

# Identificación de influyentes en twitter a través del análisis textual y la elaboración de grafos

## ***Dra. Rocío Abascal Mena***

Universidad Autónoma Metropolitana - Cuajimalpa, Av. Vasco de Quiroga 4871, Cuajimalpa Santa Fe,  
C.P. 05348, Ciudad de México, Distrito Federal, México, Teléfono: 5814-6502  
*mabascal@correo.cua.uam.mx*

## ***Dr. Erick López Ornelas***

Universidad Autónoma Metropolitana - Cuajimalpa, Av. Vasco de Quiroga 4871, Cuajimalpa Santa Fe,  
C.P. 05348, Ciudad de México, Distrito Federal, México, Teléfono: 5814-6502  
*elopez@correo.cua.uam.mx*

## ***Dr. Sergio Zepeda Hernández***

Universidad Autónoma Metropolitana - Cuajimalpa, Av. Vasco de Quiroga 4871, Cuajimalpa Santa Fe,  
C.P. 05348, Ciudad de México, Distrito Federal, México, Teléfono: 5814-6502  
*jzepeda@correo.cua.uam.mx*

## **Resumen**

Las redes sociales digitales se han vuelto imprescindibles en la participación social y política creando espacios que se contraponen a los ya tradicionales medios de comunicación. En este artículo se presentan los resultados de la investigación realizada a partir de la extracción de un gran volumen de datos del caso UBER en México. La investigación sobre el fenómeno *#UberSeQueda* aborda distintos aspectos en el que la transparencia y la democratización de la información son el resultado del uso de Twitter. En este artículo se aplica un proceso computacional para extraer actores influyentes a partir de los mensajes enviados. La metodología utilizada se base en estudios que analizan las redes complejas explorando la interdisciplinariedad entre la Computación y las Humanidades para detectar actores influyentes. Se concluye, a partir del uso de

grafos, que la preponderancia de los actores involucrados en un movimiento social basado en las redes sociales cambia constantemente en cortos lapsos de tiempo y que es necesaria la combinación de métodos que involucren tanto el perfil del usuario como el contexto actual.

**Palabra(s) Clave(s):** análisis de redes sociales, grafos, actores influyentes, interdisciplina.

## 1. Introducción

En la actualidad es posible tener a la mano nuevas formas de comunicación mediadas por la computadora como lo son las llamadas redes sociales digitales. Estas son estructuras sociales formadas por nodos -personas, grupos u organizaciones-, que están vinculados respecto a intereses y/o puntos de vista en común. Una de las redes con mayor participación a nivel mundial, y en México, es Twitter cuyo éxito se debe primordialmente a su facilidad de uso. Básicamente los usuarios envían mensajes de máximo 140 caracteres, llamados *tuits* (desde el 2012 la Real Academia Española incluyó los términos tuitear, tuit y tuitero en su diccionario).

En Twitter, existen tres tipos de relaciones posibles entre un usuario A y un B: 1) A sigue a B, significa que los tuits posteados por B aparecen en la línea de tiempo (*timeline*) del usuario A; 2) B sigue a A y 3) ambos se siguen. Desde luego, también pueden no seguirse. Los usuarios también tienen la posibilidad de retuitear (RT), contestar un tuit (con un *reply*) o añadir una marca de favorito a un tuit de manera a difundir a sus seguidores lo que piensan y que es particularmente digno de leerse (lo que conduce a una difusión capital de los tuits).

La idea de compartir y transformarse en protagonistas principales, productores y generadores de información y contenidos, de dejar una postura pasiva, de consumo exclusivo; para transformarse en elementos activos y participantes clave, es una tendencia que no admite discusión y que se vislumbra irreversible. En este contexto,

encontramos el uso de Twitter como un medio para la discusión política en contextos locales y nacionales (como las elecciones del 2010 en Australia, 2011 en México o 2012 en Estados Unidos, por nombrar algunas) hasta llegar a protestas y movimientos activistas como las ocurridas en Egipto, Túnez y Yemen. Twitter se ha convertido en un canal legítimo de comunicación en la arena política como un resultado de la campaña presidencial del 2008 realizada en Estados Unidos [1].

