



Doctorado de Ingeniería
Tesis doctoral

Develando patrones de innovación en redes
longitudinales de citación de patentes por medio de
leyes de potencia y entropía de información

Jorge E.Mejía Q

Director: PhD Nelson Obregón Neira

6 de julio de 2015

Este trabajo doctoral está dedicado a las personas que me han apoyado durante estos años. En primer lugar a mi esposa Margarita por su infinito amor. A mis hijas María José y Silvia quienes me concienten a toda hora. A mi madre Stella, con quien compartí muchas lecturas durante este período. A mi fallecido padre, de quien heredé la paciencia para esperar que los resultados se presentaran sin tener que forzarlos. Y finalmente, a los 2 profesores que guiaron mi trabajo. Mi director Nelson Obregón, quien desde muy temprano identificó el camino a seguir, pero que esperó con paciencia a que yo mismo lo encontrara. Y Rafael Hurtado quien me mostró desde la física, la belleza del mundo de las redes.

Índice general

1. Introducción	10
2. Estado del arte	12
2.1. Economía de la innovación	12
2.2. Análisis de Redes Sociales	14
2.3. Física	15
2.3.1. Leyes de potencia	15
2.3.2. Entropía de información	16
2.4. Ingeniería de sistemas complejos	17
2.5. Cienciometría y patentes	18
2.5.1. Algunas críticas a los estudios de patentes	18
2.5.2. Redes de citación entre patentes y análisis de palabras	19
2.5.2.1. Identificación de patentes similares	19
2.5.2.2. Redes de palabras con títulos de las patentes	20
2.5.2.3. Relación entre patentes y referencias en publicaciones científicas	20
2.5.2.4. Redes neuronales, redes de publicaciones y patentes	21
2.5.2.5. Transacción entre patentamiento y producción de artículos	21
2.5.2.6. Uso de la entropía en cienciometría	21
2.6. Conclusión	22
3. Objetivos e hipótesis	23
3.1. Pregunta de investigación	23
3.2. Objetivos	25
3.2.1. Objetivo General	25
3.2.2. Objetivos Específicos	25
3.2.3. Hipótesis	25
3.2.3.1. Hipótesis Dinámicas y Macro	25
3.2.3.2. Hipótesis Dinámicas y Micro	26
3.2.3.3. Hipótesis Estáticas-Micro	27
4. Metodología	29
4.1. Tesis	29
4.2. Tratamiento de datos	31
4.3. Análisis estático de redes	31
4.4. Centralidad de autoridad y concentrador	34
4.5. Análisis longitudinal de redes	37
5. Resultados	39
5.1. Resultados del análisis estático: Aplicación de leyes de potencia	39
5.1.1. Grado entrante de las patentes, d_I	39
5.1.2. Centralidad de autoridad, C_A	40
5.1.3. Modularidad de la red Q	40
5.1.4. Distribución de grado de las redes de citaciones, $P(X_T)$	41

5.2. Resultados longitudinales-Aplicación de la entropía	52
5.2.1. Rol de las patentes en la red	52
5.2.2. Duración del tiempo de protección de una patente	53
5.2.3. Difusión de conocimiento	54
5.2.4. Entropía y antigüedad de la patente	55
5.2.5. Ganancia de información de la red	56
5.2.6. Potencial de la red	57
6. Discusión y proyección investigativa	59
6.1. Discusión de resultados	59
6.1.1. Sobre las leyes de potencia	59
6.1.2. Sobre la entropía	60
6.1.3. Sobre la innovación	61
6.2. Proyección investigativa	61
6.2.1. Estudios sobre la innovación en Colombia	61
6.2.2. Investigaciones en la línea de la tesis doctoral	64
6.2.2.1. Hipótesis pendientes de probar	64
6.2.2.2. Profundización en los otros países CIVETS	65
6.2.3. Consolidación del formalismo matemático	65
6.2.4. La dinámica de redes socio-técnicas	65
6.2.4.1. Historias de innovación y patentamiento	65
6.2.4.2. Redes jurisprudenciales	66
6.2.4.3. Seguridad	66
6.2.4.4. Redes de apoyo social	66
Bibliografía	81

Índice de figuras

0.0.1.Redes de apoyo	9
2.0.1.Estado del arte: Áreas de conocimiento y conceptos	12
2.4.1.Ingeniería de sistemas complejos.	18
2.5.1.Metared de una patente	19
3.1.1.Pregunta de Investigación	24
3.2.1.Modelo de red de citasiones de una patente innovadora	27
3.2.2.Contenido conocimiento de patentes	28
4.1.1.Combinación de los análisis estático y dinámico en la metodología.	29
4.1.2.Diagrama de bloques de la metodología	30
4.3.1.Patentes y competitividad.	32
4.3.2.Patrones de competitividad en países CIVETS	33
4.3.3.Series de tiempo de patentes de los países CIVETS registradas en la USPTO.	34
4.4.1.Roles de autoridad y concentrador	36
4.5.1.Redes longitudinales y sus marcos de tiempo	37
5.1.1.Prueba de leyes de potencia para el grado (Citasiones directas k) de todas las patentes	39
5.1.2.Prueba de leyes de potencia para la Centralidad de autoridad C_A de todas las patentes	40
5.1.3.Prueba de leyes de potencia para la Modularidad	40
5.1.4.Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 1	42
5.1.5.Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 2	43
5.1.6.Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 3	44
5.1.7.Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 4	45
5.1.8.Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 1	47
5.1.9.Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 2	48
5.1.10. Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 3	49
5.1.11. Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 4	50
5.1.12. Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_I . Grupo 5	51
5.2.1.Evidencia de comportamiento emergente en la red de la patente US3753243	52
5.2.2.Explicación de la emergencia de autoridad	53
5.2.3.Red de citasiones generada por la patente US4347294.	53
5.2.4.Mínimos Locales	54
5.2.5.Curva en forma de “S” de la patente US5133726	54
5.2.6.Difusión de conocimiento en patentes colombianas.	55
5.2.7.Difusión de conocimiento en patentes turcas	55
5.2.8.Entropía y envejecimiento de redes de patentes.	56
5.2.9.Ganancia positiva de información de la Patente US5133726.	56
5.2.10. No ganancia de información para la patente US3753243	57
5.2.11 Entropía Relativa	58
6.2.1.Mapa de grupos y áreas de investigación sobre la innovación en Colombia	63
6.2.2.Contenido conocimiento de patentes	64

Resumen

Se propone el modelo integrado por leyes de potencia y entropía de información para comprender rasgos ocultos de la innovación en países emergentes. Se han estudiado las patentes de países emergentes registradas ante la oficina de patentes de los Estados Unidos y mediante técnicas de minería de datos y análisis de redes sociales se han elaborado sus redes de citación entre patentes. Se ha determinado cuáles son las que poseen mayor autoridad por medio del cálculo de sus medidas de centralidad y para cada una de éstas se han construido sus redes anuales de citaciones desde el momento de su publicación hasta el año 2013 y se les ha aplicado el modelo propuesto. Se evidenció que existen redes de citación de patentes que exhiben comportamiento de leyes de potencia. Esta característica fue explorada con mayor profundidad mediante un análisis dinámico de las redes y se descubrió que está asociado a la difusión del conocimiento. Además, se encontró que la entropía de información permite evidenciar comportamientos emergentes en las redes, identificar la duración de la protección a los derechos de propiedad intelectual, determinar el rol de las patentes en la red, relacionar la entropía y el envejecimiento de las patentes, calcular la información ganada por la red y evaluar el nivel de incertidumbre alcanzado por cada red. La variedad de los resultados es una evidencia del poder que las leyes de potencia combinadas con la entropía de información entrañan cuando son adaptada a las redes.

Palabras claves: Patentes. Innovación. Entropía. Leyes de potencia. Redes longitudinales. Nuevos países emergentes

Agradecimientos

Haber logrado la culminación exitosa de mi formación doctoral se debe a la red de apoyo que conformé (Ver gráfica 0.0.1a). El agradecimiento a cada uno ellos se convierte en algo muy importante para mí y a esto dedico este aparte de mi documento. Combino en esta parte la prosa con la ilustración de las redes que conforman mi red de apoyo. Hablo en detalle de lo que han significado mi familia, mis amigos, mis compañeros de trabajo de la Universidad Central cuando ejercí como director del departamento de Ingeniería Industrial, mis compañeros de doctorado de la Universidad Javeriana y mis directores y jurados. En la gráfica puede observarse en la esquina superior izquierda la red completa con todas las subredes integradas. Y para facilidad visual, las he separado para efectos de la explicación que sigue.

En primer lugar está mi familia (Figura 0.0.1b). Hacia finales de 2008 cuando empecé a considerar la posibilidad de realizar el doctorado lo primero que hice fue una conversación con mi esposa Margarita y mis hijas Silvia y María José. La decisión de ingresar al programa de la Javeriana fue por lo tanto, una decisión de carácter familiar. Tenía claro desde aquel entonces que debía ser capaz de combinar el tiempo como estudiante con el tiempo familiar. Durante estos 6 años así fue y ello fue fundamental para haberlo logrado. Incluyo en este grupo a Betty quien trabaja en mi casa y siempre estuvo pendiente mi alimentación. Por supuesto que también hacen parte de mi familia mi mamá Stella, mi papá Roque y mis hermanos Pilar y Alejandro. Durante estos años de formación me volví un gran lector, eso se lo debo a mi madre, incluso con ella compartí muchos libros, los cuales devoré con pasión con el firme propósito de poder escribir mejor. Mi fallecido padre Roque merece un lugar especial en este agradecimiento. Algunas de sus características de personalidad las heredé. Compartimos la tranquilidad, la paciencia, el no dar el brazo a torcer y sobre todo, el aprender que las cosas que uno se merece le llegan sin ser forzadas. Con respecto a mi hermano Alejandro, ha estado pendiente de mi trabajo y a veces me aterrizó con sus inteligentes preguntas. Con mi hermana a pesar de la distancia, pues vive en Frankfurt, fortalecí la relación. Gracias a ella, que trabaja en el Instituto Max Planck, conseguí artículos y libros completos que no pude obtener gratuitamente desde las bibliotecas en Colombia. Por último, menciono a la familia de mi esposa. Don Mario, Doña Teresa y sus hijos. Ellos siempre estuvieron ahí y me brindaron su voz de aliento para avanzar.

En segundo lugar están algunos de mis más grandes amigos (Figura 0.0.1c). Dos de ellos, Iván y Mario, que viven en los Estados Unidos fueron esenciales para la realización de parte de mi pasantía en ese país. Iván me recibió en Houston y me brindó su hogar como si fuese el mío. Allí realicé un curso de inmersión en inglés en la Universidad de Rice. Mario a su turno, también me recibió en su hogar en la ciudad de Weston y durante un mes me dió todas las facilidades en su oficina para sentarme a escribir y estudiar. Gracias debo a ellos, mis dos grandes amigos desde la infancia. Incluyo en este listado a Camilo y Hector Pablo, mis dos grandes amigos de la Universidad de los Andes en donde estude mi pregrado. Ellos me acompañaron en la sustentación. Gracias por haber asistido.

En tercer puesto, menciono a mis compañeros de Doctorado de la Javeriana (Figura 0.0.1d). Con ellos estrené oficinas, tuve discusiones científicas y socialicé durante estos años. Mi compañera de pupitre fue Eddi Herrera. Su cercanía influyó en darle hermosura a mi trabajo incluyendo las ecuaciones matemáticas de mis planteamientos. A Julio Alberto García lo recuerdo porque siempre estaba preguntándome por mi familia, especialmente por mis niñas. Ambos hablábamos de nuestros padres y de la importancia que representamos para ellos. Mauricio Gonzalez se volvió un gran interlocutor. Con él me vi forzado a comprender muy bien el concepto de difusión desde el punto de vista físico. Pero también dialogamos mucho sobre las instituciones que nos patrocinaron el doctorado y lo que debemos hacer para transformarlas. Su compañera Paula Villegas me contagió con su gran entusiasmo para todo lo que realiza. Orlando López, el más cercano desde el punto de vista cognitivo, me debe aún unos buenos tragos para discutir sobre la naturaleza de nuestros trabajos investigativos. De Pedro García conservo su permanente caballerosidad y don de gentes. César Cardona, con su pinta de gran científico, barbudo

y sereno es alguien a quien respeto. Con César Valdés compartí mi perfil de estudiante doctoral mayor de 40 años, con familia, hijas adolescentes y responsabilidades financieras. A todos ellos les doy gracias por haberme acompañado en las diferentes simulaciones que hice de los exámenes de competencias y de candidatura. Sus preguntas o silencios me dieron respuestas a mis inquietudes.

El cuarto grupo de personas lo integran mis compañeros y amigos de la Universidad Central (Figura 0.0.1e). Primero las mujeres. Ellas son Deyanira Perdomo, Mariluz Osorio, Natalia Malaver, Eliana Pardo, Nubia Patarroyo, Paola Valdiri y Constanza Pérez. Siguen los hombres. Ellos son Jimmy Vega, Diego Suero, Oscar Buitrago, Fernando Guerra, Pablo Sánchez, Víctor Díaz, Hernando Gutierrez, Delimiro Visbal, Juan Pablo Madiedo, Jorge Ernesto Mendoza y Jair Rocha. A todos ellos presenté mi idea de proyecto doctoral, la cual aprobaron para poder acceder al apoyo de la Universidad Central. Los lazos que construimos son muy fuertes y ahora que están la mayoría en otras latitudes, les agradezco la relación que poseemos.

El quinto grupo lo componen mis directores, evaluadores y jurados (Figura 0.0.1f). Ellos pertenecen a distintas universidades y estuvieron conmigo en alguno de los exámenes que presenté. Nelson Obregón desde que llegué a la Javeriana me ofreció su apoyo y me brindó su corazón. Algún día me dijo: *“Jorge esta relación debe ser como un buen matrimonio, construido sobre la confianza y el respeto mutuo”*. Él acolitó mis caprichos y también me exigió. Le dió alas a mis sueños y abrigo a mis insatisfacciones. Voy a extrañar las madrugadas a las 6 am para nuestra reunión de seguimiento. Durante estos años, su acompañamiento fue permanente. Siempre sentí un gran respaldo de parte suya. Ahora empieza una relación de padre e hijo con él. Rafael Hurtado (Rafael Germán Hurtado Heredia) merece también un lugar en estos reconocimientos. Desde el día que lo contacté apelando a nuestro origen común en el colegio Refous, supe que construiríamos una excelente relación. Me invitó a tomar su curso de Socio-Física en la maestría de física de la Universidad Nacional y en éste encontré la inspiración para abordar mi trabajo desde el campo de la redes sociales. Me abrió por completo el panorama, reforzó mi gusto por la física aplicada a los temas sociales y me señaló el camino para asistir a dos espectaculares cursos de formación en MIT y Carnegie Mellon. Al calor de nuestro almuerzo compartido en su oficina, escuchó mis planteamientos y le realizó observaciones de la manera más académica posible. La última sesión que tuvimos antes de la presentación de la disertación me dió la seguridad y tranquilidad necesaria para exponer mi tesis. Rafael me acompañó como jurado en el examen de competencias y de candidatura. La primera prueba que afronté se denomina examen de competencias. En la misma estuvieron presentes Jorge Robledo (Jorge Robledo Velásquez), Florentino Malaver y Rafael Hurtado. Agradezco a ellos porque elaboraron un examen exigente y sus comentarios y observaciones fueron constructivos. Posteriormente, presenté el examen de candidatura y mis jurados fueron Eric Quintane, Antonio Carlos Olivera Barroso, Iván Darío Hernández. El primero es francés y trabaja actualmente con las universidades de los Andes y Melbourne (Eric Quintane). Lo consulté varias veces y me brindó su asesoría con gran generosidad. El segundo es brasileño, es investigador de la Universidad de San Pablo (Antonio Carlos Olivera Barroso) y nos conocimos mientras tomábamos uno de los cursos de la pasantía en la Universidad de Carnegie Mellon en Pittsburg. Nos hicimos buenos amigos y me dió la señal que avanzar con el concepto de entropía en las redes podía ser prometedor. A Iván Darío Hernández (Iván Darío Hernández) lo conocí desde 2008. Me presenté a su oficina de director del Doctorado de Economía de la Universidad Nacional buscando un director para mi proyecto doctoral y con una sola conversación me dió su carta de recomendación. Posteriormente, tuve varias charlas con él, al calor de un buen café, me escuchó con empatía y me aconsejó. El último examen que presenté fue la sustentación doctoral y los jurados fueron Dominique Vinck, Guillermo Parra y Rafael Gonzalez. El primero (Dominique Vinck) es un gran experto en el tema de mi investigación. Se leyó el documento de la tesis con gran detalle y me realizó preguntas muy pertinentes que me permitieron enriquecer la presente versión final. Las respuestas a sus comentarios hacen parte del capítulo de estado del arte y de la metodología. Me sentí muy feliz cuando supe que la Javeriana lo traería para el examen. Sus preguntas durante la sustentación me hicieron sentir muy seguro de mi defensa. El otro jurado fue Guillermo Parra (Guillermo Parra). Lo contacté inicialmente por su conocimiento en el tema de patentes y posteriormente, recibí sus observaciones actuando como jurado. La respuesta a sus inquietudes se manifiestan en el diagrama de bloques que me pidió realizar para resumir la metodología de mi investigación. El último jurado fue Rafael González (Rafael Gonzalez). Él presidió el jurado y su pregunta final sobre que le diría a Loet Leydesdorff (Loet Leydesdorff) si pudiera trabajar con él, me dió la tranquilidad que mi defensa había sido exitosa. En resumen, todos mis jurados fueron importantes en esta investigación. Actuaron como profesores y dieron lo mejor para mi proceso de formación.

Institucionalmente, quiero dar un reconocimiento a la Universidad Javeriana. Me brindó un lugar de trabajo cómodo y agradable. Me apoyó con los recursos financieros y humanos para la realización del proyecto interno Uso de herramientas de la complejidad para estudiar la innovación. Este proyecto me posibilitó explorar mi tema

y fue importante para mi investigación. La Javeriana me patrocinó la asistencia a varios eventos y la financiación de parte de la pasantía en las universidades de MIT y Carnegie Mellon. En la biblioteca conté con todo el apoyo para obtener los documentos que necesitaba y fue clave el acceso que me dieron al programa VantagePoint para la minería de datos. En el campus de la universidad reforcé mi fé en Dios.

Por último debo agradecer a la Universidad Central que me patrocinó los estudios. Lugar especial merecen en este reconocimiento los directivos que diseñaron la política institucional de apoyo a estudios de postgrado. El rector Guillermo Páramo, que supo darle un nuevo rumbo a la universidad y la encaminó en la dirección de la acreditación. Una de las políticas que hacen parte de dicho camino, es precisamente el apoyo a procesos de formación de sus profesores. Pablo Leyva fue en algún momento mi jefe directo cuando fue decano de la facultad de ingeniería y posteriormente, tuvo participación directa como vicerector académico en el diseño de dicha política. Así mismo, debo incluir a Julio Mario Rodríguez quien fue mi jefe como decano de la facultad y conté con su apoyo mientras estuvo en dicho cargo. Igualmente, hice parte del grupo Complexus dirigido por él y allí encontré mi tema de investigación. Quiero registrar en esta lista a Martha Baracaldo de la Escuela de Pedagogía y responsable del seguimiento a los profesores beneficiarios del apoyo. Ella me asesoró y defendió ante la universidad en los momentos que tuve que justificar el haber sobrepasado los 4 años para la culminación de los estudios. Finalmente, agradezco al nuevo decano, Oscar Herrera, quien también fue beneficiario del programa de apoyo y que estuvo presente en mi defensa doctoral y con quien debemos abrir camino en la universidad para encontrar el lugar que merece la investigación en nuestra institución.

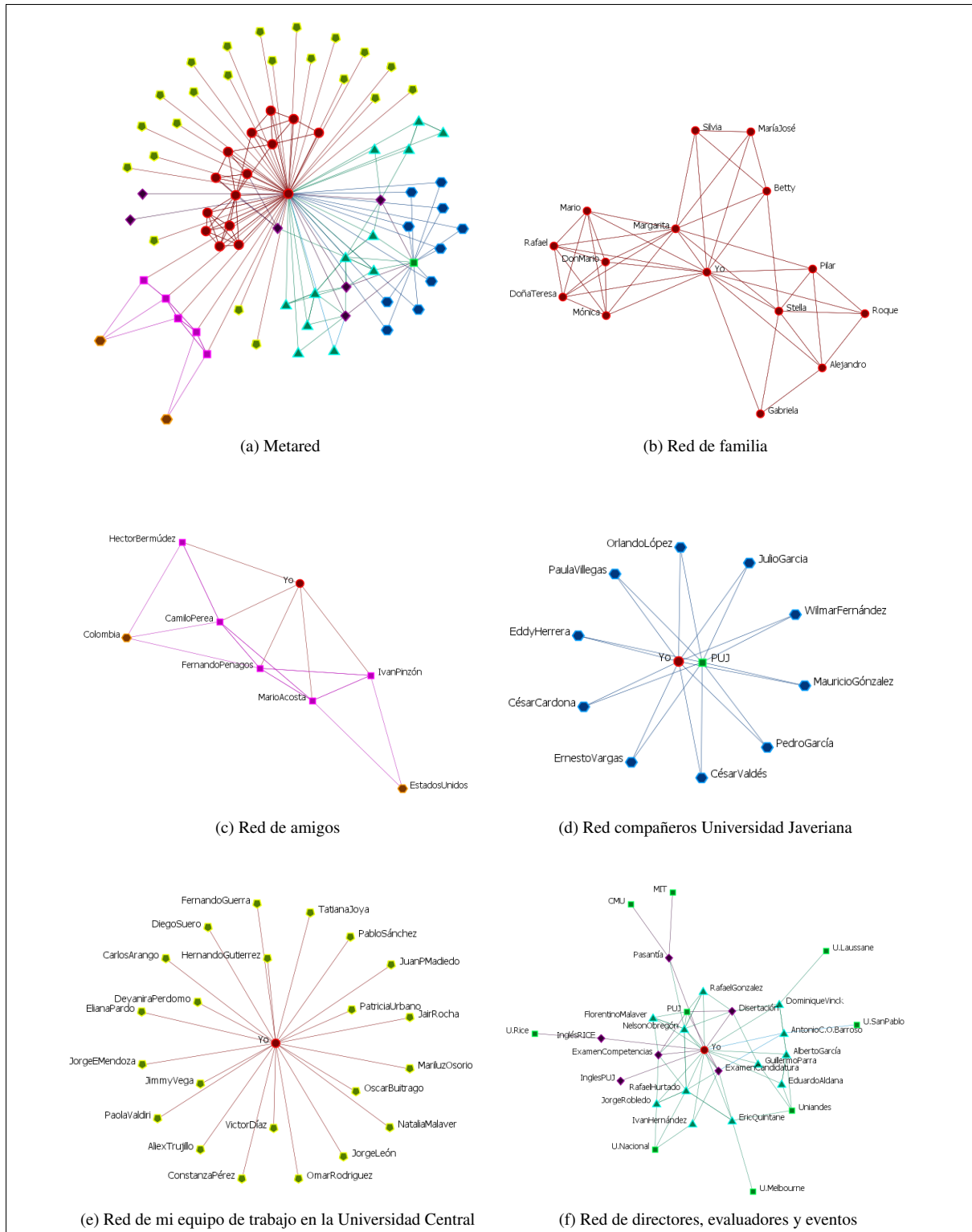


Figura 0.0.1: Redes de apoyo

Capítulo 1

Introducción

Humans interact with one another. Is it possible that the dynamics of human societies are self organized? (Bak, 1997)

Esta frase ha sido parafraseada por el autor de esta tesis doctoral de la siguiente forma: Los inventores y evaluadores de patentes interactúan entre sí por medio de las patentes. ¿Es posible que el conjunto de redes de patentes generado por ellos sea auto-organizado? La respuesta a esta pregunta se aborda desde dos conceptos centrales: Las leyes de potencia y la entropía de información (Shannon, 1948) aplicados a redes longitudinales de citación de patentes. Trabajar en conjunto con estos dos conceptos aplicados a un conjunto de redes que poseen una dinámica temporal permitió formular la tesis doctoral que se enuncia así: **Un modelo para análisis de redes de citación de patentes integrado por los conceptos de leyes de potencia y entropía de la información permite descubrir patrones de la dinámica de difusión de la innovación.** De esta forma, el estudio de la tesis doctoral en el presente trabajo está organizado de la siguiente forma: Un estado del arte, seguido por la metodología ejecutada, los resultados y la discusión de los mismos.

El estado del arte cubre los principales temas que proveen el marco conceptual de la investigación. Se inicia con la economía de la innovación haciendo un recuento desde la innovación schumpeteriana pasando por la economía del conocimiento tan popular en nuestros días y llegando al concepto de redes de innovación. Éstas a su vez, son estudiadas desde el Análisis de Redes Sociales (ARS). Este campo constituye en el elemento central de la modelación de la información relacional de las redes de patentes. Así mismo, el ARS permitió el vínculo con la física. Desde la física se profundizó en la comprensión de la entropía y las leyes de potencia que se convirtieron en los conceptos tanto físicos como matemáticos que permitieron la formalización matemática del modelo propuesto. También hace parte del estado del arte el campo de la ingeniería de sistemas complejos. Al final, se incluye una revisión de la literatura correspondiente a la ciencia de redes y a las patentes, que es el área en donde se ubica específicamente mi investigación doctoral.

El capítulo 3 se concentró en la explicación de la metodología usada para la definición de los objetivos e hipótesis de la investigación. Se destaca entonces en detalle los pasos que se siguieron para la formulación de los objetivos general y específicos. Así mismo, se incluyeron en este capítulo las hipótesis de trabajo que guiaron la investigación. Este corto capítulo recopila la formulación del problema y las preguntas que condujeron al estudio del mismo.

En el capítulo 4 se aborda el modelo de análisis propuesto en el trabajo. Se inicia el capítulo con la enunciación de la tesis de la investigación y sigue con una explicación muy específica del modelo propuesto. Se explican los componentes del mismo. Es decir, el análisis estático y el análisis longitudinal de redes. Con respecto al primero, se escogen un grupo de países mediante el análisis de su competitividad. Posteriormente, se precede a la minería de datos de patentes de los países previamente seleccionados. Se continúa con la limpieza de los datos obtenidos. Y se remata el análisis estático con el ARS aplicado a las redes de citación de patentes que emergen de las patentes de los países seleccionados. Enseguida, se aborda el análisis longitudinal de las redes. Este análisis se compone de la aplicación de los conceptos de entropía y sus métricas derivadas al conjunto de redes de citación que cambian en el tiempo. En resumen, este capítulo explica en detalle la forma cómo se aplicó la tesis al grupo de datos seleccionados.

El capítulo 5 presenta los resultados de la aplicación del modelo. Se subdivide el capítulo en 2 partes. La primera dedicada a los resultados que son fruto del análisis estático. Estos resultados están particularmente

asociados al concepto de leyes de potencia. Se realizó medición de leyes de potencia a cuatro conceptos: El grado entrante de la patentes, la centralidad de autoridad de cada patente en la red generada, las modularidades de las redes y por último, la distribución de grado de las redes. La segunda parte del capítulo se dedicó a los resultados de la aplicación de la entropía. Los resultados se clasificaron de la siguiente forma: Rol de las patentes en la red, Duración del tiempo de protección de una patente, difusión de conocimiento, entropía, antigüedad de la patente, ganancia de información de la red y potencial de la red.

Por último, en el capítulo 6 se realizan los comentarios a los resultados y la proyección investigativa. La parte inicial se focaliza en la discusión de los resultados que arrojó el uso del concepto de leyes de potencia. A continuación, se destacan los resultados alrededor la aplicación de la entropía. Se finaliza esta parte con las implicaciones sobre la innovación. En cuanto a la proyección investigativa, se realiza una reflexión y formulación de las líneas de acción investigativa que se pueden desprender de la tesis. En primera instancia, se ubicó al autor de la tesis en el panorama nacional. Para ello se recurrió a un análisis de redes sociales destancado los investigadores, los grupos, los temas alrededor del campo de la innovación en Colombia, así como el hueco estructural que puede ocupar el autor de la tesis. Posterior a este ejercicio, se esbozaron algunas de las líneas de investigación que se pueden derivar de la tesis. Genéricamente, las líneas se componen de las hipótesis que no se pudieron probar, el formalismo matemático del modelo y los estudios de dinámica de redes socio-técnicas. Este capítulo señala el futuro de la investigación y la forma que debe adoptar para su consolidación como proyecto de investigación.

Capítulo 2

Estado del arte

¿Por qué un estado del arte con artículos y conceptos tanto centenarios como clásicos? es la pregunta que busca responder este apartado. El estilo del estado del arte que se presenta retoma temas clásicos y los contrasta con aplicaciones actuales. Desde el principio y hasta el final de mi formación doctoral me encontré con temas clásicos y además, desarrollé un placer muy especial por comprender los postulados consultando los escritos realizados por los autores originales. Es más, solamente pocas personas pueden comprender la alegría que me da encontrar en medio magnético copias de los artículos originales escritos en máquina de escribir y tenerlos en mis manos para estudiarlos. En consecuencia, la búsqueda de este tipo de trabajos se volvió algo muy importante para mí, pues los temas que conforman mi trabajo son en verdad clásicos, pero han adquirido una renovada actualidad. Los temas a los cuales hago referencia son: Innovación, entropía, análisis de redes sociales y leyes de potencia. Los mismos se muestran en el siguiente gráfico relacionados con el campo o área de conocimiento de donde son originarios.



Figura 2.0.1: Estado del arte: Áreas de conocimiento y conceptos

2.1. Economía de la innovación

Las contribuciones de la economía tienen un espectro conceptual muy amplio que va desde la primera definición de innovación, pasando por conceptos asociados a la idea de red hasta recientes estudios sobre patentes. A comienzos del siglo pasado, fue publicado uno de los trabajos más influyentes en el pensamiento económico sobre la *innovación* (OECD, 2006, p.56). Se trata del libro escrito por Schumpeter (1911), donde se plantea su idea de la *destrucción creativa* de los aportes de los empresarios al crecimiento económico por medio de sus

innovaciones. Aunque esta posición parece hoy natural en nuestro contexto, y está incorporada en los 2 últimos planes de desarrollo (DNP, 2007, 2011), duró muchos años olvidada por los economistas y, en consecuencia, no fue tomada en cuenta como parte de las políticas económicas de muchos países. Sólo varias décadas después, la economía volvió a interesarse en el tema de la innovación (Freeman, 1991, p.499) y Nelson and Winter (1982) exponen sus planteamientos desde la *teoría evolucionista* para explicar la *innovación tecnológica* (OECD, 2006) y la competencia entre las empresas, su texto es una fuerte crítica al enfoque de la economía clásica para explicar los mismos fenómenos.

Por esa misma época, Freeman (1991) relanza una definición de *redes de innovadores* propuesta por Imai and Baba (1989). A mediados de los años 90, Economides (1996) estudió la similitud entre la estructura económica de redes y la estructura vertical de sectores económicos relacionados. Posteriormente, Powell and Grodal (2005) proponen la idea de *redes formales e informales de innovadores* y exponen una distinción entre el tipo de vínculos, dándoles una gradación entre formales, informales y flexibles.

Hacia comienzos de los 90's, aparecen en el campo de la economía, otros planteamientos sobre la innovación asociados a la idea de red. Cabral (1990) relaciona el tema de la innovación con los conceptos de la *curva S de adopción de la innovación* y de *externalidad de red* (Oster, 2000, p.395). En primer lugar, argumenta que la innovación puede ser interpretada como una aproximación a un punto discontinuo (catástrofe) de una trayectoria de equilibrio. En segundo lugar, sustenta la forma cómo las externalidades de red favorecen la difusión de innovaciones, incluidas las normas de calidad (Cabral, 2000). Partiendo de esta idea del tratamiento de la introducción de normas técnicas de calidad en un ambiente socio-económico como una clase de innovación, Albuquerque et al. (2007) realizan un estudio de difusión espacio temporal de normas técnicas.

Otro enfoque sobre la innovación, ha sido liderado por Lundvall (1992); Lundvall and Vinding (2004) quienes se interesan por investigar los *sistemas nacionales de innovación*. En este caso se estudia la influencia de las organizaciones externas sobre las actividades de innovación de las empresas y otros actores. Se enfatiza la importancia de la transferencia y la difusión de ideas, conocimiento, información y signos. Los canales y redes por medio de los cuales éstos circulan están embebidos en un telón social, político y cultural que guía y restringe las actividades y capacidades de innovación.

Ya en el siglo actual, Leydesdorff (2006); Etzkowitz (2000) comparan la *economía basada en el conocimiento* (OECD, 2005, p.28) con la triple hélice del ADN en donde la estrecha colaboración entre empresa-universidad-gobierno es fundamental para el favorecimiento de la innovación y proponen entonces el *modelo de la triple hélice de la economía basada en conocimiento*. Por su lado, Goyal (2009) se ubica en el área de la *economía de redes* especialmente en los sectores que tienen estructura en forma de red tales como las aerolíneas, comunicaciones, electricidad y la industria del software. Otros autores que investigan sobre el tema de redes y economía son Ten Kate and Niels (2006) quienes desarrollan en el ambiente de redes una relación entre las *curvas FEED* y determinados momentos en los cuales se requiere establecer precios por debajo del costo marginal para que la curva de la demanda alcance un punto crítico con pendiente positiva.

Así mismo, desde la economía el estudio de las *patentes* (SIC, 2008) permite realizar una mirada a la economía del conocimiento y la innovación. Al principio de la actual década Jaffe and Trajtenberg (2002) ejecutaron un proyecto para convertir la información de la Oficina de Patentes de los Estados Unidos (www.uspto.gov) en una fuente de conocimiento para la realización de estudios que usan las patentes como aproximación a la innovación. Al final de la misma década, Acemoglu et al. (2011) demuestran que, bajo ciertas circunstancias, las patentes motivan la experimentación por parte de potenciales innovadores y permiten la transmisión socialmente beneficiosa del conocimiento a través de las empresas. En primer término, el principal argumento en favor de las patentes es que favorecen la innovación *ex ante* al crear rentas monopolísticas *ex post*. Por otro lado, desde el punto de vista tecnológico, las patentes han sido usadas para medir la contribución industrial en cluster tecnológicos (He and Fallah, 2011), (Owen-Smith and Powell, 2004). Y también han sido estudiadas para analizar el tema de la cooperación entre empresas (Allarakhia et al., 2007). Por lo tanto, como objeto de estudio permiten flexibilidad para observar situaciones económicas, sectoriales y organizacionales.

En resumen, la amplitud de campos de investigación que ofrece la economía ha permitido la ubicación de definiciones muy robustas sobre innovación, así mismo, permite diversas asociaciones entre las ideas red e innovación y el trabajo desde la economía de conocimiento basado en el estudio de las patentes y su contribución a la innovación.

2.2. Análisis de Redes Sociales

El campo del Análisis de Redes Sociales (ARS) se ha desarrollado y crecido desde hace 80 años con la incorporación de aportes desde diferentes áreas de conocimiento como la sociología, la sociometría, la psicología, la antropología, la teoría de grafos, la informática y la ciencia de las redes (Newman, 2006, 2010). Del interior de la sociología, surgen el campo de investigación denominado Análisis de Redes Sociales, la idea de redes tecno-económicas y la teoría del actor-red, los cuales contribuyen a la comprensión del concepto de redes de innovación.

Las primeras inquietudes surgen de la observación cuidadosa de las interacciones entre grupos de personas, especialmente de aquellas con manifestaciones de problemas psiquiátricos. El pionero en el análisis de redes sociales para estas problemáticas fue Moreno and Jennings (1938), considerado como el padre de la sociometría y del sociodrama. Él combinó la estadística y la sociometría para estudiar lo que denominaron *configuraciones* de los grupos sociales, que hoy conocemos bajo el concepto de redes sociales. Leer su documento original sobre la sociometría fue una lección de matemáticas aplicada a las relaciones entre las personas.

