

## **Análisis de redes sociales aplicado a redes de investigación en ciencia y tecnología**

*Luis A. Navarro Sánchez*

Universidad Austral de Chile, Valdivia, Chile. Egresado de Ingeniería Civil en Informática  
e-mail: [luisns@gmail.com](mailto:luisns@gmail.com)

*Juan P. Salazar Fernández*

Universidad Austral de Chile, Valdivia, Chile. Ingeniero Civil en Informática  
e-mail: [juansalazar@uach.cl](mailto:juansalazar@uach.cl)

El presente documento expone un sistema que toma los currículos de investigadores en ciencia y tecnología de la plataforma SICTI para generar archivos de redes sociales, los que sirven como entrada para software de análisis y visualización de redes sociales (Pajek y NetDraw). Primeramente, se exponen conceptos básicos de análisis de redes sociales y en qué consiste la plataforma SICTI. Luego se describe los software Pajek y NetDraw y se justifica la elección de usarlos. A continuación se presenta la información útil en los currículos para generar redes y un modelo de análisis apropiado. Luego se explica el análisis, diseño e implementación del sistema que genera las redes. Finalmente, se realiza un análisis sobre la red completa de investigadores y se presentan las conclusiones.

*Palabras clave:* análisis de redes sociales, SNA, SICTI, redes de investigación científica, currículos de investigadores, Pajek, NetDraw.

The present document expounds a system that takes science and technology researchers curriculums from the SICTI platform in order to generate social networks files, which are used as input to social networks analysis and visualization software (Pajek y NetDraw). First, basic concepts from social networks analysis theory are introduced and what the SICTI platform is about. Next, Pajek and NetDraw software are described and the choice of using them is justified. Next, useful information from the curriculums is laid out, which is useful to generate networks and an appropriated analysis model. Next, analysis, design and implementation of the system which generates networks is explained. Finally, an analysis of the entire researchers' networks is performed and conclusions are presented.

*Key words:* social networks analysis, SNA, SICTI, scientific research networks, researchers curriculums, Pajek, NetDraw.

## 1. INTRODUCCIÓN

El análisis de redes sociales o SNA (del inglés Social Networks Analysis) es un área del conocimiento que es usada para estudiar cómo se relacionan las personas u otras entidades como instituciones, empresas, etc. La plataforma SICTI permite a los investigadores en ciencia y tecnología almacenar sus currículos, postular y guardar información sobre sus proyectos. Actualmente SICTI no cuenta con una herramienta que permita ver las relaciones de colaboración entre las personas registradas en el sistema o entre las instituciones o disciplinas. Aquí entra en juego el análisis de redes sociales, el cual nos permite estudiar estas relaciones que de otra forma estarían ocultas. Para generar las redes se trabaja con los currículos de los investigadores, pero para llegar a generar una red se requiere la implementación de un sistema que se encargue de la obtención de los currículos, su procesamiento para extraer la información relevante y generación de las redes en base a una serie de parámetros proporcionados. Este sistema genera archivos de redes que pueden ser leídos por software de SNA y así poder realizar una serie de análisis y visualizar las redes formadas. La información resultante puede servir, por ejemplo, para encontrar las personas o instituciones más influyentes, y descubrir fortalezas y debilidades en las redes.

## 2. CONCEPTOS

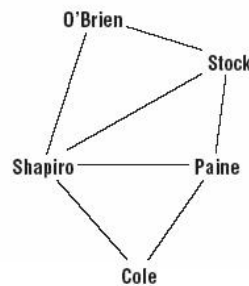
### 2.1. *Análisis de redes sociales*

El análisis de redes es el estudio de las relaciones y flujos entre actores tales como personas, grupos, organizaciones u otras entidades procesadoras de información y/o conocimiento [1]. El análisis de redes sociales provee un análisis matemático y visual de sistemas humanos complejos [2].

El análisis de redes sociales tiene múltiples aplicaciones en los más diversos campos, es usado para mejorar la eficiencia de la estructura organizacional de las empresas [3], desarrollo exitoso de nuevos productos, análisis de transmisión de enfermedades, propagación de virus informáticos, análisis del comportamiento de comunidades, estudio del capital social, entre otros. Gráficamente una red social se representa mediante grafos. Los nodos del grafo corresponden a los actores; y los vínculos entre éstos, muestran las relaciones o flujos entre los nodos. A los vínculos generalmente se le asigna un peso que es la fuerza o importancia de la relación.

