

## IDENTIFICATION OF INFLUENTIAL USERS IN FACEBOOK<sup>i</sup>

IDENTIFICACIÓN DE USUARIOS INFLUYENTES EN FACEBOOK

IDENTIFICAÇÃO OS USUÁRIOS INFLUENTES DO FACEBOOK

MARISA ANALÍA SÁNCHEZ<sup>ii</sup>

### CITATION

Sánchez, Marisa A. (2018) Identification of influential users in facebook. *Dimensión Empresarial*, 16(2), 7-20. DOI: <http://dx.doi.org/10.15665/rde.v15i1+E1.865>

JEL: M3, M39

### ABSTRACT

Social networks are a mode of internet users to communicate and share information. From the area of marketing there is an interest in identifying individuals who may have influence to promote products or services. The purpose of this research work is to present a method on how to identify influential users of Facebook Fan Page. The proposal is based on the works of Khobzi y Teimourpour (2015) and Weng y Lento (2014) and takes into account pragmatic issues derived from data Access restrictions. A real case is described. The results allow identifying user segments and the interest of different topics. This research is relevant to Fan Page managers and to the marketing area of organisations with social media presence.

**Key words:** Social networks, Facebook, influential users.

### RESUMEN

Las redes sociales constituyen un medio para que los usuarios de Internet se comuniquen y compartan información. Desde el área de marketing existe un interés por identificar individuos que puedan tener influencia para promocionar sus productos o servicios. El objetivo de este trabajo de investigación es presentar un método para identificar usuarios influyentes en una Fan Page de Facebook. La propuesta está basada en los trabajos de Khobzi y Teimourpour (2015) y Weng y Lento (2014) y considera aspectos pragmáticos derivados de las restricciones en el acceso a datos privados. Se describe un caso real. Los resultados permiten identificar segmentos de usuarios y el interés de diferentes tópicos. Esta investigación es relevante para los administradores de una Fan Page y para el área de marketing de organizaciones con una presencia en los medios sociales.

Palabras clave: Redes Sociales, Facebook, usuarios influyentes.

### RESUMO

A rede social é uma maneira para os usuários da Internet para se comunicar e compartilhar informações. Desde a área de marketing há um interesse na identificação de indivíduos que podem ter influência para promover seus produtos ou serviços. O objetivo desta pesquisa é apresentar um método para identificar os usuários influentes em um Facebook Fan Page. A proposta baseia-se no trabalho de Khobzi e Teimourpour (2015) e Weng e Lento (2014) e considera aspectos pragmáticos decorrentes das restrições de acesso a dados privados. Ele descreve um caso real. Os resultados permitem identificar os segmentos de usuários e o interesse de diferentes temas. Esta pesquisa é relevante para os gestores de uma Fan Page e para área de marketing de organizações com uma presença nos meios de comunicação social.

Palavras chave: redes sociais, Facebook, usuários influentes.

## INTRODUCCIÓN

Muchas redes sociales tales como Facebook hacen posible que entidades tales como empresas, marcas y figuras públicas puedan crear una Fan Page e interactúen con sus fans o clientes en un mundo virtual (Khobzi & Teimourpour, 2015). De esta forma, los medios sociales tienen el potencial de convertirse en un nuevo canal de comunicación con los clientes. Algunos autores indican que las redes sociales son particularmente promisorias para identificar individuos influyentes, habilitar una distribución efectiva de mensajes personalizados e interactuar con segmentos de clientes numerosos de diferentes formas (Canhoto, Clark, & Fennemore, 2013). En particular, en este trabajo nos concentramos en el uso de las redes sociales como herramienta para posicionar la oferta de una empresa. A tal efecto, las redes sociales permiten la distribución de propuestas altamente diferenciadas, y la interacción con los clientes permite un compromiso más profundo. Una de las preguntas que un administrador de medios sociales se hace es cómo hacer una gestión de la Fan Page más eficiente y efectiva. Para responder este interrogante es necesario analizar la conducta de los usuarios incluyendo cómo los usuarios se involucran con los posts de una Fan Page; qué usuarios resultan más promisorios para incrementar la popularidad de una Fan Page; y detectar si las interacciones dependen de los temas asociados a los posts.

El objetivo de este trabajo es elaborar un marco para identificar usuarios influyentes en una Fan Page de Facebook. Se definirá una propuesta basada en el análisis de Redes Sociales y cuya aplicación resulte factible a la luz de restricciones de acceso a información privada. Para evaluar la propuesta se presenta un caso basado en datos recopilados de una Fan Page de un gimnasio multidisciplinario.

El artículo se organiza de la siguiente forma. Después de presentar un breve marco teórico y un análisis de la

literatura relacionada, se describe la metodología. Luego, se presenta un ejemplo. Finalmente, se describen las contribuciones y las limitaciones de la propuesta.

## MARCO TEÓRICO

### Plataformas de redes sociales

La era moderna de las redes sociales comenzó con la mejora de la performance de Internet a partir de 1995. En el periodo 2002-2004 aparecieron y se promocionaron Cyworld, Friendster, Plaxo, Reunion.com, Hi5, LinkedIn, MySpace, Orkut, Facebook, y Live Spaces (Skeels & Grudin, 2009). Facebook comenzó como un sitio orientado a los jóvenes en ambientes universitarios y actualmente es utilizado para compartir fotos o vínculos con amigos o publicar mensajes en la “biografía” de otra persona. Dado la próspera tendencia de la utilización Facebook muchas organizaciones se han convertido en unos de los principales usuarios para sus estrategias de marketing (Narayanan, y otros, 2012). Las empresas pagan post promocionales, utilizan servicios de publicidad o crean una “Fan page” en forma gratuita. Los posts de las Fan page se difunden al público en general y a usuarios que suscriben a una Fan page seleccionando el ícono “Me gusta” en la Fan page de la compañía (Shin, Lee, & Hall, 2014).

