

UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

FACULTAD 3

Herramienta para el análisis de redes de innovación tecnológica



Método para inferir redes de interacción social a partir de las relaciones recíprocas entre los actores

AUTORA:

Hany Bello Rodríguez

TUTORES:

Ing. Vladimir Milián Núñez

Lic. Raynel Batista Téllez

CO-TUTORA:

MsC. Mailen Edith Escobar

La Habana, junio de 2019
“Año 61 de la Revolución”



“Solo llegarán a la cima, los que en su empeño pongan fé”

Ché

DECLARACIÓN DE AUTORÍA

Declaro ser autora de la presente tesis que tiene por título: Método para inferir redes de interacción social a partir de las relaciones recíprocas entre los actores y reconozco a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo. Para que así conste firmo la presente a los ____ días del mes de _____ del año _____.

Vladimir Milián Núñez

Firma del Tutor

Raynel Batista Téllez

Firma del Tutor

Mailen Edith Escobar Pompa

Firma del Cotutor

Hany Bello Rodríguez

Firma del Autor

DATOS DE CONTACTO

Tutor: Ing. Vladimir Milián Núñez, Ingeniero en Ciencias Informáticas, de la Universidad de las Ciencias Informáticas, 2008. Máster en Ciencias en la Universidad de las Ciencias Informáticas, 2018, Cuba. Profesor Asistente de las disciplinas Física y Matemática Aplicada de la Universidad de las Ciencias Informáticas. Miembro del grupo de Investigación de Inteligencia Artificial y Reconocimiento de Patrones desde el 2008. Líder de proyecto y líder técnico (Arquitecto de software principal) de varios proyectos del centro de desarrollo de Telemática de la UCI en el periodo entre 2008 – 2011. Miembro de la Asociación Cubana de Reconocimiento de Patrones (ACRP) desde el 2013 y presidente en funciones de la delegación de base de la UCI. Editor Asociado de la Revista Cubana de Ciencias Informáticas. Miembro del comité organizador y del comité científico de varios eventos. Autor y co-autor de varias publicaciones en revistas y memorias de eventos. Profesor de un curso de postgrado. Tutor de 24 Tesis de pre-grado para Ingeniería.

Correo electrónico: vmilian@uci.cu Researchgate: [Vladimir Milián Núñez](#) Twitter: [@vmiliann](#)

Tutor: Lic. Raynel Batista Téllez, Sociólogo, editor, Coordinador General de Ediciones Futuro, sello editorial de la Universidad de las Ciencias Informáticas. Profesor Auxiliar. Ha impulsado proyectos interdisciplinarios que vinculan las ciencias sociales y el cálculo computacional, así como las humanidades digitales, en colaboración con actores de América Latina, Europa y Asia. Ha dirigido varios proyectos en temas como sociocibernética, interacción hombre-computadora y social media. Posee varias publicaciones, participaciones en eventos, proyectos, distinciones y premios. Ha asesorado trabajos de diploma, tesis de maestría y doctorado. Es miembro de la Junta Nacional Editorial del Ministerio de Cultura, Organización Internacional de Sociología, Sociedad Cubana de Derecho e Informática, Asociación Cubana de Comunicadores Sociales, Organización Internacional de Editores, Junta Editorial de la Revista Cubana de Ciencias Informáticas. Ha mostrado interés por los nuevos entornos de lectura, las redes académicas de interacción social y la ciencia de los datos. Sus principales resultados se enmarcan en la sociología de la innovación y la antropología digital, con especial interés en la curación de contenidos digitales.

Correo electrónico: rainer@uci.cu Researchgate: [Raynel Batista Téllez](#) Twitter: [@RBatell](#)

Co-tutor: Mailen Edith Escobar Pompa. Ingeniera en Ciencias Informáticas, de la Universidad de las Ciencias Informáticas, 2008. Máster en Gestión de Proyectos Informáticos, 2016, Universidad de las Ciencias Informáticas, Cuba. Profesor Asistente de la disciplina Inteligencia Artificial de la Universidad de las Ciencias Informáticas. Miembro del grupo de Investigación de Gestión de Proyectos Informáticos desde el 2013. Planificadora de varios proyectos de los centros de desarrollos de Gestión Empresarial y Gobierno Electrónico 2008 – 2015. Miembro de la Asociación Cubana de Reconocimiento de

DATOS DE CONTACTO

Patrones (ACRP) desde el 2016. Autora y co-autora de varias publicaciones en revistas y memorias de eventos. Profesora de un curso de postgrado. Tutor de 5 Tesis de pre-grado para Ingeniería.

DEDICATORIA

A Yunelis y Joel, mis padres, mejores amigos y maestros.

A Humberto, el abuelo más bueno del mundo.

A Daniel, por ser el novio más comprensivo

A toda mi familia y amigos.

AGRADECIMIENTOS

Quisiera agradecerle de todo corazón a mi mamá por estar siempre para mí, por ser mi guía, por apoyarme y motivarme, por confiar tanto en mí, por cuidarme siempre, escucharme y quererme, por ser más que una madre. Te agradezco todo lo que soy. Te quiero con todo mi corazón.

A Laffi por ser como un padre para mí, por apoyarme en todas mis decisiones, y ayudarme a convencer a mi mamá cuando no me quiere dar permiso para salir por ahí.

A mi abuelo por ser un amigo, un cómplice, por apoyarme, por quererme tanto, por ser su niña consentida.

A mi novio Daniel por ser mi amigo, mi confidente, mi pañuelo de lágrimas, mi soporte, por darme tanto amor y tanta paciencia te prometo que voy hacer lo mismo contigo el año que viene, gracias por todo Te amo.

A mi amigo Ginarte que es como mi hermano, por estar ahí para mí en todos los momentos, en lo felices y tristes también, hemos discutido, pero a pesar de todo nuestro cariño va por encima de cualquier cosa, tkm mi titi.

A mis tíos, Yunedis, Aldito, Yahima que es la esposa de mi tío Aldito, pero se ha ganado todo mi cariño les quiero agradecer por todo su cariño y amor.

A mi primo Josué por darme tantos momentos lindos, por ser tan travieso, tkm.

También quisiera agradecerle a todos los que han estado ahí para mí como, por ejemplo, Alfredo, Luis Rubén, Ramón, Eric, Yoansel, Mi papa.

Quisiera darles las gracias a todos mis profesores desde primer año por ayudarme a llegar aquí, principalmente a nuestra jefa de año Dariela, a Mailen porque a pesar de ser mi tutora es mi amiga, a el gordo Osvaldo por aguantarme casi todas las noches en su casa hasta tarde, y ser siempre el mismo, a mis tutores por ayudarme de una forma u otra, y al tribunal porque en momento determinado fueron mis profesores que me enseñaron, educaron y ayudaron para estar aquí, gracias a todos.

A todos los que de alguna manera han compartido conmigo esta parte tan especial de mi vida, gracias.

RESUMEN

El Análisis de Redes Sociales es un área de investigación que estudia las redes como grafos para predecir su comportamiento. El estudio de influencia en las redes de interacción social ha demostrado importantes aplicaciones en este sentido. El presente trabajo expone los resultados de un método propuesto para inferir redes de interacción social en un entorno informal a partir de la reciprocidad, considerándose uno de los aspectos que caracteriza el comportamiento de la influencia en una red y demandado por la comunidad científica. La implementación del método propuesto considera aspectos que caracterizan los estudios sobre influencia social para inferir redes de interacción a partir de las relaciones recíprocas entre los nodos en un entorno específico de datos. La aplicación experimental del método en una base de datos bibliográfica obtuvo una red de coautoría capaz de reflejar la influencia que ejercen los autores entre sí a través de relaciones de reciprocidad. El empleo de técnicas sociométricas y pruebas estadísticas para evaluar los resultados del experimento comprobó la correspondencia entre los datos obtenidos con el sociograma y aquellos mediante el método desarrollado demostrando su efectividad y validez. La solución obtenida para la inferencia de redes de interacción social a partir de las relaciones recíprocas entre los actores contribuye sustancialmente al estudio del comportamiento de la influencia social en una red.

Palabras claves: análisis de redes sociales, influencia social, reciprocidad, método.

ABSTRACT

Social Network Analysis is an area of research that studies networks as graphs to predict their behavior. The study of influence in social interaction networks has shown important applications in this regard. This paper presents the results of a proposed method to infer social interaction networks in an informal environment based on reciprocity, considering one of the aspects that characterizes the behavior of influence in a network and demanded by the scientific community. The implementation of the proposed method considers aspects that characterize the studies on social influence to infer interaction networks from the reciprocal relationships between the nodes in a specific data environment. The experimental application of the method in a bibliographic database obtained a co-authorship network capable of reflecting the influence exerted by the authors on each other through reciprocal relationships. The use of sociometric techniques and statistical tests to evaluate the results of the experiment proved the correspondence between the data obtained with the sociogram and those using the developed method demonstrating its effectiveness and validity. The solution obtained for the inference of networks of social interaction from the reciprocal relationships between the actors contributes substantially to the study of the behavior of social influence in a network.

Keywords: analysis of social networks, social influence, reciprocity, method.

INTRODUCCIÓN	1
CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA	6
1.1 Conceptos básicos de redes sociales	6
1.1.1 Teoría de grafos.....	6
1.1.2 Análisis de redes sociales	7
1.1.3 Representación de las redes sociales	8
1.1.4 Clasificación de las redes sociales	9
1.1.5 Inferencia de redes sociales.....	11
1.2 Análisis de Influencia Social	11
1.2.1 Patrones de comportamiento	13
1.2.2 Análisis de la Reciprocidad	15
1.3 Redes de colaboración y coautoría	17
1.4 Herramientas para el análisis de redes sociales.....	20
1.5 Herramientas y tecnologías a utilizar	22
1.5.1 Modelo de datos	22
1.5.2 Lenguaje de programación y lenguaje de consultas para grafos RDF	23
1.5.3 Entorno de Desarrollo Integrado	23
1.5.4 Bibliotecas	24
1.5.5 Plataforma	25
Conclusiones parciales	25
CAPÍTULO 2. DESCRIPCIÓN DE LA PROPUESTA	26
2.1 Introducción	26
2.2 Métrica para el cálculo de la reciprocidad.....	26
2.2.1 Definición de métrica para el cálculo de la reciprocidad	26
2.3 Componentes del método propuesto	26
2.3.1 Base conceptual	27
2.3.2 Inferencia de redes sociales.....	28
2.3.3 Cálculo de la influencia social	29
2.4 Representación de las redes de coautoría	30
2.5 Clasificación del resultado del cálculo de la reciprocidad.....	31
2.5.1 Definición de la variable difusa Reciprocidad.	38
2.6 Conclusiones parciales	41
CAPÍTULO 3. VALIDACIÓN DE LA PROPUESTA	42
3.1 Introducción	42
3.2 Pruebas.....	42

3.2.1 Pruebas de unidad	42
3.2.2 Pruebas de aceptación	44
3.3 Utilización de un sociograma para la validación del método propuesto	45
3.4 Aplicación del sociograma	45
3.5 Correlación entre los resultados del sociograma y los resultados del método propuesto	46
3.5.1 Coeficiente de correlación de Spearman.....	46
3.5.2 Coeficiente de correlación lineal o Coeficiente de Pearson	47
3.5.3 Diagrama de dispersión	47
3.6 Aplicación del coeficiente de correlación de Spearman y coeficiente de Pearson	48
3.7 Conclusiones parciales	53
CONCLUSIONES GENERALES	54
REFERENCIAS BIBLIOGRÁFICAS	55
ANEXOS	59
Anexo 1. Encuesta realizada para la aplicación del sociograma.....	59
Anexo 2. Acta de aceptación	61

ÍNDICE DE FIGURAS

Figura 1. Ejemplo de red social como un grafo (Informe de la Unesco sobre la ciencia 2010 ,2010).... 9

Figura 2. Componentes del método propuesto..... 27

Figura 3. Prototipo de interfaz para el cálculo de la reciprocidad..... 30

Figura 4. Ejemplo del cálculo de la reciprocidad..... 30

Figura 5. Representación de la red de coautoría 31

Figura 6. Representación detallada de una red de coautoría..... 31

Figura 7. Representación de la función de pertenencia Función de pertenencia puntual. Fuente:
(Martín del Brío y Sanz Molina 2001). 33

Figura 8. Representación de la función de pertenencia Saltos escalones estrictos. Fuente:(Martín del
Brío y Sanz Molina 2001). 34

Figura 9. Representación de la función de pertenencia Función Gamma. Fuente:(Martín del Brío y
Sanz Molina 2001). 34

Figura 10. Representación de la función de pertenencia Función L. Fuente:(Martín del Brío y Sanz
Molina 2001). 35

Figura 11. Representación de la función de pertenencia Función Lambda. Fuente:(Martín del Brío y
Sanz Molina 2001). 36

Figura 12. Representación de la función de pertenencia Función Pi. Fuente:(Martín del Brío y Sanz
Molina 2001). 36

Figura 13. Ejemplo de la representación de la variable lingüística Altura. 37

Figura 14. Conjuntos difusos de la variable lingüística Reciprocidad..... 38

Figura 15. Representación del valor 0.309 en los conjuntos difusos de la variable lingüística
Reciprocidad. 40

Figura 16 Código de la prueba unitaria 1 del algoritmo implementado (Fuente: Elaboración propia).. 43

Figura 17 Cuadro de respuesta de la prueba unitaria 1 (Fuente: Elaboración propia) 43

Figura 18 Código de la prueba unitaria 2 del algoritmo implementado (Fuente: Elaboración propia).. 43

Figura 19 Cuadro de respuesta de la prueba unitaria 2 (Fuente: Elaboración propia) 44

Figura 20. Resultados de las pruebas de aceptación 44

Figura 21. Diagrama de dispersión (Escalante et al. 2005). 48

Figura 22.Comparación de los resultados obtenidos de correlación. 52

Figura 23.Diagrama de dispersión sociograma / método computacional 52

Figura 24.Diagrama de dispersión sociograma / método computacional (Polinómica)..... 53

ÍNDICE DE TABLAS

Tabla 1. Clasificación de las redes sociales 9

Tabla 2. Cuadro comparativo de las herramientas para el análisis y representación de redes sociales	22
Tabla 3. Pseudocódigo del algoritmo 1.....	28
Tabla 4. Pseudocódigo del algoritmo 2.....	28
Tabla 5. Pseudocódigo del algoritmo 3.....	29
Tabla 6. Coeficiente de correlación del registro de eventos Fiesta.	49
Tabla 7. Coeficiente de correlación del registro de eventos Deporte.	49
Tabla 8. Coeficiente de correlación del registro de eventos Investigación.	50
Tabla 9. Coeficiente de correlación de la unión de todos los registros de eventos	50
Tabla 10. Coeficiente de correlación entre el método tradicional y el método computacional de todos los registros de eventos.....	51

ÍNDICE DE FÓRMULAS

Fórmula 1. Cálculo de la reciprocidad	26
Fórmula 2. Coeficiente de correlación de Spearman (Sarabia Alegría y Pascual Saéz 2005).	46
Fórmula 3. Coeficiente de correlación de Pearson (Sarabia Alegría y Pascual Saéz 2005).	47

INTRODUCCIÓN

Con el avance de las Tecnologías de la Informática y la Comunicación (TIC) en todas las esferas de la vida social se han incrementado considerablemente los datos electrónicos que registran las aplicaciones y sistemas de gestión de flujos. Desafortunadamente, la información contenida en estos registros no siempre es empleada adecuadamente para analizar procesos subyacentes en los sistemas (Capote y Viera 2014).

Los registros de eventos revelan cómo los usuarios ejecutan las actividades almacenadas en las trazas del sistema (van der Aalst y Song 2004). Así mismo, estas relaciones pueden ser representadas mediante grafos o matrices empleando el enfoque de la sociometría o sociografía, la cual es una técnica de la Sociología, surgida en la década de 1930 que permite representar las relaciones interpersonales (Aalst 2011), sentando las bases para el análisis de las redes sociales.

El estudio mediante métodos formales de la manera en que las personas construyen sus relaciones y cómo estas interacciones permiten descubrir diversas potencialidades y patrones, contribuyendo no solo a explicar la ocurrencia de ciertos fenómenos sociales, sino además a emplear los análisis en favor de detectar, estimular, predecir o controlar la magnitud de los comportamientos humanos en diversos contextos y escenarios. En la medida que las personas despliegan sus actividades en la web y las organizaciones emplean aplicaciones para gestionar las actividades de sus negocios, se confiere una oportunidad para inferir interacciones sociales a partir de los registros de eventos que guardan los sistemas, teniendo en cuenta que los eventos contienen la información sobre las actividades que ejecutan los usuarios del sistema (Capote y Viera 2014).

El ambiente competitivo de los negocios exige una administración eficaz de los recursos humanos y de la información como activos fundamentales en la empresa. Estudios revelan que algunos patrones de relaciones encontrados en las empresas no aparecen en los organigramas y se desarrollan a partir de la interacción impuesta y determinada por la organización formal del trabajo (Chiavenato 2006). Los resultados de este estudio demostraron que no siempre la organización social de una empresa se corresponde exactamente a su estructura organizativa y que existen otros procesos espontáneos de evolución social, ignorados con frecuencia en la administración, que también participan directamente en la realización del objetivo de la empresa. El análisis de los patrones informales de relación que se establecen en el tejido social empresarial complementa la explicación del comportamiento de los individuos en el trabajo y fortalece el proceso de toma de decisiones en la empresa.

El Análisis de Redes Sociales (ARS) es aplicable a redes sociales dentro de un entorno, viéndose con mayor influencia en las empresas a nivel mundial. El ARS es una actividad que surge con necesidad en los procesos de toma de decisiones a diferentes niveles.

Una **red social** se define como un conjunto de puntos (actores) vinculados por una serie de relaciones que cumplen determinadas propiedades debiendo poseer o estar vinculadas a ciertas particularidades o rasgos comunes con el fin de poder interactuar entre sí. Para la representación de las redes sociales generalmente se utilizan métodos formales matemáticos, ya que permiten representar las descripciones de las redes de forma compacta y sistemática. Igualmente, las estructuras más conocidas para la representación de sus relaciones, son los grafos y las matrices. La teoría de grafos ha sido aceptada como herramienta de análisis de redes y algoritmos que se utilizan, además de tener varios años de uso y aceptación (Rouco 2016).

La inferencia de las redes sociales influye en el cálculo de la influencia social que se encuentra presente en todos los ámbitos de la vida del ser humano. La sociedad influye sobre las percepciones, actitudes, juicios, opiniones o comportamientos de las personas.

La **influencia social** son los cambios de conducta o pensamiento que se producen en un sujeto debido a una serie de procesos mentales derivados de la comunicación con otros seres o medios. Esta influencia puede ser dirigida a un fin o simplemente debida a la presión de grupo, derivándose de lo que el propio sujeto considere que se le pide o de lo que directamente le sea comunicado (Mimenza 2017). La influencia social, o persuasión, se basa en ciertos principios psicológicos que sirven de guía en determinadas conductas humanas como por ejemplo el principio de escasez, principio de validación social, principio de autoridad, principio de simpatía, principio de coherencia y el principio de reciprocidad.

Los estudios de la influencia se emplean generalmente en entornos formales de datos, donde las actividades son predeterminadas para los usuarios y las interacciones que se producen son impositivas. En este tipo de entornos, una persona no es capaz de elegir la actividad que desee, puesto que existen procesos específicos que limitan el grado de libertad del usuario para realizar una iteración propia.

Se ha comprobado que existe documentación del uso de las técnicas de influencia social mayormente aplicables para entornos formales de datos. Sin embargo, el estudio de entornos informales consta de escasos resultados de investigación. Esta realidad invita a la experimentación en entornos informales de datos, donde las relaciones que se producen tienen un carácter espontáneo, teniendo los usuarios la capacidad de elegir con quien establecer una relación. Lo anterior puede evidenciarse en aplicaciones sociales como por ejemplo las de mensajería instantánea. En este tipo de aplicaciones se registran un conjunto de conversaciones donde un emisor establece una comunicación con un receptor, en ambas direcciones (bidireccional). Por tanto, la técnica de influencia social más apropiada para su estudio es la reciprocidad (Chiavenato 2006).