En la siguiente figura se aprecia una red social:

Figura 2.1.1.  
Red Social



Pueden ser encuestas, tipos de observación directa, registros escritos y experimentos. Se pueden generar distintos tipos de redes que permiten estudiar interacciones a distintos niveles [3]: redes de confianza, redes de colaboración, redes de comunicación, entre otras; con esto es posible analizar los tipos de relaciones que puedan existir entre los actores de la red, como por ejemplo: sus lazos afectivos, sus roles sociales y flujo de información entre éstos.

Existe una serie de métricas para entender las redes y sus actores [4]. Estas medidas ayudan a determinar la importancia y el rol de un actor en la red. Las más usadas se clasifican en métricas de centralidad y poder, y métricas de grupos.

### 2.1.1. *Centralidad y poder*

El poder de un actor es la dependencia que tienen los demás actores sobre éste. El poder se puede medir en términos del concepto de centralidad. La centralidad es una forma de medir el poder, se refiere a que tan cerca un actor está del centro de una red, es decir de las posiciones que pueden dar mayor dominancia e influencia.

Las métricas más usadas son:

- 1) *Grado de centralidad (o grado)*: es el número de conexiones directas (distancia igual a 1) que tiene un actor con los demás.
- 2) *Cercanía*: indica, como un promedio, qué tan cercano está un actor al resto de los actores en una red.
- 3) *Intermediación*: se mide como la proporción de veces que un actor está en la ruta entre distintos pares de actores.

### 2.1.2. *Grupos*

El número, tamaño y conexiones entre los grupos o subgrupos pueden explicar la conducta de la red como un todo. Se puede observar las conductas de los actores en función de su ubicación en los subgrupos. Éstos pueden actuar de puentes entre grupos, ser actores aislados o actores cosmopolitas.

Las métricas más usadas son:

- a) *Clique*: es un subgrupo de una red en que sus actores están más cercanos y unidos que el resto de los actores de la red. Formalmente, un clique es un subgrupo de actores en el cual están presentes todos los vínculos posibles entre éstos.
- b) *N-clique*: es un subgrupo en el que todos los actores están separados a lo más por una distancia de N de los demás actores.
- c) *Componentes*: grupo en que todos sus miembros están conectados con al menos otro miembro.
- d) *Puntos de corte*: son nodos que, si se remueven, dividen la red en dos o más componentes.

### 2.1.3. *Plataforma SICTI*

SICTI (*Sistema de Información en Ciencia, Tecnología e Innovación*) es una plataforma que permite administrar información de proyectos, y la producción generada a partir de ellos: información de instituciones, currículos, publicaciones y patentes.

Nacido de un concurso lanzado desde el Programa Fondef de CONICYT (Comisión Nacional de Investigación Científica y Tecnológica de Chile), este proyecto es ejecutado con fondos provenientes del Programa de Desarrollo e Innovación Tecnológica (PDIT) del Ministerio de

Economía, financiado por el Banco Interamericano de Desarrollo y con el aporte de la empresa privada. Como plataforma de información integrada, SICTI permite saber quién hace ciencia en Chile, compartir estándares, elaborar indicadores válidos para todo el Sistema Científico Nacional, acceder a currículos en línea de todos los investigadores chilenos y extranjeros y construir e implementar nuevas herramientas de formulación, evaluación y gestión de proyectos.

Por otra parte, SICTI no sólo ofrece herramientas para el intercambio y extracción de información, sino que también, contempla un diseño modular y transferible para que las instituciones y empresas que desarrollan investigación e innovación puedan adoptar uno o más sistemas a fin de crear una red interoperable y entregar respuestas a las necesidades locales de cada una de ellas.

