del revés, otorgando al administrador el problema de propagar la información deseada en esta misma pequeña red. El análisis sugiere entonces una estrategia eficaz para hacer esto: se empieza con los centros (nodos 9 y 13), y se dispone la información deseada para transmitirla desde allí.

Por supuesto, es de esperar que la regla de la distancia y la regla de ascenso pronunciado den resultados en conflicto para algunos nodos. Un punto importante a considerar a partir de las Figuras 2 a 7 es que la imagen cualitativa general es bastante insensible a la elección de la regla para definir regiones. Puede esperarse que este sea el caso para la mayoría de los gráficos. La elección de centros es independiente de que se use la regla; y estos centros a su vez existen precisamente porque se encuentran en una región del gráfico que tiene algún “peso” - es decir, algún número de nodos que están mejor conectados entre sí que a sus “alrededores”. Resumiendo, las diferentes reglas, que aparentemente definen regiones, realmente se diferencian principalmente de acuerdo con dónde ponen los límites entre regiones - mientras que las propias regiones son objetos bastante estables.

### *Sumario de las definiciones de papeles y regiones en las redes*

El criterio básico para definir una región (y su centro) ha sido una buena conectividad, medida mediante la función de gráfico “regular”, la centralidad del vector propio o EVC. Además de definir las agrupaciones naturales de un gráfico, nuestro procedimiento asigna también un único papel a cada nodo en el gráfico.

## ES 2 285 564 T3

Las dos reglas que definen regiones dan imágenes cualitativamente similares para la estructura del gráfico en su totalidad, pero imágenes bastante diferentes en términos de qué papeles asignan a los nodos que están presentes en el análisis.

5 Es decir, la Regla 1 -que asocia nodos con regiones basándose simplemente en su distancia, en los saltos de trayectoria más corta, desde los centros- pone un número significativo de nodos en el conjunto del borde. Estos nodos a su vez pueden ponerse en dos papeles diferentes: nodos puente, y oscilantes (véase la Figura 2). La Regla 2 se mantiene más próxima al espíritu “topográfico” del procedimiento según se describe dentro de la presente solicitud, asociando los nodos con los centros a los que están enlazados mediante una trayectoria de ascenso pronunciado. Esta  
10 regla normalmente (en ausencia de simetría especial) no pone nodos en el conjunto del borde - de manera que, con la Regla 2, los dos papeles en el conjunto del borde (nodos puente y oscilantes) se excluyen esencialmente, y todos los nodos son centros de una región, o miembros de una región.

15 Pueden imaginarse otras reglas para definir regiones. El aspecto principal del procedimiento de acuerdo con la presente invención es identificar centros en primer lugar, y después permitir que las regiones “crezcan” hacia fuera desde esos centros. Ambas reglas de acuerdo con la presente invención se ajustan a esta imagen. El enfoque de Girvan/Newman permite la descomposición jerárquica de un gráfico, descomponiendo agrupaciones en sub-agrupaciones, etc. Una descomposición jerárquica similar podría realizarse de acuerdo con la presente invención, eliminando los nodos y enlaces del borde, y aplicando el análisis de acuerdo con la presente invención a las regiones aisladas re-  
20 sultantes. Pueden definirse papeles adicionales basándose en los presentes procedimientos de análisis. En un ejemplo muy sencillo, puede asignarse el papel de “Borde de la región” a aquellos nodos que están conectados a elementos de borde (nodos o enlaces). Puede asignarse un tipo diferente de papel de Borde a aquellos nodos que están “más lejos” del centro, pero que no están enlazados con ningún elemento de borde.

### 25 Aplicaciones

A continuación, se dan aplicaciones del procedimiento y sistema de acuerdo con la presente invención. Claramente, puede destacarse que ambos nodos altamente centrales, y puentes (enlaces o nodos) merecen atención y cuidado extra para evitar la propagación de averías. Los nodos muy centrales tienen más probabilidades de infectar a sus regiones; y  
30 los puentes a su vez deben protegerse para que la infección no se propague de una región a las otras. Por lo tanto sería práctico inmunizar ciertos elementos, y asegurar así que se aísla cualquier infección a una única región. Para regiones más grandes, sería práctico inmunizar los nodos más centrales en cada región -priorizando por supuesto aquellas regiones con el mayor número de nodos. Algunos casos como los sistemas de igual a igual muy bien conectados, por otro lado, son difíciles de proteger, porque están demasiado bien conectados. Esto significa que hay muchos nodos en  
35 cada región con prácticamente la misma centralidad, y que hay muchos puentes entre regiones (para aquellos casos en los que haya más de una región).