Siguiendo la línea de las matemáticas de las redes, o teoría de grafos, encontré otro par de clásicos. En primer lugar, se encuentra Erdos and Renyi (1960) quien es un campeón y un ícono para los amantes de las matemáticas y que además se miden su número de Erdos (Oakland_University, 2015), para determinar qué tan cerca se encuentran de él. Este matemático, es también un ejemplo para la ciencia de redes, que es el área cognoscitiva en la cual mi tesis está ubicada. Su modelo ER model (Ver Definición 1) es usado como referencia para comparar las estadísticas producidas por diferentes tipos de redes, pues su topología corresponde a aquella que posee mayor grado de aleatoriedad, todas las redes son equiprobables y por lo tanto, es el modelo que posee mayor grado de *entropía*. En los años 60, el ARS sigue creciendo y enriqueciéndose con las contribuciones realizadas por las matemáticas aplicadas a las redes, conocidas como Teoría de redes o Teoría de grafos. Para la conformación de esta teoría se destacan los aportes realizados por Erdos and Renyi (1960). Igualmente, se tiene que reconocer la importancia de los trabajos de Harary (1969) para el completo establecimiento de la teoría de grafos con la aparición de textos, journals, conferencias especializadas, etc. Unos años después, Milgram (1969) con su famoso experimento del problema del mundo pequeño (Ver Definición 2) popularizó las redes de mundo pequeño. Posteriormente, Granovetter (1978, 1983) mostró que en una red con presencia de *lazos fuertes* y *lazos débiles* los últimos tienen mayor importancia para los procesos de difusión en la red.

En la siguiente década, el otro clásico es Wasserman and Faust (1994) quien compendió en su texto los desarrollos matemáticos para la realización del análisis de redes sociales. Su libro es referencia obligada para los que escriben sobre análisis de redes sociales. Para mí fue importante, porque en la sección de futuras áreas de desarrollo, él expuso el tema de las redes longitudinales, en concreto dijo:

Good, easy-to-use methods for longitudinal network data would be an important addition to the literature (Wasserman and Faust, 1994, p. 731)

Esta frase me motivó a leer a quienes se pusieron a trabajar en la dirección señalada por Wasserman. En esa vía, estudié los modelos de Kossinets and Watts (2006b); Butts (2008); de Nooy (2011) (ver Definición 3, 4 y 5) y a decir verdad, aún no cumplen con lo exigido por Wasserman es decir fáciles de usar. Por eso, consideré estimulante el avanzar en la dirección de un modelo bueno y sencillo de usar.

Posteriormente, Callon (1991) plantea la idea de las *redes tecno-económicas* las cuales son empleadas para describir las interacciones entre investigadores, tecnólogos, ingenieros, usuarios e industriales para la creación científica y técnica, su difusión y consolidación de resultados. Su aproximación teórica se conoce como *sociología de la traducción* y consiste en que por medio de procesos de *problematización*, *interesamiento*, *enrolamiento* y *movilización* se puede construir redes con nodos heterogéneos (actores-redes) para generar y difundir innovaciones de manera interactiva, en un mercado que también es un arreglo socio técnico (Callon, 1986, 1991, 1995, 2001b, 2007). Durante la pasada década, Callon (2001a) contribuye nuevamente con su *teoría del actor-red (ANT)* a la comprensión de la manera cómo el emprendedor configura todo un mundo de relaciones que requiere armar para la obtención de resultados innovadores.

En conclusión, la sociología usa el paradigma de las redes para dar cuenta de las redes sociales, las interacciones en las redes de carácter tecno-económico, así como, las actividades del emprendedor. Por lo tanto, las redes se convirtieron en el concepto central para el modelamiento de mi trabajo doctoral.

2.3. Física

Por el lado de las ciencias básicas, la física ha hecho especiales contribuciones al tema de interés. Sus aportes incluyen las aplicaciones alrededor del concepto de entropía, la idea de redes complejas y la creación de campos de investigación como son la Socio-Física y la Econo-Física. Desde un punto de vista muy personal y que comparto con uds, me costó mucho esfuerzo intelectual comprender los conceptos de entropía y leyes de potencia y su aplicación a un objeto muy sencillo como lo es una red. Es posible que por esta razón me centrara en la búsqueda de artículos científicos en el área de la física que es donde nacen estas ideas y donde estuvieran juntas las palabras claves de entropía, leyes de potencia y redes.

2.3.1. Leyes de potencia

¿Qué significado tiene encontrar leyes de potencia? o en otras palabras ¿qué se encuentra detrás de la expresión $p(k) \sim k^{-\alpha}$?

Hoy en día, un área que ha cobrado un especial interés entre muchos científicos son las denominadas ciencias de la complejidad. En estas ciencias, cuando se demuestra la existencia de comportamiento de leyes de potencia (Auerback, 1913; Zipf, 1949) en algún fenómeno de interés se dice que hay presencia de un comportamiento emergente y por lo tanto, muy interesante de analizar con mayor profundidad. Pues a pesar que no soy un seguidor de esta corriente de pensamiento científico, si tengo a mi alrededor muchos expertos que la comprenden y practican. En consecuencia, tomé la decisión de medir este comportamiento en las redes que construí y con gran sorpresa encontré que en una muestra de ellas se presentaba dicho comportamiento.

Los fenómenos ya sean sociales o naturales que exhiben comportamiento de leyes de potencia, en alguna de sus variables descriptivas, poseen propiedades intrigantes o se comportan de manera sorprendente (Newman, 2010, p. 249, 255). En cuanto a los sistemas sociales, la comprensión de ciertos fenómenos que se presentan en sistemas de origen humano ha estado asociada a la presencia de leyes de potencia. Recientemente, Bohorquez et al. (2009) encontraron comportamientos de leyes de potencia en algunas estadísticas de la guerra en Colombia, que han permitido comprender el carácter sui-generis de nuestro conflicto. De la misma forma, se asocia la aparición de ciertas regularidades en sistemas de origen humano, a la presencia de leyes de potencia. Algunos de estos mecanismos sorprendentes, como la difusión, están asociados a la estructura del sistema en el cual ocurre. Es entonces, importante decir que *“la auto-organización ha sido usada por muchos años para describir la habilidad de determinados sistemas que están en desequilibrio para desarrollar estructuras y patrones en la ausencia de control o manipulación por parte de un agente externo*¹ (Nicolis, 1989) citado por Bak (1997, p.2).” Para los sistemas organizados en forma de red, como las redes de citación de patentes, Barabasi and Bonabeau (2003) demostraron que las denominadas redes libres de escala ofrecen la estructura necesaria para que fenómenos como la difusión ocurran.

Con respecto a las patentes, éstas han sido estudiadas desde las leyes de potencia por diversos autores. En primer lugar, O’ Neale and Hendy (2012) investigaron el tema y encontraron asociación entre el número de solicitudes de patentes y las características innovadores de un grupo de países pertenecientes a la OECD (Organisation for Economic Co-operation and Development). Chen et al. (2011), a su vez se concentraron en la búsqueda de vínculos ocultos para identificar patentes relevantes, para ello usó el principio de Pareto, similar al concepto de leyes de potencia, para determinar el umbral de la fuerza de acoplamiento bibliográfico. Otros autores usaron el modelo de leyes de potencia para analizar redes similares a las citaciones entre patentes. Por ejemplo, Xu et al. (2005) se focalizaron en las redes que crecen con el tiempo y encontraron en el mecanismo de vinculación preferencial la explicación a dicho fenómeno. Sheridan et al. (2012) también se interesaron en las redes que crecen y se dedicó a medir el mecanismo de vinculación preferencial usando simulación de Monte Carlo en cadenas de Markov. Por su parte, Ravasz and Barabasi (2003), estudiaron redes complejas que presentan estructura jerarquizada y modular y descubrieron que los nodos concentradores juegan un papel importante para efectos de comportamientos dinámicos de las redes. Igualmente, el trabajo de Rahmandad and Sterman (2008) se centró en los temas de heterogeneidad y la estructura de red para relacionarlo con el tema de difusión.

De forma complementaria, los temas de redes libres de escala (Scale Free Networks) o que exhiben comportamiento de leyes de potencia también han sido ampliamente estudiados. Dentro de este grupo se encuentran los siguientes trabajos: En primer lugar, se menciona el trabajo pionero de Barabasi et al. (1999)

¹Traducción propia

que centró su interés en las redes aleatorias y metodológicamente empleó la teoría del campo medio (Mean Field Theory) para demostrar que estas redes exhiben comportamiento de leyes de potencia. Posteriormente y de manera similar, Mitzenmacher (2004, p.5) planteó que un criterio importante para un modelo de redes aleatorias es que éste produzca distribuciones de leyes de potencia para el grado entrante y saliente. El mismo Mitzenmacher (2004, p.8) estudió redes que tienen estructura comunitaria o modular jerarquizada y propuso que esta es una propiedad que no se encuentra en las redes aleatorias. Estas redes que presentan estructura modular y jeraquizada también han sido muy estudiadas. Por ejemplo, Girvan and Newman (2002) las analizaron y propusieron un algoritmo basado en los indicadores de centralidad para identificar los agrupamientos o módulos. A su vez, Wu et al. (2012), encontraron que las transiciones de fase permiten detectar comunidades o agrupamientos. Por otra parte, Bommarito II et al. (2010), estudiaron un tipo particular de estas redes, denominadas por ellos como redes dinámicas de citación, entre las cuales se consideran las redes entre artículos, patentes o documentos legales. Su trabajo, propuso medidas sobre la distancia existente entre los nodos tomando como base las relaciones compartidas entre nodos. En consecuencia, la estructura modular y jeraquizada está estrechamente asociada a las redes que presentan comportamiento de leyes de potencia.

Por lo tanto, el estudio de la difusión de la innovación a través de redes de citación entre patentes está relacionado con la aparición de leyes de potencia. Y esta tesis doctoral evalúa la aparición de este comportamiento en las redes de citación y devela qué oculta mediante el uso de la entropía de información.

2.3.2. Entropía de información

La entropía me persiguió desde que ingresé al programa de doctorado en ingeniería de la Javeriana. Al final del 2008, cuando me presenté al programa en búsqueda de un director, me recibió el profesor Obregón, director en aquel entonces tanto del Doctorado como del Instituto Geofísico. Este dato es importante, pues señala una cercanía a conceptos provenientes de la física. Una vez aceptado en el programa empecé a tomar los cursos y en uno de ellos tuve la ocasión de asistir a una clase de José Daniel Muñoz profesor del departamento de física de la Universidad Nacional (<http://www.docentes.unal.edu.co/jdmunozc/>). En su curso, él explicó el concepto de la entropía y algunas de sus distintas interpretaciones, que van desde la idea de calor hasta la idea de incertidumbre. Esta versatilidad me pareció muy interesante. Posteriormente, tomé el curso del profesor Vizcaya de la Javeriana (CvLac) sobre teoría de la información en donde tuve la oportunidad de aprender sobre la entropía de información (Shannon, 1948) (Ver Definición 6.2.15). En ese momento, segundo semestre de 2009, ya empezaba yo a acercarme a la idea de redes y me pareció muy curioso que en el texto guía del curso (Cover, 2006) no hubiese ni una sola referencia a la entropía de una red. En 2010 tomé uno de los Seminarios de Investigación del profesor Obregón y entre los muchos temas que aprendí, me gustó bastante el concepto de ganancia de información (Quinlan, 1986) y yo ya empezaba a esbozar mi idea de una red en donde se gana información a medida que crece. A mediados de 2011, tuve la ocasión de asistir a un curso de análisis de redes en Carnegie Mellon University y aunque es tan sólo un dato curioso, la cafetería se llama Entropy. En ese instante, entendí la frase del profesor Obregón: “El mundo conspira contra uno” y tomé la decisión de avanzar en la construcción de la métrica basándome en la entropía. De manera que se reafirmó mi interés por seguir encontrándole aplicaciones a este clásico concepto acuñado hace mucho tiempo por Clausius (1865) y asociado al famoso Boltzmann (1872) según explica Muller (2007, p.72-100) en su libro sobre la historia de la termodinámica².

Posteriormente, empecé a buscar referencias sobre el tema entropía de una red. Hacia 2013, una vez conversando con Rafael Hurtado (<http://www.docentes.unal.edu.co/rghurtadoh/>) sobre mi interés en la entropía, él me dijo que consultara los trabajos más recientes sobre el tema relativo a redes y entropía en la base de datos artículos del laboratorio de los Alamos (<http://xxx.lanl.gov/>). La mayor parte de referencias provenían de artículos o investigaciones realizados por físicos y algunas de éstas, estaban relacionadas con el tema de redes complejas. Y allí me encontré con los más recientes trabajos de Ginestra Bianconi (2009; 2011; 2008; 2011; 2013) sobre este tema. La dinámica de las redes de patentes ha sido estudiada principalmente desde el paradigma de la física por medio de técnicas asociadas a la mecánica estadística (Bianconi et al., 2008)(Ge et al., 2013). Estas técnicas implican la disponibilidad de grandes volúmenes de datos, que se transforman en redes muy grandes, con las cuales se hacen análisis mediante la construcción de ensamblados (Dorogovtsev, 2003). En estos ensamblados, el uso de la entropía ha demostrado su utilidad para grandes redes y permite ofrecer conclusiones con validez estadística (Bianconi et al., 2008; Anand and Bianconi, 2009; Anand et al., 2011; Zhao et al., 2011, 2013).

²Se sugiere (Jaynes, 1978) para una revisión del uso histórico del concepto de entropía

La naturaleza de mi trabajo, me llevó a concentrarme en el uso de la entropía asociado a la información. Con el fin de analizar la información contenida en la redes de citación de patentes, la entropía de información (Shannon, 1948; Jaynes, 1957, 1978; Cover, 2006) ofrece posibilidades para su descripción cuantitativa. En primer lugar, cada vez que se agrega una patente a la red de citación de patentes se produce un aumento en la entropía del sistema o como explica Kelley (1969, p.179) “*cuando un investigador adquiere datos y hechos, él está mejorando el orden dentro de su propia esfera. La entropía del experimentador, sus datos y sus registros está aumentando*”³. En segunda instancia, como la entropía puede ser entendida como sinónimo de incertidumbre, para disminuir la propia incertidumbre del investigador se ha recogido una cantidad creciente de información, “*pero, es muy usual que esto en sí mismo, puede contribuir a un aumento de la incertidumbre en lugar de tener el efecto contrario*” (Kapur, 1992, p.2). Tercero, el principio de Máxima Entropía de Jaynes (Jaynes, 1957, 1978) (Kapur, 1992) establece que se debe usar sólo la información suministrada, evitar información no suministrada y escoger la función de probabilidad que maximiza la entropía. Por estas razones, la entropía ofrece la oportunidad de analizar la incertidumbre oculta en las redes de citación de patentes a través del tiempo.

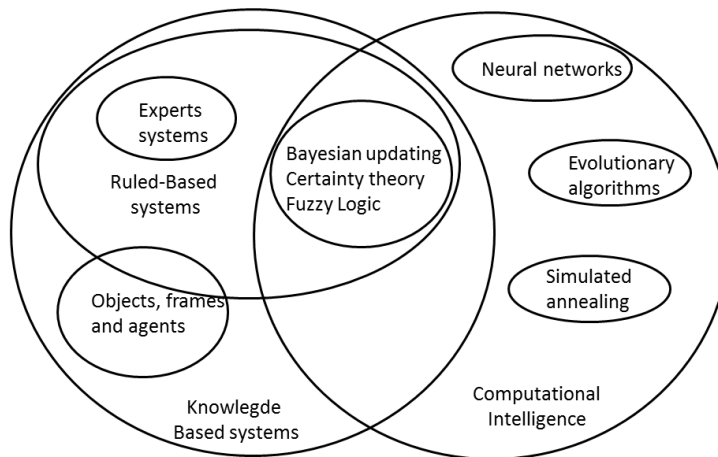
Al profundizar en los vínculos entre física e innovación, uno de los primeros autores que usa conceptos de la mecánica estadística para estudiar la innovación es Kapur (1992). En concreto, él usa el postulado del Principio de Máxima Entropía (Ver Definición 7) (Jaynes, 1957, 1978) y lo aplica en problemas de difusión de la innovación. Desde la **complejidad**, las **redes complejas** se estudian para identificar las interacciones que están por detrás de todo sistema complejo, los autores más citados son: Watts and Strogatz (1998) y Barabasi (1999). Este último, ha estudiado muchas de las redes reales existentes en la naturaleza y en la vida diaria y que responden al modelo conocido como **red libre de escala** o SF model (Scale Free Model). Él demostró que estas redes son libres de escala por dos motivos principales: i) las redes se expanden porque permanentemente están agregando nuevos nodos y ii) los nuevos vértices se agregan preferencialmente a vértices que están bien conectados. A esta última característica se le denomina actualmente como el modelo **BA model** (Barabasi, 1999) o mecanismo de **vinculación preferencia**. Posteriormente, Girvan and Newman (2002) propusieron un algoritmo para encontrar la estructura comunitaria existente en sistemas sociales y biológicos. Otra de las aproximaciones, que es el resultado de una mirada conjunta de la sociología y la física, se da desde el campo conocido como **Socio-Física** (Galam et al., 1982; Galam, 2004, 2008, 2012). Este campo de investigación ha construido puentes entre las área de la sociología y la física (Roehner, 2007). Uno de esos puentes lo constituye el tema de la redes y el paso por dicho puente ha llevado al surgimiento de la ciencia de las redes, que ha encontrado en Barabasi (2002), Strogatz (2003), Watts (2003) y Newmann (2006; 2010) a 4 de sus precursores.

En consecuencia, una de las preguntas abordada por este trabajo es ¿cómo cuantificar la información oculta en una patente para convertirla en un indicador de innovación que sea más potente que el conteo de patentes individuales? Se propone como respuesta que la entropía de información de la red que emerge de las patentes evaluada año tras año, $\{H(X_t)\} = \{-\sum_{x_t \in \mathcal{X}_t} p(x_t) \log_2 p(x_t)\}$, satisface este propósito.

2.4. Ingeniería de sistemas complejos

En cuanto a mi propia disciplina, es decir la ingeniería, Ronald and Sipper (2000) clasifican la ingeniería en 3 clases: clásica, emergente y de la vida artificial. La primera es la usualmente enseñada en las programas de pregrado de las universidades y corresponde con la aplicación de las ciencias básicas y la matemáticas por medio de las cuales las propiedades de la materia y las fuentes de energía se hacen útiles para la gente. Los resultados de aplicar este tipo de ingeniería produce resultados que no sorprenden, pues son esperados. La segunda clase incluye el uso de nuevas tecnologías de la inteligencia artificial que son comunes hoy en día o no sorprendentes pero que producen resultados sorprendentes. Mientras que la última, aborda los problemas de los sistemas complejos y produce sorpresas sorprendentes. Para explicar con mayor detalle en qué consiste la vida artificial o ingeniería de sistemas complejos se ha retomado la gráfica 2.4.1

³La entropía en este caso es interpretada como contenido informacional (Kelley, 1969, p.180)



Fuente: Hopgood, A. (2000). *Intelligent Systems for Engineers and Scientist* (2nd ed.).

Figura 2.4.1: Ingeniería de sistemas complejos.

propuesta por Hopgood (2000, p.17) quien comenta que este tipo de ingeniería ha desarrollado o empleado software denominado como inteligente, que se puede clasificar en dos grandes subconjuntos: Software de Inteligencia Computacional y Sistemas basados en conocimiento. Dentro de los primeros se destacan las redes neuronales, los algoritmos evolutivos y el recocido simulado, mientras que en los segundos se encuentran los sistemas basados en reglas, como los sistemas expertos, los objetos, marcos, los *agentes* (Peres et al., 2010, p.14) y los *autómatas celulares* (Wolfram, 1984). Estos últimos han tomado bastante fuerza para estudiar la innovación, algunos de los temas considerados son: el efecto de enfriamiento de la externalidades de red (Goldenberg et al., 2010), la dinámica de difusión en redes de pequeño mundo con consumidores heterogéneos (Delre et al., 2006), el modelamiento de la difusión de innovación con autómatas probabilísticos (Boccara and Fuks, 2008; Boccara, 1997; Zhang et al., 2005), la construcción de una función para estimar el mercado potencial de una innovación usando un autómata bidimensional (Guseo and Guidolin, 2009) y el efecto de la resistencia hacia la innovación (Moldovan, 2004). Concluyendo, esta perspectiva, ofrece un amplio espectro tanto de metodologías, como métodos para tratar situaciones que pueden ser convencionales, menos convencionales y sobre todo sorprendentes, si se trata de situaciones que corresponde a los denominados sistemas complejos.

2.5. Cienciometría y patentes

Al final del año 2014, estuve buscando una revista para publicar mis resultados relativos a la entropía y la sugerencia dada por el portal de autores de springer (www.springer.com) fue la revista *Scientometrics* (<http://link.springer.com/journal/11192>). Esto me condujo a estudiar artículos sobre mi tema en dicha revista y a reconocer otra área de investigación conocida como la **cienciometría**. Por ese motivo y gracias a los comentarios de Dominique Vinck (<http://www.unil.ch/unisciences/DominiqueVinck>), uno de mis jurados, decidí incluir esta sección sobre artículos provenientes de esta área y que estuvieran relacionados con el análisis de patentes.

2.5.1. Algunas críticas a los estudios de patentes

El estudio de las patentes presenta cuestionamientos de variada naturaleza. En primer lugar, según Cozzens et al. (2010) las patentes retrasan la invención de tecnologías específicas y por lo tanto, su estudio no es útil para identificar tecnologías emergentes. De otra parte, Newman (2010, p.71) afirma que las patentes registran, en muchas ocasiones, citas a otras patentes de manera defensiva con el fin de diferenciarse de anteriores invenciones y demostrar su propio valor para ser considerada como una nueva patente. Aún más, varios estudios, entre ellos Leydesdorff and Meyer (2010), han demostrado que el patentamiento por parte de las universidades de los países más avanzados ha tenido un fuerte declive y en consecuencia, el grado de innovación de una nación

puede estar subestimado. Finalmente, Harris et al. (2007) plantean que la producción de patentes como indicador de innovación corresponde a una mirada lineal pero internacionalmente aceptada y también muy criticada sobre la innovación como sinónimo de la invención científica y tecnológica y a la cual hay que encontrarle alternativas de medición. Al respecto, Chesbrough et al. (2006) afirma que las medidas tradicionales de la innovación, como el porcentaje de inversión en Investigación&Desarrollo, o el porcentaje de ventas resultado de la inversión anterior, sólo capturan el gasto interno de la empresa y dejan totalmente de lado la inversión realizada por otras partes como proveedores, consumidores o terceras partes. Otra medida tradicionalmente usada como las patentes generadas por la I&D interna son obsoletas dado que también recogen sólo el esfuerzo interno. Al contrario, sería más útil, valorar o hacer seguimiento a las patentes internas y externas incorporadas en el servicio o producto propio o de otras empresas.

2.5.2. Redes de citación entre patentes y análisis de palabras

Uno de los enfoques que ha primado en el análisis cuantitativo está relacionado con comprender las patentes mediante el análisis de las palabras relacionadas con la patente. Algunos de las líneas de investigación originadas por este enfoque se exponen a continuación.

2.5.2.1. Identificación de patentes similares

El artículo escrito por Wu et al. (2010) propone un método para encontrar la similitud entre un par de patentes al buscar primero entre las redes de citas directas e indirectas de las patentes y las palabras claves al hacer minería de texto en los resúmenes del texto aprobado de la patente. Al respecto, nunca pensé en la realización de análisis de las palabras claves porque quise focalizarme solamente en redes de patentes que fuesen de modo 1. Es decir sólo un tipo de nodos, las patentes citando a otras patentes. Tuve la oportunidad de haber avanzado en el estudio de redes que incluyeran diversidad de nodos, autores, empresa cesionaria, países, códigos de clasificación, fechas, etc, puesto que estuve tomando un curso en Carnegie Mellon con el grupo de investigación CASOS, que diseñó el software ORA (Carley et al., 2009) el cual permite la realización de este tipo de redes con una gran variedad de nodos. El concepto que permite esta articulación de nodos de diferente naturaleza se conoce como **metared**. En este caso la metared de una patente luciría como se muestra en la figura 2.5.1.

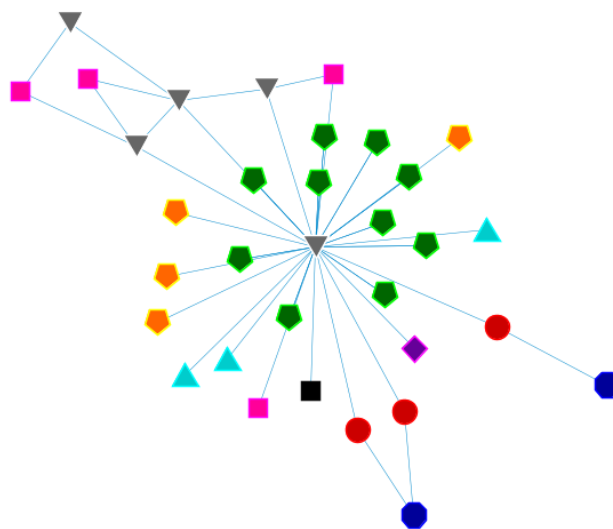


Figura 2.5.1: Metared de una patente

Hexágonos azules=Países inventores, Círculos rojos=Inventores, Rombo Violeta =Fecha publicación, Cuadrado=Cesionario, Pentágonos verdes=Citaciones recibidas, Triángulos grises invertidos=Citaciones hechas, Pentágonos naranjas=Palabras claves, Cuadrados magenta=Códigos clasificación, Triángulos azules claro=Otras publicaciones

El avance hubiese podido ser desde un enfoque socio-técnico en donde todos estos nodos son considerados de igual importancia y a las redes construidas se les puede aplicar el mismo tipo de indicadores que se usan en el análisis de redes sociales. Sin embargo, siempre me pareció difícil explicarme a mi mismo la validez de “mezclar peras con manzanas” y contarlas como si fuesen lo mismo. Uno de los artículos que me leí sobre lo importante de trabajar en la dirección de las palabras incluidas sobre todo en el título fue escrito por Cozzens et al. (2010) y aunque me pareció interesante esa dirección, decidí escoger otro camino.

2.5.2.2. Redes de palabras con títulos de las patentes

Uno de los artículos centrales en esta dirección fue escrito por *Courtial et al. (1993)* inspirado en los planteamientos de Callon, Law & Rip (1985). Este interesante artículo, además de realizar una excelente revisión de la literatura relativa a temas de complejidad (fractalidad, emergencia, auto-organización, ley de Zipf, etc...), la cual le debe encantar a mi director, se centra en estudiar los títulos de las patentes y a partir de éstos crea redes temáticas y las asocia a áreas tecnológicas. Percibo detrás de este trabajo un acompañamiento con expertos de las áreas tecnológicas con el fin de darle sentido a los resultados. Me recuerda metodológicamente lo explicado por Charum (2009) sobre la secuencia dato, información y conocimiento. Al respecto y con relación a mi alcance, éste está limitado a la parte de producción de información y no tanto de conocimiento. La razón principal radicó en que yo seleccioné patentes de múltiples áreas tecnológicas y en consecuencia, la disponibilidad de expertos para todas las áreas tecnológicas se convertiría en una limitante de mi enfoque, si hubiese escogido la alternativa de estudiar los títulos de las patentes.

2.5.2.3. Relación entre patentes y referencias en publicaciones científicas

La relación entre patentes y referencias en publicaciones científicas fue analizada entre otros por *Callaert et al. (2006)*. En un principio, esta dirección no me interesa porque no me parece fácil la interpretación de indicadores o métricas en donde se mezclan artículos con patentes, los cuales son 2 objetos diferentes.

Por otra parte, este artículo escrito por *Shin et al. (2006)* sí me parece pertinente y cercano a mi problema de investigación. De una u otra forma, parte de mi interés central radica en la comparación entre países que son similares a nivel de competitividad, pero que es muy probable sean diferentes en el tipo de conocimiento que producen en forma de patentes y que por lo tanto, la inversión extranjera que se realice en ellos esté mediada por la información que provee el conocimiento que producen. Durante mi examen de candidatura mi exposición partió de proponerle a mis jurados la siguiente pregunta: *Ud. es un inversionista (tomador de riesgo) que quiere invertir su dinero en algunos de los nuevos países emergentes. Para tomar su decisión, además de los análisis tradicionales, ud va a basar su decisión en el criterio de la producción de conocimiento en forma de patentes. Este trabajo busca arrojar luces para tomar una decisión.* Es decir, en mi proyecto de investigación existe una inquietud por comparar países desde una perspectiva competitiva, agregándole un elemento adicional para la decisión, la producción de conocimiento en forma de patentes.

El estudio desarrollado por *Atallah and Rodriguez (2006)* es similar a mi trabajo doctoral en la metodología. Ambos tomamos como objeto de análisis la red de citas que emerge de la patente bajo estudio. En dicha red se consideran tanto las citas directas recibidas por la patente como las que reciben las patentes citantes. De esta forma, se conforma una red que incluye tanto citas directas como indirectas. Dicho trabajo es también similar en cuanto al objetivo de aprovechar la información de todo tipo de citas. En sus propias palabras: *“Using all the direct and indirect citations of a patent makes maximum use of the available information”*. Ambos trabajos son similares también en que construyen un indicador de calidad de las citas basadas en conteos de citas directas e indirectas. Sin embargo, son diferentes en qué se cuenta. El trabajo de Atallah and Rodriguez (2006) cuenta la cantidad de citas directas e indirectas y construye el indicador de calidad. Por mi parte, lo primero que hago es transformar los datos de citas directas e indirectas en redes. Lo segundo, es resumir estadísticamente la red en su función de distribución de grado. Y finalmente, a esta estadística calcularle la entropía de información durante el tiempo y ese es el indicador que se propuso en mi trabajo.

El trabajo realizado por *Guan and He (2007)* tomó como objeto de estudio las referencias en journals a las patentes de interés, así como las clasificaciones tecnológicas de las patentes. El propósito de la investigación fue distinguir los conceptos de ciencia y tecnología desde una perspectiva de la construcción de indicadores. De esta forma, las patentes fueron asociadas al concepto de tecnología, mientras que las citas en publicaciones

fueron usadas para determinar la relación entre ciencia y tecnología. En mi caso particular, por ahora no me interesa entrar en la discusión sobre qué es ciencia o qué es tecnología, por lo tanto, los trabajos centrados en dicha distinción no han sido significativos para mí.

Por último, la investigación de *Gao and Guan (2009)* intenta mapear la relación entre patentes y ciencia mediante la identificación de las publicaciones científicas que son citadas en las patentes.

2.5.2.4. Redes neuronales, redes de publicaciones y patentes

Mi idea inicial arrancó tratando de medir el retorno de la inversión en innovación. Esto se debió a la lectura del libro *Payback (Andrew and Sirkin, 2006)* que me recomendó uno de los mentores en mi interés investigativo llamado Alberto García. Mas rápidamente, llegué a la conclusión que la información para medir el retorno implicaría tocar las puertas de las empresas y ésto en un país como Colombia no sería muy fructífero. Por éso todo lo referente a valoración de la patente quedó excluido de mis intereses.

Sin embargo, aproximaciones a la medición del valor en el mercado de una patente se han realizado desde distintas perspectivas. Aquí señalo la tomada por *Chen and Chang (2009)* y *Polanco et al. (2001)*. Ellos trabajaron con redes neuronales que son capaces de aprender y esta propiedad las hace muy útiles. Por el contrario, yo me interesé por las redes de citación entre patentes, las cuales son redes dirigidas y en dónde no existen ciclos. Por esta última característica estas redes no pueden aprender, puesto que no hay retroalimentación o feed back. Debido a esta razón nunca contemplé trabajar con redes neuronales en mi investigación.

2.5.2.5. Transacción entre patentamiento y producción de artículos

La investigación desarrollada por *Meyer (2006)* explora la transacción entre publicaciones científicas y patentamiento por parte de investigadores universitarios. Por lo tanto, metodológicamente se concentra en los nombres de los investigadores que producen patentes o artículos. Así mismo, la investigación escogió 3 países de la UE (Alemania, Gran Bretaña y Bélgica) para la realización del estudio de caso. Con respecto al estudio, en alguna parte de mi trabajo señalé que *Leydesdorff and Meyer (2010)* explicaron que la transacción entre patentar y publicar artículos está asociado a la culminación del incentivo que ofrecía, por lo menos en los Estados Unidos, la ley Bayh-Dole. Así mismo, señalaron que parte de la culpa se debe también a la importancia de los rankings académicos, en donde el producir patentes y artículos son valorados de manera muy similar, en perjuicio del patentamiento que es más costoso tanto en tiempo como en uso de recursos.

2.5.2.6. Uso de la entropía en cienciometría

El objeto de estudio de esta investigación de *Grupp (1990)* es el involucramiento institucional en las decisiones sectoriales de inversión en investigación y el desarrollo. El planteamiento es el siguiente: sea I_i para $i = 1, 2, \dots, n$ el conjunto de instituciones de un país y sea F_j para $j = 1, 2, \dots, m$ el conjunto de campos de I&D. Así mismo, se define p_{ij} como el porcentaje de publicaciones o patentes de la empresa I_i en el campo F_j . De esta forma, se aplica el concepto de entropía en primer lugar para la inversión institucional en un campo F_j como $S_j = -\sum_i p_{ij} \ln p_{ij}$ con $\sum_i p_{ij} = 1$ y la entropía institucional $S_i = -\sum_j p_{ij} \ln p_{ij}$ con $\sum_j p_{ij} = 1$. Con este planteamiento, se puede evaluar la entropía tanto institucional como por campos de I&D de un país y así realizar comparaciones entre países. Este estudio es por lo tanto, un ejemplo de evaluación a nivel macro y además comparativo. Adicionalmente, la idea usada de entropía tiene que ver con su interpretación como sinónimo de incertidumbre contenida en este caso en los valores p_{ij} .

De otra parte, esta investigación fue de las primeras que consulté referente al tema de patentes y entropía. Sin embargo, deseché replicar el camino tomado por dicho estudio porque la experiencia de trabajar con los datos del DANE (*Hurtado and Mejia (2011, 2014)*) me decepcionó por varios motivos. El primero tiene que ver con la desactualización de los datos, pues las encuestas EDIT III, IV y V se ponen disponibles al público muy tarde y contienen datos con 5 años de retraso en promedio. Por otra parte, para tener acceso a los datos se requiere que la institución tenga convenios con el DANE y cuando se encuentra quien administra el convenio, él se cree dueño de la información y la acapara y comparte con mucha dificultad. Por estas 2 razones tomé la decisión de acceder a fuentes de información más actualizadas y que no necesitaran convenios ni permisos para consultar. Siendo así, encontré que la USPTO tiene a disposición de todo el mundo su base de datos de patentes con información actualizada al día y disponible por internet sin mayores inconvenientes. Esto me llevó a interesarme

en la información de patentes a nivel micro y posteriormente, a la medición de la entropía de información de la red generada por la patente.

2.6. Conclusión

Se ha presentado el estado del arte de la presente investigación doctoral. Del mismo se concluye que la investigación tiene un carácter multidisciplinar. Este carácter multidisciplinar se nutre de conceptos fundamentales originados en las disciplinas de la economía, sociología, física y las matemáticas, los cuales son complementados con desarrollos conceptuales producidos en áreas de conocimiento como son el análisis de redes sociales, la ciencia de las redes o la ingeniería de sistemas complejos. Como conclusión esencial, se tiene que para el estudio de la innovación, las patentes representan una fuente de información muy interesante, que ha sido muy usada en países muy desarrollados, pero no así en países que se encuentran en el tránsito entre ser movidos por la eficiencia de los factores hacia movidos por el conocimiento, como es el caso de Colombia. Igualmente, se ha encontrado que los estudios sobre redes de citación de patentes han sido realizados en esencia por científicos provenientes de la física y que han mostrado su interés en descubrir leyes que gobiernan el comportamiento de las mismas. Finalmente, se ha concluido que el estudio de las redes de citación de patentes para descubrir el comportamiento de la innovación en países similares a Colombia representa una interesante posibilidad de generación de conocimiento pertinente.