SICTI está conformado por tres módulos base:

- a) *Sistema de Currículo*: sistema actualizado en línea, enlazado con el resto de las bases de datos del programa, disponible para investigadores, estudiantes, gestores y emprendedores de la comunidad científica y tecnológica. Este sistema contiene información de currículos, los que pueden ser completados por los usuarios en Internet, y sirve de base para ofrecer información pública sobre un área central de ciencia y tecnología.
- b) *Sistema de Proyectos*: sistema en donde se presenta, completa y mantiene la información asociada a los proyectos, como: identificación del proyecto, resumen, financiamiento, participantes, disciplinas, sector de aplicación, producción.
- c) *Sistema de Instituciones*: sistema que permite mantener la información actualizada de datos asociados de las instituciones de la CTIT.

SICTI está funcionando activamente, tiene todos sus módulos ya implementados. Actualmente cuenta con más de 11.000 currículos activos. Es usado principalmente por investigadores, científicos y profesionales de la innovación para almacenar sus currículos, y para acceder y participar en concursos y fondos en línea. Se espera que siga creciendo el número de investigadores junto con la información disponible de currículos, proyectos y fondos para la ejecución de proyectos.

#### 2.1.4. *Software de análisis y visualización de redes sociales*

Existe una gran cantidad de software para análisis de redes sociales. Se evaluaron varios de éstos en base a los siguientes parámetros:

“disponible”).

- a) *Dis. Disponibilidad del software*: gratis, pagado o *shareware*. Los valores están en USD.
- b) *Int. Interfaz*: claridad de la interfaz de usuario.
- c) *Mét. Métricas*: cantidad y facilidad de uso de las métricas y sus resultados
- d) *Imp./Exp. Importar/Exportar*: capacidad para importar o exportar otros formatos de archivo, ya sea de otros software de análisis, Excel u otros.
- e) *Gráf. Gráfica*. Capacidad para graficar redes.
- f) *Docs. Documentación*: disponibilidad y claridad de la documentación.
- g) *Actual*: frecuencia de actualización del software y nuevas versiones recientes.
- h) *Total*: promedio aritmético de los parámetros (no se incluye parámetro)

Tabla 2.1.4.1.  
Evaluación de software de SNA

Nombre	Disp	Int	Mét	Imp. /Exp.	Gráf	Docs	Actual	Total
<i>Anthropac</i>	Desde 30	3	4	2	3	5	3	3.33
<i>Blanche</i>	Gratis	4	3	4	4	5	4	4
<i>FATCAT</i>	Gratis	3	3	2	2	4	2	2.66
<i>Iknow</i>	Gratis	4	3	2	2	4	3	3
<i>Inflow</i>	-	4	4	*	3	*	4	3.75
<i>Krackplot</i>	Gratis	3	3	3	3	5	2	3.16
<i>MultiNet</i>	-	4	3	3	3	4	4	3.5
<i>Negopy</i>	Shareware	1	3	2	1	5	3	2.5
<i>Netdraw</i>	Gratis	5	2	5	5	3	5	4.16
<i>NetMiner</i>	Desde 35	5	5	5	5	5	5	5
<i>NetViz</i>	Gratis	4	4	3	4	4	4	3.83
<i>Pajek</i>	Gratis	4	5	4	4	4	5	4.33
<i>Stocnet</i>	Gratis	4	3	4	2	5	5	3.83
<i>UCINET</i>	Desde 40	5	5	5	5	5 <sup>†</sup>	5	5
<i>Visone</i>	Gratis	4	4	4	4	3	4	3.83

### 2.1.5. Elección de los software de análisis

Como se ve en la tabla 2.1.4.1. los software mejor puntuados resultaron ser UCINET, NetMiner, Pajek y Netdraw.

Para decidir qué software ocupar se usó el parámetro “disponibilidad del software”. NetMiner y UCINET requieren licencia pagada, por lo que se descartan. Es importante mencionar que a pesar de que UCINET y NetMiner tienen un precio mínimo similar, NetMiner no ofrece todas sus funcionalidades con la licencia de USD 35.

Como Pajek y NetDraw son gratuitos, éstos serán los software a utilizar en la tesis. Pajek se utilizará para realizar los análisis y Netdraw, para graficar las redes. Pajek [5] puede manejar cómodamente redes grandes (cientos de miles de vértices), posee una gran cantidad de algoritmos para análisis de redes y puede exportar las redes a una gran variedad de formatos.