La reciprocidad es uno de estos principios o criterios que posibilitan determinar cuán influyente es una persona sobre otra. La reciprocidad generalizada se da entre parientes cercanos y tiene un nivel más alto de obligación moral, compromiso ideológico es mayor, reciprocidad equilibrada implica una

obligación de devolver en un tiempo determinado y con un valor determinado dependiendo del bien, reciprocidad negativa es un intercambio en el que se produce un beneficio e incluso un engaño (Gómez 2013), la reciprocidad directa es aquella estrategia de cooperación mutua, donde los participantes asumen costes en el presente para mejorar los resultados en el corto, medio y largo plazo y la reciprocidad indirecta se entiende como aquella actitud recíproca cuya motivación no surge por la ayuda percibida, sino de la observación y comportamientos cooperativos en otros (Cabezas y Esteban 2012). Varias investigaciones en la Universidad de las Ciencias Informáticas sobre la temática, han llegado hasta los estudios del comportamiento, la centralidad, y la importancia de los nodos en una red (Ortiz Muñoz 2015),(Capote y Viera 2014),(Rodríguez y Ruiz 2014).

Actualmente la reciprocidad como técnica de influencia social, se ha visto calculada en entornos formales de datos, no siendo utilizada en entornos informales de datos. En los entornos informales de datos se encuentran las redes de colaboración. Los criterios de medición científica para formar redes de colaboración sin importar el enfoque que se utilice, ya sea la web semántica, la minería de datos o la minería de procesos, están basados en la coautoría, llegando a formarse la red solamente teniendo en cuenta la misma. Sin embargo, la reciprocidad ha demostrado ser un aspecto esencial para determinar la influencia, logrando ir un paso más allá de la identificación de las redes de colaboración. Por tanto, el uso de la reciprocidad va a permitir inferir redes a partir de la interacción social, donde se tendrá en cuenta la coautoría y la influencia social que genera cada autor con respecto a la red. Esto va a permitir identificar hacia donde está evolucionando la red que se analiza y que grado de evolución presenta.

Partiendo del contexto anterior se identificó el siguiente **problema**: ¿Cómo determinar la influencia social a partir del principio de reciprocidad en una red de interacción social informal?

Se delimita como **objeto de estudio**: Análisis de influencia en redes de interacción social.

Para dar solución al problema planteado, se define como **objetivo general** de la investigación: Desarrollar un método de inferencia de redes de interacción social a partir de las relaciones recíprocas entre actores.

El objetivo general se desglosa en los siguientes **objetivos específicos**:

- Caracterizar los procesos de influencia en redes de interacción social a partir de la reciprocidad.
- Implementar un método que a partir de las relaciones recíprocas entre actores permita calcular la inferencia de redes de interacción social.
- Validar el algoritmo propuesto mediante un caso de estudio.

Enmarcado en el **campo de acción**: Análisis de reciprocidad en redes de interacción social.

A partir de un análisis preliminar se enuncia la siguiente **idea a defender**: si se desarrolla un método para inferir redes de interacción social a partir de relaciones recíprocas, entonces se podría estimar el comportamiento de la influencia social en una red.

Tareas a cumplir para el desarrollo de la investigación:

- Revisión bibliográfica acerca de los procesos de análisis de influencia en redes de interacción social para determinar los principales conceptos asociados a los mismos.
- Elaboración del marco teórico de la investigación.
- Estudio de las herramientas para el desarrollo de algoritmos para el cálculo de inferencia de redes de interacción social.
- Análisis para definir las herramientas, métodos y metodologías de desarrollo de software para la implementación del algoritmo.
- Identificación de los elementos a tener en cuenta en el cálculo de inferencia de redes de interacción social.
- Diseño de la fórmula para el cálculo de inferencia de redes de interacción social.
- Implementación de un algoritmo que dada una red de interacción social realice el cálculo de la inferencia.
- Validación de la propuesta de solución mediante la aplicación en un caso de estudio.

Para el desarrollo del presente trabajo se emplean los siguientes métodos científicos:

Métodos teóricos:

- **Analítico-sintético:** permitió analizar individualmente los principales conceptos relacionados con el área de estudio a tratar, posibilitando un análisis profundo de cada uno, para luego llevar a cabo el estudio de las relaciones que se establecen entre ellos.
- **Hipotético-deductivo:** se utilizó para guiar la investigación desde el planteamiento del problema hasta la verificación de la solución a partir de las validaciones, orientando la secuencia lógica de las tareas que se realizaron.
- **Análisis documental:** se utilizó para el apoyo de la investigación, donde el objetivo es la revisión de artículos, libros, tesis, publicaciones, documentos, que ayude al desarrollo del tema propuesto.
- **Modelación:** ha permitido la descripción de la propuesta de solución, basándose en la aplicación de restricciones y modelado del método a partir de redes de coautoría (RC) en grafos RDF.

Métodos empíricos:

- **Caso de estudio:** se empleó con el objetivo de probar la validez de la propuesta de solución permitiendo obtener resultados medibles. Es decir, comprobar que se cumplieran cada una de las restricciones del método propuesto a partir de su implementación.
- **Experimento:** se empleó para verificar la correspondencia entre los resultados obtenidos con la aplicación del método propuesto y los resultados de la aplicación de técnicas tradicionales de la Sociología.

El presente documento está estructurado en tres capítulos cuyos contenidos son:

- **Capítulo 1:** Fundamentación teórica: se exponen los principales conceptos relacionados con el análisis de redes sociales, influencia social. Se realiza un análisis sobre una de las técnicas de la influencia social, la reciprocidad, exponiendo los tipos de reciprocidad que existen y como se evidencian.
- **Capítulo 2:** Descripción de la propuesta: se describen los principales algoritmos diseñados. Se presenta la implementación del módulo para el cálculo de la reciprocidad. Visualización detallada de la red de coautoría. Corroborar en qué nivel se encuentra la red.
- **Capítulo 3:** Validación del método implementado: se realiza la validación del método desarrollado mediante varios criterios para determinar que la solución propuesta responde satisfactoriamente a los objetivos que se trazaron. Se realizan varias pruebas al código para asegurar la calidad del método propuesto verificando que no existan fallas.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

En este capítulo se expone el marco conceptual enfocado en el estudio del análisis de redes sociales, e influencia social, y la reciprocidad.

1.1 Conceptos básicos de redes sociales

En términos analíticos, una red social es una estructura social compuesta por un conjunto finito de actores y configurada en torno a una serie de relaciones entre ellos (Aguirre 2011).

Una forma de representar una red social es a través de un grafo donde los actores o nodos pueden ser personas o grupos de éstas (por ejemplo: empresas, comunidades, organizaciones, países, ciudades, y demás). Los vínculos o aristas son las relaciones que se manifiesten entre los actores (por ejemplo: amor, poder, alianzas, amistad, parentesco familiar, contacto por correo electrónico, creencias religiosas comunes, rivalidad, entre otras). En términos teóricos, las redes sociales configuran contextos de comunicación e intercambio entre actores, marcan pautas operativas, normas y valores que condicionan la conducta de los actores en ellas, y posiciones funcionalmente diferenciables que son claves para entender el comportamiento de los nodos dentro de cada red y el desempeño de la red en su conjunto (Aguirre 2011).

La sociometría, también conocida por sociografía se refiere a métodos de representar datos en las relaciones interpersonales en forma de gráfico o matriz (van der Aalst, Reijers y Song 2005). El término sociometría fue acuñado por Jacob Levy Moreno quien llevó a cabo el primer estudio sociométrico en la década de 1930 en la universidad de Columbia en Nueva York (Estados Unidos). Jacob L. Moreno decidió realizar una radiografía grupal, para mostrar las distintas relaciones que se establecen entre los sujetos que conforman el grupo, exponiendo los lazos de influencia, preferencia y las fuerzas dinámicas que mantenían en movimiento al grupo.

1.1.1 Teoría de grafos

En este epígrafe se presentan definiciones y notaciones utilizadas en los algoritmos que se proponen. Las definiciones que siguen son tomadas de (Tucker 2004).

DEFINICIÓN 1.1.1 Un grafo no dirigido, $G = (V, E)$ se define como un conjunto V finito y no vacío de vértices y un multiconjunto E de aristas, donde cada arista $(vi, vj) \in E, vi, vj \in V$ es un par no ordenado de vértices. Opcionalmente una arista puede tener un valor que la identifique y una lista de atributos. Cuando los elementos de E tienen multiplicidad uno, el grafo se denomina grafo simple.

DEFINICIÓN 1.1.2 Se denomina grado de un vértice v a la cantidad de aristas que inciden en v y se representa como $g(v)$.

DEFINICIÓN 1.1.3 En un grafo dirigido, se denomina grado de entrada de un vértice v a la cantidad de aristas que tienen como vértice final a v y se representa como $g^-(v)$. De forma similar, se denomina grado de salida a la cantidad de aristas que tienen como vértice inicial a v y se representa como $g^+(v)$.

DEFINICIÓN 1.1.4 Se define un grafo ponderado como una estructura $G = (V, E, fc)$ donde:

- ❖ V representa el conjunto de vértices del grafo.
- ❖ E representa un multiconjunto de aristas del grafo.
- ❖ La función $fc: E \rightarrow \mathbb{R}^+$ le hace corresponder a cada arista un valor real positivo.

DEFINICIÓN 1.1.5 Se denomina camino desde el vértice vi al vértice vj en un grafo $G = (V, E)$ a la secuencia de vértices $CA = v_{k_1}, v_{k_2}, \dots, v_{k_t}$ si $\exists (v_{k_1}, v_{k_2}), (v_{k_2}, v_{k_3}), \dots, (v_{k_t-1}, v_{k_t}) \in E, vi = v_{k_1}, v_j = v_{k_t}$. Además, una arista solo puede aparecer una vez.

DEFINICIÓN 1.1.6 Se denomina longitud de un camino $CA = v_{k_1}, v_{k_2}, \dots, v_{k_t}$ $t > 1$, a la suma de los costos de todas las aristas presentes en el mismo y se representa como $|CA| = \sum_{h=1}^{t-1} c_{kh \ k_{h+1}}$. En el caso de grafos no ponderados, la longitud del camino se puede calcular como la cantidad de aristas presentes en el mismo, o sea $|CA| = t - 1$.

1.1.2 Análisis de redes sociales

El Análisis de Redes Sociales (ARS) es un acercamiento al estudio de las interacciones sociales humanas. Brinda una colección de métodos, técnicas y herramientas de propósitos sociométricos para analizar redes sociales (van der Aalst, Reijers y Song 2005). El objeto de estudio del ARS son las redes sociales, su morfología, sus patrones de comportamiento general y su relación con los individuos que las componen (Aguirre 2011). El análisis de redes sociales puede ser utilizado para investigar patrones de afinidad, estructuras de comunidades, o de organización de redes sociales formales e informales. Estas redes sociales pueden estar asociadas a corporaciones, grupos familiares, grupos filiales, estructuras de orden y control, u organizaciones encubiertas (Capote y Viera 2014).

El ARS detecta posiciones dentro de las redes de relaciones y las características funcionales de éstas en relación a toda la red. Permite diferenciar posiciones, estrategias, flujos de transacciones y distribución de poder entre un conjunto de actores en un espacio y tiempo determinados; además de dar cuenta de cómo las redes asisten en configurar identidades, normas y valores compartidos en el ámbito que se esté aplicando (Aguirre 2011).

Uno de los problemas más importantes en el análisis de redes sociales es determinar los roles funcionales de los individuos en la red. Diferentes individuos en la red social a menudo desempeñan diferentes roles. Ejemplo de los roles intuitivos incluyen a los líderes, seguidores, reguladores, personas populares, entre otros (Capote y Viera 2014). Los resultados del ARS pueden también identificar otros

roles menos intuitivos, aunque estos sigan siendo comunes, como son los nodos puentes los cuales conectan sub-grupos en la red.

En sus inicios la Teoría de Redes solo contaba con una base metodológica pues sus aportes prácticos eran locales y limitados. La incapacidad de explicar la complejidad en las relaciones a gran escala solo pudo ser resuelta en la década de 1970 con la llegada de los algoritmos de computación, orientando la formalización y medición de los datos con mayor precisión (Aguirre 2011) .

1.1.3 Representación de las redes sociales

Una red social se puede representar como un grafo, donde cada nodo, también llamado actor o vértice, puede ser representado como un individuo o grupo de individuos. Una arista, también llamada relación o vínculo, conecta a dos nodos y representa el enlace entre dos individuos en una red social. El grafo puede ser dirigido o no dirigido, por ejemplo: a una persona A le gusta la persona B pero no viceversa (Mejia Olivares 2010). Además, las relaciones pueden ser binarias (existe una relación o no) o con peso. El peso de las aristas es usado para cuantificar la relación en dependencia de lo que se esté midiendo. El grafo resultante es nombrado sociograma (van der Aalst y Song 2004).

Las redes pueden tener pocos o muchos actores y uno o más tipos de relaciones entre pares de actores. El ARS utiliza herramientas de las matemáticas para representar información de las relaciones entre actores sociales, los grafos y las matrices. Una razón para usar métodos formales para representar redes sociales es que la representación matemática permite utilizar computadoras y elevar la precisión de dicho análisis (Mejia Olivares 2010).

Formalmente un sociograma es un grafo (P, R) donde P es el conjunto de individuos (o actores) y R son las relaciones que se establecen entre ellos ($R \subseteq P \times P$). Si el grafo es no dirigido, R es simétrico. Si el grafo es con peso, entonces existe adicionalmente una función W la cual asigna un valor a todos los elementos de R (van der Aalst y Song 2004).

La figura 1 representa una red social como un grafo, obtenido del Informe General del Estado de la Ciencia, la Tecnología y la Innovación realizado por la UNESCO en el año 2010. Los nodos representan países y las aristas las relaciones entre ellos, el tamaño de los nodos es proporcional al avance científico técnico que presentan estos países.

	lo que se quiera representar con la información.	
Redes basadas en la evolución	Este tipo de redes depende de los cambios que sufre la red a través del tiempo. Estas redes pueden ser de cualquier tamaño y tener distintas formas.	-Redes Estáticas. -Redes Dinámicas.
Redes basadas en su topología	Este tipo de redes depende de la complejidad de la red.	-Redes Simples. -Redes Complejas.
Redes basadas en su origen	Este tipo de redes depende de su fuente de datos de origen. Muchas de estas redes pueden representar comunidades virtuales y/o del mundo real	-Redes fuera de línea (Off-line se traduce a sin conexión). -Redes en línea (On-line se traduce a con conexión).

En las redes fuera de línea las relaciones sociales son establecidas sin la intervención de un medio electrónico. La administración y conocimiento de las relaciones recae exclusivamente en el conocimiento del individuo sin ayuda de un sistema informático que le permita llevar la gestión de contactos (Mejía Olivares 2010).

Las redes en línea son redes que dependen altamente de medios electrónicos y se mantienen ligadas a los cambios en la tecnología de los sistemas.

En sus inicios la mayoría de los estudios sobre redes sociales estaban basados en redes a pequeña escala. En años recientes se han incrementado los estudios basados en la estructura de las redes sociales. Esto se debe al alto número de sistemas en la web que permiten generar comunidades virtuales que pueden ser representadas como una estructura de red social. Estos sistemas web brindan una nueva fuente de información para poder estudiar las propiedades de las redes sociales del mundo real dada la cantidad de información que poseen y la cantidad de usuarios que interactúan (Mejía Olivares 2010).

Por lo general las redes sociales en línea involucran un gran número de nodos y poseen propiedades que no pueden ser obtenidas con un simple escalamiento de la red (Capote y Viera 2014).

1.1.5 Inferencia de redes sociales

Los vínculos entre los actores de una red social permiten la construcción de un grafo que representa las relaciones que se establecen entre ellos. El enriquecimiento de estas conexiones brinda un mayor nivel de entendimiento de los actores dentro de la red para futuros análisis. Diversas investigaciones avalan la importancia de la inferencia de información para el establecimiento de estas relaciones y el descubrimiento de nuevos conocimientos dentro de una red social (Lasso-Sambony, Ortega-Ponce y Corrales 2013).

La inferencia según la Real Academia de la Lengua Española es la acción y efecto de inferir, es decir, de sacar una consecuencia o deducir alguna cosa. Por tanto, la inferencia en el contexto de una red social es la acción de deducir una nueva relación. Esta relación se fortalece debido a una conclusión o consecuencia que conduce a un nuevo resultado. Una inferencia suele generarse a partir de un análisis de características y probabilidades con las cuales se puede generar un nuevo conocimiento. Debido a que son múltiples las relaciones que se pueden representar en una red, el proceso de inferir nuevos conocimientos y relaciones es muy complejo. Por lo que para su configuración se utilizan disímiles algoritmos que son capaces de generar un grafo a partir de un análisis del comportamiento de los actores dentro de la red social (Méndez y Vega 2015).

Los algoritmos para la generación de las relaciones entre los actores dentro de una red social tienen en cuenta comportamientos, actitudes y conocimientos con los que es capaz de inferir las conexiones entre los actores de una red. Con ello es posible obtener una representación más cercana a la realidad de las conexiones existentes dentro de la red (Aguirre 2011). Los algoritmos que generalmente se utilizan para inferir una red social se encargan de extraer el conocimiento a partir del agrupamiento de los actores teniendo en cuenta su comportamiento dentro de la red.

1.2 Análisis de Influencia Social

La influencia social estimula el comportamiento humano fuera de línea y presenta un papel especialmente importante en las ciencias sociales. En el análisis de influencia social, se opina que las actitudes y los gustos de los individuos están influenciados por otros. Los primeros ejemplos de influencia social están inspirados en estudios de enfermedades (virus), de cómo podría contagiar una persona a otra por medios de vínculos sociales (Ovejero Bernal y Digitalia (Firm) 2007).

Existen individuos que, en sus trabajos existentes, están integrados en las redes sociales por lo que llegan a ser influenciados a través de las interacciones sociales. Un ejemplo es que, si A compra un producto X entonces su amigo B está influenciado a comprar el mismo producto X, esta influencia social es más bien conocida como influencia local. Otro tipo de influencia de importancia es la influencia de popularidad puesto que no depende necesariamente de presentar una red social para llegar a tener

impactos globales en los individuos, por ejemplo, una persona famosa puede influir en el andar de vida de una persona sin tener vínculos sociales directos o no.

La influencia y la formación de enlaces están estrechamente interrelacionadas en el proceso de evolución de la red. El enlace entre dos nodos se debe a la razón de que su influencia mutua es lo suficientemente fuerte, mientras que a su vez la formación del enlace mejora la influencia entre estos dos nodos.

El campo de la influencia social es amplio, de forma que puede ser relacionado con la práctica total de la psicología social. De forma tal que las modalidades de influencia social puedan ser casi tan numerosas como se quiere por ejemplo imitación, contagio, facilitación social, entre otras. Se distinguen, por el escenario en que tiene lugar la influencia, entre influencia interpersonal, persuasión e influencia a través de los medios de comunicación.

Todas las personas, a lo largo del día, asisten a múltiples procesos de influencia social. Por ejemplo, *al levantarse se puede intentar convencer a algún familiar para que prepare el desayuno, o alguien tratará de vender un producto en la calle, o es posible que se le haya pedido a un compañero que nos preste sus apuntes.* Pero el proceso de persuasión no se limita a estos procesos de influencia en los que la interacción es cara a cara.

A través de esos procesos de influencia y persuasión se configuran sus afectos, creencias, actitudes, intenciones y conductas. La intención de influir, va dirigida siempre a conseguir un cambio en la conducta de los otros, individuos o grupos. Algunas veces está claro que el objetivo es conseguir un comportamiento concreto como, que se prepare el desayuno. Por otra parte, los intentos de influencia se pueden dar en procesos cara a cara (cuando se quiere convencer a un familiar o a un cliente) o a través de los medios de comunicación.

Se han propuesto muchas tipologías en el estudio de la influencia como:

En función del objetivo del agente de influencia, se encuentran con dos tipos de metas: conseguir un comportamiento concreto en el receptor o conseguir que éste cambie sus actitudes para producir, a la larga, un cambio conductual. La campaña de ventas de un determinado producto sólo va dirigida al comportamiento de comprarlo. Otras campañas publicitarias, como las centradas en la seguridad vial o en la prevención del SIDA, constituyen intentos de influencia dirigidos a ampliar y modificar los conocimientos sobre estas cuestiones, y a cambiar los aspectos afectivos y evaluativos relacionados con ellas, de forma que esos cambios afecten a múltiples conductas que se intentan modificar o implantar.

En función del escenario en que tiene lugar el intento de influir. Los tres tipos de escenarios más frecuentes son: comunicación interpersonal directa o cara a cara, comunicación directa dirigida a una audiencia o comunicación de masas.