Capítulo 3

Objetivos e hipótesis

3.1. Pregunta de investigación

¿Cómo se llegó a la pregunta de investigación? Para hacerlo se usó la metodología resumida en la figura 3.1.1 y que ha sido tomada de SeminairePolibius (2007). Este método posibilita la definición de la pregunta de investigación partiendo de las áreas de interés y después de pasar por el tópic. Enseguida, transforma la pregunta en una proposición, este es un proceso de traducción de la pregunta de investigación en constructos, de forma que la proposición explica a un nivel abstracto cuál es la relación entre los constructos. A su vez los constructos son una construcción teórica para resolver un problema científico determinado. Posteriormente, mediante el planteamiento de hipótesis de investigación, se operacionaliza la proposición en términos de un fenómeno medible.

Para el caso de la presente investigación, el problema es multidisciplinario y las áreas desde las cuales se plantea la pregunta de investigación son: la economía, la socio-física y la ingeniería. En primer lugar, existe un interés desde la economía por el tema del patentamiento debido a su directa asociación con lo que hoy se denomina la economía basada en el conocimiento. Una patente representa conocimiento empaquetado y además, es una muy buena aproximación de una innovación. En segundo lugar, se menciona el área de la socio-física, puesto que se está interesado en el tema de la difusión de innovación y ésta se da en un medio social, de una persona a otra y muchos de los modelos más ampliamente aceptados han sido desarrollados desde la física. Dentro de esta área, se profundiza mucho más y se realiza una concentración en la subárea del Análisis de Redes Sociales. Por último, se trata de una mirada desde la ingeniería, con el fin de dar aplicación a conceptos robustos de las anteriores áreas en un objeto de estudio concreto.

Ahora bien, el tópic de estudio es la dinámica de difusión de innovación. Existe interés genérico por el tema de la innovación, mas no de una forma estática, todo lo contrario, interesa comprender la difusión de la innovación en su carácter dinámico. Es decir, existe una motivación científica por comprender la forma cómo ha cambiado en el tiempo este fenómeno.

Por otro lado, para poder entender el fenómeno se ha tomado la decisión de estudiarlo desde 2 posturas teóricas que se espera puedan complementarse. La primera es la naciente ciencia de la redes y la segunda es la minería de datos. La ciencia de las redes permitirá investigar sobre las redes de citación de patentes, mientras que la minería de datos, buscará ser un complemento, para encontrar patrones en las redes. De esta manera, la pregunta de investigación es:

¿Cómo es el comportamiento de la dinámica de difusión de innovación de los nuevos países emergentes a partir de la red de citación de sus patentes?

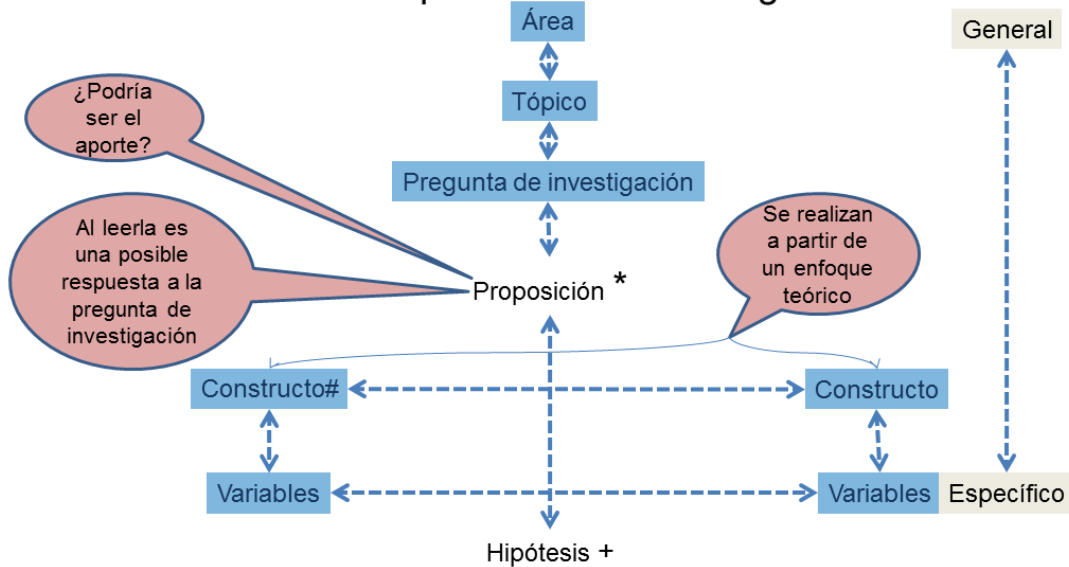
Siguiendo con la metodología propuesta, el siguiente paso es transformar la pregunta en una proposición. Para lograrlo, se plantean los siguientes 2 constructos.

Primer constructo: **Un modelo para análisis de redes de citación de patentes integrado por conceptos del ARS & Teoría de la Información**

Segundo constructo: **Permita descubrir patrones de la dinámica de la difusión de la innovación**

Los cuales combinados conducen a la proposición: **Un modelo para análisis de redes de citación de patentes integrado por conceptos del ARS & Teoría de la Información permite descubrir patrones de la dinámica de difusión de la innovación.**

Definiendo el problema de investigación



1 Tomado de: Defining a research question, Seminaire Polibius, Katholieke Universiteit LEUVEN, 2007

* Proceso de traducción de la pregunta de investigación en constructos. Explica a un nivel abstracto cual es la relación entre los constructos.

+ Operacionalización de la proposición en términos de un fenómeno medible

Construcción teórica para **resolver** un problema científico determinado.

Definiendo el problema de investigación

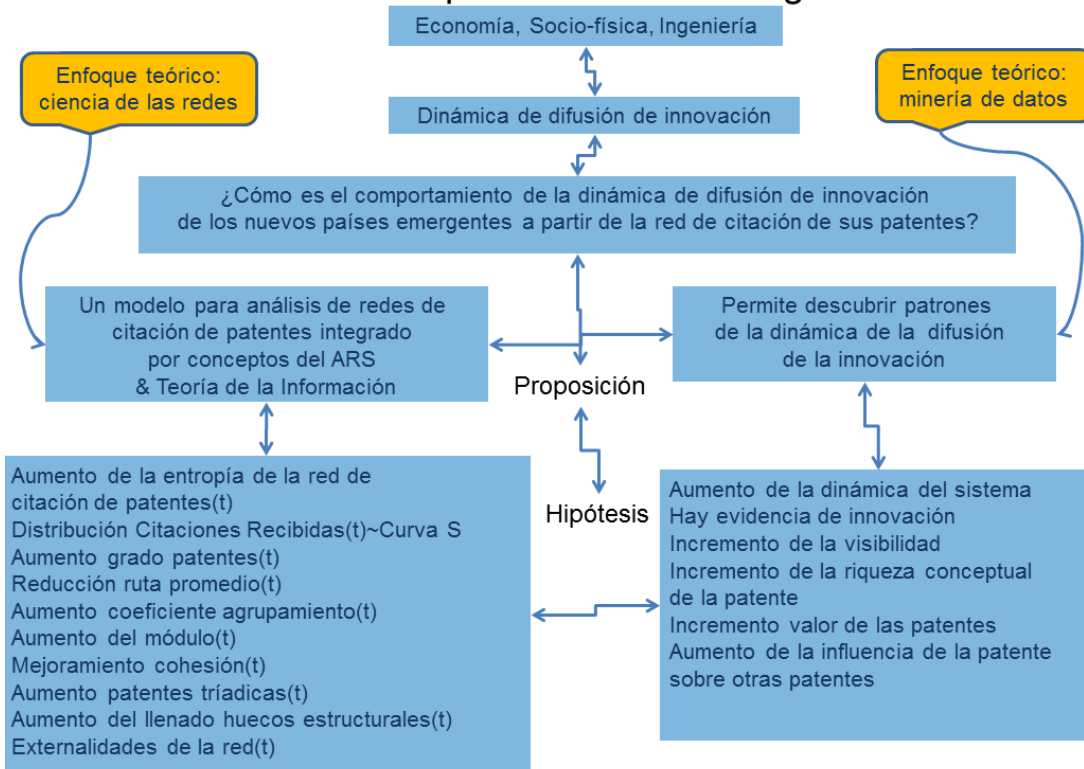


Figura 3.1.1: Pregunta de Investigación

3.2. Objetivos

3.2.1. Objetivo General

Describir la dinámica de difusión de innovaciones de los países del grupo CIVETS mediante la caracterización de las redes de citación de sus patentes.

3.2.2. Objetivos Específicos

- Proponer un modelo de análisis de la dinámica de difusión de patentes integrado por conceptos del ARS y Teoría de la Información
- Determinar patrones de las estructuras de redes de citación de patentes de países CIVETS mediante la aplicación del modelo propuesto
- Comparar la dinámica de las redes de citación de patentes de los países CIVETS a partir de la implementación del modelo propuesto

3.2.3. Hipótesis

A continuación, el aspecto a abordar es la caracterización estadística de las redes de patentes y de su función de distribución. Para hacerlo, se recurrirá a probar algunas hipótesis que se focalizan en aspectos de carácter dinámico, estático y con mirada macro o micro de las redes de patentes. El primer criterio de clasificación es la dinámica, definida como la evaluación del cambio temporal en la estructura de la red. En contraste, la estática de la red se considera como su situación en un momento particular. De manera complementaria para los dos criterios anteriores, se puede tener una mirada micro y otra macro. La mirada micro corresponde a la consideración de la conformación de nodos y vínculos entre éstos. Mientras que la mirada macro busca una mirada general sobre la red. Por lo tanto, las hipótesis han sido clasificadas bajo tres tipologías: Dinámicas-Macro, Dinámicas-Micro y Estáticas Micro.

3.2.3.1. Hipótesis Dinámicas y Macro

El primer grupo los constituyen las hipótesis que tienen carácter dinámico y simultáneamente apuntan a un fenómeno macroscópico.

$$H_H : H(X(t_n)) \geq H(X(t_{n-1})) \cdots \forall t \quad (3.2.1)$$

Que se leería como: La entropía de la red de citaciones de patentes de países CIVETS aumenta con el paso del tiempo. Donde

$$H(X) \equiv \text{Entropía de la red} = - \sum_{x \in \mathcal{X}} p(x) \log_2 p(x) \quad (3.2.2)$$

Derivado de lo anterior, si $H(X)$ es la entropía de la red completa de citación de patentes y $H(Y)$ es la entropía de una red construida a partir de la anterior mediante procesos de agrupamiento, es altamente probable que desde un punto de vista de contenido informacional se pueda presentar pérdida de información, es decir

$$\text{Information Loss} \equiv IL \equiv H(X) - H(Y) \quad (3.2.3)$$

De manera que, se necesita minimizar IL para un grado dado de compresión de la información. Esto se logra escogiendo aquel Y que asegura un grado determinado de simplificación de tal forma que $H(Y)$ sea máxima (Kaplan, 1992, p.209), lo anterior es conocido como, el problema de reconocimiento de patrones. En conclusión, la entropía es un concepto robusto que será usado en la investigación sobre citación de patentes.

3.2.3.2. Hipótesis Dinámicas y Micro

¿Cómo identificar las etapas de la innovación a partir de una red de citación de patentes? o en otras palabras cómo se puede identificar el proceso de difusión a partir de la red de citación de patentes? La segunda opción para medir la dinámica del proceso de difusión se ha construido con la información a nivel micro que es factible recoger para cada una de las patentes de interés. Dado que se puede disponer de las fechas exactas de las citaciones recibidas por una patente, con estos datos se pueden construir histogramas de frecuencia acumuladas de estas citaciones y así la hipótesis puede quedar planteada de la siguiente forma:

$$H_S : G(n_i) \text{ función acumulada probabilidad citaciones recibidas} \sim \text{Curva "S"} = \frac{a}{1 + e^{-(b+ct)}} + d \quad (3.2.4)$$

la cual se leería como: Existe evidencia de innovación en la estructura de la red de citación de patentes colombianas.

Wang et al. (2006) presentan un modelo basado en ecuaciones diferenciales, (una modificación del modelo de Bass (1969, 2004)), el cual ha sido bastante empleado (Mahajan and Kerin, 1984) para entender la manera en que se propaga un producto innovador y que ha capturado una amplia variedad de patrones de difusión observados en la práctica. Este modelo describe el proceso de innovación simplificando las 5 etapas mencionadas por Rogers (1962): Apropriación (Individuo expuesto a la innovación), interés (individuo busca más información), Evaluación (el individuo aplica la innovación a una serie de situaciones), pruebas (el individuo emplea la innovación en pequeña escala), Adopción (el individuo hace pleno uso de la innovación), a 3 etapas: apropiación, evaluación y toma de decisiones. Este modelo se caracteriza por presentar bajo ciertas condiciones, retardos de tiempo y un umbral a partir del cual la innovación ocurre. Se propone mediante esta hipótesis clasificar las patentes que citan a la supuesta patente innovadora, basado en las fechas de aprobación y realizar una evaluación estadística sobre su ajuste a la curva "S" de innovación.

$$H_{CI} : \text{La Centralidad de intermediación es una variable apropiada para identificar patentes innovadoras} \quad (3.2.5)$$

Cuando una patente es aprobada, en su texto se registran las patentes que han referenciado tanto el autor del invento como el evaluador de la patente y éstas pueden ser consideradas como conocimiento empaquetado que se requiere para "construir" o elaborar el artefacto o proceso que es objeto de patentamiento. De esta forma, se configura una primera parte de la red de citaciones, es decir la red de patentes citadas, $x_i(t)$ (hacia la izquierda en el gráfico). En lenguaje de redes la dirección de las flechas corresponde al concepto de outdegree o grado saliente del nodo. Complementariamente, una vez la patente ha sido aprobada, y sí ésta despierta interés o más exactamente si resulta ser una patente que conduce a productos o procesos innovadores, empieza a ser citada por otras patentes de manera masiva y se configura la otra porción de la red de citaciones, $x_i(0)$ (hacia la derecha en el gráfico) en concreto, la red de patentes que la han citado en un momento concreto. En términos de redes, corresponde al concepto de indegree o grado entrante del nodo. Todo esto resultaría en una configuración similar a la siguiente red, donde en el centro se ubica la patente "innovadora", a su izquierda se ubican las patentes citadas y a su derecha, las patentes que la citan. Lo anterior configura lo que se denomina un nodo con alto grado de **centralidad de intermediación** (Wasserman and Faust, 1994, p.189-191) y así podría identificarse una patente innovadora

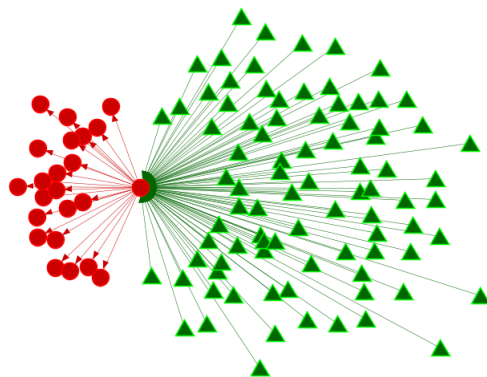


Figura 3.2.1: Modelo de red de citas de una patente innovadora
Círculos son patentes citadas en el texto aprobado. Triángulos son patentes que referencian a la patente bajo estudio

3.2.3.3. Hipótesis Estáticas-Micro

$H_{pc}: P(\text{Patente CIVETS} = \text{Innovación}) = f(\text{Entropía Conocimiento, Internacionalización Referencias, Indegree})$

Que se lee: La probabilidad que una patente de un país CIVETS se convierta en innovación es una función de su contenido de conocimiento, el grado de internacionalización de sus referencias y las citas generadas. Esta hipótesis puede conducir a la construcción de una métrica que combine los estadísticos clásicos de redes como el grado nodal, la centralidad de intermediación y la cohesión, con medidas como la entropía, evaluadas en su dinámica temporal.

H_{CAT} : Una patente catalizadora puede conducir a patentes innovadoras (3.2.6)

Es muy probable que una patente resuelva el hueco estructural (Burt, 1992) existente en un área tecnológica y este hecho puede posibilitar que más adelante, otra patente use este conocimiento y sea esta última la que reciba un alto porcentaje de citas. Esta idea está inspirada en Sanjay and Sandeep (2006, p.561) quienes propusieron que el cambio en la estructura de una red compleja al aparecer un nuevo tipo de nodo puede ser un indicador de innovación en la red. Su idea usa la metáfora de las sustancias catalizadoras en las reacciones químicas y este sería el papel jugado por una patente catalizadora.

H_{EI} : La entropía de información (Ver definición 6) contribuye a la identificación de las patentes con mayor contenido de información.

Pudiera suponerse que una medida del valor de una patente está asociada con su contenido de información. Este contenido está registrado en las patentes y las citas recibidas. De esta forma, las patentes pueden ser clasificadas según su contenido de información y una patente será más valiosa que otra si su contenido es mayor. A manera de ejemplo (Gráfica 3.2.2), la entropía de información de la red 1 es menor que la entropía de información de la red 2 y por lo tanto, la patente 2 es más interesante que la 1 por tener un contenido de información mayor.

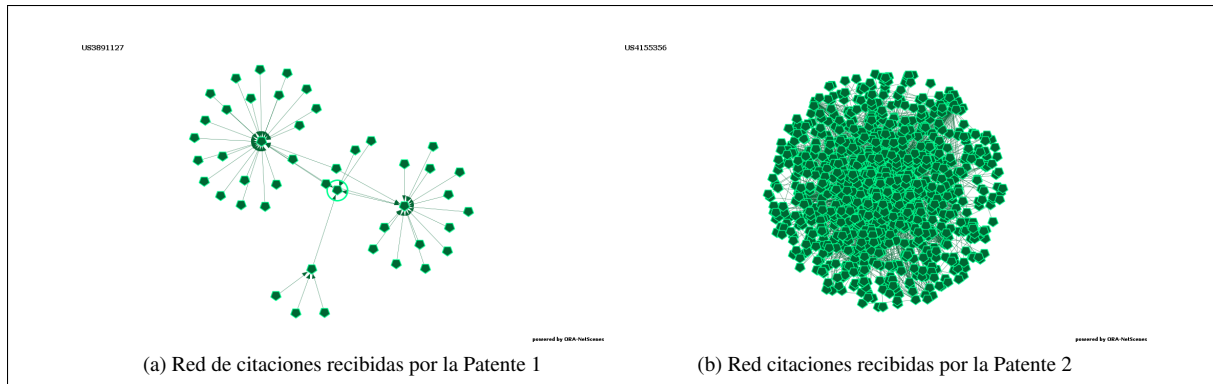


Figura 3.2.2: Contenido conocimiento de patentes
 Nodos representan patentes citantes de la patente bajo estudio

H_{PS} : Parábola de la semilla (3.2.7)

Esta hipótesis se basa en la importancia del papel del evaluador de patentes. Él es quien confirma las palabras claves sugeridas por el inventor y asigna unas nuevas palabras claves asociadas a la patente y en consecuencia, su trabajo es el que posibilita encontrar la patente por medio de las claves de búsqueda. Un invento puede ser muy bueno en sí mismo, al igual que una semilla mas su germinación está condicionada a la calidad del terreno en el cual es cultivada. En esta situación, las palabras son estratégicas para la clasificación y posterior búsqueda de la patente en las bases de datos de patentes.

Capítulo 4

Metodología

4.1. Tesis

Para comprender el presente a veces es necesario regresarse al pasado. Metodológicamente esta investigación se posiciona inicialmente al final del año 2013, encuentra unos resultados que sorprenden y para explicarlos da un salto hasta diferentes pasados y se devuelve dinámicamente hasta el presente. Al seguir el camino trazado por la Fig.4.1.1, la investigación integra el análisis estático y longitudinal de redes. El primero de éstos corresponde al estudio de las redes de citación al final del año T que en el caso de esta investigación es 2013. Por otra parte, el segundo análisis observa los cambios en las redes desde el año de publicación hasta el año T . De esta forma, se explica en este apartado cómo se llegó a la tesis de la investigación: En resumen, este trabajo parte del análisis de los datos, para producir información y conocimiento, bajo un esquema conocido como análisis guiado por los datos (*Data Driven Analysis* en inglés). Este sistema de indagación es aplicado a un conjunto de patentes y se concentra metodológicamente en el estudio de las redes aplicando los modelos de las leyes de potencia y de la entropía de información. El estudio empieza explorando la presencia de leyes de potencia en las redes y posteriormente, para comprender a qué fenómeno está asociado este comportamiento se analiza cada red desde la fecha de publicación de la patente hasta finales del año 2013 y a cada una de estas redes se le estima la entropía de información.

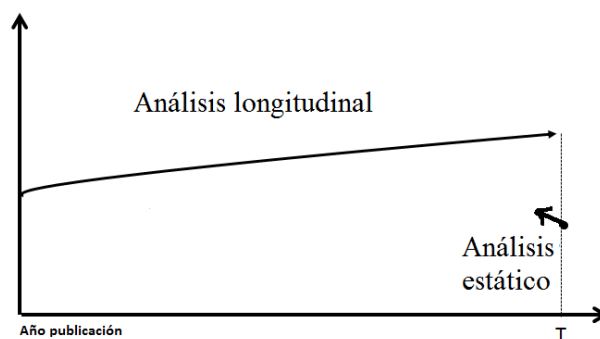


Figura 4.1.1: Combinación de los análisis estático y dinámico en la metodología.

Por lo tanto, la tesis doctoral se enuncia de la siguiente manera: **Un modelo para análisis de redes de citación de patentes integrado por los conceptos de leyes de potencia y entropía de la información permite descubrir patrones de la dinámica de difusión de la innovación.**

El procedimiento detallado se describe en el siguiente diagrama de bloques.

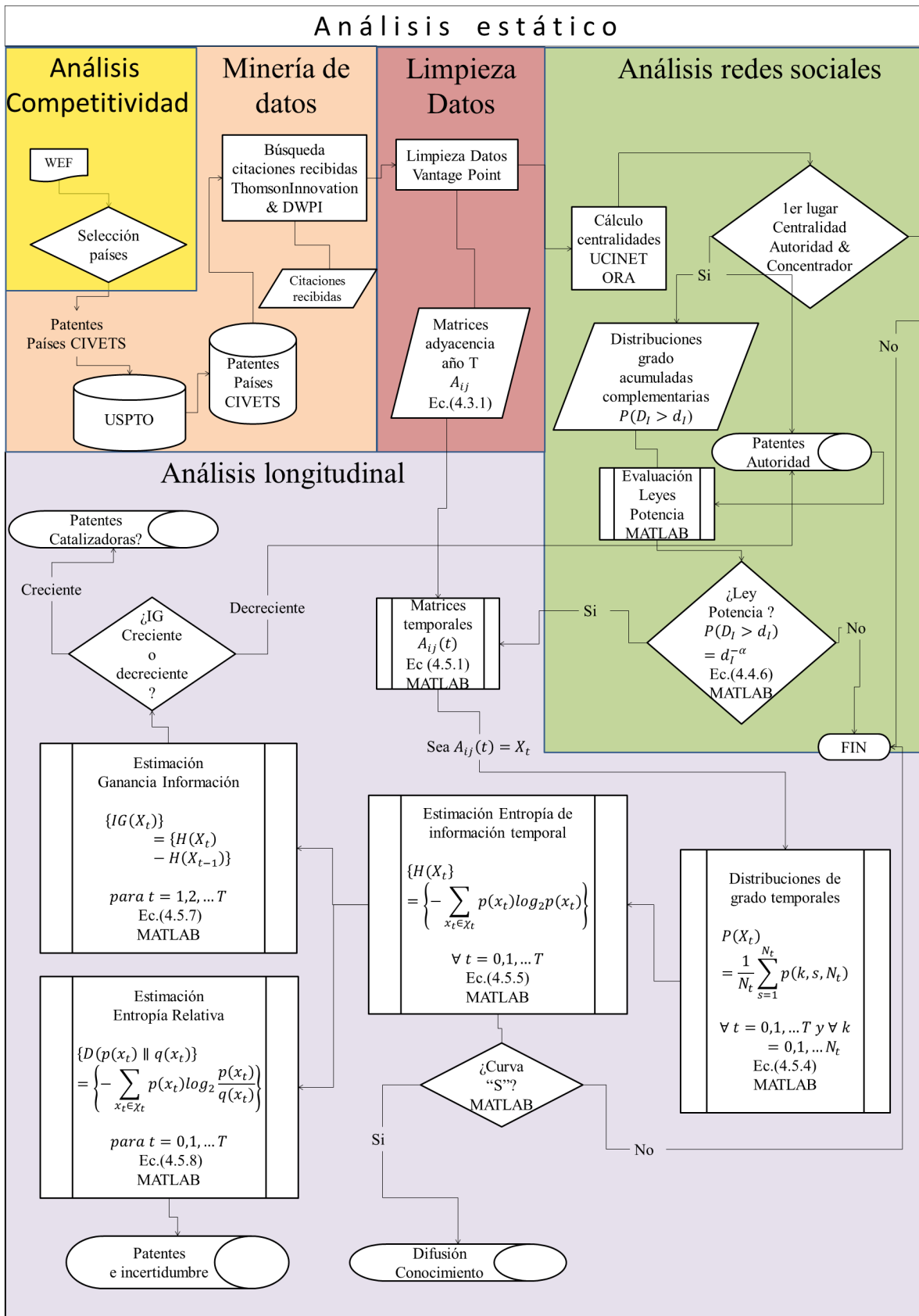


Figura 4.1.2: Diagrama de bloques de la metodología

Los números de las ecuaciones corresponden a los presentados en el capítulo 4. MATLAB, VantagePoint, ThomsonInnovation, UCINET y ORA son los programas con los cuales se ejecutaron los pasos descritos

4.2. Tratamiento de datos

Mi trabajo doctoral necesitó de un gran esfuerzo en el tratamiento de datos para convertirlos en información y en conocimiento. Una vez que encontré “mis joyas de la corona” (Silva, 1908; Molinares, 1912; Mayor Mora et al., 2005) tomé la decisión de profundizar en el estudio de patentes desde países emergentes para intentar identificar el posible impacto global de las mismas. Para lograrlo, primero extraje de la oficina de patentamiento de los Estados Unidos todas las patentes que incluyeran al menos un inventor de nacionalidad de los países CIVETS o que la empresa cesionaria fuese también de alguno de estos países. Con estos datos disponibles procedí a buscar todas las citas recibidas por las mismas. Para ello, usé la aplicación ThomsonInnovation (<http://info.thomsoninnovation.com/>) que por medio de su índice Derwent (Reuters, 2012) busca las citas en las mejores bases de datos de patentes internacionales. Sin embargo, para poder usar el programa tuve que lograr que la empresa ThomsonReuters instalara el programa en la Universidad Central, ya que no estaba disponible en la Javeriana. Este programa tiene un costo de US\$50,000 y me lo instalaron para evaluar y usarlo durante 2 meses. Durante ese corto período extraje las citas que necesitaba. Posteriormente, los archivos de las citas fueron sometidos al proceso de limpieza, que consiste en unificar nombres y números de las patentes. El programa que empleé se llama VantagePoint (<https://www.thevantagepoint.com/>) y sólo estaba disponible en la biblioteca de la Universidad Javeriana. Después de tener los datos limpios, pude proceder a armar las redes de citación de patentes. En esta parte usé distintas aplicaciones, empezando por UCINET (Borgatti et al., 2002) pasando por Pajek (Mrvar and Batagelj, 2013), ORA (Carley et al., 2009) hasta Visone (KonstanzUniversity, 2001) principalmente. Cada una de ellas tiene rutinas comunes pero algunas que necesitaba no estaban en las otras, de manera que aprendí a “cacharrear” con cada una de éstas. Finalmente, me tuve que enfrentar al tema de realizar mis propios programas para poder “tener el control” sobre las métricas que me interesaban. Ésto me obligó a recuperar de mi memoria las bases de programación que obtuve en el pregrado, pues hace 28 años había elaborado el último programa que me corrió. Las rutinas que diseñé fueron realizadas usando el programa Matlab (Mathworks, 2014). Debo reconocer que el haber reaprendido a programar es otro de los logros personales y que significó un gran esfuerzo en tiempo y sudor para mí.

4.3. Análisis estático de redes

El análisis estático de redes implicó la selección de un grupo de patentes que fueran autoridad. Para ello, las principales actividades metodológicas fueron: análisis de la competitividad y de la innovación de un conjunto de países, minería de datos, limpieza de datos, selección patentes autoridad mediante análisis de redes sociales y por último, la estimación de leyes de potencia para la distribución de grado de las redes generadas por estas patentes.

La primera de las etapas de este estudio fue el análisis de la competitividad para una selección de nuevos países emergentes. La selección estuvo basada en tres etapas secuenciales. Inicialmente, la Fig. 4.3.1 fue elaborada mostrando la correlación existente entre competitividad y producción de patentes. De esta imagen concluimos que hay dos grupos evidentes de países. Aquellos con alta producción de patentes y alta competitividad (Cluster 1) y los demás (Cluster 2) que incluye una gran variedad de otras naciones. Como resultado de esta etapa, concluimos que la selección debía estar complementada por la búsqueda de un conjunto de países con índices de competitividad similares, dado el interés por hacer visibles a patentes prestigiosas originarias de países pertenecientes al cluster 2.

Colombia necesita la rápida entrada en vigencia del tratado de libre comercio con los Estados Unidos, por su parte Turquía, persiste en su larga lucha por lograr hacer parte de la Unión Europea mientras que, Vietnam e Indonesia, buscan ponerse a la altura de los antiguos y nuevos dragones asiáticos, además de lograr el nivel exigido para hacer parte del vigoroso crecimiento de la China. Además, estas naciones han sido agrupadas porque sus patrones de competitividad presentan un comportamiento muy similar. En general, la competitividad se mide por medio de los 12 pilares mostrados en la Fig.4.3.2 y en todos ellos los países CIVETS son similares, luego sus puntajes de competitividad son casi los mismos. Uno de los pilares es la innovación. Ésta es medida en función de: La capacidad de innovación, la calidad de las instituciones de investigación científica, el gastos empresarial en investigación y desarrollo, la colaboración entre las universidades y la industria en I&D, las compras del gobierno de productos avanzados tecnológicamente, la disponibilidad de científicos e ingenieros y finalmente, las solicitudes por millón de habitantes de patentes(WEF, 2010, 2013) por medio de tratado PCT¹. Este trabajo está centrado en la producción de patentes.

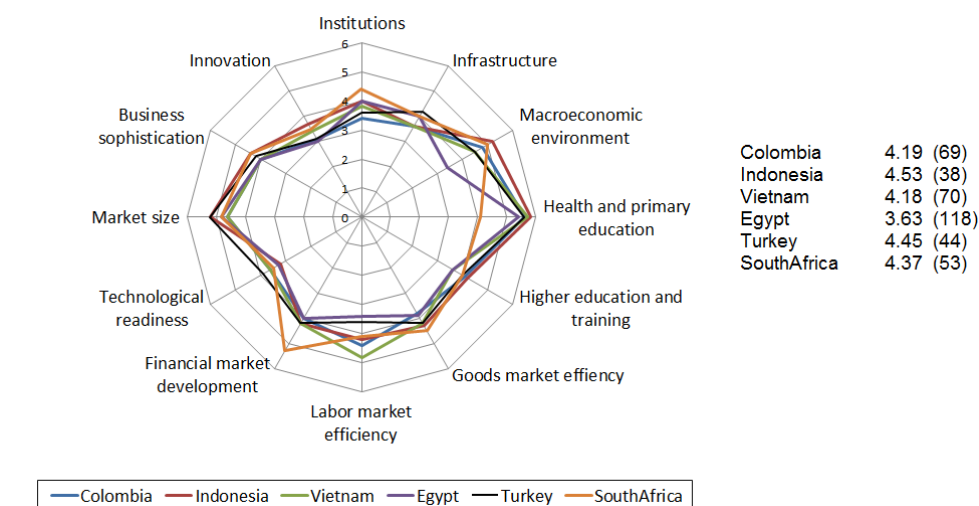


Figura 4.3.2: Patrones de competitividad en países CIVETS
Entre paréntesis la posición ocupada entre 148 naciones. Fuente WEF (2013)

El tercer paso consistió en concentrarse en la producción histórica de patentes que constituye el elemento esencial para medir la innovación (WEF, 2013, 2010). Por medio de la Fig. 4.3.3 se puede afirmar que los países CIVETS presentan un comportamiento similar, a excepción de Suráfrica y Vietnam, en cuanto a la serie histórica de patentes registradas en la base de datos de USPTO y por lo tanto, son susceptibles para realizar comparaciones o validaciones de los modelos y análisis del presente trabajo. La evaluación de este grupo incluyó el conteo del total de patentes/per capita en cada nación. Para nuestro caso, Colombia y Turquía tienen el mismo número y esta es la razón por la cual fueron seleccionados para estudiar sus redes de patentes con mayor profundidad.

¹Patent Cooperative Treaty

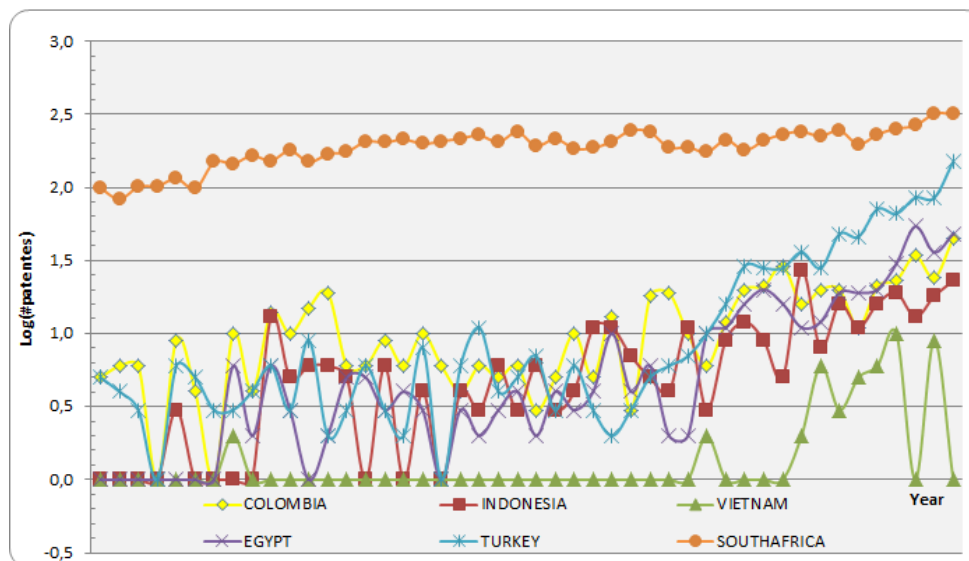


Figura 4.3.3: Series de tiempo de patentes de los países CIVETS registradas en la USPTO.
Fuente USPTO (2011)

El siguiente paso fue minería de datos (Witten, 2005, p.5) para buscar las patentes de estos 2 países. La base de datos usada para encontrar estas patentes fue la Oficina de patentes de los Estados Unidos, USPTO, que según Leydesdorff and Meyer (2010, p.6) es la que ofrece mayor estabilidad de un año al otro en la información registrada. Enseguida, se buscaron las citaciones recibidas por cada una de las patentes. Para ello se usó la aplicación ThomsonInnovation® que busca las patentes en múltiples oficinas de patentes a nivel mundial y genera confianza en la calidad de la información mediante el índice Derwent Patent Citation® (Reuters, 2012). A continuación, se ejecutó el procedimiento de limpieza de los datos (Witten, 2005, p.52-60) por medio del programa VantagePoint®. Al final de este proceso de minería y limpieza de datos se dispuso de una base de datos de cerca de 7000 patentes con sus correspondientes citaciones para un total de casi 100,000 patentes. La base de datos incluye patentes desde 1900 hasta Diciembre de 2013 que corresponde al año T del estudio. Estos datos quedaron resumidos en las matrices de adyacencia (Wasserman and Faust, 1994, p.150) (Ec.4.3.1) de cada una de las redes generadas por las patentes de interés.

$$A_{ij} = \begin{cases} 1 & \text{si patente } i \text{ cita patente } j \\ 0 & \text{de lo contrario} \end{cases} \quad (4.3.1)$$

A continuación, se realizó un análisis estándar de las redes de citación de patentes. Su objetivo es identificar las patentes de los dos países seleccionados que son autoridad y concentrador dentro de la red.