NetDraw [6] también posee algoritmos para análisis, pero su fuerte está en la visualización de redes, permite fácilmente manipular una red y tiene variadas opciones para visualización.

## 3. ANÁLISIS DE LOS CURRÍCULOS

Cada currículum se obtiene desde el SICTI en formato XML. Cada uno es identificado mediante un número entero único.

\* No hay datos sobre el parámetro.

† UCINET no grafica por sí solo, sino que llama a NetDraw, por lo que se le otorgó la misma puntuación que a NetDraw.

El currículo tiene una estructura jerárquica que se compone de las varias secciones y sub-secciones. Dentro de cada sub-sección hay una serie de campos con datos. Se pretende identificar cuáles son los más relevantes para realizar un análisis de redes sociales sobre redes generadas a partir de la información contenida en éstos.

Para crear redes sociales se necesita información que permita vincular actores. Se pueden distinguir dos formas de crear vínculos entre actores: mediante asociaciones o similitudes.

### 3.1. Asociaciones

Una asociación se refiere a las relaciones que implican algún tipo de colaboración entre personas, por ejemplo: se crea una asociación al colaborar con otras personas en una publicación de un libro, un artículo en una revista, un acta de congreso, etc. Una asociación tiene asociada un peso que depende del criterio de la persona que va analizar la red; es decir, la importancia que se le da a la relación. Las asociaciones se obtienen de las siguientes secciones del currículo:

- a) *Producción*: se manifiesta como una lista de nombres de personas en el campo de autores de la publicación, esto indica que estas personas son coautores de alguna publicación, por ejemplo: en un libro, tesis, actas de congreso, etc.

Tabla 3.1.1.  
Asociaciones en Producción

Nombre sección	Nombre campo
Libro	AutoresLibro
ArtículoRevistasDiarios	AutoresArticulo
ActasCongr	Autores
ProduccionArtisticaCultural	AutoresProduccionArtistica
OtraProduccionBibliografica	AutoresOtraProduccion
CapituloLibro	AutoresCapitulo
CapituloLibro	AutoresLibro
ProduccionTecnica	AutoresProduccionTecnica
TesisTitulosGrados	AutoresTesis
TesisTítulosGrados	ProfesorTutor
TesisTitulosGrados	OtroProfesorTutor

En “TesisTitulosGrados” también se puede obtener los profesores con los trabajó una persona mediante el campo “ProfesorTutor” y “OtroProfesorTutor”.

- b) *Experiencia laboral y profesional*: en las sub-secciones relativas a las direcciones de tesis de pregrado, magíster y doctorado, se puede obtener el nombre del tesistas o los tesistas que una persona tuvo a su cargo como profesor tutor, mediante el campo “nombre tesista”.

Tabla 3.1.2.  
Asociaciones en Dirección de Tesis

Nombre sección	Nombre campo
DirPregradoTesis	NomTesista
DirMagisterTesis	NomTesista
DirDoctoradoTesis	NomTesista

- c) *Propiedad intelectual*: aquí se puede obtener relaciones de asociación con los nombres de las personas con las cuales se han creado distintos tipos de propiedad intelectual. Los campos que sirven son:

Tabla 3.1.3.  
Asociaciones en Propiedad Intelectual

Nombre sección	Nombre campo
Patente	NombreAutorPatente
DerechoAutor	NombreAutorRegistro
OtraPropiedadIndustrial	NombreAutorPropiedadIndustrial

### 3.2. Similitudes

Una similitud es una relación que se crea sólo por el hecho de compartir atributos comunes, por ejemplo: misma comuna de residencia, misma institución de estudio, mismo lugar de trabajo, etc.

Una similitud puede generarse a partir de cualquier campo del currículo. Lo ideal es que sea un campo que pueda ser llenado por la mayoría de las personas y cuya información se elija desde un conjunto de valores definidos; y de este modo, hacer fácilmente las comparaciones. Por ejemplo: Nombre de universidades, profesiones, regiones de Chile, etc.

Las similitudes se pueden considerar como ponderadores de las relaciones de asociación, en cuyo caso, el sistema sólo debe crear una similitud entre dos actores, si existe antes una asociación.

Al igual que las asociaciones, las similitudes tienen un peso asociado.