### Análisis de redes sociales

El Análisis de Redes Sociales aborda el tratamiento de redes de dos formas. Una utiliza redes egocéntricas que proveen una vista desde la perspectiva del actor en la red, y la otra conceptualiza en la estructura completa de la red (Hanneman & Riddle, 2005), (Edwards, 2010), (Haythornthwaite, 1996). En el primer caso, los vínculos que los actores individuales mantienen con otros describen su propia red egocéntrica. Describen una foto de un actor típico en un ambiente en particular

y muestran cuántos vínculos tiene un actor con otro, qué tipo de vínculos mantienen, y qué tipo de información brindan o reciben de otros actores en la red. Este análisis es útil cuando el tamaño de la población es grande o cuando es difícil definir los límites de la población.

En el segundo caso, las redes socio-céntricas o completas describen los vínculos que todos los miembros de un ambiente mantienen con el resto de los miembros del ambiente. En principio, este enfoque requiere respuestas de todos los miembros de un ambiente sobre el resto de los miembros. Este requerimiento limita el tamaño de las redes que se pueden examinar.

Los datos relacionales que se recopilan en un estudio de redes sociales se representan con grafos. Los actores son nodos en el grafo, y las relaciones se representan con arcos entre los nodos. En los grafos dirigidos, un arco es un par ordenado, y el par ordenado representa la dirección del arco que vincula dos vértices. En los grafos no-dirigidos, no existe una dirección asociada al arco.

La posición de un nodo dentro de su red permite inferir su importancia en el intercambio de información y conocimiento. Se utilizan diferentes indicadores basados en los vértices. El Grado está dado por la cantidad de arcos que inciden en un vértice. Los vértices que están incluidos en muchos de los caminos más cortos entre otros vértices poseen la Centralidad de Intermediación más alta y tienen la capacidad de conectar dos nodos que de otra forma no se vincularían. La Centralidad de Cercanía se define como el promedio de las distancias más cortas de cada vértice hacia el resto. Los vértices con una centralidad baja son capaces de obtener o distribuir mensajes al resto en forma rápida (con pocos pasos). La Centralidad de Autovector considera el grado de un vértice y el grado de sus vecinos. De esta forma, constituye un indicador de la popularidad de los vecinos de un vértice.

Algunos de los indicadores globales del grafo están dados por la distancia geodésica (la distancia más corta entre dos vértices); y la densidad del grafo que indica cuán interconectados están los vértices entre sí. Una red con una densidad del 100% se lograría si todos los nodos estuvieran conectados entre sí. Algunos estudios indican que una estructura poco densa (no redundante) favorece el flujo de información y conocimiento dentro de la red. Por otro lado, la densidad puede ser clave para la colaboración dado que facilita la construcción de confianza.

## REVISIÓN DE LA LITERATURA

La literatura vinculada con el uso de redes sociales para dar apoyo a objetivos de marketing es muy amplia. A continuación, consideramos con particular atención las contribuciones que se concentran en el problema de identificar a los usuarios más influyentes. Jucaityte y Mašcinskiene (2014) presentan una revisión muy completa del estado del arte sobre la integración de los medios sociales en la comunicación de marketing. Los autores señalan que la diferencia esencial que distingue a los medios sociales de los tradicionales reside en la participación creciente del consumidor. Esto se logra a través del estímulo de la actividad del consumidor; apertura a la participación; participación mediante votos, comentarios o información compartida; la comunicación bilateral; y la posibilidad de crear comunidades con intereses distintivos. Además, recopilan varios de los problemas que las empresas enfrentan al utilizar los medios sociales, a saber, dado que los consumidores podrían molestarse por el exceso de publicidad, entonces la comunicación debe ser sensible y sutil. Estos autores agregan que los criterios de segmentación tradicionales (tales como edad, lugar de residencia o género) son difíciles de implementar en los medios sociales dado que los usuarios tienden a agruparse de acuerdo a intereses o posturas.

Muruganatham y Gandhi (2005) consideran el problema de identificar a los usuarios más influyentes que pueden generar el mayor beneficio con respecto a tópicos o situaciones de negocio específicas. Proponen un método que utiliza métricas de centralidad y luego define un orden de preferencia de los usuarios de mayor influencia de acuerdo a características de homofilia tales como género, estado civil y localidad de residencia. La homofilia es el principio que indica que el contacto entre individuos similares ocurre a una tasa más alta que entre individuos no similares (McPherson, Smith-Lovin, & Cook, 2001). La metodología de Muruganatham y Gandhi debe aplicarse sobre una muestra representativa de posts. Algunas restricciones para aplicar la propuesta resultan de la posibilidad de poder recolectar los datos necesarios. Por ejemplo, los usuarios de Facebook pueden ajustar la configuración de privacidad de sus datos, y de ese modo no siempre es posible acceder a datos demográficos como la edad, el lugar de residencia, los intereses, entre otros.

Mochalova y Nanopulos (2014) proponen un método para detectar a usuarios influyentes en base al conocimiento previo del mercado y métricas de centralidad para identificar usuarios que tienen una alta probabilidad de alcanzar y activar muchos usuarios del mercado potencial (los denominan “semillas”). El término activar se refiere a la adopción del objetivo de una campaña de marketing viral tal como comprar un producto, recuperar información de este, entre otros. Para un sistema de redes sociales que permite hacer calificaciones de productos, los autores proponen seleccionar los 50 productos con mayor cantidad de comentarios. A partir de esos productos definen 50 mercados diferentes y los usuarios se corresponden con aquellos que mostraron preferencias por los productos. De esos usuarios se seleccionan las semillas en base a métricas de centralidad.

Majumdar, Saha y Dasgupta (2015) proponen un método para identificar potenciales embajadores de página para marcas de proveedores de servicios móviles. Definen un índice denominado “valor de

asociación” compuesto por la proximidad, similitud e interacción. La proximidad se mide considerando el número de saltos de separación entre el usuario y la página de interés. La similitud mide el grado en que el perfil de un usuario coincide con el conjunto de atributos definidos para una marca. La interacción se mide en términos de la frecuencia de una interacción (un usuario indica que le gusta un post, lo comenta o comparte) y su antigüedad. A efectos de validar la propuesta los autores debieron complementar los datos con la información que no pudieron recuperar debido a restricciones de acceso por parte de Facebook.