4.4. Centralidad de autoridad y concentrador

La decisión de usar la centralidad de autoridad y concentrador estuvo basada en una comparación de las centralidades principales y los resultados concretos de las patentes bajo estudio. En primer lugar, para todos los nodos (patentes) de las redes surgidas de cada patente, es decir, las redes hacia adelante, se calcularon las siguientes centralidades: Centralidad de intermediación (Ver Definición 11), Centralidad de cercanía (Ver Definición 12), Centralidad de vectores propios (Eigenvector Centrality), Centralidad de Vectores Propios por Componente (Eigenvector Centrality per component), Centralidad de grado entrante (InDegree Centrality), Centralidad de grado saliente (OutDegree Centrality), Centralidad total de grado (Ver Definición 10) y Centralidad interna de cercanía (In-Closeness Centrality) (Wasserman and Faust, 1994). Para ello se usó el programa ORA (Carley et al., 2009) que tiene una rutina que realiza esta comparación. La rutina corre los diferentes algoritmos y produce un cuadro comparativo con todas los valores de las centralidades. Adicionalmente, en el cuadro se registra el top 10 de las patentes según el valor obtenido de las centralidades

de nodo. Posteriormente, con los resultados para todas las patentes de Colombia y Turquía, cerca de 1000, se procedió a analizar los resultados y tomé la decisión de observar en cuales de las centralidades las patentes de estos países habían ocupado el primer lugar del top 10 (Por ejemplo la patente en **negrilla** del cuadro 4.1). Lo anterior, me permitió reducir el conjunto de patentes y concentrarme en las casi 70 que ocuparon el primer lugar. Lo interesante de lo anterior, radicó en la evidencia que los primeros lugares se lograron en la centralidad de vectores propios.

Ranking	Betweenness	Closeness	Eigenvector	Eigenv.Compon
1	USD644441S1	USD677905S1	USD426888S1	USD426888S1
2	USD544215S1	USD666612S1	USD449936S1	USD449936S1
3	USD642808S1	USD683969S1	USD644441S1	USD644441S1
4	USD493525S1	USD683970S1	USD450191S1	USD450191S1
5	USD644442S1	USD683971S1	USD449935S1	USD449935S1
6	USD652632S1	USD670686S1	USD544215S1	USD544215S1
7	USD556462S1	USD674612S1	USD647310S1	USD647310S1
8	USD449936S1	USD652631S1	USD556462S1	USD556462S1
9	USD475849S1	USD652632S1	USD642808S1	USD642808S1
10	USD647310S1	USD669889S1	USD644841S1	USD644841S1
Ranking	InDegree	In-Closeness	Out Degree	Total Degree
1	USD426888S1	USD426888S1	USD544215S1	USD426888S1
2	USD493525S1	USD449936S1	USD556462S1	USD493525S1
3	USD449936S1	USD450191S1	USD642808S1	USD644441S1
4	USD644441S1	USD449935S1	USD644441S1	USD449936S1
5	USD544215S1	USD493525S1	USD644442S1	USD544215S1
6	USD450191S1	USD644441S1	USD644841S1	USD642808S1
7	USD449935S1	USD450190S1	USD644842S1	USD450191S1
8	USD642808S1	USD544215S1	USD645257S1	USD449935S1
9	USD475849S1	USD642808S1	USD647310S1	USD556462S1
10	USD556462S1	USD475849S1	USD652631S1	USD644442S1

Cuadro 4.1: Comparación de centralidades

Esto me pareció bastante curioso, pues la centralidad de vectores propios es la única de las centralidades que realiza el cálculo para cada nodo como función de la centralidad de los demás nodos. O sea, un nodo resulta más central en la medida que está relacionado con otros nodos también centrales (Newman, 2010, p.179-181). Una vez comprendido ésto, estudié con mayor profundidad el tema y encontré que para redes dirigidas, como las de citas, es más apropiado estudiar el comportamiento de la centralidad de autoridad y concentrador. Esta centralidad es similar en su forma de cálculo a la realizada por la centralidad de vectores propios, con la diferencia que para cada nodo(patente) le evalúa 2 cantidades, la centralidad de autoridad y la centralidad de concentrador. De manera que un nodo es considerado como la autoridad sí está relacionado con otros nodos que son los máximos concentradores (Ver gráfica 4.4.1)



Figura 4.4.1: Roles de autoridad y concentrador

De forma que decidí concentrar mi atención tan sólo en las patentes de Colombia y Turquía que hubiesen ocupado el 1er lugar en la centralidad de autoridad. Concluí que este tipo de centralidad permite revelar la estructura principal de la red como función de los nodos que son autoridad y concentrador. Esto es muy importante, pues esta estructura es por la cual se puede afirmar que se realiza la difusión del conocimiento registrado en la patente que es autoridad. Es decir, la patente autoridad es citada por las patentes que son concentradores y éstas a su vez están relacionadas (citadas) por muchas otras patentes que no son tan importantes, pero que resultan usando el conocimiento producido por la patente autoridad. Esto se convirtió en el resultado que yo estaba esperando, en concreto poder identificar aquellas pocas patentes de países emergentes que son influyentes en términos de autoridad en su tema y que además en su mayoría son desconocidas.

Para ello, se construyeron cada una de las ego-redes de diámetro 2 (Wasserman and Faust, 1994, p.134) generadas por las patentes de interés. Estas redes de citación de patentes suelen presentar comportamientos esperados a medida que transcurre el tiempo. El comportamiento típico que emerge se llama estructura modular y jerarquizada (Ravasz and Barabasi, 2003). Este tipo de estructura se destaca por la aparición de patentes que se convierten en **concentradores** y otras en **autoridades** (Leicht et al., 2007). La identificación de estos roles (Newman, 2010, p.179) para cada nodo es posible mediante el cálculo de la centralidades de autoridad, C_A y de concentrador, C_H para cada patente perteneciente a la red. Estas centralidades fueron calculadas usando el algoritmo propuesto por Kleinberg (1999) que resuelve las ecuaciones simultáneas 4.4.2. Este algoritmo está incorporado en el software UCINET (Borgatti et al., 2002).

En términos matemáticos lo anterior se expresa como:

Sean \mathbf{z} y \mathbf{y} dos variables que describen la centralidad de autoridad y de concentrador, C_A y C_H de la red X_T , de forma que

$$\mathbf{z} = C_A \quad \mathbf{y} = C_H \tag{4.4.1}$$

Sí se toma $\mathbf{A}_{ij}(\mathbf{T}) = \mathbf{A}$ por medio de la siguientes ecuaciones simultáneas

$$\mathbf{A}\mathbf{A}'\mathbf{z} = \lambda\mathbf{z}, \quad \mathbf{A}'\mathbf{A}\mathbf{y} = \lambda\mathbf{y} \tag{4.4.2}$$

se encuentran los vectores y valores propios que permiten identificar las patentes con mayores valores para estas centralidades.

De igual forma, con el fin de conocer con mayor detalle la estructura modular, se ha estimado la **modularidad** (Valverde et al., 2007, p.2) de las redes seleccionadas por medio de la Ec.4.4.3. El módulo Q de una red es una medida asociada al número de agrupamientos que mejor describe la red.

$$Q = \frac{1}{2m} \sum_{i,j} \left[a_{ij} - \frac{k_i k_j}{2m} \right] \delta(C_i, C_j) \tag{4.4.3}$$

La modularidad fue estimada usando el algoritmo propuesto por Clauset et al. (2004); Girvan and Newman (2002) y que está incorporado en el programa ORA (Carley et al., 2009)

Como último paso del análisis estático, se realizó la prueba de leyes de potencia propuesta por Clauset et al. (2009). Esta prueba se aplicó a la centralidad de autoridad, C_A , la modularidad, Q y al grado entrante en el año T , $d_i(T)$, para todas las patentes prestigiosas encontradas. Es decir se estimaron las siguientes cantidades:

$$p(C_A \geq c_A) = c_A^{-\alpha} \text{ para } c_A \geq c_{Amin} \tag{4.4.4}$$

$$p(Q \geq q) = q^{-\alpha} \text{ para } q \geq q_{min} \tag{4.4.5}$$

$$p(D_I \geq d_I) = d_I^{-\alpha} \text{ para } d_I \geq d_{Imin} \tag{4.4.6}$$

Las ecuaciones anteriores definen las distribuciones acumuladas complementarias. Este calculo se realizó en lugar de la tradicional $p(K = k)$ con el fin de mejorar la visualización de la distribución de leyes de potencia. Lo anterior es explicado en detalle tanto por Clauset et al. (2009, p.667) como por Newman (2010, p.250-254). El único cuidado que debe tenerse es que al valor α obtenido se le debe restar 1 para obtener el verdadero valor.

Al final del análisis estático se dispone de dos tipos de redes. Aquellas que exhiben leyes de potencia en la distribución de grado y aquellas que no. Por lo tanto, las redes que pasaron la prueba fueron sometidas al análisis longitudinal de redes. El objetivo fue descubrir cuál es el fenómeno “escondido” y que está asociado al comportamiento de leyes de potencia.

4.5. Análisis longitudinal de redes

El comienzo de la etapa para la construcción de las redes longitudinales incluyó el diseño de una matriz de adyacencia en donde el vínculo entre las patentes es el año, t , en el cual se dieron las citaciones y que corresponde con el año de publicación del texto de la patente. De esta forma, el vínculo entre dos patentes es el año de la citación en lugar del valor 1 registrado en la Ec. 4.3.1, con el fin de tener en un solo archivo toda la historia de la red y para poder generar posteriormente la matriz de adyacencia para cualquier año t . Se identificó a esta estructura de información con el nombre de Matriz Anual de Citaciones de Patentes, APCM²:

Sí la citación de i a j se produjo en el año t se define $A_{ij}(t)$ como la APCM $\forall i, j = 1, \dots, N_t$ y para $t = 1, \dots, T$

$$A_{ij}(t) = \begin{cases} t & \text{patente } i \text{ cita } \text{patente } j \text{ antes o en año } t \\ 0 & \text{de lo contrario} \end{cases} \quad t = 1, \dots, T \tag{4.5.1}$$

Con esta matriz se resume en una sola estructura de información toda la historia de citaciones generada por la patente de interés, lo cual posibilita estudiar la información por años. Este modelo está inspirado en la idea de marcos de tiempo propuesto por Brodka et al. (2012) (Fig.4.5.1).

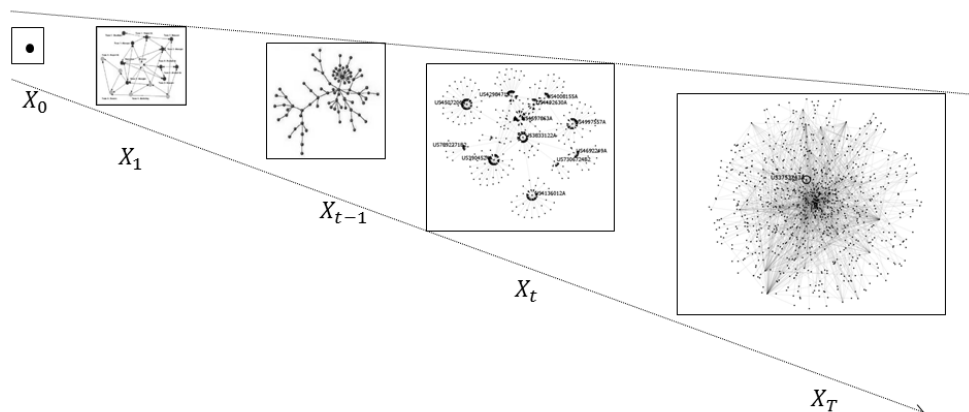


Figura 4.5.1: Redes longitudinales y sus marcos de tiempo

En este momento, se está listo para realizar el análisis de la entropía de información para cada una de las redes longitudinales de patentes. Para este propósito, se tomaron los postulados de Shannon (1948) sobre la entropía de

²Por su sigla en inglés

información y las propuestas de Leydesdorff (1991, 2012) sobre el análisis estático y dinámico de redes usando la teoría de información. El supuesto esencial es :

“En la teoría de la información, los supuestos acerca de la forma de las distribuciones no son necesarias, ya que estos cálculos pueden ser realizados con cualquier función de distribución relativa³” (Bar-Hillel, 1955) citado por Leydesdorff (2012).

Supóngase ahora que la variable X se refiere a una red. Si X cambia en el tiempo desde $t = 0$ hasta un tiempo T máximo se obtiene la serie de la variable X

$$X_0, X_1, \dots, X_t, \dots, X_T \quad (4.5.2)$$

Sea p la probabilidad que el vértice s en la red X_t de tamaño N_t tenga k conexiones en el tiempo t (Dorogovtsev, 2003)

$$p(k, s, N_t) = p(x_t) \quad \text{para } t = 0, 1, \dots, T \quad (4.5.3)$$

y sea $P(X_t)$ la función de probabilidad del indegree de la red X_t de tamaño N_t calculada a partir de $p(x_t)$

$$P(X_t) = \frac{1}{N_t} \sum_{s=1}^{N_t} p(k, s, N_t) \quad \text{para } t = 0, 1, \dots, T \quad \text{y para } k = 0, 1, \dots, N_t \quad (4.5.4)$$

Enseguida, se estimó la entropía de información para todos los momentos t y así se obtuvo la serie propuesta como modelo de análisis en la Ec.4.5.5. En consecuencia, en esta parte se puede calcular la serie temporal de entropías H de la red X de citación de patentes en el tiempo t , a partir de su función de probabilidad del indegree $p(x_t)$.

$$\{H(X_t)\} = \left\{ - \sum_{x_t \in \mathcal{X}_t} p(x_t) \log_2 p(x_t) \right\} \quad \text{para } t = 0, 1, \dots, T \quad (4.5.5)$$

Esta serie de tiempo fue usada como base para calcular otras dos series: Ganancia de Información y Entropía Relativa. Sea IG la ganancia de información (Quinlan, 1986) de la variable X entre una pareja de momentos consecutivos $t - 1, t$ definida como :

$$IG(X_t) = H(X_t) - H(X_{t-1}) \quad (4.5.6)$$

al realizar la operación para $t = 1, \dots, T$ se obtiene la serie de tiempo:

$$\{IG(X_1), \dots, IG(X_t), \dots, IG(X_T)\} \equiv \{IG(X_t)\} \quad \text{para } t = 1, 2, \dots, T \quad (4.5.7)$$

A continuación, sea $D(p \parallel q)$ la entropía relativa o distancia de Kullback-Liebler (Kullback and Leibler, 1951). Si ésta es medida $\forall t$ se obtiene una serie temporal de entropías relativas.

$$\begin{aligned} & \{D(p(x_0) \parallel q(x_0)), D(p(x_1) \parallel q(x_1)), \dots \\ & \quad D(p(x_t) \parallel q(x_t)), \dots, D(p(x_T) \parallel q(x_T))\} \\ & \equiv \{D(p(x_t) \parallel q(x_t))\} = \left\{ - \sum_{x_t \in \mathcal{X}_t} p(x_t) \log_2 \frac{p(x_t)}{q(x_t)} \right\} \\ & \quad \text{para } t = 0, 1 \dots T \quad (4.5.8) \end{aligned}$$

En donde $q(x_t)$ es la distribución de Erdos Renyi.

En consecuencia, se aprovechó la potencia de la idea de incertidumbre que implica el uso de la entropía para descubrir los patrones de innovación oculta en la dinámica de las redes de citación de patentes.

³Traducción propia

Capítulo 5

Resultados

Los resultados de la investigación se clasifican en dos grandes grupos. Los que se derivan del concepto de leyes de potencia y los que son resultado de la aplicación de entropía.

5.1. Resultados del análisis estático: Aplicación de leyes de potencia

El análisis de leyes de potencia fue aplicado a tres variables diferentes y a una función. Ellas son: el grado de las citas directas, d_I , recibidas por cada una de las patentes estudiadas, la centralidad de autoridad, C_A , la modularidad, Q y finalmente, la distribución de grado de la red, $P(X_T)$, surgida de cada patente.

La aplicación del modelo de leyes de potencia arrojó diversos resultados. Los mismos han sido registrados en los cuadros 6.2 y 6.3.

5.1.1. Grado entrante de las patentes, d_I

El primer resultado de la aplicación del modelo de leyes de potencia fue el realizado al número de citas directas de todas las patentes de Colombia y Turquía. La tabla 6.4 que contiene los datos puede consultarse en los Anexos de la investigación. El resultado se muestra en la Fig.5.1.1. En esta puede apreciarse que el modelo de leyes de potencia no es el más apropiado para describir los datos. Los parámetros obtenidos fueron $k_{min} = 27$, $p - value = 0,0120$ y $\alpha = 1,65$. En conclusión puede afirmarse que el grado d_I no exhibe comportamiento de leyes de potencia, puesto que el $p - value \leq 0,1$.

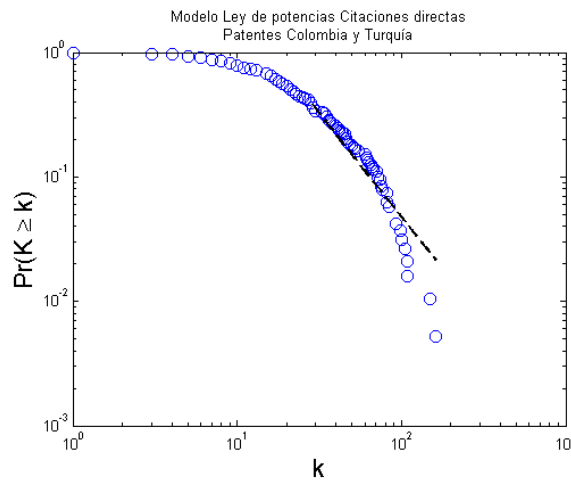


Figura 5.1.1: Prueba de leyes de potencia para el grado (Citas directas k) de todas las patentes

5.1.2. Centralidad de autoridad, C_A

La segunda variable a la cual se le aplicó el modelo de leyes de potencia es la centralidad de autoridad de las patentes. Los valores de esta centralidad están registrados en los cuadros 6.8 y 6.9. El resultado gráfico se muestra en la Fig. 5.1.2. Los parámetros obtenidos fueron $k_{min} = 0,4250$, $p - value = 0,0070$ y $\alpha = 1,8648$. Se concluye entonces que esta variable tampoco exhibe comportamiento de leyes de potencia puesto que $p - value \leq 0,1$

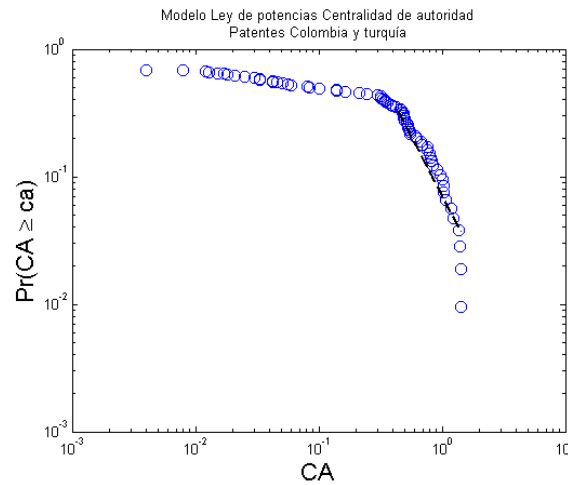


Figura 5.1.2: Prueba de leyes de potencia para la Centralidad de autoridad C_A de todas las patentes

5.1.3. Modularidad de la red Q

Todas las redes estudiadas en esta investigación presentan comportamiento jerarquizado y organizado que se manifiesta en agrupamientos (ver tablas 6.5 y 6.6). Sin embargo, al someter esta variable a la prueba de leyes de potencia, tampoco se percibe dicho comportamiento. Los parámetros obtenidos fueron $Q_{min} = 0,76771$, $p - value = 0,9380$ y $\alpha = 28,16$. Se concluye entonces que esta variable tampoco exhibe comportamiento de leyes de potencia puesto que $p - value \sim 1$, α muy por afuera de los valores reportados en otras investigaciones y además, la muestra ≤ 100 . (Fig.5.1.3)

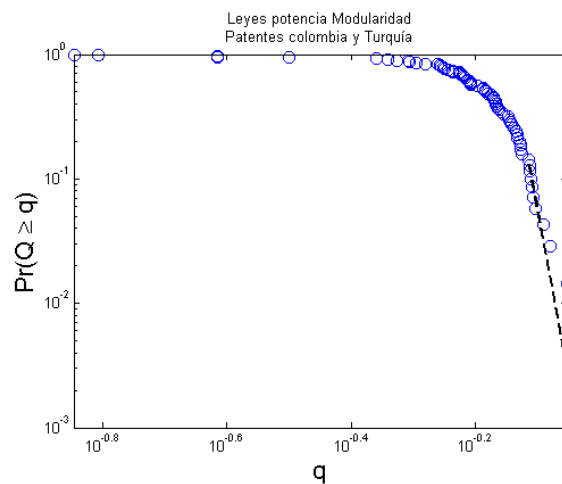


Figura 5.1.3: Prueba de leyes de potencia para la Modularidad

5.1.4. Distribución de grado de las redes de citaciones, $P(X_T)$

El parámetro α o **factor de escalamiento** permite clasificar y caracterizar los fenómenos que exhiben comportamiento de leyes de potencia. Diversos estudios (Clauset et al., 2009; Ravasz and Barabasi, 2003) han medido este exponente para múltiples tipos de variables. De estos estudios los que realizaron medición en redes de citación de patentes son los que nos interesan. Los resultados reportados por algunos de estos estudios son:

$1,12 \leq \alpha \leq 1,38$ estudio de Sheridan et al. (2012) para un subconjunto de patentes de la USPTO altamente citadas. Y $1,66 \leq \alpha \leq 2,37$ estudio de O' Neale and Hendy (2012) sobre distribución de patentes entre los solicitantes de un grupo de países de la OECD.

Para el caso de la presente investigación se encontró que tanto para Colombia como para Turquía el valor tuvo el siguiente rango $0,52 \leq \alpha \leq 2,50$

Los valores que pasan la prueba del $p - value$ y que están registradas en el cuadro 6.7 se muestran gráficamente en las Fig.5.1.4 a 5.1.7 para el caso de Colombia y en las Fig.5.1.8 a 5.1.12 para Turquía. En ambos casos, los resultados se explican por separado.

Los resultados del caso colombiano se clasifican en dos grupos. Uno de ellos corresponde a las patentes que originan redes con más de 500 nodos y donde la organización modular y jeraquizada no es observable mediante los grafos (señaladas con * en el cuadro 6.7). El otro grupo, son redes más pequeñas y donde si es observable la estructura modular jeraquizada. En el primer grupo las redes tienen en común varios aspectos. En todas ellas las patentes que dieron origen a las redes, se encuentran ubicadas en la periferia. Así mismo, en 3 de ellas el valor de escalamiento se ubicó en rangos similares a los valores reportados en la literatura ($1,04 \leq \alpha \leq 1,77$). Por el contrario, en la red de la patente US5133726 $\alpha = 2,37$, que está por encima de los valores reportados. Este comportamiento, aparentemente es explicado porque dicha patente corresponde a un invento que logró difusión de su conocimiento y por esto esta reportada en el cuadro 6.10. Con respecto al otro grupo, en sus redes es observable la estructura modular y jeraquizada a simple vista. Y en todas las redes, la patente generatriz de la red, se ubica en el centro. Con respecto a los parámetros α , Q y m no se observa ningún patrón, aunque varias de ellas también lograron difusión de su conocimiento (Ver Cuadro.6.10). En conclusión, este grupo de redes pasó la prueba del examen de leyes de potencia y esto parece ser indicativo de un comportamiento que debe ser analizado con mayor profundidad. Lo cual es abordado en la sección 5.2.

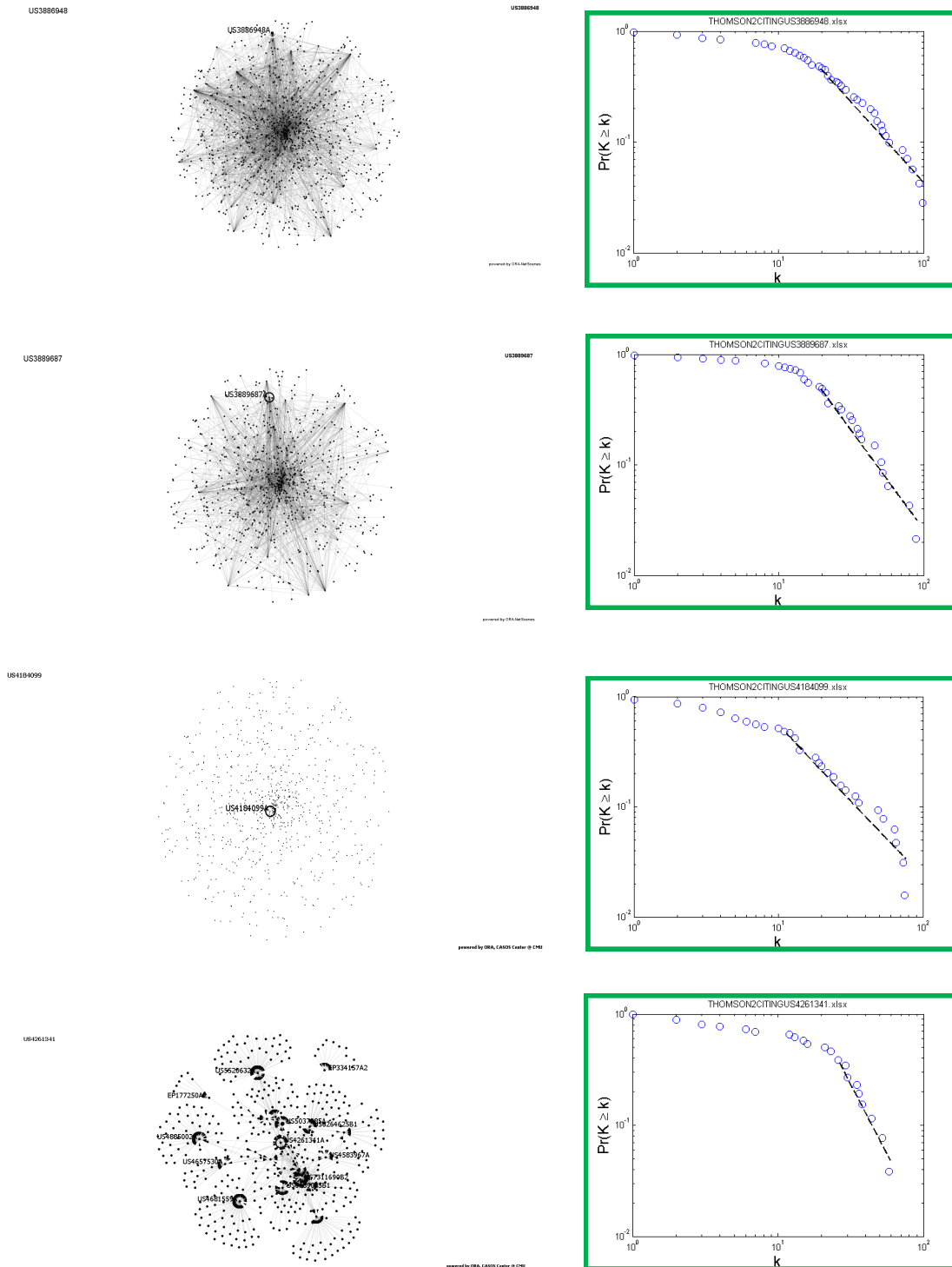


Figura 5.1.4: Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 1
 En orden descendente patentes US38886948, US3889687, US4184099 y US4261341

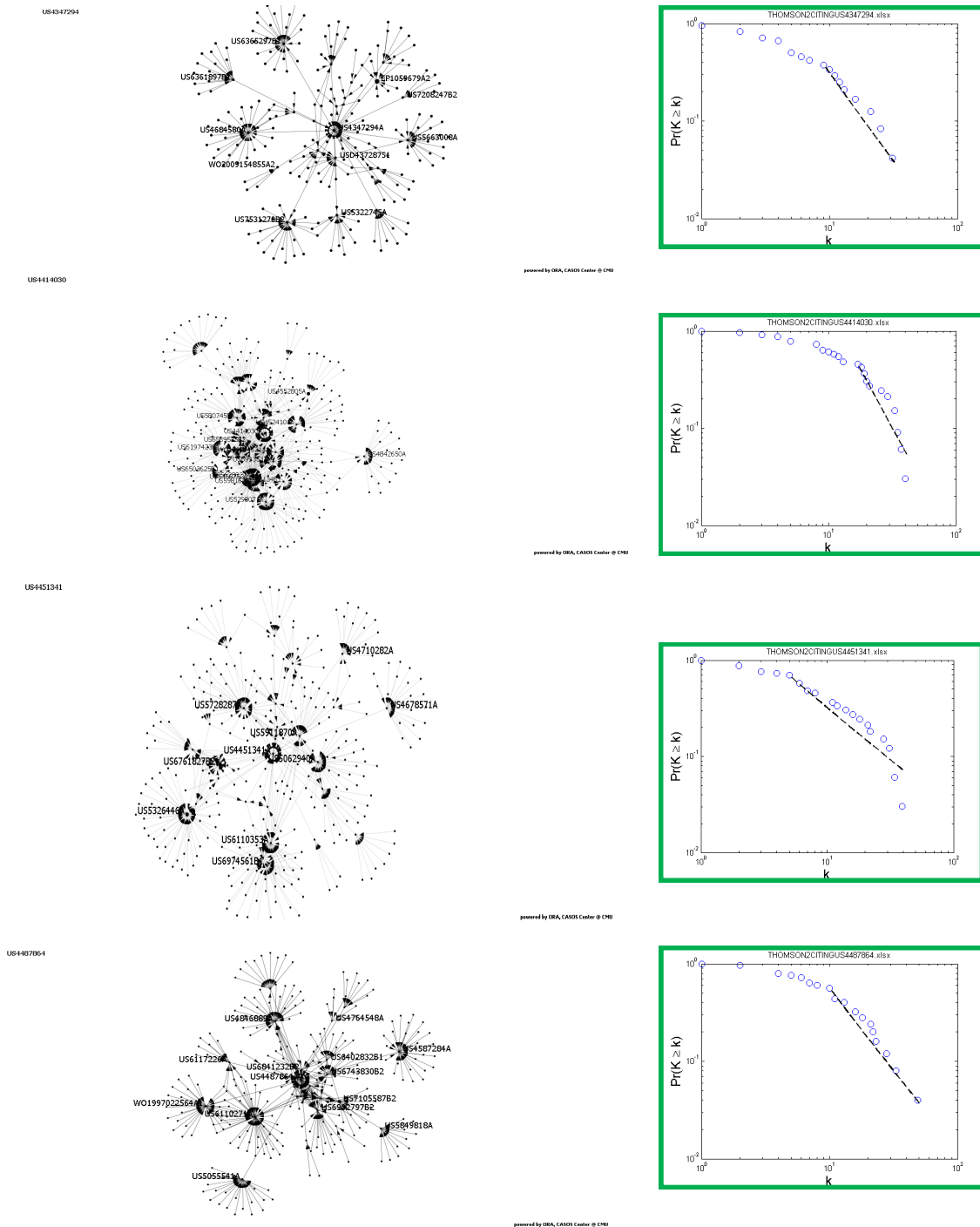


Figura 5.1.5: Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 2
 En orden descendente patentes US4347294, US4414030, US4451341 y US4487864

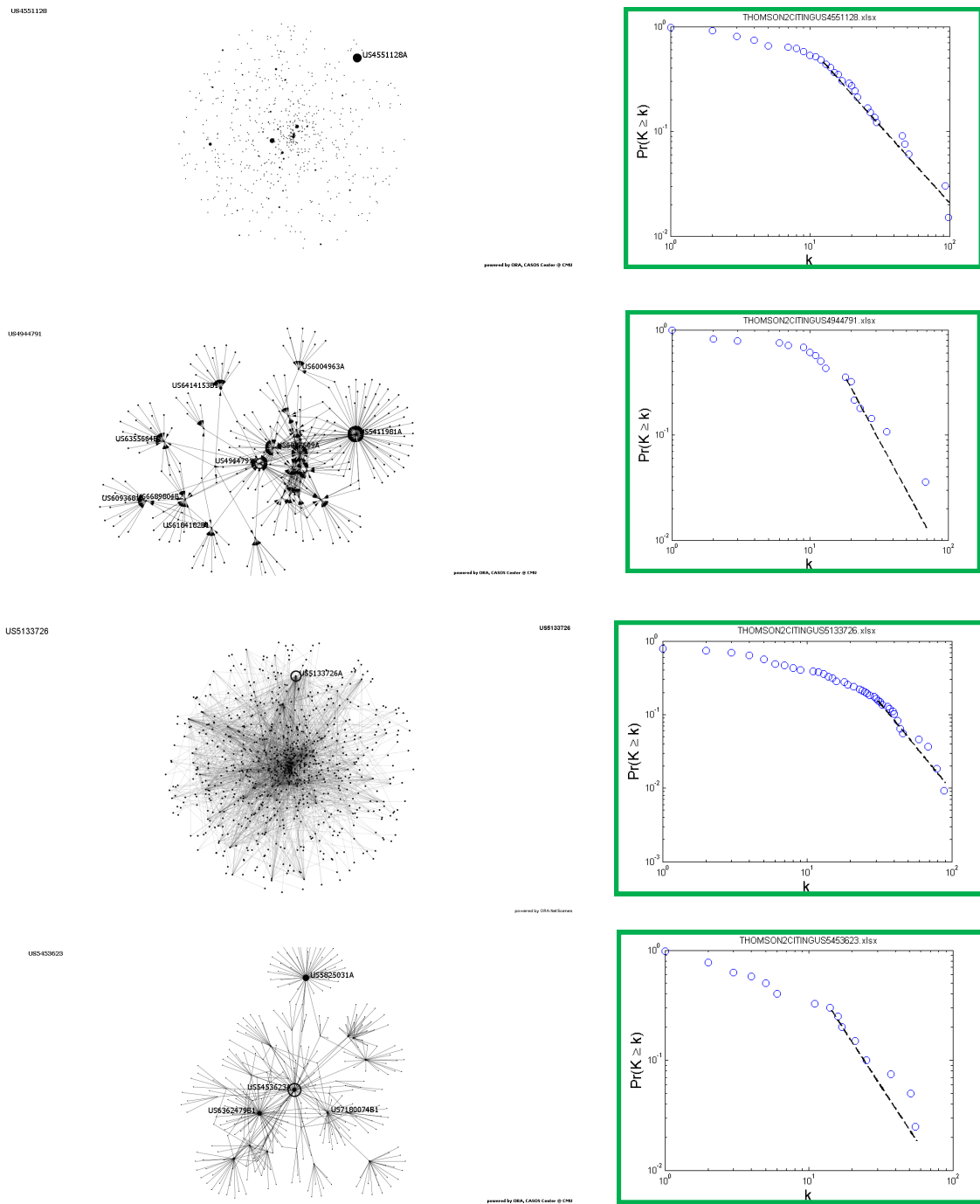


Figura 5.1.6: Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 3
En orden descendente patentes US4551128, US4944791, US5133726 y US453623

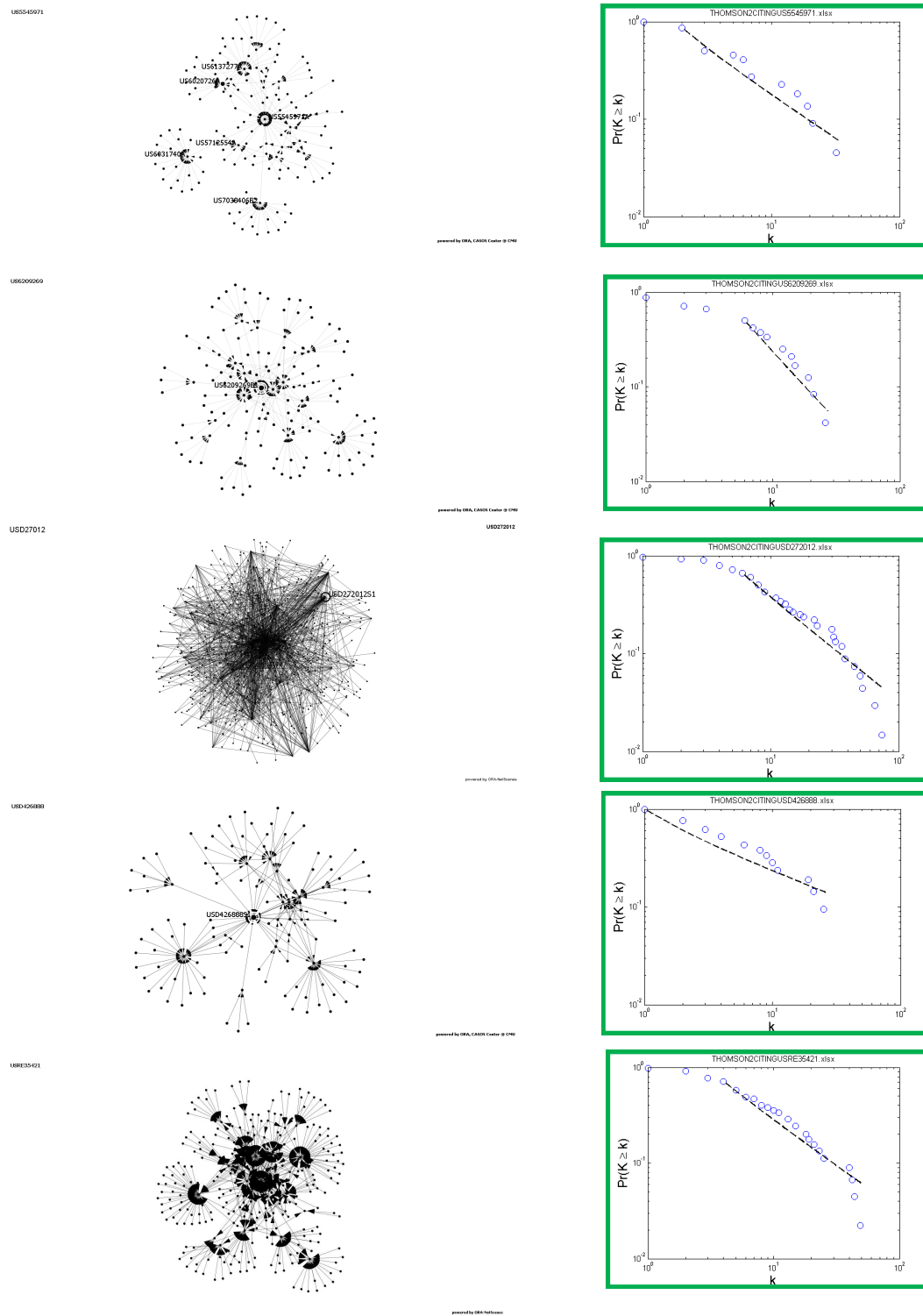


Figura 5.1.7: Patentes colombianas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 4
 En orden descendente patentes US5545971, US6209269, USD27012, USD426888 y USRE35421

Los resultados para Turquía son similares a los colombianos. Existen nuevamente 2 tipos diferentes de redes. Aquellas grandes, densas y donde no es visible la organización típica de este tipo de redes. Las otras, opuestas en tamaño y densidad pero sí visibles en la organización. Las redes marcadas con * en esta ocasión se caracterizan por $1,42 \leq \alpha \leq 2,48$. De estas redes 4 de 5 presentan un $\alpha \geq 2$ valor que está por encima de los valores reportados en la literatura.