El uso del sistema y procedimiento es aplicable a otros muchos tipos de gráficos -en principio a cualquier gráfico que sea no dirigido. El procedimiento se modifica fácilmente -como se describe en la primera realización- también para permitir los pesos (distintos de 0 ó 1) para los enlaces entre nodos. El procedimiento y sistema de acuerdo con la presente invención demostrará que es útil en el análisis de redes sociales - que pueden tener (de nuevo) una fuerza (positiva) asociada con cada enlace.

45 Cuando una innovación -un nuevo producto o servicio- se introduce en una población, la difusión de la innovación sigue un patrón típico. La innovación normalmente es descubierta por un pequeño grupo de adoptadores prematuros, y después de un tiempo, dependiendo de la aprobación de los adoptadores prematuros, los líderes de opinión (o líderes adoptadores) adoptan la innovación. Este es el punto crítico del proceso de difusión, porque la adopción de la innovación por la mayoría de la población normalmente depende de la aceptación de los líderes de opinión [6]. En otras palabras, la adopción de una innovación empieza a tener éxito cuando los líderes de opinión o distribuidores de  
50 la red social aprueban y adoptan la innovación.

El procedimiento como se ha descrito en la realización y sus ejemplos adjuntos de la presente invención, usa una matriz de contigüidad modificada, basada en datos de flujo, para computar una centralidad medida para cada nodo en una red social. Este índice de centralidad permite a la mayoría de los nodos centrales de la red social que esta matriz  
55 de contigüidad represente para ser elegida. Estos nodos -los distribuidores de la red- son, en términos sociales de la red, líderes de opinión. De esta manera son buenos objetivos para la propagación de información etc., porque pueden contribuir potencialmente a la aceleración de la difusión de dicha información. Una aplicación obvia del procedimiento es de esta manera en el área de difusión de innovación.

60 En la parte preliminar se ha hecho referencia a la epidemiología, telecomunicación, redes de datos, sistemas de energía eléctrica, etc. Puede añadirse que el resultado del análisis de acuerdo con la presente invención tiene adicionalmente un amplio intervalo de aplicaciones. Un ejemplo es la planificación de los horarios dentro del transporte, o sistemas de transmisión y distribución. Analizando el flujo de tráfico en una red de carreteras o en un sistema de ferrocarriles, podría encontrarse la mejor temporización para la distribución para evitar la congestión de tráfico. Aná-  
65 logamente, la planificación de las rutas de tráfico en redes de telecomunicaciones y redes de datos, así como la planificación del tráfico en una base más general, es una aplicación obvia de la presente invención, porque el procedimiento puede identificar fácilmente los puntos congestionados o las rutas buenas. Aún más a un nivel más microscópico puede usarse en el diseño de ordenadores, para analizar el tráfico interno y de esta manera optimizar sus componentes y sus

enlaces comunes. Esto último es particularmente útil dentro del área del procesado paralelo, para reducir el tráfico entre procesadores/ordenadores.

Obsérvese que aunque en lo anterior se ha proporcionado una descripción detallada de la presente invención, debe entenderse que los equivalentes se incluyen dentro del alcance de la invención de acuerdo con lo reivindicado. La descripción detallada trata en una gran extensión con la teoría detrás de la presente invención, sin embargo el uso de estas teorías tiene un amplio campo de aplicaciones, con la condición de que los gráficos sean no dirigidos.

De esta manera en una base general el procedimiento de acuerdo con la presente invención es aplicable dentro de una amplia área de campos y puede aplicarse para resolver problemas dentro de estas áreas. Otras realizaciones ventajosas de la presente invención resultarán evidentes a partir de las reivindicaciones dependientes adjuntas.