Una investigación próxima a los objetivos de este trabajo es la propuesta por Khobzi y Teimourpour (2015). Los autores proponen un método para segmentar a los usuarios de una Fan Page en base a indicadores de la popularidad de un post tales como la cantidad de “Me gusta”, comentarios y un puntaje que muestra el grado en que los comentarios de un usuario resultan positivos o negativos. Recopilan datos correspondientes a la ventana temporal de interés (por ejemplo, los datos de los últimos 100 posts de una Fan Page) y aplican un algoritmo de agrupamiento o *clustering* basado en los indicadores mencionados. Presentan un caso de estudio para el cual determinan cuatro aglomerados que denominan los fans apáticos (es el grupo mayoritario y participan con pocos “Me gusta” y comentarios); los fans firmes (es el grupo minoritario con menos del 0,02% del total, y se observa la media de cantidad de “Me gusta” y comentarios, y polaridad muy superior al resto); los fans comunes (si bien participan poco los comentarios son positivos); y los fans perezosos (participan indicando que les gusta un post pero difícilmente realizan comentarios).

Weng y Lento (2014) estudian cómo los individuos se comunican en línea en base a los tópicos de conversación desde una perspectiva egocéntrica. La metodología se basa en recopilar una muestra de perfiles de la red Facebook y se reconstruye una red con los amigos que han tenido al menos una referencia al post del ego en la ventana temporal analizada. A

partir de esa red se definen otras redes basadas en diferentes tópicos. Luego, para cada red basada en un tópico se analizan tres aspectos: homogeneidad basada en la densidad; crecimiento durante la ventana temporal de estudio; y el solapamiento de los usuarios que participan en más de un tópico. En resumen, para un conjunto de perfiles se estudia la relación entre la estructura de amistad y los tópicos de conversación.

## METODOLOGÍA

En esta sección se presenta la metodología de trabajo para identificar el conjunto de usuarios activos de una Fan Page que resultan más promisorios para incrementar la popularidad de la misma. La metodología se basa en las contribuciones de los trabajos de Khobzi y Teimourpour (2015) y Weng y Lento (2014). En el primer caso los autores indican cómo analizar una Fan Page en base a la participación de los usuarios y la polaridad de los comentarios. Pero no consideran el comportamiento para diferentes tópicos. Weng *et al.* indican cómo analizar la estructura de la red en base a tópicos, pero a diferencia del problema planteado en este trabajo, construyen la red a partir de usuarios que forman parte de una red de amistad.

### Definiciones

A continuación, se presentan las definiciones de varios términos tal como se utilizan en el trabajo. Las definiciones están basadas en el trabajo de Weng *et al.* (2014).

**Red ego.** Una red ego  $R_u$  es una red de conexiones centrada en un actor ego  $u$ .  $R_u$  contiene  $n_u$  usuarios activos de  $u$ , representados como  $V_u = \{v_u^i | 1 \leq i \leq n_u\}$ . Existe un arco entre  $v_u^i$  y  $v_u^j$  ( $v_u^i, v_u^j \in V_u$  y tal que  $i \neq j$ ) si  $v_u^i$  indicó que le gusta o comentó un post que le gusta, comentó o compartió el

usuario  $v_u^j$ . Existen  $a_u$  arcos en total entre todos los  $a_u$  usuarios.

**Tópico.** El ego  $u$  crea un post en Facebook, y a cada post se le asigna una etiqueta de tópico  $t_u^k \in T_u$ , donde  $T_u$  incluye todos los tópicos sobre los cuales  $u$  ha posteado.

**Respuesta.** Un alter  $v_u^i \in V_u$  se considera interesado en un tópico  $t_u^k$  si indica que le gusta, comenta o comparte posts sobre  $t_u^k$ . La intensidad de  $v_u^i$  respondiendo a  $t_u^k$  se cuantifica con la suma de “Me gusta”, “Comentar” y “Compartir” y se denota como  $r_u^{i,k}$ .

**Grupo basado en un tópico.** Un grupo o aglomerado  $A_u^k$  es un sub-grafo de  $R_u$ , compuesto de alters con respuestas al tópico  $t_u^k$ ,  $\{v_u^i | r_u^{i,k} > 0, 1 \leq i \leq n_u\}$ . El aglomerado  $A_u^k$  contiene  $n_u^k$  nodos y  $a_u^k$  arcos creados en la ventana temporal de observación ( $n_u^k \leq n_u, a_u^k \leq a_u$ ).

La principal diferencia con el trabajo de Weng y Lento (2014) radica en la definición de la red ego, dado que los autores mencionados conectan usuarios si son amigos. Esto implica que para el resto del análisis, estudian sub-grafos del grafo que representa a la red de amistad.

### Definición del problema

El problema de investigación se refiere a identificar un conjunto  $I \subset V_u$  de usuarios activos de  $u$  que resultan más promisorios para incrementar la popularidad de una Fan Page, representados como  $I_u = \{i_u^j | 1 \leq j \leq n_u\}$ . Se asume que la intensidad de  $i_u^j$  respondiendo a  $t_u^k$  refleja la influencia de  $i_u^j$ . Asimismo, dados dos grupos  $A_u^k$  y  $A_u^l$ , los usuarios  $i_u^j$  tal que  $i_u^j \in A_u^k \wedge i_u^j \in A_u^l$  reflejan la influencia de  $i_u^j$  para incrementar la popularidad de más de un tópico. En la Fan Page participan diferentes personas indicando que les gusta, comentando o compartiendo posts. El

administrador de la Fan Page crea posts correspondientes a diferentes tópicos, por lo cual, podría esperarse que para cada tópico haya diferentes conjuntos de individuos influyentes.