El inmenso flujo de información que la tecnología provee a través de dispositivos digitales y las redes sociales (Twitter, Facebook) permite que los usuarios conozcan la actualidad de manera inmediata y, a su vez, sean partícipes de los sucesos. [2] menciona que los movimientos sociales, a lo largo de la historia, son los productores de nuevos valores y objetivos en torno al cual las instituciones de la sociedad se transforman para crear nuevas normas para organizar la vida social. En el caso de las redes sociales digitales los usuarios son variados y sus objetivos también lo son. [3] afirma que la clave de cada cambio de conducta, moda, ideología o tendencia está en un núcleo de personas pioneras que adoptan, comunican y contagian ideas, conductas o modas. En definitiva, en líderes de opinión. Ahora bien, la influencia de estos líderes no deriva de las apariciones en los medios, sino de lo que los sociométricos llaman "*la centralidad de red*":

ellos son nodos en la red del boca a oreja, que conectan a todo el mundo con un máximo de seis grados de separación (en referencia a la teoría de los seis grados de Stanley Milgram).

La teoría de [3] puede discutirse porque hay factores que explicarían por qué una persona que no está reconocida como líder de opinión también puede expandir mensajes boca a oreja con frecuencia. Uno de ellos, de carácter endógeno, sería la pertenencia a una red social con más o menos vínculos débiles [4].

El análisis de redes sociales ARS (Social Network Analysis) es un conjunto de técnicas que permiten el estudio formal de las relaciones entre actores y las estructuras sociales

que surgen de la recurrencia de esas relaciones o de la ocurrencia de determinados eventos. Algunos trabajos en el área de la Computación han intentado determinar por qué un mensaje se vuelve viral. Tales son los trabajos de [5 y 6] en los que se cuantifica el efecto causal de las redes sociales a partir de la identificación de quién influye en quién además de otras señales sociales que influyen en la propagación masiva de la información a gran escala.

Twitter ha sido analizado para detectar patrones de influencia en mensajes sociales pero también en el *marketing*. [7] se basa en el estudio de las ligas directas, que representan desde una amistad hasta intereses en común entre los usuarios, para determinar la influencia de otros a partir del grado de entrada (*indegree*), los retuits y las menciones. Por su parte, [8] se basa en el uso de los seguidores (*followers*) para determinar las tendencias en influencia. A su vez [9] explora quién dice qué en la red Twitter a partir del estudio de las listas que conforman grupos selectos de tuiteros. Patrones, diferencias en los mecanismos de difusión entre diferentes temas en los que se ven involucrados los idiomas y los hashtags políticos así como las probabilidades de influencia son estudiados por [10, 11 y 12]. Sin embargo, ninguno de los trabajos anteriores analiza la topología de la red, quedándose solamente en un estudio sobre los hashtags, las menciones o el número de seguidores.

Con el creciente uso de Twitter, se vuelve fundamental contar con herramientas que permitan extraer información para analizar cómo se difunde la información y quiénes son los usuarios influyentes a partir del uso de grafos en el que se considere la topología de la red y no únicamente, por ejemplo, el número de tuits o seguidores. Por lo tanto, la hipótesis del presente trabajo se resume en:

*H<sub>0</sub>*. La aplicación de una metodología para extraer mensajes provenientes de Twitter considerando la topología de la red permite obtener a los actores influyentes diluyendo a los *bots* a partir de la concentración en grupos con características similares.

La validación de la hipótesis se realizó a partir de la extracción de tuits y la visualización de grafos, a lo largo del tiempo, aplicando medidas de centralidad, grado, comunidad y excentricidad. El caso de estudio es aplicado en un corpus recuperado con *#UberSeQueda*. El objetivo del artículo se centra en la identificación de actores influyentes a partir de la aplicación de medidas tales como el grado, la excentricidad, la centralidad de intermediación y de proximidad así como la conformación de comunidades pertenecientes a la red. Para ello se utilizaron los mensajes enviados con el hashtag *#UberSeQueda* en las redes sociales, tomando en cuenta el tiempo. Este ejemplo de aplicación, enfocado en tuits políticos y sociales, intenta contestar a la siguiente pregunta:

*P<sub>1</sub>*. ¿Es posible detectar actores influyentes en movimientos sociales, en Twitter, considerando la topología de la red durante un periodo de tiempo?

Este artículo presenta una alternativa en el área de la lingüística computacional al utilizar la topología de red para extraer actores influyentes.

El artículo está organizado de la siguiente manera: en la Sección 2 se presenta el contexto del caso de estudio *#UberSeQueda*. En la sección 3 se describe la metodología que explica paso a paso el proceso para la extracción y análisis de los actores involucrados en los tuits. En la Sección 4 se presenta un caso de estudio utilizando el hashtag *#UberSeQueda* y sus resultados. En la Sección 5 se hace una discusión de los resultados y finalmente en la Sección 6 se abordan las conclusiones y el trabajo a futuro.