#### Abreviaturas y referencias

1. G. D. **BATISTA**, P. **EADES**, R. **TAMASSIA**, Y I. G. **TOLLIS**, *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, Upper Saddle River, New Jersey, 1999.

2. P. **BONACICH**, *Factoring and weighting approaches to status scores and clique identificación*, *Journal of Mathematical Sociology*, 2 (1972), pág. 113-120.

3. M. **BURGESS**, G. **CANRIGHT**, Y K. **ENGØ**, *A graph theoretical model of computer security: from file access to social engineering*, *Internacional Journal of Information Security*, (2003). Presentada a publicación.

4. G. **CANRIGHT**, K. **ENGØ-MONSEN**, Y Å. **WELTZIEN**, *Multiplex structure of the communications network in a small working group*, *Social Networks - An Internacional Journal of Structural Analysis*, (2003). Presentada a publicación.

5. M. **GIRVAN** Y M. **NEWMAN**, *Community structure in social and biological networks*, *Proc. Natl. Acad. Sci. USA*, 99 (2002), pág. 8271-8276,

6. E. M. **ROGERS**, *Diffusion of Innovations*. Free Press, Quinta Edición, 2003.

7. J. M. **KLEINBERG**, *Authoritative sources in a hyperlinked environment*, *Journal of the ACM*, 46 (1999), pág. 604-632.

8. M. **NEWMAN**, *The structure and function of complex networks*, *SIAM Review*, 45 (2003), pág. 167-256.

9. A. Y. **NG**, A. X. **ZHENG**, Y M. I. **JORDAN**, *Stable algorithms for link analysis*, en Proc. 24th Annual Intl. ACM SIGIR Conference, ACM, 2001.

10. A. **ORAM**, ed., *Peer-to-peer: Harnessing the Power of Disruptive Technologies*, O'Reilly, Sebas-topol, California, 2001,

11. L. **PAGE**, S. **BRIN**, R. **MOTWANI**, Y T. **WINOGRAD**, *The pagerank citación ranking: Bringing order to the web, tech. report*, Stanford Digital Library Technologies Project, 1998,

12. R. **PASTOR-SATORRAS** Y A. **VESPIGNANI**, *Epidemic propagation in scale-free networks*, *Phys. Rev. Lett.*, 86 (2001), pág. 3200-3203.

13. T. H. **STANG**, F. **POURBAYAT**, M. **BURGESS**, G. **CANRIGHT**, K. **ENGØ**, Y Å. **WELTZIEN**, *Archipelago: A network security analysis tool*, en Proceedings of The 17<sup>th</sup> Annual Large Installation Systems Administration Conference (LISA 2003), San Diego, California, EE.UU., octubre 2003.

14. G. H. **GOLUB** Y C. H. **VAN LOAN**, *Matrix Computations*. The Johns Hopkins University Press, Segunda Edición, 1989.

## REIVINDICACIONES

5 1. Un procedimiento para determinar la capacidad de una red para propagar información o tráfico físico, incluyendo dicha red numerosos nodos de red interconectados mediante enlaces, incluyendo dicho procedimiento las etapas de

- mapear la topología de una red,
- computar un valor para la fuerza de enlace entre los nodos,
- 10 • computar un índice de Centralidad del Vector Propio para todos los nodos, basado en dichos valores de la fuerza de enlace
- identificar nodos que son máximos locales del índice de Centralidad del Vector Propio como nodos centrales,
- 15 • agrupar los nodos en regiones que rodean cada nodo central identificado,
- asignar un papel a cada nodo a partir de su posición en una región, como nodos centrales, nodos miembro de una región, nodos de borde, nodos puente, nodos oscilantes,
- 20 • determinar la capacidad de la red para propagar información o tráfico físico, basada en el número de regiones, su tamaño, y cómo están conectadas

**caracterizado** por

- 25 • asignar el papel de nodos miembro de una región en una región dada a todos los nodos para los que una trayectoria de enlace de ascenso pronunciado en el mapa topológico termina exclusivamente en el nodo central de esa región.

30 2. Un procedimiento de acuerdo con la reivindicación 1, **caracterizado** por computar dicho valor de la fuerza de enlace contando el número de tipos de enlaces diferentes que usa cualquier par de nodos en su interacción, usando el número de enlaces como una medida para la fuerza de enlace.