A continuación, se describen en detalle cada una de ellas. La **comunicación interpersonal o cara a cara** se caracteriza porque la interacción es bidireccional y dialéctica. La interacción es en la que el agente de influencia y su receptor intervienen al mismo tiempo. El agente de influencia trata de imaginar cuáles son las representaciones que el otro tiene de la situación para tratar de adaptar su conducta a ellas, pero, a su vez, el receptor de influencia participa imponiendo su propia posición (Ovejero Bernal y Digitalia (Firm) 2007).

La **comunicación directa** dirigida a una audiencia está ejemplificada por un mitin. Se trata de un tipo de interacción que podríamos calificar de unidireccional y poco recíproca. Aunque el receptor de influencia se puede expresar mediante reacciones, como aplausos y abucheos, su influencia en la conducta del agente es mucho menor que en la comunicación interpersonal.

La **comunicación de masas** es la comunicación característica de los medios de comunicación. No existe contacto directo entre el comunicador y la audiencia, por lo que claramente la influencia es unidireccional. El blanco de influencia sólo puede influir sobre el agente aceptando o rechazando el medio de comunicación. El estudio de la influencia social pone en evidencia el poder de los otros para condicionar nuestra conducta, lo que sin duda despierta un cierto rechazo.

La Psicología social analiza los procesos psicológicos implicados en la influencia interpersonal y las tácticas de influencia más efectivas. El hacer una reflexión sobre la ética de estas tácticas, que en último término dependerá de las intenciones del agente de influencia.

Sintetizando lo expuesto anteriormente: los procesos de influencia intencionada pueden presentarse en muchos contextos diferentes; la persona o grupo que trata de influir puede intentar convencer a los otros en beneficio propio o para ayudarlos; se puede buscar conseguir un determinado comportamiento o cambiar las actitudes para que se refleje en varias conductas relacionadas con esa actitud.

(Ovejero Bernal y Digitalia (Firm) 2007).

1.2.1 Patrones de comportamiento

El comportamiento es la manera de proceder de un individuo con relación a su entorno, su reacción a los estímulos y su respuesta a un tejido de interacciones. Para caracterizar la influencia en una red social, se aprovechan algunos patrones de conducta interiorizados de manera colectiva en la psiquis humana. Estos patrones están enunciados en seis principios definidos por (Cialdini PH.D. 2009).

Coherencia

También llamado del compromiso o de la consistencia. El ser humano necesita ser consecuente en lo que hace, lo que compra y en general en su comportamiento. De una forma obsesiva, se tiene el deseo de parecer coherente con respecto a los demás. A la hora de tomar cualquier decisión, se notará la “presión” de comportarse de acuerdo a los principios propios.

La coherencia se basa en el deseo de ser y parecer una persona de actitudes y comportamientos consecuentes a lo largo del tiempo. Los individuos tratan de ser coherentes ante los demás, pero también de cara a ellos mismos. La coherencia es un rasgo de personalidad muy valorado socialmente, que se asocia a otros rasgos como el ser lógico, racional, estable y honrado. Por el contrario, a una persona poco coherente se la considera superficial, poco inteligente, insegura y débil.

Aprobación social

Es un principio irreprochable en psicología, postula que el individuo actúa de la misma manera que lo hace la sociedad (o subgrupos sociales) para lograr la aceptación de la misma, a menudo, aunque considere que la sociedad esté equivocada. Según el economista John Maynard Keynes “Es más fácil equivocarse con la multitud que enfrentarse a la multitud y decir la verdad”.

Simpatía

La simpatía se divide en cuatro instrumentos básicos: los elogios y halagos, el atractivo físico, la semejanza y la familiaridad.

Raramente un individuo realizará un regalo a otro que no es de su agrado. La simpatía es clave para vender, aunque la simpatía en extremo produce un efecto de rechazo. En pocos campos esto se hace tan evidente como en la política. Frecuentemente, los políticos aparecen rodeados de actores y famosos para apropiarse de parte de la simpatía que estas personas ocasionan.

Un factor que aumenta la simpatía es la familiaridad. La exposición repetida a un estímulo aumenta la atracción hacia él. La tendencia a evaluar de manera más positiva los objetos que nos son familiares que los nuevos se producen incluso cuando la persona no es consciente de haber visto ese estímulo.

Autoridad

La autoridad tiene muchas manifestaciones diferentes, no siempre relacionadas con el poder directo sino también con la credibilidad. En el principio de autoridad entran en juego dos elementos, jerarquía y símbolos. La jerarquía se basa en la creencia de que las personas que llegan a puestos superiores tienen más conocimiento y experiencia que el resto. Los símbolos aportan credibilidad y pueden ser el uniforme de un policía, la bata de un médico, los títulos que posea un académico, entre otros.

Escasez

Las oportunidades parecen más valiosas cuanto más difíciles resultan conseguir las. Si un cliente percibe una baja oferta, o una elevada demanda de un bien, inmediatamente se mostrará interesado y estará dispuesto a pagar un precio más alto.

Reciprocidad

La reciprocidad se refiere a la necesidad que existe en las relaciones sociales de restaurar el equilibrio. Por ejemplo, cuando se recibe algo, se siente la necesidad de devolver algo a cambio. La reciprocidad se fundamenta en que dos personas se tratan de acuerdo a los comportamientos anteriores que haya tenido una con respecto a la otra. Las normas sociales impelen a corresponder al “benefactor”, cuando esa persona ha ayudado, obsequiado o realizado cualquier tipo de favor. A la persona que no se comporta de acuerdo con esta norma se le considera habitualmente como ingrata o aprovechada.

La norma de reciprocidad es uno de los pilares para el funcionamiento de las sociedades. Es importante porque beneficia a los individuos y al grupo en su conjunto. Garantiza que cualquier recurso que se comparta en un momento determinado le será devuelto por los demás cuando lo necesite.

1.2.2 Análisis de la Reciprocidad

La norma de reciprocidad es una de las más importantes sobre la que se sustenta el funcionamiento de las sociedades humanas. Es importante porque beneficia a los individuos y al grupo en su conjunto. Garantiza que cualquier recurso que uno comparta en un momento determinado le será devuelto por los demás cuando lo necesite. (López-Sáez 2018).

En la vida cotidiana se pueden encontrar multitud de ejemplos de influencia en los que funciona este principio. Son frecuentes los favores y regalos no solicitados otorgados por empresas o individuos a funcionarios públicos, a políticos o a ejecutivos de otras empresas de los que piensan que pueden necesitar algún favor, aunque no sea inmediatamente. La norma de reciprocidad se aplica a una amplia gama de comportamientos. Esa actitud de complacencia del receptor de influencia se daba independientemente de que el cómplice del experimentador se mostrara posteriormente antipático y grosero o amable y simpático (López-Sáez 2018).

El principio de reciprocidad es la base de dos importantes tácticas de influencia: la técnica de «esto no es todo» y la técnica de «el portazo en la cara».

La técnica de «esto no es todo»

Esta técnica consiste, en síntesis, en ofrecer un regalo añadido al producto, hacer un obsequio o proporcionar algún tipo de beneficio como paso previo a la solicitud de un favor. Ésta es una técnica muy extendida en el sector de ventas. Es muy frecuente el que se añada un obsequio como, por ejemplo, un desodorante al comprar una colonia, ya que esta práctica aumenta el interés por el producto (López-Sáez 2018).

La técnica del «portazo en la cara»

Esta estrategia de influencia se basa en la obligación de hacer una concesión a aquellos que han cedido previamente en sus requerimientos. El procedimiento consiste en comenzar con una petición elevada y cuando ésta se rechaza (nos dan con la puerta en la cara), hacer una petición mucho menor, que suele ser aceptada. El principio de reciprocidad funciona sobre la base de las concesiones mutuas: puesto que tú has cedido en tu petición, yo debo también ceder en algo (Cialdini 1990).

La técnica del portazo en la cara se utiliza en negociaciones y suele ser más eficaz que el camino directo (Garrido, Herrero y Tabernero, 1997). El procedimiento consiste en situar la petición inicial en el extremo de las reivindicaciones que se intentan conseguir, aun a sabiendas que no se va a lograr tanto. Una vez que esa petición ha sido rechazada, se retira y se sustituye por una inferior, en la que se presenta lo que en realidad se deseaba desde el principio (López-Sáez 2018).

1.2.3 Tipos de reciprocidad

En este epígrafe se presentan los diferentes tipos de reciprocidad que existen y sus definiciones.

Reciprocidad generalizada:

Los recursos circulan sin esperar en principio nada a cambio, no hay un registro específico que valore lo que se da para que después sea devuelto. Este tipo de altruismo se da dentro del grupo de parentesco, principalmente entre padre e hijo, donde existe una obligación moral de darle cosas al hijo por el hecho de ser su padre, y lo mismo pasa con el hijo. Lo que se espera con este intercambio es reconocimiento, respeto, gratitud, no hay una contrapartida realmente económica. Se da en sociedades donde el parentesco rige la organización social, así, entre los cazadores-recolectores el que sacrifica el animal es el propietario para compartirlo entre los demás cazadores y cada cazador se encargaba de repartirlo entre los miembros de su familia (Gómez 2013).

También hay normas ideológicas que premian esas conductas y sancionan las contrarias: La generosidad es un valor en estas sociedades que permite acumular prestigio, aunque no sea algo explícito. Lo contrario está castigado, si uno no reparte es que es un individuo asocial, embrujado al que se aparta del grupo (Gómez 2013).

Reciprocidad equilibrada:

La reciprocidad equilibrada es una relación económica que implica el cálculo del valor del regalo, este intercambio logra crear alianzas, se intercambia con los grupos que son importantes los unos para los otros. Al aparecer estas relaciones sociales y alianzas le quitan el carácter económico, el capital que importa en estas sociedades es el de los parientes y amigos en conjunto ya que de ellos van a obtener todo lo necesario para vivir (Gómez 2013).

Reciprocidad negativa:

Es un intercambio en el que se produce un beneficio e incluso un engaño. Se da en pocos casos y siempre entre pueblos enemigos o extraños. No entran contacto directo ni en el momento del intercambio (Gómez 2013).

Reciprocidad directa:

La reciprocidad directa se define como aquella estrategia de la cooperación mutua donde los participantes asumen costes en el presente para mejorar los resultados a corto, medio o largo plazo. Significa que A da o hace alguna cosa por B, y B responde equitativamente dando o haciendo alguna cosa por A, ya sea en cantidad o en cualidad (Cabezas y Esteban 2012).

Reciprocidad indirecta:

La reciprocidad indirecta fue propuesta por el biólogo Richard Alexander en (Alexander 2017) en un intento por cubrir alguno de los huecos dejados por el 'altruismo de parentesco' (se define como aquellos actos beneficiosos para la familia u otros individuos genéticamente relacionados), y el 'altruismo recíproco' (se define como aquellas acciones que confieren un beneficio a otros con un coste para sí mismo pero con la perspectiva de un beneficio recíproco subsiguiente suficiente para cubrir costos). A pesar del alto desarrollo de ambos enfoques, continuaba sin hallarse una explicación plausible para aquellas actitudes recíprocas observadas entre agentes desconocidos.

La 'reciprocidad indirecta' se entiende como aquella actitud recíproca cuya motivación no surge por la ayuda percibida, sino de la observación de comportamientos cooperativos en otros. Esto es, una predisposición a ayudar desinteresadamente a aquellas personas que ayudan desinteresadamente a los demás: A ayuda a B, y C que ve la acción de A, decide cooperar con él a pesar de no haber sido ni el beneficiario de la acción observada ni tampoco esperar respuesta alguna por su acción. Se trata, pues, de un enfoque de reciprocidad cuya estructura adquiere la forma de una cadena de favores (Cabezas y Esteban 2012). Este tipo de reciprocidad es la que se pone de manifiesto en el contexto de la presente investigación.

1.3 Redes de colaboración y coautoría

La aplicación de la teoría de redes al estudio de la coautoría parte de la idea de entender a la ciencia como un sistema auto-organizado en el que la selección de colaboradores y la localización de la investigación se da a partir de las elecciones de los propios investigadores, más que a través de incentivos institucionales, nacionales o de cualquier otra índole, y sin que exista autoridad alguna en el

seno de la organización científica mundial encargada de arbitrar o resolver posibles disputas que puedan surgir durante los intercambios que en ella se realizan.

En el siglo XXI con el desarrollo de las Tecnologías de la Información y las Comunicación se han producido avances en la perspectiva científica de los estudios estructurales, se ha convertido la teoría de redes en un área de investigación interdisciplinaria emergente. Las investigaciones están orientadas al desarrollo de teorías y técnicas para aumentar el conocimiento existente tanto de las redes biológicas o tecnológicas, como de las académicas y de los sistemas complejos (He, Ding y Ni 2011).

Estos estudios tienen su origen en disciplinas como las matemáticas discretas y la teoría de grafos, la sociología, la matemática, la psicología de grupos o la biología, pero también en la bibliometría, la infometría y la cibermetría, y más recientemente en la física. Los estudios de redes de colaboración y coautoría comprenden múltiples enfoques, que se centran en la descripción y el análisis de sus propiedades, la investigación de su modelado y sus dinámicas y el establecimiento de nuevas técnicas para su visualización. Su ámbito de análisis engloba problemas derivados de la existencia de redes sociales, biológicas, de información o tecnológicas, como los sistemas de comunicación (Internet, redes telefónicas), las infraestructuras de transporte (carreteras, líneas ferroviarias o aéreas), los sistemas biológicos (interacción de proteínas, ácido desoxirribonucleico (ADN), evolución de epidemias) y un amplio abanico de estructuras de interacción social, entre las que se encuentran las redes de colaboración entre científicos (Donetti y Muñoz 2008).

Revisiones bibliográficas hacen referencia al aumento de artículos científicos donde se pone en práctica la coautoría. La colaboración científica es considerada como un requisito previo para la coautoría (Melin y Persson 1996). La misma puede manifestarse en proyectos científicos, publicaciones, contratos de investigación, contactos informales, patentes, formación a través de cursos, seminarios, tesis de pregrado, maestrías, doctorados, intercambios de investigadores y la participación en comités científicos (Ortiz Muñoz 2015).

La estructura y evolución del desarrollo científico desde dos modelos diferentes: el modelo descriptivo, destinado a definir los principales rasgos de un conjunto de datos casi siempre estáticos, y el modelo de proceso, que trata de determinar los mecanismos y dinámicas temporales mediante el empleo de redes del mundo real, como las redes de colaboración basadas en coautorías (Borner, Maru y Goldstone 2004).

Las redes de filiación de científicos, en las que los enlaces se obtienen a partir de la coautoría en una o más publicaciones, son más reales desde el punto de vista de red social que muchas redes de filiación institucional, puesto que este tipo de redes refleja de forma genuina la interacción profesional entre científicos y se postula como las redes sociales más grandes hasta ahora analizadas (Martin et al. 2013). Aunque similares a las redes de citación y co-citación, las redes de colaboración implican lazos sociales más fuertes que los existentes en las primeras. Las citas ocurren sin que los autores se

conozcan entre sí y pueden prolongarse a lo largo del tiempo, las redes de coautoría (RC), en cambio, implican una relación colegiada y temporal que tiene lugar en el ámbito de los estudios de análisis de redes sociales (Liu et al. 2005).

El estudio de la colaboración científica ayuda a establecer grupos y redes de trabajo, pudiendo ser analizada y evaluada mediante el empleo de técnicas bibliométricas y representada mediante lo que algunos autores denominan redes de coautoría (RC) o mapas bibliométricos. A lo largo de la investigación se utiliza el término “coautoría” para hacer referencia a la firma conjunta de un trabajo científico por dos o más autores.

Las RC hacen referencia a una clase importante de las redes sociales y se han utilizado ampliamente para determinar la estructura de colaboraciones científicas y la situación de los investigadores individuales. Ejemplos de RC se muestran en el Número del Proyecto Erdos, el menor número de enlaces de coautoría entre cualquier individuo matemático y el matemático húngaro Erdos (Castro y Jerrold W. Grossman 1999).

Las asociaciones de coautoría entre científicos reconocen tanto las deudas intelectuales, como las personales y ofrecen la posibilidad de identificar y medir la actividad social y la influencia entre distintas especialidades científicas. El examen de los enlaces de coautoría entre científicos muestra a aquellos investigadores que trabajan en la misma área de conocimiento, aunque no necesariamente en conjunto. El resultado es un número de grupos colaboradores de tamaño variable, conectados o aislados los unos de los otros, dentro de los cuales algunos científicos juegan un papel principal, otros son los que sirven de nexo, de unión entre grupos y otros desempeñan ambos papeles simultáneamente (T. D. Stokes and J. A. Hartley 1989). Newman abunda en el concepto de RC indicando que se está en presencia de redes que además de describir la sociedad académica, muestran la estructura del conocimiento (Newman y Girvan 2004).

Medir la colaboración científica implica incluir varios factores. El estudio de los intercambios informales entre científicos requiere observación directa. Este contacto puede definirse en forma de artículos científicos que son en los que los especialistas de la colaboración ponen su atención. Sin embargo, además de la coautoría, también se podrían tener en cuenta aspectos como la movilidad de los investigadores, el desarrollo de proyectos científicos o la creación de centros de investigación compartidos (Ortiz Muñoz 2015).

El análisis de coautoría, también ha sido aplicado a diversas conferencias de ACM (*acrónimo de Association for Computing Machinery*), Recuperación de Información (SIGIR) y Gestión de Datos (SIGMOD) (Nascimento, Sander y Pound 2003) ,así como las matemáticas y la neurociencia y el campo de análisis de redes sociales (Ball y Newman 2013) (Azizifard 2014). Las RC internacionales se han estudiado en la JASIST (*acrónimo de Journal of American Society for Information Science & Technology*) y en Science Citation Index (Wagner, Bornmann y Leydesdorff 2015).

Los elementos anteriores permiten aseverar que el trabajo con redes de colaboración y coautoría es un área de investigación en desarrollo. Posibilita el análisis de la colaboración e intercambio científico de los investigadores. Definiéndose así que la coautoría constituye un indicador de impacto a tener en cuenta para la modelación y análisis de redes de colaboración científica.

1.4 Herramientas para el análisis de redes sociales

Dado los programas más empleados en los trabajos de investigación sobre la visualización y el análisis de redes sociales según (Mejía Olivares 2010), se hizo una selección de los que actualmente se mantienen activos. A continuación, se describen cada una de estas herramientas.

R

“R” es un entorno de trabajo para computación estadística y gráfica. Fue desarrollado por Bell Laboratories (*Formalmente AT&T, actualmente Lucent Technologies*) por John Chambers y un grupo de desarrolladores. Unas de las ventajas de “R” es la facilidad con la que se puede desarrollar un diseño con calidad, incluyendo símbolos y fórmulas matemáticas. El mismo se ocupa de los valores predeterminados de las opciones de diseño de menor importancia en los gráficos, pero el usuario conserva el control total de estos. Está disponible como software libre bajo los términos de la Licencia Pública General de GNU de la Free Software Foundation. Compila y ejecuta en una amplia variedad de plataformas UNIX y sistemas similares (incluyendo FreeBSD y Linux), Windows y MacOS (Chambers, y otros 2018).

“R” es también un lenguaje de programación. Una de sus características más importantes es su extensibilidad. Numerosos investigadores han creado paquetes de “R”, que son básicamente librerías con implementaciones de diversos algoritmos, y que son publicadas de forma pública en la red CRAN (The Comprehensive “R” Archive Network, por su denominación en inglés). El repositorio de paquetes CRAN cuenta actualmente con 3.725 paquetes disponibles. Entre estos paquetes destacan el paquete “network”, el paquete “sna”, el paquete “statnet” y el paquete “igraph”, para análisis de redes sociales. Los dos primeros están disponibles en CRAN, el tercero actualmente disponible en statnet (WebMining Consultores 2012).

El paquete “sna” versión 2.3-2 fue distribuido en febrero del 2014, el cual presenta una gama de herramientas para el análisis de redes sociales. Entre sus funcionalidades incluye los índices de nodo y de nivel gráfico, distancia estructural y de covarianza, métodos de detección de equivalencia estructural, la regresión de la red, generación de gráficos al azar y la visualización de la red en 2 y 3 dimensiones (Carter T. Butts 2014).