## Etapas

Para realizar un análisis completo de la Fan Page se propone realizar cuatro pasos que incluyen (a) un entendimiento de los objetivos de la empresa y clasificación de tópicos; (b) recolección y preparación de datos; (c) análisis de grupos; y (d) determinación de aglomerados. A continuación, se describe cada etapa.

**Entendimiento de los objetivos de la empresa y clasificación de tópicos.** A efectos de analizar los posts se decide clasificarlos en tópicos, para entender qué tipo de tópicos resultan más atractivos. Se deben entender los objetivos de la organización a los cuales los administradores de la Fan Page pretenden dar apoyo. De esta forma, se pueden determinar tópicos relevantes para esos objetivos y analizar la conducta de los usuarios con respecto a cada uno. Si los creadores de los posts son los administradores de la Fan Page, entonces, ellos debieran definir qué tópicos desean analizar a partir del conocimiento de la organización. Si se tratara de una Fan Page en donde muchos usuarios contribuyen en la creación de posts, entonces se debieran identificar los tópicos a partir del análisis de los posts. Por ejemplo, Weng *et al.* (2014) recolectan datos de 65.000 usuarios activos de Facebook seleccionados en forma aleatoria. Considerando que muchos posts no están asociados a un texto descriptivo o son breves e informales, los autores se concentraron en URL compartidos dado que cada URL se refiere a una página web que provee información de contexto. De esta forma, identificaron los dominios más populares y los etiquetaron utilizando una lista de veinte etiquetas predefinidas.

**Recolección y preparación de datos.** Existen diversas herramientas que permiten extraer información de una Fan Page de Facebook incluyendo las

interacciones de “Me gusta” y “Comentar” para cada post de interés. Para efectuar el análisis de sentimientos de los comentarios es necesario analizar cada comentario para computar un puntaje. Si bien existen varias herramientas disponibles, y caracterizadas por distintos niveles de sofisticación (Pang & Lee, 2008), la posibilidad de aplicación depende del idioma en el cual están escritos los textos a analizar. Martínez-Cámara, Martín-Valdivia, Ureña-López y Mitkov (2014) analizan varios algoritmos de análisis de sentimientos para tweets en español.

**Análisis de grupos basados en tópicos.** El estudio de los aglomerados incluye el análisis de su homogeneidad, la ventana temporal de interés, la segmentación basada en características demográficas (en este trabajo solo se considera el género), y el solapamiento.

La homogeneidad de un aglomerado  $A_u^k$  se cuantifica calculando la densidad. Si en un grupo se incluyen varios posts de un mismo tópico, entonces se puede calcular la media de la densidad y el desvío estándar. De esta forma, la media de la densidad de cada aglomerado se contrasta con la densidad de  $R_u$  para testear si las diferencias son estadísticamente significativas. Si la densidad de un tópico resulta significativamente superior a la de  $R_u$ , existe evidencia de homofilia basada en el tópico, dado que los usuarios que están interesados en un mismo tópico están más conectados. De acuerdo con la homofilia, los individuos similares tienen más predisposición a tener contacto entre sí que con individuos diferentes. La existencia de homofilia en ambientes virtuales ha sido observada en varios estudios empíricos. La similitud entre los individuos puede cuantificarse en términos de características demográficas, lugares de residencia o tópicos de interés. Si no se consideran suficientes posts en un grupo como para realizar pruebas estadísticas, se observa si la densidad del grupo es superior o no con respecto a la densidad de  $R_u$ .

La ventana temporal de interés para un aglomerado  $A_u^k$  se define por la cantidad de días entre la fecha más

temprana de creación de un post vinculado al tópico  $t_u^k$  y la fecha de la última referencia (“Me gusta”, “Comentar” o “Compartir”) observada para algún post vinculado a  $t_u^k$ .

El solapamiento de aglomerados constituye una forma para analizar la heterogeneidad entre aglomerados. En este trabajo nos interesa analizar si los usuarios que participan en la Fan Page comparten los mismos tópicos, y determinar cuáles usuarios participan de más de un tópico. Aquellos usuarios presentes en los solapamientos podrían representar a los usuarios más activos y que participan en la difusión de la mayoría de los tópicos. Un bajo grado de solapamiento podría sugerir heterogeneidad entre los aglomerados. Tal como propone Weng et al. (2014), se utiliza el índice de similitud de Jaccard entre nodos de cada par de aglomerados para cuantificar el solapamiento. Dados dos grupos  $A_u^k$  y  $A_u^l$ , el solapamiento se define como:

$$J(A_u^k, A_u^l) = \frac{|\{v_u^i | v_u^i \in A_u^k \wedge v_u^i \in A_u^l, 1 \leq i \leq n_u\}|}{|\{v_u^i | v_u^i \in A_u^k \vee v_u^i \in A_u^l, 1 \leq i \leq n_u\}|}$$

Para interpretar el índice, se puede comparar el índice para cada aglomerado con respecto a un índice de referencia. El índice de referencia se calcula para el grupo que representa la intersección de todos los grupos. Podría suceder que el conjunto de la intersección de todos los grupos resulte vacío, es decir, ningún usuario referencia a todos los tópicos. En tal caso, se puede comparar el índice para un par de aglomerados con respecto al promedio de los índices de solapamiento y determinar si existe una diferencia significativa.

### **Conglomerados basados en los valores de “Me gusta”, “Comentar”, “Compartir”, y Percepción.**

Un alter  $v_u^i \in V_u$  se considera interesado en un tópico  $t_u^k$  si indica que le gusta, comenta o comparte posts sobre  $t_u^k$ . Además, los comentarios pueden puntuarse de acuerdo al grado de comentarios positivos realizados. Tal como propone Khobzi *et al.*, se aplica un algoritmo de *clustering* considerando cuatro variables para segmentar a los usuarios activos de  $R_u$ . Las

variables representan la cantidad de “Me gusta”, “Comentar”, “Compartir” y polaridad de los comentarios.