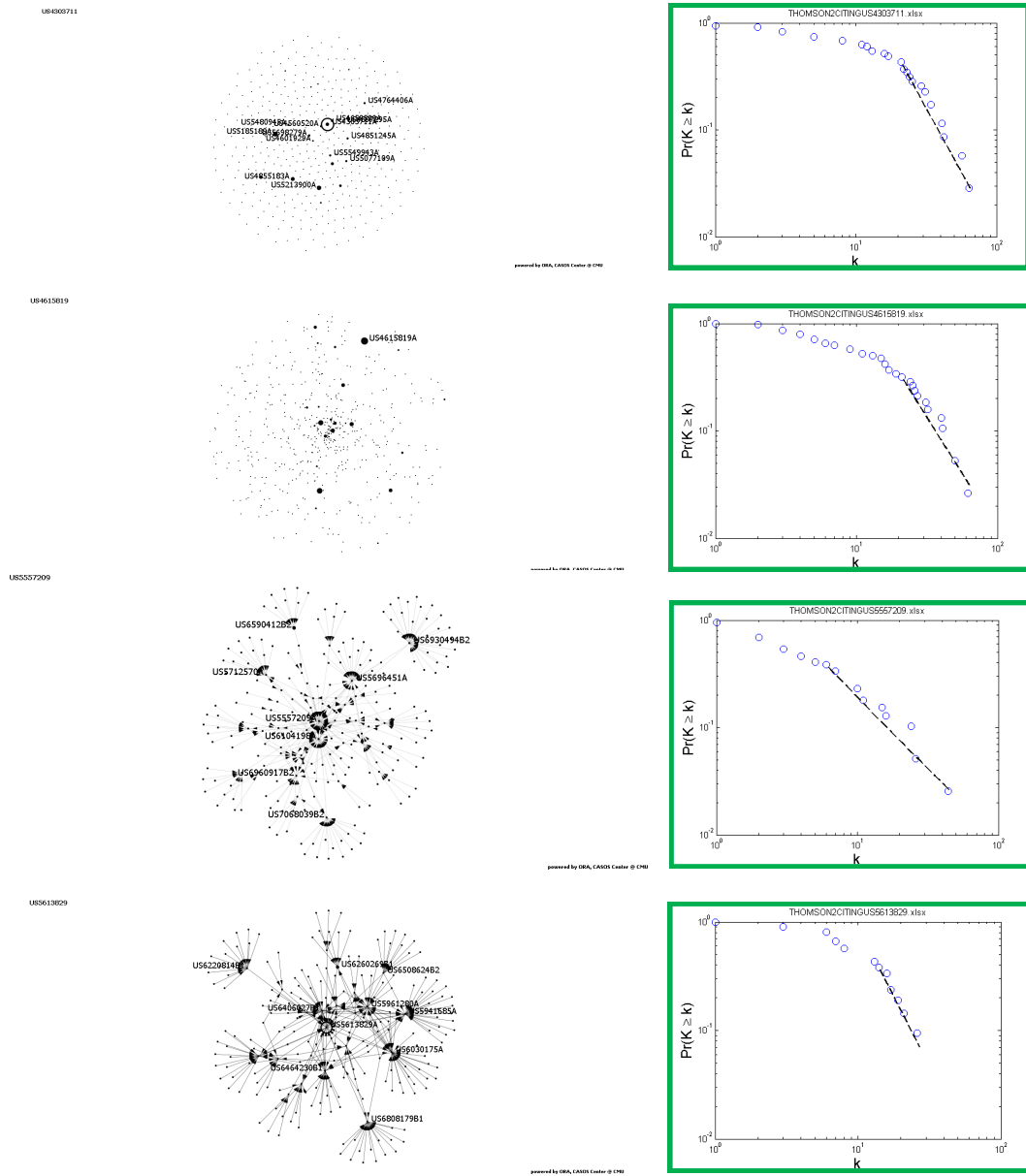


Figura 5.1.8: Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 1
 En orden descendente patentes US4303711, US4615819, US5557209 y US5613829

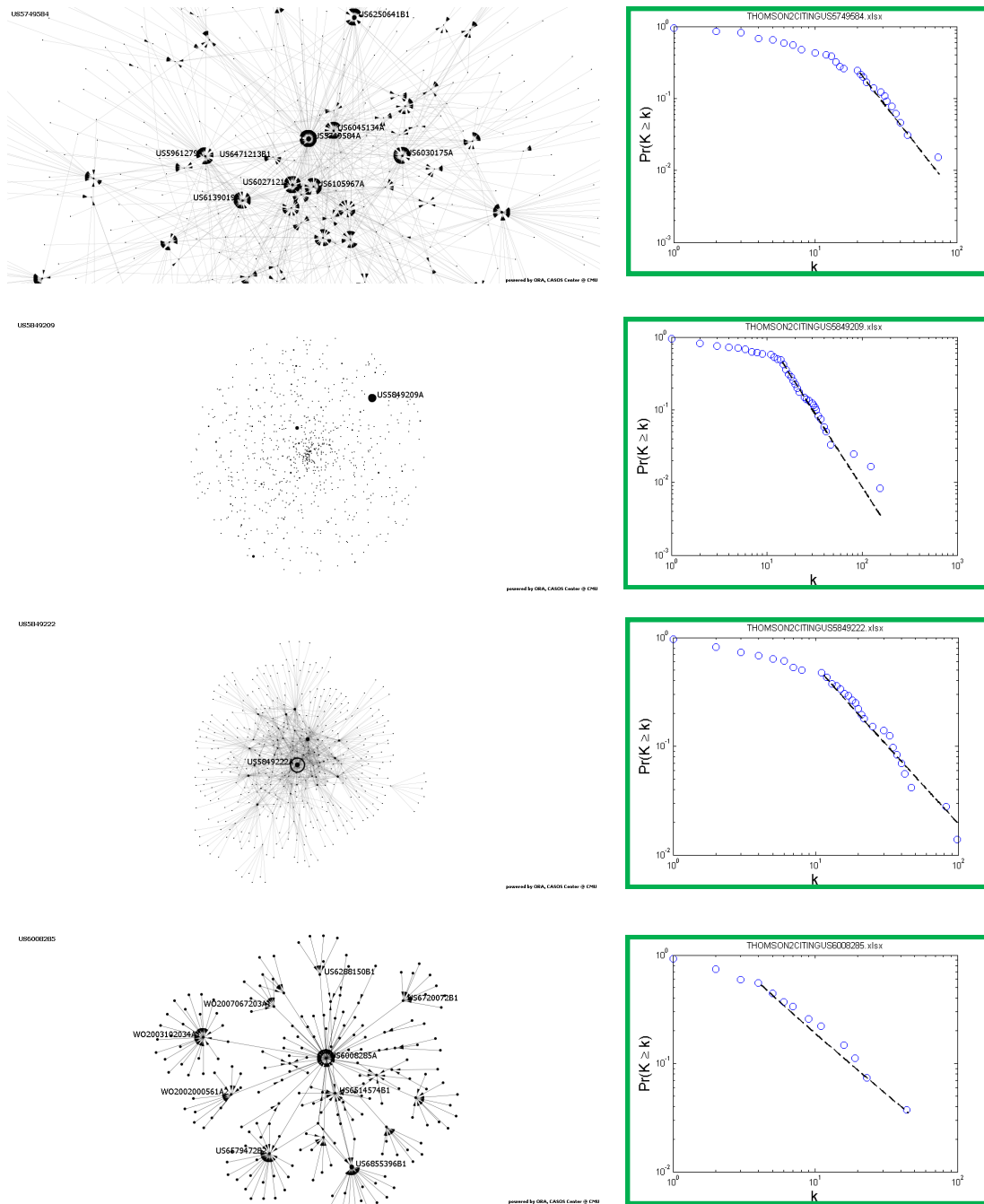


Figura 5.1.9: Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 2
 En orden descendente patentes US5749584, US5849209, US5849222 y US6008285

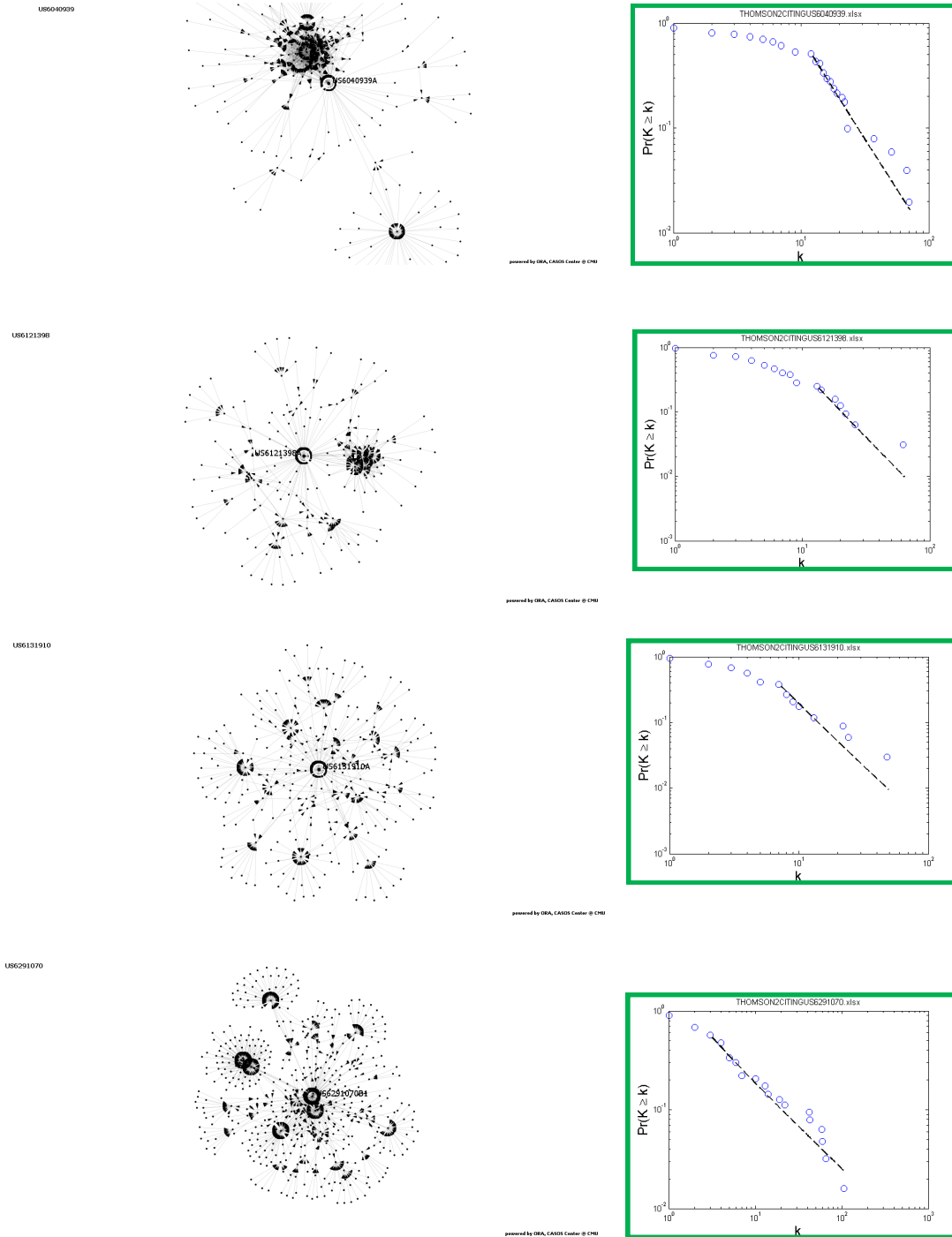


Figura 5.1.10: Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 3
 En orden descendente patentes US6040939, US6121398, US6131910 y US6291070

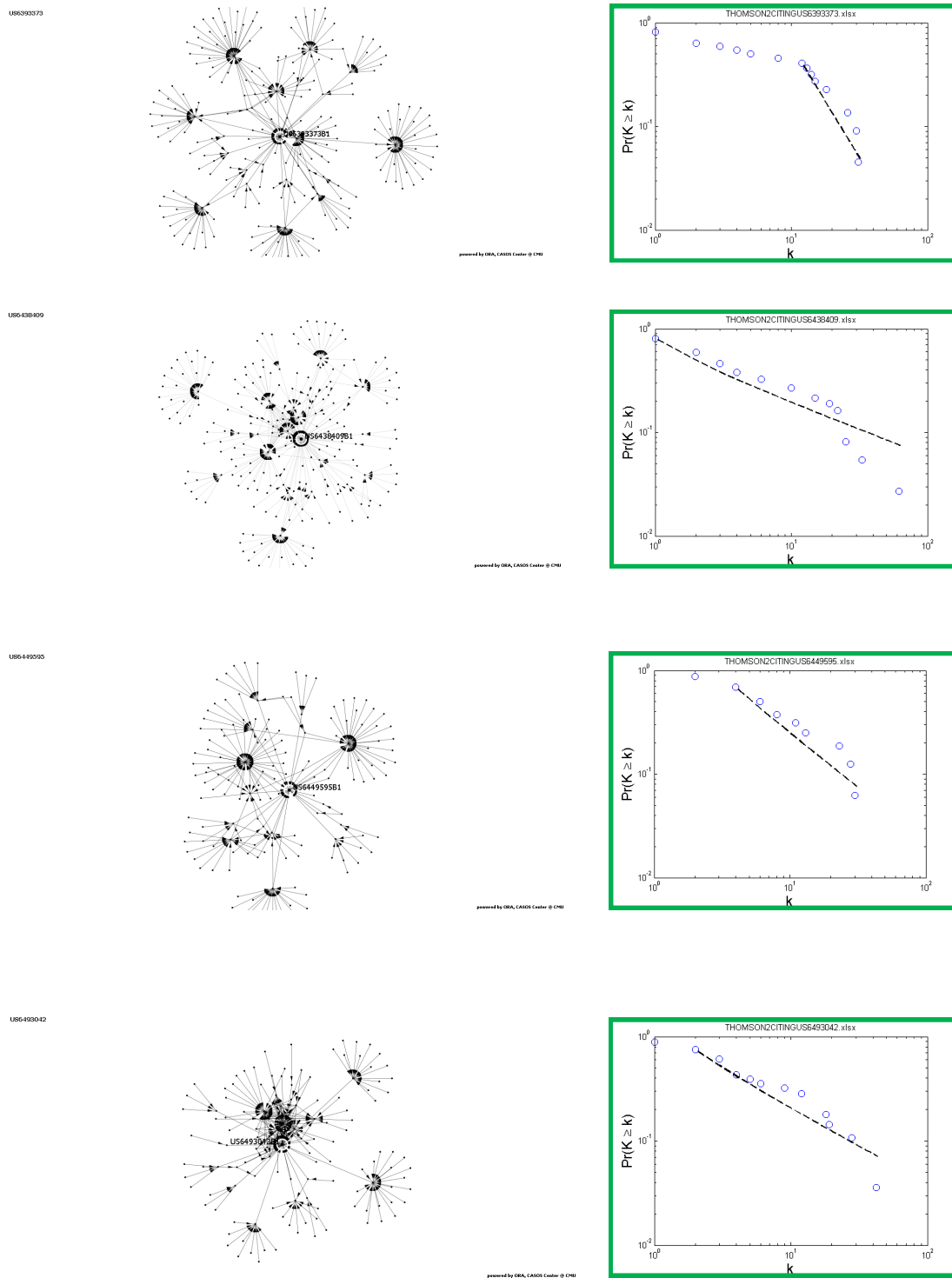


Figura 5.1.11: Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 4 En orden descendente patentes US6393373, US6438409, US6449595 y US6493042

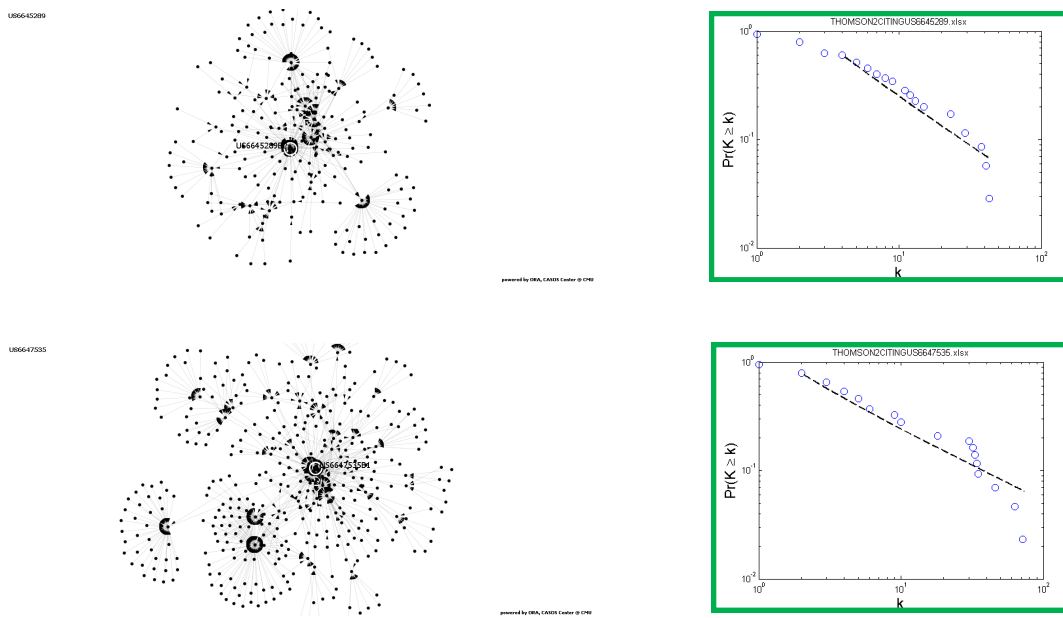


Figura 5.1.12: Patentes turcas que exhiben comportamiento de ley de potencia en el grado d_i . Grupo 5 En orden descendente patentes US6645289 y US6647535

Nuevamente, el comportamiento de estas redes parece estar asociado al fenómeno de la difusión, lo cual será ampliado en la siguiente sección del informe.

5.2. Resultados longitudinales-Aplicación de la entropía

Los resultados están organizados de la siguiente forma. En primer lugar, la evidencia de comportamiento emergente es explicada. Basados en este resultado, la relación con la duración de la protección de la patente es presentada. A continuación, la identificación de los roles de las patentes en la red son evaluados. Después de esto, se muestra la presencia de la difusión de conocimiento. Enseguida, la asociación entre entropía y envejecimiento es develada y seguida por los resultados acerca de la ganancia de información y finalmente, se expone la dinámica de la entropía relativa.

5.2.1. Rol de las patentes en la red

Las series de tiempo de las entropías temporales propuestas mediante la Ec.(4.5.5) permitieron en algunos casos encontrar evidencias de comportamientos emergentes en las redes. Cuando la serie, $\{H(X_t)\}$, muestra la existencia de mínimos locales, implica que la entropía cayó para esos años (Ver Fig. 5.2.1). Esto sucede porque aumentan las citas recibidas por patentes que se van convirtiendo en autoridad, lo cual en términos de entropía de información significa una reducción en el valor.

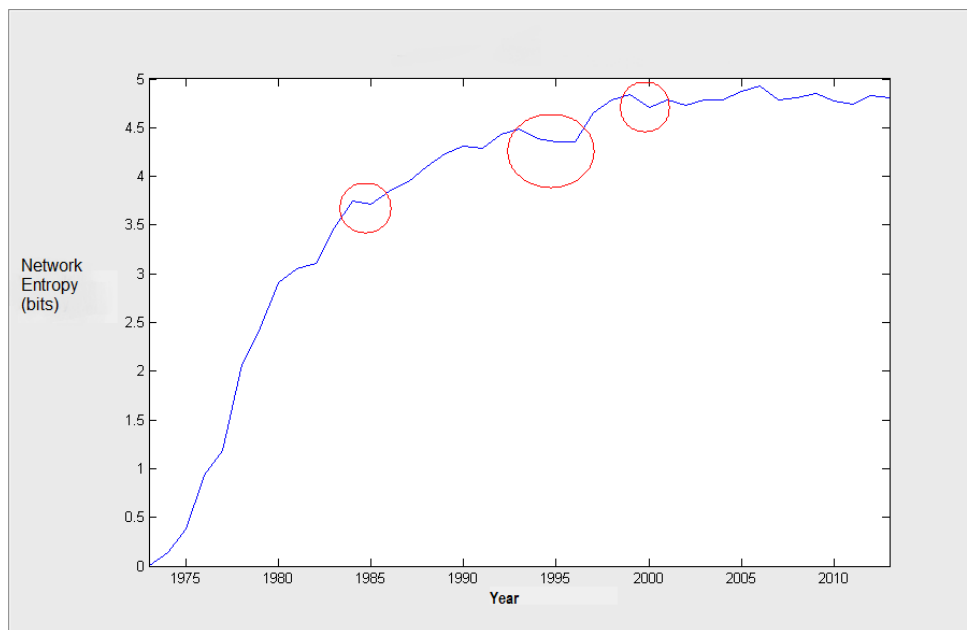


Figura 5.2.1: Evidencia de comportamiento emergente en la red de la patente US3753243

Esto es explicado en la Fig.5.2.2. Para cada tiempo t desde 0 hasta 4 se muestra el cambio en la red, así como los cambios en el grado y la distribución de grado. El nodo a consigue citas de sus nodos vecinos b , c , y d . Como resultado, su grado entrante $d_I(a)$ aumenta de 0 a 3. En consecuencia, la función de probabilidad $P(X_t)$ también cambia. Finalmente, la fórmula de la entropía es usada y la reducción en su valor es el resultado observado. Inicialmente, éste fue un resultado inesperado y parecía incorrecto basado en el hecho que la entropía siempre aumenta (Cover, 2006, p.11). Pero esto sucede solamente en sistema cerrados y las redes de citas de patentes son un sistema abierto. De manera que el comportamiento es correcto.

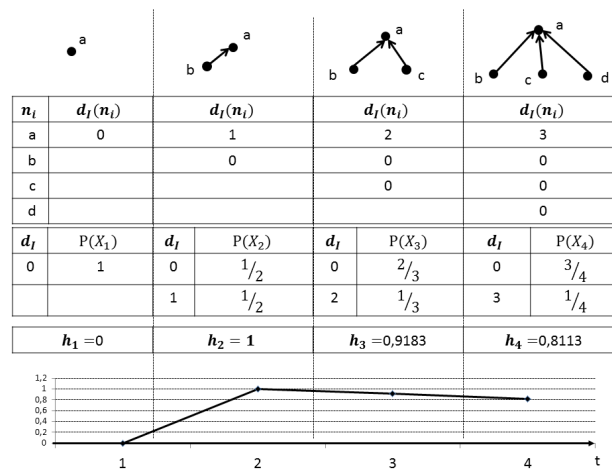


Figure 5.2.2: Explicación de la emergencia de autoridad

$$d_I = \text{Grado Entrante}, p(d_I = x) = \text{Probabilidad}(d_I = x), \text{Entropía}(t) = h(t) = -\sum p(d_I) * \log(p(d_I))$$

Basado en la anterior explicación, se pudo descubrir el rol de las patentes en las redes. Se confirmó, que algunas de las patentes se convirtieron en autoridad dentro de su red, al lograr los más altos valores de centralidad de autoridad. Los valores de la centralidad de autoridad, C_A , para los dos países seleccionados se muestran en los Cuadros 6.8 y 6.9 respectivamente. A manera de ejemplo visual, en la Fig.5.2.3 se ha incluido una red de citación de patentes en donde se puede observar la configuración característica que este tipo de redes asume, con la patente autoridad en el centro y siendo citada por varias patentes que son concentradoras y están ubicadas en la periferia.

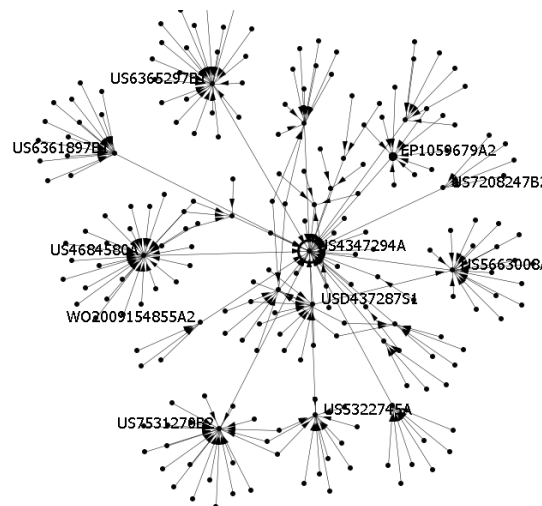


Figura 5.2.3: Red de citaciones generada por la patente US4347294. Patentes en la periferia son concentradores. La patente en el centro es autoridad.

5.2.2. Duración del tiempo de protección de una patente

Al analizar las series de tiempo, se encontró que en promedio el primer mínimo local ocurre τ años después de la publicación de las patentes en donde $10 \leq \tau \leq 20$. Este dato parece coincidir con lo señalado por Oster (2000, p.90) con respecto a la vida promedio de una patente en los Estados Unidos. Adicionalmente, el número de mínimos locales $2 \leq ml \leq 4$, toma valores cercanos a 2 para redes con una antigüedad de 20 años y 4 para redes con antigüedad mayor a los 40 (Ver Fig.5.2.4).

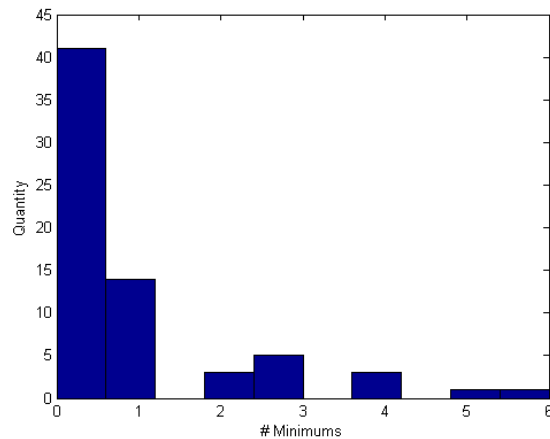


Figura 5.2.4: Mínimos Locales

5.2.3. Difusión de conocimiento

La dinámica de la entropía de información puede ser usada para descubrir el uso o difusión del conocimiento. Con este propósito en mente, una patente entendida como conocimiento empaquetado y codificado genera interés en la medida que es citada por otras patentes. En consecuencia, la red de citas recibidas de forma directa e indirecta se constituye en un objeto que contiene conocimiento que se difunde por medio de los vínculos entre las patentes que conforman la red. A nivel de innovación, Rogers (1962) planteó que la difusión de innovaciones está asociada a la aparición de una curva en forma de “S” que mide el número acumulado de adoptantes de una innovación en particular. De forma similar, en el transcurso de este estudio se encontró que para algunos casos, la serie de la entropía informacional de la red también adopta la forma de la curva logística, conocida popularmente como curva “S”. Este concepto en términos matemáticos se expresa mediante la curva logística, es decir, $\{H_t\} = \frac{a}{1+e^{-(b+ct)}} + d$ (Lilien, 2003, p.36). En conclusión, el conocimiento contenido en una patente se difunde por la red que genera y éste puede ser evaluado mediante la estimación de la entropía informacional de la red a lo largo del tiempo (Fig.5.2.5). Los casos donde se encontró este tipo de curva se relacionan en la tabla 6.10

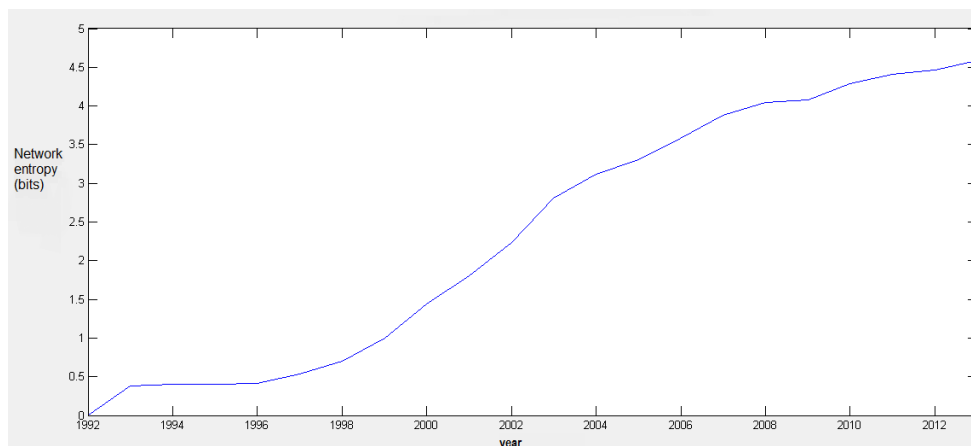


Figura 5.2.5: Curva en forma de “S” de la patente US5133726

La difusión del conocimiento incluido en las patentes se debe a dos razones diferentes para cada uno de los países. En primer lugar, para Colombia (Fig.5.2.6) está posiblemente asociado a un esfuerzo individual de los inventores (Nodos azules), mientras que en Turquía (Fig.5.2.7), la difusión puede estar más asociada a las

empresas multinacionales con sede en grandes países desarrollados (Nodos verdes). Usualmente, estas empresas son las empresas cesionarias de la patente.

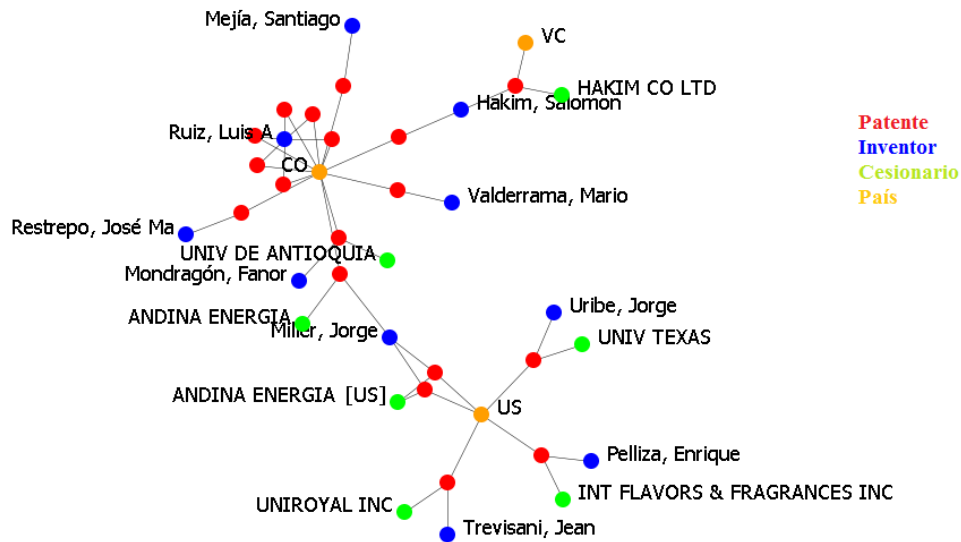


Figura 5.2.6: Difusión de conocimiento en patentes colombianas.