El análisis de este entorno permitió identificar que este permite importar archivos XML, pero no adecuadamente un registro de eventos en formato Extensible Event Stream (XES). Además, éste solo

tiene en cuenta un contexto de interacción cuando realiza la inferencia de redes sociales (Capote y Viera 2014).

SoNIA

“SoNIA” (Social Network Image Animator) es un proyecto de código abierto en forma de módulos, está basado en Java y concebido para la visualización de datos de una red social, tanto dinámicos como longitudinales. Por dinámica se refiere además de información sobre las relaciones (aristas) entre varias entidades (actores, nodos) también hay información sobre cuándo se producen estas relaciones, o por lo menos el orden relativo en que se producen (McFarland y otros 2014).

El objetivo de esta herramienta es leer la información de una red dinámica de varios formatos, ayudar al usuario en la construcción de los diseños de la red y exportar los resultados en imágenes o animaciones. Además, proporciona información sobre las técnicas y ajustes de los parámetros utilizados para la construcción de los diseños y algún tipo de estadística que indica la "exactitud" o grado de distorsión presente en el diseño de la red.

“SoNIA” puede ser utilizada como plataforma de trabajo para desarrollar, probar y comparar varias técnicas de diseños de redes sociales. También se puede concebir como un navegador de datos de la red social basada en el tiempo, ya que tiene opciones para inspeccionar visualmente un conjunto de relaciones en grandes conjuntos de datos a diferentes secciones de tiempo y niveles de agregación (McFarland y otros 2014).

La herramienta solo permite la importación de archivos en texto plano, con la extensión “.net” o su versión nativa “.son”. Además, ésta solo tiene en cuenta un contexto de interacción cuando realiza la inferencia de redes sociales (Capote y Viera 2014).

Marco de trabajo ProM

“ProM” es un marco de trabajo capaz de importar registros de evento en formato XES o MXML además de poder cargar modelos de procesos de diferentes estándares. Algunas de sus principales funcionalidades son: el descubrimiento de la perspectiva de control de flujo de un proceso, la minería de procesos poco estructurados, el análisis de los recursos y la perspectiva de datos de un proceso, la inspección de los parámetros de rendimiento específicos de un proceso y la verificación del cumplimiento de un estándar dado un registro de eventos y un modelo de proceso (Agarwal y Singh 2014).

Adicionalmente la herramienta soporta una gran variedad de técnicas de filtrado las cuales pueden ser usadas en un registro de eventos en el paso de preprocesamiento. Dicha herramienta provee una gran cantidad de formatos de exportación, desde representación visual de los resultados (ejemplo PNG) a diferentes formatos de registros (ejemplo CSV).

“ProM” es un marco de trabajo de código abierto desarrollado para dar soporte a una amplia variedad de técnicas de Minería de Procesos. Diferentes implementaciones de algoritmos al servicio de diferentes funcionalidades están disponibles en forma de complementos, donde investigadores de todo el mundo pueden hacer su aporte.

Como una plataforma académica este marco de trabajo está en la vanguardia en la investigación de Minería de Procesos. Esta hace fácil el desarrollo y prueba de nuevos algoritmos, aunque se requieren de ciertos niveles de conocimiento previos en Minería de Procesos para usarlo.

Esta herramienta permite la importación de registros de eventos en formato XES, pero los complementos desarrollados para el análisis de redes sociales solo tienen en cuenta un contexto de interacción. Esto se evidencia mediante la limitación de importar solo un registro de eventos para realizar los análisis.

Tabla 2. Cuadro comparativo de las herramientas para el análisis y representación de redes sociales

Nombre	Tipo de herramienta	Empresa	Análisis de redes sociales
GNU R	Código abierto	Lucent Technologies (www.r-project.org)	Si
ProM	Código abierto	Process Mining Group (managed by the AIS group at TU/e) (www.processmining.org)	Si
SoNIA	Código abierto	University of Stanford (www.stanford.edu)	Si

A partir del análisis de estas herramientas se concluye que a pesar de que todas se usan para el análisis de redes sociales ninguna realiza el cálculo de la influencia social haciendo uso del criterio de reciprocidad, por lo que se hace necesario implementar el método que permita realizar este cálculo.

1.5 Herramientas y tecnologías a utilizar

Las herramientas y tecnologías que se describen en el presente capítulo son las que serán utilizadas para el desarrollo de la solución propuesta.

1.5.1 Modelo de datos

RDF (Resource Description Framework):

Modelo de datos que utiliza XML, N-triples, Turtle, entre otros, como lenguaje para serializar los datos y metadatos de los recursos de la web. Permite la interoperabilidad entre aplicaciones que intercambian información comprensible por la página web, para proporcionar una infraestructura que soporte

actividades de metadatos. Dicho modelo de datos constituye la fuente principal de información que almacena los metadatos bibliográficos necesarios para el desarrollo del prototipo funcional. Al ser un estándar a nivel mundial, posibilita la utilización de cualquier grafo RDF que almacene metadatos bibliográficos como fuente de datos del prototipo funcional (Lapuente y Lapuente 2018).

1.5.2 Lenguaje de programación y lenguaje de consultas para grafos RDF

Python es un lenguaje de programación interpretado o de script, con tipado dinámico, fuertemente tipado, multiplataforma y orientado a objetos (González Duque 2015). Sobresale en la manipulación de datos y la programación de red por la cantidad de librerías que contiene y que facilitan el trabajo para los desarrolladores. Python ha desarrollado una gran y activa comunidad científica de computación y análisis de datos. Ha pasado de ser un lenguaje informático científico a ser uno de los lenguajes más importantes para la ciencia, el aprendizaje automático y el desarrollo general de software en la academia y la industria (McKinney 2013). En el presente trabajo se utiliza Python 3.6 para realizar la implementación de la solución propuesta por la variedad de librerías que contiene para el trabajo con redes colaborativas.

SPARQL

Lenguaje de consulta para Resource Description Framework (RDF). Se puede utilizar para expresar consultas que permiten interrogar diversas fuentes de datos, si los datos se almacenan de forma nativa como RDF o son definidos mediante vistas RDF a través de algún sistema middleware. SPARQL también soporta la ampliación o restricciones del ámbito de las consultas indicando los grafos sobre los que se opera. Es aplicado como el lenguaje para obtener información de la fuente de datos principal del prototipo funcional. A partir del uso de la biblioteca Apache Jena este lenguaje de consulta constituye la herramienta esencial para la interoperabilidad con las fuentes de datos que puede usar el prototipo funcional. Permite acceder a un grafo RDF a partir de un archivo físico o a través de un SPARQL Endpoint (Prud'hommeaux y Seaborne 2008).

1.5.3 Entorno de Desarrollo Integrado

Un entorno de desarrollo integrado (IDE, por sus siglas en inglés) es un software que facilita el desarrollo de aplicaciones. Están diseñados para abarcar todas las tareas de programación en una aplicación y ofrecen una interfaz que presenta todas las herramientas que un desarrollador necesita, incluyendo entre otras: compilador y depurador de errores (JetBrains 2018).

PyCharm

Se decidió utilizar PyCharm 2018.1 para el desarrollo del método ya que es un IDE que proporciona el completamiento inteligente de códigos, inspecciones de códigos, resaltado de errores sobre la marcha y soluciones rápidas, junto con refactorizaciones automáticas de códigos y capacidades de navegación avanzadas. Se integra con IPython Notebook, tiene una consola interactiva de Python y es compatible con Anaconda, así como con múltiples paquetes científicos, incluidos matplotlib y NumPy. PyCharm también admite JavaScript, CoffeeScript, TypeScript, Cython, SQL, HTML/CSS, lenguajes de plantilla, AngularJS, Node.js entre otros. También posee una gran colección de herramientas lista para usar, un depurador integrado y un corredor de prueba; Python Profiler; una terminal incorporada; e integración con sistemas de control de versiones y herramientas de base de datos incorporadas (JetBrains 2018).

1.5.4 Bibliotecas

Apache Jena

Marco de trabajo o biblioteca para Java que permite la construcción de aplicaciones para la Web Semántica, ofreciendo una colección de herramientas y bibliotecas Java. Incluye la utilización del lenguaje de consultas para grafos RDF SPARQL 1.1 Update, elemento que lo destaca entre otros marcos de trabajo con funcionalidades afines. Entre sus principales características se encuentran:

- Posee una interfaz de programación de aplicaciones o API para RDF, la misma que soporta la creación, manipulación y consulta de grafos RDF.
- Posee un API para el manejo de ontologías.
- Permite realizar lectura y escritura de documentos en formato RDF/XML, N3 y NTriples.
- Posee un almacenamiento persistente para poder recopilar gran cantidad de tripletas RDF (Panchal 2017).

NetworkX

Con NetworkX, se puede cargar y almacenar redes en formatos de datos estándar y no estándar, generar muchos tipos de redes aleatorias y clásicas, analizar la estructura de la red, construir modelos de red, diseñar nuevos algoritmos de red, dibujar redes, entre otras.

NetworkX está destinado a proporcionar:

- Herramientas para el estudio de la estructura y dinámica de redes sociales, biológicas y de infraestructura.
- Una interfaz de programación estándar e implementación gráfica que es adecuada para muchas aplicaciones.

- Un entorno de rápido desarrollo para proyectos colaborativos y multidisciplinares.
- Una interfaz para los algoritmos numéricos existentes y el código escrito en C, C ++ y FORTRAN («NetworkX - Python Graph Library» 2018).

1.5.5 Plataforma

Anaconda

La distribución de código abierto Anaconda facilita el proceso de implementar soluciones relacionadas con manipulación de datos en Python y cuenta con más de seis millones de usuarios. Incluye cientos de paquetes de código abierto, así como el paquete conda y el administrador de entorno virtual para Windows, Linux y MacOS. Conda facilita y agiliza el proceso de instalar, ejecutar y actualizar entornos complejos como SciPy. Anaconda es la base de millones de proyectos, así como de aprendizaje automático de Amazon Web Services y Anaconda para Microsoft en Azure y Windows (Anaconda 2019). Se decidió que la versión Anaconda2-2.5.0 es la más adecuada para trabajar con la versión lenguaje de programación seleccionado, debido a que contiene una amplia colección de librerías empleadas para facilitar el trabajo con Python.

Conclusiones parciales

- El estudio de influencia en la red permite determinar la capacidad que tiene un nodo de cambiar el comportamiento de otro.
- El análisis de las principales fuentes bibliográficas permitió identificar los aspectos que caracterizan la influencia en la red, entre los que se señalan la autoridad, reciprocidad, coherencia, simpatía, aprobación social y escasez.
- A partir del cálculo de la reciprocidad se puede hacer un análisis minucioso del grado de evolución de las redes de coautoría.

CAPÍTULO 2. DESCRIPCIÓN DE LA PROPUESTA

2.1 Introducción

En este capítulo se presenta un método para la inferencia de redes sociales a partir de las redes de colaboración orientado a las redes de coautoría explicándose en detalle cada componente que lo integra. Mediante su uso se calcula la reciprocidad, que constituye una de las técnicas de influencia social.

2.2 Métrica para el cálculo de la reciprocidad

Para el desarrollo del algoritmo planteado se define una métrica para el cálculo de la reciprocidad en una red de coautoría para el cual se entiende como premisa que, si un individuo X realiza una acción hacia un individuo Y, logrando que este reaccione de la misma manera hacia el primero, entonces se infiere que el individuo X y el individuo Y poseen un grado de reciprocidad respecto al resto de la red.

2.2.1 Definición de métrica para el cálculo de la reciprocidad

Sea R el valor correspondiente a la evaluación de la reciprocidad del individuo A sobre el individuo B en el conjunto de contextos a analizar, se tiene que:

$$R = \left(\frac{w(A, B)}{g(A)} \right) * \left(\frac{w(B, A)}{g(B)} \right)$$

Fórmula 1. Cálculo de la reciprocidad

donde:

$w(A, B)$: peso de la arista de ab

$g(A)$: grado de la arista a

$w(B, A)$: peso de la arista ba

$g(B)$: grado de la arista b

La relación de reciprocidad se establece si $w(A, B), w(B, A) > 0$

Si $w(A, B) > 0$, entonces $g(A) > 0$

Si $w(B, A) > 0$, entonces $g(B) > 0$.

2.3 Componentes del método propuesto

En este epígrafe se presenta la descripción del método propuesto. En la Figura 2 se muestra un esquema que representa los componentes del mismo.

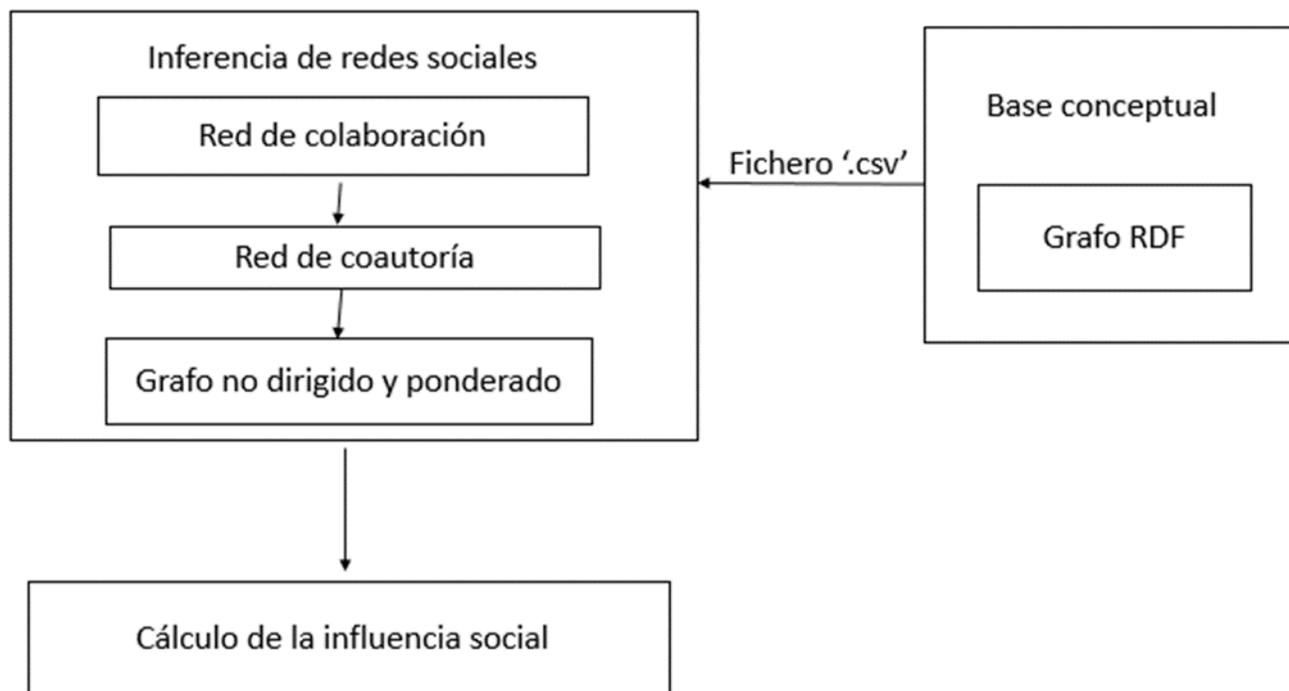


Figura 2. Componentes del método propuesto

A continuación, se exponen en detalle cada uno de los componentes de este método, así como las relaciones existentes entre los mismos.

2.3.1 Base conceptual

Para comenzar la ejecución del método se parte como base conceptual de un grafo RDF que ya contiene la información referente a las publicaciones y autores de la universidad. Este grafo cuenta con 8011 nodos y 35820 aristas. Tiene como atributos título de la publicación, año en que fue realizada la publicación, nombre del autor, la afiliación a la que pertenecen los autores, fuente de datos, resumen, y la URI donde se pueden encontrar las publicaciones. De los atributos antes mencionados solo se utilizarán el nombre del autor y el título de las publicaciones, por lo que es necesario crear un fichero csv a partir de este grafo RDF con los datos necesarios para crear posteriormente una red de coautoría. Se carga como entrada en el Apache Jena Fuseki versión 3.7.0 este grafo RDF para luego realizar una consulta en el SPARKQL que devuelva ordenadamente todas las publicaciones por autor, para luego importar estos datos a un fichero CSV. Esta operación es realizada por el algoritmo 1, el cual se explica a continuación.

El algoritmo **ObtenerCSV** hace uso de una consulta en SPARKQL para obtener a partir del grafo RDF de entrada el fichero CSV con la relación de autores y publicaciones. A continuación, se describen los pasos que sigue el mismo.

Algoritmo 1 ObtenerCSV

Entrada: Fichero 'grafo.ttl' el cual contiene un grafo RDF.

Salida: Fichero 'queryResults.csv'.

Tabla 3. Pseudocódigo del algoritmo 1.

1:	Seleccionar los atributos nombre y título
2:	Para cada título de publicación
	Buscar autores y devolver los nombres
3:	Ordenar por título

Luego se importa el fichero CSV, el cual se utilizará para crear la red de coautoría. Este fichero presenta por cada una de sus filas un par autor, publicación.

2.3.2 Inferencia de redes sociales

Para realizar la inferencia de redes sociales se recibe un fichero CSV que contiene los nombres de los autores con sus respectivas publicaciones obtenido en la ejecución del algoritmo 1. Con la entrada de este fichero se crea una red de coautoría donde los autores son los nodos y las aristas representan la colaboración entre ellos, bajo la condición de ser una red no dirigida y ponderada. De realizar este proceso se encarga el algoritmo 2, el cual se detalla a continuación.

El algoritmo **ObtenerRC** fue desarrollado para modelar una red de coautoría, la cual va a ser representada como un grafo ponderado y no dirigido. El algoritmo tiene como entrada un fichero '.csv' de donde extrae la información de todos los autores y sus publicaciones. Genera las relaciones correspondientes entre los autores y devuelve una red de coautoría ponderada y no dirigida.

Algoritmo 2 ObtenerRC

Entrada: Fichero 'queryResults.csv'.

Salida: RC (red de coautoría).

Tabla 4. Pseudocódigo del algoritmo 2.

1:	Se <i>carga</i> el fichero csv
2:	Se crea un grafo no dirigido
3:	Se crea un diccionario
4:	Para cada fila en el fichero hacer
5:	Separar por comas los autores de las publicaciones

6:	Crea un nodo autor
6:	Si se tiene la publicación en el diccionario entonces
7:	Se obtiene el arreglo de autores de esa publicación
8:	Para cada autor del arreglo de autores hacer
9:	Crear una arista entre el nuevo autor y los otros autores
10:	Se añade el autor al arreglo de autores
11:	Se actualiza el diccionario
12:	Si no hacer
//	si no está la publicación en el diccionario
14:	Añadir la publicación con su respectivo autor
15:	cerrar fichero

2.3.3 Cálculo de la influencia social

Después de obtener la red de coautoría se pasa al cálculo de una de las técnicas de la influencia social, la reciprocidad. Primeramente, es necesario escoger dos nodos cualesquiera para pasar a calcular su reciprocidad con respecto a toda la red de coautoría. Para ello se propone el algoritmo 3, el cual se explica a continuación.

El algoritmo **Reciprocidad** permite calcular la reciprocidad entre dos nodos cualesquiera. Al estar representada la red por un grafo no dirigido este brinda información sobre la dirección de la relación entre dos nodos, por lo que se asume que si entre A y B hay una relación entonces existe reciprocidad.

Algoritmo 3 Reciprocidad

Entrada: nodo A, nodo B

Salida: cálculo de reciprocidad

Tabla 5. Pseudocódigo del algoritmo 3.

1:	Entrar el peso de la arista de AB
2:	Entrar el grado de la arista A
3:	Entrar el peso de la arista de BA
4:	Entrar el grado de la arista B
5:	Calcular reciprocidad

La Figura 3 muestra la interfaz que conecta estos elementos y la Figura 4 muestra un ejemplo del cálculo de la reciprocidad.