35 3. Un procedimiento de acuerdo con la reivindicación 1, **caracterizado** por computar dicho valor de la fuerza de enlace midiendo el tráfico entre dos nodos cualquiera, usando la medida de tráfico como una medida para la fuerza de enlace.

40 4. Un procedimiento de acuerdo con la reivindicación 1, **caracterizado** por computar dicho valor de la fuerza de enlace midiendo el tráfico entre cada par de nodos para cada tipo de enlace diferente, dividiendo la cantidad de tráfico en cada tipo de enlace por el tráfico total para ese tipo de enlace, usando la suma de las fracciones resultantes como una medida de la fuerza de enlace.

45 5. Un procedimiento de acuerdo con cualquiera de las reivindicaciones 1-4, **caracterizado** por organizar dichos valores de la fuerza de enlace en una matriz, la matriz de contigüidad, y computar el índice de Centralidad del Vector Propio como el vector propio principal de dicha matriz de contigüidad.

50 6. Un procedimiento de acuerdo con la reivindicación 1, **caracterizado** por asignar el papel de nodos de borde a todos los nodos que no tienen una asociación única a cualquier nodo central.

55 7. Un procedimiento de acuerdo con la reivindicación 1, **caracterizado** por asignar el papel de nodos puente a todos los nodos de borde que se encuentran en al menos una trayectoria de enlace no auto-retrazable que conecta dos nodos centrales.

8. Un procedimiento de acuerdo con la reivindicación 1, **caracterizado** por asignar el papel de nodos oscilantes a todos los nodos de borde y que se encuentran en una trayectoria de enlace no auto-retrazable que conecta dos nodos centrales.

9. Uso del procedimiento de acuerdo con cualquiera de las reivindicaciones 1-8 para evitar la propagación de virus o información perjudicial en una red.

60 10. Uso del procedimiento de acuerdo con cualquiera de las reivindicaciones 1-8 para mejorar la propagación de información en una red.

11. Uso del procedimiento de acuerdo con cualquiera de las reivindicaciones 1-8 para planificar la arquitectura de una red, para mejorar la robustez y/o seguridad y/o eficacia de comunicación en dicha red.

65 12. Uso del procedimiento de acuerdo con cualquiera de las reivindicaciones 1-8 para planificar la arquitectura de una red de energía para mejorar la robustez de dicha red.

## ES 2 285 564 T3

13. Uso del procedimiento de acuerdo con cualquiera de las reivindicaciones 1-8 para planificar una red de distribución de mercancías.

5 14. Uso del procedimiento de acuerdo con cualquiera de las reivindicaciones 1-8 para planificar una red de transporte.

10

15

20

25

30

35

40

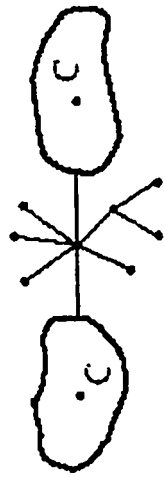
45

50

55

60

65



(derecha)



(izquierda)