Código de colores: Azul oscuro=Inventor. Verde =Cesionario. Rojo =Patente. Naranja =País del cesionario

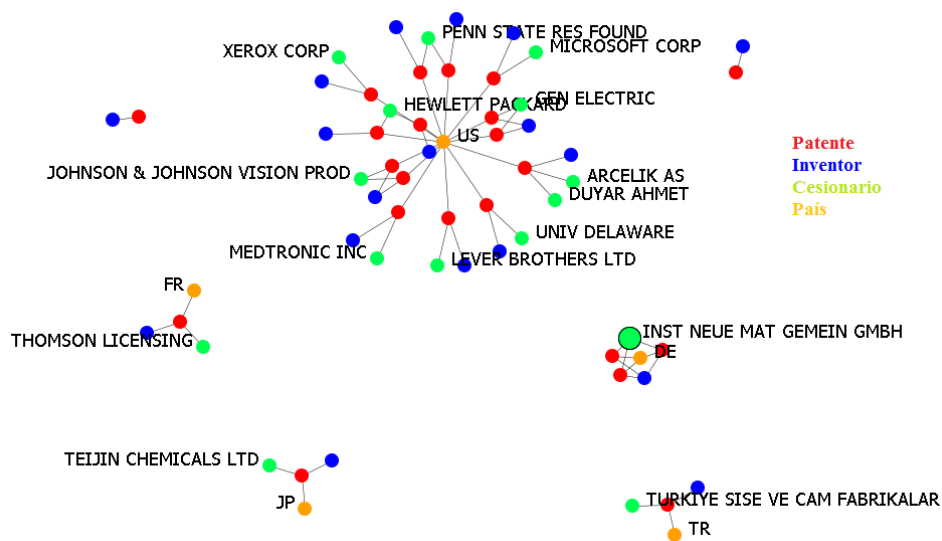


Figura 5.2.7: Difusión de conocimiento en patentes turcas

Código de colores: Azul oscuro=Inventor. Verde =Cesionario. Rojo =Patente. Naranja =País del cesionario

5.2.4. Entropía y antigüedad de la patente

Otro resultado del tipo de información propuesto por la presente investigación tiene que ver con la relación entre entropía y antigüedad de la patente. En este etapa, una vez seleccionadas las patentes que son autoridad, se trabajó con un subgrupo de patentes publicadas desde el año 1973 hasta Diciembre de 2013. Es decir, donde la antigüedad en años considerada como variable de análisis presenta el siguiente rango: $0 \leq Antigüedad \leq 41$,

que permite considerar patentes tanto maduras como jóvenes. El cálculo de la entropía temporal para cada una de las redes consideradas arrojó como resultados : $0 \leq \{H(X_t)\} \leq 5$. Finalmente, en la Fig.5.2.8 se muestra la dispersión por edad de las entropías.

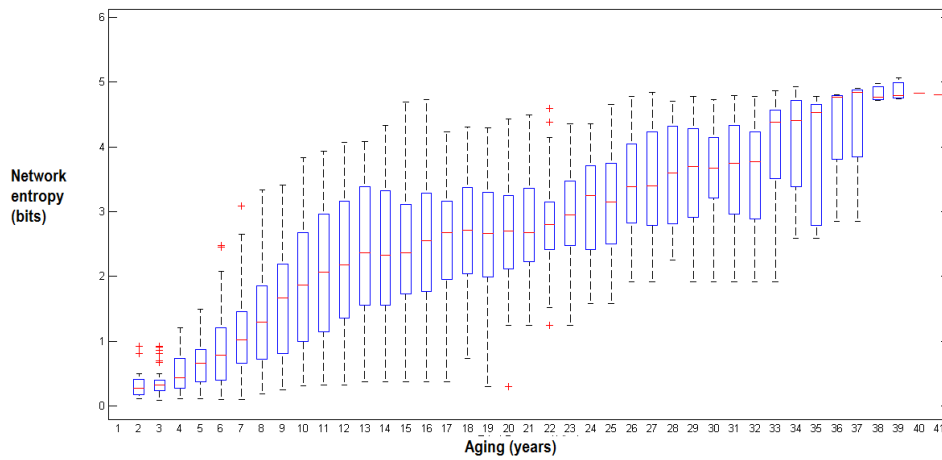


Figura 5.2.8: Entropía y envejecimiento de redes de patentes.

5.2.5. Ganancia de información de la red

La entropía también puede ser usada para estimar el grado de enriquecimiento que obtiene la red a medida que van aumentando las citaciones en el tiempo. Para estimar este variable, se usó la Ec.4.5.7 que calcula el delta de la entropía entre un año y el siguiente. Para este cálculo pueden esperarse dos tipos de resultados: El primero, que a medida que la red crece vaya ganando en información (Fig.5.2.9).

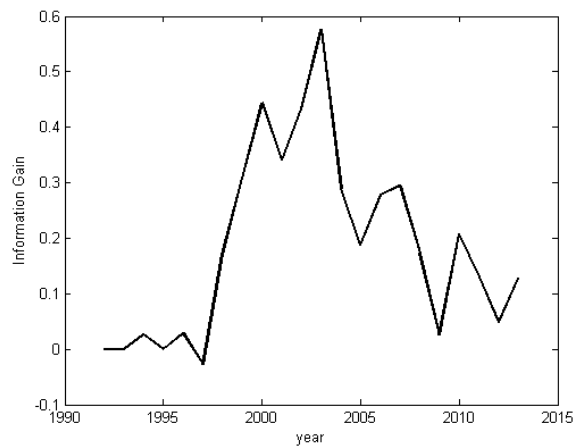


Figure 5.2.9: Ganancia positiva de información de la Patente US5133726.

Esto quiere decir, que se van sumando más patentes que se citan entre sí y por lo tanto, no se concentran las citaciones en unas cuantas patentes consideradas como autoridad.

El segundo resultado, es el contrario del anterior (Fig.5.2.10).

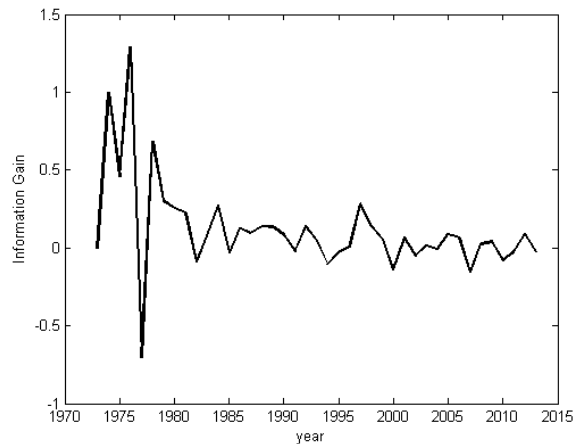
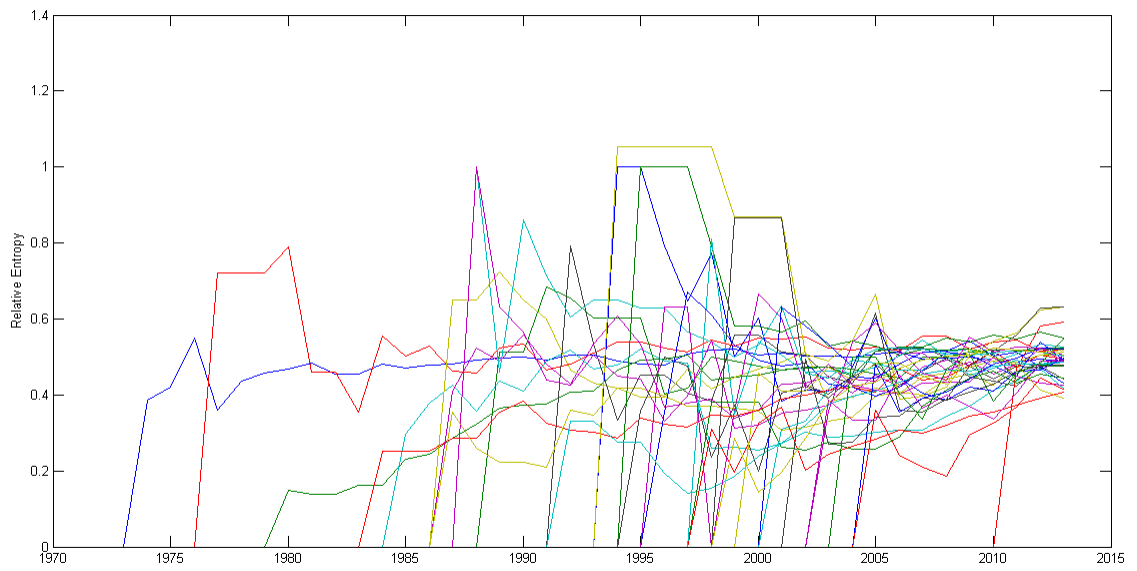


Figure 5.2.10: No ganancia de información para la patente US3753243

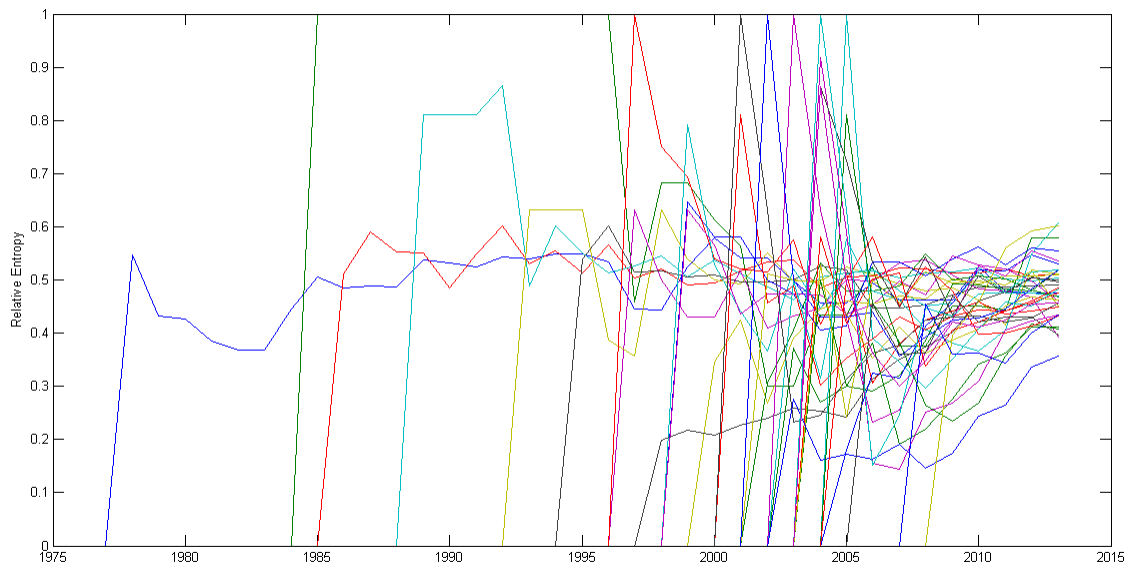
En este caso, en determinados momentos las citaciones recibidas se concentran en unas cuantas patentes que se convierten en autoridad por lo tanto, a medida que crecen las citaciones a estas patentes la variedad en la red no aumenta y por lo tanto, se dice que no hay ganancia en la información, pues todo está concentrado en las patentes mencionadas.

5.2.6. Potencial de la red

El último resultado es una posible respuesta al interrogante: ¿Qué tan alejadas de su máximo nivel de incertidumbre están las redes de patentes? Una aproximación a esta respuesta se plantea mediante la estimación de la entropía relativa o distancia de Kullback-Liebler. Para el caso de la presente investigación, se calculó la entropía relativa tomando como referencia la distribución uniforme, que es la que tiene máxima entropía (Renyi, 1961). La variante usada implica estimar la entropía relativa para todos los años t de la red de citación de patentes y obtener así la serie de tiempo propuesta mediante la Ec. 4.5.8. El resultado obtenido es $:0,4 \leq \{D_t\} \leq 0,6$. Es decir, en promedio las redes alcanzan entre el 40% y el 60% de la entropía máxima para cada uno de los años desde su publicación hasta el presente. El resultado para los dos casos estudiados de Colombia y Turquía arroja resultados similares (Fig.5.2.11)



(a) Caso colombiano



(b) Caso turco

Figura 5.2.11: Entropía Relativa

Capítulo 6

Discusión y proyección investigativa

Podemos explicar con detalle lo que ha sucedido, pero no podemos predecir lo que sucederá en el futuro Bak (1997, p.7)¹

En este capítulo se concluye sobre la veracidad de la proposición presentada en la introducción del trabajo, es decir **el modelo para análisis de redes de citación de patentes integrado por los conceptos de leyes de potencia y entropía de la información si permite descubrir patrones de la dinámica de difusión de la innovación**. La sección se divide en dos. La primera parte, incluye la discusión sobre los resultados y la segunda sección referente a la proyección investigativa.

6.1. Discusión de resultados

6.1.1. Sobre las leyes de potencia

Usualmente, los estudios que proponen la existencia de leyes de potencia se realizan con una cantidad suficientemente grande de datos en términos estadísticos con el fin de poder realizar inferencias que tengan validez estadística. Este estudio tomó una postura diferente y está fundamentada en el trabajo realizado por Clauset et al. (2009) en el cual se realizaron múltiples ejemplos con datos empíricos y sobre una diversidad de tamaño de las variables sobre las cuales se pretendía realizar las pruebas de leyes de potencia. Dicho estudio concluyó que como regla basada en la experiencia, cuando se tienen muestras con $n \gtrsim 50$ se obtienen resultados con un nivel de exactitud cercano al 99% en la estimación del parámetro α de escalamiento. Lo anterior, permitió la realización de las pruebas de leyes de las redes generadas por las patentes de Colombia y Turquía cuyos tamaños se encuentran en el rango propuesto como regla.

El análisis sobre la modularidad Q quedó abierto. En posteriores investigaciones derivadas de este trabajo, es posible realizar la asociación de los módulos con las áreas temáticas que lograron articular las diferentes patentes. Esto es posible, al considerar los códigos de clasificación asignados a cada una de las patentes y los cuales quedan registrados en el texto de la patente, así como en las bases de datos de las diferentes oficinas de patentamiento.

Haber encontrado comportamiento de leyes de potencia en las redes objeto de estudio de esta tesis doctoral no era suficiente. Hubo la necesidad de identificar qué ocultaba este comportamiento para poder afirmar que las redes con este comportamiento son importantes en términos de difusión innovación.

Las implicaciones de lo anterior conducirán a explorar en concreto el contenido de conocimiento registrado en patentes de países emergentes, el cual puede ser apropiadas para otros países emergentes y aún para naciones de mayor grado de competitividad en donde problemáticas propias de países emergentes son cada vez más frecuentes.

¹Traducción propia

6.1.2. Sobre la entropía

Esta investigación examinó la potencia de la entropía de información como una alternativa al conteo de patentes como medida de la innovación. Aún más, la tesis doctoral reveló la innovación oculta de patentes tal y como está definida en (OECD, 2005, p.46). Esto fue realizado por medio de la difusión o uso del conocimiento de una patente hecho por las patentes citantes. Esta forma de entender la innovación excluye patentes sin citas como evidencia de innovación. La sección discute en primer lugar acerca de la metodología y enseguida acerca de cada uno de los resultados presentados en el capítulo 5.

La metodología ofrece una ventaja. Está construida sobre una de las estadísticas esenciales del análisis de redes (*Degree*). Así mismo, da una idea sobre el contenido de información de la red ya que está medida en bits.

La tesis doctoral muestra que el uso diferente de la entropía informacional medida en el tiempo brinda información relevante sobre la patente de interés y la red que surgió a partir de ella. Adicionalmente, se considera que es una medida pertinente para patentes cuyas redes tienen un tamaño entre pequeño y mediano ($10 \leq \text{nodos} \leq 1000$).

El segundo resultado sobre la cantidad de mínimos locales (Fig. 5.2.4) es coherente con la ley de protección a la propiedad intelectual. Esta ley protege patentes por períodos específicos que varían entre 10 y 20 años. Luego, las citas recibidas por una patente aumentan inmediatamente después que la protección finaliza y ésta es la razón para convertirse en autoridad para algunas patentes. Esto implica que para las patentes estudiadas la protección a la invención está yendo más allá de evitar la comercialización o reproducción no autorizada y está incluyendo también el conocimiento incluido en la patente.

Los nodos desempeñan diferentes roles en una red. En el caso de las redes de citación entre patentes es lo mismo. En esta situación, las patentes pueden ser autoridad o concentrador o pueden tener roles aún menos importantes. Una patente que es autoridad posee conocimiento que es importante para otras patentes, no importa si es para propósitos positivos o negativos. Un propósito positivo podría ser que la citación reconoce el valor de la patente como autoridad para tomar ventaja del conocimiento incluido para su propio beneficio. En lugar de esto, un propósito negativo consiste en el rechazo del conocimiento de la patente que es autoridad para apartarse de este. De cualquier manera y a pesar de las críticas hechas a las redes de citas entre patentes, la estructura de la red de citas en términos de autoridades y concentradores es relevante para el análisis de patentes desde las áreas de conocimiento.

La producción de conocimiento (Charum, 2009) consiste en la obtención de datos, los cuales al ser procesados ofrecen resultados en la forma de información, para finalmente ser estudiados en conjunto con expertos temáticos para producir conocimiento. En el caso colombiano, mediante la entrevista a un experto en patentes colombianas (Buitrago, 2014, 2012), se pudo confirmar la importancia de estas patentes que eran desconocidas. Es decir, se confirmó que las patentes son importantes, su conocimiento se está difundiendo y sobre todo, los inventores siguen siendo activos en su campo y presentando más patentes alrededor de las analizadas.

La dispersión de la entropía es alta en términos de envejecimiento (Fig. 5.2.8). Ésta varía desde valores bajos tanto en patentes nuevas como maduras hasta valores altos cuando las patentes tienen entre 7 y 30 años de publicación. Este resultado, está por supuesto, relacionado con la gran variedad de patentes existentes.

Lo anterior confirma lo manifestado por Drucker (2002) *Transcurre un largo tiempo entre la aparición de un nuevo conocimiento y su destilación en una tecnología utilizable. Luego hay otro período prolongado antes que esta nueva tecnología aparezca en el mercado en forma de productos, procesos o servicios. En total, el tiempo de desarrollo involucrado es algo así como 50 años, un lapso que no se ha acortado apreciablemente a lo largo de la historia.*

El resultado sobre ganancia de información parece contraintuitivo (Figs. 5.2.10 y 5.2.9). Si una patente se convierte en autoridad en su campo de estudio, recibe más y más citas de otras patentes. Pero esto implica que no hay nueva información acerca de la patente, su importancia ya es conocida por la comunidad. Por otro lado, cuando una red emerge de una patente y está ganando información, esta situación puede estar relacionada con un área de investigación dinámica o muy activa.

En resumen, las patentes fueron escogidas de manera determinística es decir no se realizó un muestreo aleatorio en el sentido estricto de la palabra. Las mismas fueron escogidas en primer lugar de un grupo de países determinado, los CIVETS. En segundo lugar, sólo se profundizó en el estudio de aquellas con alto nivel de centralidad de autoridad y tercero, se indagó más a fondo el comportamiento de las que exhibieron comportamiento de leyes de potencia. Así mismo, la red que se conforma también es determinística, puesto que

las citas a otras patentes son escogidas directamente por los inventores y en algunos casos sugeridas por los evaluadores de patentes, pero a medida que el tiempo transcurre la red alcanza algún tipo de aleatoriedad (Figs.5.2.11b y 5.2.11b). ¿Por qué es esto posible? será materia de mayor investigación.

6.1.3. Sobre la innovación

Una de las intenciones de este trabajo fue encontrar patentes provenientes de países de menor desarrollo relativo que hubieran tenido algún tipo de impacto. Para lograrlo, se escogieron patentes con alto valor de centralidad de autoridad, como indicador de su importancia y prestigio. Del universo bajo estudio, efectivamente sólo una proporción reducida de patentes se destacó por dicha centralidad. Sin embargo, al considerar posteriormente, cuales de éstas presentaron comportamientos de leyes de potencia en la distribución de grado, se encontraron patentes cuyo conocimiento se difundió por medio de su red asociada y esta idea de uso de conocimiento es lo que se propone como patente innovadora. El concepto detrás de esta propuesta se tomó del Manual de Oslo (OECD, 2005), tomando como base la palabra *implementation* en lugar de la traducción al español como introducción. El listado de las mismas se incluye como parte de los resultados de este trabajo de investigación.

Schumpeter (1911) propuso hace más de un siglo que normalmente la economía se encuentra en un estado de equilibrio y sólo esporádicamente este equilibrio es roto. La innovación está asociada al rompimiento de los estados de equilibrio de la economía cuando un empresario introduce en el mercado dicha innovación. Esta idea schumpeteriana encuentra en el presente trabajo una analogía. A medida que el conocimiento registrado en una patente se difunde por medio de la red de citas, ésta puede convertirse en una autoridad en el momento que la pendiente de la curva “S” supera un valor crítico conocido como el *tipping point* Lee et al. (2011). En dicho instante se produce una “avalancha” de citas o el proceso termina abruptamente. La analogía se expresa en la siguiente cita:

“I will argue that complex behavior in nature reflects the tendency of large systems with many components to evolve into a poised, "critical" state, way out of balance, where minor disturbances may lead to events, called avalanches, of all sizes. Most of the changes take place through catastrophic events rather than by following a smooth gradual path.(Bak, 1997, p.1)”

Al final, Colombia y Turquía son muy similares bajo los parámetros de este estudio. Las proporciones en la producción de patentes, las que son autoridad y las que difunden su conocimiento son casi idénticas. Es decir, éstos 2 países compiten muy estrechamente en términos de innovación desde la mirada de las patentes.

6.2. Proyección investigativa

¿Qué va a ser de mi vida como investigador? La respuesta a esta pregunta es lo que pretendo delinear. Hace unos años en una reunión con el rector de la Universidad Central de ese entonces, profesor Guillermo Páramo (Restrepo and Vasquez (2004)) se estaba hablando sobre el status de investigador en la universidad. La discusión era ardua y había muchas posiciones encontradas. Después de varias horas y cuando los ánimos estaban caldeados el profesor Páramo se paró y dijo palabras más palabras menos lo siguiente:

“A veces las tautologías son necesarias y contundentes. Para todos la expresión $A = A$ es válida e incuestionable. De manera que un investigador es aquella persona que tiene un proyecto de investigación a su cargo y esta persona es casi igual a su proyecto de investigación, puesto que éste representa una opción de vida”.

La discusión terminó después de sus palabras. Por mi parte en este apartado del documento expongo lo que considero puedo abordar como mi proyecto de investigación a corto y mediano plazo. En primer lugar, presento un mapa del terreno de la investigación sobre innovación en Colombia y mi ubicación en el mismo. Posteriormente, expongo posibles líneas de investigación que se desprenden del trabajo doctoral.

6.2.1. Estudios sobre la innovación en Colombia

¿Cómo luce el escenario y qué lugar puede ocupar esta investigación en el espacio de la investigación sobre innovación en Colombia? Se pretende dar una respuesta a estas dos preguntas basado en los documentos que

se han ido recopilando, analizando e incorporando como insumos en el presente trabajo doctoral. Aunque no se intentó realizar un panorama exhaustivo, si se busca presentar un bosquejo del mismo, con el fin de identificar el nicho de investigación concreto que se puede empezar a desarrollar mediante este trabajo. En consecuencia, se muestra a continuación un resumen de los principales investigadores, sus grupos y áreas de interés alrededor del tema de la innovación. Se acompaña el texto con una metared (Krackhardt and Carley, 1998)(Ver gráfica 6.2.1) que sintetiza el numeral. Se han organizado los grupos por las disciplinas en las cuales se ubican.

En primer lugar desde la economía, el profesor Iván Darío Hernández produjo recientemente su libro (Hernandez Umana, 2008) y es en Colombia uno de los principales investigadores, él ha centrado sus reflexiones en el campo llamado economía evolucionista mediante la creación del grupo de investigación de Economía Evolucionista e Institucional (www.virtual.unal.edu.co/GrupoEEI/). En segundo lugar, en la Universidad Javeriana, el grupo de investigación CINNCO (Ver *Web del grupo*), cuyos investigadores principales son Florentino Malaver y Marisela Vargas, ha ubicado su nicho de estudio en la investigación alrededor del *Sistema Nacional de Ciencia Tecnología e innovación (SNCTi)*, la competitividad, la gestión de la tecnología, la innovación, el desarrollo industrial y por último, las políticas industriales de ciencia y tecnología a nivel nacional y regional (Malaver and Vargas, 2007). De manera complementaria, Sanchez et al. (2007) realizaron un exhaustivo estudio sobre las patentes colombianas del período 1968 a 2007 cuyo objetivo principal fue preguntarse cómo podría Colombia insertarse en la socioeconomía del conocimiento. Ellos describieron las áreas de conocimiento, los mayores inventores e inclusive las universidades que produjeron patentes para ese período. A continuación se mencionan 2 trabajos que no han sido realizados por grupos de investigación sino por parte de entidades del Estado pertenecientes al SNCTi. Uno de éstos, es el estudio coordinado por DNP (Maloney, 2008) quien elaboró un trabajo sobre la innovación en Colombia en el marco de las discusiones sobre política de ciencia tecnología e innovación. Y el más reciente proyecto en el tema de patentamiento, fue realizado por la SIC en colaboración con las universidades Rosario, Externado y Sergio Arboleda (SIC et al., 2012), en el marco del programa Hoja de Ruta de la OMPI. Este trabajo se focalizó en la reestructuración del área de patentes de la SIC

Enseguida se localizan los grupos que se han afincado en las facultades de ingeniería. Primero, el profesor Jorge Robledo, de la Universidad Nacional-sede Medellín, se ha focalizado en la investigación sobre Políticas de desarrollo tecnológico e innovación, Capacidades organizacionales y sistémicas para la innovación y Gestión de la I+D. Desde su grupo de investigación Innovación y Gestión Tecnológica (Ver *GrupLac*) ha realizado investigaciones usando técnicas como los sistemas neurodifusos, redes neuronales y dinámica de sistemas (Robledo Velasquez et al., 2009; Robledo Velasquez and Ceballos, 2008; Robledo Velasquez, 2007, 2010; Aguirre R et al., 2009; Duque et al., 2009). De los pocos estudios que se han realizado en Colombia sobre el tema de patentes se encuentra el adelantado por Rivera (2009) estudiante del profesor Jorge Robledo. Este estudio se focaliza en una mirada al sistema de patentamiento en Colombia y propone una reorganización del mismo, apartir del papel que pueden jugar las cámaras de comercio. El estudio recoge parte de las problemáticas del sistema de patentamiento, las cuales condujeron a la aprobación de un documento CONPES (DNP, 2008). También desde la Universidad Nacional en la sede Bogotá, el grupo de Biogestión (www.biogestion.unal.edu.co) ha centrado su interés investigativo en las cadenas productivas y los estudios de *vigilancia tecnológica (Ashton and Stacey, 1995)* (Castellanos et al., 2011; Uribe G et al., 2011). Su mirada se ha dado desde la relación entre la biología y la gestión (Jimenez and Castellanos, 2009) y de ahí el nombre del grupo. En sus diferentes investigaciones han encontrado que la *adaptación* (Stepanic et al., 2005) y la evolución son dos de los conceptos centrales traídos desde la biología para comprender el área genérica de la gestión tecnológica y dentro de ésta la *innovación tecnológica*. Por su parte, en la Universidad de los Andes y desde el grupo de investigación TESO (<http://teso.uniandes.edu.co/>) se ha empezado a realizar modelamiento de procesos de difusión de innovación en los cuales el voz a voz se simula usando el enfoque de redes sociales (Montes et al., 2011). Por último, se encuentra el grupo Complexus (Ver *Gruplac Complexus*) que ha fijado su mirada en los estudios de la innovación desde una perspectiva de la *complejidad*. Su líder ha sido Julio Mario Rodríguez y algunos de los trabajos relacionados con el tema de redes de innovación son:(Rodriguez Devis, 2006), (Perez, 2010; Guerra et al., 2011; Torres, 2007, 2010).

Por otro lado y desde una disciplina diferente a las anteriores de la economía y de la ingeniería, los profesores Rafael Hurtado y Carlos Quimbay de la Universidad Nacional-sede Bogotá, crearon el grupo de investigación Econo-Física y Socio física (Ver *GrupLac*) y se han dedicado a estudiar el tema de la innovación usando los paradigmas de la socio física y la econo-física. Un ejemplo de este enfoque es el trabajo de Otalora et al. (2009) quienes usaron el enfoque de redes complejas para estudiar la información de fuentes de financiamiento registradas en la EDIT y su relación con el riesgo asumido por los empresarios.

6.2.2. Investigaciones en la línea de la tesis doctoral

6.2.2.1. Hipótesis pendientes de probar

¿Qué quedó pendiente por probar? En el capítulo 3 se registraron las hipótesis que guiaron el trabajo doctoral. De éstas hubo algunas que se exploraron y otras que no. Por lo tanto, en esta sección se aborda el trabajo investigativo que se puede derivar de las hipótesis que quedaron esbozadas pero que no han sido trabajadas. Las mismas son:

H_{PC} : $P(\text{Patente CIVETS}=\text{Innovación})=f(\text{Entropía Conocimiento, Internacionalización Referencias, Indegree})$

Esta primera hipótesis se lee de la siguiente forma: H_{PC} : La probabilidad que una patente de un país CIVETS se convierta en innovación es una función de su contenido de conocimiento, el grado de internacionalización de sus referencias y las citaciones generadas. Se decía en el capítulo respectivo que: “Esta hipótesis puede conducir a la construcción de una métrica que combine los estadísticos clásicos de redes como el grado nodal, la centralidad de intermediación y la cohesión, con medidas como la entropía, evaluadas en su dinámica temporal” Avanzar en esta dirección implica combinar los conceptos de teoría de grafos, teoría de la información, con la teoría de la probabilidad. En consecuencia, seguir esta ruta conlleva un énfasis en análisis matemático acompañado de búsqueda de casos exitosos para poder validar las formulaciones. Un ejemplo de posibles trabajos en esta dirección, lo señala Felice et al. (2013).

H_{CAT} : Una patente catalizadora puede conducir a patentes innovadoras.

Es muy probable que una patente resuelva el hueco estructural (Burt, 1992) existente en un área tecnológica y este hecho puede posibilitar que más adelante, otras patentes usen este conocimiento y sean algunas de éstas las que se conviertan en innovaciones. Esta idea está inspirada en Sanjay and Sandeep (2006, p.561) quienes propusieron que el cambio en la estructura de una red compleja al aparecer un nuevo tipo de nodo puede ser un indicador de innovación en la red. Su idea usa la metáfora de las sustancias catalizadoras en las reacciones químicas y este sería el papel jugado por una patente catalizadora. La contribución de la presente tesis en dicha dirección es haber encontrado evidencias que la estructura de la red de citaciones cambia al aparecer un nodo que se convierte en autoridad y otros en concentradores. Éste tipo de patentes pueden ser las innovadoras que plantea la hipótesis.

H_{EI} : La entropía de información contribuye a la identificación de las patentes con mayor contenido de información.

Pudiera suponerse que una medida del valor de una patente está asociada con su contenido de información. Este contenido está registrado en las patentes y las citaciones recibidas. De esta forma, las patentes pueden ser clasificadas según su contenido de información y una patente será más valiosa que otra si su contenido es mayor. A manera de ejemplo, la entropía de información de la red 1 es menor que la entropía de información de la red 2 y por lo tanto, la patente 2 es más interesante que la 1 por tener un contenido de información mayor.

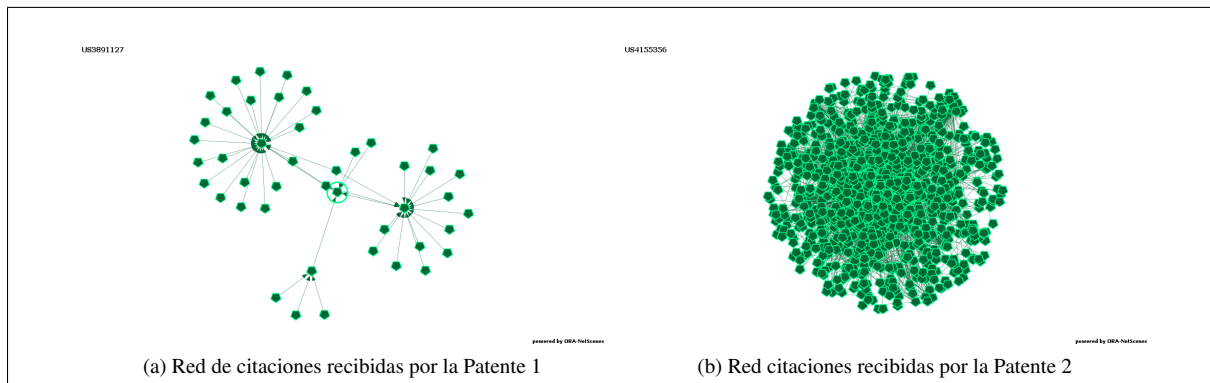


Figura 6.2.2: Contenido conocimiento de patentes
Nodos representan patentes citantes de la patente bajo estudio

H_{PS} : La parábola de la semilla

Esta hipótesis se basa en la importancia del papel del evaluador de patentes. Él es quien confirma las palabras claves sugeridas por el inventor y asigna unas nuevas palabras claves asociadas a la patente y en consecuencia, su trabajo es el que posibilita encontrar la patente por medio de las claves de búsqueda. Un invento puede ser muy bueno en sí mismo, al igual que una semilla mas su germinación está condicionada a la calidad del terreno en el cual es cultivada. En esta situación, las palabras registradas en el documento publicado de la patente, son estratégicas para la búsqueda de la patente.

6.2.2.2. Profundización en los otros países CIVETS

Así como se mostró que los países CIVETS poseen un patrón de competitividad casi idéntico, quedó abierta la pregunta, ¿tendrán los demás países CIVETS en el tema de patentes un comportamiento similar al mostrado por Colombia y Turquía? En el capítulo 5 se mostraron los resultados de aplicar la tesis a las patentes de Colombia y Turquía, por lo tanto, el camino a seguir es aplicar el modelo a Indonesia, Vietnam, Egipto y Suráfrica.

6.2.3. Consolidación del formalismo matemático

¿Cómo se relacionan matemáticamente las expresiones de leyes de potencia, entropía de información y centralidad de autoridad?

Es decir, que relación tienen las expresiones $p(K > k) = k^{-\alpha}$ y $\{H(X_t)\} = \left\{ - \sum_{x_t \in \mathcal{X}_t} p(x_t) \log_2 p(x_t) \right\}$ para $t = 0, 1, \dots, T$

En donde p es la probabilidad que el vértice s en la red X_t de tamaño N_t tenga k conexiones en el tiempo t (Dorogovtsev, 2003)

$$p(k, s, N_t) = p(x_t) \quad \text{para } t = 0, 1, \dots, T \quad (6.2.1)$$

y $P(X_t)$ es la función de probabilidad del indegree de la red X_t de tamaño N_t calculada a partir de $p(x_t)$

es decir $P(X_t) = \frac{1}{N_t} \sum_{s=1}^{N_t} p(k, s, N_t)$ para $t = 0, 1, \dots, T$.

Es más que relación matemática tiene un comportamiento de leyes de potencia con el hecho que en la red haya unos pocos nodos que se comporten como autoridad y concentrador?

Lo anterior tomando en consideración que en términos matemáticos para encontrar las centralidades de autoridad y concentrador se deben resolver las siguientes ecuaciones simultáneas, que representan una clase particular de valores y vectores propios de la red dirigida.

Sean \mathbf{z} y \mathbf{y} dos variables que describen la centralidad de autoridad y de concentrador, C_A y C_H de la red X_T , de forma que

$$\mathbf{z} = C_A \quad \mathbf{y} = C_H \quad (6.2.2)$$

Sí se toma $\mathbf{A}_{ij}(\mathbf{T}) = \mathbf{A}$ por medio de la siguientes ecuaciones simultáneas

$$\mathbf{A}\mathbf{A}'\mathbf{z} = \lambda\mathbf{z}, \quad \mathbf{A}'\mathbf{A}\mathbf{y} = \lambda\mathbf{y} \quad (6.2.3)$$

En conclusión, esta línea de trabajo investigativo busca ahondar en las relaciones existentes entre 3 conceptos: Leyes de potencia, Entropía de Información y Vectores-Valores propios.