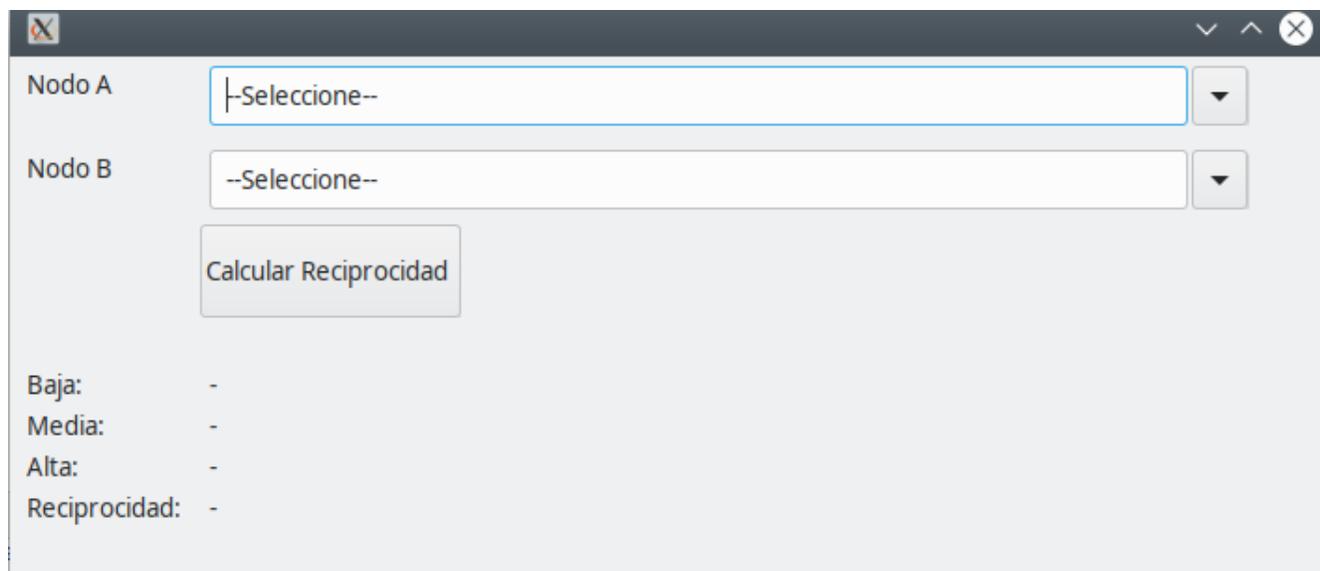


Figura 3. Prototipo de interfaz para el cálculo de la reciprocidad.

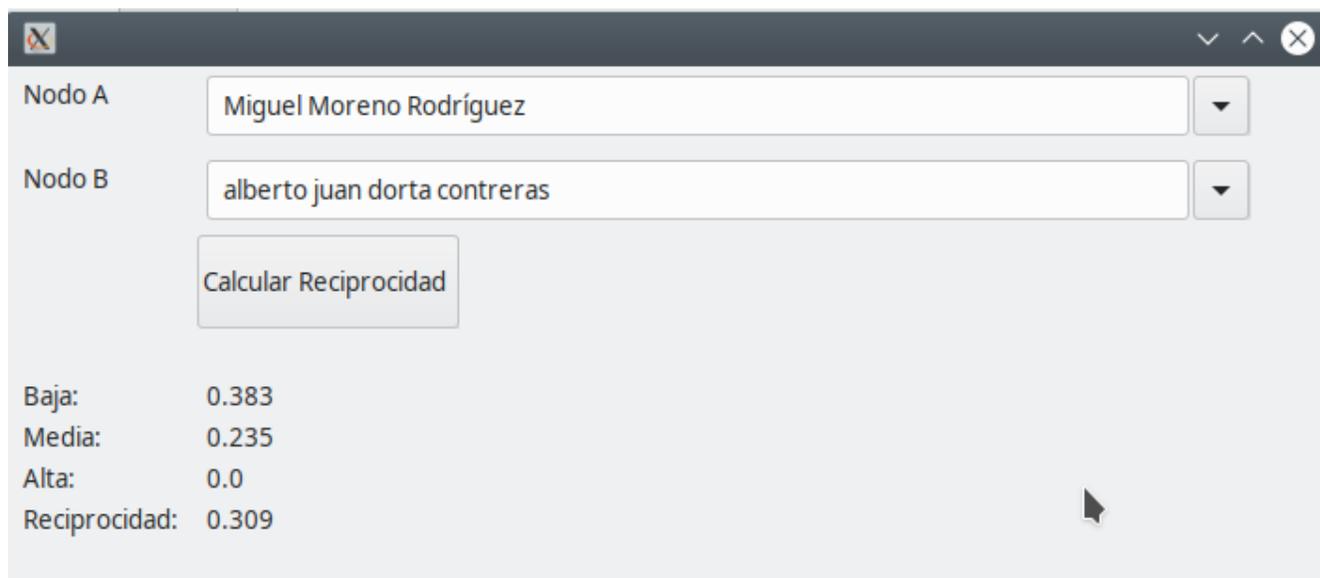


Figura 4. Ejemplo del cálculo de la reciprocidad

2.4 Representación de las redes de coautoría

Para el módulo de graficación desarrollado se utilizó la biblioteca NetworkX. En la Figura 5 se muestra la representación de la red de coautoría creada por el método. Como se puede observar en esta figura, a simple vista no se observa ninguna arista, debido a la cantidad de nodos que fueron creados. Sin embargo, al realizar un acercamiento a esta red, como se muestra en la Figura 6 se observa que hay presencia de aristas que interconectan a los nodos.

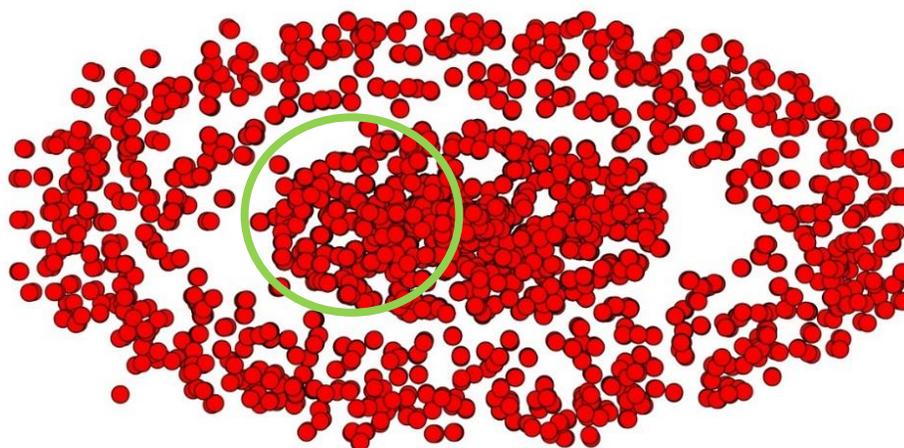


Figura 5. Representación de la red de coautoría

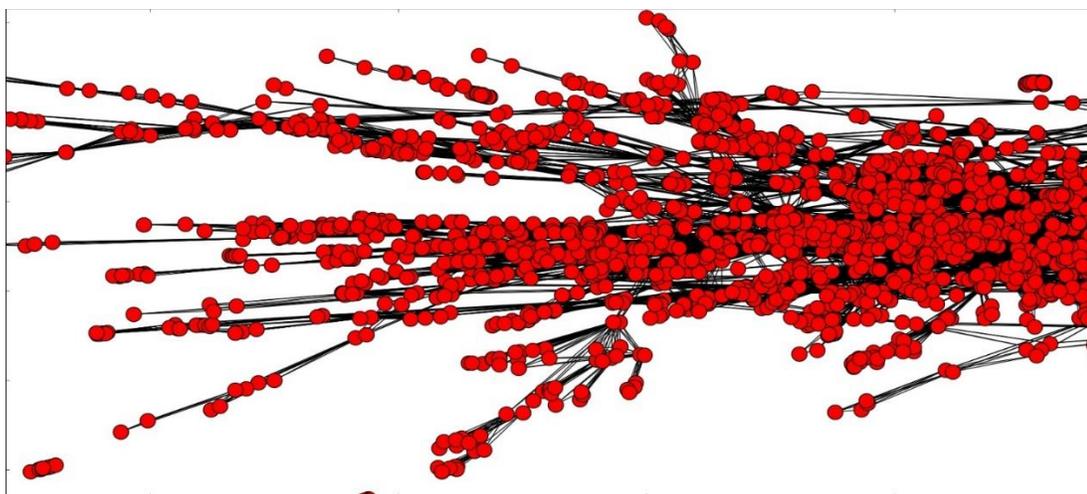


Figura 6. Representación detallada de una red de coautoría.

2.5 Clasificación del resultado del cálculo de la reciprocidad

Para la clasificación del resultado del cálculo de la reciprocidad se hace uso de la lógica difusa. La lógica difusa es definida como un sistema matemático que modela funciones no lineales, convirtiendo las entradas en salidas acordes con los planteamientos lógicos que usan el razonamiento aproximado. En cierto nivel, la lógica difusa puede ser vista como un lenguaje que permite trasladar sentencias sofisticadas en lenguaje natural a un lenguaje matemático formal (Martín del Brío y Sanz Molina 2001). Se fundamenta en los denominados conjuntos difusos y un sistema de inferencia difuso basado en Reglas de Producción (RP) de la forma "si P entonces h " o " $e \rightarrow h$ ", donde los valores lingüísticos de la

premisa y el consecuente están definidos por conjuntos difusos, es así como las reglas siempre convierten un conjunto difuso en otro (Martín del Brío y Sanz Molina 2001).

En general la lógica difusa se aplica a sistemas de control y para modelar cualquier sistema continuo de ingeniería, física, biología o economía. Las aplicaciones van desde procesos tan complejos como el control de un reactor, el movimiento de una grúa, o la modelación de la decisión de un experto (Sistemas Expertos Difusos), hasta objetos tan cercanos como son las lavadoras que tienen control automático del agua, el detergente y el desagüe, y que pueden identificarse claramente porque dicen "Fuzzy Logic". Este es quizás una de las aplicaciones más conocidas y populares de las técnicas de Inteligencia Artificial (Martín del Brío y Sanz Molina 2001).

Una variable lingüística constituye el concepto que se califica de forma difusa siendo estas el centro de las técnicas de modelado difuso. Una variable lingüística es el nombre de uno o varios conjuntos difusos. Se le aplica el adjetivo "lingüística" porque se definen sus características mediante el lenguaje hablado (Martín del Brío y Sanz Molina 2001). Ejemplos de estas variables lingüísticas son:

- "altura"
- "longitud"
- "velocidad"

Una variable lingüística encapsula las propiedades que se quiere clasificar o definir. Esto generalmente se traduce en que las variables lingüísticas son sustantivos que reflejan determinada propiedad.

Una etiqueta constituye una de las diferentes clasificaciones que se efectúan sobre la variable lingüística. Cada etiqueta se asocia por tanto a una variable lingüística y tendrá un conjunto difuso asociado (Martín del Brío y Sanz Molina 2001). Ejemplos de estas etiquetas pueden ser:

- "alto", "pequeño" o "mediano" que podrían estar asociados a una variable lingüística "altura"
- "largo" o "corto" asociados a una variable lingüística "longitud".
- "rápido", "moderado" o "lento" para la variable lingüística "velocidad".

Una etiqueta encapsula una clasificación posible de los valores de la variable lingüística. Esto generalmente se traduce en que las variables lingüísticas son adjetivos que caracterizan a un sustantivo que sería la variable lingüística (Martín del Brío y Sanz Molina 2001).

El universo de discurso constituye el rango de valores que pueden tomar los elementos que poseen la propiedad expresada por la variable lingüística. Generalmente, este universo del discurso puede venir dado por el rango de valores que toma una variable numérica para la variable lingüística. Por ejemplo, para la variable lingüística "longitud", el universo del discurso puede venir dado por los valores reales positivos, ejemplo 0.7 m, 123 m, etc. Los elementos del dominio del discurso son, por tanto, los valores numéricos que puede tomar el concepto definido como variable lingüística, y que pueden ser clasificados en cada una de las etiquetas que se han asociado a la variable lingüística en cuestión (Martín del Brío y Sanz Molina 2001).

La función de pertenencia es un elemento esencial en la definición de los conjuntos difusos. Asocia a cada elemento de un conjunto difuso el grado con que pertenece a la etiqueta asociada (Martín del Brío y Sanz Molina 2001). Existen varias funciones de pertenencia las cuales se explican a continuación.

Función de pertenencia puntual

Uno de los casos más simples para modelar evaluaciones estrictas. Un tipo de función frecuentemente usada es la lineal o puntual definida como se muestra en la Figura 7.



Figura 7. Representación de la función de pertenencia Función de pertenencia puntual. Fuente: (Martín del Brío y Sanz Molina 2001).

Formalmente quedaría:

$$M(X)= 1 \text{ si } X = A$$

$$M(X)=0 \text{ si } X \neq A$$

Saltos escalones estrictos

Esta función sirve para modelar conjuntos clásicos que se expresarían con etiquetas como las siguientes:

- Mayor o Igual que
- Menor o Igual que

La Figura 8 muestra cómo se representa esta función.

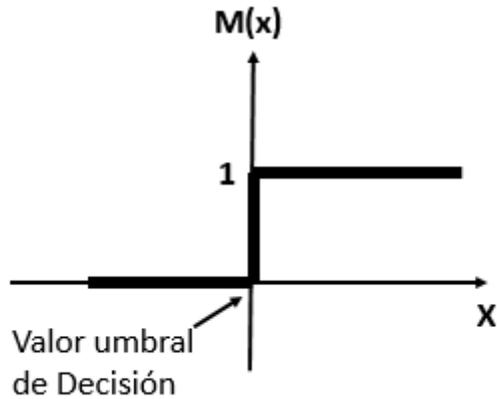


Figura 8. Representación de la función de pertenencia Saltos escalones estrictos. Fuente:(Martín del Brío y Sanz Molina 2001).

La definición formal de un escalón estricto ascendente quedaría:

1. $M(X) = 0$ si $X < \text{Umbral}$
2. $M(X) = 1$ si $X \geq \text{Umbral}$

En el caso de un escalón estricto descendente formalmente quedaría definido como:

1. $M(X) = 1$ si $X \leq \text{Umbral}$
2. $M(X) = 0$ si $X > \text{Umbral}$

Función Gamma: Escalones crecientes.

La función Gamma se emplea para las etiquetas que caracterizan los valores más altos del intervalo.

La Figura 9 muestra cómo se representa esta función.

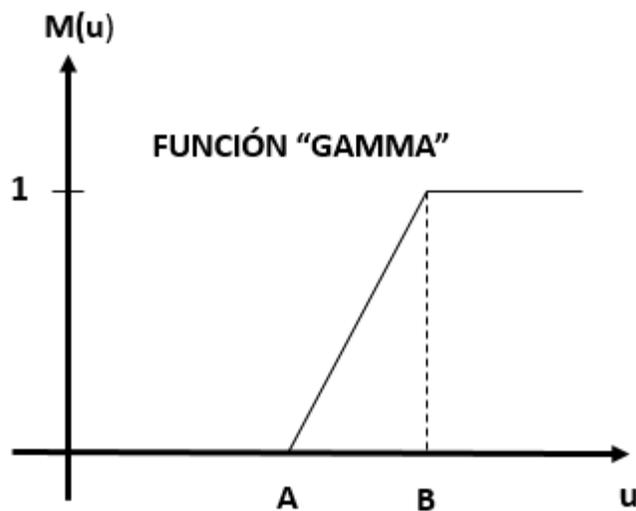


Figura 9. Representación de la función de pertenencia Función Gamma. Fuente:(Martín del Brío y Sanz Molina 2001).

Formalmente quedaría:

$$M(X) = 0 \text{ si } X \leq A$$

$$M(X) = (X-A)/(B-A) \text{ si } X > A \text{ y si } X < B$$

$$M(X) = 1 \text{ si } X \geq B$$

Función L: Escalones decrecientes

La función L se utiliza para las etiquetas que caracterizan los valores más bajos del intervalo. La Figura 10 muestra cómo se representa esta función.

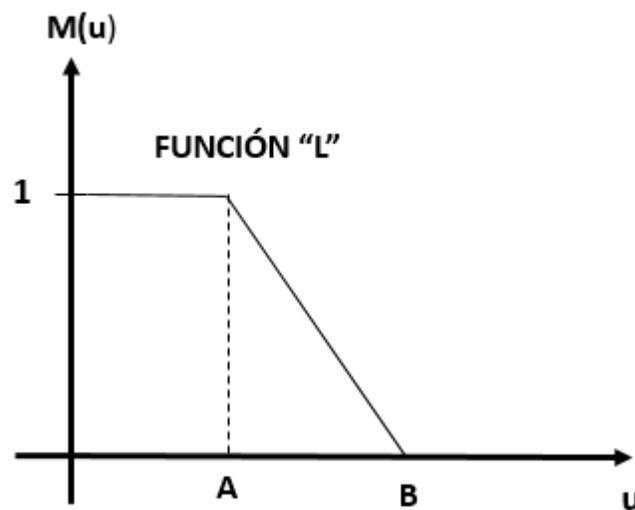


Figura 10. Representación de la función de pertenencia Función L. Fuente:(Martín del Brío y Sanz Molina 2001).

Formalmente quedaría:

$$M(X) = 1 \text{ si } X \leq A$$

$$M(X) = (B-X)/(B-A) \text{ si } X > A \text{ y si } X < B$$

$$M(X) = 0 \text{ si } X \geq B$$

Función LAMBDA: Funciones triangulares

La función LAMBDA se usa para las etiquetas que caracterizan los valores centrales del intervalo. La Figura 11 muestra cómo se representa esta función.

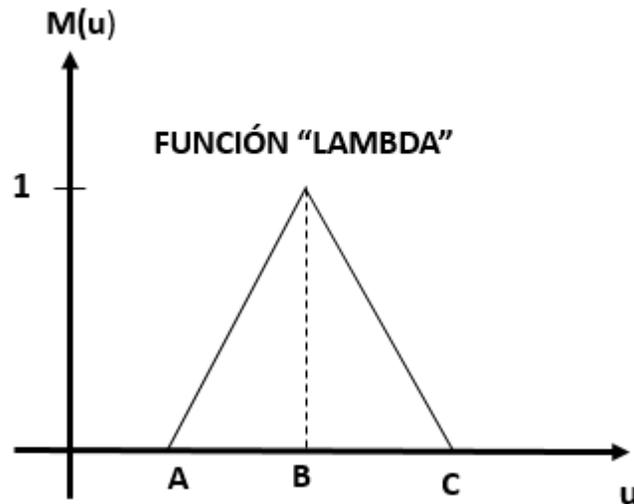


Figura 11. Representación de la función de pertenencia Función Lambda. Fuente:(Martín del Brío y Sanz Molina 2001).

Formalmente quedaría:

$$M(X) = 0 \text{ si } X \leq A$$

$$M(X) = (X-A)/(B-A) \text{ si } X > A \text{ y si } X \leq B$$

$$M(X) = (C-X)/(C-B) \text{ si } X > B \text{ y si } X < C$$

$$M(X) = 0 \text{ si } X \geq C$$

Función Pi: Funciones trapezoidales

La función Pi se usa para las etiquetas que caracterizan los valores centrales del intervalo. La Figura 12 muestra cómo se representa esta función.

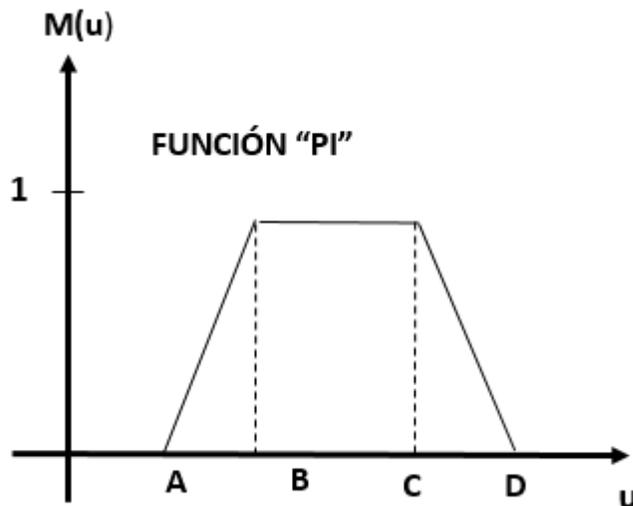


Figura 12. Representación de la función de pertenencia Función Pi. Fuente:(Martín del Brío y Sanz Molina 2001).

Formalmente quedaría:

$$M(X) = 0 \text{ si } X \leq A$$

$$M(X) = \frac{(X-A)}{(B-A)} \text{ si } X > A \text{ y si } X < B$$

$$M(X) = 1 \text{ si } X \geq B \text{ y si } X \leq C$$

$$M(X) = \frac{(D-X)}{(D-C)} \text{ si } X > C \text{ y si } X < D$$

$$M(X) = 0 \text{ si } X \geq D$$

A continuación, se presenta un ejemplo completo donde están presentes los elementos anteriores, para la caracterización de la altura de las personas.