## **ANÁLISIS DE UN CASO**

El caso considera una Fan Page asociada a un gimnasio multidisciplinario que tiene como objetivo promover la calidad de vida y el vivir bien, que abarca los conceptos de salud, belleza y diversión. El administrador de la Fan Page observa que los posts no se difunden y comparten demasiado. Por otro lado, desearía identificar a los usuarios que tienen mayor participación y potencial de influir en la red.

### **Entendimiento de los objetivos de la empresa y clasificación de tópicos**

A partir de una entrevista con un responsable de la organización se definen aspectos de interés para la organización. La Fan Page tiene como objetivo hacer que se conozca la organización. El administrador crea posts vinculados con capacitaciones, la difusión de nuevas tecnologías y servicios. En general, los posts tienen asociados texto e imágenes creadas por profesionales de diseño a fin de transmitir una imagen de excelencia y calidad. Cabe observar que los posts no están personalizados. Los tópicos de interés para los administradores están vinculados con el lanzamiento de una nueva actividad; la incorporación de tecnología; campañas solidarias; Campaña de promoción de pago; y concursos.

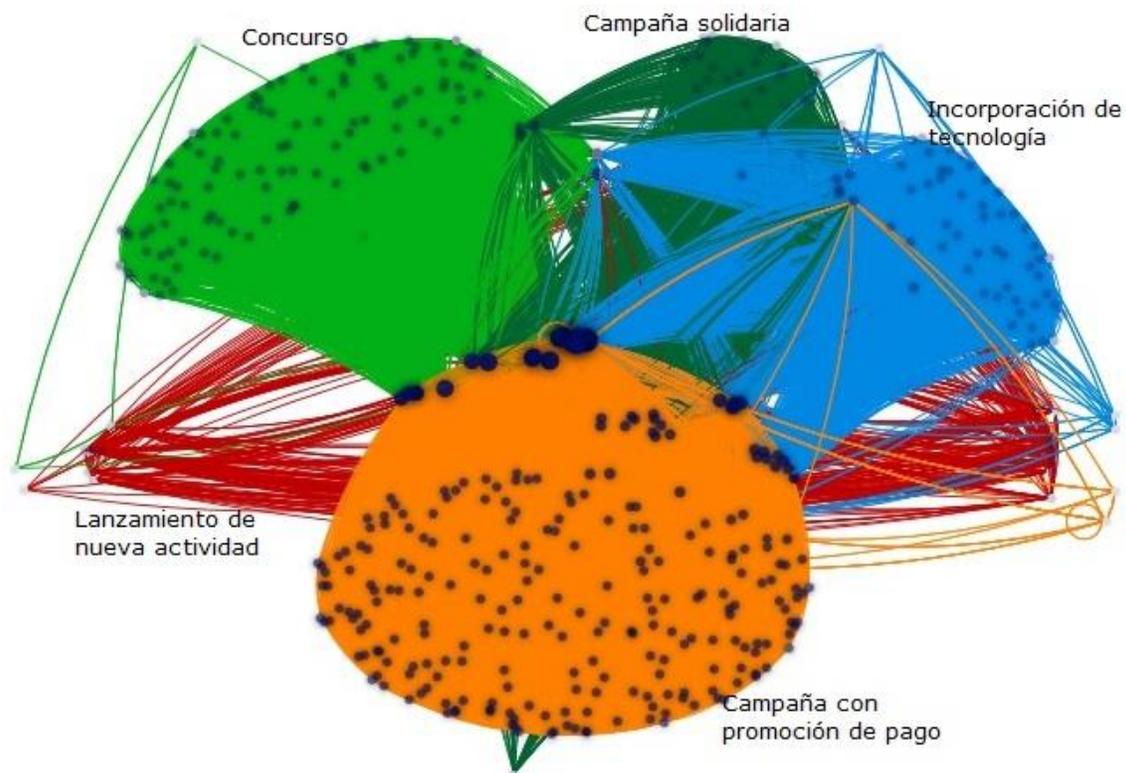
### **Recolección de datos**

A partir de la Fan Page de la empresa se recopiló la información correspondiente a cinco posts y cada post se refiere a una temática diferente. Cabe aclarar que se podrían haber recopilado información de varios posts

por tópicos. No se hizo dado que aún no se han registrado suficientes posts de cada tema en la Fan Page analizada. En base a cada post se elaboró una red en la cual los nodos corresponden con usuarios y se conectan los usuarios que indicaron “Me gusta” o comentaron

sobre el mismo post de acuerdo con las definiciones brindadas en la sección “Definiciones”. Para efectuar la recolección de datos y calcular métricas para los grafos de interés se utilizó la herramienta NodeXL (Hansen, Shneiderman, & Smith, 2011).

Figura 1. Grafo para la red completa. El color de los arcos está basado en el contenido del post. El tamaño de los vértices se corresponde con la centralidad de cercanía, y opacidad con el grado.



Fuente: elaboración propia

### Resultados del análisis de grupos o aglomerados

En la Tabla 1 se describen para cada uno de los aglomerados la cantidad de “Me gusta”, “Comentar”, y “Compartir”; la ventana temporal de interés; y características del grafo que los representa incluyendo la densidad. La densidad del grafo que representa a la red conformada por todos los posts analizados es de 0,39446127. Todos los posts (excepto “Lanzamiento de nueva actividad”) tienen una densidad superior, lo cual

sugiere que la propiedad de homofilia basada en los temas tratados está presente.

Si se observa la ventana temporal de interés de los tópicos concluimos que se mantienen activos por muy poco tiempo. Se podría esperar que el tópico “Concurso” y “Campaña solidaria” hubiera tenido referencia durante todo el período del concurso o la campaña.

Tabla 1. Descripción global de los sub-grafos correspondientes a cada post.