6.2.4. La dinámica de redes socio-técnicas

6.2.4.1. Historias de innovación y patentamiento

Uno de los sueños que tengo es poder organizar un evento periódico en el cual se premie a los colombianos por sus patentes. El objetivo de esta actividad es motivar el patentamiento por medio de la fuerza del ejemplo. Poder cristalizar esta idea implica de todas maneras avanzar en una línea de investigación concentrada en el

rescate de las historias existentes detrás del patentamiento. La metodología es muy similar a la construcción de la dinámica de una red, la cual está ejemplificada en la gráfica 4.5.1. La diferencia radica en que para esta línea de investigación los nodos son actores socio-técnicos y no sólo patentes. Uno de los primeros avances en esta dirección es la reconstrucción de la historia de patentamiento del profesor Omar Rodríguez de la Universidad Central (Mejía and Rodríguez, 2012). Esta línea de investigación requiere combinar el modelamiento con la buena narrativa y en consecuencia, explorar alianzas con departamentos de literatura o creación literaria. En el caso concreto de la Universidad Central, la alianza habría que construirla con grupo Heterolalia (<http://www.ucentral.edu.co/investigacion/grupos-y-semilleros>)

6.2.4.2. Redes jurisprudenciales

¿Será el sistema de citaciones de sentencias jurídicas de las altas cortes colombianas auto-organizado?

Leicht et al. (2007) y Bommarito II et al. (2010) en sus investigaciones construyeron redes de citación entre textos de la corte suprema de justicia de los Estados Unidos. La base de datos que usó implicó la revisión de toda la historia de este importante órgano judicial. La investigación encontró quienes eran los magistrados más influyentes en términos de autoridad. En Colombia también es posible realizar este tipo de estudios, puesto que gran parte de las sentencias están sistematizadas.

6.2.4.3. Seguridad

Colombia es un país caracterizado por altos niveles de violencia. Una de las mayores fuentes de violencia es la realizada por grupos terroristas, delincuencia común y en general bandas organizadas. El estudio de la estructura de sus redes es posible gracias a las tecnologías de la información y TIC's. En Colombia, muchas de los organismos de seguridad del Estado es altamente probable que posean información para poder realizar este tipo de investigaciones. A manera de ejemplo Krebs (2002) aplicó el análisis de redes sociales para comprender la estructura de la red de Al Qaeda que derribó las torres gemelas. Avanzar en esta línea de investigación puede ser riesgoso, pero no hay duda que puede ser una fuente de recursos para realizar investigación.

6.2.4.4. Redes de apoyo social

¿Cómo lucen las redes de la felicidad? Christakis (2009) investigador de Harvard mapeó las redes de apoyo de los pacientes enfermos de corazón y confirmó la incidencia positiva que tienen la recuperación de cualquier paciente los grupos de apoyo que se conforman a su alrededor. Su trabajo mostró que las personas infelices y felices tienden a agruparse en grupos diferentes. Construir este tipo de redes es posible mediante el seguimiento a los pacientes sus amigos y familiares.

Anexos

Armenia(AM)	Estonia(EE)	Lithuania(LT)	Rwanda(RW)
Australia(AU)	Finland(FI)	Luxembourg(LU)	SaudiArabia(SA)
Austria(AT)	France(FR)	MacedoniaFYR(MK)	Seychelles(SC)
Bahrain(BH)	Georgia(GE)	Malaysia(MY)	Singapore(SG)
Barbados(BB)	Germany(DE)	Malta(MT)	SlovakRepublic(SK)
Belgium(BE)	Guatemala(GT)	Mauritius(PU)	Slovenia(SI)
Bosnia & Herzegovina(BA)	HongKongSAR(HK)	Mexico(MX)	SouthAfrica(ZA)
Botswana(BW)	Hungary(HU)	Montenegro(ME)	Spain(ES)
Brazil(BR)	Iceland(IS)	Morocco(MA)	SriLanka(LK)
BruneiDarussalam(BN)	India(IN)	Netherlands(NL)	Sweden(SE)
Bulgaria(BG)	Indonesia(ID)	NewZealand(NZ)	Switzerland(CH)
Cambodia(KH)	Iran(IR)	Norway(NO)	TaiwanChina(TW)
Canada(CA)	Ireland(IE)	Oman(OM)	Thailand(TH)
Chile(CL)	Israel(IL)	Panama(PA)	Tunisia(TN)
China(CN)	Italy(IT)	Peru(PE)	Turkey(TR)
Colombia(CO)	Japan(JP)	Philippines(PH)	Ukrain(UA)
CostaRica(CR)	Jordan(JO)	Poland(PL)	UnitedArabEmirates(AE)
Croatia(HR)	Kazakhstan(KZ)	Portugal(PT)	UnitedKingdom(GB)
Cyprus(CY)	KoreaRep(KR)	PuertoRico(PR)	UnitedStates(US)
CzechRepublic(CZ)	Kuwait(KW)	Qatar(QA)	Uruguay(UY)
Denmark(DK)	LaoPDR(LA)	Romania(RO)	Vietnam(VN)
Ecuador(EC)	Latvia(LV)	RussianFederation(RU)	

Cuadro 6.1: Códigos de países

Patente #	N_T	Q	α	x_{min}	$p - value$	Patente #	N_T	Q	α	x_{min}	$p - value$	
US3753243	843	0.833559	2,41	14	0,026	+	US5133726	673	0.523415	3,37	30	0,219
US3886948	823	0.885813	2,43	19	0,144		US5189026	54	0.690716	1,61	1	0,0420
US3889687	608	0.471315	2,77	19	0,156		US5243098	12	0.156248	3,28	20	0,0950
US4184099	664	0.783384	2,35	11	0,601		US5334777	92	0.142485	3,50	22	0,205
US4261341	360	0.649605	3,50	26	0,680		US5439225	61	0.59465	1,59	1	0,245
US4347294	183	0.813474	2,75	9	0,773		US5453623	291	0.721301	2,97	14	0,848
US4350470	71	0.660919	3,07	6	0,263	*	US5533997	500	0.649603	1,61	3	0
US4363737	94	0.736394	3,19	7	0,225	*	US5545971	129	0.738432	1,87	2	0,135
US4414030	235	0.49147	3,37	17	0,135		US5614459	64	0.681728	2,80	7	0,297
US4451341	264	0.776158	2,05	5	0,318		US5622580	102	0.719775	1,74	2	0,155
US4465421	86	0.674246	3,50	10	0,637	*	US6129722	324	0.685046	1,79	4	0,00900
US4487864	241	0.717531	2,64	10	0,338		US6209269	114	0.597393	2,40	6	0,464
US451128	547	0.613525	2,46	12	0,271		US6302877	142	0.580717	1,58	2	0,0980
US4802991	102	0.743785	1,74	3	0,100	*+	US6607527	104	0.694893	2,16	3	0,0520
US4944791	233	0.569602	3,40	18	0,743		USD272012	493	0.620613	2,04	6	0,198
US5056841	77	0.619845	3,47	10	0,571	*	USD426888	111	0.579942	1,52	1	0,115
US5068478	27	0.315793	2,93	6	0,435	*	USRE35421	268	0.601502	1,93	4	0,211

Cuadro 6.2: Patentes colombianas. Existencia leyes potencia. N_T = Número final nodos al año T . Q = modularidad de la red. α = exponente de la ley potencias. x_{min} = Valor a partir del cual puede existir ley potencias. + Significa redes excluidas del análisis por tener $p - value \leq 0,1$ * Significa Precaución $N_T \leq 100$ y $p - value \geq 0,1$ (Clauset et al., 2009, p.678)

Patente #	N_T	Q	α	x_{min}	$p - value$		Patente #	N_T	Q	α	x_{min}	$p - value$
US4047130	64	0.772312	2,11	3	0,0370	*+	US6121398	129	0.550044	3,01	13	0,722
US4141517	22	0.607639	3,50	4	0,769	*	US6131910	191	0.700623	2,84	7	0,581
US4303711	394	0.605599	3,37	21	0,524		US6185469	123	0.773364	1,81	2	0,0340
US4615819	468	0.735388	3,07	21	0,180		US6215231	80	0.713036	1,88	2	0,672
US4796240	42	0.636161	2,80	4	0,315	*	US6291070	431	0.748363	1,85	3	0,524
US5002399	66	0.686548	2,74	5	0,746	*	US6393373	183	0.76771	3,11	12	0,733
US5254953	435	0.656715	1,79	4	0,00500	+	US6397095	83	0.617025	1,74	2	0,171
US5356809	58	0.507493	3,31	7	0,573	*	US6438409	195	0.620594	1,51	1	0,195
US5433060	69	0.436849	1,89	2	0,823	*	US6449595	128	0.746157	2,02	4	0,642
US5557209	205	0.681795	2,29	6	0,462		US6493042	118	0.494266	1,71	2	0,125
US580614	195	0.788129	1,74	3	0,00400	+	US6507947	387	0.668861	1,66	1	0,0670
US5613829	182	0.678511	3,50	14	0,336		US6615174	93	0.65616	2,21	3	0,924
US5749584	425	0.56386	3,48	20	0,796		US6620514	57	0.561851	1,96	2	0,520
US5849209	709	0.595938	3,02	14	0,832		US6645289	192	0.553719	1,88	4	0,413
US5849222	380	0.456288	2,42	11	0,472		US6647535	316	0.673792	1,65	2	0,345
US5858498	89	0.728889	2,33	4	0,345	*	US6713655	32	0.243202	1,58	1	0,127
US6008285	180	0.778639	2,09	4	0,767		US6930981	106	0.751675	1,65	1	0,224
US6040939	136	0.242759	2,89	12	0,444		US7150029	44	0.555357	2,37	2	0,392

Cuadro 6.3: Patentes turcas. Existencia leyes potencia + Significa redes excluidas del análisis por tener $p - value \leq 0,1$ * Significa Precaución $N_T \leq 100$ y $p - value \geq 0,1$ (Clauset et al., 2009, p.678)

#Patente	d_I	#Patente	d_I	#Patente	d_I	#Patente	d_I
US902920A	4	US4802991	23	US6129722	60	US5613829	24
US969500	6	US4822247	4	US6204851	34	US5617828	23
US1271022	1	US4850866	4	US6209269	28	US5674557	13
US3201105	0	US4889591	3	US6299309	21	US5675147	82
US3272618	1	USD256966	6	US6302877	47	US5713113	12
US3527226	81	USD272012	71	US6339720	30	US5749584	84
US3578886	8	USD297811	9	US6605079	28	US5832105	7
US3586197	4	US4932651	27	US6607527	45	US5849209	161
US3633849	16	US4936750	6	US6740078	17	US5849222	100
US3638916	13	US4944791	37	USD426888	26	US5858498	15
US3675656	43	US4956354	8	US4025711	18	US5899385	10
US3750280	19	US4957738	10	US4047130	11	US5908981	40
US3753243	52	US5027650	18	US4070289	22	US6008285	51
US3758908	10	US5040553	11	US4141517	8	US6014598	74
US3774193	4	US5056841	21	US4170723	15	US6040939	72
US3782387	20	US5068478	23	US4177771	5	US6121398	65
US3783538	1	US5133726	108	US4274341	9	US6131910	60
US3805299	4	US5154489	8	US4303711	48	US6136033	25
US3829677	34	US5171839	6	US4320109	16	US6185469	27
US4184099	81	US5189026	13	US4480592	12	US6209120	19
US4214384	71	US5222822	5	US4510618	28	US6215231	28
US4248945	9	US5238042	8	US4544178	35	US6228921	17
US4254002	19	US5243098	61	US4563893	21	US6252975	41
US4261341	33	US5248678	13	US4592347	20	US6272178	38
US4387715	77	US5298329	5	US4592740	15	US6291070	106
US4347294	34	US5302583	7	US4603585	6	US6304349	15
US4350470	17	US5334777	50	US4613986	28	US6337917	28
US4363737	19	US5374411	8	US4615819	42	US6357646	16
US4370543	11	US5411182	13	US4630010	4	US6393373	33
US4396144	16	US5439225	20	US4706048	6	US6397095	29
US4407676	28	US5453623	55	US4768529	17	US6438409	66
US4414030	38	US5458894	22	US4779241	7	US6449595	23
US4451341	36	US5463929	27	US4796240	16	US6475215	41
US4465421	15	US5508046	62	US4870418	7	US6493042	44
US4472673	10	US5511213	20	US5002399	16	US6493288	15
US4487864	37	US5533997	92	US5021085	11	US6507947	150
US4496419	19	US5545971	35	US5115414	9	US6615174	46
US4515450	8	US5579784	13	US5227351	12	US6620514	35
US4551128	109	US5614459	18	US5254953	67	US6645289	45
US4595390	98	US5619718	27	US5320023	7	US6647535	74
US4608992	35	US5622580	29	US5356809	16	US6713655	25
US4615691	84	US5710487	13	US5357040	17	US6930981	29
US4616111	9	US5722657	9	US5422117	11	US7150029	29
US4668300	5	US5863301	20	US5433060	46	US7354909	18
US4696158	63	US5928129	40	US5470910	84	USD570620S	21
US4727701	6	US5980543	45	US5557209	46	USD577869S	24
US4728525	7	USRE35421	55	US5580614	20		
US4772257	75	US6016446	22	US5587307	27		

Cuadro 6.4: Grado entrante d_I , patentes de Colombia y Turquía

Patente #	N_T	Q	m	Patente #	N_T	Q	m
US3753243	843	0.833559	24	US5133726	673	0.523415	17
US3886948	823	0.885813	12	US5189026	54	0.690716	6
US3889687	608	0.471315	5	US5243098	12	0.156248	6
US4184099	664	0.783384	17	US5334777	92	0.142485	6
US4261341	360	0.649605	12	US5439225	61	0.59465	8
US4347294	183	0.813474	14	US5453623	291	0.721301	13
US4350470	71	0.660919	9	US5533997	500	0.649603	11
US4363737	94	0.736394	11	US5545971	129	0.738432	9
US4414030	235	0.49147	8	US5614459	64	0.681728	7
US4451341	264	0.776158	12	US5622580	102	0.719775	9
US4465421	86	0.674246	8	US6129722	324	0.685046	18
US4487864	241	0.717531	12	US6209269	114	0.597393	10
US4551128	547	0.613525	18	US6302877	142	0.580717	5
US4802991	102	0.743785	8	US6607527	104	0.694893	12
US4944791	233	0.569602	9	USD272012	493	0.620613	15
US5056841	77	0.619845	7	USD426888	111	0.579942	8
US5068478	27	0.315793	3	USRE35421	268	0.601502	14

Cuadro 6.5: Modularidad redes patentes de Colombia.

Patente #	N_T	Q	m	Patente #	N_T	Q	m
US4047130	64	0.772312	9	US6121398	129	0.550044	9
US4141517	22	0.607639	4	US6131910	191	0.700623	14
US4303711	394	0.605599	11	US6185469	123	0.773364	11
US4615819	468	0.735388	17	US6215231	80	0.713036	11
US4796240	42	0.636161	7	US6291070	431	0.748363	18
US5002399	66	0.686548	8	US6393373	183	0.76771	13
US5254953	435	0.656715	19	US6397095	83	0.617025	8
US5356809	58	0.507493	7	US6438409	195	0.620594	12
US5433060	69	0.436849	7	US6449595	128	0.746157	10
US5557209	205	0.681795	15	US6493042	118	0.494266	10
US5580614	195	0.788129	10	US6507947	387	0.668861	26
US5613829	182	0.678511	11	US6615174	93	0.65616	12
US5749584	425	0.56386	15	US6620514	57	0.561851	12
US5849209	709	0.595938	10	US6645289	192	0.553719	12
US5849222	380	0.456288	10	US6647535	316	0.673792	16
US5858498	89	0.728889	8	US6713655	32	0.243202	4
US6008285	180	0.778639	18	US6930981	106	0.751675	15
US6040939	136	0.242759	7	US7150029	44	0.555357	8

Cuadro 6.6: Modularidad redes patentes de Turquía.

Patente #	N_r	Q	m	α	x_{min}	$p - value$	Patente #	N_r	Q	m	α	x_{min}	$p - value$	
US3886948	823	0.885813	12	1,43	19	0,144	*	US4944791	233	0.569602	9	2,40	18	0,743
US3889687	608	0.471315	5	1,77	19	0,156	*	US5133726	673	0.523415	17	2,37	30	0,219
US4184099	664	0.783384	17	1,35	11	0,601		US5453623	291	0.721301	13	1,97	14	0,848
US4261341	360	0.649605	12	2,50	26	0,680		US5545971	129	0.738432	9	0,87	2	0,135
US4347294	183	0.813474	14	1,75	9	0,773		US6209269	114	0.597393	10	1,40	6	0,464
US4414030	235	0.49147	8	2,37	17	0,135		USD272012	493	0.620613	15	1,04	6	0,198
US4451341	264	0.776158	12	1,05	5	0,318		USD426888	111	0.579942	8	0,52	1	0,115
US4487864	241	0.717531	12	1,64	10	0,338		USRE35421	268	0.601502	14	0,93	4	0,211
US4551128	547	0.613525	18	1,46	12	0,271								

(a) Caso Colombia.

Patente #	N_r	Q	m	α	x_{min}	$p - value$	Patente #	N_r	Q	m	α	x_{min}	$p - value$	
US4303711	394	0.605599	11	2,37	21	0,524	*	US6121398	129	0.550044	9	2,01	13	0,722
US4615819	468	0.735388	17	2,07	21	0,180	*	US6131910	191	0.700623	14	1,84	7	0,581
US5557209	205	0.681795	15	1,29	6	0,462		US6291070	431	0.748363	18	0,85	3	0,524
US5613829	182	0.678511	11	2,50	14	0,336		US6393373	183	0.767710	13	2,11	12	0,733
US5749584	425	0.56386	15	2,48	20	0,796	*	US6438409	195	0.620594	12	0,51	1	0,195
US5849209	709	0.595938	10	2,02	14	0,832	*	US6449595	128	0.746157	10	1,02	4	0,642
US5849222	380	0.456288	10	1,42	11	0,472	*	US6493042	118	0.494266	10	0,71	2	0,125
US6008285	180	0.778639	18	1,09	4	0,767		US6645289	192	0.553719	12	0,88	4	0,413
US6040939	136	0.242759	7	1,89	12	0,444		US6647535	316	0.673792	16	0,65	2	0,345

(b) Caso Turquía

Cuadro 6.7: Patentes que exhiben comportamiento de leyes de potencia en el grado. N_r = Número final nodos al año T . Q = modularidad de la red. m = #módulos α = exponente de la ley potencias. x_{min} = Valor a partir del cual puede existir ley potencias.

Patente	CA	Patente	CA	Patente	CA	Patente	CA
US4175868	1.414	US5508046	0.616	US5243098	0.346	US4254002	0.051
US6607527	1.413	US5133726	0.605	US3675656	0.335	US5248678	0.046
US4347294	1.411	US3886948	0.565	US4214384	0.32	US4117468	0.045
US5622580	1.39	US4487864	0.546	US4261341	0.318	US5579784	0.042
US5545971	1.374	USD272012	0.541	US6016446	0.303	US5710487	0.042
US4184099	1.24	US4615691	0.53	US4081288	0.286	US3638916	0.033
US5439225	1.176	US3758908	0.526	US4451341	0.246	US4772257	0.033
US6129722	1.083	US3774193	0.516	US3829677	0.214	US3527226	0.03
US4136417	1.033	US3889687	0.503	US4106510	0.211	US3877137	0.026
US3578886	1.029	US3924635	0.502	US4155356	0.181	US4109865	0.026
US5056841	1.023	US5463929	0.499	US5619718	0.17	US3782387	0.025
US3783538	1	US6740078	0.499	US5040553	0.165	US4078699	0.025
US3972985	1	US5928129	0.493	US4696158	0.139	US5863301	0.025
US5453623	0.973	US4944791	0.49	US5411182	0.139	US4496419	0.021
USD426888	0.904	US4396144	0.485	US4098864	0.137	US3633849	0.018
US5068478	0.841	US4957738	0.474	US4608992	0.1	US3991912	0.018
US4350470	0.828	US4414030	0.472	US3835354	0.099	US4932651	0.015
US4551128	0.815	US4595390	0.455	US6299309	0.085	US4472673	0.013
US6302877	0.79	US5980543	0.425	US6204851	0.081	US5511213	0.012
USRE35421	0.757	US4363737	0.397	US4051751	0.069	US5722657	0.008
US5533997	0.754	US5027650	0.385	US6339720	0.06	US4407676	0.004
US6209269	0.684	US5334777	0.358	US3833122	0.056		
US5614459	0.68	US5458894	0.349	US5238042	0.056		

Cuadro 6.8: Valores de la centralidad de autoridad para patentes colombianas.

Patente	CA	Patente	CA	Patente	CA	Patente	CA
US4141517	1,414	US6228921	0,877	US5675147	0,391	US4480592	0,114
US6620514	1,411	US5849222	0,775	US4025711	0,374	US4544178	0,112
US6008285	1,407	US5749584	0,76	US6272178	0,363	US5899385	0,099
US7150029	1,405	US6490342	0,743	US6493288	0,363	US5115414	0,094
US6131910	1,4	US6647535	0,743	USD577869	0,333	US5021085	0,063
US6615174	1,393	US6136033	0,734	US4563893	0,328	US5908981	0,062
US6215231	1,387	US6713655	0,718	USD570620	0,327	US5470910	0,043
US4796240	1,383	US6357646	0,712	US4070289	0,306	US4592347	0,033
US5557209	1,219	US5858498	0,701	US4615819	0,25	US4768529	0,03
US6507947	1,178	US6645289	0,595	US6475215	0,206	US5422117	0,016
US6291070	1,122	US6121398	0,577	US5617828	0,176	US5587307	0,008
US6393373	1,027	US5849209	0,543	US6252975	0,175	US4592740	0,002
US5433060	1,008	US5356809A	0,522	US5580614	0,151	US5320023	0,002
US6438409	0,956	US6337917	0,459	US5613829	0,15		
US5254953	0,939	US6040939	0,4	US5832105	0,129		
US6397095	0,924	US4303711	0,396	US5674557	0,121		

Cuadro 6.9: Valores de la centralidad de autoridad para patentes turcas.

PATENTE #	CÓDIGO PAÍS	LEY POTENCIA?	PATENTE #	CÓDIGO PAÍS	LEY POTENCIA?
US3886948	CO	Si	US4615819	TR	Si
US4184099	CO	Si	US5254953	TR	
US4347294	CO	Si	US5433060	TR	Si
US4414030	CO	Si	US5749584	TR	Si
US4551128	CO	Si	US5849209	TR	Si
US5068478	CO		US5849222	TR	Si
US5133726	CO	Si	US6008285	TR	Si
US5243098	CO		US6040939	TR	Si
US5334777	CO		US9121398	TR	Si
US5453623	CO	Si	US6131910	TR	Si
US5533997	CO		US6215231	TR	
US5614459	CO		US6291070	TR	Si
US6129722	CO	Si	US6393373	TR	Si
US6209269	CO	Si	US9397095	TR	
US6302877	CO		US6438409	TR	Si
US6607527	CO		US6507947	TR	
USD272012	CO		US6615174	TR	
USRE35421	CO		US6620514	TR	
			US6645289	TR	Si
			US6647535	TR	Si
			US6713655	TR	
			US6930981	TR	
			US7150029	TR	

Cuadro 6.10: Patentes que muestran difusión de conocimiento.

Definiciones

Definición 1. Modelo de Erdos & Renyi: Sea $E_{n,N}$ el conjunto de todos los grafos que tienen n vértices etiquetados como V_1, V_2, \dots, V_n y N bordes. Los grafos considerados se suponen no orientados, sin bordes paralelos y sin ondas. Así un grafo que pertenece al conjunto $E_{n,N}$ se obtiene seleccionando N de los posibles $\binom{n}{2}$ aristas entre los puntos V_1, V_2, \dots, V_n y de esta forma el número de elementos de $E_{n,N}$ es igual a $\binom{\binom{n}{2}}{N}$. Un grafo aleatorio $\Gamma_{n,N}$ puede ser definido como un elemento de $E_{n,N}$ seleccionado al azar, de forma que todos los elementos de $E_{n,N}$, tienen la misma probabilidad $1/\binom{\binom{n}{2}}{N}$ de ser escogidos. (Erdos and Renyi, 1960)

Definición 2. Problema de mundo pequeño. La formulación del problema se inicia con la pregunta ¿Cuál es la probabilidad que cualquiera par de personas, seleccionadas al azar de una gran población, como los EU, se conozcan entre sí? Sin embargo, una formulación más interesante, toma en cuenta que mientras dos personas a y z puede que no se conozcan entre sí directamente, pueden compartir una o más personas conocidas; es decir, puede existir un conjunto de individuos, B (compuesto por b_1, b_2, \dots, b_n) quienes conocen tanto a a como z y por lo tanto, los vinculan entre sí. De forma más general, a y z pueden estar conectados no por un sólo conocido común, sino por una serie de tales intermediarios, $a - b - c - \dots - y - z$, i.e a conoce a b ; b conoce a a además de c , c a su vez conoce a d , etc. (Milgram, 1969)

Definición 3. Modelo de Kossinets y Watts. Este modelo, desarrollado por Kossinets and Watts 2006a calcula la probabilidad p_{new} de una nueva relación entre los individuos i, j , en función de la longitud del camino más corto que los une d_{ij} y de la cantidad de afiliaciones compartidas entre ellos s_{ij} :

$$p_{new}(d_{ij}, s_{ij}) = \frac{\sum_{t=61}^{270} M_{new}(d_{ij}, s_{ij}, t)}{\sum_{t=61}^{270} M(d_{ij}, s_{ij}, t)} \quad (6.2.4)$$

donde $M(d_{ij}, s_{ij}, t)$ es la cantidad de parejas de vértices (Individuos) en la categoría (d_{ij}, s_{ij}) , al momento t y $M_{new}(d_{ij}, s_{ij}, t)$ es la cantidad de nuevas parejas de la misma categoría desde el período anterior $t - 1$.

Definición 4. Modelo de Butts (Butts, 2008) propuso un modelo donde los eventos siguen una distribución de Poisson. Su modelo es el siguiente:

$$p(A_t) = \left[\prod_{i=1}^M \left[h(\tau(a_i) | s(a_i), r(a_i), c(a_i), \mathbf{X}_{a_i}, \mathbf{A}_{\tau(a_{i-1})}) \times \right. \right. \quad (6.2.5)$$

$$\left. \left. \prod_{a_t \in \mathbb{A}(\mathbf{A}_{\tau(a_i)})} S(\tau(a_i) - \tau(a_{i-1}) | s(a_t), r(a_t), c(a_t), \mathbf{X}_{a_t}, \mathbf{A}_{\tau(a_{i-1})}) \right] \right] \quad (6.2.6)$$

$$\times \prod_{a_t \in \mathbb{A}(A_t)} S(t - \tau(a_M) | s(a_t), r(a_t), c(a_t), \mathbf{X}_{a_t}, \mathbf{A}_t) \quad (6.2.7)$$

$$Si \lambda = h(t) y S(t) = \exp(-\lambda(t - t')) \quad (6.2.8)$$

$$\lambda_{a_t, \theta} = \lambda(s(a), r(a), c(a), X_a, A_t, \theta) \quad (6.2.9)$$

$$p(A_t | \theta) = \left[\prod_{i=1}^M \left(\lambda_{a_i A_{\tau(a_{i-1})} \theta} \prod_{a_t \in \mathbb{A}(A_{\tau(a_i)})} \exp(-\lambda_{a_t A_{\tau(a_{i-1})} \theta} (\tau(a_i) - \tau(a_{i-1}))) \right) \right] \quad (6.2.10)$$

$$\times \left[\prod_{a_t \in \mathbb{A}(A_t)} \exp(-\lambda_{a_t A_t \theta} (t - \tau(a_M))) \right] \quad (6.2.11)$$

$$p(A_t | \theta) = \prod_{i=1}^M \left[\frac{\lambda_{a_i A_{\tau(a_{i-1})} \theta}}{\sum_{a_t \in \mathbb{A}(A_{\tau(a_i)})} \lambda_{a_t A_{\tau(a_{i-1})} \theta}} \right] \quad (6.2.12)$$

Definición 5. Modelo de Nooy

Propuesto por de Nooy (2011) quien estudió como evento (y) los momentos en los cuales se conforman los vínculos entre autores de libros y sus críticos literarios. El modelo define como nodos a los autores (g), los críticos (f), los libros llamados episodios (j), la popularidad de los autores (x_{tjfg}), los períodos de las revisiones (t) y una identificación del sexo de cada uno de los autores y críticos (x_{jfg}). Por medio de la penúltima variable incorporaron la dinámica en el modelo y con la última posibilitaron la realización de análisis sobre la incidencia de la **homofilia** en las críticas. Matemáticamente, el autor desarrolló un modelo estocástico basado en la clase de modelos denominados como **modelos de eventos históricos** que define el riesgo (h) que aparezca una crítica literaria en el tiempo (t) por parte del crítico (f) del libro (j) escrito por el autor (g) por medio de la ecuación:

$$h_{tjfg} = Pr[y_{tjfg} \neq 0 / y_{sjfg} = 0, s < t] \quad (6.2.13)$$

donde con una función logit busca predecir la ocurrencia del evento

$$\text{logit}(h_{tjfg}) = \log\left(\frac{h_{tjfg}}{1 - h_{tjfg}}\right) = \alpha(t) + \beta^T \mathbf{x}_{tjfg} + u_j + v_f + w_g \quad (6.2.14)$$

Definición 6. Entropía de información (Entropía de Shannon) Sea X una variable discreta con alfabeto χ y una función de probabilidad $p(x) = Pr\{X = x\}$, $x \in \chi$. La entropía de la variable X se define como

$$H(X) = - \sum_{x \in \chi} p(x) \log p(x) \quad (6.2.15)$$

(Shannon, 1948)

Definición 7. Principio de Máxima Entropía (MaxEnt): Método para encontrar la distribución de probabilidad para describir los estados (micro) de un sistema, basado en las mediciones (i.e valores promedios de funciones) hechos a nivel macro. El método también es conocido como Principio de Máxima Incertidumbre, Método de Máxima Entropía o simplemente como el formalismo de Jaynes (Jaynes, 1957, 1978). El principio motiva la selección de la distribución que maximiza la medida de entropía de Shannon y que respete las restricciones sobre los valores promedios dados.

Matemáticamente se formula como:

Maximize

$$-\sum_{i=1}^n p_i \ln(p_i)$$

sujeto a

$$\sum_{i=1}^n p_i = 1 \tag{6.2.16}$$

$$\sum_{i=1}^n p_i g_r(x_i) = \sum_{i=1}^n p_i g_{ri} = a_r, \quad r = 1, 2, \dots, m; \tag{6.2.17}$$

y

$$p_i \geq 0 \quad \forall i$$

Definición 8. Redes libres de escala (Scale-free networks) En general si los vértices de una red son distingibles, para cada vértice se define $p(k, s, N)$ como la probabilidad que el vértice s en la red de tamaño N tenga k conexiones (los k vecinos más cercanos).

Cuando la distribución es de la forma

$$P(k) \propto k^{-\gamma}, k \neq 0 \tag{6.2.18}$$

donde γ es el exponente de la distribución, se denomina que sigue una distribución de ley de potencia, la cual no tiene una escala natural y por eso estas redes se les denomina redes libre de escala (Scale-free networks).

Definición 9. Matriz de adyacencia. También conocida como sociomatriz, X , de tamaño $g \times g$ (g filas y g columnas) para redes de modo 1 (Un solo tipo de nodos). Existe una fila y una columna para cada nodo con etiquetas $1, 2, \dots, g$. Las entradas en la sociomatriz, x_{ij} registran cuales nodos son adyacentes. Para el caso de un digrafo (grafo dirigido) tiene elementos x_{ij} igual a 1 sí hay un arco del nodo fila n_i a la columna nodo n_j y 0 de lo contrario y la matriz no es simétrica. (Wasserman and Faust, 1994)

Definición 10. Centralidad de grado

Sí definimos $C_D(p_k) = \sum_{i=1}^n a(p_i, p_k)$ donde $a(p_i, p_k) = \begin{cases} 1 & \Leftrightarrow \text{están conectados por un borde} \\ 0 & \text{dlc} \end{cases}$

entonces $C_D(p_k)$ es el conteo del grado o número de adyacencias del punto p_k . C_D es grande sí p_k es adyacente a muchos puntos y C_D es pequeño sí está alejado de muchos puntos.

Cómo $C_D \propto n$ entonces para estandarizar la medida y poder comparar con otros grafos se debe relacionar cada C_D contra el máximo posible que en este caso es $n - 1$

Entonces

$$C'_D(p_k) = \frac{\sum_{i=1}^n a(p_i, p_k)}{n - 1} \tag{6.2.19}$$

es la proporción de otros puntos que son adyacentes a p_k . (Freeman, 1977, 1979)

Definición 11. Centralidad de intermediación

La máxima intermediación se presenta cuando existen $n - 1$ bordes y existe un punto p_k que está en todas las geodésicas de distancia > 1 . Esta figura es una estrella o rueda y define el límite superior de $C(p_k)$:

$$\max C_B(p_k) = \frac{n^2 - 3n + 2}{2}$$

Luego la medida estandarizada de centralidad de intermediación es: (Freeman, 1977, 1979)

$$C'_B(p_k) = \frac{C_B(p_k)}{\max C_B(p_k)} = \frac{2 \sum^n \sum^n_{i < j} b_{ij}(p_k)}{n^2 - 3n + 2} \tag{6.2.20}$$

Definición 12. Centralidad de cercanía

Si se define $d(p_i, p_k)$ = número de bordes en la geodésica que une a p_i con p_k entonces un punto es más central si esta distancia es más corta.

Y se define $C_c(p_k)^{-1} = \sum_{i=1}^n d(p_i, p_k)$ como la centralidad de cercanía del punto p_k . De manera que si $\sum_{i=1}^n d(p_i, p_k) \Delta \uparrow$ entonces $C_c(p_k)^{-1} \Delta \uparrow$

Sin embargo, esta medida nuevamente depende de n . Entonces para estandarizarla se compara contra la máxima $C_c(p_k) = n - 1$

De forma que la centralidad de cercanía es según Freeman (1977, 1979):

$$C'_c(p_k)^{-1} = \frac{\sum_{i=1}^n d(p_i, p_k)}{n - 1} \quad (6.2.21)$$

$$C'_c(p_k) = \frac{n - 1}{\sum_{i=1}^n d(p_i, p_k)} \quad (6.2.22)$$

Nomenclatura

adaptación: Es la respuesta relativamente exitosa por parte de un sistema a las influencias externas a las cuales está sometido

agentes: Un agente es un elemento de un sistema que interactúa con otros por medio de vínculos y cuyo comportamiento está definido por una regla de decisión.

autoridades: Son nodos que contienen información valiosa sobre un tema de interés

autómatas celulares: Los autómatas celulares son modelos matemáticos para sistemas naturales complejos que contienen gran cantidad de componentes sencillos e idénticos con interacciones locales. Ellos consisten en una rejilla de lugares, cada uno con un conjunto finito de posibles valores. Los valores de los lugares evolucionan de manera sincrónica en períodos discretos de acuerdo a unas reglas idénticas. El valor de un determinado lugar está determinado por los valores previos de una vecindad de lugares alrededor de este.

BA model: Modelo de redes propuesto por Barabasi y Albert para explicar el mecanismo de vinculación preferencial

centralidad de intermediación: La centralidad de intermediación de un nodo corresponde a determinar que tanto está ubicado en la distancia geodésica de cualquier otro par de nodos y por lo tanto, se convierte en punto de pasaje obligado para conectarlos.

cienciometría: La ciencia de la ciencia se encarga de las características cuantitativas de la ciencia y la investigación científica. El énfasis se ubica en investigaciones en las cuales el desarrollo y mecanismos de la ciencia son estudiados mediante métodos estadísticos y matemáticos.

CIVETS: Nombre asignado por la Unidad de Inteligencia de la revista The Economist, EIU al conjunto de países integrado por Colombia, Indonesia, Vietnam, Egipto, Turquía y Suráfrica

complejidad: Afirma que algunos sistemas exhiben fenómenos que son completamente inexplicables por cualquiera de los análisis convencionales de las partes que constituyen el sistema. Estos fenómenos, poseen comportamientos no lineales, exhiben comportamiento emergente y auto-organización y parece que ocurren en muchos sistemas complejos incluyendo tanto organismos vivos, el mercado de acciones o aún el cerebro humano. Fuente John L. Casti, Encyclopedia Britannica

concentradores: Son nodos que indican donde se encuentran las mejores autoridades de una red

configuraciones: Formas de red

curva S de adopción de la innovación: Registra en función del tiempo el número de adoptantes acumulados de la innovación

curvas FEED: Estas curvas, conocidas en inglés como Fulfilled-Expectations Equilibrium Demand (FEED curves) expresan el precio de un bien o servicio con externalidades de red en función del tamaño de la red (cantidad de nodos)

destrucción creativa: La economía habitualmente se encuentra en una situación de equilibrio, en un mundo sin incertidumbre ni beneficios, pero también sin desarrollo. Sólo periódicamente una innovación rompe ese equilibrio y conduce a la economía a una etapa de desarrollo

economía basada en el conocimiento: Expresión acuñada para describir la tendencia en las economías avanzadas hacia mayores dependencias en el conocimiento, la información y altos niveles de capacidades, así como la necesidad del acceso a éstos por parte de las empresas privadas y el sector público.

economía de las redes: Área de investigación que busca respuestas a dos preguntas:

(i) ¿Cuáles son los efectos económicos del patrón de conexiones que se dan en una sociedad?

(ii) ¿Cuál es la arquitectura de las conexiones que se dan cuando entidades individuales forman vínculos entre sí basadas en consideraciones de costos personales y beneficios?

enrolamiento: Concepto del ANT (ver más adelante) que consiste en un conjunto de estrategias mediante las cuales los investigadores buscan definir e interrelacionar los diversos papeles asignados a los otros.

entropía: Es una medida de la incertidumbre de una variable aleatoria

externalidades de red: Una externalidad de red ocurre cuando los consumidores valoran más un bien a medida que se incrementa el número de usuarios compatibles

homofilia: Es la tendencia a interactuar selectivamente con otros actores que son similares culturalmente

innovación tecnológica: Innovación referida a productos y procesos

innovación: Es la introducción [al mercado] de un nuevo o significativamente mejorado producto (bien o servicio), de un proceso, de un nuevo método de comercialización o método organizativo, de las prácticas internas de la empresa, la organización del lugar de trabajo o las relaciones externas.

interesamiento: Concepto del ANT (Ver más adelante) consistente en una serie de procesos mediante los cuales los investigadores buscan asegurar a los otros actores en los papeles que han sido propuestos para ellos en el programa.

lazos débiles: Relaciones establecidas con personas que se consideran simplemente como conocidos

lazos fuertes: Relaciones establecidas con los amigos muy cercanos

metared: Superposición de redes con diferentes tipos de nodos: Agentes, organizaciones, fechas, lugares, roles, conocimiento, tareas

modelo triple hélice del conocimiento: Metáfora inspirada en la estructura de hélice del ADN para explicar la evolución, e interacción entre empresa-universidad y estado como elementos portadores de la innovación

modelos de eventos históricos: Estos modelos se activan en un instante inicial conocido cuando un actor está en riesgo de experimentar un evento y se detienen cuando el evento ocurre o la observación finaliza

movilización: Concepto del ANT (Ver más adelante) que contempla un conjunto de métodos usado por los investigadores para asegurar que el supuesto vocero de diversas colectividades relevantes esté apropiadamente capacitado para representar aquellas colectividades y no sea traicionado por la colectividad

patente: Es un título de propiedad otorgado por el gobierno de un país, que da a su titular el derecho a impedir por un tiempo determinado a otros la fabricación, venta y/o utilización comercial de la invención protegida

problematización: Concepto de ANT (Ver más adelante) en donde los investigadores buscan volverse indispensables a otros actores, al definir la naturaleza de los problemas y los problemas de los primeros y sugiriendo que éstos podrán ser resueltos si los actores negocian un punto de pasaje obligatorio por el programa de investigación de los investigadores.

red libre de escala: Modelo de redes propuesto que exhibe comportamiento de leyes de potencia

redes complejas: Redes que están detrás de los sistemas complejos, las cual define las interacciones entre los componentes.

redes de innovadores: La organización en red es un arreglo institucional esencial para estar al frente de la innovación sistémica. Las redes pueden ser vistas como una forma compenetrada del mercado y la organización. Empíricamente , ellas son organizaciones pobremente acopladas que tienen un centro con lazos fuertes y débiles entre sus miembros constituyentes. . . . Nosotros enfatizamos la importancia de las relaciones cooperativas entre las empresas como un mecanismo clave de vinculación de las configuraciones de red. Estos incluyen joint ventures, acuerdos de licenciamiento, contratos de administración, subcontratación, producción compartida y colaboración en la IyD

redes formales e informales de innovadores: Son redes basadas en relaciones contractuales formales, como las relaciones de subcontratación, las alianzas estratégicas, la participación en un consorcio industrial amplio de investigación y también en relaciones informales, basadas en la membresía a una asociación profesional o comercial, o aún una asociación más flexible a una comunidad tecnológica

redes tecno-económicas: Red Tecno-Económica (RTE) es un conjunto coordinado de actores heterogéneos: laboratorios públicos, centros de investigación técnica, empresas, organismos financieros, usuarios y poderes públicos que participan colectivamente en la concepción, elaboración, producción y distribución de procedimientos de producción, bienes y servicios, algunos de los cuales dan lugar a una transacción mercantil

sistemas nacionales de innovación: Conjunto de instituciones que interactúan con relación a la innovación

SNCTi: Corresponde al nombre del sistema nacional de ciencia y tecnología colombiano que le agrega el tema de innovación a la sigla genérica

Socio-Física: Debe decirse con fuerza que el objetivo no es lograr una mirada desde la física a todos los aspectos de la vida humana. Como primer paso esencial, se busca clarificar la existencia de leyes cuantitativas que gobiernan el comportamiento humano. En un segundo paso, aún por concretarse, los sociofísicos pueden convertirse a una ciencia cuantitativa para descubrir las leyes actuales del comportamiento social en la misma forma que la física lo ha hecho con las leyes que rigen la materia inerte. Esto no significa que las leyes sean idénticas. Los sociofísicos están convencidos que la comprensión de las leyes del comportamiento humano pueden traer un gran beneficio a la vida humana y a la humanidad.

sociología de la traducción: En los intentos de los investigadores por hacerse aceptar ellos mismos y sus definiciones de las situaciones de otros, cuatro momentos se producen (a) problematización; (b) interesamiento; (c) enrolamiento; (d) movilización. Ver ANT

teoría del actor-red (ANT): ANT, también conocida como sociología de la traducción. Su propósito es mostrar como se construye la sociedad y provee herramientas para analizar el proceso. Uno de sus supuestos centrales es que lo usualmente llamado sociedad es un logro sobre la marcha. ANT es un intento por proveer herramientas analíticas para explicar el proceso por medio del cual la sociedad es reconfigurada permanentemente. Lo que distingue este enfoque del constructivismo es la explicación de la sociedad en construcción, en la cual la ciencia y la tecnología juegan un papel clave.

teoría evolucionista: La explicación del desempeño económico agregado en la economía evolucionista yace en dos fuerzas: la selección y la generación de novedad. En el tiempo, la variedad presente en el sistema es reducida por la selección. La novedad es constantemente agregada al sistema. La innovación es un importante proceso generador de novedad y el mercado y otras instituciones económicas están entre los más importantes mecanismos de selección en las economías modernas.

vigilancia tecnológica: Es el proceso de búsqueda, obtención, análisis y empleo de la información sobre desarrollos y tendencias en el ámbito científico y tecnológico, que es de alto valor para la competitividad de la empresa y, por tanto, es útil para la toma de decisiones estratégicas. Este proceso orienta los esfuerzos para adquirir, desarrollar, explotar y retirar tecnología en las empresas.

vinculación preferencial: Mecanismo de agregación de nuevos nodos a aquellos que tienen mayor relaciones establecidas

Bibliografía

- Acemoglu, D., Bimpikis, K., and Ozdaglar, A. (2011). Experimentation, Patents, and Innovation. *American Economic Journal: Microeconomics*, 3(1):37–77.
- Aguirre R, J., Robledo Velasquez, J., and Perez P, A. (2009). Metodología para medir y evaluar las capacidades tecnológicas de innovación aplicando sistemas de lógica difusa, caso fábricas de software. Cartagena (Colombia).
- Albuquerque, P., Bronnenberg, B. J., and Corbett, C. J. (2007). A Spatiotemporal Analysis of the Global Diffusion of ISO 9000 and ISO 14000 Certification. *Management Science*, 53(3):451–468.
- Allarakhia, M., Walsh, S., and Wensley, A. (2007). Models of Cooperation and Knowledge Management: The Case of Biomedical Technology Management. *Management of Engineering and Technology, Portland International Center for*, pages 438–447.
- Anand, K. and Bianconi, G. (2009). Entropy measures for networks: Toward an information theory of complex topologies. *Physical Review E*, 80(4):045102.
- Anand, K., Bianconi, G., and Severini, S. (2011). Shannon and von Neumann entropy of random networks with heterogeneous expected degree. *Physical Review E*, 83(3):036109.
- Andrew, J. and Sirkin, H. L. (2006). *Payback : reaping the rewards of innovation*. Harvard Business School Press, Boston Mass.
- Ashton, W. and Stacey, G. (1995). Technical intelligence in business: understanding technology threats and opportunities. *International Journal of Technology Management*, 10(1):79–104.
- Atallah, G. and Rodriguez, G. (2006). Indirect patent citations. *Scientometrics*, 67(3):437–465.
- Auerback, F. (1913). Das Gesetz der Bevölkerungskonzentration. *Geographischer Anstalt*, 59:74–76.
- Bak, P. (1997). *How nature works*. Oxford University Press.
- Bar-Hillel, Y. (1955). An examination of information theory. *Philosophy of Science*, pages 86–105.
- Barabasi, A.-L. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512.
- Barabasi, A.-L. (2002). *Linked : the new science of networks*. Perseus Pub., Cambridge Mass.
- Barabasi, A.-L., Albert, R., and Jeong, H. (1999). Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1):173–187.
- Barabasi, A.-L. and Bonabeau, E. (2003). Scale-Free. *Scientific American*.
- Bass, F. M. (1969). A New Product Growth for Model Consumer Durables. *Management Science*, 15(5):215–227.
- Bass, F. M. (2004). JSTOR: Management Science, Vol. 50, No. 12 (Dec., 2004), pp. 1833-1840. 50(12):1833–1840.
- Bianconi, G., Coolen, A. C., and Vicente, C. J. P. (2008). Entropies of complex networks with hierarchically constrained topologies. *Physical Review E*, 78(1):016114.

- Boccaro, N. (1997). A new class of automata networks. *Physica D: Nonlinear Phenomena*, 103(1-4):145–154.
- Boccaro, N. and Fuks, H. (2008). Modeling diffusion of innovations with probabilistic cellular automata. *1-17*.
- Bohorquez, J. C., Gourley, S., Dixon, A. R., Spagat, M., and Johnson, N. F. (2009). Common ecology quantifies human insurgency. *Nature*, 462(7275):911–914.
- Boltzmann, L. (1872). Further studies about the heat equilibrium among gas molecules. *Sitzungsberichte der Akademie der Wissenschaften Wien*, II(66):275–370.
- Bommarito II, M. J., Katz, D. M., Zelner, J. L., and Fowler, J. H. (2010). Distance measures for dynamic citation networks. *Physica A: Statistical Mechanics and its Applications*, 389(19):4201–4208.
- Borgatti, S. P., Everett, M., and Freeman, L. (2002). UCINET 6 For Windows: Software for Social Network Analysis.
- Brodka, P., Saganowski, S., and Kazienko, P. (2012). GED: the method for group evolution discovery in social networks. *Social Network Analysis and Mining*.
- Buitrago, R. (2012). Sugerencias para avanzar en un estudio sobre patentes colombianas.
- Buitrago, R. (2014). Importancia patentes colombianas.
- Burt, R. (1992). *Structural holes : the social structure of competition*. Harvard University Press, Cambridge Mass.
- Butts, C. T. (2008). A relational event framework for social action. *Social Metodology*, 38:155–200.
- Cabral, L. (1990). On the adoption of innovations with network externalities. *Mathematical Social Sciences*, 19(3):299–308.
- Cabral, L. (2000). *Introduction to industrial organization*. MIT Press, Cambridge Mass.
- Callaert, J., Van Looy, B., Verbeek, A., Debackere, K., and Thijs, B. (2006). Traces of prior art: An analysis of non-patent references found in patent documents. *Scientometrics*, 69(1):3–20.
- Callon, M. (1986). Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. *Power, action and belief: A new sociology of knowledge*, 32:196 a 233.
- Callon, M. (1991). Techno-economic networks and irreversibility. In *A sociology of monsters ? Essays on Power, Technology and Dominations*. J.Law Ed.
- Callon, M. (1995). Algunos elementos para una sociología de la traducción. La domesticación de las vieiras y los pescadores de la bahía de St. Brieuc. In *Sociología de la ciencia y la tecnología*, pages 259–282.
- Callon, M. (2001a). Actor-Network Theory.
- Callon, M. (2001b). Redes tecno-economicas e irreversibilidad. *Redes Revista de Estudios Sociales de la Ciencia*, 8(17):85–126.
- Callon, M. (2007). An Essay on the Growing Contribution of Economic Markets to the Proliferation of the Social. *Theory, Culture & Society*, 24((7-8)):139–163.
- Carley, K. M., Reminga, J., Storrick, J., and De Reno, M. (2009). ORA user 's Guide 2009.
- Castellanos, O., Fuquene, A., and Ramirez, D. (2011). *Análisis de tendencias: de la información hacia la innovación*. Universidad Nacional de Colombia, Bogotá D.C. Colombia.
- Charum, J. (2009). Sobre las nociones de dato, información y conocimiento. In *Encuestas, datos y descubrimiento de conocimiento sobre la innovación en Colombia*, page 366. Fundación Cultural Javeriana de Artes Gráficas, Bogotá D.C. Colombia.

- Chen, D.-Z., Huang, M.-H., Hsieh, H.-C., and Lin, C.-P. (2011). Identifying missing relevant patent citation links by using bibliographic coupling in LED illuminating technology. *Journal of Informetrics*, 5(3):400–412.
- Chen, Y.-S. and Chang, K.-C. (2009). Using neural network to analyze the influence of the patent performance upon the market value of the US pharmaceutical companies. *Scientometrics*, 80(3):637–655.
- Chesbrough, H., Vanhaverbeke, W., and West, J. (2006). *Open innovation : Researching a new paradigm*. Oxford University Press, New York.
- Christakis, N. (2009). *Connected : the surprising power of our social networks and how they shape our lives*. Little Brown and Co., New York, 1st ed. edition.
- Clauset, A., Newman, M., and Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70(6).
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM*.
- Clausius, R. (1865). On different forms of the fundamental equations of the mechanical theory of heat and their convenience for application. *Poggendorffs Annalen der Physik*, 125.
- Courtial, J. P., Callon, M., and Sigogneau, A. (1993). The use of patent titles for identifying the topics of invention and forecasting trends. *Scientometrics*, 26(2):231–242.
- Cover, T. (2006). *Elements of information theory*. Wiley-Interscience, Hoboken N.J., 2nd ed. edition.
- Cozzens, S., Gatchair, S., Kang, J., Kim, K.-S., Lee, H. J., Ordóñez, G., and Porter, A. (2010). Emerging technologies: quantitative identification and measurement. *Technology Analysis & Strategic Management*, 22(3):361–376.
- de Nooy, W. (2011). Networks of action and events over time. A multilevel discrete-time event history model for longitudinal network data. *Social Networks*, 33(1):31–40.
- Delre, S. A., Jager, W., and Janssen, M. A. (2006). Diffusion dynamics in small-world networks with heterogeneous consumers. *Computational and Mathematical Organization Theory*, 13(2):185–202.
- DNP (2007). *Plan Nacional de Desarrollo 2006 - 2010*.
- DNP (2008). Bases de un plan de acción para la adecuación del sistema de propiedad intelectual a la competitividad y productividad nacional 2008-2010.
- DNP (2011). *Plan Nacional de Desarrollo 2010 - 2014*.
- Dorogovtsev, S. (2003). *Evolution of networks : from biological nets to the Internet and WWW*. Oxford University Press, Oxford.
- Drucker, P. (2002). The discipline of innovation. *Harvard Business Review*.
- Duque, L., Manrique, J., and Robledo Velasquez, J. (2009). Descubrimiento de conocimiento a partir de la encuesta anual manufacturera colombiana empleando inteligencia artificial. Cartagena (Colombia).
- Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization*, 14(6):673–699.
- Erdos, P. and Renyi, A. (1960). *On the evolution of random graphs*. Citeseer.
- Etzkowitz, H. (2000). The dynamics of innovation: from National Systems and Mode 2 to a Triple Helix of university -industry - government relations. *Research Policy*, 29(2):109–123.
- Felice, D., Mancini, S., and Pettini, M. (2013). Quantifying Networks Complexity from Information Geometry Viewpoint. *arXiv preprint arXiv:1310.7825*.
- Freeman, C. (1991). Networks of innovators: A synthesis of research issues. *Research Policy*, 20(5):499–514.

- Freeman, L. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40:35–41.
- Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239.
- Galam, S. (2004). Sociophysics: a personal testimony. *Physica A: Statistical and Theoretical Physics*, 336(1-2):49–55.
- Galam, S. (2008). SOCIOPHYSICS: A REVIEW OF GALAM MODELS. *International Journal of Modern Physics C*, 19(03):409.
- Galam, S. (2012). The Question: Do Humans Behave like Atoms? In *Sociophysics*, pages 21–39. Springer US, Boston, MA.
- Galam, S., Gefen (Feigenblat), Y., and Shapir, Y. (1982). Sociophysics: A new approach of sociological collective behaviour. A mean behaviour description of a strike. *The Journal of Mathematical Sociology*, 9(1):1–13.
- Gao, X. and Guan, J. (2009). Networks of scientific journals: An exploration of Chinese patent data. *Scientometrics*, 80(1):283–302.
- Ge, Y., Song, Z., Qiu, X., Song, H., and Wang, Y. (2013). Modular and hierarchical structure of social contact networks. *Physica A: Statistical Mechanics and its Applications*.
- Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.
- Goldenberg, J., Libai, B., and Muller, E. (2010). The chilling effects of network externalities. *International Journal of Research in Marketing*, 27(1):4–15.
- Goyal, S. (2009). *Connections : an introduction to the network economy*. Princeton University Press, Princeton N.J. ; Woodstock.
- Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443.
- Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1:201.
- Grupp, H. (1990). The concept of entropy in scientometrics and innovation research. *Scientometrics*, 18(3):219–239.
- Guan, J. and He, Y. (2007). Patent-bibliometric analysis on the Chinese science-technology linkages. *Scientometrics*, 72(3):403–425.
- Guerra, F., Trujillo, A., Mejia, J., Munoz, R., and Perez, C. (2011). La heterogeneidad en las fuentes de innovacion: Un estudio de 4 casos del premio innova 2008. In *Redes de innovacion en empresas colombianas*, pages 87–100. Bogota.
- Guseo, R. and Guidolin, M. (2009). Modelling a dynamic market potential: A class of automata networks for diffusion of innovations. *Technological Forecasting and Social Change*, 76(6):806–820.
- Harary, F. (1969). *Graph theory*. Perseus Books, Reading, Mass.
- Harris, M., Halkett, R., and Cox, J. (2007). Hidden Innovation. Technical report, NESTA.
- He, J. and Fallah, M. H. (2011). The typology of technology clusters and its evolution - Evidence from the hi-tech industries. *Technological Forecasting and Social Change*.
- Hernandez Umana, I. D. (2008). *Empresa, innovacion y desarrollo*. Universidad Nacional de Colombia. Facultad de Ciencias Economicas, Bogota.
- Hopgood, A. (2000). *Intelligent Systems for Engineers and Scientist*. 2 edition.

- Hurtado, R. and Mejia, J. (2011). Estructura de la Inversion en Actividades de innovacion y Desarrollo Tecnologico de la Industria Manufacturera Colombiana y la Posicion de la Ingenieria y el Diseno Industrial. Lima - Peru.
- Hurtado, R. and Mejia, J. (2014). Estructura de la inversion de la industria manufacturera colombiana en actividades de innovacion y desarrollo tecnologico. *Revista Innovar Journal*, 24:33–40.
- Imai, K. and Baba, Y. (1989). Systemic Innovation and Cross-Border Networks: Transcending Markets and Hierarchies to create a new techno-economic system. Paris (Fr).
- Jaffe, A. and Trajtenberg, M. (2002). *Patents, citations, and innovations : a window on the knowledge economy*. MIT Press, Cambridge Mass.
- Jaynes, E. (1957). Information theory and statistical mechaniscs. *Physical Review*, 106(4).
- Jaynes, E. (1978). Where do we stand on maximun entropy. MIT.
- Jimenez, C. and Castellanos, O. (2009). Exploring the use of biological metaphor upon Technology Management research within the new paradigm of ongoing change. In *Management of Engineering & Technology, 2009. PICMET 2009. Portland International Conference on*, page 37 a 46, Portland.
- Kapur, J. (1992). *Entropy optimization principles with applications*. Academic Press, Boston.
- Kelley, H. J. (1969). Entropy of knowledge. *Philosophy of science*, 36(2):178–196.
- Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632.
- KonstanzUniversity (2001). Visone.
- Kossinets, G. and Watts, D. (2006a). Supporting Online Material for Empirical Analysis of an Evolving Social Network.
- Kossinets, G. and Watts, D. J. (2006b). Empirical Analysis of an Evolving Social Network. *Science*, 311(5757):88–90.
- Krackhardt, D. and Carley, K. M. (1998). A PCANS model of structure in organization. In *Proceedings of the 1998 International Symposium on Command and Control Research and Technology*, volume 113, page 119.
- Krebs, V. E. (2002). Mapping networks of terrorist cells. *Connections*, 24(3):43–52.
- Kullback, S. and Leibler, R. (1951). On information and sufficiency. *Ann.Math.Stat*, 22:79–86.
- Lee, S., Kim, H., and Jung, K. (2011). Analysis of tipping points in social networks for diffusion of innovations. MIT.
- Leicht, E. A., Clarkson, G., Shedden, K., and Newman, M. E. (2007). Large-scale structure of time evolving citation networks. *The European Physical Journal B*, 59(1):75–83.
- Leydesdorff, L. (1991). The static and dynamic analysis of network data using information theory. *Social networks*, 13(4):301–345.
- Leydesdorff, L. (2006). *The knowledge based economy : modeled, measured, simulated*. Universal publishers, Boca Raton Florida.
- Leydesdorff, L. (2012). Statistics for the Dynamic Analysis of Scientometric Data: The evolution of the sciences in terms of trajectories and regimes. *Scientometrics*, pages 1–11.
- Leydesdorff, L. and Meyer, M. (2010). The Decline of University Patenting and the End of the Bayh-Dole Effect. *Scientometrics*.

- Lilien, G. L. (2003). *Marketing engineering: computer-assisted marketing analysis and planning*. Prentice Hall, Upper Saddle River, N.J, 2nd ed edition.
- Lundvall, B.-A., editor (1992). *National Systems of Innovation - Towards a Theory of innovation and Interactive Learning*. Printer Publishers, London ;New York.
- Lundvall, B. A. and Vinding, A. L. (2004). Product Innovation and Economic Theory - User - Producer Interactions in the Learning Economy. In *Research on Technological Innovation, Management and Policy*, volume 8, pages 101–128. Emerald (MCB UP), Bingley.
- Mahajan, V. and Kerin, R. A. (1984). Introduction Strategy for New Products with Positive and Negative Word of Mouth. *Management Science*, 30(12):1389–1404.
- Malaver, F. and Vargas, M. (2007). *Vigilancia tecnologica y competitividad sectorial: Lecciones y resultados de cinco estudios*. Fundacion Cultural Javeriana de Artes Graficas.
- Maloney, W. (2008). La innovacion colombiana en contexto internacional, Colciencias, memorias Bogota 14 y 15 de Abril de 2008.
- Mathworks (2014). MatLab.
- Mayor Mora, A., Jimenez Gomez, S. I., and Cardona Bueno, H. (2005). *Inventos y patentes en Colombia, 1930-2000 : de los limites de las herramientas a las fronteras del conocimiento*. I.T.M., Medellin, Colombia.
- Mejia, J. and Rodriguez, O. (2012). De la esperanza a la decepcion: Un caso de patentamiento en una mediana universidad colombiana.
- Meyer, M. (2006). Knowledge integrators or weak links? An exploratory comparison of patenting researchers with their non-inventing peers in nano-science and technology. *Scientometrics*, 68(3):545–560.
- Milgram, S. (1969). An Experimental of the Small World Problem. *Sociometry*, 32(4):425–443.
- Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet mathematics*, 1(2):226–251.
- Moldovan, S. (2004). Cellular automata modeling of resistance to innovations: Effects and solutions. *Technological Forecasting and Social Change*, 71(5):425–442.
- Molinares, A. (1912). Amusement device.
- Montes, F., Sarmiento, O. L., Zarama, R., Pratt, M., Wang, G., Jacoby, E., Schmid, T. L., Ramos, M., Ruiz, O., Vargas, O., Michel, G., Zieff, S. G., Valdivia, J. A., Cavill, N., and Kahlmeier, S. (2011). Do Health Benefits Outweigh the Costs of Mass Recreational Programs? An Economic Analysis of Four Ciclovía Programs. *Journal of Urban Health*, 89(1):153–170.
- Moreno, J. L. and Jennings, H. H. (1938). Statistics of Social Configurations. *Sociometry*, 1(3/4):342.
- Mrvar, A. and Batagelj, V. (2013). *Pajek and Pajek-XXL Programs for Analysis and Visualization of Very Large Networks*.
- Muller, I. (2007). *A history of thermodynamics the doctrine of energy and entropy*. Springer, Berlin; London.
- Nelson, R. R. and Winter, S. G. (1982). *An evolutionary theory of economic change*. Belknap Press of Harvard University Press, Cambridge, Mass.
- Newman, M. (2006). *The structure and dynamics of networks*. Princeton University Press, Princeton N.J. ;;Oxford.
- Newman, M. (2010). *Networks : an introduction*. Oxford University Press, Oxford ;New York.
- Nicolis, G. (1989). Physics of far from equilibrium systems and self organisation. In Davids, P., editor, *The new physics*. Cambridge University Press.

- O' Neale, D. R. and Hendy, S. C. (2012). Power Law Distributions of Patents as Indicators of Innovation. *PloS one*, 7(12):e49501.
- Oakland_University (2015). Erdos Project Number.
- OECD (2005). *Oslo manual : guidelines for collecting and interpreting innovation data*. Organisation for Economic Co-operation and Development ;Statistical Office of the European Communities, Paris, 3rd ed. edition.
- OECD (2006). *Manual de Oslo : guía para la recogida e interpretacion de datos sobre innovacion : la medida de las actividades científicas y tecnológicas*. Grupo Tragsa, Madrid, 3a. ed. edition.
- Oster, S. (2000). *Analisis moderno de la competitividad*. Oxford University Press, Mexico.
- Otalora, D., Hurtado, R., and Quimbay, C. (2009). Un Modelo simple para capturar la Dinamica del Sistema Empresas-Fuentes de Financiacion:El caso del Conocimiento de la Existencia de Fuentes de Financiacion para Innovacion y Desarrollo Tecnologico en el Sector Industrial Manufacturero Colombiano. Cartagena (Colombia).
- Owen-Smith, J. and Powell, W. W. (2004). Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. *Organization Science*, 15(1):5–21.
- Peres, R., Muller, E., and Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, 27(2):91–106.
- Perez, C. (2010). Algunas nociones sobre el analisis de redes sociales. In *Redes*, page 77. Bogota.
- Polanco, X., Francois, C., and Lamirel, J.-C. (2001). Using artificial neural networks for mapping of scienceand technology: A multi-self-organizing-maps approach. *Scientometrics*, 51(1):267–292.
- Powell, W. and Grodal, S. (2005). Networks of Innovators. In *The Oxford handbook of innovation*. Oxford University Press, Oxford; New York.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1):81–106.
- Rahmandad, H. and Sterman, J. (2008). Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models. *Management Science*, 54(5):998–1014.
- Ravasz, E. and Barabasi, A.-L. (2003). Hierarchical organization in complex networks. *Physical Review E*, 67(2).
- Renyi, A. (1961). On measures of entropy and information. In *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 547–561.
- Restrepo, L. P. and Vasquez, U. E. (2004). Actualidad y sentido de la universidad frente a un mundo enajenado: entrevista a Guillermo Paramo. *Nomadas*, (21):205–216.
- Reuters, T. (2012). Derwent World Patents Index.
- Rivera, J. (2009). *Una propuesta de innovacion del sistema de registro de marcas y patentes en Colombia*. PhD thesis, Universidad Nacional De Colombia - Sede Medellin., Medellin, Colombia.
- Robledo Velasquez, J. (2007). De los grupos consolidados de investigacion a los sistemas dinamicos de innovacion: el desafio actual del desarrollo científico y tecnológico colombiano. *Dyna*, 152:1 – 7.
- Robledo Velasquez, J. (2010). *Introduccion a la gestion tecnologica*.
- Robledo Velasquez, J. and Ceballos, Y. (2008). Estudio de un proceso de innovacion utilizando la dinamica de sistemas. *Cuadernos de Administracion*, 21(35):127 – 159.
- Robledo Velasquez, J., Malaver, F., and Vargas, M. (2009). *Encuestas, datos y descubrimiento de conocimiento sobre la innovacion en Colombia*. Universidad Nacional de Colombia, Bogota.

- Rodriguez Devis, J. M. (2006). La innovacion tecnologica: mas que un proyecto exitoso. A proposito del premio nacional de innovacion. *Ingenieria e Investigacion*, 26(2):84–93.
- Roehner, B. M. (2007). *Driving forces in physical, biological and socio-economic phenomena a network science investigation of social bonds and interactions*. Cambridge University Press, Cambridge; New York.
- Rogers, E. (1962). *Diffusion of innovations*. Free Press, New York, 5th ed. edition.
- Ronald, E. and Sipper, M. (2000). Engineering, Emergent Engineering and Artificial Life: Unsurprise, Unsurprising Surprise, and Surprising Surprise. In Bedau, M., Caskill, J., Packard, N., and Rasmussen, S., editors, *Artificial Life VII*, pages 523 – 528.
- Sanchez, J. M., Medina, J. E., and Leon, A. M. (2007). Publicacion internacional de patentes por organizaciones inventores de origen colombiano. *Cuadernos de Economia*, 26(47):247–270.
- Sanjay, J. and Sandeep, K. (2006). Can we recognize an innovation? Perspective form an evolving network model. In Chakrabarti, B., Ckarakoti, A., and Chatterjee, A., editors, *Econophysics and sociophysics : trends and perspectives*. Wiley-VCH, Weinheim.
- Schumpeter, J. (1911). *The theory of economic development : an inquiry into profits, capital, credit, interest, and the business cycle*. Transaction Books, New Brunswick N.J.
- SeminairePolibius (2007). Defining a research question.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst.Tech.Journal*, 27(1):379–423.
- Sheridan, P., Yagahara, Y., and Shimodaira, H. (2012). Measuring preferential attachment in growing networks with missing-timelines using Markov chain Monte Carlo. *Physica A: Statistical Mechanics and its Applications*, 391(20):5031–5040.
- Shin, J., Lee, W., and Park, Y. (2006). On the benchmarking method of patent-based knowledge flow structure: Comparison of Korea and Taiwan with USA. *Scientometrics*, 69(3):551–574.
- SIC, S. I. y. C. (2008). Guia de propiedad intelectual.
- SIC, S. I. y. C., Rosario, U., Externado, U., and Arboleda, U. S. (2012). *Propiedad Industrial 2020*. Bogota.
- Silva, L. (1908). Bicycle.
- Stepanic, J., Sabol, G., and Åœebec, M. S. (2005). Describing social systems using social free energy and social entropy. *Kybernetes*, 34(6):857–868.
- Strogatz, S. H. (2003). *Sync : the emerging science of spontaneous order*. Hyperion, New York.
- Ten Kate, A. and Niels, G. (2006). Fulfilled-expectations Equilibria in Network Demand. *Review of Network Economics*, 5(3).
- Torres, L. (2007). *Complejidad*. Bogota.
- Torres, L., editor (2010). *Memorias*. Bogota.
- Uribe G, C., Fonseca R, S., Bernal R, G., Contreras P, C., and Castellanos, O. (2011). *Sembrando innovacion para la competitividad del sector agropecuario*. Giro Editores Ltda, Bogota D.C. Colombia.
- USPTO (2011). Number of Utility Patent Applications filed in the United States by Country of origin. Technical report, USPTO, USA.
- Valverde, S., Sole, R., Bedau, M., and Packard, N. (2007). Topology and evolution of technology innovation networks. *Physical Review E*, 76(5).
- Wang, W., Fergola, P., Lombardo, S., and Mulone, G. (2006). Mathematical models of innovation diffusion with stage structure. *Applied Mathematical Modelling*, 30(1):129–146.

- Ward, R. (2009). BRICS and BICIS. *The Economist*.
- Wasserman, S. and Faust, K. (1994). *Social network analysis : methods and applications*. Cambridge University Press, Cambridge ;New York.
- Watts, D. J. (2003). *Six degrees the science of a connected age*. Norton, New York.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.
- WEF (2010). *The Global Competitiveness Report 2010 - 2011: Highlights*.
- WEF (2013). *WEF Global Competitiveness Report 2013-2014*.
- Witten, I. (2005). *Data mining : practical machine learning tools and techniques*. Morgan Kaufman, Amsterdam ;;Boston MA, 2nd ed. edition.
- Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1-2):1–35.
- Wu, H.-C., Chen, H.-Y., Lee, K.-Y., and Liu, Y.-C. (2010). A method for assessing patent similarity using direct and indirect citation links. In *Industrial Engineering and Engineering Management (IEEM), 2010 IEEE International Conference on*, pages 149–152. IEEE.
- Wu, J., Lu, R., Jiao, L., Liu, F., Yu, X., Wang, D., and Sun, B. (2012). Phase transition model for community detection. *Physica A: Statistical Mechanics and its Applications*.
- Xu, X., Liu, F., and Liu, L. (2005). Mechanism for linear preferential attachment in growing networks. *Physica A: Statistical Mechanics and its Applications*, 356(2):662–670.
- Yi, Y., Qi, W., and Wu, D. (2013). Are CIVETS the next BRICs? A comparative analysis from scientometrics perspective. *Scientometrics*, 94(2):615–628.
- Zhang, T., Xuan, H., and Gao, B. (2005). Modeling diffusion of innovation with cellular automata. In *Proceedings of ICSSSM '05. 2005 International Conference on Services Systems and Services Management, 2005.*, pages 976–980 Vol. 2, Chongqing, China.
- Zhao, K., Karsai, M., and Bianconi, G. (2011). Entropy of dynamical social networks. *PLoS one*, 6(12):e28116.
- Zhao, K., Karsai, M., and Bianconi, G. (2013). Models, Entropy and Information of Temporal Social Networks. In *Temporal Networks*, pages 95–117. Springer.
- Zipf, G. (1949). *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, Reading, MA.