- Variable lingüística: Altura
- Etiquetas asociadas (son tres, lo cual implica definir tres conjuntos difusos, uno para cada una de ellas):

- a. Bajo
- b. Mediano
- c. Alto

Para cada una de estas etiquetas, se muestra el resto de los elementos en el gráfico siguiente:

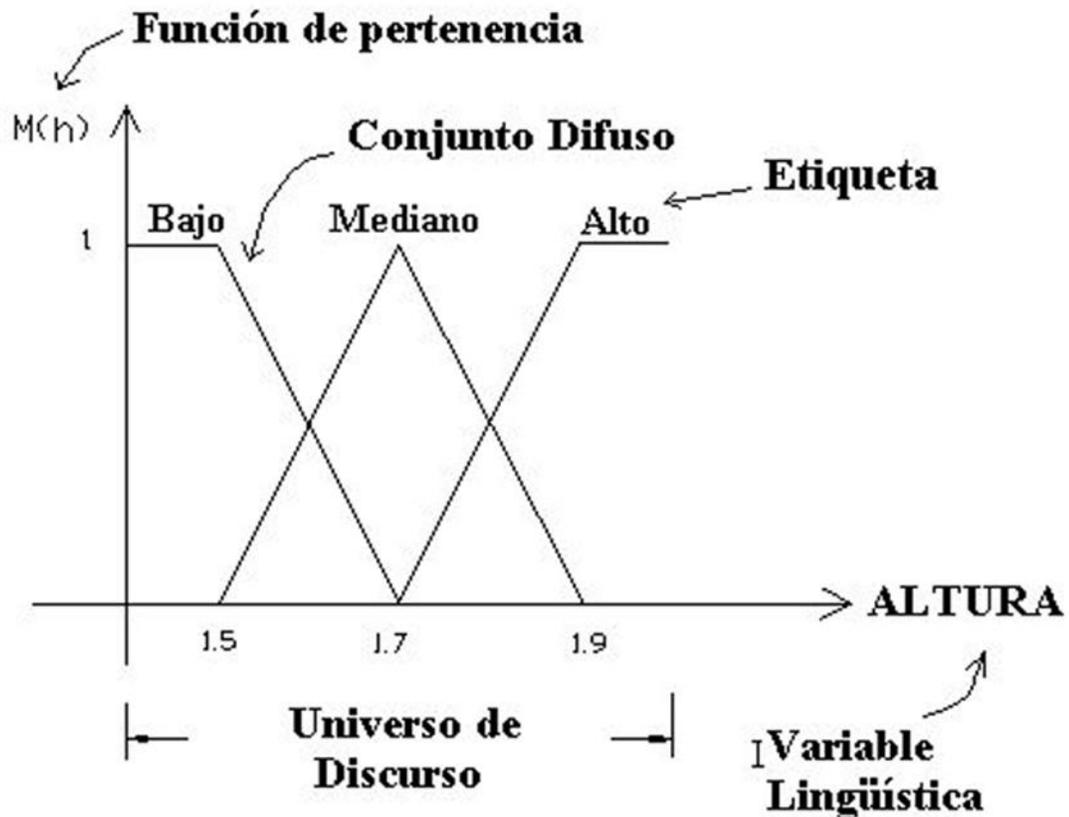


Figura 13. Ejemplo de la representación de la variable lingüística Altura.

La principal diferencia entre la lógica tradicional y la lógica difusa consiste que mientras los valores de la función de pertenencia de la primera son 0 o 1, la lógica difusa se mueve en todo el intervalo $[0,1]$. Aunque es posible trabajar con otros valores, es decir en un intervalo diferente a $[0,1]$, esto no es lo usual. Así queda expresado que mientras un elemento puede estar dentro de un determinado conjunto, puede no cumplir totalmente las especificaciones de dicho conjunto (por ejemplo, en el caso de un individuo de altura 1.8, haciendo uso de los conjuntos difusos expresados en la figura anterior, podría afirmarse que es “un poco mediano y más bien alto”). Este proceso de convertir valores numéricos en grados de pertenencias (que son de hecho muy parecidos a los valores de certeza) es una característica muy particular de los sistemas difusos.

2.5.1 Definición de la variable difusa Reciprocidad.

En la solución propuesta los conjuntos difusos definidos van a permitir determinar en qué nivel se encuentra la reciprocidad entre dos nodos cualesquiera, a partir del valor de reciprocidad calculado en el método. La Figura 14 muestra los conjuntos difusos que se proponen.

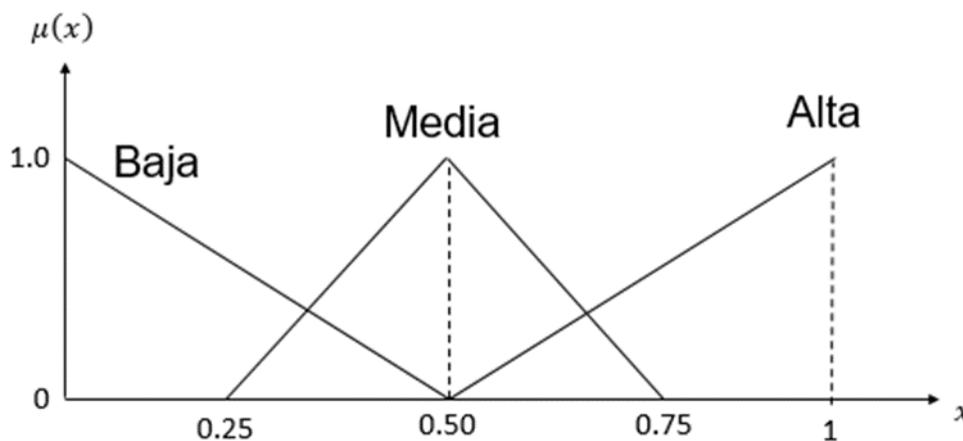


Figura 14. Conjuntos difusos de la variable lingüística Reciprocidad.

La variable lingüística Reciprocidad cuenta con tres etiquetas: Baja, Media y Alta. Su universo de discurso va desde el 0 hasta el 1, ya que los valores que se van a obtener en la reciprocidad están dentro de este rango. La función de pertenencia está definida por $M(X)$. A continuación, se describen formalmente cada uno de los conjuntos difusos asociados a cada una de las etiquetas.

Conjunto difuso para la etiqueta Baja.

Para la definición de este conjunto difuso se hizo uso de la Función de pertenencia L, la cual determina un escalón decreciente, comenzado en 0 y finalizado en 0.5. El cálculo del factor de pertenencia de

cada valor X que puede tomar la reciprocidad a la etiqueta Baja queda determinado de la siguiente forma:

$$M(X)=1 \text{ si } X=0.$$

$$M(X)=(0.5-X) / 0.5 \text{ si } X>0 \text{ y si } X<0.5$$

$$M(X)=0 \text{ si } X \geq 0.5$$

Conjunto difuso para la etiqueta Media.

Para la definición de este conjunto difuso se hizo uso de la Función de pertenencia LAMBDA o función triangular. Como su nombre indica se crea un triángulo, comenzando en 0.25, toma su mayor valor en 0.5 y termina en 0.75. El cálculo del factor de pertenencia de cada valor X que puede tomar la reciprocidad en la etiqueta Media queda determinado de la siguiente forma:

$$M(X)= 0 \text{ si } X \leq 0.25$$

$$M(X)= (X-0.25) / 0.25 \text{ si } X>0.25 \text{ y si } X \leq 0.5$$

$$M(X)= (0.75-X) / 0.25 \text{ si } X>0.5 \text{ y si } X<0.75$$

$$M(X)= 0 \text{ si } X \geq 0.75$$

Conjunto difuso para la etiqueta Alta.

Para la definición de este conjunto difuso se hizo uso de la Función de pertenencia GAMMA, la cual determina un escalón creciente, comenzando en 0.5 y terminando en 1. El cálculo del factor de pertenencia de cada valor X que puede tomar la reciprocidad en la etiqueta Alta queda determinado de la siguiente forma:

$$M(X)= 0 \text{ si } X \leq 0.5$$

$$M(X)= (X-0.5) / 0.5 \text{ si } X>0.5 \text{ y si } X<1$$

$$M(X)= 1 \text{ si } X= 1$$

Una vez calculado el factor de pertenencia a cada una de las etiquetas, son mostrados estos valores a los usuarios, como se muestra en la Figura 4, determinando como etiqueta final aquella que tenga mayor factor de pertenencia asociado a ella. En caso que haya dos factores de pertenencia iguales, en el caso de las etiquetas Baja y Media, se escogerá la etiqueta Media, y en el caso de Media y Alta se elegirá la etiqueta Alta.

Para demostrar lo anteriormente expuesto se realizarán los cálculos detallados sobre el ejemplo de la Figura 14. Luego de obtener el valor del cálculo de reciprocidad se pasa a calcular el factor de pertenencia en cada etiqueta. En el caso del valor 0.309 de reciprocidad que se presenta en la Figura 14 queda representado en los conjuntos difusos como se muestra en la Figura 15.

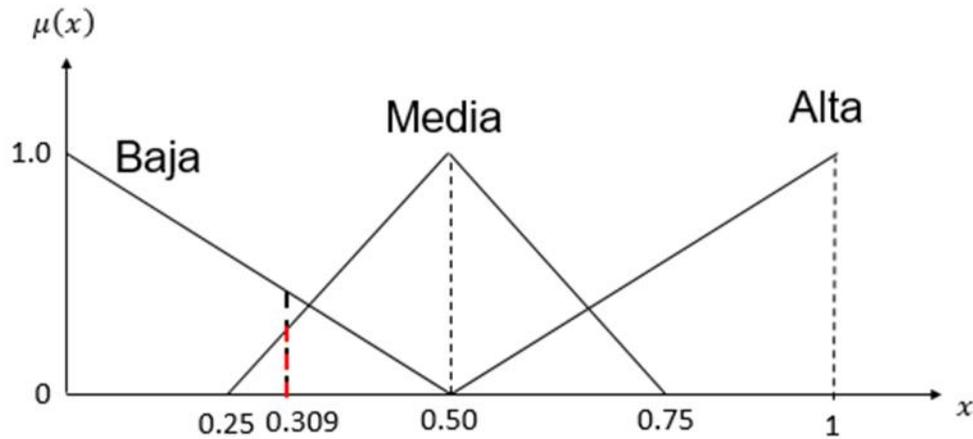


Figura 15. Representación del valor 0.309 en los conjuntos difusos de la variable lingüística Reciprocidad.

Para calcular el factor de pertenencia a la etiqueta baja, como el valor se encuentra entre 0 y 0.5 se utiliza la función:

$$M(X) = (0.5 - X) / 0.5 \text{ si } X > 0 \text{ y si } X < 0.5$$

$$M(X) = (0.50 - 0.309) / 0.50$$

$$M(X) = 0.191 / 0.50$$

$$M(X) = 0.382$$

Para el caso de la etiqueta media, como el valor se encuentra entre 0.25 y 0.5 la función que se utiliza es la siguiente:

$$M(X) = (X - 0.25) / 0.25 \text{ si } X > 0.25 \text{ y si } X \leq 0.5$$

$$M(X) = (0.309 - 0.25) / 0.25$$

$$M(X) = 0.059 / 0.25$$

$$M(X) = 0.236$$

En el caso de la etiqueta alta, como se puede observar en la figura 16 el valor 0.309 no corta en ningún momento el conjunto definido para esta etiqueta, por lo que su valor de pertenencia para ella es 0, como se determina a continuación:

$$M(X) = 0 \text{ si } X \leq 0.5$$

El grado de pertenencia más alto le corresponde al conjunto difuso de baja, por lo que se puede concluir que el nivel de reciprocidad entre los nodos Miguel Moreno Rodríguez y alberto juan dorta contreras es baja.

2.6 Conclusiones parciales

- La implementación del algoritmo **Reciprocidad** a partir del método diseñado para la inferencia de redes sociales, facilitó el estudio del comportamiento de los nodos en la red a través de la reciprocidad que unos influyen sobre otros.
- La representación de los componentes permitió modelar el funcionamiento del método propuesto, relacionando los principales conceptos asociados a la investigación.

CAPÍTULO 3. VALIDACIÓN DE LA PROPUESTA

3.1 Introducción

En el presente capítulo se realiza la validación del método desarrollado mediante varios criterios para determinar que la solución propuesta responde satisfactoriamente a los objetivos que se trazaron. Se realizan varias pruebas al código para asegurar la calidad del algoritmo propuesto verificando que no existan fallas.

3.2 Pruebas

La (IEEE 2015), define las pruebas de software como una actividad en la que un sistema o un componente es ejecutado bajo condiciones especificadas. El objetivo es diseñar una serie de casos de pruebas que tengan una alta probabilidad de encontrar errores. Para ello se aplican las técnicas de pruebas del software, las cuales facilitan una guía sistemática para diseñar pruebas que demuestren la lógica interna de los componentes de software y verifiquen los dominios de entrada y salida del programa para descubrir errores en la funcionalidad, el comportamiento y rendimiento.

Las pruebas tienen como objetivo valorar y mejorar la calidad de los productos del trabajo generado durante el desarrollo y modificación del software. Según (Pressman et al. 2010) *verificación es el conjunto de actividades que aseguran que el software implemente correctamente una función específica, y la validación es un conjunto diferente de actividades que aseguran que el software construido corresponde y satisface los requisitos del cliente.* Se tuvieron en cuenta los siguientes niveles de pruebas:

- Nivel de Unidad
- Nivel de Aceptación

Para la realización de las pruebas de software a la solución desarrollada se trabajó con las tecnologías y herramientas seleccionadas (ver epígrafe 1.5) empleando como plataforma el sistema operativo Windows 10 Pro de 64 bits. Las características de hardware sobre el que se trabajó fue, con un microprocesador Intel(R) Core(TM) i3-4005U a 1.70 GHz y una memoria RAM DDR 3 de 6GB.

3.2.1 Pruebas de unidad

El alcance de las pruebas en este nivel está dado por la intención de probar el código fuente. Las pruebas se realizaron de forma automática. Para realizar las pruebas unitarias automáticamente en el lenguaje de programación Python se utilizó el módulo “unittest” al cual se accedió mediante Anaconda2-2.5.0.

Donde ***unittest** ofrece la posibilidad de crear las pruebas en el propio código implementando una clase llamada **unittest.TestCase** cada una de las pruebas puede devolver tres respuestas en función del resultado:

- **OK**: Indica que la prueba se ha pasado exitosamente.
- **FAIL**: Indica que la prueba no ha pasado exitosamente.
- **ERROR**: Indica que la prueba no ha pasado exitosamente, pero el resultado en lugar de ser una asección es otro error.

La figura 16 muestra el resultado de aplicar las pruebas unitarias de forma automática al código implementado:

```

1 import unittest
2 def dfs(url):
3     f=open(url)
4     return f
5 class Test_URL_dfs(unittest.TestCase):
6     def test_dfs2(self):
7         self.assertIsNotNone(dfs('D:/UNIVERSIDAD/tesis/new.csv'))
8
9 if __name__ == "__main__":
10     unittest.main()
11
12

```

Figura 16 Código de la prueba unitaria 1 del algoritmo implementado (Fuente: Elaboración propia)

```

Python 2.7.11 [Anaconda 2.5.0 (64-bit)] (default, Jan 29 2016, 14:26:21) [MSC v.1500 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
Anaconda is brought to you by Continuum Analytics.
Please check out: http://continuum.io/thanks and https://anaconda.org
>>> runfile('C:/Users/Hany/Desktop/prueba 1.py', wdir='C:/Users/Hany/Desktop')
.
-----
Ran 1 test in 0.001s

OK

```

Figura 17 Cuadro de respuesta de la prueba unitaria 1 (Fuente: Elaboración propia)

```

1 from __future__ import division
2 import unittest
3
4 def calcular(e_a,e_b,e_ab,e_ba):return (e_ab/e_a)*(e_ba/e_b)
5 class reciprocity(unittest.TestCase):
6     def test_reciprocity(self):
7         self.assertEqual(calcular(5,9,4,4),0.35555555555555557)
8
9
10 if __name__ == "__main__":
11     unittest.main()
12

```

Figura 18 Código de la prueba unitaria 2 del algoritmo implementado (Fuente: Elaboración propia)

```
Python 2.7.11 [Anaconda 2.5.0 (64-bit)] (default, Jan 29 2016, 14:26:21) [MSC v.1500 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
Anaconda is brought to you by Continuum Analytics.
Please check out: http://continuum.io/thanks and https://anaconda.org
>>> runfile('C:/Users/Hany/Documents/prueba 2.py', wdir='C:/Users/Hany/Documents')
.
-----
Ran 1 test in 0.000s

OK
```

Figura 19 Cuadro de respuesta de la prueba unitaria 2 (Fuente: Elaboración propia)

Para la presente investigación se realizaron un total de 2 pruebas unitarias de las cuales no se obtuvieron no conformidades logrando así un 100% de resultados satisfactorios en la primera iteración.

3.2.2 Pruebas de aceptación

Según (IEEE 2008) las pruebas de aceptación son realizadas para establecer si un sistema satisface sus criterios de aceptación y para permitirle al cliente determinar si acepta el sistema. También se define como una prueba formal realizada para permitir que un usuario, cliente u otra entidad autorizada determine si acepta un sistema o componente.

Para realizar las mencionadas pruebas fue empleado el mismo principio que se utilizó en la realización de las pruebas internas, permitiéndole al cliente que pueda realizar las mismas. Para esto, fueron realizadas 2 iteraciones y los resultados obtenidos se evidencian a continuación:

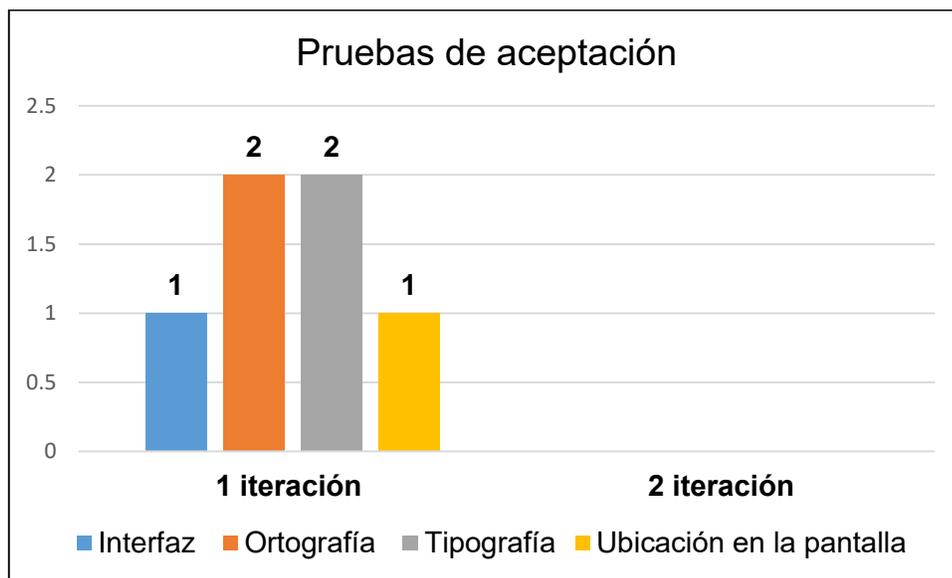


Figura 20. Resultados de las pruebas de aceptación

Se realizaron dos iteraciones, en la primera iteración el cliente detectó problemas de interfaz, ortografía, tipografía y ubicación en la pantalla, quedando resueltas en la segunda iteración, de las cuales no se obtuvo no conformidades, logrando así un 100 por ciento de resultados satisfactorios.

3.3 Utilización de un sociograma para la validación del método propuesto

Para la validación del método implementado se aplicó una técnica tradicional de la sociología: el sociograma. Se conoce como sociograma a la técnica que, a través de la observación y la evaluación de un contexto, refleja en un gráfico los diferentes vínculos entre los miembros de un grupo. De esta manera, logra dejar evidenciado los grados de influencia y los lazos de preferencia que se presentan en él. Los sociogramas son recursos útiles para graficar la estructura de una organización, ya sea grande o pequeña (Ander-Egg 1982).