### 3.3. Limitaciones de la información

- 1) *se confía en lo que escribe*: toda la información del currículo es ingresada por el usuario. No se puede verificar la validez de la información o posibles errores.
- 2) *información omitida*: no siempre el usuario provee toda la información que se desea analizar, lo que hace que en algunos campos haya poca información.
- 3) *no hay identificador único*: un nombre representando a una entidad puede estar escrito de distintas formas, lo que hace que se requiera procesamiento adicional para identificarlo correctamente.
- 4) *nombres repetidos*: no hay una forma única de identificar a una persona, y si dos o más personas se llaman se podría dar el caso de confundirlas.

## 4. MODELO DE ANALISIS

En este apartado se definirá un modelo de análisis de redes sociales en base al análisis de los currículos del apartado anterior y la revisión de estudios de redes similares. Se decidirá qué tipos de redes analizar y qué tipos de métricas usar.

Se escogió dos formas de generar las redes: Red Filtro y Red *Snowball*.

#### 4.1. *Red Filtro*

Una red generada mediante filtros la compone un subconjunto de personas que tienen atributos y relaciones en común. Mediante este método se genera una red compuesta de dos tipos de personas:

- a) El primer tipo son personas que cumplen con los filtros, es decir, sus atributos son iguales a los que definen los filtros.
- b) El segundo tipo, son actores que no cumplen con los filtros, pero tienen vínculos directos con los del primer tipo.

Por ejemplo, una red puede definir un filtro que defina disciplina igual a biología, de esta manera la forma estaría compuesta de personas cuya disciplina es biología más las personas que están directamente relacionados con éstas.

#### 4.2. *Red Snowball*

Mediante el método de *snowball* (bola de nieve), la red se genera a partir de un conjunto de personas iniciales; es decir, se buscan las relaciones directas a otras personas con estas personas iniciales y luego, a estas personas encontradas, también se les buscan las relaciones con otras personas y así sucesivamente hasta llegar a un límite de iteraciones, que es la profundidad o tamaño de la red, o que no hayan más personas que recorrer.

Las personas que están más alejadas de los actores iniciales se consideran menos importantes en relación a la influencia que pueden tener sobre éstos. Por eso se considera conveniente implementar un algoritmo de “poda” que permita eliminar relaciones más alejadas, si así se desea.

#### 4.3. *Métricas*

Las métricas que se ocuparán sobre las redes generadas son las mismas ya explicadas. Además se consideran las siguientes métricas:

- a) *distancia promedio en la red*: promedio de las distancias más cortas entre pares de nodos.
- b) *número de pares no alcanzables*: el número de pares de actores entre los cuales no existen un camino para llegar desde uno hasta el otro. Un número alto indica falta de conexiones en la red.
- c) *diámetro de la red*: es la distancia más larga en la red, tomada de las distancias más cortas entre cada par de nodos en los que existen un camino.

#### 4.4. *Tipos de análisis*

Se presentan los tipos de análisis que se realizarán:

##### 4.4.1. *Red completa*

Se incluyen todas las personas presentes en el módulo Currículo del SICTI. Este nivel de análisis provee un cuadro completo de la red de investigadores. Es muy útil para ver cuáles son las áreas más importantes de investigación, en las que hay más y menos comunicación o por cuáles atributos se tienden a agrupar las personas.



#### 4.4.2. *Subconjunto de red*

Usando uno o varios filtros se obtiene un subconjunto de la red completa, como por ejemplo de una universidad o área en particular. Es útil para acotar el área de análisis si se quiere analizar un conjunto específico de actores.

#### 4.4.3. *Redes individuales*

Aquí se genera una red *snowball*. Es útil si se quiere analizar el entorno que rodea a una o varias personas; para ver cómo están influenciadas e influyen a personas con las que están relacionadas o que quizá no conocen. Se puede ver cómo se relacionan esas personas iniciales con personas de otras instituciones u otras disciplinas.

Debido a la forma en que se genera la red, la utilidad de las medidas de centralidad cambia, ya que los actores iniciales siempre tendrán altos valores en las métricas de poder.

#### 4.4.4. *Redes de nivel superior*

En este tipo de redes, los nodos son grupos de personas. En base a los atributos de las personas se pueden armar redes en que cada nodo es el valor de un atributo y el valor de un vínculo entre dos atributos es la suma de los valores de los vínculos de las personas que tienen esos atributos. De esta forma se puede ver cómo se relacionan los atributos. Por ejemplo: se puede observar las relaciones entre instituciones o entre disciplinas del conocimiento.

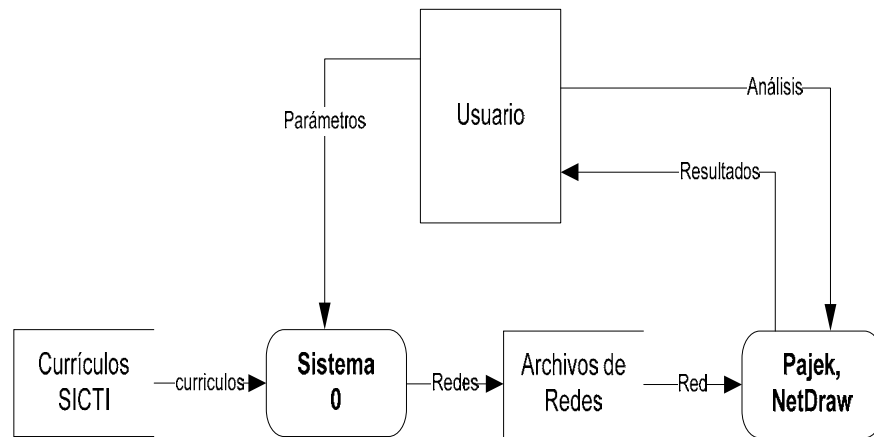
## 5. ANALISIS, DISEÑO E IMPLEMENTACIÓN DEL SISTEMA

Como se vio en apartados anteriores; por un lado, se tiene a SICTI que provee el módulo Currículo y con este módulo se tiene la red de investigadores; por otro lado, para realizar el análisis de redes sociales se eligió a Pajek y NetDraw. Entonces lo que falta es una herramienta que permita tomar toda la información disponible en el módulo Currículo, y la transforme de modo que pueda ser leída por los software para análisis de redes sociales, para luego realizar los análisis apropiados a las redes generadas aplicando el modelo de análisis creado. La creación de esta herramienta se verá en este apartado.

### 5.1. *Requisitos*

Se necesita un sistema que tenga como entrada currículos de investigadores más una serie de parámetros; y que como salida genere una o varias redes sociales que serán leídas por Pajek y NetDraw, con el cual se podrán realizar análisis a las redes creadas. Esto se visualiza en el siguiente diagrama de contexto:

Tabla 5.1.1.  
DFD Nivel 0



*Los requisitos más importantes del sistema son:*

- a) Definir una serie de filtros para una Red Filtro
- b) Definir una serie de actores para una red *snowball*.
- c) Definir las asociaciones que componen la red con sus respectivos pesos; definir las similitudes y sus respectivos pesos.
- d) Definir profundidad de la red para red *snowball*
- e) Generar una red por cada relación definida.
- f) Generar una red igual a la suma de todas las relaciones
- g) Generar redes de nivel superior; todas las redes se generan en formato Pajek y NetDraw.
- h) Traer los currículos remotamente desde SICTI y almacenarlos localmente.
- i) Crear una base de datos relacionar para leer la eficientemente la información de los currículos.