Fig. 1

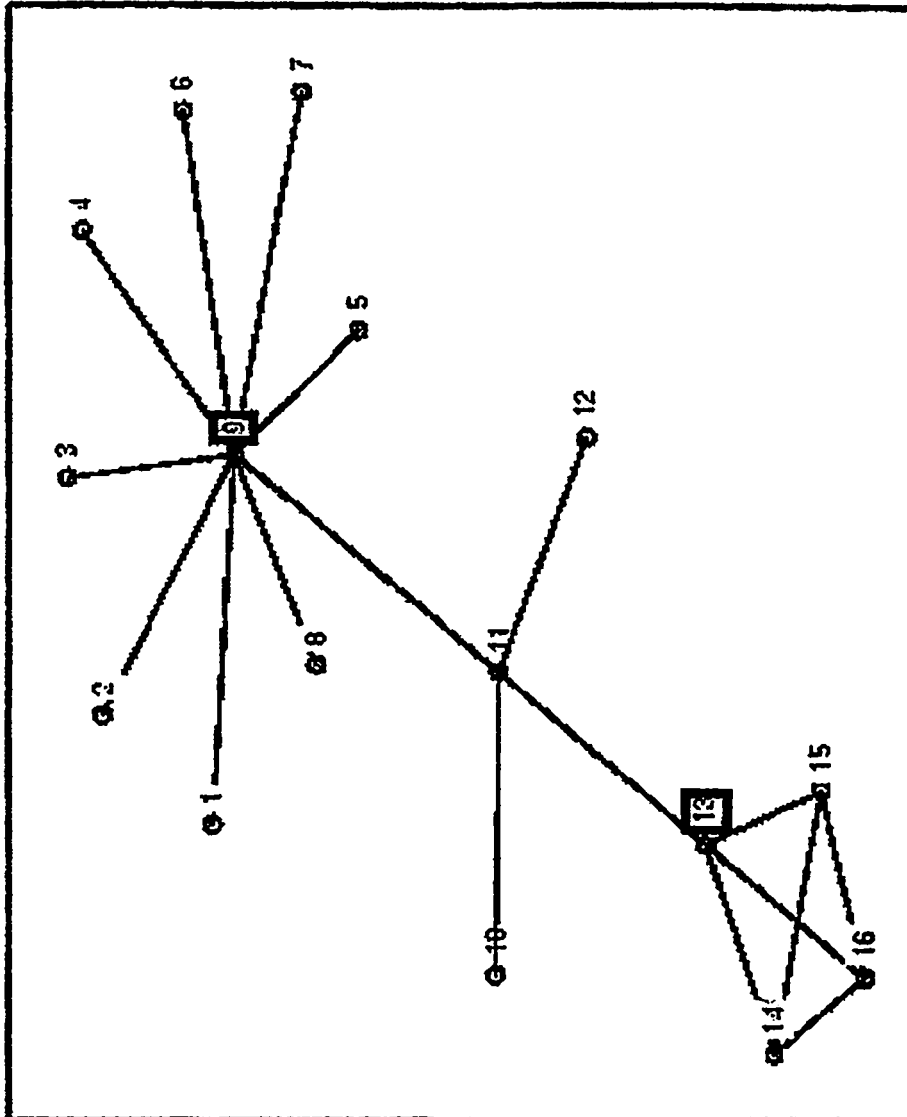


Fig. 2