El sociograma está considerado como una técnica de carácter sociométrico, es decir, un método que permite medir las relaciones sociales entre los integrantes de un grupo humano, donde sus elementos se conocen, poseen objetivos comunes e influyen mutuamente. A nivel gráfico, un sociograma representa las relaciones interpersonales por medio de nodos (los individuos) que aparecen unidos por una o más aristas (las relaciones interindividuales).

El gráfico obtenido permite apreciar afinidades, detectar subgrupos dentro del conjunto principal y ubicar a los líderes sociométricos (las personas más influyentes). De esta forma, es posible actuar sobre el grupo para activar vínculos potenciales o desactivar aquellos existentes.

Todos estos conocimientos se consiguen a través de la elaboración de una serie de preguntas que conformarán un cuestionario. En concreto, esas cuestiones serán de dos tipos. Por un lado, estarán las intelectuales que son las que se refieren a la actividad productiva de la clase y por otro, estarán las afectivas que se refieren a las elecciones que realizarían las personas para desarrollar actividades de tipo lúdico.

Entre las preguntas intelectuales se encontrarían algunas como “¿A quién elegirías para realizar deporte?”. Mientras, una muestra de lo que son las cuestiones afectivas sería la siguiente: “¿A qué compañero elegirías para estudiar?”.

El sociograma es una herramienta muy usual en el campo de la Sociología. Con estos tipos de métodos, esta ciencia investiga las relaciones que los sujetos mantienen entre sí y con el sistema, y el grado de cohesión que existe en la estructura social. Los métodos de la sociología pueden ser cualitativos (con descripciones detalladas de situaciones, comportamientos y personas) o cuantitativos (que se encarga de las características y variables que pueden ser representadas por valores numéricos). El objetivo de la aplicación de esta técnica en el presente trabajo es verificar la utilidad y el nivel de exactitud que presenta el método implementado en el algoritmo (Ander-Egg 1982).

3.4 Aplicación del sociograma

Para la aplicación del sociograma se realizó un sistema de encuesta el cual se les aplicó a los diez usuarios detectados con los mayores niveles de reciprocidad en la biblioteca de la Univesidad de las Ciencias Informáticas. La encuesta contó con 3 preguntas, cada pregunta tenía como objetivo que cada

uno de los diez usuarios escogiera a 3 de sus compañeros para la realización de determinadas actividades (ver anexo 1). El objetivo final ha sido detectar por quien se sentían más influenciados en cada uno de los escenarios encuestados. El sistema de preguntas fue desarrollado por un sociólogo de la Universidad de las Ciencias Informáticas.

3.5 Correlación entre los resultados del sociograma y los resultados del método propuesto

Se entiende como correlación al grado de relación existente entre dos variables, en el contexto de la investigación y para su análisis éstas van a estar representadas por el sociograma y el método computacional. Para verificar la correspondencia de los resultados obtenidos con el sociograma y con el método propuesto se aplicaron dos pruebas: el coeficiente de correlación de Spearman y el coeficiente de correlación de Pearson. Se aplicó también un diagrama de dispersión para poder observar de forma gráfica la dispersión de los datos (Capote y Viera 2014).

3.5.1 Coeficiente de correlación de Spearman

El coeficiente de correlación de Spearman permite medir la relación entre atributos ordenados. Se aplica para obtener correlaciones entre ordenaciones. Sea A y B dos atributos en escala ordinal y a_1, \dots, a_n y b_1, \dots, b_n sus diferentes categorías. Se suponen ordenadas estas categorías de menor a mayor. Se denomina rango a la posición de cada categoría en la muestra ordenada. Los rangos toman entonces los valores $\{1, 2, \dots, n\}$. Se denota por x_i e y_i , $i = 1, 2, \dots, n$ los rangos correspondientes a los dos atributos (Sarabia Alegría y Pascual Saéz 2005).

Se denomina coeficiente de correlación ordinal de Spearman o coeficiente de correlación por rango de Spearman al coeficiente de correlación lineal entre los dos rangos (Sarabia Alegría y Pascual Saéz 2005).

Si se suponen dos atributos A y B en escala ordinal, cuyos valores han sido transformados en los rangos $(x_i, y_i), i = \{1, 2, \dots, n\}$, el coeficiente de correlación de rangos de Spearman viene dado por:

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

Fórmula 2. Coeficiente de correlación de Spearman (Sarabia Alegría y Pascual Saéz 2005).

Donde $d_i = x_i - y_i, i = 1, 2, \dots, n$

Al tratarse de un coeficiente de correlación su valor varía entre -1 y +1. Si la concordancia entre los rangos es perfecta entonces $d_i = 0, i = 1, 2, \dots, n$ y por lo tanto $r = 1$. Cuando la discordancia entre los rangos es perfecta entonces $r = -1$. Si los dos criterios de clasificación son independientes entonces $r = 0$ (Sarabia Alegría y Pascual Saéz 2005).

3.5.2 Coeficiente de correlación lineal o Coeficiente de Pearson

El coeficiente de correlación lineal o Coeficiente de Pearson es una medida de dependencia lineal adimensional entre dos variables cuantitativas x e y (Sarabia Alegría y Pascual Saéz 2005). La cual se define como:

$$r = \frac{S_{xy}}{S_x * S_y}$$

Fórmula 3. Coeficiente de correlación de Pearson (Sarabia Alegría y Pascual Saéz 2005).

Donde S_{XY} representa la covarianza y S_x y S_y las desviaciones típicas de x e y , respectivamente.

Propiedades del Coeficiente de correlación de Pearson:

- ❖ El coeficiente de correlación lineal es adimensional, es decir, carece de unidades de medida.
- ❖ Es invariante frente a transformaciones lineales en los datos. En concreto, si a partir de una variable bidimensional (x, y) se construye una nueva variable $(a_1 + b_1X, a_2 + b_2Y)$ se verifica que $(b_1, b_2 = 0)$: $r(a_1 + b_1X, a_2 + b_2Y) = \text{signo}(b_1 * b_2) * r(X, Y)$
- ❖ Es acotado: $-1 \leq r \leq 1$.

Cuando $|r|$ está cercano a 1, existe una relación aproximadamente lineal entre las variables X e Y . Si el valor de r está cercano a 0, no existwe relación de tipo lineal entre las variables (Sarabia Alegría y Pascual Saéz 2005).

3.5.3 Diagrama de dispersión

Un diagrama de dispersión es una representación gráfica de la relación entre dos variables, muy utilizada en las fases de comprobación de teorías e identificación de causas raíz en el diseño de soluciones y mantenimiento de los resultados obtenidos. Estudios revelan que resulta más fácil identificar las relaciones en un diagrama de dispersión que en una tabla de números.

El análisis de un diagrama de dispersión consta de un proceso de cuatro pasos: se elabora una teoría razonable, se obtienen los pares de valores y se dibuja el diagrama, se identifica la pauta de correlación y se estudian las posibles explicaciones. Las pautas de correlación más comunes son: correlación fuerte positiva (Y aumenta claramente con X), correlación fuerte negativa (Y disminuye claramente con X), correlación débil positiva (Y aumenta algo con X), correlación débil negativa (Y disminuye algo con X), correlación compleja (Y parece relacionarse con X pero no de un modo lineal) y correlación nula (no hay relación entre X e Y) (Escalante et al. 2005).

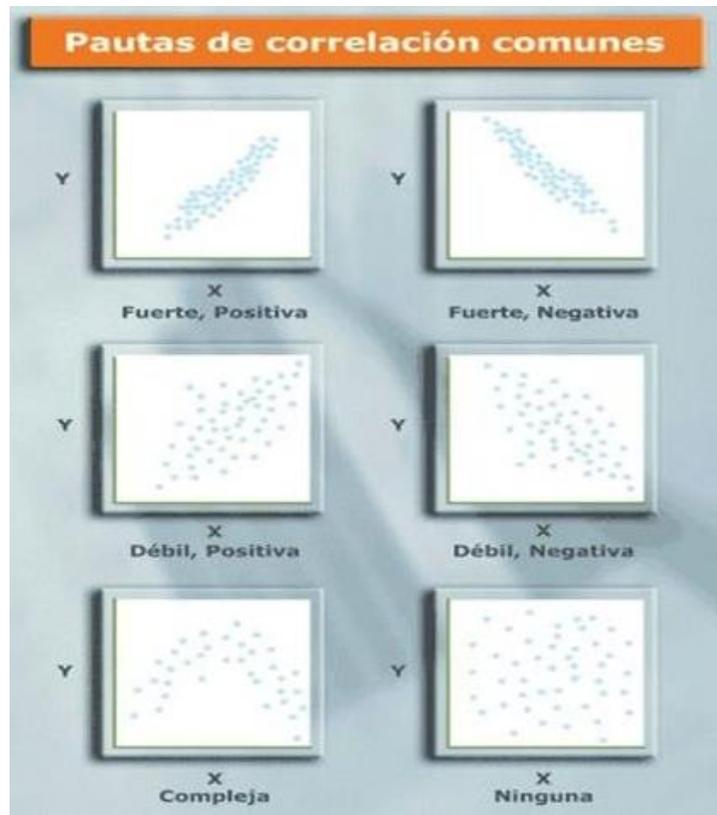


Figura 21. Diagrama de dispersión (Escalante et al. 2005).

3.6 Aplicación del coeficiente de correlación de Spearman y coeficiente de Pearson

Para la aplicación del coeficiente se definieron las siguientes premisas: si se obtiene un valor de correspondencia mayor o igual que 0,6 y menor que 0,8, entonces el valor de correspondencia es adecuado; si se obtiene un valor de correspondencia mayor o igual que 0,8, entonces el valor de correspondencia es muy confiable (Capote y Viera 2014).

$0.6 \leq x < 0.8$ Adecuado

$x \geq 0.8$ Muy confiable

Para validar el supuesto de que al analizar varios contextos de interacción se eleva la precisión de los análisis del comportamiento de los individuos, se aplicó el coeficiente de correlación de Spearman y de Pearson a cada uno de los contextos por separado y luego se procede a la unión de todos ellos.

Las tablas 6, 7 y 8 fueron desarrolladas para aplicar el coeficiente de Spearman y Pearson en tres contextos diferentes. La primera columna representa los usuarios seleccionados como muestra para realizar la prueba. La segunda columna, los valores asociados a cada usuario obtenidos al aplicar la técnica del sociograma. La tercera, los valores obtenidos del componente desarrollado. X e Y representan los datos ordenados por rango, X el orden de los valores del complemento e Y el orden de los valores del sociograma.

Contexto Fiesta.

Tabla 6. Coeficiente de correlación del registro de eventos Fiesta.

ID	Sociograma	Método	X	Y	$D = x - y$	D^2
O	7	0.2178	7.5	2.5	5	25
Y	11	0	2.5	7	-4.5	20.25
S	7	0	2.54	2.5	0.04	0.0016
M	8	0.3267	10	4.5	5.5	30.25
B	8	0	2.5	4.5	-2	4
H	13	0.2178	7.5	9	-1.5	2.25
S	12	0.2178	7.5	8	-0.5	0.25
A	11	0.1089	5	7	-2	4
D	11	0.2178	7.5	7	0.5	0.25
M	2	0	2.5	1	1.5	2.25
Total						88.5016

Para los datos obtenidos en el contexto Fiesta se obtuvo un valor de correspondencia de Spearman de 0.46362667 y un valor de correspondencia Pearson de 0.32329885.

Contexto Deporte

Tabla 7. Coeficiente de correlación del registro de eventos Deporte.

ID	Sociograma	Método	X	Y	$D = x - y$	D^2
O	7	0.1089	6	2.5	3.5	12.25
Y	11	0.1089	6	7	-1	1
S	7	0.3267	10	2.5	7.5	56.25
M	8	0	2.5	4.5	-2	4
B	8	0	2.5	4.5	-2	4
H	13	0.2178	8.5	9	-0.5	0.25
S	12	0.1089	6	8	-2	4
A	11	0	2.5	7	-4.5	20.25
D	11	0.2178	8.5	7	1.5	2.25
M	2	0	2.5	1	1.5	2.25
Total						106.5

Para los datos obtenidos en el contexto Deporte se obtuvo un valor de correspondencia de Spearman de 0.35454545 y un valor de correspondencia Pearson de 0.21896278.

Contexto de Investigación

Tabla 8. Coeficiente de correlación del registro de eventos Investigación.

ID	Sociograma	Método	X	Y	$D = x - y$	D^2
O	7	0	3.5	2.5	1	1
Y	11	0.2178	8	7	1	1
S	7	0	3.5	2.5	1	1
M	8	0	3.5	4.5	-1	1
B	8	0.2178	8	4.5	3.5	12.25
H	13	0	3.5	9	-5.5	30.25
S	13	0	3.5	8	-4.5	20.25
A	11	0.3267	10	7	3	9
D	11	0	3.5	7	-3.5	12.25
M	2	0.2178	8	1	7	49
Total						137

Para los datos obtenidos en el contexto Estudio se obtuvo un valor de correspondencia de Spearman de 0.169697 y un valor de correspondencia Pearson de -0.059727.

Tabla 9. Coeficiente de correlación de la unión de todos los registros de eventos

ID	Sociograma	Método	X	Y	$D = x - y$	D^2
O	7	0.4356	3	2.5	0.5	0.25
Y	11	0.8712	10	7	3	9
S	7	0.5445	6	2.5	3.5	12.25
M	8	0.4356	3	4.5	-1.5	2.25
B	8	0.4356	3	4.5	-1.5	2.25
H	13	0.7625	9	9	0	0
S	12	0.5445	6	8	-2	4
A	11	0.6534	8	7	1	1
D	11	0.5445	6	7	-1	1
M	2	0.2178	1	1	0	0
Total						32

Al aplicar el método de Spearman se obtuvo un valor de correspondencia de 0.80606061 y al aplicar el método de Pearson se obtuvo un valor de correspondencia de 0.70502613.

Tabla 10. Coeficiente de correlación entre el método tradicional y el método computacional de todos los registros de eventos

ID	Sociograma	Método computacional	X	Y	$D = x - y$	D^2
O	0.23	0.37	2	3	-1	1
Y	0.36	0.46	7	6.5	0.5	0.25
S	0.21	0.26	1	2	-1	1
M	0.26	0.41	4	4	0	0
B	0.27	0.45	6	5	1	1
H	0.42	0.4	3	10	-7	49
S	0.4	0.56	8.5	9	-0.5	0.25
A	0.36	0.56	8.5	6.5	2	4
D	0.38	0.62	10	8	2	4
M	0.06	0.43	5	1	4	16
Total						76.5

Al aplicar el método de Spearman se obtuvo un valor de correspondencia de 0.53636364 y al aplicar el método de Pearson se obtuvo un valor de correspondencia de 0.53353659.

Se puede concluir que según el valor de correspondencia de Spearman y el de Pearson la correlación del sociograma con respecto a la del método desarrollado en cada contexto por separado, es menor que la correlación de todos los registros de eventos unidos.

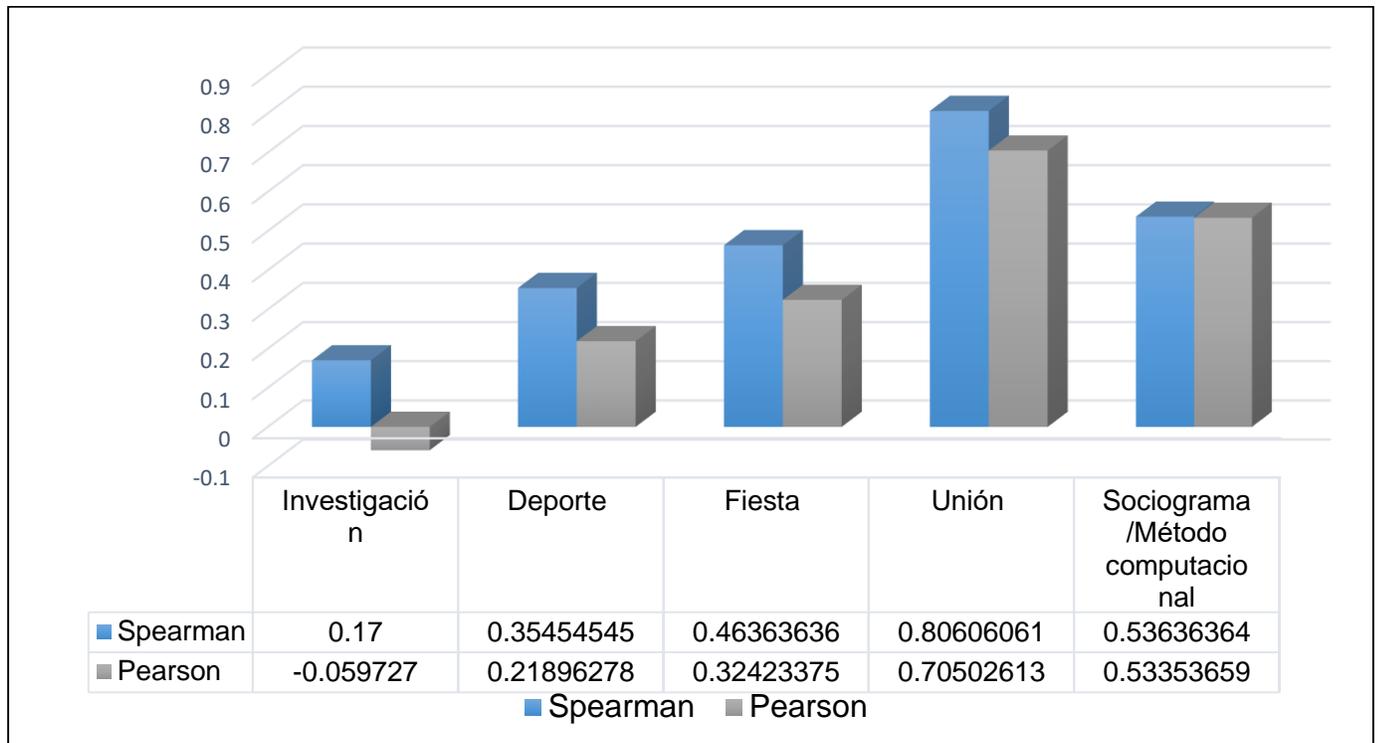


Figura 22. Comparación de los resultados obtenidos de correlación.

Avalándose que: si se desarrolla un método para la inferencia de redes de interacción social a partir de las actividades realizadas por los usuarios, sus datos personales e institucionales, entonces se eleva la precisión de los análisis del comportamiento de los individuos en sus contextos de interacción.

Diagrama de dispersión

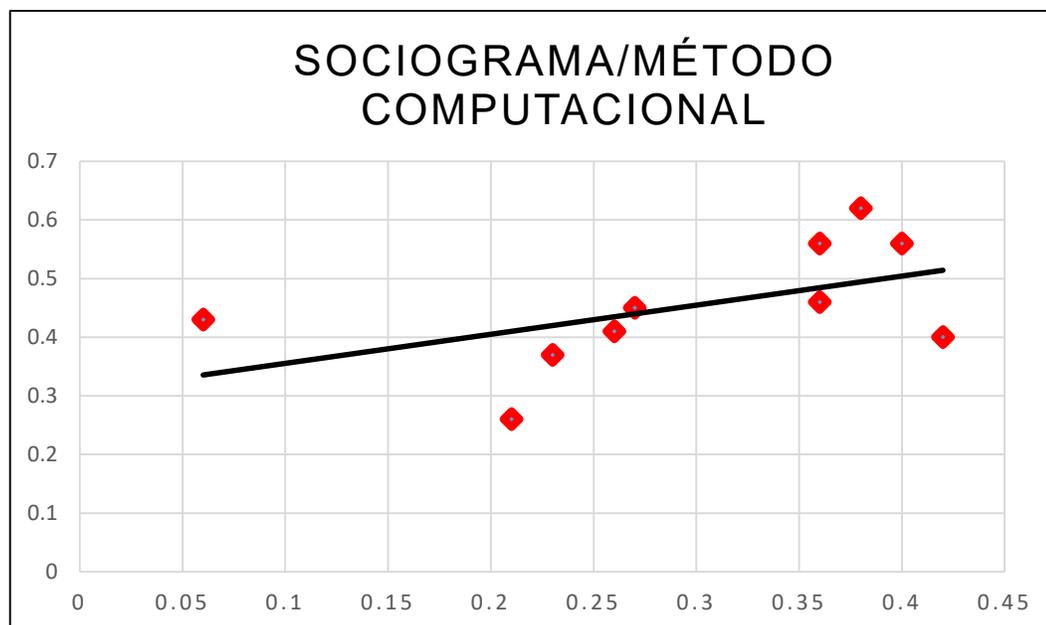


Figura 23. Diagrama de dispersión sociograma / método computacional

Al observar la gráfica se puede concluir que el sistema no es linealmente correlacionado se buscan otras variantes como por ejemplo la polinómica.