## **2. Contexto**

Las Tecnologías de la Información y la Comunicación (TICs) han ido permitiendo, en los últimos años, una transparencia y democratización de la información. Los retos que se enfrentan de manera diaria se apoyan en el uso de la tecnología. Por ejemplo, existen nuevas opciones de movilidad que resultan de innovaciones tecnológicas que ponen en

contacto a personas con conductores de vehículos privados. Ese es el caso de Uber y Cabify donde la contratación se hace mediante una aplicación de teléfono inteligente. Se pacta origen y destino, la tarifa depende de la disponibilidad del tipo de vehículo que se escoge, un cobro base, tiempo y distancia. La empresa Uber que hace la intermediación cobra un porcentaje del costo del viaje, y es la que liquida la diferencia al propietario del vehículo que lo renta y al conductor. En toda esta cadena todos pagan impuestos, además de que cada uno puede ser totalmente fiscalizado por hacerse los pagos mediante bancos. Uber exige que el propietario del vehículo cuente con seguro de cobertura amplia; que los vehículos sean de modelos recientes en buenas condiciones; que los conductores pasen por exámenes de conocimiento de la ciudad de México, que no tengan antecedentes penales y estén libres de drogas, entre muchas otras características que los diferencian de los taxis libres, de los de sitio y los radio taxis.

Así, las manifestaciones de taxistas que se dieron en el mes de mayo de 2015, en México, es un ejemplo muy clásico de cómo el avance tecnológico tarda en cambiar las estructuras sociales que son muy complejas. La posibilidad de solicitar un taxi a partir de Internet y obtener un servicio, aparentemente, de mejor calidad, más seguro y mucho más confiable se contrapone al ya tradicional sistema de taxis en México. Este es corporativista, en el cual existen gremios y en el que, evidentemente, quienes se benefician son solo los líderes. Las manifestaciones de los taxis en distintos puntos del Distrito Federal originaron inconformidades por parte de los usuarios, los cuales se hicieron presentes en las redes sociales digitales mediante mensajes etiquetados con *#MexicoNoPara*, *#UberNoPara* y *#UberSeQueda*. Además de ser tendencia mundial *#UberSeQueda*, UBER registró un aumento del 800 por ciento en las descargas de esta aplicación en teléfonos inteligentes durante las manifestaciones.

Los mensajes extraídos corresponden a los etiquetados con *#UberSeQueda* recopilados entre el 25 y 27 de mayo de 2015. El marco tiempo fue una opción

metodológica que permite analizar el rápido aumento en el volumen de mensajes enviados de un día a otro.

### 3. Metodología

Para realizar el análisis de grandes cantidades de información se seleccionó la red social digital Twitter. Esta permite el envío de mensajes que se caracterizan por su limitación de 140 caracteres y su etiquetado con *hashtags*. Indudablemente, Twitter debido a su capacidad por llegar a un gran público y su anonimato ha permitido que, en México, su crecimiento sea exponencial y que sea un medio de información alternativo a los ya tradicionales.

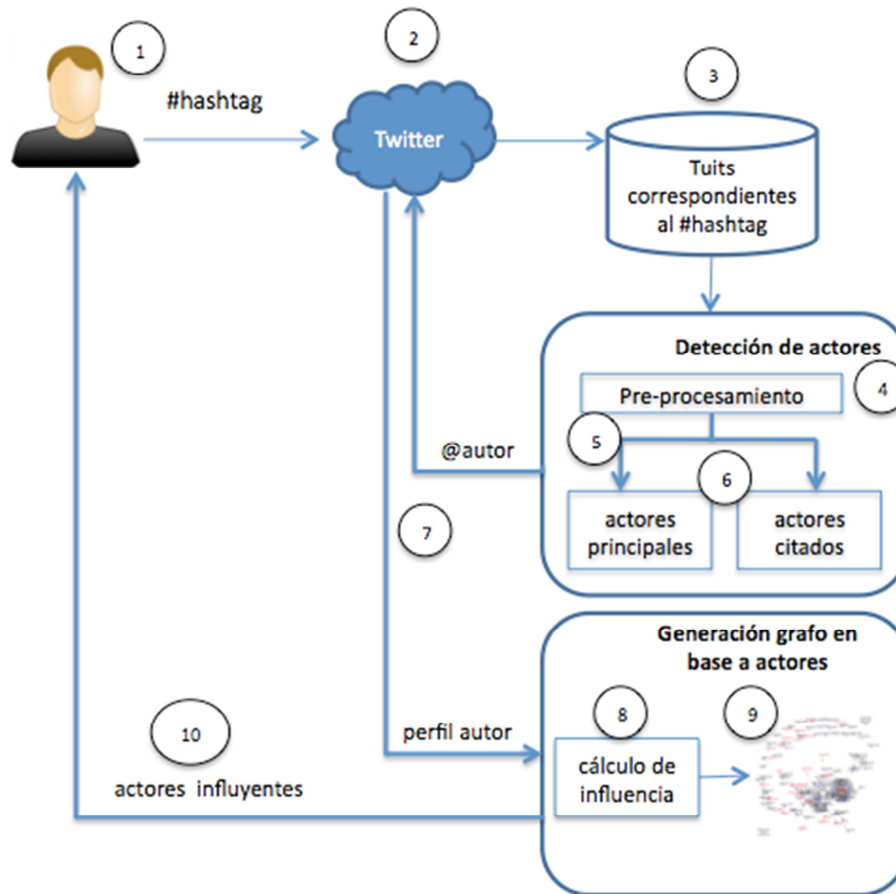


Fig. 1. Proceso de extracción y análisis de actores provenientes de Twitter.

Twitter provee una API (Application Programming Interface) abierta la cual permite el desarrollo de aplicaciones que pueden tomar ventaja de la información disponible en red. De este modo, se hace posible hacer una búsqueda en Twitter y recopilar todos los mensajes vinculados a ciertos términos de búsqueda. Por tal motivo, se utilizó la API de Twitter a partir del lenguaje de programación R (<http://www.r-project.org/>).

El proceso de extracción de información y análisis de los mensajes está compuesto por varias etapas que se describen a continuación (ver Fig. 1):

1. El usuario hace una búsqueda a partir de un hashtag, por ejemplo *#UberSeQueda*.
2. El sistema recupera a partir del API de Twitter y utilizando R (a través de su librería *twitterR*) los tuits asociados al hashtag insertado por el usuario.
3. Los tuits son guardados en una base de datos. Estos tuits extraídos conforman el corpus inicial y contienen 16 parámetros entre los que se encuentra el ID del tuitero, si el tuit es un retuit (RT), la localización geográfica, el URL del tuit inicial, etc.
4. Una primera etapa corresponde a la detección de actores en las que se realiza un pre-procesamiento que consiste en la eliminación de palabras sin sentido (*stop words*) y la construcción del corpus a partir de fechas indicadas.
5. Después del pre-procesamiento se recorre cada uno de los tuits para extraer al autor principal y a aquellos que están mencionados. Es decir que, se recuperan todos los autores que están precedidos por el símbolo de @.
6. Los actores son catalogados en actores autores y en actores citados.
7. A partir de los actores autores se hace una nueva extracción en Twitter para conocer otros datos como son el número de tuits favoritos, el número de tuits retuiteados, el número de amigos y seguidores, etc.
8. El cálculo de la influencia de cada uno de los actores está dado por el número de amigos y seguidores pero también por el número de mensajes que le fueron



retuiteados o puestos como favoritos, el número de menciones y sus propios contactos con otros actores influyentes.

9. La influencia de los actores se presenta en forma de grafo en el cual cada nodo tiene un tamaño de acuerdo a su posición dentro de toda la red. Esta posición depende de qué tan influyente es con respecto a los demás nodos (actores) y los mensajes que tuitea.
10. El usuario recibe como resultado el grafo generado junto con una lista en orden descendente de los usuarios más influyentes.

En la siguiente sección se describe un caso de estudio y los resultados obtenidos.

#### 4. Resultados

Para el presente estudio se utilizó el hashtag *#UberSeQueda* para recuperar todos los tuits entre el 25 y 27 de mayo de 2015. La Tabla 1 muestra algunas de las características principales del corpus.

Número de tuits	Usuarios únicos	Total de palabras	Términos más importantes	Longitud máxima de términos
10,000	6,460	163,120	7,505	29

**Tabla 1. Características principales del corpus *#UberSeQueda* del 25 al 27 de mayo.**

El proceso de detección de actores consiste en identificar a los autores de cada tuit que corresponden a uno de los parámetros extraídos junto con todo el corpus. Entre los parámetros que pueden recuperarse y que son interesantes están el ID del usuario, la fecha de creación (*created*), si es un reply y a qué ID se le hizo (*replyToUID*), el nombre del usuario (*screenName*), el número de retuits (*retweetCount*), etc. El ID del usuario es muy importante debido a que el nombre que utiliza puede cambiar constantemente y, en cambio, el ID es permanente. Por lo tanto, aunque se muestran los nombres de los autores o de aquellos que están citados es importante asociar el ID a cada uno de ellos.

El corpus es recorrido en su totalidad para buscar los autores de los tuits y son guardados para que sean contabilizados una sola vez. Es decir, lo que es fundamental es encontrar cuál es el número de autores únicos. En este caso, el corpus está compuesto por 6,460 autores únicos. Aunque algunos de estos son autores de múltiples tuits, el guardarlos una única vez permite definir las interacciones entre los autores y tuiteros citados o mencionados en una sola ocasión aunque la interacción se realice en múltiples ocasiones.

Para cada uno de los diez mil tuits se busca que exista una cadena precedida por el símbolo de @ el cual indica la cita o mención de algún usuario. Por ejemplo, en el siguiente tuit creado por el autor @DenunciasDF están citados @CDMX\_Semovi @HLeonRufino @MEYERKG:

*Haz patria. Dale RT #UberSeQueda @CDMX\_Semovi @HLeonRufino @MEYERKG  
#MueraPRD #NiUnVotoAIPrd <http://t.co/AsEYdhFOza>*

A partir de cada tuit en el que se encuentran autores citados se crea o complementa una tabla en la que se muestran las relaciones. Por ejemplo, para el tuit anterior la tabla correspondiente (ver Tabla 2) consiste de una columna del autor y otra del actor mencionado.

Autor	Mención
@DenunciasDF	@CDMX_Semovi
@DenunciasDF	@HLeonRufino
@DenunciasDF	@MEYERKG

**Tabla 2. Relación del autor con el actor mencionado.**

La generación de las relaciones entre los actores y la lista de actores únicos permite construir dos tablas distintas: 1) tabla de nodos y 2) tabla de aristas; a partir de las cuales es posible recuperar algunos datos que son importantes para la detección de la

influencia. La construcción de un grafo (formado por nodos y relaciones entre estos) permite la visualización de características importantes que difícilmente pueden analizarse únicamente en una tabla de datos.

Para cada uno de los actores únicos se recuperan datos como el número de retuits (RT), el número de tuits favoritos, el número de seguidores y de amigos. A partir de ello se construye un grafo que muestra como puntos centrales a los actores influyentes y los agrupa en comunidades de “amistad”. Estas comunidades agrupan a los actores de acuerdo a las relaciones que tienen tanto de comunicación como de amistad. En el siguiente grafo (ver Fig. 2) se muestran ya solo los actores influyentes con tal de minimizar en lo más posible la lectura del mismo.

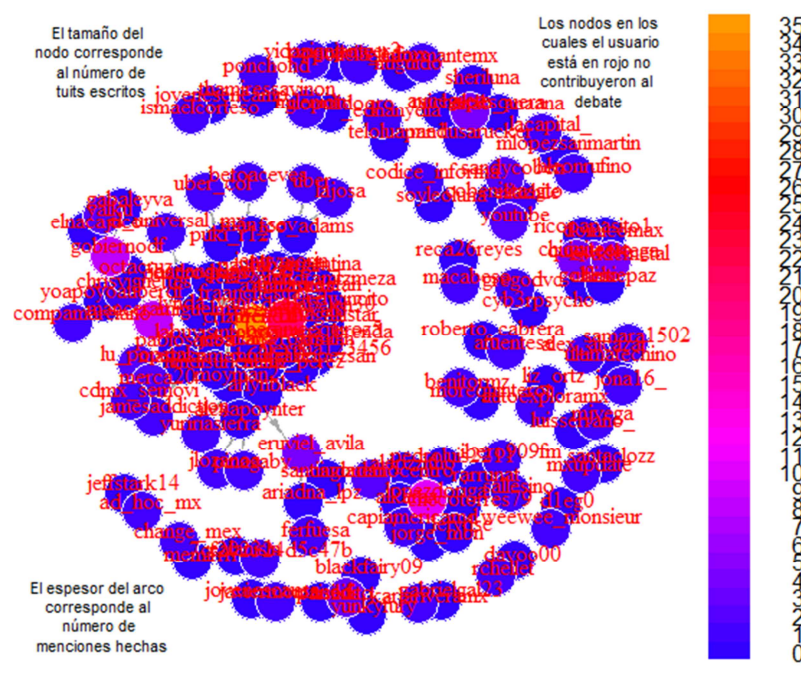


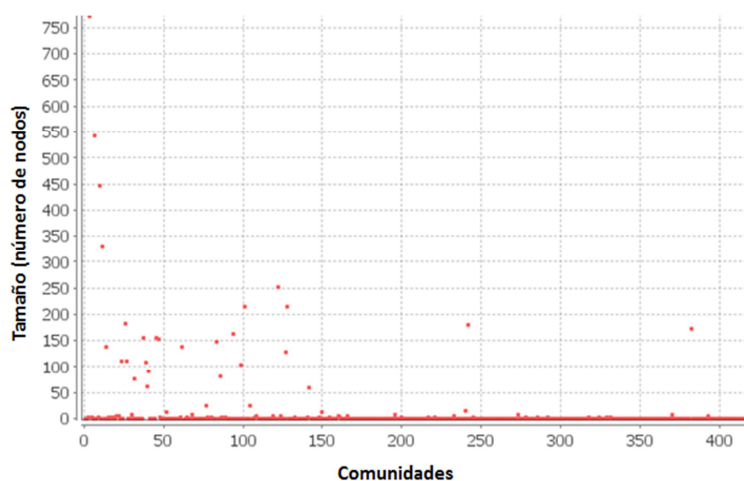
Fig. 2. Grafo con identificación de actores influyentes.

Para evaluar la metodología se optó por utilizar el mismo corpus y el uso de Gephi la cual es una plataforma de código abierto para la visualización, creación y consulta de grafos. Las etapas descritas previamente se siguieron hasta la creación de las tablas de autores y autores citados. Estas son el insumo principal en Gephi a partir de las cuales

es posible crear un grafo que muestra las relaciones entre los nodos existentes. Asimismo, en Gephi se aplicaron los siguientes algoritmos:

- *Grado*: es el número de conexiones de un vértice o nodo con otros nodos.
- *Centralidad de intermediación*: La intermediación es la frecuencia con la cual un nodo (o actor) se encuentra entre un par de otros nodos en el camino más corto o geodésico conectando a ellos. Lo que sugiere [13] es que un individuo puede muy bien estar ligeramente conectado a los otros (es decir, centralidad de grado baja) y sin embargo ser un intermediario esencial en los intercambios
- *Modularidad*: permite medir la fuerza de la división de una red en módulos (también llamados grupos, agrupamientos o comunidades). Las redes con alta modularidad tienen conexiones sólidas entre los nodos dentro de los módulos, pero escasas conexiones entre nodos en diferentes módulos.

En la siguiente figura (ver Fig. 3) se muestra el número de comunidades detectadas que para este caso son 422. En realidad, lo interesante es hacer un filtrado de los resultados y solo visualizar el grafo que contiene las comunidades más aglomeradas tales como aquéllas cuyo número de nodos es mayor a 100.



**Fig. 3. Detección de 422 comunidades que agrupan autores de acuerdo a la comunicación establecida en los tuits.**

En la Fig. 4 se encuentra el grafo completo con los nodos más importantes etiquetados. Además de conocer las comunidades con mayor concentración; la centralidad intermediaria y el grado son fundamentales para conocer a los usuarios influyentes. Para todos los actores se tienen los siguientes datos:

- **Número de comunidad a la que pertenece**
- **Centralidad de intermediación**
- **Grado**
- **Excentricidad:** el número de asociación, también llamado así, de un nodo es la distancia geodésica mayor entre ese nodo y cualquier otro nodo.
- **Centralidad de proximidad:** otra manera de medir la centralidad de un individuo (nodo) consiste en medir su grado de proximidad con respecto a todos los otros individuos. Se trata de una medida más global utilizando no solamente las conexiones de un individuo a su vecindario (*neighbourhood*), sino que su proximidad con el conjunto de los miembros de la red.

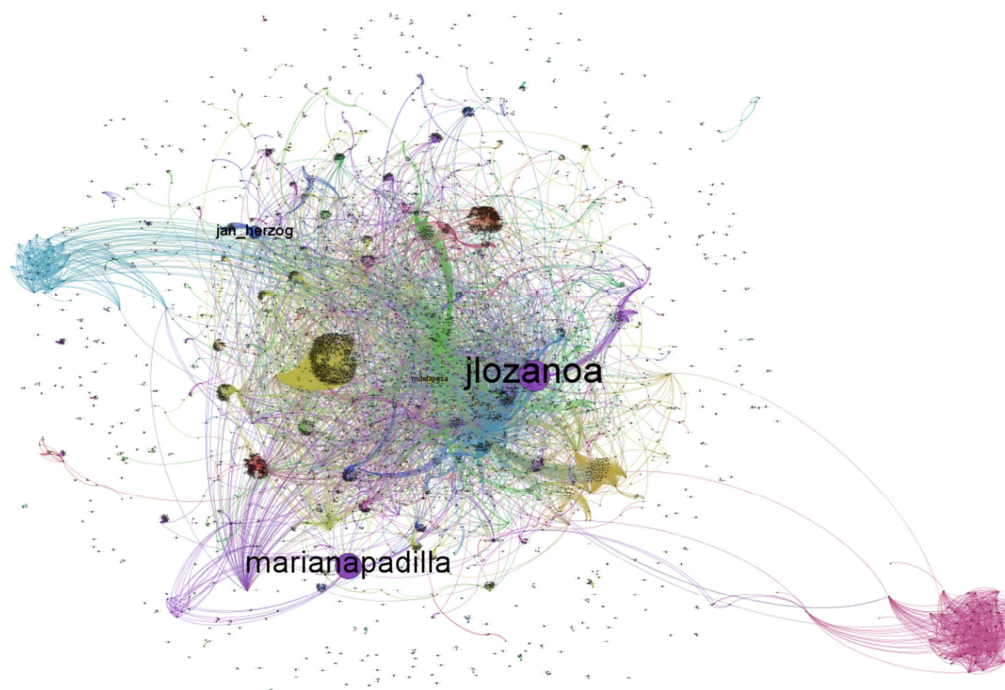


Fig. 4. Grafo completo a partir del uso de Gephi.

En la siguiente tabla (ver Tabla 3) se muestran los datos obtenidos para los principales autores.

<b>@autor</b>	<b>Grado</b>	<b>Excentricidad</b>	<b>Centralidad de intermediación</b>	<b>Centralidad de proximidad</b>	<b>Comunidad</b>
@carce55	29	1	1,0681	1	121
@change_mex	80	1	1,384.333	1	9
@jan_herzog	78	2	4,839.5	1.176	26
@jlozanao	228	4	11,807.5	2.491	100
@marianapadilla	31	3	9,934	1.849	100
@mdelapeza	11	5	2,283.5	2.887	381

**Tabla 3. Atributos para cada uno de la muestra de actores influyentes.**

En la Tabla 3 se muestra el grado el cual si es tomado, únicamente, como característica principal para la determinación de la influencia habrá un error. Aun cuando un usuario tenga muchos contactos no quiere decir que sea influyente. Por lo tanto, es necesario contar con atributos tales como la centralidad de proximidad que toman en cuenta la topología no solo del nodo principal pero también de sus vecinos.

Con ambas experimentaciones los resultados se complementan. Por una parte, el primer ejercicio muestra los actores de acuerdo a su número histórico de menciones, de favoritos y al número de seguidores y amistades con los que cuenta cada tuitero. Por otra parte, el segundo ejercicio está basado en el corpus actual en el cual el grafo es determinante para indicar si un usuario es influyente o no. Los resultados del segundo ejercicio se basan en la red que resulta de las menciones entre los autores del corpus extraído. Una combinación de ambos métodos nos permite seleccionar a los actores influyentes tomando en cuenta el histórico como su participación en un momento

preciso del movimiento social. En ambos casos es posible apreciar la agrupación de nodos con respecto a las relaciones existentes con otros nodos.

## **5. Discusión**

Los resultados anteriores evidencian que una solución para la detección adecuada de actores influyentes debe darse en base a la combinación del perfil del usuario (menciones, retuits, número de seguidores, listas, etc.) además del contexto actual (tuits enviados durante un periodo o movimiento preciso).

En este artículo se presentan dos métodos por separado. El uso del perfil del usuario puede dar información o, incluso, categorizar a los tipos de usuarios de acuerdo a un histórico. En cambio, el uso del contexto actual permitiría dirigir mensajes claros a usuarios específicos o dar seguimiento a los acontecimientos a partir de los usuarios detectados como influyentes.

## **6. Conclusiones**

El presente artículo muestra una metodología específica basada en la detección de los autores de los tuits y las menciones que se hacen de otros actores. El propósito del artículo es evidenciar la necesidad de fusionar métodos existentes para considerar el perfil del usuario así como el contexto actual de un movimiento social.

El ARS permite apoyarse en la topología de la red para detectar patrones y características que no pueden valorarse solamente con la información extraída de Twitter.

Siendo Twitter una de las redes sociales más importantes a nivel mundial es necesario contar con herramientas que permitan una rápida extracción de información.

El trabajo futuro se centra en la creación de una metodología que reúna tanto el histórico como el contexto actual. De igual manera, es importante definir una tipología de usuarios.

## **7. Referencias**

- [1] A. Tumasjan, T. Sprenger, et al., "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment". International AAAI Conference on Weblogs and Social Media. Vol. 4. 2010. 178-185 pp.
- [2] M. Castells, *Redes de indignación y esperanza: los movimientos sociales en la era de Internet*. 2012. Cambridge: Polity Press.
- [3] M. Gladwell, *The Tipping Point*. 2002. Little, Brown and Company. Nueva York.
- [4] M. Granovetter, "The strength of weak ties". *American Journal of Sociology*. Vol. 78. 1973. 1360-1380 pp.
- [5] J. Berger, K. L. Milkman, "What makes online content viral?" *Journal of Marketing Research*. Vol. 49. No. 2. 192-205 pp.
- [6] E. Bakshy, I. Rosenn, C. Marlow, L. Adamic, "The role of social networks in information diffusion". *Proceedings of the 21st international conference on World Wide Web*. ACM. 2012. 519-528 pp.
- [7] M. Cha, Haddadi, F. Benevenuto, P. K. Gummadi, "Measuring User Influence in Twitter: The Million Follower Fallacy". *ICWSM*. 2010. 10-17 pp.
- [8] J. Weng, E. P. Lim, J. Jiang, Q. He, "Twitterrank: finding topic-sensitive influential twitterers". *Proceedings of the third ACM international conference on Web search and data mining*. ACM. 2010. 261-270 pp.



- [9] S. Wu, J. M. Hofman, W. A. Mason, D.J. Watts, "Who says what to whom on twitter". Proceedings of the 20th international conference on World Wide Web. ACM. 2011. 705-714 pp.
- [10] D. M. Romero, W. Galuba, S. Asur, B. A. Huberman, "Influence and passivity in social media". Machine learning and knowledge discovery in databases. Springer Berlin Heidelberg. 2011. 18-33 pp.
- [11] A. Goyal, F. Bonchi, L. V. Lakshmanan, "Learning influence probabilities in social networks". Proceedings of the third ACM international conference on Web search and data mining. 2010. ACM. 241-250 pp.
- [12] G. Rattanaritnont, M. Toyoda, M. Kitsuregawa, "Analyzing patterns of information cascades based on users' influence and posting behaviors". Proceedings of the 2nd Temporal Web Analytics Workshop. 2012. ACM. 1-8 pp.
- [13] L. C. Freeman, "Centrality in social networks: I Conceptual clarification". Social Networks, Vol. 1. 1979. 215- 239 pp.

## **8. Autores**

La Dra. Rocío Abascal Mena es profesor-investigador en el Departamento de Tecnologías de la Información de la Universidad Autónoma Metropolitana – Cuajimalpa en la Ciudad de México. Realizó su doctorado en el Instituto Nacional de Ciencias Aplicadas (INSA) de Lyon, Francia, en el 2005. Sus intereses actuales de investigación incluyen: web semántica, lingüística computacional, recuperación y análisis de información, Interacción Humano Computadora, análisis de redes sociales.

El Dr. Erick López-Ornelas es profesor-investigador en el Departamento de Tecnologías de la Información de la Universidad Autónoma Metropolitana - Cuajimalpa en la Ciudad de México. Recibió su doctorado en Informática en la Universidad Paul Sabatier de Toulouse Francia, en el 2005. Sus intereses de investigación incluyen: análisis de redes sociales, sistemas de información geográfica, visualización geográfica, percepción remota y geomática. También ha trabajado en aplicaciones de HCI y sistemas contextuales.

El Dr. Sergio Zepeda-Hernández realizó su doctorado en Ingeniería Eléctrica en la especialidad de Informática del Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional (México) en 2009. Actualmente es profesor-investigador del Departamento de Tecnologías de la Información de la Universidad Autónoma Metropolitana - Cuajimalpa en la Ciudad de México. Sus intereses de investigación incluyen la ingeniería web, análisis de redes sociales, recuperación de información, Interacción Humano Computadora, usabilidad y diseño de interfaces.