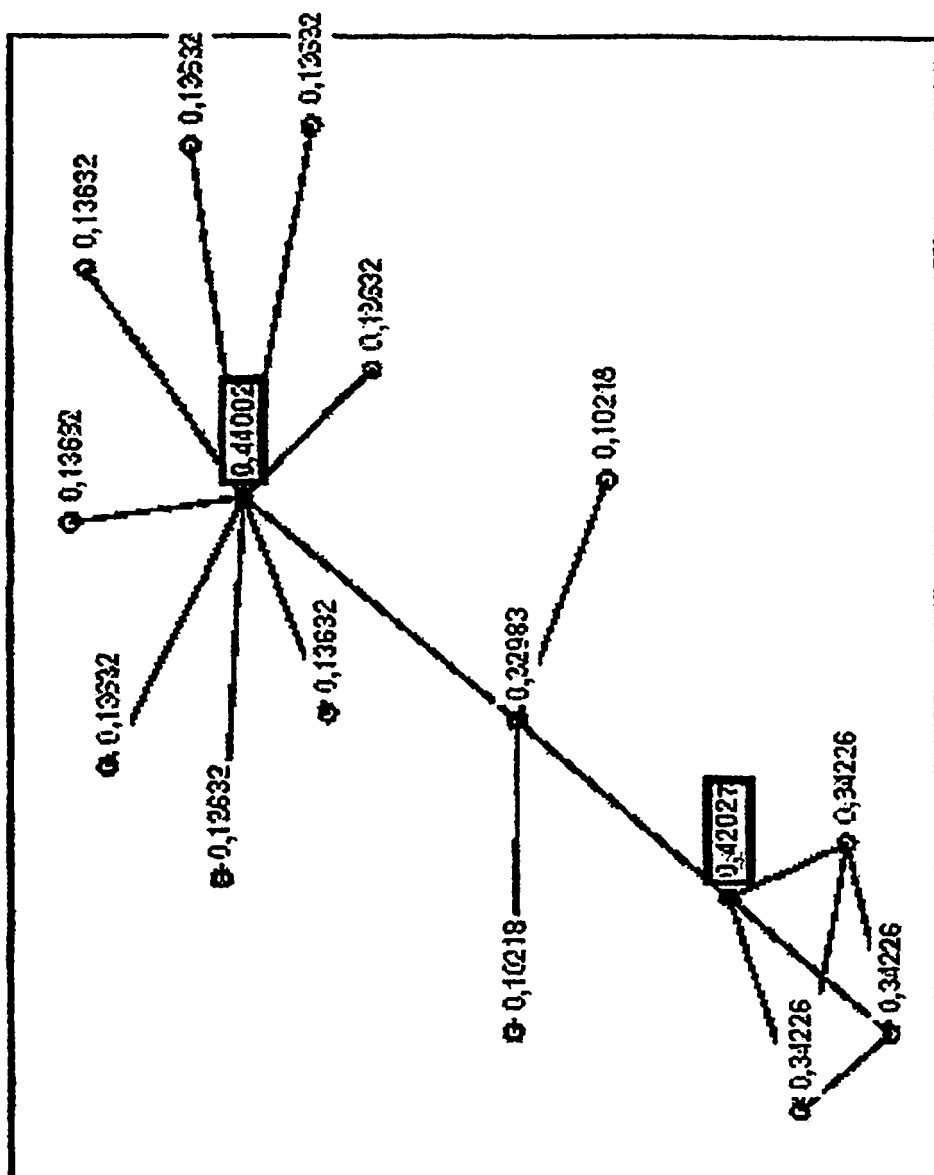


Fig. 3

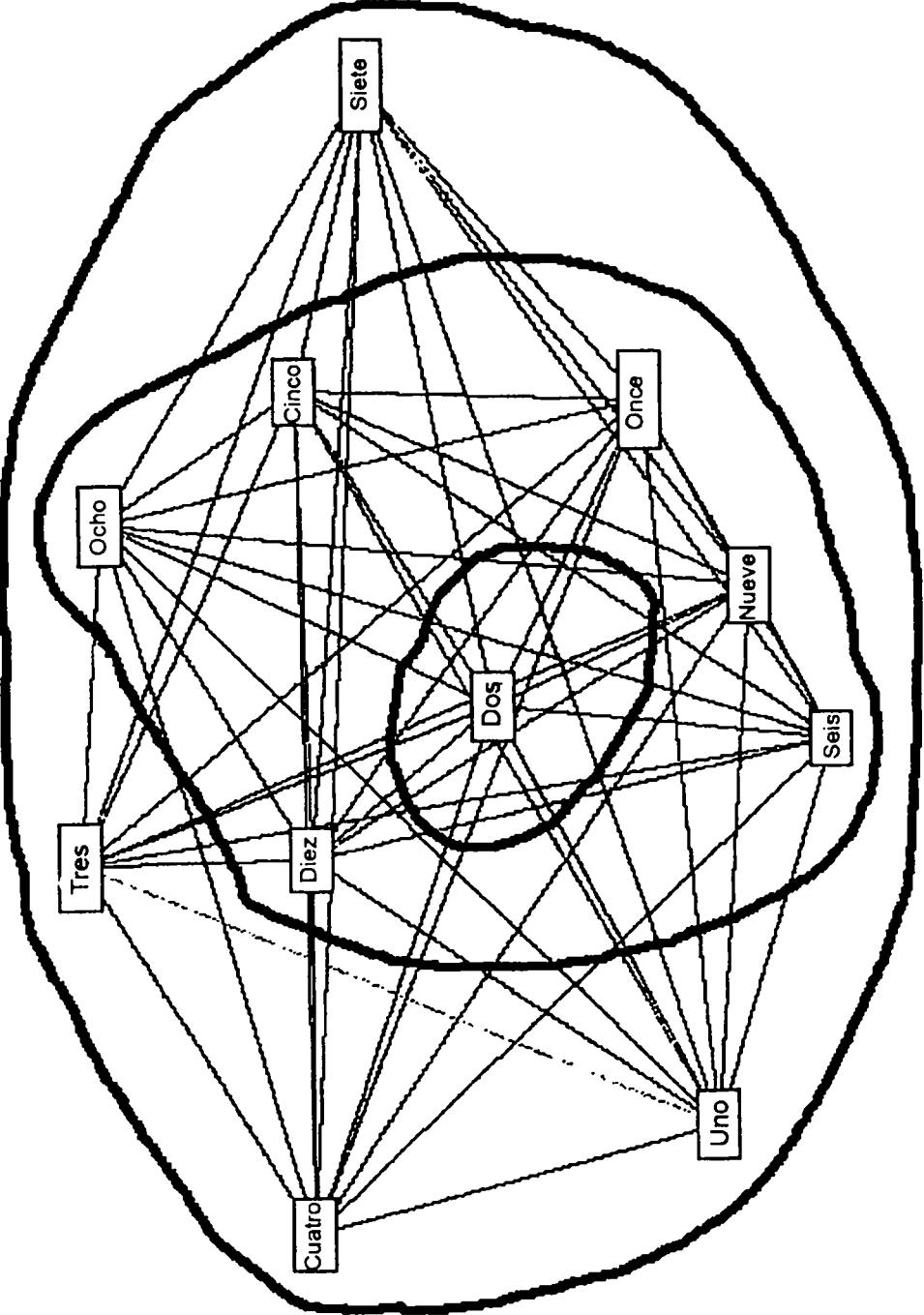