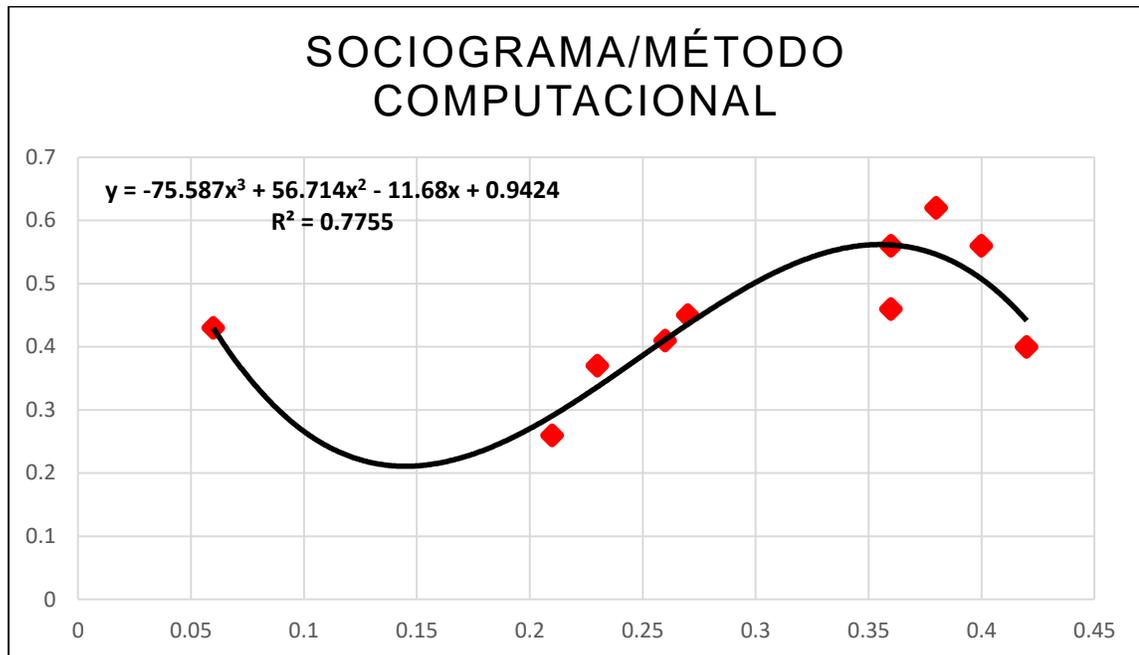


Figura 24. Diagrama de dispersión sociograma / método computacional (Polinómica).

Al observar la gráfica se puede concluir que la dispersión obtenida presenta una correlación polinómica de grado 3, que está descrito por la función $y = -75.587x^3 + 56.714x^2 - 11.68x + 0.9424$, obteniéndose un coeficiente de correlación $R^2 = 0.7755$ este resultado es capaz de distinguir un 77.55 por ciento de la aparición de los datos, por lo tanto teniendo un valor de correlación de 0.88062478, por lo que se concluye que el valor de correspondencia es muy confiable y según las pautas de correlación después de haber dibujado el diagrama de dispersión se puede clasificar en fuerte positiva donde las Y aumenta claramente con X, quiere decir que son directamente proporcionales.

3.7 Conclusiones parciales

- La validación realizada a través de pruebas unitarias y de aceptación permitieron corregir las no conformidades detectadas, probando la calidad de la solución propuesta.
- Se aplicó una técnica tradicional de carácter sociométrico para evaluar el algoritmo implementado, evidenciando los grados de influencia y lazos de preferencia que se representan en la red.
- La aplicación del coeficiente de correlación de Spearman y el coeficiente de correlación de Pearson verificó la correspondencia entre los datos obtenidos con el sociograma y aquellos mediante el método desarrollado.

CONCLUSIONES GENERALES

Al concluir la presente investigación se puede arribar a las siguientes conclusiones que dan solución al objetivo general de la misma:

- El análisis de las principales fuentes bibliográficas permitió identificar los aspectos que caracterizan la influencia social en la red, y el rol que desempeñan las relaciones recíprocas entre actores.
- La implementación del método propuesto permitió calcular la inferencia de redes de interacción social a partir de las relaciones recíprocas entre actores, lo cual facilitó el estudio del comportamiento de la influencia en la red.
- La aplicación experimental del método propuesto en un caso de estudio y el empleo de técnicas sociométricas y pruebas estadísticas para evaluar los resultados, permitió verificar la correspondencia entre los datos obtenidos con el sociograma y aquellos mediante el método desarrollado, lo cual demostró su efectividad.

Por tanto, el método propuesto para la inferencia de redes de interacción social a partir de las relaciones recíprocas entre los actores contribuye sustancialmente al estudio del comportamiento de la influencia social en una red.

REFERENCIAS BIBLIOGRÁFICAS

- AALST, W. van der, 2011. *Process mining: discovery, conformance and enhancement of business processes*. Berlin: Springer. ISBN 978-3-642-19344-6.
- AGARWAL, N. y SINGH, L., 2014. Process Mining Tools: A comparative Analysis and Review. *Advances in Computer Science and Information Technology*, vol. 1, no. 2, pp. 4.
- AGUIRRE, J.L., 2011. *J. Aguirre. Introducción al Análisis de Redes Sociales.pdf*. 2011. S.l.: s.n.
- ALEXANDER, R.D., 2017. *The biology of moral systems*. Hawthorne, N.Y: A. de Gruyter. Foundations of human behavior. ISBN 978-0-202-01173-8. BJ58 .A43 1987
- ANACONDA, 2019. Anaconda (distribución de Python). En: Page Version ID: 115607635.
- ANDER-EGG, E., 1982. *Técnicas de Investigación Social.pdf*. Buenos Aires: Universidad Nacional Abierta. Dirección de Investigaciones y Postgrado: s.n.
- AZIZIFARD, N., 2014. Social Network Clustering. *International Journal of Information Technology and Computer Science* [en línea], vol. 6, no. 1, pp. 76-81. [Consulta: 7 diciembre 2018]. ISSN 20749007, 20749015. DOI 10.5815/ijitcs.2014.01.09. Disponible en: <http://www.mecspress.org/ijitcs/ijitcs-v6-n1/v6n1-9.html>.
- BALL, B. y NEWMAN, M.E.J., 2013. Friendship networks and social status. *Network Science* [en línea], vol. 1, no. 01, pp. 16-30. [Consulta: 7 diciembre 2018]. ISSN 2050-1242, 2050-1250. DOI 10.1017/nws.2012.4. Disponible en: http://www.journals.cambridge.org/abstract_S2050124212000045.
- BORNER, K., MARU, J.T. y GOLDSTONE, R.L., 2004. The simultaneous evolution of author and paper networks. *Proceedings of the National Academy of Sciences* [en línea], vol. 101, no. Supplement 1, pp. 5266-5273. [Consulta: 7 diciembre 2018]. ISSN 0027-8424, 1091-6490. DOI 10.1073/pnas.0307625100. Disponible en: <http://www.pnas.org/cgi/doi/10.1073/pnas.0307625100>.
- CABEZAS, Dr.P.C. y ESTEBAN, Dra.E.G., 2012. *Racionalidad económica: Aspectos éticos de la reciprocidad*. S.l.: s.n.
- CAPOTE, R.R. y VIERA, T.S., 2014. *Título: Algoritmo para la inferencia de redes sociales a partir de las actividades realizadas por los usuarios en diferentes contextos de interacción*. S.l.: s.n.
- CARTER T. BUTTS, 2014. *Tools for Social Network Analysis.pdf*. S.l.: s.n.
- CASTRO, R.D. y JERROLD W. GROSSMAN, 1999. *Famous trails to Paul Erdos.pdf*. 4 julio 1999. S.l.: s.n.
- CHAMBERS , Y OTROS, 2018. *R Data Import/Export*. S.l.: s.n.
- CHIAVENATO, I., 2006. *Introducción a la teoría general de la administración*. México, D.F.: McGraw-Hill. ISBN 978-970-10-5500-7.

- CIALDINI PH.D., R.B., 2009. Influence: The Psychology of Persuasion. , pp. 279.
- CIALDINI, R.B., 1990. *Influencia: ciencia y práctica : cuáles son los factores determinantes para que una persona diga sí a otra persona*. Madrid: Prodilu. ISBN 978-84-87736-00-1.
- DONETTI, L. y MUÑOZ, M.A., 2008. Detecting network communities: a new systematic and efficient algorithm. *Journal of Statistical Mechanics: Theory and Experiment* [en línea], vol. 2004, no. 10, pp. P10012. [Consulta: 7 diciembre 2018]. ISSN 1742-5468. DOI 10.1088/1742-5468/2004/10/P10012. Disponible en: <http://stacks.iop.org/1742-5468/2004/i=10/a=P10012?key=crossref.151f49bf83774dafd01f3d759a70e139>.
- ESCALANTE, G.A.F., PARADA, M.J.R., MARTINEZ, M.R., MENDOZA, N.J.S., MORALES, J.C.S. y MUÑOZ, V.S., 2005. INFORME FINAL QUE PARA OBTENER EL TITULO DE CONTADOR PUBLICO PRESENTAN: , pp. 151.
- GÓMEZ, S., 2013. ANTROPOLOGÍA ECONÓMICA. [en línea]. [Consulta: 29 noviembre 2018]. Disponible en: <https://www.docsity.com/es/antropologia-economica-apuntes-antropologia-social/340379/>.
- GONZÁLEZ DUQUE, R., 2015. *Python para todos.pdf* [en línea]. S.l.: s.n. Disponible en: <http://mundogeek.net/tutorial-python/>.
- HE, B., DING, Y. y NI, C., 2011. Mining enriched contextual information of scientific collaboration: A meso perspective. *Journal of the American Society for Information Science and Technology* [en línea], vol. 62, no. 5, pp. 831-845. [Consulta: 7 diciembre 2018]. ISSN 15322882. DOI 10.1002/asi.21510. Disponible en: <http://doi.wiley.com/10.1002/asi.21510>.
- IEEE, 2008. IEEE Std 754™-2008 (Revision of IEEE Std 754-1985), IEEE Standard for Floating-Point Arithmetic. , pp. 70.
- IEEE, 2015. *ieee_annual_report.pdf*. S.l.: s.n.
- INFORME DE LA UNESCO SOBRE LA CIENCIA 2010, 2010. *INFORME DE LA UNESCO SOBRE LA CIENCIA 2010 (1).pdf*. 2010. S.l.: s.n.
- JETBRAINS, 2018. PyCharm Edu: The Python IDE to Learn Programming Quickly & Efficiently. *JetBrains* [en línea]. [Consulta: 21 mayo 2019]. Disponible en: <https://www.jetbrains.com/pycharm-edu/>.
- LAPUENTE, M.J.L. y LAPUENTE, C.L., 2018. RDF. [en línea]. [Consulta: 21 mayo 2019]. Disponible en: <http://www.hipertexto.info/documentos/rdf.htm>.
- LASSO-SAMBONY, E.G., ORTEGA-PONCE, S.M. y CORRALES, J.C., 2013. Enriquecimiento semántico e inferencia de relaciones en una red social en línea2. , pp. 21.
- LIU, X., BOLLEN, J., NELSON, M.L. y VAN DE SOMPEL, H., 2005. Co-authorship networks in the digital library research community. *Information Processing & Management* [en línea], vol. 41, no. 6, pp. 1462-1480. [Consulta: 7 diciembre 2018]. ISSN 03064573. DOI 10.1016/j.ipm.2005.03.012. Disponible en: <http://linkinghub.elsevier.com/retrieve/pii/S0306457305000336>.
- LÓPEZ-SÁEZ, M., 2018. *7.1_influencia_social_y_tacticas_de_influencia.doc*. 7 diciembre 2018. S.l.: s.n.

- MARTÍN DEL BRÍO, B. y SANZ MOLINA, A., 2001. *Redes Neuronales y Sistemas Difusos.pdf*. 2001. S.l.: s.n.
- MARTIN, T., BALL, B., KARRER, B. y NEWMAN, M.E.J., 2013. Coauthorship and citation patterns in the Physical Review. *Physical Review E* [en línea], vol. 88, no. 1. [Consulta: 7 diciembre 2018]. ISSN 1539-3755, 1550-2376. DOI 10.1103/PhysRevE.88.012814. Disponible en: <https://link.aps.org/doi/10.1103/PhysRevE.88.012814>.
- MCFARLAND Y OTROS, 2014. SoNIA - Social Network Image Animator. [en línea]. [Consulta: 30 noviembre 2018]. Disponible en: <http://web.stanford.edu/group/sonia/>.
- MCKINNEY, W., 2013. *Python for data analysis*. First Edition. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly. ISBN 978-1-4493-1979-3.
- MEJIA OLIVARES, C.P., 2010. Análisis de Redes Sociales a Gran Escala. , pp. 117.
- MELIN, G. y PERSSON, O., 1996. Studying research collaboration using co-authorships. *Scientometrics* [en línea], vol. 36, no. 3, pp. 363-377. [Consulta: 7 diciembre 2018]. ISSN 0138-9130, 1588-2861. DOI 10.1007/BF02129600. Disponible en: <http://link.springer.com/10.1007/BF02129600>.
- MÉNDEZ, J.J.F. y VEGA, Y.P., 2015. Trabajo de diploma para optar por el título de Ingeniero en Ciencias Informáticas. , pp. 80.
- MIMENZA, O.C., 2017. Teoría de la Influencia Social: sus aportes psicológicos. [en línea]. [Consulta: 29 noviembre 2018]. Disponible en: <https://psicologiyamente.com/psicologia/teoria-influencia-social>.
- NASCIMENTO, M.A., SANDER, J. y POUND, J., 2003. Analysis of SIGMOD's co-authorship graph. *ACM SIGMOD Record* [en línea], vol. 32, no. 3, pp. 8-10. [Consulta: 7 diciembre 2018]. ISSN 01635808. DOI 10.1145/945721.945722. Disponible en: <http://portal.acm.org/citation.cfm?doid=945721.945722>.
- NetworkX - Python Graph Library. *JournalDev* [en línea], 2018. [Consulta: 1 febrero 2019]. Disponible en: <https://www.journaldev.com/19410/networkx-python-graph-library>.
- NEWMAN, M.E.J. y GIRVAN, M., 2004. Finding and evaluating community structure in networks. *Physical Review E* [en línea], vol. 69. [Consulta: 7 diciembre 2018]. DOI 10.1103/physreve.69.026113. Disponible en: <https://link.aps.org/doi/10.1103/PhysRevE.69.026113>.
- ORTIZ MUÑOZ, E., 2015. *Detección de comunidades a partir de redes de coautoría en grafos RDF*. S.l.: s.n.
- OVEJERO BERNAL, A. y DIGITALIA (FIRM), 2007. *Las relaciones humanas: psicología social teórica y aplicada* [en línea]. Madrid: Biblioteca Nueva. [Consulta: 7 diciembre 2018]. ISBN 978-84-7030-565-8. Disponible en: <http://www.digitaliapublishing.com/a/4491>.
- PANCHAL, R., 2017. Execution of SPARQL Query using Apache Jena Fuseki Server in AISHE domain. *International Journal of Advance Engineering and Research Development*, vol. 4, no. 9, pp. 9.

- PRESSMAN, R.S., CAMPOS OLGUÍN, V., ENRÍQUEZ BRITO, J., VILLEGAS QUEZADA, C. y FERRO CASTRO, B.J., 2010. *Ingeniería del software: un enfoque práctico*. S.l.: s.n. ISBN 978-607-15-0314-5.
- PRUD'HOMMEAUX, E. y SEABORNE, A., 2008. SPARQL Lenguaje de consulta para RDF. [en línea]. [Consulta: 21 mayo 2019]. Disponible en: <http://skos.um.es/TR/rdf-sparql-query/>.
- RODRÍGUEZ, A.W.F. y RUIZ, J.J.H., 2014. *Análisis de influencia social en el contexto universitario*. S.l.: s.n.
- ROUCO, L.V., 2016. *Algoritmo para la inferencia de redes sociales a partir de la interacción de los recursos en la ejecución de las actividades de un proceso*. S.l.: s.n.
- SARABIA ALEGRÍA, J.M. y PASCUAL SAÉZ, M., 2005. *Curso básico de estadística para economía y administración de empresas* [en línea]. S.l.: s.n. [Consulta: 20 mayo 2019]. Disponible en: https://books.google.com/books/about/Curso_b%C3%A1sico_de_estad%C3%ADstica_para_econo.html?hl=es&id=QEp1dROP0OAC.
- T. D. STOKES AND J. A. HARTLEY, 1989. *Coauthorship, Social Atructure and Influence Within Specialties.pdf*. 1989. S.l.: s.n.
- TUCKER, A.B., 2004. *Computer science handbook*. 2. ed. Boca Raton, Fla: Chapman & Hall/CRC [u.a.]. ISBN 978-1-58488-360-9.
- VAN DER AALST, W.M.P., REIJERS, H.A. y SONG, M., 2005. Discovering Social Networks from Event Logs. *Computer Supported Cooperative Work (CSCW)* [en línea], vol. 14, no. 6, pp. 549-593. [Consulta: 7 diciembre 2018]. ISSN 0925-9724, 1573-7551. DOI 10.1007/s10606-005-9005-9. Disponible en: <http://link.springer.com/10.1007/s10606-005-9005-9>.
- VAN DER AALST, W.M.P. y SONG, M., 2004. Mining Social Networks: Uncovering Interaction Patterns in Business Processes. En: J. DESEL, B. PERNICI y M. WESKE (eds.), *Business Process Management* [en línea]. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 244-260. [Consulta: 7 diciembre 2018]. ISBN 978-3-540-22235-4. Disponible en: http://link.springer.com/10.1007/978-3-540-25970-1_16.
- WAGNER, C.S., BORNMANN, L. y LEYDESDORFF, L., 2015. Recent Developments in China–U.S. Cooperation in Science. *Minerva* [en línea], vol. 53, no. 3, pp. 199-214. [Consulta: 7 diciembre 2018]. ISSN 0026-4695, 1573-1871. DOI 10.1007/s11024-015-9273-6. Disponible en: <http://link.springer.com/10.1007/s11024-015-9273-6>.
- WEBMINING CONSULTORES, 2012. Análisis de Redes Sociales con R | WebMining. [en línea]. [Consulta: 29 noviembre 2018]. Disponible en: <http://www.webmining.cl/2012/04/analisis-de-redes-sociales-con-r/>.

ANEXOS

Anexo 1. Encuesta realizada para la aplicación del sociograma.

<p>El grupo de investigación "Herramienta para el análisis de redes de innovación tecnológica" está realizando un estudio que permitirá calcular una de las técnicas de la influencia social, la reciprocidad que se establece entre sus miembros. Pedimos su colaboración y la sinceridad en las respuestas. Garantizamos su anonimato y la confidencialidad de los datos. Muchas gracias por su contribución.</p>	0	↓
<p>¿Con quién te gustaría investigar?</p> <p>1-</p> <p>2-</p> <p>3-</p>	1	↓
<p>¿Con quién te gustaría practicar deporte?</p> <p>1-</p> <p>2-</p> <p>3-</p>	2	↓

¿Con quién te gustaría compartir en fiestas?

1-

2-

3-

Fin

Anexo 2. Acta de aceptación

**ACTA DE ACEPTACIÓN**

En cumplimiento de los objetivos y tareas especificadas para el Proyecto de Investigación y Desarrollo: **Herramienta para el análisis de redes de innovación tecnológica** y en función de la ejecución del proyecto: Método para inferir redes de interacción social a partir de las relaciones recíprocas entre los actores, se hace entrega del producto que se relaciona a continuación, son de la propia autoría de la estudiante Hany Bello Rodríguez.

- Método para inferir redes de interacción social a partir de las relaciones recíprocas entre los actores.

El tutor Ing. Vladimir Milián Núñez y el Lic. Raynel Batista Téllez, luego de haber revisado los productos de trabajo determinan la validez de los mismos.

Recibe:

Raynel Batista Téllez

Nombre y Apellidos:

Entrega:

Hany Bello Rodríguez

Nombre y Apellidos:

Recibe:

Vladimir Milián Núñez

Nombre y Apellidos: