



Universidad de las Ciencias Informáticas
Facultad 4

**MÉTRICA DE AUTORIDAD PARA DETERMINAR EL NIVEL DE
INFLUENCIA DE LAS PERSONAS EN LAS ORGANIZACIONES
TENIENDO EN CUENTA SU INTERACCIÓN EN REDES DE
MENSAJERÍA INSTANTÁNEA**

Trabajo final presentado en opción al título de
Máster en Informática Avanzada

Autor:

Ing. Dariel Corchado López del Castillo

Tutor:

Dr. C. José Felipe Ramírez Pérez, Profesor Auxiliar

La Habana, 13 de diciembre de 2018

AGRADECIMIENTOS

Agradezco a la mujer de mi vida por su paciencia, apoyo e impulso incondicionales, que lograron convertir desilusiones en fuerza de voluntad.

A mi mamá y mi papá por su ejemplo, apoyo y preocupación.

A mi hermana y mi hermano por su apoyo incondicional cuando tuve algún problema.

A mis suegros, mis tíos, mis tías, mi cuñada y mis cuñados por su apoyo brindado.

A mi hermana y mi cuñado por la laptop que me facilitaron, fue de gran utilidad para el desarrollo de la maestría.

A mi tutor por sus críticas acertadas, sus sugerencias y sus ideas para encontrar nuevos puntos de vista. Sin su ayuda esta investigación no se hubiera realizado.

A Raynel Batista por su colaboración vital en esta maestría.

A los profesores de la maestría, que con alto nivel y profesionalidad nos enseñaron temas técnicos complejos y modos de actuación coherentes.

A la Universidad de las Ciencias Informáticas por formarme como Ingeniero y ahora como Máster en Informática Avanzada.

A todas aquellas personas que, de una forma u otra contribuyeron y colaboraron con el desarrollo de esta investigación.

A los amigos y compañeros que siempre estuvieron presentes.

A todas las personas que desde pequeño influyeron en mi formación como profesional y como ser humano.

¡MUCHAS GRACIAS!

DEDICATORIA

A toda mi familia, porque son parte esencial de todo lo que soy, por su apoyo, comprensión y dedicación. A la Universidad de las Ciencias Informáticas por formarme profesionalmente

DECLARACIÓN DE AUTORÍA

Yo, Dariel Corchado López del Castillo, con carné de identidad 91033040802, declaro que soy el autor principal del resultado que expongo en la presente memoria titulada “Métrica de autoridad para determinar el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea”, para optar por el título de Máster en Informática Avanzada. Autorizo a la Universidad de las Ciencias Informáticas a hacer uso de la misma en su beneficio, así como los derechos patrimoniales con carácter exclusivo.

Para que así conste, firmo la presente declaración jurada de autoría en La Habana a los ____ días del mes de _____ del año 2018.

Firma del autor

RESUMEN

Para un correcto análisis de influencia y autoridad en una red de mensajería instantánea hay que tener en cuenta tanto el análisis estructural del grafo representativo de dicha red, como las reacciones que provoca un usuario sobre otro, cuando le escribe un mensaje. La autoridad de una persona en dichas redes está muy relacionada con su jerarquía y credibilidad. En esta investigación se desarrolla una métrica de autoridad, denominada Autoridad Total, para determinar el nivel de influencia de las personas en las organizaciones, teniendo en cuenta sus interacciones en redes de mensajería instantánea. Se exponen las bases teóricas que sustentan la investigación, las métricas utilizadas y las fórmulas y algoritmos desarrollados. Se presenta además a GRAPH-CHAT, herramienta desarrollada para implementar y probar la métrica y que permite visualizar la estructura del grafo que representa las relaciones de los usuarios de la mensajería instantánea, mostrando los que poseen mayor autoridad como los nodos de mayor tamaño, la fortaleza de sus relaciones con el grosor de la arista y los grupos sociales por el color de los nodos y las aristas. Se explica cómo GRAPH-CHAT genera el grafo a partir de los registros de eventos de los servidores de mensajería instantánea y cómo lo almacena en una base de datos orientada a grafos. Se presenta además cómo la herramienta utiliza esta información para sugerir equipos de trabajo. Se valida la métrica Autoridad Total mediante Criterios de Expertos y un Estudio de Caso. GRAPH-CHAT es validada mediante la satisfacción de sus usuarios potenciales.

Palabras clave: autoridad, bases de datos orientadas a grafos, GRAPH-CHAT, influencia, mensajería instantánea, redes sociales.

ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO 1. MARCO TEÓRICO REFERENCIAL DE LA INVESTIGACIÓN RELACIONADO CON LA DETECCIÓN DE LOS NIVELES DE AUTORIDAD PRESENTES EN LAS REDES DE MENSAJERÍA INSTANTÁNEA	9
1.1. Teoría de grafos	9
1.2. Influencia en redes de interacción social. Bases conceptuales.	12
1.2.1. Análisis de influencia social. Trabajos relacionados.....	13
1.3. Métricas de centralidad.....	16
1.3.1. Centralidad de grado (<i>degree</i>)	16
1.3.2. Centralidad de intermediación (<i>betweenness</i>).....	17
1.3.3. Centralidad de cercanía (<i>closeness</i>)	17
1.3.4. Centralidad de vector propio (<i>eigenvector</i>).....	18
1.3.5. PageRank	18
1.3.6. HITS.....	19
1.4. Detección de comunidades.....	19
1.4.1. Propagación de etiquetas (<i>Label Propagation</i>)	20
1.4.2. Componentes conectados (<i>Union find</i>)	20
1.4.3. Modularidad de Louvain.....	20
1.5. Herramientas para determinar influencia	21
1.6. Bases de Datos Orientadas a Grafos	23
Conclusiones del capítulo.....	25
CAPÍTULO 2. MÉTRICA DE AUTORIDAD PARA DETERMINAR EL NIVEL DE INFLUENCIA DE LAS PERSONAS EN LA MENSAJERÍA INSTANTÁNEA	26
2.1. Cálculo de Autoridad Total (AT), la Autoridad Estructural (AE) y Autoridad de Reacción (AR) en la mensajería instantánea.....	26
2.1.1. Autoridad Estructural (AE)	27
2.1.2. Autoridad de Reacción (AR)	29
2.2. GRAPH-CHAT: Herramienta para la visualización de la autoridad de los usuarios de la mensajería instantánea y la creación de equipos de trabajo	31

2.2.1. Importación de los datos.....	34
2.2.2. Procesamiento.....	36
2.2.3. Análisis	37
Conclusiones del capítulo.....	44
CAPÍTULO 3. VALIDACIÓN DE LA INVESTIGACIÓN.....	45
3.1 Aplicación del método Criterio de Expertos	45
3.2 Aplicación de la técnica Iadov para medir satisfacción	47
3.3 Estudio de Caso	50
3.4. Triangulación metodológica de los métodos científicos aplicados	55
Conclusiones del capítulo.....	56
CONCLUSIONES	58
RECOMENDACIONES	59
REFERENCIAS BIBLIOGRÁFICAS	60
ANEXOS.....	67
Anexo 1. Diagnóstico preliminar	67
Anexo 2: Encuesta del diagnóstico preliminar.	68
Anexo 3. Cuestionario a expertos.....	69
Anexo 4. Encuesta para determinar el coeficiente de competencias de los expertos	70
Anexo 5. Procedimiento empleado para determinar el coeficiente de competencia de los candidatos a expertos	70
Anexo 6. Resultados de la encuesta aplicada a los candidatos a expertos para determinar su nivel de competencia	71
Anexo 7. Respuestas dadas por cada uno de los expertos para cada indicador	71
Anexo 8: Encuesta aplicada para medir satisfacción del usuario con la herramienta GRAPH-CHAT.....	72
Anexo 9: Sociograma.	73
Anexo 10: Procedimiento para computar el sociograma.....	73

INTRODUCCIÓN

Desde tiempos inmemoriales las personas tienen la tendencia a interactuar entre ellas a partir de sus intereses en común, conformando redes sociales. Con el surgimiento del internet y las webs 2.0 fueron revolucionadas las redes sociales, operando en el ciberespacio y permitiendo así la participación activa de los usuarios en el proceso comunicativo generado. En la actualidad las personas constantemente interactúan y se comunican a través de algún dispositivo o medio tecnológico.

Los beneficios para aquellos que forman parte de las redes sociales son varios, entre ellos se encuentran aspectos de suma importancia para el ser humano, como pertenecer a un grupo con el que se tiene afinidad, hacer o renovar amistades, en lo laboral, colaborar y compartir conocimientos, así como buscar trabajo, promoviendo los conocimientos y habilidades con los que se cuenta (Valenzuela, 2013). En muchas empresas, el uso de las redes sociales fortalece y mejora la comunicación, promoviendo la interacción e impacto positivo entre los empleados, predominando una cultura de apoyo y de compartir información entre los usuarios mediante el intercambio. El conocimiento que se adquiere de dichas redes llega a ser considerado para la toma de decisiones (Sandoval-Almazan, Gómez-Díaz, & Demuner Flores, 2012).

Se pueden definir a las redes sociales como estructuras compuestas por personas conectadas por uno o varios tipos de relaciones (de amistad, de parentesco, de trabajo, ideológicas) con intereses comunes (Islas & Ricaurte, 2013). A nivel general, es una estructura social formada por nodos, habitualmente individuos u organizaciones, que están vinculados por uno o más tipos de interdependencia (Valenzuela, 2013). Una red social de internet es un espacio virtual de comunicación entre personas con algún rasgo en común en la que sus usuarios pueden contactar con familiares, amigos o desconocidos, y compartir contenidos sin que importe el tiempo, el espacio o la ubicación. Un sistema abierto de interacción social, en permanente construcción gracias a las continuas aportaciones de cada usuario (Pérez, 2013).

Es factible vincular dichos conceptos mediante la Teoría de Grafos, que expone que la red se constituye por nodos conectados por aristas, donde los nodos son los individuos y las aristas, las relaciones que les unen. La representación de la información correspondiente a los patrones de relaciones entre actores sociales se realiza generalmente mediante el uso de grafos, donde un grafo es una representación de un conjunto de objetos en el que los pares de objetos están conectados mediante relaciones entre ellos. (Kuz, Falco, & Giandini, 2016).

Las redes sociales de internet, también conocidas como Medios Sociales o Social Medias (Stieglitz, Mirbabaie, Ross, & Neuberger, 2018), pueden conectar personas y ayudarlas a mantener sus relaciones en toda su vida. Les permiten compartir e interactuar unas a otras de incontables maneras. Las personas pueden usar estos medios para construirse y exponer su identidad online, compartiendo los mejores eventos de su vida (Golbeck, 2015). Estos medios se han vuelto muy populares en la actualidad destacándose entre ellos Facebook, Twitter, YouTube e Instagram como los más utilizados.

Un componente importante de los medios sociales es la mensajería instantánea. Su utilización permite la comunicación activa entre los usuarios de la red. Consiste en una forma de comunicación en tiempo real, o en tiempos relativamente muy cortos, entre dos o más personas basada en texto, el cual es enviado a través de dispositivos conectados a una red como Internet. Para profundizar en su concepto se debe mencionar el término Chat como el origen de las diferentes modalidades de comunicación síncrona de la actualidad, siendo utilizado ambos términos como sinónimos al compartir la mayoría de las características básicas de este tipo de comunicación (Martín Fernández, 2017).

La mensajería instantánea ha pasado a formar parte inevitable de los hábitos de una gran mayoría de la población, forja una nueva forma de comprender la comunicación y unas nuevas expectativas en torno al intercambio entre emisor y receptor (Carrión, 2016; Mancinas-Chávez & Reig, 2016). Permite el intercambio de mensajes de manera sincrónica y en tiempo real, donde los mensajes son escritos y leídos al mismo tiempo, a pesar de la distancia que media entre los usuarios (Mancinas-Chávez & Reig, 2016; Torres-Fernández & Pérez-Sánchez, 2013). Con los grupos de contactos de estos usuarios, se puede llegar a conformar una pequeña o gran red social. La magnitud de su tamaño depende de la cantidad de contactos que posea cada uno.

En muchas organizaciones este servicio se utiliza como medio informal de comunicación entre los trabajadores. Teniendo en cuenta que la influencia procede de la comunicación informal e interpersonal más que de los medios formales (García, Daly, & Sánchez-Cabezudo, 2016), se puede concluir que su estudio constituye una fuente valiosa de información para identificar trabajadores influyentes. Para ello las instituciones pueden auxiliarse del Análisis de Redes Sociales (ARS).

El ARS es un área del conocimiento surgido desde la sociología que provee métodos y herramientas para estudiar las relaciones humanas y predecir comportamientos con el objetivo de apoyar la toma de decisiones organizacionales (Ramírez, 2016). Cuantifica las relaciones entre los actores con el objetivo de crear matrices y redes gráficas que representen

esas relaciones como un todo, y de esa forma analizar las distintas características del sistema de relaciones bajo estudio. Se concentra en la red de relaciones, las posiciones funcionalmente diferenciables dentro de éstas, sus procesos dinámicos de adaptación, sus flujos y transacciones, entre otras (Aguirre, 2011).

El análisis de dichas relaciones refleja, en principio, la capacidad de algunas personas de influir en el comportamiento de otras, esto es conocido como influencia social (Cialdini, 2009). El estudio de la influencia social no solo resulta difícil de medir, sino que los criterios de influencia varían según el contexto y objetivo de la medición. Para caracterizar la influencia, Cialdini (2009) sintetiza los patrones de conducta y comportamientos de las personas en seis principios psicológicos: reciprocidad, coherencia, aprobación social, empatía, escasez y autoridad.

La influencia social y los fenómenos relacionados con ella, como el liderazgo y el poder social, son intrínsecos a cualquier proceso de organización humana. Ha sido concebida como un proceso relacional para el cual la influencia interpersonal que ocurre en el nivel de diada constituye un elemento básico (Carter, Dechurch, Braun, & Contractor, 2015; Saint-Charles & Mongeau, 2018). En los grupos de trabajo, la influencia social afecta cómo toman decisiones, llevan a cabo su trabajo o dan forma a su comprensión de la tarea y el medio ambiente (Johnson, van de Schoot, Delmar, & Crano, 2015; Saint-Charles & Mongeau, 2018). Las interacciones y la comunicación son fundamentales para la influencia social (Saint-Charles & Mongeau, 2018).

Identificar las personas más influyentes (conocidos comúnmente como influencers) tiene muchas aplicaciones para cualquier organización. Ayuda a comprender mejor por qué ciertas tendencias o innovaciones se adoptan más rápidamente que otras (Cha, Hamed, Benevenuto, & Gummadi, 2010). Un influencer es una persona que cuenta con una gran presencia y credibilidad en las redes sociales (seguidores), y, por tanto, una gran influencia en ellas. El influencer tiene la capacidad de generar opiniones y reacciones en los usuarios cuando habla de un cierto tema, conectando así con el público objetivo. Esto hace que sea atractivo para las empresas y les busquen para trabajar con sus marcas y productos (Rodríguez-Colomo, 2017). Su favorable posición en una red social les permite iniciar diversos procesos como la transmisión de creencias, chismes, publicidad viral, etc. En estos casos, el proceso empieza en el influencer y se distribuye a sus vecinos, los cuales redistribuyen a sus propios vecinos, sucesivamente (Monsalve, 2008).

En Cuba también se investiga acerca de la influencia en las redes sociales. El proyecto de investigación “Herramientas para el Análisis de Influencia en Redes de Innovación

Tecnológica” de la Universidad de las Ciencias Informáticas (UCI), al cual pertenece el autor de este documento, es un vivo ejemplo de esto. Bajo el enfoque de este proyecto se han desarrollado investigaciones (Bello-Lara, 2018; Estevez & Antigua, 2015; Fernández Rodríguez, Hidalgo Ruiz, & Rodríguez Puente, 2014; Ramírez, 2016; Sánchez Viera & Ruana Capote, 2014) que utilizan técnicas del Análisis de Redes Sociales en sus soluciones. Estos estudios tienen un marcado carácter transdisciplinar y representan campos emergentes de la ciencia como la Sociocibernética y la Minería de Procesos.

Aunque se han realizado investigaciones acerca de las redes sociales y los medios sociales, no se ha investigado la mensajería instantánea. Este tipo de sistemas son muy utilizados por parte de los trabajadores de muchas organizaciones del país. Constituyen uno de los mayores entornos de datos donde las personas interactúan de forma espontánea, una red informal de interacción social y es una fuente valiosa de información que perfectamente se puede utilizar para determinar cuáles son las personas de más influencia social. Conocer esta información posibilitaría tomar decisiones relacionadas con:

- Determinación de liderazgo.
- Fortaleza o cohesión de los grupos sociales.
- Conocimiento de los actores claves en la propagación de una información (Brókeres de la información).

Para corroborar esto, fue aplicada una encuesta a trabajadores y directivos de la Universidad de las Ciencias Informáticas y la empresa Viajero Servicios de Reservación (ver Anexo 1 y Anexo 2). Luego de computar los resultados se pudo constatar lo siguiente:

- El 100% de los encuestados utiliza la mensajería Instantánea y el 85% la usa frecuentemente.
- El 88,3% cree que un trabajador influyente puede ser un buen líder.
- El 98,3% desconoce la existencia de alguna herramienta en su organización que permita determinar el liderazgo o influencia que ejerce un trabajador sobre los demás, en determinada área o especialidad.
- El 98.3% plantea que no es posible conocer los principales grupos sociales que interactúan en su área de trabajo, o la fortaleza que tienen estos.
- El 96% desconoce si en su organización para tomar decisiones, se tiene en cuenta la influencia y liderazgo que ejercen los trabajadores en los distintos grupos sociales.
- El 100% plantea que no disponen de alguna herramienta que, teniendo en cuenta el liderazgo o influencia que ejercen los trabajadores entre ellos, así como su información

conocida (datos personales y/o laborales), le apoye la toma de decisiones para un mejor cumplimiento de los objetivos de su organización.

Estos planteamientos evidencian las siguientes insuficiencias:

- El desconocimiento del liderazgo o influencia que ejerce un trabajador, en determinada área o especialidad, dificulta identificar cuáles son más propensos a modificar el comportamiento de los demás, generando opiniones y reacciones cuando hablan de un cierto tema.
- Al no conocer la cohesión de las interacciones de los principales grupos sociales se imposibilita predecir cómo estos colaborarán y compartirán conocimiento para realizar alguna tarea, así como quiénes serán los trabajadores claves e influyentes en la comunicación entre estos grupos.

Por lo antes expuesto se identifica el siguiente **problema a resolver**: ¿Cómo determinar el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea?

Se delimita como **objeto de estudio**: análisis de influencia en redes de interacción social, específicamente en el **campo de acción**: autoridad presente en las redes de mensajería instantánea.

Para solucionar el problema planteado se define como **objetivo general**: Desarrollar una métrica de autoridad, utilizando técnicas de análisis de redes sociales, que permita determinar el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea.

Para dar cumplimiento al objetivo general se definen los siguientes **objetivos específicos**:

1. Elaborar el marco teórico de la investigación relacionado con el análisis de influencia en redes de interacción social y la autoridad presente en las redes de mensajería instantánea.
2. Diagnosticar el estado actual en las organizaciones asociado con la determinación del nivel de influencia de sus trabajadores, a partir de la interacción en redes de mensajería instantánea.
3. Desarrollar una métrica de autoridad para determinar el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea.

4. Desarrollar una herramienta informática que incorpore la métrica de autoridad para mostrar el nivel de influencia de los usuarios en las redes de mensajería instantánea.
5. Validar la métrica y la herramienta desarrollada a partir de los métodos científicos definidos.

Para guiar la investigación se propone la siguiente **hipótesis científica**: Si se desarrolla y aplica una métrica de autoridad, utilizando técnicas de análisis de redes sociales, se determinará el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea.

Variable independiente: métrica de autoridad.

Variable dependiente: nivel de influencia.

Durante la investigación se utilizaron los siguientes métodos científicos:

Métodos teóricos:

- Método analítico-sintético: permitió analizar individualmente los principales conceptos relacionados con la determinación de influencia en redes sociales, posibilitando un análisis profundo de cada uno, para luego llevar a cabo el estudio de las relaciones que se establecen entre ellos.
- Método histórico-lógico: se empleó para investigar los estudios existentes acerca de la influencia en redes de interacción social y los niveles de autoridad presentes en estas redes.
- Método inductivo-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 realizan.
- Estudio de caso: se empleó para verificar la correspondencia entre los resultados obtenidos y los esperados.

Métodos empíricos:

- Análisis documental: Se utilizó para el estudio de los referentes teóricos de la investigación, de forma tal que la propuesta tuviese relevancia científica y aporte práctico. Se realizó consulta de libros y de artículos científicos digitales.
- Observación: Se empleó para analizar los procesos y necesidades existentes a la hora de tomar decisiones que estuvieran relacionadas con la influencia de las personas en las organizaciones.
- Encuestas: Se utilizaron para diagnosticar la situación problemática y para validar la propuesta.

- Criterio de expertos: A partir de la consulta de expertos en sociología, psicología, comunicación social e informática se evaluaron los elementos teóricos que fundamentan la investigación y posibilitan el cálculo de la métrica de autoridad. Se tuvieron en cuenta sus conocimientos y experiencias.
- Técnica ladov: Se utilizó para validar y obtener retroalimentación de los usuarios sobre su nivel de satisfacción con la herramienta GRAPH-CHAT.

Además, se realizó una triangulación metodológica como procedimiento de control para disminuir el sesgo que puede producirse en la aplicación de los métodos científicos mencionados.

La **novedad científica** de la investigación se expresa en los siguientes **aportes prácticos**:

- Una métrica de autoridad que, en la red social formada por las interacciones los usuarios de la mensajería instantánea, combina la jerarquía (Autoridad Estructural) y credibilidad (Autoridad de Reacción) de cada usuario en la red para mostrar cuáles son los más influyentes.
- Una herramienta informática que visualiza esta red mediante un grafo, resaltando sus usuarios de mayor autoridad y sus grupos sociales.
- Un algoritmo que, basándose en las competencias profesionales e interacciones de los usuarios de la mensajería instantánea, de la Universidad de las Ciencias Informáticas, sugiere equipos de trabajo teniendo en cuenta tres contextos: Colectivo de asignatura, Proyecto de investigación, Proyecto productivo, los cuales representan los procesos fundamentales de la Universidad (Formación, Investigación y Producción).

El documento está estructurado en tres capítulos:

CAPÍTULO 1. Marco teórico referencial de la investigación, relacionado con la detección de los niveles de autoridad presentes en las redes de mensajería instantánea: Se abordan los elementos que conforman el marco teórico referencial y que componen el objeto de estudio y campo de acción de la investigación. Se analiza la importancia que tiene conocer cuán influyente es una persona en una red de interacción social y los beneficios que aporta este conocimiento. Se indaga acerca de las principales soluciones, herramientas, algoritmos y métricas del Análisis de Redes Sociales y la teoría de grafos existentes para el cálculo de autoridad e influencia.

CAPÍTULO 2. Métrica de autoridad para determinar el nivel de influencia de las personas en la mensajería instantánea: Se presenta la métrica de autoridad definida (Autoridad Total), sus componentes (Autoridad Estructural y Autoridad de Reacción) y los algoritmos desarrollados para su cálculo. Se presenta además a GRAPH-CHAT, herramienta que fue desarrollada para visualizar mediante un grafo la autoridad de los usuarios de la mensajería instantánea, exponiendo sus funcionalidades y resaltando sus potencialidades para crear equipos de trabajo.

CAPÍTULO 3. Validación de la investigación: Se valida el cumplimiento del objetivo general de la investigación, a partir de los métodos científicos definidos. Se valida la pertinencia, efectividad y utilidad de la métrica de autoridad desarrollada en la determinación de personas influyentes, mediante el método cualitativo Criterio de Expertos y un Estudio de Caso aplicado en la Facultad 4 de la Universidad de las Ciencias Informáticas. Se valida además la satisfacción de los usuarios potenciales de herramienta la GRAPH-CHAT mediante la técnica ladov. El proceso de validación es concluido realizando una triangulación metodológica inter-métodos e intra-métodos que permite evaluar el sesgo que se produce al aplicar cada uno de estos métodos y técnicas.

CAPÍTULO 1. MARCO TEÓRICO REFERENCIAL DE LA INVESTIGACIÓN RELACIONADO CON LA DETECCIÓN DE LOS NIVELES DE AUTORIDAD PRESENTES EN LAS REDES DE MENSAJERÍA INSTANTÁNEA

En el presente capítulo se precisan un conjunto de elementos que conforman la fundamentación teórica de la investigación. Se abordan los conceptos fundamentales de la teoría de grafos y el análisis de influencia en redes de interacción social. Se realiza un análisis de las métricas, algoritmos y herramientas más populares para detectar autoridad e influencia en redes sociales. Se exponen además los beneficios de la utilización de las bases de datos orientadas a grafos para almacenar las relaciones sociales de forma eficiente.

1.1. Teoría de grafos

Un grafo $G(V, E)$ es un conjunto finito de elementos no vacíos $V(G)$, llamados vértices o nodos, y una familia $E(G)$ de pares no ordenados (no necesariamente distintos) de elementos de $V(G)$, llamada aristas. Una arista $\{v, w\}$ consiste en la unión de los vértices v y w , y se puede abreviar como vw . Por ejemplo, en la Figura 1, $V(G)$ es el conjunto $\{u, v, w, z\}$ y $E(G)$ consiste en las aristas uv , vv (dos veces), vw (tres veces), uw (dos veces) y wz (Wilson, 1996).

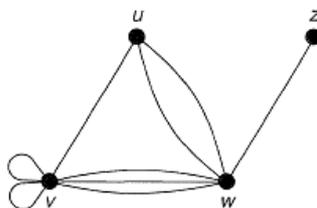


Figura 1: Ejemplo de un grafo. Fuente: (Wilson, 1996).

A continuación, se muestran un conjunto de definiciones relacionadas con la teoría de grafos que se consideran útiles para esta investigación. Fueron tomadas de (Heileman, 1998; Wilson, 1996) y se enumeran a continuación:

1. Un grafo simple G consiste en un conjunto finito no vacío $V(G)$ de elementos llamados vértices o nodos, un conjunto finito $E(G)$ de distintos pares ordenados de elementos de $V(G)$ llamados aristas. Se llama $V(G)$ el conjunto de vértices y $E(G)$ el conjunto de aristas de G . Una arista $\{v, w\}$ consiste en la unión de los vértices v y w , y es usualmente abreviado como vw . La Figura 2 representa el grafo simple G cuyo conjunto de nodos $V(G)$ es $\{u, v, w, z\}$ y cuyo conjunto de aristas $E(G)$ consiste en las aristas uv , uw , vw y wz .

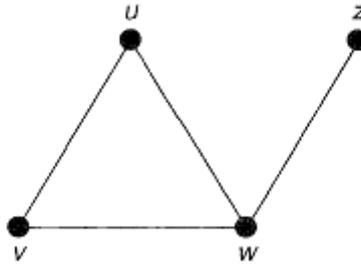


Figura 2: Ejemplo de grafo simple. Fuente: (Wilson, 1996).

- Se dice que dos vértices v y w de un grafo G son adyacentes o vecinos si existe una arista vw que los une, y los vértices v y w son incidentes con tal arista. Similarmente dos aristas distintas son adyacentes si tienen un vértice en común (ver Figura 3).



Figura 3: Los vértices v y w son adyacentes. Las aristas e y f son adyacentes también. Fuente: (Wilson, 1996).

Un subgrafo de un grafo G es un grafo, cuyos vértices pertenecen a $V(G)$ y cuyas aristas pertenecen a $E(G)$. Por ejemplo, en la Figura 4: el grafo A es un subgrafo de B, pero no un subgrafo de C porque C no contiene triángulos.

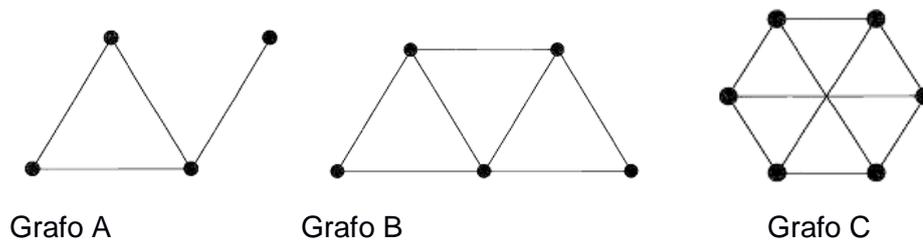


Figura 4: Representación de grafos y subgrafos. Fuente: (Wilson, 1996).

- Si G es un grafo con los vértices etiquetados por $\{1, 2, \dots, n\}$, su matriz de adyacencia A es la matriz de $n \times n$ cuya entrada $ij - th$ es el número de aristas que unen el vértice i con el j . Si, en adicción, las aristas son etiquetadas por $\{1, 2, \dots, m\}$, su matriz de incidencia M es la matriz $n \times m$ cuyas $ij - th$ entradas es 1, si el vértice i incide en la arista j y 0 de otra forma. La Figura 5 muestra un grafo G con sus matrices de adyacencia e incidencia.

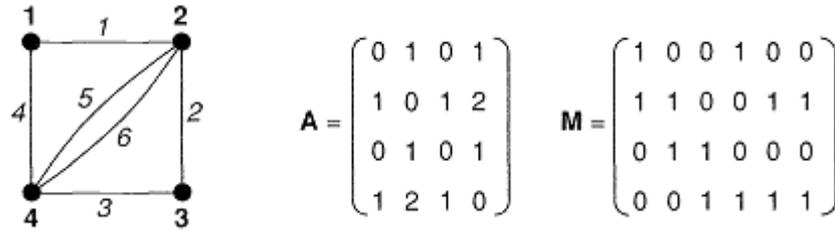


Figura 5: Grafo G con su matriz de adyacencia A y matriz de incidencia M . Fuente: (Wilson, 1996).

- Un grafo dirigido o digrafo $G = (V, E)$ también consiste en un conjunto de nodos y un conjunto de aristas. Sin embargo, en los grafos dirigidos, el conjunto de aristas consiste en pares ordenados de nodos de $V(G)$. En otras palabras, la dirección de la arista es importante en los grafos dirigidos. De este modo la arista uv que une los nodos u y v en el grafo dirigido G se dice que es incidente desde el nodo u e incidente hacia el nodo v . En la representación de un grafo dirigido las aristas se dibujan con flecha para mostrar su dirección (ver Figura 6).



Figura 6: Grafo dirigido incidente desde u hacia v . Fuente: elaboración propia.

- Cuando se usan grafos dirigidos o no dirigidos para modelar ciertas relaciones, a menudo resulta útil asociar un peso a cada arista de un grafo. Se usa $w(u, v)$ para denotar el peso asociado a la arista uv . Una asignación de pesos a un grafo $G(V, E)$ viene dada por una aplicación $w: \{(u, v) \in E\} \rightarrow \mathbb{R}$. Los grafos que tienen una asignación como esta constituyen grafos ponderados.
- Un paseo en un grafo se define como una secuencia alternada de nodos y aristas, comenzando y terminando con nodos, tal que cada arista es incidente en los nodos que la preceden y la siguen.
- Un paseo que comienza y termina en el mismo nodo se llama paseo cerrado. Un paseo que no está cerrado se llama paseo abierto. Un paseo simple es un paseo en el cual no se repite ninguna arista, mientras que un camino es un paseo en el cual no se repite ningún nodo. Si hay un camino desde el nodo u hasta el nodo v , entonces se dirá que v es accesible desde u .
- Un paseo simple cerrado en un grafo se conoce como un circuito y un circuito cuyos nodos son diferentes se denomina ciclo (se excluye en esta definición el ultimo nodo al que se debe regresar con el fin de forma el circuito).
- Un grafo o grafo dirigido que no contenga ningún ciclo se dice que es acíclico.
- Un grafo se dice que es conexo si existe al menos un camino entre cada par de nodos del grafo. Un grafo dirigido se dice que es fuertemente conexo si cualesquiera dos nodos del grafo dirigido son accesibles el uno desde el otro.

1.2. Influencia en redes de interacción social. Bases conceptuales.

La influencia se define como el “*poder, valimiento, autoridad de alguien para con otra u otras personas*” (Real Academia Española, 2018). La influencia social de un usuario es su habilidad para inducir reacciones en otros usuarios (Rao, Spasojevic, Li, & Dsouza, 2015). Ocurre cuando la opinión, las emociones y el comportamiento de una persona son afectados por otra (Li, Zhang, & Huang, 2018). En los grupos de trabajo, la influencia social afecta cómo toman decisiones, realizan su trabajo o comprenden una tarea y su entorno (Saint-Charles & Mongeau, 2018).

La señal principal de una interacción influyente es cuando una acción de un usuario provoca reacciones entre otros usuarios. Cada una de estas reacciones indica una unidad de flujo de información. Una reacción puede estar representada por una tupla de dimensiones (Quién, Cuándo, Dónde, Qué, Cómo), dichas dimensiones se explican a continuación (Rao et al., 2015):

- **Quién:** las características de la audiencia que reaccionó a la publicación original del usuario.
- **Cuándo:** la diferencia entre la hora actual y el momento en que ocurrió la reacción.
- **Dónde:** la red social en la que se produce la reacción.
- **Qué:** la unidad de contenido original o acción en que provocó la reacción.
- **Cómo:** el tipo de reacción.

En el ejercicio y desarrollo de la influencia y de la autoridad intervienen numerosos elementos que escapan a las mediciones. Es el factor humano, tan imprevisible como rico en valores intangibles (Serrano-Puche, 2012). 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 psicológicos definidos por Cialdini (2009), los cuales se muestran a continuación:

1. **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 más a cambio. Se fundamenta en que dos personas se tratan de acuerdo a los comportamientos anteriores que haya tenido una con respecto a la otra.
2. **Coherencia:** La coherencia se basa en el deseo de ser y parecer una persona de actitudes y comportamientos consecuentes a lo largo del tiempo. La coherencia es un rasgo de personalidad muy valorado socialmente, que se asocia a otros rasgos como

el ser lógico, racional, estable y honrado. A una persona poco coherente se la considera superficial, poco inteligente, indecisa y débil.

3. **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, aunque considere que la sociedad esté equivocada.
4. **Empatía:** 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. Es clave para vender, aunque la empatía en extremo produce un efecto de rechazo. Un factor que aumenta la empatía 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.
5. **Escasez:** las oportunidades parecen más valiosas cuanto más difíciles resultan conseguirlas. 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.
6. **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.

En el contexto de esta investigación, el autor define influencia social como la habilidad de una persona de inducir reacciones en otras por medio de su autoridad. Además, la autoridad es definida como la jerarquía y credibilidad que posee la persona en la red de interacción social.

1.2.1. Análisis de influencia social. Trabajos relacionados

El análisis de influencia social (AIS) estudia principalmente la manera de modelar el proceso de difusión de influencia en las redes, y cómo proponer un método eficiente para identificar un grupo de nodos objetivo en una red. Las principales preguntas de estos estudios incluyen: ¿quién influencia a quién?; ¿quién es influenciado?, ¿quiénes son los usuarios más influyentes?, entre otras. AIS tiene un significado social importante y ha sido aplicado en

muchos campos como el Marketing viral, recomendación en línea, comunidades de asistencia médica, descubrimiento de expertos, propagación de un rumor, y otras aplicaciones que dependen del efecto de la influencia social. Puede ayudar a comprender los comportamientos sociales de personas, proveer soporte teórico para hacer decisiones públicas e influenciar opinión pública, y puede promover cambios y disseminación de actividades diversas (Li et al., 2018).

La búsqueda de actores influyentes en las redes sociales ha interesado mucho a los investigadores. Por ejemplo, una búsqueda en *ScienceDirect*¹ de las palabras clave “social”, “network”, “influence”, “analysis” devuelve más de 200 mil resultados. En la Figura 7 se muestran estos resultados agrupados por el año de publicación. Se puede apreciar que los estudios relacionados con esta área del conocimiento han tenido un crecimiento exponencial



Figura 7: Exponencial crecimiento de publicaciones en ScienceDirect que contienen las palabras clave “social”, “network”, “influence”, “analysis”. Fuente: elaboración propia.

Como ejemplo de estos estudios se pueden citar los trabajos de los siguientes autores:

- Sudhahar, Veltri y Cristianini (2015) realizan un escaneo automático de 130 mil 213 artículos acerca de las elecciones de Estados Unidos en el 2012 produciendo un grafo donde los nodos son sus actores políticos y las cuestiones tratadas, las aristas representan las relaciones de apoyo u oposición de los actores a las cuestiones. El uso de la centralidad de grado en esta investigación reveló los actores y las cuestiones más sobresalientes, la intermediación identificó aquellos actores y asuntos que tienen

¹ <https://www.sciencedirect.com/>

función de puentes entre los diferentes grupos de actores y cuestiones identificadas en el grafo, la centralidad de cercanía ayudó a identificar los nodos que actuaron como contextualizadores. La autoridad (medida con el algoritmo HITS) permitió identificar aquellas cuestiones que estuvieron presentes en la agenda de campaña de cada partido.

- Beveridge y Shan (2016) grafican la red social formada por las interacciones de los personajes de la popular serie "Game of Thrones" en su tercer libro ("Storm of Swords"). Detectan sus principales comunidades y; mediante el cálculo para cada nodo de las centralidades de grado, de grado con peso, de intermediación, de cercanía, de vector propio y el valor del PageRank; definen sus personajes más importantes. En el grafo que obtienen aparecen las distintas comunidades existentes representadas del mismo color, el tamaño de cada vértice corresponde a su PageRank, sus etiquetas son escaladas al valor de la centralidad de intermediación y el ancho de las aristas del grafo corresponde al peso de la relación que cada una representa.
- García et al. (2016) presentan una metodología donde identifican a los actores influyentes como los nodos con mayor centralidad de grado e intermediación y los clasifican como diseminadores, relacionales y líderes.
- Un enfoque para la identificación de usuarios influyentes, que permite analizar la evolución de la influencia de dichos usuarios en tiempo real, en Twitter, es propuesto por Alonso et al. (2016). En este trabajo utilizaron el modelo Credit Distribution² de maximización de influencia.
- Ríos et al. (2017) proponen el uso de análisis semántico para filtrar los tipos de interacciones de las redes sociales en línea, alcanzando un grafo simplificado que preserva sus principales características permitiendo la detección de los verdaderos influencers en un costo computacional reducido significativamente.
- Riquelme et al. (2018), exploran las posibilidades del Modelo Linear Threshold (Kempe, Kleinberg, & Tardos, 2015) para definir métricas de centralidad en redes sociales donde las relaciones tienen pesos y proponen una nueva métrica de centralidad para asignar un ranking a los usuarios de la red: el Ranking de Umbral

² Credit Distribution: Modelo que calcula la propagación de influencia mediante la utilización de datos históricos.

Lineal (LTR). Proponen además el Umbral Lineal de Centralización (LTC) el cuál determina en que magnitud la red social tiene una estructura centralizada.

- He et al. (2018) formulan un modelo de maximización de influencia para formación de opinión (IMOF) como un problema de optimización, reflejan la influencia entre dos nodos adyacentes basándose en sus centralidades de grado, de cercanía y de intermediación y proponen un algoritmo heurístico de tres pasos para calcular los nodos más influyentes.
- Zhao, Lee y Chen (2018) realizan un ranking de usuarios en redes sociales a partir de una modificación que proponen al algoritmo PageRank.

1.3. Métricas de centralidad

Como se ha observado, la mayoría de las investigaciones calculan la influencia combinando diversas métricas de centralidad presentes en las ciencias sociales y la teoría de grafos. Esto se debe a que una mayor centralidad en una red, supone una forma de ventaja estructural, rango, poder, influencia o dominio (García et al., 2016). Las métricas de centralidad más utilizadas son el Grado, la Intermediación, la Cercanía, la Centralidad de Vector Propio, el PageRank y los valores del HITS. Estas determinan cuán relevante es un actor a partir de la estructura de la red social a la que pertenecen por lo que están muy relacionadas con la topología del grafo que la representa. (Freeman, 1978; García et al., 2016; Riquelme et al., 2017). Por su importancia son explicadas a continuación.

1.3.1. Centralidad de grado (degree)

Coincide con el grado de un vértice del grafo. El grado de un vértice v es el número de líneas que tienen a v como nodo de uno de sus extremos. El grado de entrada de un vértice v es el número de líneas que poseen a v como nodo terminal y el grado de salida de un vértice v es el número de líneas que poseen a v como nodo inicial (Kuz et al., 2016).

En términos de grafos, la centralidad de grado de un actor se calcula como su número de vecinos (Monsalve, 2008), también llamadas conexiones o aristas incidentes. La centralidad de grado con peso se define similarmente, pero sumando los pesos de las aristas incidentes (Beveridge & Shan, 2016). Si se está modelando una red social de amigos, la centralidad de cada actor consiste en su número de amigos (Monsalve, 2008). Las personas que tienen muchas conexiones con terceros pueden tener más influencia, más acceso a la información o más prestigio que los que tienen menos conexiones (García et al., 2016).

1.3.2. Centralidad de intermediación (betweenness)

Se define como el número de rutas mínimas en las que el actor participa (Monsalve, 2008). Es una métrica muy importante porque se puede utilizar para identificar brókeres³ de información en la red o nodos que pueden conectar grupos separados (Lyon, 2016). Un nodo con alto betweenness tiene una gran influencia en la transferencia de información entre los nodos a través de la red, asumiendo que cada nodo transfiere siguiendo los caminos más cortos (Freeman, 1978; Kuz et al., 2016). Posee un potencial control sobre la información, ya que, si se opone a una idea, su difusión a otros segmentos del grupo puede ser bloqueada. Puede frenar los flujos, distorsionar lo que se transmite para servir a sus intereses. Su eliminación de la red provocaría una gran interrupción en la comunicación (García et al., 2016).

1.3.3. Centralidad de cercanía (closeness)

Se define como el promedio de las distancias de un vértice a todos los vértices, teniendo en cuenta que valores pequeños constituyen una mayor importancia (Beveridge & Shan, 2016). Representa los nodos que, a pesar de tener pocas conexiones, sus arcos permiten llegar a todos los puntos de la red más rápidamente que desde cualquier otro punto y que poseen una excelente posición para monitorear el flujo de información de toda la red (Kuz et al., 2016). En resumen, el valor de la cercanía entre dos nodos no es más que la longitud del camino mínimo que los une (cantidad de aristas que separan un nodo del otro) y el valor bruto total de la cercanía de un nodo es igual al promedio de las cercanías para cada nodo. Es importante notar que el valor bruto total de la cercanía es una medida inversa de centralidad porque los nodos con menor valor son considerados los más centrales (Needham & Hodler, 2018).

En esta investigación se utiliza la implementación de esta métrica desarrollada por la Comunidad de Neo4j (Needham & Hodler, 2018; Neo4j, 2018), donde su valor es normalizado teniendo en cuenta la cantidad de usuarios del grafo (ver Fórmula 1). De esta forma el valor de la Cercanía calculado no constituye una métrica inversa por lo que los nodos con mayores valores son los más centrales.

$$SN = (N - 1)/S \quad (1)$$

- SN es el valor normalizado de la cercanía, N la cantidad de nodos y S el valor bruto total de la cercanía.

³ Personas que conectan varios grupos sociales, por lo que constituyen el puente entre la información que circula entre esos grupos.

1.3.4. Centralidad de vector propio (*eigenvector*)

Se define como una centralidad recursiva. Un actor central es aquel que tiene un vecindario con buena centralidad. Por ejemplo, un feriante puede conocer a mucha gente común mientras que un político puede conocer menos gente, pero que son personas influyentes. Al final, el político está mejor posicionado en influencia que el feriante, aunque conozca menos gente. Ahora, no todo el vecindario de un actor contará con la misma centralidad, por lo que hay que considerar que los actores con mayor centralidad son más influyentes que el resto (Monsalve, 2008).

Siguiendo el concepto de centralidad recursiva, se dirá que la centralidad de un actor es proporcional a la suma de las centralidades de sus vecinos en el grafo. Matemáticamente, esto se expresa como $c_i = \lambda \sum_j a_{ij} c_j$, donde λ es la constante de proporcionalidad y a_{ij} es el elemento de la fila i y la columna j de la matriz de adyacencia A de la red social. En álgebra lineal $\vec{c} = \lambda A \vec{c}$, o sea, \vec{c} es un vector propio de A (y λ^{-1} es su correspondiente valor propio) (Monsalve, 2008).

1.3.5. PageRank

PageRank es el popular algoritmo de búsqueda de google. Trabaja contando el número y calidad de links a una página para determinar cuán importante es. Constituye una variación de la Centralidad de vector propio (Needham & Hodler, 2018). En este algoritmo cada vértice del grafo adquiere un valor de importancia a partir de sus vecinos que, a diferencia de la Centralidad de vector propio, un vértice no adquiere el total de la importancia, en lugar de ello, esta importancia es dividida equitativamente entre sus conexiones directas. En otras palabras un vértice con un alto grado de relevancia solo pasa una pequeña fracción de su importancia a cada vecino (Beveridge and Shan, 2016).

PageRank fue inicialmente introducido para calcular el ranking de las páginas web en internet (Page, Brin, Motwani, & Winograd, 1999). También puede ayudar a seleccionar nodos influyentes en las redes. Ha sido usado en diversos dominios como análisis de redes de citas, predicción de enlaces, identificación de líderes de opinión y en modelos de influencia social que muestran que la autoridad es equivalente a la influencia (Zhao et al., 2018). Se basa en el supuesto de que un hipervínculo desde una página que apunta a otra página es un indicador implícito de la autoridad de la página de destino. Así, cuanto más enlace recibe una página i , más prestigio posee. Al mismo tiempo, una página se vuelve más importante si es apuntado por otras páginas importantes, generando una definición recursiva (Corbellini, 2015).

Para formular las ideas anteriores, dado el grafo dirigido $G = (V, E)$ el valor de PageRank de la página i , denotado como $PR(i)$, se define como:

$$PR(i) = (1 - \alpha)e + \alpha \sum_{(j,i) \in E} \frac{PR(j)}{O_j} \quad (2)$$

donde O_j es el número de links salientes del vértice j . El parámetro α es llamado "*factor de reducción*", y modela el hecho de que en cada paso el navegante puede saltar a una página Web al azar con probabilidad α , o puede seguir alguno de los links en la página actual con probabilidad $(1 - \alpha)e$. En otras palabras, el algoritmo es una caminata aleatoria con reinicios al azar sobre G . Si e es todos unos, entonces existe igual chance de saltar a cualquier página o vértice en G . Personalizar e permite ajustar el comportamiento del caminante aleatorio (Corbellini, 2015).

1.3.6. HITS

Kleinberg (1999) presenta el algoritmo HITS para detectar concentradores ("*hubs*") y autoridades ("*authorities*") dentro de las redes de enlaces que se forman entre las diversas páginas webs de la World Wide Web (www). Este algoritmo define como concentradores aquellas páginas que tienen enlaces salientes hacia un gran número de páginas y como autoridades aquellas que son ampliamente referenciadas (enlaces entrantes) por otras páginas. La noción de concentrador expresa la calidad de una página Web como un acceso a recursos útiles (por ejemplo, los directorios Web actúan como buenos "*hubs*"), y la noción de autoridad establece la calidad de la página como un recurso en sí. De esta manera, a cada página de la Web se le asignan puntajes de concentrador y de autoridad (Corbellini, 2015).

HITS tiene sus raíces en la noción de relación de refuerzo mutuo entre concentradores y autoridades. Un buen concentrador es una página que apunta a buenas autoridades, mientras que una buena autoridad es una página apuntada por buenos concentradores. A la luz de esto, el algoritmo tiene como objetivo la búsqueda de sitios web de autoridad. Ambos resultados, concentradores y autoridades, se calculan mediante un algoritmo iterativo que actualiza los resultados de una página basada en las puntuaciones de las páginas de su vecindad inmediata (Corbellini, 2015).

1.4. Detección de comunidades

Además de las métricas de centralidad, es de interés para esta investigación conocer acerca de algoritmos para realizar detección de comunidades en redes sociales pues permitirá visualizar a que grupos sociales pertenecen los actores de mayor autoridad que se

identifiquen. A continuación, se muestran los principales algoritmos para detección de comunidades identificados.

1.4.1. Propagación de etiquetas (Label Propagation)

Propagación de etiquetas es un algoritmo rápido para buscar comunidades en un grafo. Detecta estas comunidades utilizando solo la estructura de la red como guía y no requiere una función objetivo predefinida o información previa sobre las comunidades (Needham & Hodler, 2018). Consiste en propagar las etiquetas (identificadores únicos) de los nodos a través de la red y formar comunidades a través del propio proceso de propagación. Según sus autores (Raghavan, Albert, & Kumara, 2007) el algoritmo trabaja de forma siguiente:

El nodo x tiene vecinos y cada vecino acarrea una etiqueta denotando a la comunidad para la cual pertenecen. Cada nodo en la red escoge unirse a la comunidad a la cual el máximo número de sus vecinos pertenece. Al comienzo, cada nodo es inicializado con una etiqueta única (identificador) y las etiquetas se propagan a través de la red. En cada iteración de propagación, el nodo actualiza su etiqueta basándose en las etiquetas de sus vecinos. Como las etiquetas se propagan rápidamente en los grupos de nodos densamente conectados, alcanzan un consenso en una etiqueta única. En el fin de la propagación, la mayoría de las etiquetas desaparecen y sólo pocas etiquetas existen en la red. La convergencia del algoritmo se alcanza cuando cada nodo tiene la etiqueta de la mayoría de sus vecinos. Al final de la convergencia, los nodos conectados con la misma etiqueta forman una comunidad.

1.4.2. Componentes conectados (Union find)

El algoritmo Componentes conectados encuentra un conjunto de nodos conectados en un grafo donde cada nodo es accesible desde otro en el mismo conjunto mediante un camino (Needham & Hodler, 2018). Permite determinar la cantidad de subgrafos que posee el grafo principal. Se utiliza con frecuencia en un análisis preliminar para entender la estructura de la red. Esto es esencial porque los algoritmos sobre grafos pueden introducir errores si este está muy desconectado. El algoritmo fue descrito por primera vez por Bernard A. Galler y Michael J. Fischer en 1964 (Needham & Hodler, 2018).

1.4.3. Modularidad de Louvain

La modularidad es una medida de la estructura de las redes, diseñada para medir la fuerza de la división de una red en módulos (también llamados grupos o comunidades). Las redes con alta modularidad tienen conexiones densas entre los nodos dentro de los módulos, pero escasas conexiones entre los nodos en diferentes módulos. Las Redes biológicas, incluyendo los cerebros animales, exhiben un alto grado de modularidad (Kuz et al., 2016).

El algoritmo de detección de comunidades Modularidad de Louvain maximiza un puntaje de modularidad para cada comunidad, donde la modularidad cuantifica la calidad de la asignación de los nodos a las comunidades, al evaluar cuánto más densamente conectados están los nodos dentro de la comunidad, en comparación con qué tan conectados estarían en una red aleatoria. (Needham & Hodler, 2018).

Louvain es uno de los algoritmos más rápidos basados en modularidad y funciona bien con grafos grandes. También revela una jerarquía de comunidades a diferentes escalas, lo que es útil para comprender el funcionamiento global de una red (Needham & Hodler, 2018). Ha sido utilizado para proveer recomendaciones (Sundaresan, Hsu, & Chang, 2014) y para la extracción de los temas más discutidos en Twitter y YouTube (Sakaji Kido, Augusto Igawa, & Barbon Jr, 2016).

1.5. Herramientas para determinar influencia

Serrano-puche (2012) identificó a Klout y PeerIndex como las herramienta más importantes del mercado para calcular la influencia de los usuarios de las redes sociales. Aunque en la actualidad ya no están disponibles porque Klout fue absorbida por Lithium Technologies (2018) y PeerIndex por Brandwatch (2018), compañías especializadas en el marketing de influencers, resulta interesante analizarlas pues estas herramientas identificaban a los influencers teniendo en cuenta su autoridad.

Klout calculaba la influencia de una persona tomando en consideración los siguientes factores (Serrano-Puche, 2012):

- **Alcance real:** número de personas sobre las que se ejercía influencia. Respondía a la cuestión “¿a cuánta gente influyes?”. Para ello analizaba cómo las personas con las que un usuario estaba conectado interactuaban con él y con su contenido.
- **Amplificación:** “¿Cuánto influyes en la gente sobre la que tienes influencia?”. Se medía tomando en cuenta cómo los mensajes provocaban conversación o eran replicados (retuiteados, marcados como “me gusta” en Facebook, merecedores de un +1 en Google+, etc.). Cada vez que uno de sus seguidores realizaba una de estas acciones, era considerada como una señal de autoridad y calidad del contenido del usuario.
- **Impacto en la Red:** determinaba la influencia que a su vez tenían los seguidores que estaban dentro del alcance real del usuario evaluado. Medía, por tanto, la calidad de sus seguidores (“¿Cómo de influyentes eran ellos?”).

Los factores de PeerIndex eran (Serrano-Puche, 2012):

- **Autoridad:** “la medida de la confianza”, el modo en que las demás personas confiaban en las recomendaciones y opiniones del usuario evaluado, de modo general y en temas concretos. Era una posición relativa respecto al resto de usuarios en un determinado tema.
- **Audiencia:** equivalente a los parámetros de alcance real y amplificación de Klout, puesto que medía el impacto de las acciones del usuario sobre sus seguidores y la manera en que estos conectan con su contenido.
- **Actividad:** representaba la frecuencia de actualización del status del usuario en relación con el resto de la comunidad en torno a un tema. Según fuera la naturaleza del tema (y exigía una conversación en torno a él más o menos continua), una actividad excesiva podría considerarse spam y penalizar al usuario, o viceversa: que su intervención en el debate fuera tan esporádica que llegara a ser irrelevante.

Ambas herramientas medían en una escala de 0 a 100 la influencia personal de los usuarios, una cifra superior a 40 ya los caracterizaba como influyentes.

Hoy en día existen numerosas herramientas (Audiense, 2018; BuzzSumo, 2018; Followerwonk, 2018; Heepsy, 2018; Klear, 2018; Kred, 2018; SocialRank, 2018; Topsy, 2018) que permiten posicionar a los usuarios por su influencia en los medios sociales. Lo interesante de todas, es que ofrecen múltiples opciones similares. En la mayoría se puede filtrar por diversas categorías: temáticas, palabras clave, idiomas, geolocalización, cantidad de seguidores y seguidos, entre otras opciones. Otra similitud que comparten es la información que visualizan en los resultados de las búsquedas y rankings. En general muestran el perfil del usuario, contactos, un puntaje de influencia, interacciones recientes, horas de mayor actividad, alcance, temáticas más recurrentes, ubicación geográfica e incluso algunas muestran información de los seguidores y seguidos por el usuario en cuestión (Alonso et al., 2016).

Una herramienta que se destaca por su transparencia es Audiense (2018), la cual utiliza el algoritmo de Kred para calcular la influencia en Twitter y Facebook. El índice de Kred está compuesto a su vez por dos medidas, la influencia y el outreach (forma de interacción con el contenido de los seguidores). Lo que hace especial a esta herramienta es que no solo indica la influencia, sino la generosidad en las interacciones propias para con otros usuarios. Es decir, el usuario visto como seguidor (outreach) y como seguido (influencia). El puntaje de influencia se mide en una escala de 0 a 1000 y el del outreach de 1 a 12. Las transparencias de las reglas a la hora de la medición también hacen particular a esta herramienta (Alonso et al., 2016). Estas reglas pueden ser consultadas en su web oficial (Kred, 2018).

Es importante destacar que, excepto Kred, las herramientas mencionadas no publican exactamente como efectúan sus cálculos. Esto se debe porque poseen un carácter comercial. Además, están enfocadas al marketing de influencers, operan sobre los medios sociales y ninguna tiene en cuenta las interacciones de las personas en la mensajería instantánea para determinar la influencia.

1.6. Bases de Datos Orientadas a Grafos

En el contexto de las redes sociales, la elección más natural para manejar grandes cantidades de datos interconectados y distribuidos son las Bases de Datos Orientadas a Grafos (BDOG) y los frameworks para el procesamiento de algoritmos basados en grafos. Estos sistemas han adquirido un gran interés por parte de empresas e instituciones académicas que necesitan manejar grandes cantidades de datos interconectados como los datos sociales y los datos geo-espaciales. La cantidad de datos almacenados y consultados por este tipo de aplicaciones hace que las bases de datos tradicionales e incluso las NoSQL sean soluciones inviables para hacer frente a los niveles de rendimiento necesarios (Corbellini, 2015).

Una base de datos orientada a grafos es aquella que permite almacenar la información como nodos de un grafo y sus respectivas relaciones con otros nodos, permitiendo así aplicar la teoría de grafos para recorrer la base de datos; son muy útiles para guardar información en modelos con muchas relaciones como redes y conexiones sociales (Pinilla, Bello, & Peña, 2017). Un Sistema de administración de Bases de Datos Orientadas a Grafos es un gestor de bases de datos en línea con métodos para la creación, lectura, inserción, actualización y eliminación (patrón CRUD) que operan sobre un modelo de datos basado en grafos. Las BDOG son generalmente construidas para ser usadas por sistemas transaccionales (OLTP) y, por tanto, son normalmente optimizadas para rendimiento transaccional (Robinson, Webber, & Eifrem, 2015).

Existen dos propiedades de las BDOG que se deben considerar cuando se investiga este tipo de tecnología (Robinson et al., 2015):

1. **El almacenamiento subyacente:** Algunas Bases de Datos Orientadas a Grafos usan almacenamiento de grafo nativo que está diseñado y optimizado para almacenar y administrar grafos. No todas las tecnologías de BDOG usan almacenamiento nativo de grafos. Algunas serializan el grafo dentro de una base de datos relacional, una base de datos orientada a objetos o algún almacén de datos de propósito general.
2. **El motor de procesamiento:** Algunas definiciones requieren que unas BDOG usen "*index-free adjacency*", significando que los nodos conectados apuntan físicamente a cada uno de los otros en la base de datos, Aquí se toma una vista ligeramente más amplia: cualquier base de datos que desde la perspectiva del usuario que se comporte

como una BDOG califica como una BDOG. Se conocen las ventajas de rendimiento significantes de *“index-free adjacency”*, y por consiguiente se usa el término de procesamiento de grafo nativo para describir las BDOG que utilizan el *“index-free adjacency”*.

La Figura 8 muestra una visión general de algunas de las BDOG que están disponibles en el mercado, ordenándolas por su almacenamiento y modelos que procesa. Al analizarla se identificaron como las bases de datos de mejor rendimiento: OrientDB, Dex y Neo4j. En la Tabla 1 se muestran las principales características de estos sistemas.

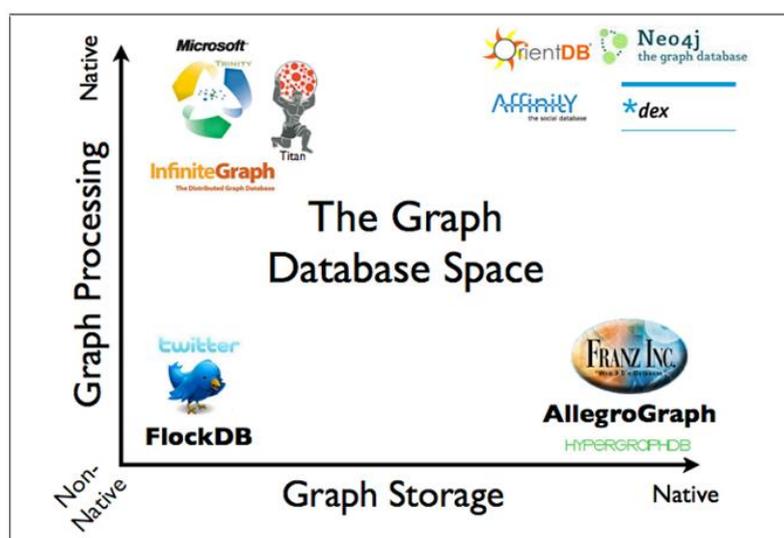


Figura 8: Visión general de las Bases de Datos Orientadas a Grafos. Fuente: (Robinson et al., 2015).

Tabla 1: Comparación de Dex, Neo4j y OrientDB. Fuente: (McColl, Ediger, Poovey, Campbell, & Bader, 2014).

	DEX	Neo4j	OrientDB
Propietario/Mantenedor	Sparsity-Technologies	Neo Technology	NuvolaBase Ltd
Licencia	Propietario	GPL/Propietario	Apache
Plataforma	x86	Java	Java
Lenguaje	C++	Java	Java
Distribución	Bin	Src/Bin	Src/Bin
Transaccional	X	X	Ambos
Basada en memoria	-	-	X
Basada en disco	X	X	X
Distribuida	-	-	X
Algoritmos de grafos	X	X	-
Lenguaje de consulta	Transversal	Cypher	Extended SQL, Gremlin
Software	-	X	X
Tipo	BDOG	BDOG/NoSQL	BDOG/NoSQL

Como se puede observar en la tabla, Neo4j y OrientDB poseen licencias libres y un software mediante el cual se puede acceder a los datos. Por otra parte, Dex y Neo4j ya tienen implementados algunos algoritmos de grafos lo cual es muy importante para esta investigación. Un dato importante es que las tres bases de datos son transaccionales lo cual

garantiza que cuando reciban varias peticiones concurrentes se conservará la atomicidad, consistencia, aislamiento y durabilidad de los datos.

Se propone utilizar Neo4J porque además posee una amplia comunidad de desarrollo que muestra ejemplos de los beneficios que ha traído a las compañías que actualmente la utilizan. Esta BDOG es capaz de manejar billones de nodos y relaciones, utiliza consultas de tipo transversales para recorrer de forma rápida las relaciones, tiene un lenguaje de consultas propio (“*Cypher*”), posee balanceador de carga para uso distribuido en diferentes servidores y provee una API que permite su uso en diferentes lenguajes de programación.

Conclusiones del capítulo

Mediante el estudio realizado se pudo constatar lo siguiente:

1. La representación de redes sociales como un grafo es muy útil para realizar análisis de influencia social pues posibilita aplicar diversos algoritmos y métricas de centralidad que analizan estructuralmente el grafo y permiten inferir conocimiento que de otra forma sería muy difícil de detectar como, los nodos más conectados, los distintos grupos sociales, los nodos que conectan esos grupos, etc.
2. Muchos trabajos relacionados con el análisis de influencia en redes de interacción social, combinan varias métricas de centralidad para determinar los actores más influyentes de la red teniendo en cuenta la estructura del grafo que representan. Sin embargo, no siempre es suficiente conocer esto para medir la influencia. En ocasiones es necesario conocer también que reacciones provocan las acciones de un usuario sobre otro, por ejemplo, si un usuario publica una foto en Facebook, aquellos usuarios que comentan la foto, la marquen como “*me gusta*” o la compartan están reaccionando a esta acción y por lo tanto están siendo influenciados.
3. En internet se encuentran herramientas disponibles que permiten calcular autoridad e influencia en los “*medios sociales*” más populares como Facebook, Twitter e Instagram, pero por su carácter comercial, la mayoría no publican exactamente como efectúan estos cálculos. Además, tampoco analizan la influencia en la mensajería instantánea.
4. El uso de bases de datos orientadas a grafos posibilita almacenar y consultar eficientemente la información presente en los grafos que se generen a partir de las interacciones sociales.
5. Para un correcto análisis de influencia y autoridad en una red de mensajería instantánea hay que tener en cuenta tanto el análisis estructural del grafo con el cual se represente dicha red (jerarquía) como las reacciones que provoca un usuario a otro cuando le escribe (credibilidad). La autoridad de una persona en una red social está muy relacionada con su jerarquía y credibilidad.

CAPÍTULO 2. MÉTRICA DE AUTORIDAD PARA DETERMINAR EL NIVEL DE INFLUENCIA DE LAS PERSONAS EN LA MENSAJERÍA INSTANTÁNEA

En este capítulo se presenta la métrica de autoridad que fue definida para calcular la influencia de los usuarios de la mensajería instantánea (Autoridad Total), sus componentes (Autoridad Estructural y Autoridad de Reacción) y las fórmulas y algoritmos desarrollados para su cálculo. Se presenta a GRAPH-CHAT, herramienta que implementa estas métricas y fue desarrollada para visualizar en un grafo las relaciones de escritura de mensaje de estos usuarios, resaltando los de mayor autoridad y los grupos sociales a los que pertenecen. Se exponen además las potencialidades de GRAPH-CHAT para apoyar la creación de equipos de trabajo.

2.1. Cálculo de Autoridad Total (AT), la Autoridad Estructural (AE) y Autoridad de Reacción (AR) en la mensajería instantánea

La Autoridad Total (AT) de un usuario en una red de la mensajería instantánea es un valor, entre cero y cien [0,100], que representa cuán influyente es esta persona en la red. Para su definición se tuvieron en cuenta aspectos que caracterizaban la autoridad desde el punto de vista de la psicología y la sociología (Cialdini, 2009), así como los criterios de autoridad que eran utilizados en, las herramientas de medición de influencia, Klout y PeerIndex (Serrano-Puche, 2012). Estos aspectos, que pueden ser consultados en el Capítulo 1, sirvieron para determinar que la autoridad de una persona en una red social está estrechamente relacionada con su jerarquía y credibilidad. Es por ello que la métrica desarrollada está compuesta por dos componentes o sub-métricas:

- **Autoridad Estructural (AE):** encargada de medir la jerarquía de los usuarios de la red de mensajería instantánea a partir de representarla como un grafo y calcular el valor de las métricas de centralidad del grafo.
- **Autoridad de Reacción (AR):** encargada de medir la credibilidad tomando en consideración la inmediatez con que un mensaje escrito, recibe respuestas en la mensajería instantánea y la cantidad de respuestas.

La Autoridad Total de un usuario se calcula hallando la media entre su Autoridad Estructural y su Autoridad de Reacción, como se muestra en la Fórmula 3.

$$AT = \frac{(AR + AE)}{2} \quad (3)$$

2.1.1. Autoridad Estructural (AE)

La Autoridad Estructural de un usuario se calcula a partir de su centralidad dentro del grafo. Tiene como base la definición de Actor Central de Freeman (1979): “(...) *un actor es considerado el más central si tiene un alto grado o es cercano a los demás o si permite interconectar otros actores en la red (...)*”, este criterio, a pesar de ser antiguo, todavía es utilizado en la actualidad por diversos investigadores (Bello-Lara, 2018; Beveridge & Shan, 2016; García et al., 2016; He et al., 2018; Kuz et al., 2016; Li et al., 2018; Needham & Hodler, 2018; Riquelme et al., 2017; Sudhahar et al., 2015; Zhao et al., 2018).

Para el cálculo de la AE se tienen en cuenta la Centralidad de Grado (D), porque representa la popularidad (cantidad de contactos) del usuario, la Centralidad de Cercanía (C), pues permite identificar cuán rápido se puede relacionar este usuario con los demás, la Centralidad de Intermediación (B), porque permite detectar aquellos usuarios que conectan las distintas comunidades que se forman en la red, así como el PageRank (PR) ya que permite identificar cuán importante es el usuario a partir de la importancia de sus contactos. Los valores de estas métricas son normalizados utilizando la técnica Normalización por el mayor elemento y luego son combinados mediante el Operador de Agregación⁴ Media Ponderada. Luego de aplicado el Operador de Agregación, el resultado final es multiplicado por 100 para que los valores de la AE se encuentren en el intervalo [0,100].

La Normalización por el mayor elemento, técnica muy utilizada por la comunidad científica, consiste en dividir cada elemento de una columna entre el mayor elemento de dicha columna (Bellver & Guijarro, 2012). Con esta variante se garantiza que el resultado se encuentre en el intervalo [0,1] y se mantenga la proporcionalidad (Bello-Lara, 2018). Matemáticamente se define como se muestra a continuación: (Bellver & Guijarro, 2012).

$$X_i \text{normalizado} = \frac{X_i}{\max X_i} \quad (4)$$

La Media Ponderada (WA, por sus siglas en inglés) es uno de los operadores de agregación más empleados en la literatura (Bello-Lara, 2018). Su definición es la siguiente:

- Un operador WA tiene asociado un vector de pesos α , con $\alpha \in [0,1]$ y $\sum_1^n \alpha_i = 1$, teniendo la siguiente forma:

⁴ Tipo función matemática empleada para la fusión de la información. Combinan n valores en un dominio D y devuelven un valor en el mismo dominio (Torra & Narukawa, 2007)

$$WA(X_1, \dots, X_2) = \sum_{i=1}^n X_i \alpha_i \quad (5)$$

donde α_i representa la importancia o relevancia de la fuente de datos X_i .

En el contexto de esta investigación X constituye el vector que almacena las métricas de centralidad, $X = \langle D, C, B, PR \rangle$ y α el vector de las ponderaciones que se aplican a cada métrica, para determinar cuánto aporta para la Autoridad Estructural (AE), $\alpha = \langle \alpha_D, \alpha_C, \alpha_B, \alpha_{PR} \rangle$. El valor de α es determinado a juicio del experto que vaya a utilizar la fórmula y puede variar tomando en cuenta el entorno de datos donde se vaya a aplicar.

La fórmula general para calcular la AE es la siguiente:

$$AE = (\sum_{i=1}^n X_i \text{normalizado} * \alpha_i) * 100 \quad (6)$$

Si despejamos los valores de X obtenemos la fórmula ampliada:

$$AE = \left(\frac{D}{\max D} * \alpha_D + \frac{C}{\max C} * \alpha_C + \frac{B}{\max B} * \alpha_B + \frac{PR}{\max PR} * \alpha_{PR} \right) * 100 \quad (7)$$

En la siguiente tabla se muestran las ponderaciones que fueron definidas para cada métrica de centralidad en el entorno de datos utilizado por esta investigación. Para su definición se tuvieron en cuenta los criterios expuestos por los autores antes mencionados, además de una consulta realizada a un sociólogo de la Universidad de las Ciencias Informáticas.

Tabla 2: Ponderaciones definidas para cada Métrica de Centralidad. Fuente: elaboración propia.

Métrica de Centralidad	Valor de α_i	Ponderación
Grado	α_D	0,2
Cercanía	α_C	0,2
Intermediación	α_B	0,2
PageRank	α_D	0,4

Para el cálculo de la AE en un grafo $G(V, E)$ se define el siguiente algoritmo:

Tabla 3: Algoritmo para calcular la AE. Fuente: elaboración propia.

Entradas	$V(G)$ donde cada nodo v_i tiene como atributos los valores de las métricas del vector X , arreglo de ponderaciones α .
Salida	$V(G)$ con sus nodos v_i enriquecidos con el valor de su AE como atributo
INICIO	
1. Inicializar $AE := 0$	
2. $\max D := \text{ObtenerMayorGrado}(V(G))$	
3. $\max C := \text{ObtenerMayorCercanía}(V(G))$	
4. $\max B := \text{ObtenerMayorIntermediación}(V(G))$	
5. $\max PR := \text{ObtenerMayorPageRank}(V(G))$	

6. Para cada $v_i \in V(G)$

$$6.1. AE := \left(\frac{D}{\max D} * \alpha_D + \frac{C}{\max C} * \alpha_C + \frac{B}{\max B} * \alpha_B + \frac{PR}{\max PR} * \alpha_{PR} \right) * 100$$

$$6.2. v_i.AE := AE$$

7. Fin para

Retornar $V(G)$

FIN

2.1.2. Autoridad de Reacción (AR)

Esta métrica se basa en el criterio de que cuando alguien recibe un mensaje de otra persona que considera importante, le responde en la menor brevedad posible. Es decir, para que un usuario tenga autoridad sobre otro en la mensajería instantánea, debe recibir respuesta de sus mensajes en un intervalo de tiempo inmediato, el cual se denotará con la letra I . El intervalo puede variar en dependencia del uso que hagan de la mensajería instantánea las personas que se estén analizando, por lo que esto se debe de tener en cuenta para su definición.

Dados dos usuarios u y v , la Autoridad de Reacción de un mensaje escrito por u a v (ARM_{uv}), que fue enviado en un instante de tiempo t_0 , se calcula contando la cantidad de respuestas que ha recibido el mensaje de u en un intervalo de tiempo que comprende desde t_0 hasta $t = t_0 + I$. Ver Fórmula 8.

$$ARM_{uv} = \sum R_{vu} \quad (8)$$

- R_{vu} es una respuesta recibida, consiste en un mensaje donde v es el emisor, u es el receptor y la fecha del mensaje ($fecha_R$) está en el intervalo de tiempo ($t_0 < fecha_R < t_0 + I$).

La Autoridad de Reacción de un usuario u sobre otro v consiste en el promedio de la ARM_{uv} de cada mensaje que ha enviado u a v . Esto se conoce como la Autoridad de Reacción Local de u sobre v (ARL_{uv}). Ver Fórmula 9.

$$ARL_{uv} = \frac{\sum ARM_{uv}}{m} \quad (9)$$

- m es la cantidad de mensajes que u ha enviado a v .

La Autoridad de Reacción total de un usuario u (AR_u), se calcula sumando todas sus ARL . Ver Fórmula 10.

$$AR_u = \sum_{i=1}^n ARL_{ui} \quad (10)$$

- n es la cantidad de vecinos de u
- i es el vecino de u que se está analizando en cada caso.

El último paso para calcular la AR es normalizar sus valores utilizando también la técnica Normalización por el mayor elemento y posteriormente multiplicarlos por 100 para que se encuentren en el intervalo $[0,100]$ y puedan ser utilizados para calcular la Autoridad Total, como se muestra en la Fórmula 11.

$$AR_{normalizada} = \frac{AR_u}{\max AR} * 100 \quad (11)$$

A continuación se presentan los algoritmos desarrollados para calcular la AR de cada nodo de un grafo $G(V, E)$:

Tabla 4: Algoritmo para calcular la AR de cada nodo de un grafo G. Fuente: elaboración propia.

Entradas	Un grafo dirigido $G(V, E)$ donde cada elemento de E representa la relación de escritura de mensaje entre los nodos, un número entero (I) que representa el rango de tiempo en el cual se van a contar los mensajes de respuesta.
Salida	Grafo $G(V, E)$ con sus nodos $V(G)$ enriquecidos con el valor de su AR como atributo
INICIO 1. $AR_u := 0$ 2. $listaVecinos := null$ 3. Para cada $u \in V(G)$ 3.1. $listaVecinos := u.obtenerVecinos$ 3.2. para cada $v \in listaVecinos$ 3.2.1. $AR_u += CalcularARL_{uv}(u, v, I)$ 3.3. Fin para 3.4. $u.AR := AR_u$ 4. Fin para 5. <i>NormalizarAR</i> FIN	

Tabla 5: Algoritmo para calcular la ARL_{uv} . Fuente: elaboración propia.

Entradas	Dos nodos vecinos u, v y un número entero (I) que representa el rango de tiempo en el cual se van a contar los mensajes de respuesta.
Salida	ARL_{uv}
INICIO 1. $ARL_{uv} := 0$ 2. $listaEnviado := u.mensajesEnviados(v)$ 3. Para cada $M \in listaEnviado$ 3.1.1. $SumaARM_{uv} += CalcularARM_{uv}(M, I)$ 3.2. Fin para	

4. $ARL_{uv} := \text{Suma}ARL_{uv} / \text{listaEnviado.cantidad}$
 5. Retornar ARL_{uv}
- FIN**

Tabla 6: Algoritmo para calcular la ARM_{uv} . Fuente: elaboración propia.

Entradas	Arista M que representa un mensaje enviado por u a v y un número entero (I) que representa el rango de tiempo en el cual se van a contar los mensajes de respuesta.
Salida	ARM_{uv}
Descripción	Cuenta los mensajes desde v a u en el rango de tiempo especificado por I .
INICIO	
1. $ARM_{uv} := 0$	
2. $\text{listaMsg} := M.\text{receptor}_v.\text{mensajesEnviados}(u)$	
3. Para cada $\text{msg} \in \text{listaMsg}$	
3.1. Si $M.\text{fecha} < \text{msg}.\text{fecha} < M.\text{fecha} + I$	
3.1.1. $\text{respuestas}++$	
3.2. Fin para	
4. Retornar respuestas	
FIN	

Tabla 7: Algoritmo para normalizar la AR de cada nodo de un grafo G . Fuente: elaboración propia.

Entradas	Un grafo dirigido $G(V, E)$ donde cada nodo $V(G)$ tiene su AR calculada.
Salida	Grafo $G(V, E)$ con la AR de cada nodo $V(G)$ normalizada.
INICIO	
1. $\text{maxAR} := \text{Buscar la mayor AR}$	
2. Para cada $u \in V(G)$	
2.1. $AR_{normalizada} := (u.AR / \text{maxAR}) * 100$	
2.2. $u.AR := AR_{normalizada}$	
3. Fin para	
FIN	

2.2. GRAPH-CHAT: Herramienta para la visualización de la autoridad de los usuarios de la mensajería instantánea y la creación de equipos de trabajo

GRPH-CHAT es una herramienta informática desarrollada para ser utilizada en la Universidad de las Ciencias Informáticas (UCI). Permite, a partir de los registros de los Servidores de mensajería Instantánea (SMI), obtener un grafo donde sus usuarios son los nodos y las aristas son las relaciones de escritura de mensaje que se establecen entre ellos. Los nodos del grafo poseen además la información de cada usuario que está disponible en el LDAP⁵ de la Universidad. El grafo muestra los distintos grupos sociales y los actores de mayor autoridad

⁵ LDAP corresponde con las siglas en inglés “*Lightweight Directory Access Protocol*”: es un protocolo que permite el acceso a través de la red a un servicio de directorio (Romano Poveda & Rubén, 2017).

en la red. La herramienta permite utilizar la información almacenada en el grafo para sugerir posibles equipos de trabajo teniendo en cuenta el contexto deseado: Colectivo de asignatura, Proyecto de investigación o Proyecto productivo. Estos tres contextos constituyen los procesos fundamentales de la UCI (Formación, Investigación y Producción). Con el uso de GRAPH-CHAT se puede apoyar la toma de decisiones relacionadas con estos aspectos.

La herramienta posee una arquitectura cliente-servidor (ver Figura 9), donde la comunicación se realiza mediante un API de servicios REST⁶. Para almacenar la información de los usuarios de la mensajería instantánea y sus relaciones, la herramienta utiliza la base de datos orientada a grafos (BDOG) Neo4j en su versión 3.3.0. Usa además la base de datos H2⁷ en su versión 1.4.194 para almacenar los recursos necesarios para las funcionalidades relacionadas con la creación de los equipos de trabajo. Ejemplo de estos recursos lo constituyen las asignaturas, las líneas, proyectos y grupos de investigación, entre otros. Es importante resaltar que, aunque GRAPH-CHAT fue desarrollada para ser utilizada en la UCI, se puede adaptar sin mucha dificultad a cualquier otra organización debido a que posee una arquitectura flexible y orienta a servicios.

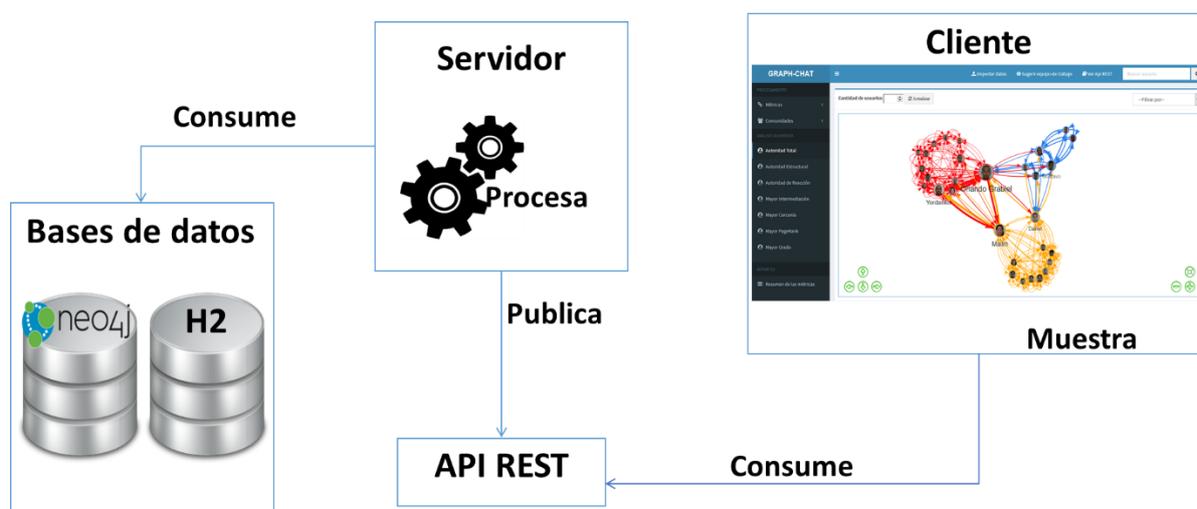


Figura 9: Arquitectura general del GRAPH-CHAT. Fuente: elaboración propia.

Para la implementación de las funcionalidades del lado del servidor se utilizó el lenguaje de programación Java en su versión 8, utilizando las tecnologías relacionadas con el Framework Spring-Boot en su versión 1.5.3. Se emplea Spring-Data-Neo4j para acceder a la base de datos y Spring-Fox 2.7 para generar la documentación del API de servicios REST (ver Figura 10).

⁶ REST: Transferencia de Estado Representacional, por sus siglas en inglés. Los Servicios Web basados en REST utilizan solicitudes HTTP para enviar y recibir datos, utilizando JSON como forma de intercambio (Otegui & Guralnick, 2016).

⁷ H2: Base de datos relacional nativa de Java.



Figura 10: API REST de GRAPH-CHAT. Fuente: elaboración propia.

El lado del cliente se implementó como una aplicación web de una sola página que consume, mediante AJAX, el API de servicios que brinda el servidor para obtener los datos que necesita en la ejecución de cada una de sus funcionalidades. En su desarrollo se utilizaron los lenguajes de programación HTML, CSS y JavaScript. Se emplearon además los siguientes frameworks de JavaScript: jQuery 1.11, en la implementación de las funcionalidades del lado del cliente, y Viz.js 4.21 para dibujar el grafo. El Framework Bootstrap 3.0 fue utilizado para maquetar y darle estilo a la página web. En la siguiente imagen se muestra el cliente desarrollado para GRAPH-CHAT.

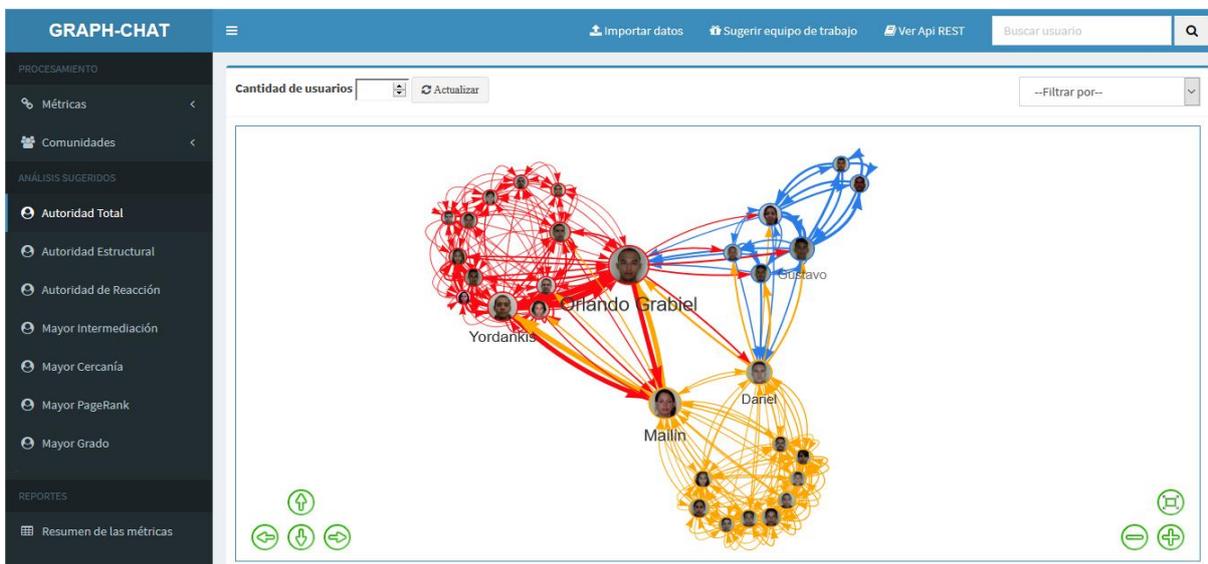


Figura 11: Interfaz principal del cliente de GRAP-CHAT. Fuente: elaboración propia.

Las funcionalidades que brinda GRAPH-CHAT están estructuradas en tres componentes: Importación de los datos, Procesamiento y Análisis. La Figura 12 muestra el flujo de

actividades básico a seguir cuando se interactúa por primera vez con la herramienta. Se aprecia claramente la relación que existe entre sus componentes.



Figura 12: Flujo de actividades básico de GRAP-CHAT. Fuente: elaboración propia.

2.2.1. Importación de los datos

Este componente está relacionado con todo lo referente a la información que se necesite gestionar para generar el grafo, calcular la autoridad que tiene cada uno de sus usuarios y sugerir los equipos de trabajo. Sus entradas son los registros del Servidor de mensajería instantánea (SMI) previamente almacenados en un archivo en formato CSV, los datos disponibles en el LDAP y la información que se registre de forma manual. Su salida lo constituye un grafo donde sus nodos representan los usuarios de la mensajería instantánea, almacenando su información básica, de carácter profesional y de carácter científico. Las aristas del grafo representan las relaciones de escritura de mensaje y almacenan el usuario emisor, el receptor, la fecha en que intercambiaron cada mensaje y la cantidad de mensajes que han intercambiado. Esta información es guardada en Neo4j en un diseño similar al de la Figura 13.



Figura 13: Diseño de la base de datos que almacenará el grafo. Fuente: elaboración propia.

Los datos que se obtienen de los registros del SMI son utilizados para generar la estructura del grafo, es decir: sus nodos y aristas. Debido a que existen diversos SMI, la estructura de estos registros puede ser diferente en dependencia del SMI que se esté utilizando. Es por ello que para su importación es necesario que primero se hayan normalizado y estandarizado en un fichero en formato CSV, donde cada línea represente un mensaje escrito. Las líneas de este fichero deben poseer la siguiente estructura:

- “usuario emisor”, “usuario receptor”, “fecha del mensaje”.

En la siguiente figura se muestra un fragmento de un fichero con este formato.

```

1 "jperezg", "jdelgado", "2018-06-20 17:34:01"
2 "dcorchado", "ogtoledano", "2018-06-20 07:52:45"
3 "jperezg", "jdelgado", "2018-06-20 10:40:27"

```

Figura 14: Fragmento del fichero datos.csv. Fuente: elaboración propia.

La información personal a almacenar de cada usuario se determinó atendiendo a la necesidad de darle semántica al grafo para poder identificar las posibles causas de la autoridad y para que pudiera sugerir mejor los equipos de trabajo. Mientras más información se posea de los nodos, más precisos serán los análisis y las sugerencias. La información se clasifica según los siguientes criterios:

- Básica (usuario, nombre, apellidos, sexo, municipio, provincia).
- Profesional (área, categoría, cargo, asignaturas que imparte, categoría docente, rol en la producción)
- Científica (grado científico, proyecto de investigación, grupo de investigación, línea de investigación).

Su obtención se puede realizar de forma automática, accediendo al fichero en formato CSV y a los datos que se encuentran publicados en el LDAP de la organización, o introduciéndola de forma manual en la herramienta. Esta última opción se desarrolló porque no siempre en el LDAP se encuentra toda la información que se necesita. El algoritmo que se utiliza para importar la información de forma automática es el siguiente:

Tabla 9: Algoritmo para importar los datos de forma automática. Fuente: elaboración propia.

Entradas	Fichero en formato CSV
Salida	Datos del grafo almacenados en Neo4j (BDOG)
INICIO	
<ol style="list-style-type: none"> 1. Cargar al servidor el fichero en formato CSV. 2. Abrir el fichero <ol style="list-style-type: none"> 2.1. Para cada línea del fichero <ol style="list-style-type: none"> 2.1.1. Leer el usuario emisor, el receptor y la fecha del mensaje 2.1.2. Buscar si el usuario emisor ya existe en la BDOG <ol style="list-style-type: none"> 2.1.2.1. Si no existe <ol style="list-style-type: none"> 2.1.2.1.1. Crear el usuario 2.1.2.1.2. Buscar sus datos personales en el LDAP 2.1.2.1.3. Guardar el usuario en la BDOG 2.1.2.2. Fin si 2.1.3. Buscar si el usuario receptor ya existe en la BDOG <ol style="list-style-type: none"> 2.1.3.1. Si no existe <ol style="list-style-type: none"> 2.1.3.1.1. Crear el usuario 2.1.3.1.2. Buscar sus datos personales en el LDAP. 2.1.3.1.3. Guardar el usuario en la BDOG. 2.1.3.2. Fin si 2.1.4. Buscar si existe la relación de escritura de mensaje entre los usuarios leídos <ol style="list-style-type: none"> 2.1.4.1. Si no existe <ol style="list-style-type: none"> 2.1.4.1.1. Crear una lista para almacenar las fechas de los mensajes 2.1.4.1.2. Crear la relación con el emisor, el receptor y la lista 	

- 2.1.4.2. Fin si
 - 2.1.5. Adicionar la fecha del mensaje a la lista que posee la relación
 - 2.1.6. Actualizar la relación
 - 2.2. Fin para
 - 3. Cerrar el fichero
- FIN**

2.2.2. Procesamiento

El componente de Procesamiento recibe como entrada el grafo que fue construido en el componente anterior. Se encarga de enriquecer las propiedades de los nodos de este grafo (Ver Figura 15) a partir de la detección de la comunidad a la que pertenecen, del cálculo de su Grado, de su Intermediación, de su Cercanía y de su PageRank, métricas de centralidad necesarias para inferir la Autoridad Estructural (AE). Calcula también la AE y la Autoridad de Reacción (AR).

