

# Recuperación de Información y Recomendaciones en la Web

OBTENER LISTA DE AMIGOS OCULTA EN FACEBOOK



Grupo 11

José Ma. Aguerre - 4.585.342-0

Sebastián Caballero - 4.456.979-3

Ignacio Fagian - 4.374.178-2

Ismael Pisano - 4.311.553-3

# Contenido

1. Introducción
  - 1.1. Definición del problema
  - 1.2. Enfoque de la solución
  - 1.3. Facebook Graph Search
2. Implementación
  - 2.1. Algoritmo
  - 2.2. Módulos
  - 2.3. Cliente
  - 2.4. Servidor
  - 2.5. Herramientas
  - 2.6. Algoritmos Parecidos
3. Casos de prueba
4. Conclusiones
5. Mejoras al algoritmo
6. Enlaces

# Introducción

## Definición del problema

Facebook permite configurar la privacidad de una cuenta, en particular ocultar la lista de amigos y que solamente la puedan ver ciertos grupos de usuarios o ninguno. La propuesta consiste en obtener la lista amigos de un usuario que haya configurado su cuenta para no mostrar su lista de amigos.

## Enfoque de la solución

Utilizando una funcionalidad que brinda Facebook llamada “graph search”, es posible consultar los amigos en común entre dos personas. Esto nos permite realizar de manera recursiva la misma búsqueda, teniendo una condición adecuada de parada, hasta llegar a obtener un listado lo más acercado posible a la totalidad de sus amigos.

Para eso es necesario dar el nombre de usuario de la persona y una consulta inicial. Un ejemplo de consulta sería “Estudiantes de Facultad de Ingeniería”. De esta manera, se ejecutaría una consulta entre el usuario inicial, del cual se desea obtener su lista de amigos, y cada uno de los usuarios que componen el grupo “Estudiantes de Facultad de Ingeniería” que tengan su lista de amigos pública. Cada una de estas consultas devolverá los amigos en común que tiene esta persona con los del resultado obtenido.

Se repite el mismo procedimiento con los resultados obtenidos de cada consulta y modificando la lista con las nuevas personas que tengan su lista de amigos pública.

## Facebook Graph Search

Es el motor de búsqueda de facebook lanzado en Marzo del 2013, sólomente se encuentra disponible en el idioma inglés. Es un motor de búsqueda semántica que permite realizar búsquedas en lenguaje natural como “Fotos de Mark Zuckerberg en Facultad de Ingeniería” o “Amigos que le gustan La Vela Puerca y Abuela Coca”.

# Implementación

Para la implementación de la solución utilizamos Java ya que existen una amplia variedad de herramientas disponibles que facilitan el trabajo de publicación de servicios, parseo de páginas, etc. Se construyó una página web para realizar la interacción con el algoritmo, la misma contiene dos campos donde se ingresarán el usuario víctima y el o los usuarios semilla y un botón para comenzar. Luego de ingresados los datos la página web envía la consulta al algoritmo y luego cada 1 segundo realiza otra petición para mostrar el progreso del mismo e ir desplegando los resultados.

Buscador de amigos ocultos de Facebook

Victima:

Semilla:

Start Cantidad: 331

Lucia Cuello ID: abigarcia22	Tambores Panizza ID: 100007696544721	Martin Alvarez Rodriguez ID: Elpeti	Tec Conectados ID: 100002342026059
Rocio Caldevilla ID: rocio.caldevilla	Ana Sosa ID: Anaa89	Marcos Müller ID: mullerrizzo	Pau Rizzo Ferrón ID: paularizzo17
María Eugenia Durán ID: maria.e.duran.33	Valeria Andrade Lallana ID: valeria.andrade.l	Pablo Ibañez ID: 100004076174598	Santiago Franco ID: santiago.franco.5015
Cecilia Miguel ID: cecilia.miguel.731	Maxi Figueroa ID: ifati.figueroa	Rodney Gonzalez ID: RDYTORKE	ArteSano Rastas-Dreadlocks ID: DreadlocksUruguay
Matias Jarovisky ID: matias.jarovisky	Cabeza De Martillo ID: cabezade.martillo.9	Christian Ramirez ID: christian.ramirez.77964201	Carla Ferrari ID: carla.ferrari.73744
Gabriel de Leon ID: gabrieldemel	Silvia Borges ID: silvia.borges.5623	Taliana Vargas ID: taliana.vargas.756	Claraa Chirico ID: clarona2003
Karen Ferreira ID: karen.ferreira.547727	Georgina Santana Guadalupe ID: georgina.santanaguadalupe	Lilian Ferreira ID: lilian.ferreira.779205	Belú Arte En Cuero ID: anabelen.araujo

## Algoritmo

El algoritmo comienza con un usuario víctima (usuario con amigos ocultos) y uno o varios usuarios semilla. Es necesario que el usuario semilla tenga al menos un amigo en común con la víctima y tenga su lista de amigos pública.

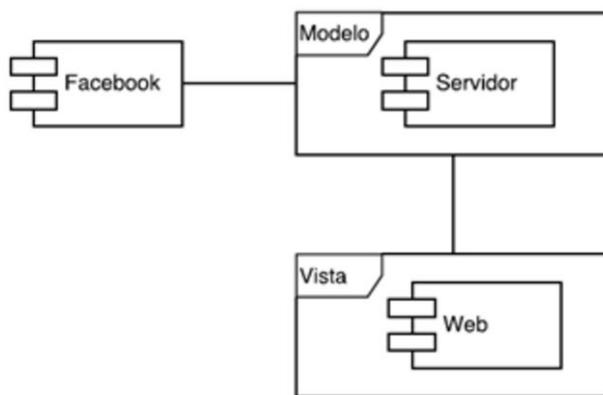
Como se deben poder hacer búsquedas dentro de Facebook es necesario iniciar sesión con una cuenta cualquiera.

Se crea una lista con un único elemento que será el usuario semilla. Se itera sobre los usuarios de la lista y se buscan los amigos en común con el usuario víctima, estos amigos en común se guardan como parte de la solución y se agregan a la lista como futuras semillas.

Este algoritmo termina cuando se vacíe la lista, esto quiere decir que o bien se encontraron todos los amigos de la víctima, o bien tiene otros conjuntos de amigos que no comparten ningún amigo con el conjunto de la semilla. Existen otra forma para que se termine el algoritmo y es una condición de parada programada, se basa en limitar las iteraciones que realiza el algoritmo.

## Módulos

Se plantea una aplicación cliente-servidor.



### Cliente:

Permitirá a un usuario utilizar el algoritmo. Para ello ingresará el nombre de usuario de Facebook, del cual se desea descubrir su lista de amigos y la semilla inicial de donde se arrancará a iterar.

### Servidor:

Será el encargado de la ejecución del algoritmo y brindará los resultados al cliente.

## Herramientas

Utilizamos Selenium web driver con HtmlUnit como navegador web para automatizar el login de la cuenta personal con Facebook, y así poder obtener las cookies de la sesión y utilizarlas para realizar las futuras peticiones.

Luego se utiliza la herramienta Jsoup para la extracción y trabajo con el contenido de consulta con el fin de recabar la información en cada iteración.

## Algoritmos Parecidos

Encontramos el siguiente algoritmo hecho en python que resuelve la parte de la búsqueda principal del problema pero no finaliza el algoritmo planteado ya que no itera entre los resultados obtenidos.

Enlace: <https://github.com/prili/fb-hfc>

## Casos de prueba

Ejecutamos el algoritmo con varios usuarios y fuimos registrando y comparando la cantidad de amigos devuelta con la cantidad de amigos reales del usuario. Seteamos la cantidad máxima de iteraciones en 100.000.

Los resultados fueron los siguientes:

Usuario	Total real	Total encontrados	Eficiencia
cachitblues	746	331	44.36%
sebastian.caballero.560	173	150	86.70%
mauro.lozov	1034	727	70.31%
veronica.pose.3	759	343	45.19%
felipe.pisano	556	239	42.99%
analialopez.sch	256	79	30.85%
ignacio.fagian.3	275	95	34.54%
agustin.peraza.3	1867	773	41.40%
germanferayala	234	56	23.93%

**Promedio de eficiencia: 46,70%**

**Desviación standard: 19,76**

# Conclusiones

Realizamos las mismas pruebas con distintas semillas iniciales pero los resultados eran prácticamente los mismos por lo que podemos inferir que a pesar de que se tome una semilla con muchos amigos en común con la víctima, el resultado será prácticamente el mismo que tomando una semilla con 1 solo amigo en común.

A partir de los datos obtenidos con las pruebas no podemos afirmar la efectividad del algoritmo dependiendo de la cantidad de amigos que el usuario tenga como era de esperar en un principio. Esto se puede deber a varios factores, el primero es que como facebook muestra los amigos en pedidos asincrónicos de 36 usuarios y nuestra implementación del algoritmo no toma en cuenta si existen más que esa cantidad ya que para ello deberíamos posicionarnos en la parte inferior de la página y así lograr que se efectúe otro pedido. Otro factor podría ser la disjunción de los grupos de amigos de facebook, por ejemplo, si la víctima tiene un grupo de amigos que no comparten ningún amigo con otro grupo, dependiendo de la semilla que se tome, uno de los grupos no aparecerá en los resultados. Un nivel de privacidad que ofrece facebook es el de mostrar la lista de amigos sólo a tus amigos, por lo que puede influenciar en los resultados si el usuario logueado con el que se hacen las consultas es amigo del usuario con la que se está iterando.

A pesar de que esperamos en un principio mejores resultados, podemos concluir que este algoritmo no es tan bueno para obtener la lista de amigos ocultos de un usuario.

## Mejoras al algoritmo

- Debido al factor explicado anteriormente sobre las peticiones asincrónicas, una mejora sería lograr ese desplazamiento hasta el fin de la página y así obtener mayor cantidad de usuarios.
- Verificación previa de configuración de privacidad: Si se llega a un usuario que tiene los amigos ocultos, éste retornará una lista vacía al compararlo con los amigos de la víctima. Se podría verificar previo a esa llamada si tiene sus amigos ocultos y así evitar el reingreso al algoritmo.
- Graph search: El algoritmo implementado utiliza algunas consultas básicas sobre el motor Graph search, como la obtención de amigos en común entre dos usuarios. Se podría explotar más esta herramienta por ejemplo utilizando como semilla una consulta del estilo “Estudiantes de Facultad de Ingeniería”.

# Enlaces

- <http://jsoup.org/>
- <http://www.seleniumhq.org/>
- <http://htmlunit.sourceforge.net/>
- <http://search.fb.com/>