Tema del post	Número de “Me gusta”	Número de Comentarios	Número de Compartidos	Intervalo temporal de comentarios (días)	Métricas del sub-grafo		
					Nodo	Arcos (con duplicados)	Densidad
Lanzamiento de nueva actividad	35	7	0	3	39	604	0,385
Campaña solidaria	61	0	0	1	61	1831	0,484
Concurso	131	3	0	3	132	8396	0,450
Campaña de promoción de pago	260	11	0	1	261	33520	0,489
Incorporación de tecnología	79	0	0	1	85	3263	0,443

Fuente: elaboración propia

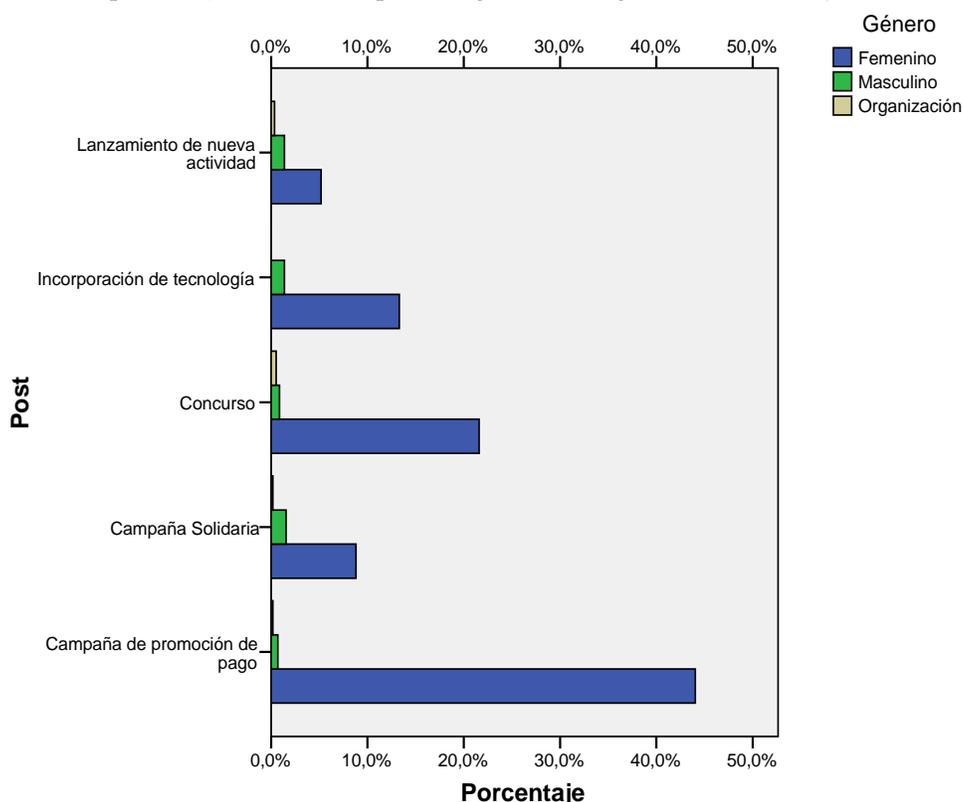
El 92,92% de los usuarios corresponden al género femenino (ver Figura 2). Cabe aclarar que, dado que la mayoría de los usuarios no completa el género en Facebook, se hizo una clasificación manual.

Para proceder con el análisis de solapamientos se deben tabular los datos correspondientes a la identificación del usuario, el post para el cual indicó “Me gusta”, comentó o compartió, e identificar el conjunto de usuarios en cada intersección. Para que sea factible desde el punto de vista práctico es necesario disponer de una herramienta que sistematice la tarea de determinar las intersecciones. Puede definirse una planilla de cálculo para facilitar el procesamiento, utilizarse algún paquete estadístico o utilizar alguna aplicación de análisis de conjuntos. En la Tabla 2 se incluye el cálculo del índice de Jaccard para cada par de aglomerados. El promedio para los índices es de 0,036227758 y este valor puede utilizarse como referencia para interpretar las intersecciones.

La primera observación es que las intersecciones no son muy numerosas. Por lo tanto, los datos reflejan que los

usuarios se agrupan por tópicos (homofilia basada en temas). En los dos solapamientos más numerosos (3 y 4) se encuentra el tópico “Campaña de promoción de pago”. Dado que es el tópico que ha recibido la mayor cantidad de “Me gusta”, resulta el más popular y en consecuencia influye en los solapamientos. El solapamiento 10 resulta más interesante dado que está conformado por dos tópicos no tan populares y se refieren a anuncios de las competencias centrales del negocio y los 12 usuarios en esa intersección podrían resultar influyentes. A partir de este análisis (de los usuarios en la intersección de los grupos), el administrador de la Fan Page puede identificar un conjunto de los usuarios activos que resultan más promisorios para incrementar la popularidad de la Fan Page. De esta forma, la empresa puede analizar si esos usuarios son clientes, empleados o potenciales clientes, y concentrar sus esfuerzos de marketing en un número reducido de individuos. En este trabajo no se listan los nombres de los usuarios dada la confidencialidad de la información.

Figura 2. Detalle del porcentaje de autor del post (Organización), géneros femenino y masculino por tópico.



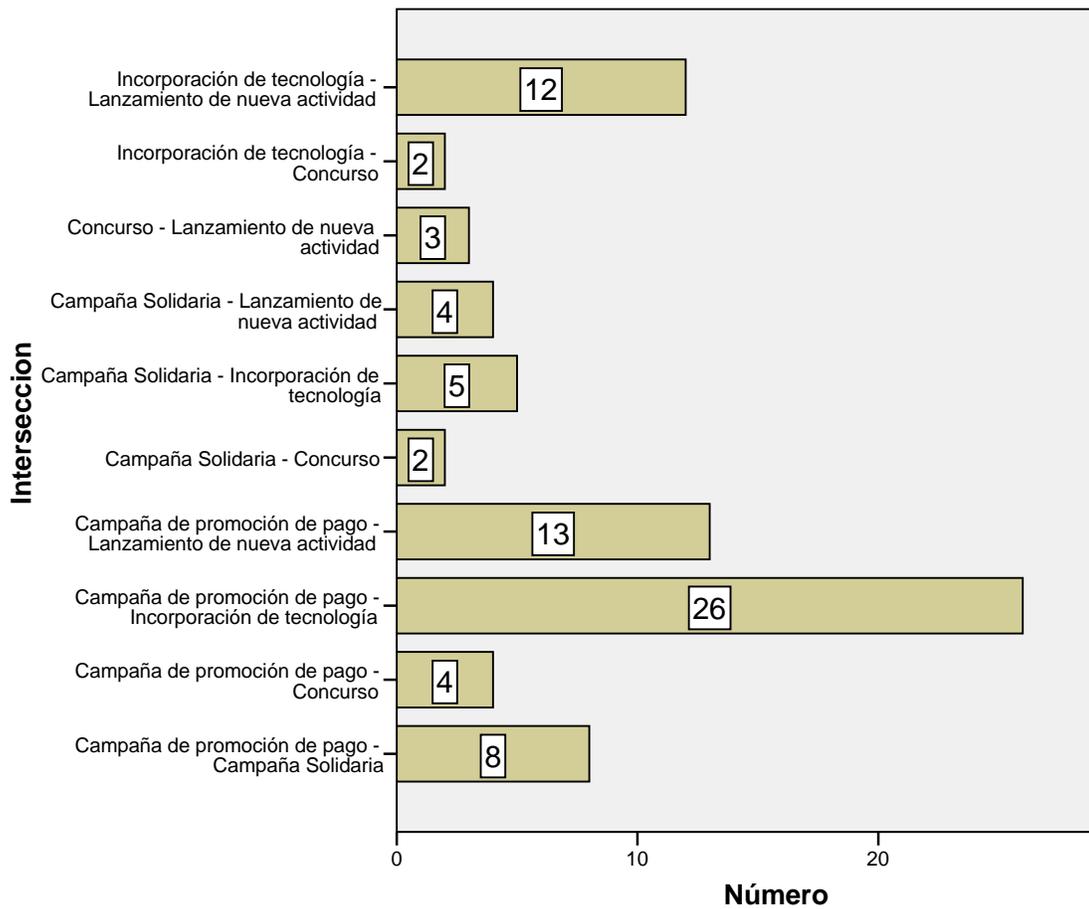
Fuente: elaboración propia.

Tabla 2. Solapamientos de usuarios en los diferentes tópicos.

#	Solapamientos	$J(A_n^k, A_n^l)$		
		Intersección		
1	Campaña de promoción de pago - Campaña Solidaria	Intersección	8	0,025477707
		Total	314	
2	Campaña de promoción de pago - Concurso	Intersección	4	0,010178117
		Total	393	
3	Campaña de promoción de pago - Incorporación de tecnología	Intersección	26	0,075144509
		Total	346	
4	Campaña de promoción de pago - Lanzamiento de nueva actividad	Intersección	13	0,043333333
		Total	300	
5	Campaña Solidaria - Concurso	Intersección	2	0,010362694
		Total	193	
6	Campaña Solidaria - Incorporación de tecnología	Intersección	5	0,034246575
		Total	146	
7	Campaña Solidaria - Lanzamiento de nueva actividad	Intersección	4	0,04
		Total	100	
8	Concurso - Lanzamiento de nueva actividad	Intersección	3	0,01754386
		Total	171	
9	Incorporación de tecnología - Concurso	Intersección	2	0,00921659
		Total	217	
10	Incorporación de tecnología - Lanzamiento de nueva actividad	Intersección	12	0,096774194
		Total	124	

Fuente: elaboración propia.

Figura 3. Detalle de los solapamientos.



Fuente: elaboración propia.

### Conglomerados basados en los valores de “Me gusta”, “Compartir”, “Comentar” y Percepción

Considerando la red conformada por todos los posts de interés, para cada usuario se computó la cantidad de “Me gusta”, “Comentar” y “Compartir”. Si bien uno de los aspectos más interesantes es la polaridad de los comentarios, en la Fan Page considerada, todos los usuarios realizan comentarios positivos, por lo tanto, no se utiliza la variable. En la Tabla 3 se incluyen algunos estadísticos descriptivos de los datos. Las Tablas 4 y 5

presentan los conglomerados donde se distinguen cuatro grupos caracterizados por un grupo numeroso con poca interacción (1); un grupo compuesto por 12 usuarios con un valor medio máximo para “Me gusta” pero no realizan comentarios (2); un grupo minoritario con el máximo valor promedio para “Comentar”; y el grupo mayoritario con un valor superior a la media para “Me gusta” (4). En general, se concluye observando que la mayoría participa indicando que les gusta un post, pero difícilmente realizan comentarios o comparten el mismo.

Tabla 3. Estadísticos descriptivos para las interacciones.

Interacción	Mínimo	Máximo	Media
“Me gusta”	0	5	1,6542
“Comentar”	0	4	0,0463
“Compartir”	0	0	0

Fuente: elaboración propia.

Tabla 4. Centros de los conglomerados finales. Se ha logrado la convergencia debido a que los centros de los conglomerados no presentan ningún cambio o este es pequeño.

Interacción	Conglomerado			
	1	2	3	4
“Me gusta”	0,96	4,42	1,00	2,08
“Comentar”	0,03	0,00	2,40	0,02
“Compartir”	0	0	0	0

Fuente: elaboración propia.

Tabla 5. Número de casos por conglomerado.

Conglomerado	Nro. de usuarios
1	193
2	12
3	5
4	244

Fuente: elaboración propia.

## CONCLUSIONES

Se han presentado las etapas para que una organización pueda determinar en forma sistemática un conjunto de usuarios influyentes en una red social. La metodología está basada en trabajos previos y para su elaboración se ha optado por una propuesta que pueda aplicarse utilizando herramientas sencillas y disponibles en forma gratuita. De esta forma, una organización que tenga en funcionamiento una Fan Page dispone de un método que le brinda información complementaria para mejorar la administración de la Fan Page y de su estrategia de marketing.

El ejemplo analizado pretende ilustrar cómo abordar cada uno de los pasos para identificar a los usuarios más influyentes y determinar si algunos tópicos favorecen la interacción. Si bien en el ejemplo no se consideran muchos posts y una ventana temporal extensa, se pudo observar un comportamiento similar a

otros estudios. Por ejemplo, los usuarios se agrupan de acuerdo con los tópicos, lo cual es un indicio de homofilia basada en los temas tratados. Tal como se describió en el caso, si bien los usuarios participan indicando “Me gusta” o comentando un post, no se observa una difusión viral de los posts. Incluso en la ventana temporal los usuarios no han compartido los posts. Un análisis basado en el texto y presentación de los posts permitió determinar que los mismos informan, pero no están personalizados. Este aspecto podría ser un indicio de la poca efectividad de los posts. De acuerdo con Dehghani y Turner (2015), la efectividad de la publicidad depende de que incluya características tales como participación, realización, personalización y retroalimentación. En base a un estudio los autores recomiendan que, con el objetivo de optimizar la publicidad en Facebook, las empresas deban proveer mensajes más personalizados y ofrecer incentivos y promociones para inducir a los clientes a que transmitan un mensaje. Considerando esto último

una recomendación natural para los administradores es incluir la personalización en los mensajes de los usuarios más influyentes. Luego, la efectividad de esta acción puede evaluarse observando si la participación y difusión de los posts se incrementa.

Durante la aplicación se presentaron algunas restricciones. Una de ellas fue que no fue posible

recuperar automáticamente las interacciones de “Compartir” un post. De todos modos, esta situación está dada por la herramienta NodeXL y puede superarse utilizando otra aplicación para procesar la red social. La propuesta se podría mejorar considerando en la definición de los clústers la popularidad de los usuarios. La misma puede definirse en base a la cantidad de amigos del usuario en la red.

## REFERENCIAS

- Canhoto, A., Clark, M., & Fennemore, P. (2013). Emerging segmentation practices in the age of the social customer. *Journal of Strategic Marketing*, 21(5), 413-428.
- Dehghani, M., & Tumer, M. (2015). A research on effectiveness of Facebook advertising on enhancing purchase intention of consumers. *Computers in Human Behavior*, 49, 597-600.
- Edwards, G. (2010). *Mixed-Method Approaches to Social Network Analysis*. ESRC National Centre for Research Methods.
- Hanneman, R., & Riddle, M. (2005). *Introduction to social network methods*. Riverside: University of California.
- Hansen, D., Shneiderman, B., & Smith, M. (2011). *Analyzing Social Media Networks with NodeXL*. Burlington: Morgan Kaufman.
- Haythornthwaite, C. (1996). Social Network analysis: An Approach and Technique for the Study of Information Exchange. *Library & Information Science Research*, 18(4), 323-342.
- Jucaityte, I., & Mašcinskiene, J. (2014). Peculiarities of social media integration into marketing communication. *Procedia - Social and Behavioral Sciences*, 156, 490-495.
- Khobzi, H., & Teimourpour, B. (2015). LCP segmentation: a framework for evaluation of user engagement in online social networks. *Computers in Human Behavior*, 50, 101-107.
- Majumdar, A., Saha, D., & Dasgupta, P. (2015). An analytical method to identify social ambassadors for a mobile service provider's brand page on Facebook. *Applications and Innovations in Mobile Computing (AIMoC)*. Kolkata: IEEE.
- Martínez-Cámara, E., Martín-Valdivia, M., López-Ureña, L., & Mitkov, R. (2014). Polarity classification for Spanish Tweets using the COST corpus. *Journal of Information Science*, 1-11.
- McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27, 415-444.
- Mochalova, A., & Nanopulos, A. (2014). A targeted approach to viral marketing. *Electronic Commerce Research and Applications*, 13, 283-294.
- Muruganatham, A., & Gandhi, M. (2005). Ranking the influence users in a social networking site using an improved Topsis method. *Journal of Theoretical and Applied Information Technology*, 73(1).
- Narayanan, M., Asur, S., Nair, A., Rao, S., Kaushik, A., Mehta, D., . . . Lalwani, R. (2012). Social Media and Business. *Vikalpa: The journal for Decision Makers*, 37(4), 69-111.
- Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Shin, S., Lee, K., & Hall, D. (2014). exploring Facebook Users' Continuous Visiting Behaviors: Conceptual Incorporation of Facebook User Perceptions toward Companies' Facebook Fan Page Usage. *Twentieth Americas Conference on Information Systems*. Savannah.
- Skeels, M., & Grudin, J. (2009). When social networks cross boundaries: a case study of workplace use of Facebook and LinkedIn. *Proceedings of the ACM 2009 International Conference on Supporting Group Work* (págs. 95-104). Florida: ACM.
- Weng, L., & Lento, T. (2014). Topic-based Clusters in Ego-centric Networks on Facebook. *AAAI Conference on Weblogs and Social Media*. Ann Arbor, Michigan: AAAI Press.

## NOTAS

---

<sup>i</sup> Research article, result of the research project "Analysis of the role of online social networks in the dissemination of information and knowledge", file number 24 / C042, subsidized by Secretaría de Ciencia y Tecnología, Universidad Nacional del Sur, Bahía Blanca, Argentina, [www.uns.edu.ar](http://www.uns.edu.ar). Project start date: January 2015, end date December 31, 2016. Filing date [07/07/2017], acceptance date [07/02/2018]

<sup>ii</sup> Doctora en Ciencias de la Computación, Profesora Asociada tiempo completo en Dpto. de Ciencias de la Administración, Universidad Nacional del Sur, Bahía Blanca, Argentina, [www.uns.edu.ar](http://www.uns.edu.ar). Intereses de investigación: El rol, el valor y la gobernanza de la digitalización en las organizaciones. Correo electrónico: [mas@uns.edu.ar](mailto:mas@uns.edu.ar)