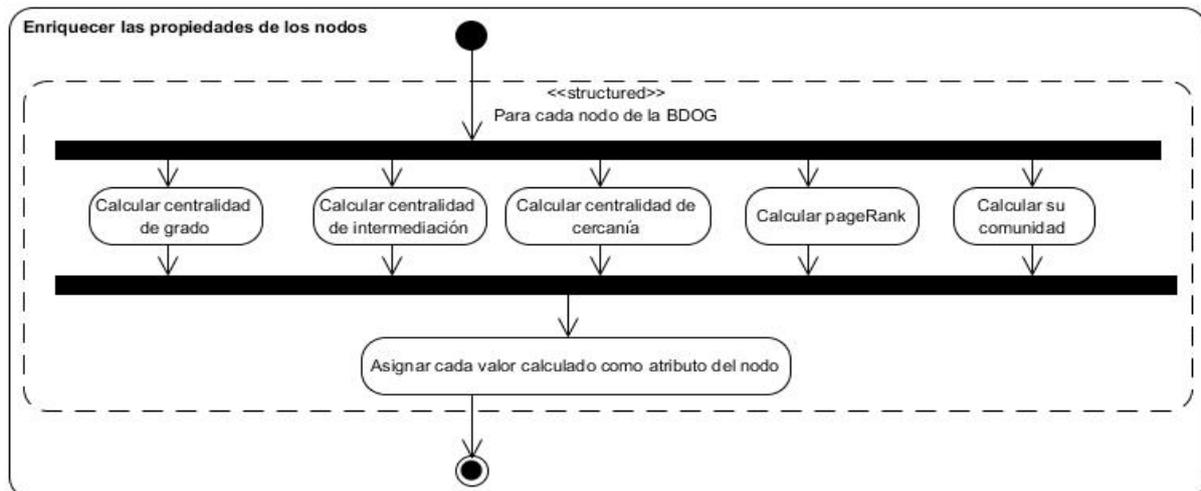


Figura 15: Flujo de actividades a seguir para enriquecer los nodos del grafo (Fuente: elaboración propia).

Para el cálculo de las métricas de centralidad y la detección de las comunidades, mediante los algoritmos Propagación de Etiquetas y Modularidad de Louvain, se utiliza una extensión de algoritmos sobre grafos de Neo4j (Needham & Hodler, 2018; Neo4j, 2018) que tiene implementada diversas funciones para realizar estos cálculos en paralelo y de forma eficiente. Las funciones de esta extensión se pueden invocar mediante consultas a Neo4j utilizando el lenguaje “*Cypher*”. La centralidad de grado, se calcula con una consulta “*Cypher*” que cuenta la cantidad de aristas que entran y salen de cada nodo.

La salida de este componente es el grafo que recibió como entrada, con la diferencia de que a sus nodos se les agregan como atributos una referencia sobre la comunidad a la que pertenece, los valores de las centralidades calculadas, su AE y su AR. Esta nueva información

recibe la clasificación de Calculada. La Figura 16 muestra el flujo básico de actividades a seguir cuando se interactúa con este componente.

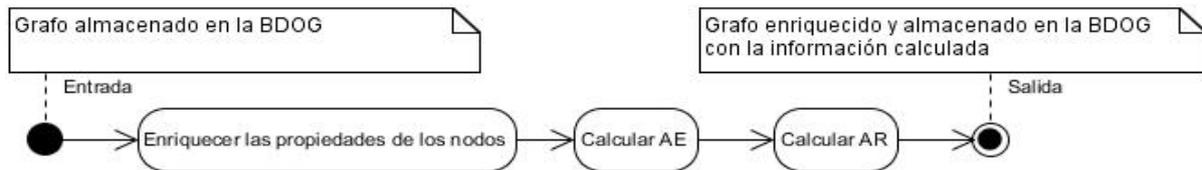


Figura 16: Flujo básico de actividades del componente de Preprocesamiento (Fuente: elaboración propia).

2.2.3. Análisis

El componente de Análisis recibe como entrada el grafo que se ha procesado y se encuentra almacenado en la BDOG. Su principal funcionalidad es mostrar la información que se ha obtenido, posibilitando que se pueda utilizar para identificar los grupos sociales, los usuarios de mayor autoridad y confeccionar equipos e trabajo. Consta de varios módulos: Visualización, Filtros, Interpretación y Sugerencia, los cuales están estrechamente relacionados. Los siguientes diagramas de actividades (Figuras 17 y 18) muestran los pasos que hay que seguir para identificar los grupos sociales, los usuarios de mayor autoridad o utilizar las potencialidades de GRAPH-CHAT para crear un equipo de trabajo teniendo en cuenta, además de las competencias profesionales de sus miembros, la existencia de buena comunicación entre ellos, medida en la mensajería instantánea.

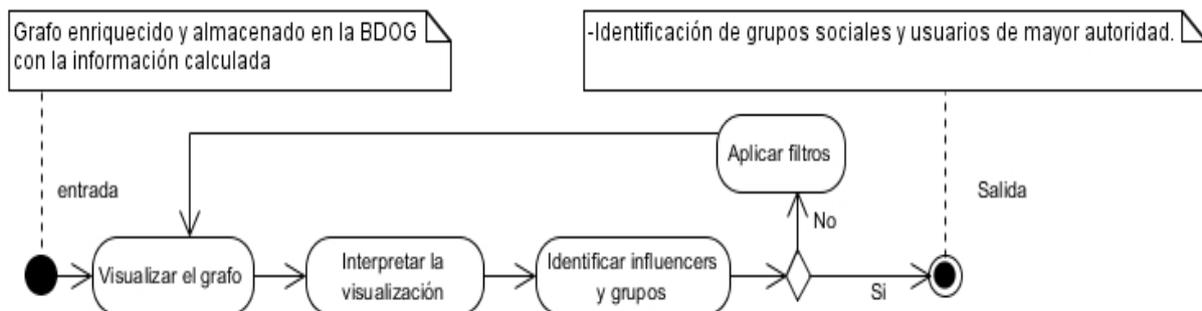


Figura 17: Flujo de actividades para identificar influencers y grupos sociales. Fuente: elaboración propia.

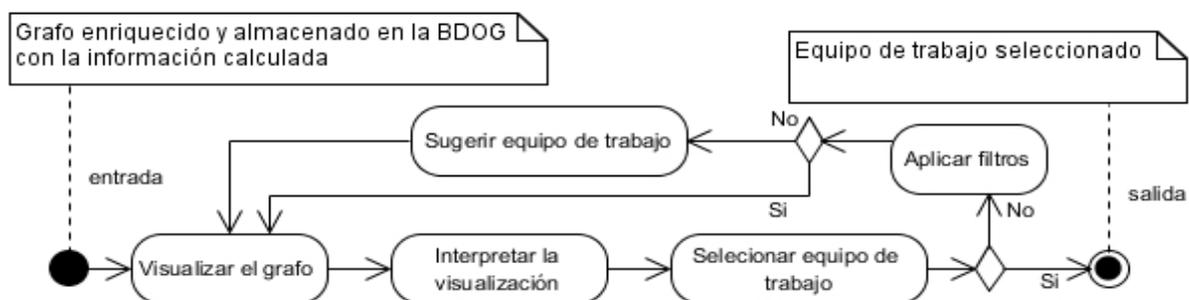


Figura 18: Flujo de actividades para seleccionar equipos de trabajo. Fuente: elaboración propia.

Visualización

El módulo de visualización es el encargado de mostrar un grafo dirigido y ponderado donde el tamaño de los nodos representa su Autoridad Total, aunque también brinda la posibilidad de acotar este tamaño solamente por la Autoridad Estructural, por la Autoridad de Reacción, por el Grado, por la cercanía, por la Intermediación o por el PageRank. Las aristas, que representan la relación de intercambio de mensaje entre los usuarios, están formadas por flechas que apuntan hacia el receptor del mensaje. El grosor de las aristas lo determina la cantidad de mensajes que han sido escritos por el usuario emisor al receptor. El color de los nodos y las aristas es determinado por la comunidad o grupo social al que pertenece el usuario.

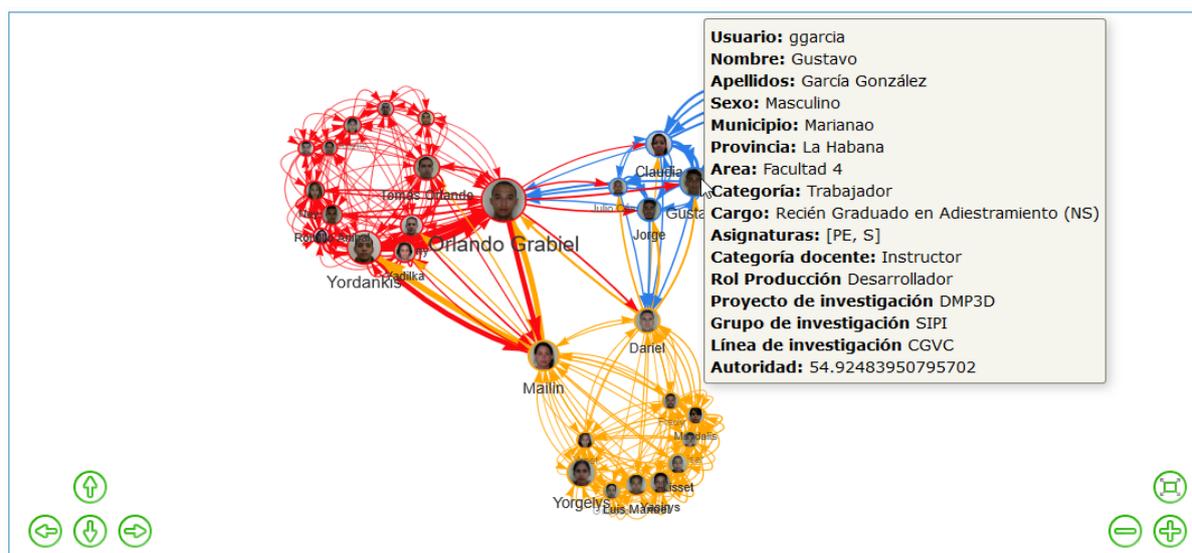


Figura 19: Grafo mostrado por el componente de visualización. Fuente: elaboración propia.

Visualización permite mostrar además la foto del usuario que representa cada nodo y sus datos personales, cuando se posiciona el puntero encima de él. Si se desea introducir manualmente alguna información del usuario que no se pueda obtener del LDAP, solo hay que hacer doble click en el nodo y se mostrará una ventana que permitirá esta acción. El grafo que se visualiza es interactivo: se puede acercar, alejar y mover a la posición de la pantalla deseada. El nombre de los nodos de mayor tamaño es mostrado por defecto, pero si se acerca el grafo se muestra también el nombre del resto de los nodos.

Filtros

La herramienta brinda la posibilidad de filtrar los nodos del grafo a partir de las propiedades de los usuarios que se encuentran almacenados en la BDOG, siempre ordenados por su autoridad. La utilización de uno o varios filtros crea una consulta dinámica que busca la información solicitada en la BDOG. Por ejemplo, si se quisiera mostrar los datos de los 10 profesores de la Facultad 4 de mayor Autoridad Total (ver Figura 20), solo habría que hacer lo siguiente:

1. Seleccionar el grafo que está acotado por la Autoridad Total.
2. Establecer en 10 la cantidad de usuarios a mostrar.
3. Filtrar los nodos por su categoría, cargo y área con los valores: “Trabajador”, “Profesor” y “Facultad 4” respectivamente.

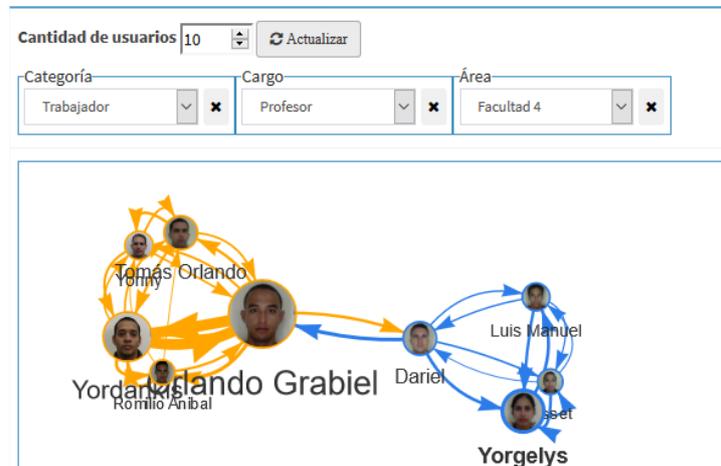


Figura 20: Resultado de la búsqueda de los 10 profesores de la Facultad 4 de mayor Autoridad Total. Fuente: elaboración propia.

Con la utilización de los filtros se pueden conformar los equipos de trabajo que se necesiten atendiendo las necesidades de diferentes procesos de la organización, los cuales a partir de ahora se identificarán como “contextos”. En esta investigación se definieron tres contextos en los que esta herramienta ayudaría en la creación de equipos de trabajo, aunque cabe la posibilidad que pueda utilizarse en otros. Estos contextos son: Colectivo de asignatura, Proyecto de investigación y Proyecto de desarrollo; los cuales coinciden respectivamente con los tres procesos fundamentales de la UCI (Formación, Investigación, Producción).

Sugerencia

El Módulo Sugerencia se desarrolló con el objetivo de facilitar la creación de equipos de trabajo en los contextos identificados. Consiste en completar un formulario (ver Figura 21) con la información que se necesita en cada contexto y luego la herramienta se encarga de crear un equipo donde se cumplan las condiciones del contexto, la relación entre sus usuarios sea fuerte y se muestre la persona de mayor categoría docente, grado científico y Autoridad Total como su posible líder.

Figura 21: Formularios para sugerir equipos de trabajo. Fuente: elaboración propia.

Si lo que se quiere generar es un equipo que constituirá un Colectivo de asignatura se debe especificar la asignatura, el área a la que pertenece el colectivo y la cantidad de miembros. Si el equipo es para un Proyecto de investigación se debe especificar la línea de investigación a la que estará ligada el proyecto y también la cantidad de miembros. Por otra parte, si se quiere conformar un equipo para un Proyecto de desarrollo se debe especificar la línea de investigación, la cantidad de analistas y la cantidad de desarrolladores. Para la creación de cada equipo de trabajo se desarrolló un algoritmo que opera realizando consultas a la BDOG y devuelve los equipos de trabajo en el formato que necesita el módulo de Visualización para mostrarlo en forma de grafo (ver Figura 22). El algoritmo es mostrado en la siguiente tabla:

Tabla 10: Algoritmo para crear un equipo de trabajo. Fuente: elaboración propia.

Entradas	Contexto, nombre de la asignatura, nombre del área, cantidad de miembros, línea de investigación, cantidad de analistas, cantidad de desarrolladores.
Salida	Colectivo de asignatura seleccionado y mostrado en forma de grafo
INICIO	
<ol style="list-style-type: none"> 1. Si <i>Contexto</i> = <i>Colectivo de asignatura</i> <ol style="list-style-type: none"> 1.1. Buscar los usuarios de la BDOG que hayan impartido la asignatura y pertenezcan al área, ordenados de mayor a menor por la categoría docente, autoridad y grado científico. 1.2. Buscar las relaciones entre estos usuarios. 1.3. Crear un grafo con la cantidad de usuarios solicitada, teniendo en cuenta el orden y la fortaleza de las relaciones. 2. Si no Si <i>Contexto</i> = <i>Proyecto de investigación</i> <ol style="list-style-type: none"> 2.1. Buscar los usuarios de la BDOG que pertenezcan a la línea de investigación, ordenados de mayor a menor por el grado científico y la autoridad. 2.2. Buscar las relaciones entre estos usuarios. 2.3. Crear un grafo con la cantidad de usuarios solicitada, teniendo en cuenta el orden y la fortaleza de las relaciones. 3. Si no Si <i>Contexto</i> = <i>Proyecto de desarrollo</i> 	

- 3.1. Buscar los usuarios de la BDOG que pertenezcan a la línea de investigación y su rol sea Analista, ordenados de mayor a menor por el grado científico y la autoridad, limitado por la cantidad de analistas.
 - 3.2. Buscar los usuarios de la BDOG que pertenezcan a la línea de investigación y su rol sea Desarrollador, ordenados de mayor a menor por el grado científico y la autoridad, limitado por la cantidad de desarrolladores.
 - 3.3. Buscar las relaciones entre estos usuarios, tanto analistas como desarrolladores.
 - 3.4. Crear un grafo con la cantidad de analistas y desarrolladores solicitada, teniendo en cuenta el orden y la fortaleza de las relaciones.
4. Retornar el Grafo.
- FIN**



Figura 22: Equipo de trabajo de 8 miembros sugerido para un nuevo Proyecto de Investigación de la línea Informática Aplicada en la Sociedad. Fuente: elaboración propia.

Interpretación

Con el módulo de Interpretación lo que se persigue es mostrar información que brinde respuesta a dos interrogantes muy importantes: ¿por qué el usuario de mayor autoridad tiene ese valor?, ¿cuán similares son los usuarios del equipo de trabajo sugerido?

Para dar solución a la primera interrogante se sigue el siguiente criterio: la autoridad de un usuario en una red informal de interacción social está muy ligado a su similitud con su medio social (Cialdini, 2009; García et al., 2016). Es por ello que la herramienta muestra una tabla (ver Figura 21) donde expone la razón de semejanza que existe entre las propiedades de los usuarios del grafo con respecto a las del usuario de mayor autoridad, llamado a partir de ahora influyente. Aquellas propiedades que posean una razón de semejanza alta son las que van a determinar la causa por la cual el usuario influyente es el de mayor autoridad.

Interpretación		
Razón de semejanza de las propiedades de los usuarios del grafo con respecto a las de Orlando Grabiél , la persona más influyente.		
Propiedad	Valor	Razón
Categoría	Trabajador	1.00
Rol en la Producción	Desarrollador	1.00
Area	Facultad 4	0.93
Sexo	Masculino	0.80
Cargo	Profesor	0.67

Figura 21: Tabla que muestra la razón de semejanza de los usuarios del grafo con respecto al usuario influyente.

Fuente: elaboración propia.

Para calcular la razón de semejanza de una propiedad (θ_p) se utiliza la siguiente fórmula:

$$\theta_p = \frac{M_p}{n-1} \quad (12)$$

- M_p es la cantidad de usuarios que tienen el mismo valor de la propiedad p que el usuario influyente.
- $n - 1$ es el total de usuarios del grafo restándole el influyente.
- Los valores que puede tomar p en la herramienta que se propone son: $\{municipio, provincia, categoría, cargo, área, rol en la producción, proyecto de investigación, grupo de investigación, línea de investigación, sexo, asignaturas\}$.

En la Tabla 10 se muestra el algoritmo a seguir para calcular la θ_p de una propiedad. Este algoritmo constituye la base para buscar las θ_p de todas las propiedades, solo habría que repetir la condición del paso 4.1 y los pasos 6 y 7 por cada propiedad.

Tabla 10: Algoritmo para calcular θ_p de una propiedad especificada por parámetro. Fuente: elaboración propia.

Entradas	Una lista de usuarios (<i>usuarios</i>) ordenada por la autoridad de los usuarios, el nombre de la propiedad (p).
Salida	Un JSON que muestra para la propiedad p , su valor y su θ_p .
INICIO 1. $M_p := 0$ 2. $influyente := usuarios.ObtenerPrimerElemento$ 3. Para cada $u_i \in usuarios$ 3.1. Si $influyente.valorPropiedad(p) = u_i.valorPropiedad(p)$ 3.1.1. $M_p ++$ 3.2. Fin Si 4. Fin para 5. $\theta_p := M_p / (usuarios.cantidad) - 1$ 6. $JSON = \{p: \{valor : influyente.valorPropiedad(p), razón : \theta_p\}\}$ 7. Retornar $JSON$ FIN	

Para dar solución a la segunda interrogante, la herramienta muestra una tabla con las propiedades comunes en todos los usuarios del grafo y la razón de semejanza global de cada propiedad (ver Figura 22). Esta razón muestra cuantas propiedades presentan valores que coinciden en todos los miembros del equipo y cuantas en la mayoría; permitiendo identificar las propiedades que determinan las relaciones mostradas en el grafo que representa al equipo de trabajo.

Interpretación	
Razón de semejanza entre todos los usuarios del grafo.	
Propiedad	Razón
Area	1.00
Cargo	1.00
Categoría	1.00
Categoría Docente	1.00
Sexo	1.00
Asignatura	1.00

Figura 22: Tabla que muestra la razón de semejanza entre todos los usuarios del grafo. Fuente: elaboración propia.

La fórmula para obtener la razón de semejanza global de una propiedad (γ_p) es la siguiente:

$$\gamma_p = \frac{\sum_{i=1}^n \delta_{pv_i}}{n} \quad (13)$$

- δ_{pv_i} es la razón de semejanza de la propiedad p para el usuario v_i y n es el total de usuarios.

Para calcular a δ_{pv_i} se emplea la siguiente fórmula:

$$\delta_{pv_i} = \frac{M_{pv_i}}{n-1} \quad (14)$$

- M_{pv_i} es la cantidad de usuarios que tienen el mismo valor de la propiedad p que el usuario v_i .

A continuación se muestran los algoritmos para calcular γ_p y δ_{pv_i} en las siguientes tablas:

Tabla 11: Algoritmo para calcular γ_p de una propiedad especificada por parámetro. Fuente: elaboración propia.

Entradas	Una Lista de usuarios (<i>usuarios</i>), el nombre de la propiedad (p)
Salida	Un JSON que muestra para la propiedad p y su γ_p .
INICIO	
1. $suma := 0$	
2. Para cada $v_i \in usuarios$	
3. Suma += δ_{pv_i}	
4. Fin para	
5. $\gamma_p = \delta_{pv_i} / usuarios.cantidad$	
6. $JSON = \{p : \gamma_p\}$	
7. Retornar $JSON$	
FIN	

Tabla 12: Algoritmo para calcular δ_{pv_i} de una propiedad especificada por parámetro. Fuente: elaboración propia.

Entradas	Una lista de usuarios (<i>usuarios</i>), el nombre de la propiedad (<i>p</i>), el usuario (<i>v_i</i>)
Salida	Un JSON que muestra para la propiedad <i>p</i> , su valor y su θ_p .
INICIO 1. $M_{pv_i} := 0$ 2. Para cada $u_i \in usuarios$ 2.1. Si $v_i.valorPropiedad(p) = u_i.valorPropiedad(p)$ 2.1.1. $M_{pv_i} ++$ 2.2. Fin Si 3. Fin para 4. $\delta_{pv_i} := M_{pv_i} / (usuarios.cantidad) - 1$ 5. Retornar δ_{pv_i} FIN	

Conclusiones del capítulo

1. En este capítulo se definió la métrica Autoridad Total, la cual permite identificar los usuarios más influyentes en la mensajería instantánea tomando en consideración su jerarquía (Autoridad Estructural) y su credibilidad (Autoridad de Reacción).
2. Se evidenció como la Autoridad Estructural utiliza el Grado, la Intermediación, la Cercanía y el PageRank para representar la jerarquía de los usuarios de la mensajería instantánea y como la Autoridad de Reacción tiene en cuenta la inmediatez con la que un mensaje enviado por un usuario es contestado para representar su credibilidad.
3. Se presentó la herramienta GRAPH-CHAT mostrándose los beneficios de su arquitectura y principales componentes. Se expuso la relación existente entre los componentes mediante Diagramas de Actividades.
4. Se evidenció cómo GRAPH-CHAT puede utilizar la información que está presente en los registros de los Servidores de mensajería instantánea y en el LDAP de las organizaciones para identificar en ellas grupos sociales, influencers y actores claves en la comunicación entre estos grupos. Se mostró como la herramienta puede utilizar esta información para apoyar el proceso de selección de equipos de trabajo.

CAPÍTULO 3. VALIDACIÓN DE LA INVESTIGACIÓN

En el presente capítulo se procede a validar los resultados de esta investigación a partir de los métodos científicos, teóricos y empíricos definidos en el marco teórico. Se valida la pertinencia, efectividad y utilidad de la Autoridad Total, mediante el método cualitativo Criterio de Expertos y un Estudio de Caso, donde es comparada con el valor de autoridad que calcula el algoritmo HITS (Kleinberg, 1999) y con un sociograma obtenido a partir de los criterios de los actores involucrados en el Estudio de Caso. Se valida además la satisfacción de los usuarios potenciales de herramienta la GRAPH-CHAT mediante la técnica ladov. Para concluir el proceso de validación, evaluar el sesgo que se produce al aplicar cada uno de los métodos y técnicas utilizados y constatar el cumplimiento de la hipótesis científica que guía la investigación, se aplica una triangulación metodológica inter-métodos. En el proceso de validación se empleó como escenario de estudio y aplicación a la Universidad de las Ciencias Informáticas

3.1 Aplicación del método Criterio de Expertos

La validación basada en el juicio de expertos permitió obtener valoraciones sobre temas relacionados con la propuesta de solución (Ramírez, 2016). Fue utilizada para validar los aspectos teóricos tomados en cuenta para el desarrollo de la métrica propuesta (Autoridad Total). Como método para el procesamiento estadístico de estos criterios o evaluaciones fue aplicada la escala psicométrica creada por Rensis Likert en 1932 a través de un cuestionario, con el objetivo de conocer el nivel de acuerdo o desacuerdo de los expertos con la métrica de autoridad desarrollada.

Se definieron 11 posibles expertos con un promedio de experiencia laboral mayor a trece años: dos licenciados en Sociología, un licenciado en Comunicación Social, un Máster en Historia con más de cinco años de experiencia en el estudio de redes sociales, dos licenciados en Psicología y cinco Doctores en Informáticas (Dr. C.). Esta selección se realizó con el objetivo de validar la métrica desde diferentes disciplinas que están asociadas a la base teórica y práctica de la investigación y representa una cantidad apropiada para garantizar la confiabilidad de los resultados. Hubo en todo momento disposición por parte de los seleccionados en colaborar con la investigación.

A estos posibles expertos se les aplicó una encuesta para determinar su coeficiente de competencia en el tema y así asegurar la confiabilidad de sus respuestas (ver Anexo 4). El procedimiento empleado para determinar el coeficiente de competencia de los candidatos puede ser consultado en el Anexo 5, así como los resultados arrojados luego de aplicada la encuesta en el Anexo 6.

Como resultado se obtiene que los 11 candidatos encuestados tienen un nivel de competencia Alto o Medio por lo que pueden ser utilizados como expertos. Los resultados de la distribución de su nivel de competencia se muestran en la Tabla 13.

Tabla 13: Distribución de expertos según su nivel de competencia. Fuente: elaboración propia.

Nivel de competencia	Cantidad	Porcentaje
Alto	9	81,8%
Medio	2	8,2%
Bajo	0	0%
Total	11	100%

Los indicadores seleccionados para ser evaluados por los expertos pueden observarse en la Sección II, del cuestionario del Anexo 3, definiéndose los siguientes aspectos a valorar:

1. Posibilidad de tener en cuenta las interacciones en la mensajería instantánea como fuente de datos para determinar el nivel de influencia de las personas en las organizaciones.
2. Utilización de un grafo para representar los usuarios y sus relaciones en la mensajería instantánea.
3. Criterio de que para un correcto análisis de autoridad e influencia en una red social hay que tener en cuenta tanto la estructura de las relaciones que se establecen en la red (jerarquía) como las reacciones que provoca un usuario sobre el otro (credibilidad).
4. Utilización de métricas y algoritmos de la teoría de grafos y la sociología para identificar la autoridad de los usuarios, tomando en consideración la estructura del grafo.
5. Identificación de las reacciones que provoca un usuario en la mensajería instantánea a partir de la cantidad de respuestas que sus mensajes reciben en un instante de tiempo inmediato.
6. Pertinencia de la métrica de autoridad, y sus sub-métricas para la determinación de los usuarios más influyentes de la mensajería instantánea
7. Utilidad de conocer las personas de mayor autoridad en la mensajería instantánea de las organizaciones.

Las preguntas del cuestionario diseñado (ver Anexo 3), están enfocadas a obtener las valoraciones en función de los indicadores definidos, representan los aspectos más relevantes de la investigación. El experto expresa su valoración de cada indicador mediante la siguiente escala: 5- muy de acuerdo (MA), 4- de acuerdo (DA), 3- ni de acuerdo ni en desacuerdo (Sí-No), 2- en desacuerdo (ED) y 1- completamente en desacuerdo (CD). Posteriormente, se procesan los resultados mediante la escala Likert. Con esta técnica son calculados los porcentajes de concordancia de los expertos con cada una de las respuestas de los planteamientos formulados (ver Anexo 5). Luego se calcula un índice porcentual de

concordancia (IP) que integra en un solo valor la aceptación de cada planteamiento por los evaluadores mediante la siguiente fórmula:

$$IP = \frac{5(\% \text{ de } MA) + 4(\% \text{ de } DA) + 3(\% \text{ de } Si-No) + 2(\% \text{ de } ED) + 1(\% \text{ de } CD)}{5} \quad (15)$$

La Figura 23 muestra que el IP relacionado con la valoración de los expertos, acerca de los aspectos planteados, en todos los casos supera al 89,1%.

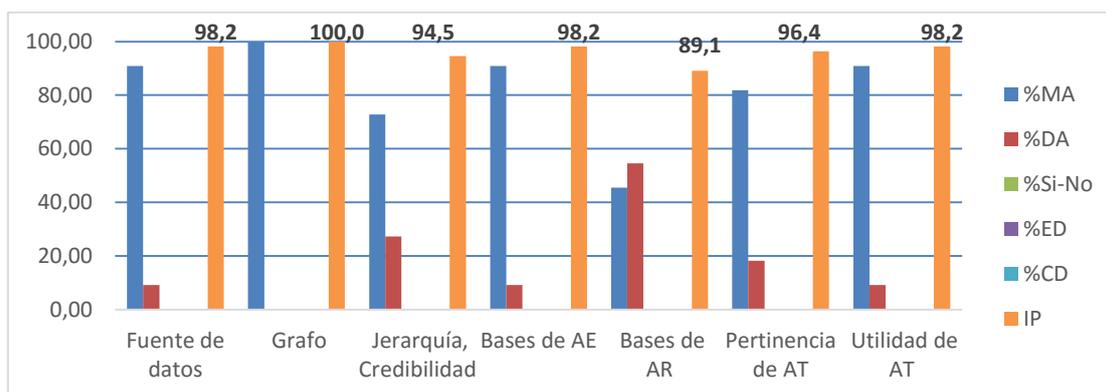


Figura 23: Valoración de los expertos sobre la Autoridad Total. Fuente: elaboración propia.

El análisis realizado a través del escalamiento de Likert evidencia que la pertinencia y utilidad de la métrica de autoridad desarrollada, tienen una alta valoración por parte de los expertos.

3.2 Aplicación de la técnica ladov para medir satisfacción

La técnica ladov constituye una vía para el estudio del grado de satisfacción de las personas. Se basa en la aplicación de un cuestionario con una estructura interna, que sigue una relación entre tres preguntas cerradas y un análisis posterior de otro conjunto de preguntas abiertas. La relación entre las preguntas cerradas se establece a través del denominado Cuadro Lógico de ladov (Ramírez, 2016), mostrado en la Tabla 14, el cual posibilita determinar el nivel de satisfacción del usuario y del grupo. La técnica fue empleada para medir la satisfacción de quince usuarios potenciales de la herramienta GRAPH-CHAT, escogiéndose personas que desempeñaban algún cargo de dirección en la Universidad de las Ciencias Informáticas (Decano, Vicedecanos, Jefes de departamentos, Jefes de proyecto).

Tabla 14: Cuadro lógico de ladov. Fuente: Elaboración propia.

¿Considera usted que las organizaciones deban continuar tomando decisiones administrativas sin tener en cuenta el liderazgo, autoridad o influencia que ejercen los trabajadores entre ellos en redes informales de interacción social como la mensajería instantánea?		
No	No sé	Sí
Si usted necesitara una herramienta para identificar los distintos grupos sociales, el liderazgo, autoridad o influencia que ejercen los trabajadores entre ellos en su organización, teniendo en cuenta las redes de mensajería instantánea, ¿utilizaría GRAPH-CHAT?		

¿Le satisface la herramienta informática desarrollada para identificar los distintos grupos sociales, el liderazgo, autoridad o influencia que ejercen los trabajadores entre ellos en redes de mensajería instantánea, como mecanismo para apoyar la toma de decisiones administrativas?	Sí	No sé	No	Sí	No sé	No	Sí	No sé	No
Me satisface mucho	1	2	6	2	2	6	6	6	6
No me satisface tanto	2	2	3	2	3	3	6	3	6
Me da lo mismo	3	3	3	3	3	3	3	3	3
Me insatisface más de lo que me satisface	6	3	6	3	4	4	3	4	4
No me satisface nada	6	6	6	6	4	4	6	4	5
No sé qué decir	2	3	6	3	3	3	6	3	4

Para obtener los resultados de la aplicación de la técnica es necesario conocer la escala de satisfacción, así como la fórmula para determinar el Índice de Satisfacción Grupal (ISG).

La escala de satisfacción responde a la siguiente estructura, en función de la puntuación obtenida luego de aplicado el cuestionario referido (Ramírez, 2016):

1. Clara satisfacción
2. Más satisfecho que insatisfecho
3. No definida
4. Más insatisfecho que satisfecho
5. Clara insatisfacción
6. Contradictoria

Luego de aplicado el cuestionario y haber triangulado las preguntas cerradas en el Cuadro Lógico de ladov, el número resultante de la interrelación de las tres preguntas cerradas indica la posición de cada una en dicha escala de satisfacción.

Para poder ponderar el ISG se establece una escala numérica entre +1 y -1, como se muestra en la siguiente tabla:

Tabla 15: Escala para ponderar el ISG. Fuente: Elaboración propia.

+1	Máximo de satisfacción
+0.5	Más satisfecho que insatisfecho
0	No definido o contradictorio
-0.5	Más insatisfecho que satisfecho
-1	Máxima insatisfacción

Luego es posible calcular el ISG a partir de la siguiente fórmula (Ramírez, 2016):

$$ISG = \frac{A(+1)+B(+0.5)+C(0)+D(-0.5)+E(-1)}{N} \quad (16)$$

- A, B, C, D, E, representan el número de sujetos con índice individual 1; 2; 3 ó 6; 4; 5
- N representa el número total de sujetos del grupo.

El ISG fluctúa entre 1 y -1. Es por ello que, una vez calculado, los valores que se encuentren comprendidos entre -1 y -0,5 indican insatisfacción; los comprendidos entre -0,49 y 0,49 evidencian contradicción y los que se ubiquen entre 0,5 y 1 indican que existe satisfacción.

En la presente investigación fue aplicada una encuesta integrada por seis preguntas a 19 usuarios potenciales, como se muestra en el Anexo 8. Para ello se tuvo en cuenta los años de experiencia laboral y el puesto de trabajo que ocupan. El valor obtenido al aplicar la técnica según fórmula 17 fue 0.97, el cual se encuentra en el intervalo de satisfacción, por lo que se puede concluir que existe un alto grado de satisfacción con la herramienta GRAPH-CHAT.

$$ISG = \frac{18(+1)+1(+0.5)+0(0)+0(-0.5)+0(-1)}{19} \quad (17)$$

En la Figura 24 se representan los porcentajes de satisfacción obtenidos luego de calculado el ISG, no encontrándose los tres últimos colores por estar en 0%.

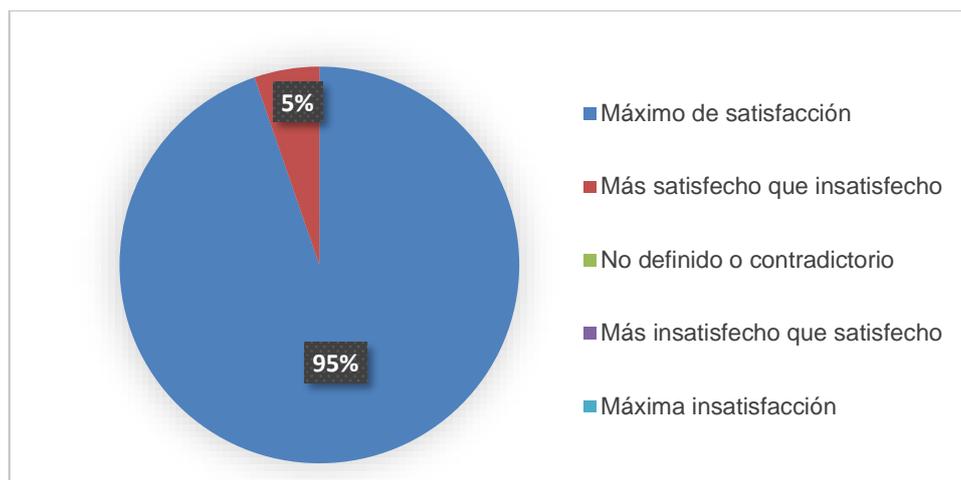


Figura 24: Satisfacción de potenciales usuarios con GRAPH-CHAT (Fuente: elaboración propia).

El cuestionario aplicado contó con tres preguntas complementarias de carácter abierto. Ello permitió profundizar en cuestiones positivas y negativas de la herramienta, así como en sugerencias para una mejor usabilidad.

Entre las principales cuestiones abordadas por los usuarios se enuncian:

- Factibilidad de GRAPH-CHAT para la identificación de usuarios influyentes en la mensajería Instantánea.
- GRAPH-CHAT es fácil de utilizar y posee una interfaz amigable.
- Utilidad de la información que brinda la herramienta en forma de grafo.

Los elementos negativos abordados fueron mínimos. Se encontraron algunas faltas de ortografías en la herramienta y errores de validación en los campos de entrada de datos. Todos fueron analizados y aceptados, siendo corregidos en la propuesta de solución final. Los encuestados aportaron además muy buenas sugerencias relacionadas con la posibilidad de exportar los equipos de trabajo formados en archivos con formato PDF o Excel. Estas se tendrán en cuenta para futuras versiones.

3.3 Estudio de Caso

El siguiente Estudio de Caso fue aplicado sobre 30 profesores de la Facultad 4 de la Universidad de las Ciencias Informáticas (UCI), distribuidos entre los departamentos docentes de Programación, Ingeniería de Software, Ciencias Básicas y Humanidades. Consiste en representar en un grafo las relaciones que se establecen entre estas personas a partir del intercambio de mensajes en la mensajería instantánea de la UCI. Luego calcular el valor de influencia de cada uno, teniendo en cuenta la métrica de autoridad desarrollada por esta investigación (Autoridad Total) y sus sub-métricas (Autoridad Estructural y Autoridad de Reacción), el valor de autoridad que brinda el algoritmo HITS y el valor de influencia determinado a partir de la aplicación de un sociograma a los profesores seleccionados. Finalmente se determinará la correlación existente entre estos valores mediante el Coeficiente de Pearson y el Coeficiente de Spearman, lo cual permitirá compararlos e identificar si existe una relación estadísticamente significativa entre ellos. Además, se mostrará la relación entre los valores mediante diagramas de dispersión.

Obtención de los datos y cálculo del valor de influencia

Para la representación del grafo (ver Figura 25) y el cálculo de la Autoridad Total se utilizó a GRAPH-CHAT. Los datos de entrada de la herramienta fueron obtenidos a partir de los historiales de conversación de los involucrados, con su previa autorización. El valor de autoridad de HITS se calculó con la ayuda de la herramienta Gephi⁸, configurándose para que se conectara a la base de datos de GRAPH-CHAT y de esta forma pudiera acceder al grafo. El sociograma aplicado a estas personas (ver Anexo 9) permitió identificar cuáles son más influyentes tomando como base sus propios criterios, fue diseñado por un experto en sociología y el procedimiento para su cómputo, se puede consultar en el Anexo 10. La tabla 16 muestra los valores de cada una de las métricas calculadas.

⁸ Gephi: herramienta de análisis y visualización de redes sociales (Gephi, 2018)

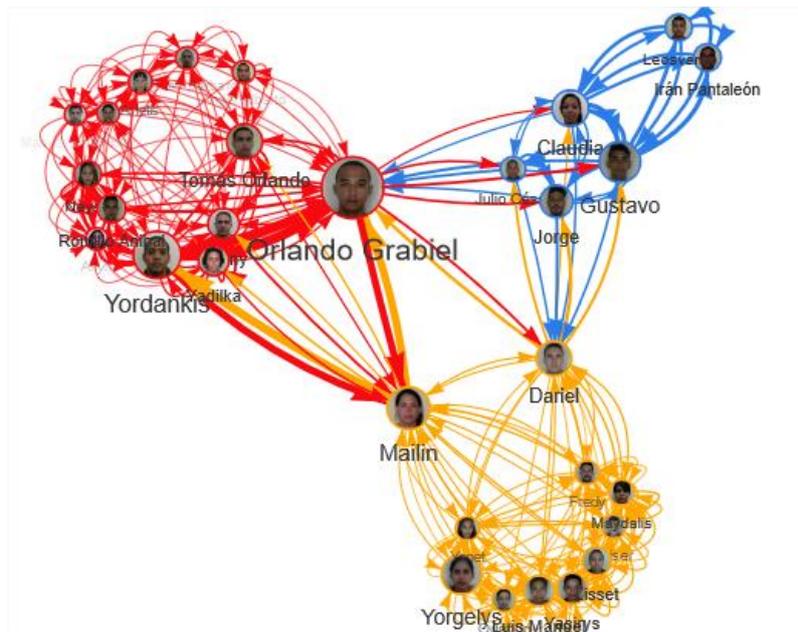


Figura 25: Grafo obtenido con GRAPH-CHAT (Fuente: elaboración propia).

Tabla 16: Valores calculados de Autoridad Total (AT), Autoridad Estructural (AE), Autoridad de Reacción (AR), Autoridad por medio del HITS e influencia por medio del sociograma para cada usuario del Estudio de Caso.

Fuente: elaboración propia.

Usuario	AE	AR	AT	HITS	Sociograma
1	30,83	0,00	15,42	34,29	22,34
2	30,45	0,00	15,22	31,26	20,21
3	36,90	0,00	18,45	36,45	26,60
4	42,76	47,63	45,19	21,58	52,13
5	63,09	21,70	42,39	75,13	48,94
6	37,36	0,00	18,68	76,37	26,60
7	42,76	67,09	54,92	21,58	62,77
8	30,48	43,15	36,82	20,75	41,49
9	37,36	0,00	18,68	76,37	26,60
10	30,48	11,46	20,97	20,75	24,47
11	30,98	0,00	15,49	33,66	22,34
12	37,36	20,68	29,02	76,37	37,23
13	37,36	28,64	33,00	76,37	41,49
14	22,69	26,91	24,80	4,61	28,72
15	28,62	0,00	14,31	30,53	19,15
16	37,15	5,73	21,44	77,05	29,79
17	70,82	47,30	59,06	100,00	63,83
18	22,69	29,97	26,33	4,61	29,79
19	35,54	15,08	25,31	37,40	32,98
20	37,01	0,00	18,51	77,10	26,60
21	100,00	100,00	100,00	69,92	100,00
22	32,85	23,99	28,42	31,79	35,11
23	50,98	35,75	43,37	56,74	53,19

24	38,80	24,67	31,74	39,31	36,17
25	42,77	95,02	68,89	47,71	77,66
26	45,73	22,44	34,08	50,78	41,49
27	34,80	71,60	53,20	70,24	62,77
28	29,10	0,00	14,55	28,88	20,21
29	37,36	0,00	18,68	76,37	26,60
30	37,36	29,66	33,51	76,37	42,55

Elementos teóricos necesarios para la aplicación del Estudio de Caso.

Antes de presentar los resultados, es importante conocer que es un sociograma, en que consiste la correlación de dos variables, los coeficientes de correlación que fueron aplicados y qué representa un diagrama de dispersión. A continuación, se explica en que consiste cada uno de ellos.

Un **Sociograma** es una técnica de la sociología que, a través de la observación y la evaluación de un contexto, refleja en un grafo los diferentes vínculos entre los miembros de un grupo, dejando en evidencia los grados de influencia y los lazos de preferencia que se presentan en él. Permite acceder a la estructura de la red social en un momento específico, tal como si se le tomara una fotografía (Urbina, López, & Cárdenas-Villalobos, 2018). A nivel gráfico, representa las relaciones interpersonales por medio de nodos (los individuos) que aparecen unidos por una o más aristas. Todos estos conocimientos se consiguen a través de la elaboración de una serie de preguntas de dos tipos que conforman un cuestionario muy completo y preparado. Por un lado, están las llamadas intelectuales que se refieren a la actividad productiva de la clase y por otro, las afectivas referidas a las elecciones que realizarían las personas para desarrollar actividades de tipo lúdico (Ander-Egg & Ezequiel, 1982).

La **Correlación**, aspecto muy estudiado por la estadística, existe entre dos variables cuando una de ellas se relaciona con la otra de alguna manera (Triola, 2004). Una forma muy común de representar la correlación entre dos variables es el **diagrama de dispersión** el cual es una gráfica en la que datos muestrales apareados (x, y) se grafican en un eje x horizontal y un eje y vertical. Cada par individual (x, y) se grafica como un solo punto (Triola, 2004).

El **Coefficiente de Pearson** o Coeficiente de Correlación Lineal (r) mide la fuerza de la relación lineal entre los valores cuantitativos apareados x y y en una muestra. Su valor se calcula con la siguiente fórmula (Triola, 2004):

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (18)$$

- n representa el número de pares de datos presentes

El **Coefficiente de Spearman** o Coeficiente de Correlación de Rangos (r_s) es una prueba no paramétrica que utiliza rangos de datos muestrales consistentes en datos apareados para probar una asociación entre dos variables (Triola, 2004). Debido a que se calcula a partir de rangos de datos y no de los propios valores, es menos sensible a valores aberrantes (outliers) que el Coeficiente de Pearson. Para su cálculo se tiene en cuenta la fórmula 19, si no existen empates en los rangos, o la fórmula 20, si existen empates (Triola, 2004).

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (19)$$

- n representa el número de rangos
- d representa la diferencia existente entre cada par de rango, restando el rango menor del mayor

$$r_s = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (20)$$

Estos coeficientes devuelven un valor entre -1 y 1. Si se acercan a 0, se concluye que no hay una correlación significativa entre x y y , pero si se acercan a -1 o a 1, se concluye que hay una fuerte correlación entre las dos variables que es directa (al aumentar x , aumenta y) si el valor es positivo o inversa (al aumentar x , disminuye y) si es negativo (Triola, 2004).

Comparación de los valores calculados y su discusión

Para comparar los valores calculados se realizó un Análisis de correlación con el software STATGRAPHICS, obteniendo dos matrices. La primera (Tabla 17) muestra los valores del Coeficiente de Pearson, entre cada par de variables. La segunda (Tabla 18), muestra los valores del Coeficiente de Spearman.

Tabla 17: Matriz que representa los valores de la Correlación de Pearson. Fuente: elaboración propia.

	AE	AR	AT	HITS	Sociograma
AE		0,5551	0,7966	0,5097	0,7797
AR	0,5551		0,9450	0,0380	0,9453
AT	0,7966	0,9450		0,2280	0,9936
HITS	0,5097	0,0380	0,2280		0,2767
Sociograma	0,7797	0,9453	0,9936	0,2767	

Tabla 18: Matriz que representa los valores de la Correlación de Spearman. Fuente: elaboración propia.

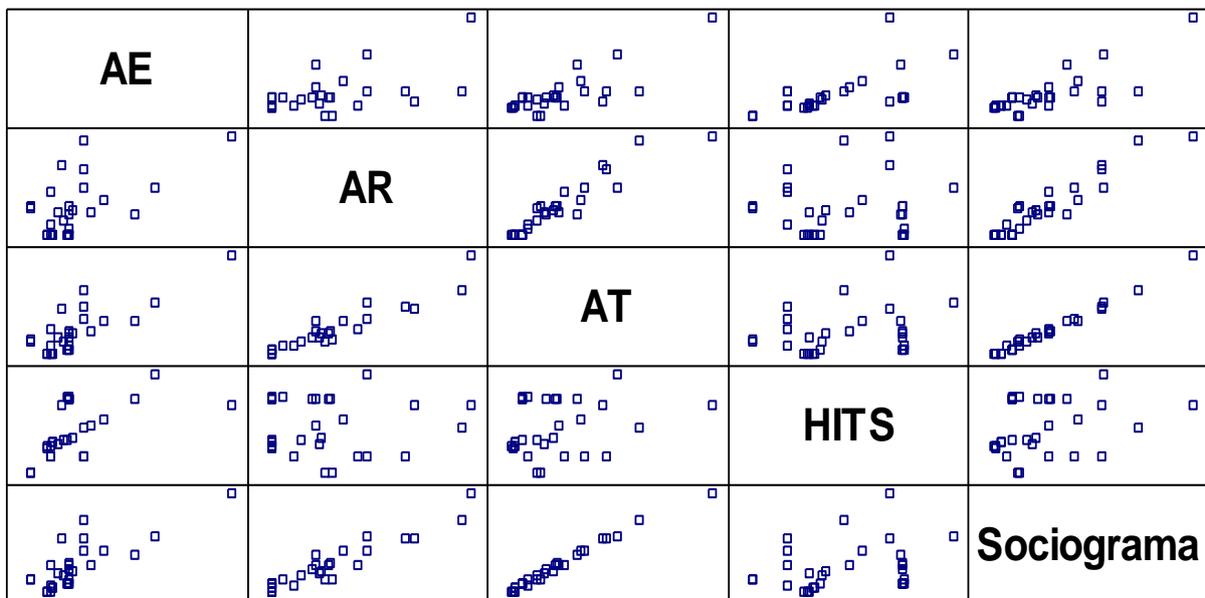
	AE	AR	AT	HITS	Sociograma
AE		0,4399	0,6931	0,5751	0,7284
AR	0,4399		0,9347	-0,0966	0,9084
AT	0,6931	0,9347		0,1477	0,9866
HITS	0,5751	-0,0966	0,1477		0,2423
Sociograma	0,7284	0,9084	0,9866	0,2423	

Cuando se analizaron los valores obtenidos de los coeficientes de correlación se observó que existe una relación relativamente débil entre la Autoridad Total (AT) y la puntuación de autoridad que otorga el HITS ($r = 0.2280$ y $r_s = 0.1477$). Sin embargo, la relación entre la

Autoridad Estructural (AE) y la autoridad del HITS es moderadamente fuerte ($r = 0.5097$ y $r_s = 0.5751$). Esto se debe a que ambas métricas tienen en cuenta solamente la estructura del grafo representativo de la red para calcularse. El HITS asigna su valor de autoridad teniendo en cuenta la cantidad y calidad de los enlaces que posee cada nodo, mientras que la AE consiste en la combinación de las principales métricas de centralidad (Grado, Cercanía, Intermediación y PageRank) por lo que también tiene en cuenta la calidad y cantidad de los enlaces, además de la posición estratégica del nodo. Por otra parte, el HITS no analiza las reacciones que provocan los usuarios de la mensajería instantánea y es por eso que su correlación con la Autoridad de Reacción (AR) es casi nula ($r = 0.0380$ y $r_s = -0.0966$).

Otro aspecto importante que resaltó en el análisis fue la fuerte correlación existente entre el valor de influencia obtenido con el sociograma y la AT ($r = 0.9936$ y $r_s = 0.9866$) así como la débil correlación entre el sociograma con la autoridad del HITS ($r = 0.2767$ y $r_s = 0.2423$). Esto evidencia la conclusión expuesta en el Capítulo 1 que planteaba: *para un correcto análisis de influencia hay que tener en cuenta tanto la estructura del grafo que representa la red social como las reacciones que provocaba un usuario sobre los demás cuando interactúan en ellas*. Además, demuestra la pertinencia, efectividad y utilidad de la AT para identificar usuarios influyentes en la mensajería instantánea y su superioridad en este contexto sobre el algoritmo HITS.

Figura 26: Matriz de grafos de dispersión Obtenidos con STATGRAPHICS para cada par de variables. Fuente: elaboración propia.



Los diagramas de dispersión, representados en la Figura 26, muestran de forma visual los análisis realizados. Se aprecia claramente la relación lineal existente entre la AT y el

sociograma pues los puntos que grafican la correlación entre estas dos variables se agrupan formando una recta. De igual forma los puntos que grafican la correlación entre la AT y el HITS y entre el HITS y el sociograma aparecen muy dispersos, lo cual reafirma la existencia de una baja correlación entre estos pares de variables.

3.4. Triangulación metodológica de los métodos científicos aplicados

La triangulación metodológica consiste en la combinación de métodos diversos en el estudio de un mismo fenómeno (Samaja, 2018). Integra métodos o técnicas con el objetivo de fortalecer los resultados y disminuir el sesgo que se produce al compararlos, aumentando su confiabilidad. Se distingue la triangulación intra-métodos y la triangulación inter-métodos. En la primera se utilizan variaciones de un mismo método, de recopilación y/o de interpretación de datos y está fundamentalmente dirigida a verificar la consistencia interna y la fiabilidad. En la segunda, se utilizan diversos métodos, cualitativos y cuantitativos, para aumentar la confianza en los resultados, mostrando que éstos no son dependientes del método (Ynoub, 2015).

A partir de la aplicación de los métodos científicos anteriormente expuestos, se realizó la triangulación metodológica inter-método e intra-métodos de los resultados. Esto permitió contrastar los resultados obtenidos, para poder determinar las coincidencias y divergencias. Constituye un criterio integrador sobre la validez de la propuesta de solución presentada. El resultado de su aplicación se muestra en la Tabla 18 (Triangulación inter-métodos) y en la Tabla 19 (Triangulación intra-métodos).

Tabla 18. Triangulación metodológica inter-métodos. Fuente: Elaboración propia.

Objetivo	Métodos cuantitativos	Métodos cualitativos	Conclusión
Evaluar la disminución del sesgo que se produce al aplicar cada una de las técnicas y métodos científicos utilizados, de manera independiente.	Estudio de Caso: se obtuvieron valores que demostraron la pertinencia, efectividad y utilidad de la Autoridad Total para determinar los niveles de influencia de los usuarios de la mensajería instantánea. Permitió validar la hipótesis de la investigación.	Criterio de expertos: alta valoración de los expertos con la métrica desarrollada. En todos los indicadores se obtuvo IP superior al 89.1%. ladov: alto grado de satisfacción de los usuarios con la información que brinda la herramienta GRAPH-CHAT (usuarios de mayor autoridad, grupos sociales de la mensajería instantánea y posibles equipos de trabajo). Se obtuvo un ISG de 0.97.	No existe contradicción en los resultados arrojados por los métodos y técnicas aplicados.

Tabla 19: Resultados de la triangulación metodológica intra-métodos. Fuente: elaboración propia.

Objetivo a evaluar	Triangulación intra-métodos	Variable comparada	Análisis realizado	Conclusión
--------------------	-----------------------------	--------------------	--------------------	------------

Evaluar la disminución del sesgo que se produce al aplicar cada una de las técnicas y métodos científicos utilizados, de manera independiente.	1.Likert (Expertos) 2.Estudio de Caso	Pertinencia	1. Pregunta 6. (IP= 96.4%) 2. Fuerte correlación existente entre los valores de la AT y el sociograma ($r = 0.9936$; $r_s = 0.9866$)	No existe contradicción en los resultados arrojados por los métodos y técnicas aplicados, teniendo en cuenta las variables: <ul style="list-style-type: none"> • Pertinencia • Efectividad • Utilidad
	1.ladov (Usuarios de GRAPCHAT) 2.Estudio de Caso	Efectividad	1. Pregunta 4. (95% de los usuarios con máxima satisfacción y el resto más satisfecho que insatisfecho) 2. Fuerte correlación existente entre los valores de la AT y el sociograma ($r = 0.9936$; $r_s = 0.9866$)	
	1.Likert (Expertos) 2.Estudio de Caso 3.ladov (Usuarios de GRAPCHAT)	Utilidad	1. Pregunta 6. (IP= 98.2%) 2. Fuerte correlación existente entre los valores de la AT y el sociograma ($r = 0.9936$; $r_s = 0.9866$) 3. Pregunta 6. (El 100% de los usuarios plantea que utilizaría GRAPH-CHAT para determinar usuarios influyentes)	

Luego de analizar cada tabla, se puede concluir que no existe contradicción en los resultados arrojados por los métodos aplicados, por lo que se constata el cumplimiento de la hipótesis de la investigación que refiere: Si se desarrolla y aplica una métrica de autoridad, utilizando técnicas de análisis de redes sociales, se determinará el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea.

Conclusiones del capítulo

1. En este capítulo fueron validados los principales aportes de esta investigación: la métrica Autoridad Total y la herramienta GRAPH-CHAT.
2. Se evidenció el alto grado de concordancia y aceptación de los expertos con los fundamentos teóricos utilizados para el desarrollo de la métrica de autoridad y sus componentes, arrojando criterios importantes para esta y futuras investigaciones.
3. Las opiniones de los potenciales usuarios de GRAPH-CHAT mostraron un grado alto de satisfacción con sus funcionalidades actuales y permitieron identificar aspectos a tener en cuenta para nuevas funcionalidades que se implementarán sus futuras versiones.

4. El Estudio de Caso aplicado permitió demostrar la efectividad, pertinencia y utilidad de la Autoridad Total para identificar usuarios influyentes en la mensajería instantánea y su superioridad, en este contexto, sobre el valor de autoridad que asigna el algoritmo HITS.
5. Se constató el cumplimiento del objetivo general y se validó la hipótesis de esta investigación.

CONCLUSIONES

Al término de esta investigación se arriban a las siguientes conclusiones:

1. Para un correcto análisis de influencia y autoridad, en un grafo representativo de una red de mensajería instantánea, hay que tener en cuenta tanto su análisis estructural (jerarquía), como las reacciones que provoca un usuario cuando le escribe a otro (credibilidad).
2. Conocer los niveles de influencia que poseen los trabajadores en la mensajería instantánea facilita identificar cuáles son más populares, más seguidos, los posibles líderes, los brókeres de información y los que son claves para difundir cualquier información. Conocimiento que puede apoyar la toma de decisiones organizacionales.
3. La métrica de autoridad desarrollada permitió identificar los niveles de influencia presentes en la mensajería instantánea mostrando los usuarios más influyentes y los influenciados.
4. Se evidenció que GRAPH-CHAT puede utilizar la información de la mensajería instantánea para determinar los usuarios de mayor autoridad, mostrarlos en un grafo y sugerir equipos de trabajo.
5. Los métodos de validación utilizados permitieron demostrar la pertinencia, utilidad y efectividad de la Autoridad Total para identificar las personas más influyentes de la mensajería instantánea, además de la satisfacción de los usuarios potenciales de la herramienta GRAPH-CHAT con sus funcionalidades.

RECOMENDACIONES

Para futuras investigaciones se recomienda:

1. Continuar profundizando en el estudio de influencia social en la mensajería instantánea, teniendo en cuenta modelos de propagación de influencia, información y predicción de enlaces.
2. Integrar la investigación con soluciones relacionadas con la Minería de Opinión que permitan realizar análisis de sentimiento y de polaridad en las opiniones generadas en la mensajería instantánea para identificar si la influencia que ejerce un usuario sobre otro es positiva o negativa.
3. Evaluar la posibilidad de utilizar elementos de la programación paralela y distribuida en los algoritmos implementados para calcular la Autoridad Total, con el objetivo de mejorar su rendimiento cuando trabajan sobre grandes volúmenes de datos.
4. Desarrollar nuevas métricas y algoritmos para determinar usuarios influyentes analizando los principios psicológicos de reciprocidad, coherencia, aprobación social, empatía y escasez, que al igual que la autoridad, caracterizan la influencia social.

REFERENCIAS BIBLIOGRÁFICAS

- Aguirre, J. L. (2011). Introducción al Análisis de Redes Sociales. *Introducción Al Análisis de Redes Sociales. Documentos de Trabajo Del Centro Interdisciplinario Para El Estudio de Políticas Públicas*, 82, 1–59.
- Alonso, D., Berdun, L., Monteserin, A., & Campo, M. (2016). Identificación de Nodos Influyentes en una Red Social de Microblogging. *ASAI, 17º Simposio Argentino de Inteligencia Artificial*. Retrieved from http://sedici.unlp.edu.ar/bitstream/handle/10915/56973/Documento_completo.pdf-PDFA.pdf?sequence=1
- Ander-Egg, & Ezequiel. (1982). Técnicas de investigación social. *Buenos Aires : Universidad Nacional Abierta. Dirección de Investigaciones y Postgrado*.
- Audiense. (2018). Audiense | Insights sobre consumidores y engagement únicos para ayudarte a crecer. Retrieved October 15, 2018, from <https://es.audiense.com/>
- Bello-Lara, R. (2018). Análisis estático de la red de colaboración científica en la Revista Cubana de Ciencias Informáticas. *Revista Cubana de Ciencias Informáticas*, 12(1), 220–232. Retrieved from <http://rcci.uci.cu>
- Bellver, J. A., & Guijarro, F. (2012). *Nuevos Métodos de Valoración de Modelos Multicriterio*. Retrieved from [https://riunet.upv.es/bitstream/handle/10251/19181/NUEVOS MÉTODOS DE VALORACIÓN - MODELOS MULTICRITERIO.pdf?sequence=1](https://riunet.upv.es/bitstream/handle/10251/19181/NUEVOS_MÉTODOS_DE_VALORACIÓN_-_MODELOS_MULTICRITERIO.pdf?sequence=1)
- Beveridge, A., & Shan, J. (2016). Network of Thrones. *Math Horizons*, 23(4), 18–22. <https://doi.org/http://dx.doi.org/10.4169/mathhorizons.23.4.18>
- Brandwatch. (2018). Brandwatch. Retrieved November 19, 2018, from <https://www.brandwatch.com/es/>
- BuzzSumo. (2018). BuzzSumo: Find the Most Shared Content and Key Influencers. Retrieved November 19, 2018, from <https://buzzsumo.com/>
- Carrión, P. (2016). Whatsapp, un asunto español. Retrieved March 7, 2018, from <https://es.kantar.com/tech/móvil/2016/junio-2016-datos-sobre-el-uso-de-la-mensajería-instantánea-en-europa/>
- Carter, D. R., Dechurch, L. A., Braun, M. T., & Contractor, N. S. (2015). Social Network Approaches to Leadership: An Integrative Conceptual Review. *Journal of Applied*

Psychology, 100(3), 597. <https://doi.org/10.1037/a0038922>

Cha, M., Hamed, H., Benevenuto, F., & Gummadi, K. P. (2010). Measuring User Influence in Twitter : The Million Follower Fallacy. *Association for the Advancement of Artificial Intelligence (Www.Aaii.Org)*.

Cialdini, R. B. (2009). *Influence. The Psychology of Persuasion*.

Corbellini, I. A. (2015). *Mecanismos de soporte para procesamiento distribuido de algoritmos de recomendación en redes sociales*. Universidad Nacional del Centro de la Provincia de Buenos Aires. Retrieved from <http://ridaa.unicen.edu.ar/xmlui/handle/123456789/583>

Estevez, E., & Antigua, D. (2015). *Plataforma Web para la gestión de información científico-técnica de la red de investigadores de la Universidad de las Ciencias Informáticas*. Universidad de las Ciencias Informáticas.

Fernández Rodríguez, A. W., Hidalgo Ruiz, J. J., & Rodríguez Puente, R. (2014). *Análisis de influencia social en el contexto universitario*. Universidad de las Ciencias Informáticas. Retrieved from <https://repositorio.uci.cu/jspui/handle/ident/9007>

Followerwonk. (2018). Followerwonk: Tools for Twitter Analytics, Bio Search and More. Retrieved October 16, 2018, from <https://followerwonk.com/>

Freeman, L. C. (1978). Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(1968), 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)

García, M. del F., Daly, A. J., & Sánchez-Cabezudo, S. S. (2016). Identifying the new Influencers in the Internet Era : Social Media and Social Network Analysis. *Revista Española de Investigaciones Sociológicas*, 153, 23–42. <https://doi.org/10.5477/cis/reis.153.23>

Gephi. (2018). Gephi - The Open Graph Viz Platform. Retrieved October 5, 2018, from <https://gephi.org/>

Golbeck, J. (2015). *Introduction to Social Media Investigation Introduction: A Hands-on Approach*. (J. L. Klavans & S. Elliot, Eds.). Elsevier Inc.

He, Q., Wang, X., Huang, M., Lv, J., & Ma, L. (2018). Heuristics-based Influence Maximization for Opinion Formation in Social Networks. *Applied Soft Computing*, 66, 330–369. <https://doi.org/10.1016/j.asoc.2018.02.016>

- Heepsy. (2018). Heepsy | Find influencers worldwide. Retrieved November 19, 2018, from <https://www.heepsy.com/>
- Heileman, G. L. (1998). *Estructuras de datos, algoritmos y programación orientada a objetos*. McGraw-Hill.
- Islas, O., & Ricaurte, P. (2013). *Investigar las redes sociales. Comunicación total en la sociedad de la ubicuidad*. México: Razón y Palabra.
- Johnson, A. R., van de Schoot, R., Delmar, F., & Crano, W. D. Social Influence Interpretation of Interpersonal Processes and Team Performance Over Time Using Bayesian Model Selection, 41 *Journal of Management* § (2015). <https://doi.org/10.1177/0149206314539351>
- Kempe, D., Kleinberg, J., & Tardos, É. (2015). Maximizing the Spread of Influence through a Social Network, 11(4), 105–147. <https://doi.org/10.4086/toc.2015.v011a004>
- Klear. (2018). Influencer Marketing Software | Klear. Retrieved October 16, 2018, from <https://es.klear.com/>
- Kleinberg, J. M. (1999). Authoritative Sources in a Hyperlinked Environment *. *Journal of the ACM*, 46(5), 604–632.
- Kred. (2018). Kred Rules. Retrieved October 15, 2018, from <http://home.kred/rules/>
- Kuz, A., Falco, M., & Giandini, R. (2016). Análisis de redes sociales : un caso práctico. *Computación y Sistemas*, 20(1), 89–106. <https://doi.org/10.13053/CyS-20-1-2321>
- Li, K., Zhang, L., & Huang, H. (2018). Social Influence Analysis: Models, Methods, and Evaluation. *Engineering*, 4(1), 40–46. <https://doi.org/10.1016/j.eng.2018.02.004>
- Lithium. (2018). Social Media Software for Community Management & Social Support | Lithium. Retrieved November 19, 2018, from <https://www.lithium.com/>
- Lyon, W. (2016). Analyzing the Graph of Thrones Import initial data into Neo4j. Retrieved March 25, 2018, from <http://www.lyonwj.com/2016/06/26/graph-of-thrones-neo4j-social-network-analysis/>
- Mancinas-Chávez, R., & Reig, R. (2016). Apuntes teórico-metodológicos para un estudio de la mensajería instantánea desde las ciencias sociales y las humanidades. *Fonseca, Journal of Communication*, 13, 87–101.

- Martín Fernández, J. M. (2017). *Mensajería instantánea y aprendizaje colaborativo : una experiencia en el Prácticum de Magisterio*. Universidad de Granada. Retrieved from <http://hdl.handle.net/10481/48347>
- McColl, R. C., Ediger, D., Poovey, J., Campbell, D., & Bader, D. A. (2014). A performance evaluation of open source graph databases. In *Proceedings of the first workshop on Parallel programming for analytics applications - PPAA '14* (pp. 11–18). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2567634.2567638>
- Monsalve, M. (2008). Análisis de redes sociales : un tutorial. *Bits de Ciencia*, 2, 1–6.
- Needham, M., & Hodler, A. E. (2018). *A Comprehensive Guide to Graph Algorithms in Neo4j*. Neo4j.
- Neo4j. (2018). Chapter 4. Algorithms - The Neo4j Graph Algorithms User Guide v3.4. Retrieved August 28, 2018, from <https://neo4j.com/docs/graph-algorithms/current/algorithms/>
- Otegui, J., & Guralnick, R. P. (2016). The geospatial data quality REST API for primary biodiversity data. *Bioinformatics*, 32(11), 1755–1757. <https://doi.org/10.1093/bioinformatics/btw057>
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking: Bringing Order to the Web. Retrieved from <http://ilpubs.stanford.edu:8090/422/>
- Pérez, Á. (2013). Álvaro Pérez García: Redes Sociales y Educación. *Revista Creativa y Sociedad*, 1–23.
- Pinilla, C., Bello, M., & Peña, C. (2017). Bases de datos orientadas a grafos. *TIA*, 5(2), 153–160.
- Raghavan, U. N., Albert, R., & Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E*, 76(3), 1–12. <https://doi.org/10.1103/PhysRevE.76.036106>
- Ramírez, J. F. (2016). *Modelo para la selección de equipos de trabajo quirúrgico en sistemas de información en salud aplicando técnicas de Inteligencia Organizacional*. Universidad de las Ciencias Informáticas.
- Rao, A., Spasojevic, N., Li, Z., & Dsouza, T. (2015). Klout Score: Measuring Influence Across Multiple Social Networks. *IEEE International Big Data Conference - Workshop*

- on Mining Big Data in Social Networks. <https://doi.org/10.1109/BigData.2015.7364017>
- Real Academia Española. (2018). influencia | Definición de influencia - Diccionario de la lengua española - Edición del Tricentenario. Retrieved November 21, 2018, from <http://dle.rae.es/srv/search?m=30&w=influencia>
- Ríos, S. A., Aguilera, F., Nuñez-Gonzalez, J. D., & Graña, M. (2017). Semantically enhanced network analysis for influencer identification in online social networks. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2017.01.123>
- Riquelme, F., Gonzalez-Cantergiani, P., Molinero, X., & Serna, M. (2017). Centrality measure in social networks based on linear threshold model. *Knowledge-Based Systems*, 140, 92–102. <https://doi.org/10.1016/j.knosys.2017.10.029>
- Robinson, I., Webber, J., & Eifrem, E. (2015). Graph Databases: new opportunities for connected data, 218.
- Rodríguez-Colomo, A. (2017). *El nacimiento de un nuevo Marketing: Influencers en España año 2017; cambios de hábito de consumo en España entre hombres y mujeres influenciados a través de las redes sociales*. Universidad de Valladolid.
- Romano Poveda, R., & Rubén. (2017). Servicio de directorio virtual del edificio de Ingeniería de Telecomunicación. Retrieved from <https://repositorio.unican.es/xmlui/handle/10902/10881>
- Saint-Charles, J., & Mongeau, P. (2018). Social influence and discourse similarity networks in workgroups. *Social Networks*, 52, 228–237 Social.
- Sakaji Kido, G., Augusto Igawa, R., & Barbon Jr, S. (2016). *Topic Modeling based on Louvain method in Online Social Networks. XII Brazilian Symposium on Information Systems-Information Systems in the Cloud Computing Era*. Retrieved from www.youtube.com
- Samaja, J. (2018). La triangulación metodológica (Pasos para una comprensión dialéctica de la combinación de métodos), 44(2), 431–443.
- Sánchez Viera, T., & Ruana Capote, R. (2014). *Algoritmo para la inferencia de redes sociales a partir de las actividades realizadas por los usuarios en diferentes contextos de interacción*. Universidad de las Ciencias Informáticas.
- Sandoval-Almazan, R., Gómez-Díaz, M. del R., & Demuner Flores, M. del R. (2012). *Redes*

Sociales en las Organizaciones. Universidad Autónoma del Estado de México.

Serrano-Puche, J. (2012). Herramientas web para la medición de la influencia digital : análisis de Klout y PeerIndex. <https://doi.org/http://dx.doi.org/10.3145/epi.2012.may.11>

SocialRank. (2018). SocialRank. Retrieved November 19, 2018, from <https://socialrank.com/>

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, 39, 156–168.
<https://doi.org/10.1016/J.IJINFOMGT.2017.12.002>

Sudhahar, S., Veltri, G. A., & Cristianini, N. (2015). Automated analysis of the US presidential elections using Big Data and network analysis, (June), 1–28.
<https://doi.org/10.1177/2053951715572916>

Sundaresan, V., Hsu, I., & Chang, D. (2014). *Subreddit Recommendations within Reddit Communities*. Retrieved from <http://snap.stanford.edu/class/cs224w-2014/projects2014/cs224w-16-final.pdf>

Topsy. (2018). Topsy. Retrieved October 16, 2018, from <http://topsy.thisisthebrigade.com/>

Torra, V., & Narukawa, Y. (2007). Modeling decisions: information fusion and aggregation operators. *Springer Science & Business Media*. Retrieved from <https://www.google.com/books?hl=es&lr=&id=vJa-65cThjYC&oi=fnd&pg=PA1&dq=TORRA,+Vicenç+y+NARUKAWA,+Yasuo.+2007.+Modeling+decisions.+Information+Fusion+and+Aggregation+Operators.+s.l.+Springer,+2007.&ots=TgSY5-62Tf&sig=-51OgfCeNbar7tjsAy1hyCPglgA>

Torres-Fernández, D., & Pérez-Sánchez, R. (2013). Predictores psicosociales de la satisfacción percibida con las relaciones interpersonales en el uso de la mensajería instantánea. *Global Media Journal México*, 10(20). Retrieved from https://journals.tdl.org/gmjei/index.php/GMJ_EI/article/view/179

Triola, M. F. (2004). *Estadística* (Novena Edición). Pearson educación.

Urbina, C., López, V., & Cárdenas-Villalobos, J. P. (2018). El uso de sociogramas en la escuela para la mejora de la convivencia. *Perfiles Educativos*, 40(160), 83–100.

Valenzuela, R. (2013). Las redes sociales y su aplicación en la educación. *Revista Digital Universitaria*, 14(4), 1–14.

Wilson, R. J. (1996). *Introduction to Graph Theory* (Fourth edi).

Ynoub, R. (2015). Revisión de algunos fundamentos lógico-metodológicos de la investigación cualitativa. *Perspectivas Metodológicas*, 2(16), 10–33. Retrieved from <http://revistas.unla.edu.ar/epistemologia/article/viewFile/768/822>

Zhao, H., Lee, D. L., & Chen, Z. (2018). Ranking Users in Social Networks with Higher-Order Structures Motif-based PageRank. *Association for the Advancement of Artificial Intelligence (Www.Aaii.Org)*.

ANEXOS

Anexo 1. Diagnóstico preliminar

Para realizar el diagnóstico preliminar se emplearon como escenario de estudio dos organizaciones muy diferentes: La Universidad de las Ciencias Informáticas (UCI) que está enfocada en la docencia y la producción y la Empresa Viajero Servicios de Reservación que tiene como principal misión la comercialización de pasajes por barco tren y ómnibus. En entrevistas realizadas a los trabajadores de estas organizaciones se pudo constatar que, en su mayoría, hacen un gran uso de la Mensajería Instantánea (MI); en la UCI para establecer conversaciones de carácter informal dentro de la universidad y en Viajero para comunicar los trabajadores de la Cede Central en La Habana con los de las diferentes UEB que poseen en el resto de las provincias, lo que evidencia que existe un gran uso de esta herramienta de comunicación.

En estas organizaciones se aplicó una encuesta (Ver Anexo 2) a directivos y trabajadores con el objetivo de que los criterios a evaluar tuvieran diversos puntos de vista.

La encuesta aplicada a Viajero se realizó en su oficina central, encuestándose una muestra de diez personas. En la UCI la muestra fue de cincuenta. Las siguientes tablas resumen las características del personal entrevistado.

Tabla 20: Características del personal entrevistado en Viajero. Fuente: elaboración propia.

Cantidad encuestados	de	Nivel escolar	Trabajador /Directivo	Experiencia
13		Universitarios: 11 Técnicos medios: 2	Directivos: 3 Trabajadores: 8	Promedio de años: 21

Tabla 21: Características del personal entrevistado en la UCI. Fuente: elaboración propia.

Cantidad encuestados	de	Grado Científico	Trabajador /Directivo	Experiencia
51		Máster: 6 Doctores: 5	Directivos: 10 Trabajadores: 41	Promedio de años: 9

Anexo 2: Encuesta del diagnóstico preliminar

Estimado colega: la presente encuesta forma parte de una investigación que está dirigida a la determinación del nivel de influencia de las personas en las organizaciones, teniendo en cuenta sus interacciones en redes de mensajería instantánea (chat).

La influencia social de un trabajador consiste en su capacidad de modificar el comportamiento de los demás. Por cuanto, sus valoraciones acerca de los asuntos que se someten a su consideración servirán de ayuda para la realización de un diagnóstico preliminar.

Datos generales del encuestado:

Nombre y Apellidos del encuestado: _____

Entidad/Área donde labora: _____

Cargo o Rol desempeñado: _____

Nivel Escolar: Universitario [] Técnico Medio []

Grado Científico: Doctor [] Máster []

Años de experiencia: ____

1. ¿Con que frecuencia usted utiliza la mensajería instantánea (chat)?

_____ Mucha _____ Poca _____ Ninguna

2. ¿Considera que un trabajador influyente puede ser un buen líder?

_____ Sí _____ No _____ Tal vez

3. ¿Conoce usted si en su organización, para tomar decisiones, se tienen en cuenta la influencia y el liderazgo que ejercen los trabajadores en los distintos grupos sociales?

_____ Sí _____ No

4. ¿Conoce alguna herramienta en su organización que permita determinar el liderazgo o influencia que ejerce un trabajador sobre los demás, en determinada área o especialidad?

_____ Sí _____ No

En caso de conocerla, por favor menciónela: _____

5. ¿Hoy es posible para usted conocer los principales grupos sociales que interactúan en su área de trabajo, o la fortaleza que tienen estos?

_____ Sí _____ No

En caso de conocerlo, por favor diga cómo: _____

6. ¿Dispone de alguna herramienta que, teniendo en cuenta el liderazgo o influencia que ejercen los trabajadores entre ellos, así como su información conocida (datos personales y/o laborales), le apoye la toma de decisiones para un mejor cumplimiento de los objetivos de su organización?

_____ Sí _____ No

En caso afirmativo, por favor menciónela: _____

Anexo 3. Cuestionario a expertos

Estimado experto (a): la presente encuesta forma parte de una investigación que está dirigida al desarrollo de una MÉTRICA DE AUTORIDAD PARA DETERMINAR EL NIVEL DE INFLUENCIA DE LAS PERSONAS EN LAS ORGANIZACIONES, TENIENDO EN CUENTA SU INTERACCIÓN EN REDES DE MENSAJERÍA INSTANTÁNEA. Por cuanto, sus valoraciones acerca de los asuntos que se someten a su consideración servirán gran ayuda de ayuda.

I- Datos generales del encuestado:

Nombre y apellidos: _____

Área donde labora: _____

Título universitario: _____

Grado científico: _____ Categoría docente: _____

Años de experiencia: _____

El objetivo de la presente encuesta consiste en que usted evalúe cada uno de los indicadores que se le presentarán en la tabla de la subsiguiente sección II, colocando el número en la casilla correspondiente y teniendo en cuenta para ello las siguientes categorías:

5: MUY DE ACUERDO (MA); 4: DE ACUERDO (DA); 3: NI DE ACUERDO NI EN DESACUERDO (Si- No); 2: EN DESACUERDO (ED); 1: COMPLETAMENTE EN DESACUERDO (CD).

II- Lista de indicadores a valorar:

No	Indicador	Valoración
1	¿Cómo evalúa usted la posibilidad de tener en cuenta las interacciones en la mensajería instantánea como fuente de datos para determinar el nivel de influencia de las personas en las organizaciones?	
2	¿Considera correcto modelar una red social como un grafo y a partir de este identificar los niveles de influencia de los usuarios?	
3	¿Está de acuerdo con la idea de que para un correcto análisis de autoridad e influencia en un red social hay que tener en cuenta tanto la estructura de las relaciones que se establecen en la red (jerarquía) como las reacciones que provoca un usuario sobre el otro (credibilidad)?	
4	¿Considera usted que mediante la utilización de métricas y algoritmos de la teoría de grafos y la sociología se puede identificar la autoridad de los usuarios, tomando en consideración la estructura del grafo?	
5	¿Concuerda con el criterio de se pueden identificar las reacciones que provoca un usuario en la Mensajería Instantánea a partir de la cantidad de respuestas que sus mensajes reciben en un instante de tiempo inmediato?	
6	¿Cómo evalúa usted la pertinencia de la métrica de autoridad, y sus sub-métricas en la determinación de los usuarios más influyentes en la mensajería instantánea?	
7	¿Considera útil conocer las personas de mayor autoridad en la Mensajería Instantánea de su organización?	

III- Si desea exponer cualquier otra opinión, por favor, expréselo en el espacio disponible a continuación.

¡MUCHAS GRACIAS POR SU COLABORACIÓN!

Anexo 4. Encuesta para determinar el coeficiente de competencias de los expertos

Compañero (a): _____

Usted ha sido seleccionado como posible experto para ser consultado por esta investigación respecto a temas relacionados con la identificación de usuarios influyentes en redes de interacción social. Agradecemos sinceramente su valiosa cooperación. Gracias.

1. Marque con una cruz (X) en la tabla siguiente el valor que se corresponde con el grado de conocimiento que usted posee sobre el Análisis de Redes Sociales, (Escala ascendente).

0	1	2	3	4	5	6	7	8	9	10

2. Realice una autoevaluación del grado de influencia que cada una de las fuentes que le presentamos a continuación ha tenido en su conocimiento y criterio sobre el Análisis de Redes Sociales. Marque con una cruz (X) según corresponda en A (alto), M (medio) o B (bajo).

No	Fuente de argumento	Grado de influencia de cada una de las fuentes.		
		A (alto)	M (medio)	B (bajo)
1	Análisis teóricos realizados			
2	Experiencia obtenida			
3	Autores nacionales			
4	Autores extranjeros			
5	Conocimiento del estado del problema en el extranjero			
6	Intuición			

Anexo 5. Procedimiento empleado para determinar el coeficiente de competencia de los candidatos a expertos

Cálculo del coeficiente de competencia de los expertos que evaluaron el componente desarrollado.

El cálculo de dicho coeficiente se realiza de la forma siguiente:

$$K_{comp} = \frac{1}{2} (K_c + K_a) \quad (14)$$

- **K_{comp}**: coeficiente de competencia.
- **K_c**: coeficiente de conocimiento o información que tiene el experto acerca del problema, calculado sobre la valoración del propio experto en una escala de 0 a 10 y multiplicado por 0,1.
- **K_a**: coeficiente de argumentación o fundamentación de los criterios del experto, obtenido como resultado de la suma de los puntos de acuerdo a la siguiente tabla patrón:

Tabla 22. Patrón a seguir para evaluar el Ka (Ramírez, 2016).

No	Fuentes de argumentación	Alto (A)	Medio (M)	Bajo (B)
1	Análisis teóricos realizados	0,30	0,20	0,10
2	Experiencia adquirida durante su vida profesional.	0,50	0,35	0,30
3	Conocimiento de investigaciones y/o publicaciones nacionales.	0,05	0,04	0,03
4	Conocimiento de investigaciones y/o publicaciones internacionales.	0,05	0,04	0,03
5	Conocimiento del estado del problema en el extranjero	0,05	0,04	0,03
6	Intuición.	0,05	0,03	0,02
	Total	1,00	0,70	0,50

Se plantea entonces que:

- La Competencia del experto es de Alta (A): si $K_{comp} > 0,7$
- La Competencia del experto es Media (M): si $0,5 < K_{comp} = < 0,7$
- La Competencia del experto es Baja (B): si $K_{comp} = < 0,5$

Anexo 6. Resultados de la encuesta aplicada a los candidatos a expertos para determinar su nivel de competencia

Expertos	Kc	1	2	3	4	5	6	Ka	Kcomp	Valor
1	0,9	0,2	0,5	0,04	0,03	0,03	0,05	0,85	0,875	Alto
2	0,9	0,3	0,5	0,04	0,04	0,04	0,05	0,97	0,935	Alto
5	0,8	0,2	0,5	0,04	0,04	0,03	0,05	0,86	0,83	Alto
4	0,9	0,2	0,5	0,05	0,03	0,03	0,05	0,86	0,88	Alto
5	0,8	0,1	0,35	0,05	0,05	0,04	0,05	0,64	0,72	Alto
6	0,6	0,2	0,35	0,03	0,05	0,04	0,03	0,7	0,65	Medio
7	0,5	0,1	0,35	0,03	0,03	0,03	0,02	0,56	0,53	Medio
8	0,8	0,3	0,5	0,04	0,05	0,05	0,03	0,97	0,885	Alto
9	0,9	0,3	0,5	0,04	0,05	0,05	0,05	0,99	0,945	Alto
10	0,7	0,3	0,35	0,04	0,04	0,05	0,05	0,83	0,765	Alto
11	0,7	0,3	0,35	0,05	0,05	0,05	0,05	0,85	0,775	Alto

Anexo 7. Respuestas dadas por cada uno de los expertos para cada indicador

Expertos	Indicador						
	1	2	3	4	5	6	7
1	5	5	5	5	4	5	5
2	5	5	5	5	5	5	5
3	5	5	4	5	5	5	5
4	5	5	5	5	4	5	5
5	5	5	4	5	5	5	5
6	4	5	5	5	4	4	5
7	5	5	5	5	4	4	5
8	5	5	5	5	4	5	5
9	5	5	5	5	4	5	5
10	5	5	5	4	5	5	4
11	5	5	4	5	5	5	5

Anexo 8: Encuesta aplicada para medir satisfacción del usuario con la herramienta GRAPH-CHAT

Nombre y apellidos: _____

Fecha de realización: _____

Especialidad: _____

Entidad: _____

1. ¿Considera usted que las organizaciones deban continuar tomando decisiones sin el apoyo de una herramienta informática que identifique el liderazgo, autoridad o influencia que ejercen los trabajadores entre ellos en redes informales de interacción social?

Sí _____ No _____

2. ¿Qué elementos consideras positivos de la herramienta presentada?

3. ¿Qué elementos consideras negativo de la herramienta presentada?

4. ¿Le satisface la herramienta informática desarrollada para identificar los distintos grupos sociales, el liderazgo, autoridad o influencia que ejercen los trabajadores entre ellos en redes de mensajería instantánea, como mecanismo para apoyar la toma de decisiones administrativas?

_____ Me satisface mucho

_____ No me satisface tanto

_____ Me da lo mismo

_____ Me insatisface más de lo que me satisface

_____ No me satisface nada

_____ No sé qué decir

5. ¿Qué sugerirías para mejorar la herramienta?

6. Si usted necesitara una herramienta para identificar los distintos grupos sociales, el liderazgo, autoridad o influencia que ejercen los trabajadores entre ellos en su organización, teniendo en cuenta las redes de mensajería instantánea, ¿utilizaría GRAPH-CHAT?

Sí _____ No _____

Anexo 9: Sociograma

Compañero (a): _____

La presente encuesta tiene como objetivo validar la métrica de autoridad propuesta por la investigación “Métrica de autoridad para determinar el nivel de influencia de las personas en las organizaciones, teniendo en cuenta su interacción en redes de mensajería instantánea”. Es por ello que pedimos su colaboración y la sinceridad en las respuestas. Garantizamos su anonimato y la confidencialidad de sus respuestas.

Teniendo en cuenta el listado de personas que se adjunta con la encuesta, responda las siguientes preguntas, escribiendo los nombres de las personas separados por comas:

1. ¿Con quiénes conversas a menudo por el chat?

_____.

2. ¿Con quiénes te gustaría realizar algún trabajo, artículo o colaborar en el desarrollo de alguna aplicación informática?

_____.

Anexo 10: Procedimiento para computar el sociograma

Para el cómputo del sociograma lo primero que se realizó fue ponderar las preguntas aplicadas en la encuesta. La primera pregunta se ponderó con una unidad ($P_1 = 1$) y la segunda con dos unidades ($P_2 = 2$). Luego se confeccionó una matriz (ver Tabla 23) para representar las respuestas de cada usuario.

Tabla 23. Matriz para representar las respuestas del sociograma. Fuente: elaboración propia.

	1	2	3	(...)	n
1				...	
2				...	
3				...	
(...)
n				...	

En la misma, las filas representan a las personas encuestadas y las columnas las del listado adjunto a la encuesta del sociograma, las cuales coinciden con las encuestadas. El valor de cada celda (X_{ij}) de la matriz corresponde con la selección de alguna persona en la primera o segunda pregunta y se calcula mediante la siguiente fórmula:

$$X_{ij} = P_1 + P_2 \quad (15)$$

El valor de la influencia social de una persona (S_i) se calcula sumando todas las celdas de su columna como se muestra en la próxima fórmula:

$$S_i = \sum_j^n X_{ij} \quad (16)$$