Fig. 4

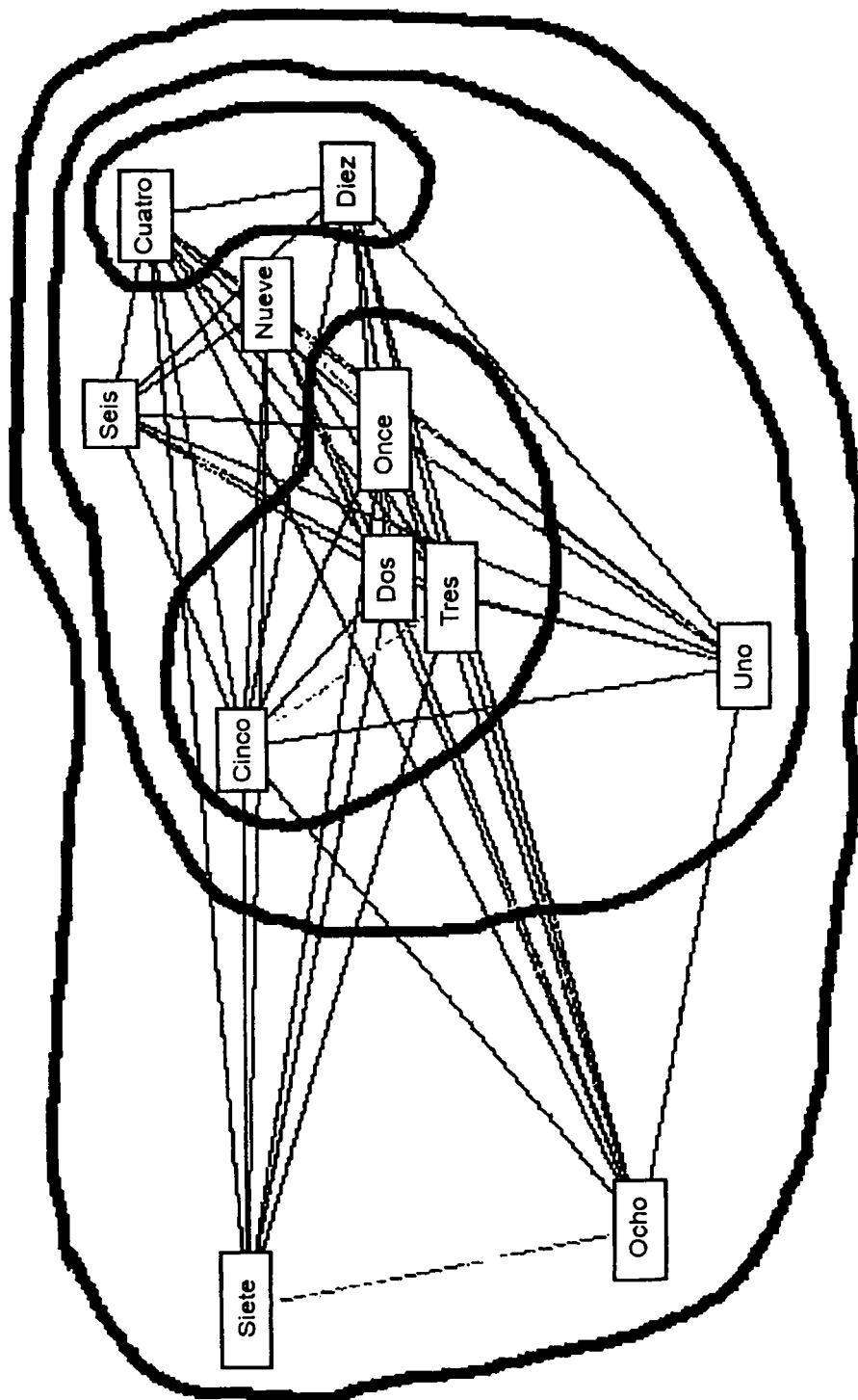


Fig. 5

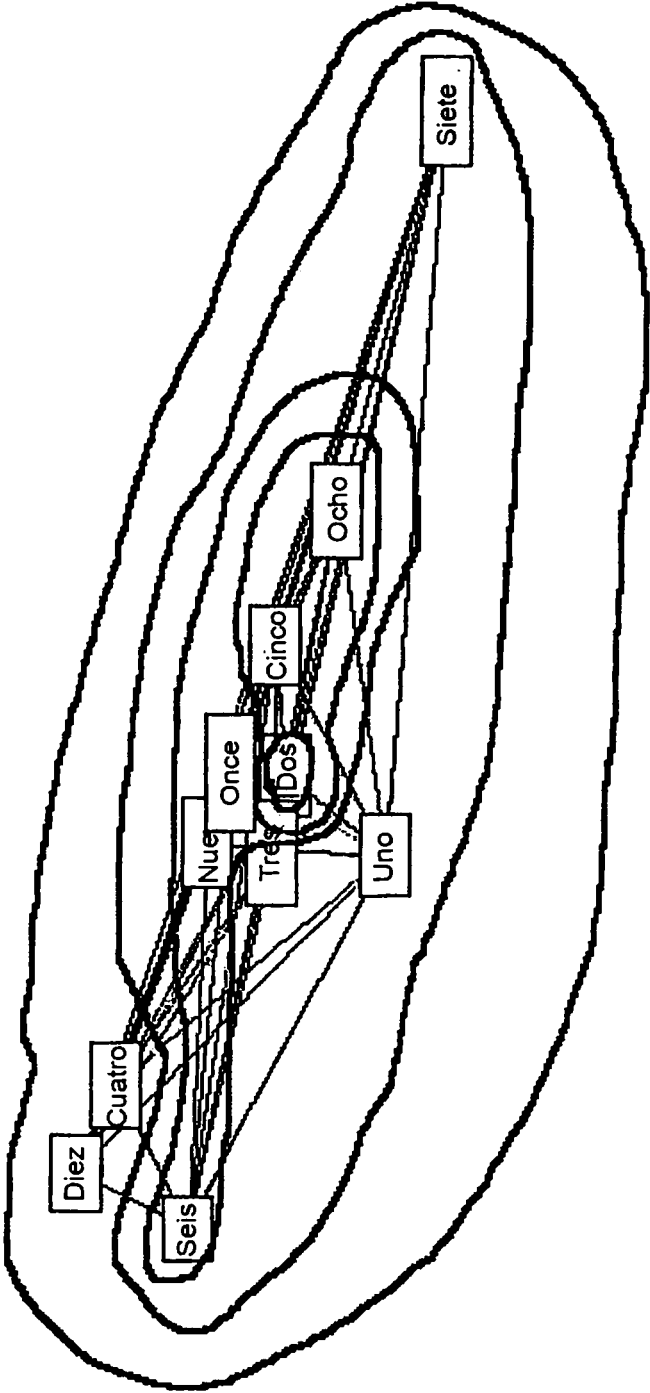


Fig. 6