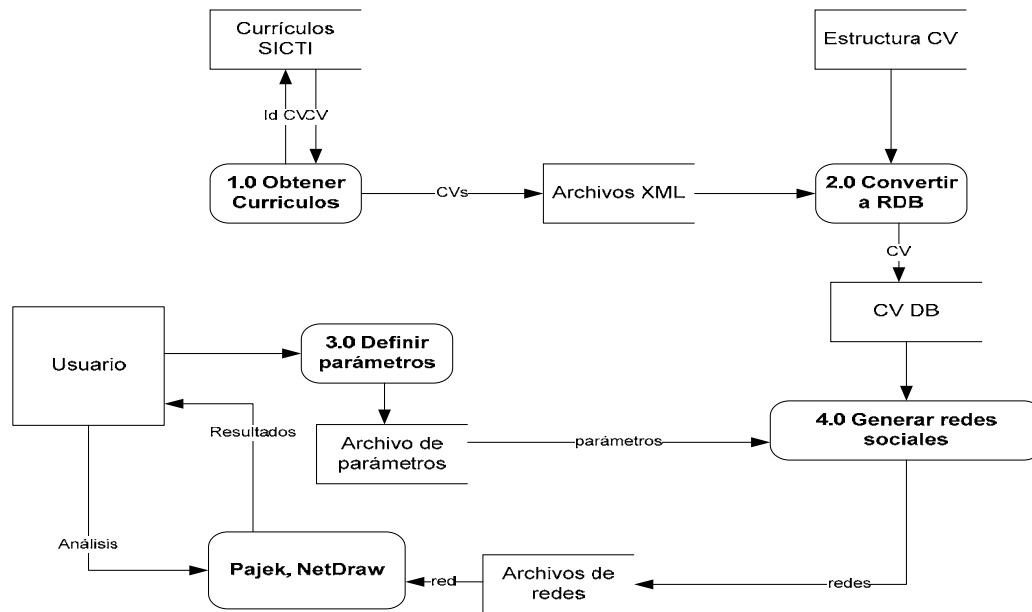
*Los requisitos no funcionales son:*

- a) El sistema debe ser simple de usar por alguien no experto en informática
- b) El sistema debe poseer mínimo impacto de instalación.

## 5.2. Análisis

Como se observa en la figura 0-1, el sistema completo consta de cuatro procesos principales. Hay dos procesos para procesar la entrada (1.0 y 2.0), un proceso que genera las redes (4.0) para ser leídas por Pajek o NetDraw y un proceso que interactúa directamente con el usuario para definir los parámetros de entrada (3.0).

Figura 5.2.1.  
DFD Nivel 1 Sistema completo



### 5.3. Proceso 1.0: Obtener currículos

Este proceso se comunica con un Web service y se encarga de obtener uno por uno, los currículos (CV) del SICTI, cada currículum es obtenido en formato XML mediante un identificador único (*id CV*), y es guardado en un directorio local (*Archivos XML*).

#### 5.3.1. Proceso 2.0: Convertir a base de datos relacional

Este proceso toma todos los currículos en formato XML guardados localmente en *Archivos XML* y crea una base de datos relacional (*CV DB*) basándose en la estructura definida en el archivo *Estructura CV*.

Cada sub-sección del currículum queda representada como una tabla. La razón de crear una base de datos relacional se debe a motivos de eficiencia de ejecución y comodidad en la programación.

#### 5.3.2. Proceso 3.0: Ingresar parámetros

Este proceso interactúa directamente con el usuario que quiere definir redes a analizar. Consiste en una interfaz gráfica, donde se ingresan todos los parámetros para generar una red. Estos parámetros se validan y se escriben en un archivo (*Archivo de parámetros*).

### 5.3.3. Proceso 4.0: Generar redes sociales

Este proceso toma como entrada el archivo de parámetros y la base de datos *CV DB*, para generar una serie de archivos de redes sociales en varios formatos, los que posteriormente sirven de entradas para software de SNA.

#### 5.3.3.1. Formatos de salida

Las redes se guardan en tres formatos distintos que corresponden a los programas para SNA. Estos son Pajek, NetDraw y UCINET.

#### 5.3.3.2. Redes generadas

Todas las redes generadas usan el parámetro “nombre de la red” como prefijo para el nombre de archivo de red. Para explicar cada red se asumirá que el valor para “nombre de la red” es *NOMBRERED*. Las redes se guardan en tres carpetas distintas llamadas Pajek, UCINET, NetDraw dependiendo del formato de salida.

- a) *NOMBRERED*. Archivo de red múltiple. Contiene una red por cada relación definida en el archivo de parámetros. Cada nodo en la red es una persona.
- b) *NOMBRERED\_total*. Contiene una red que es la suma de todas las relaciones definidas en el archivo de parámetros. Es decir, el valor de un vínculo entre dos actores es igual a la suma de los valores de ese vínculo en cada relación.
- c) *NOMBRERED\_seccion-atributo*. Contiene una red que corresponde únicamente a la relación seccion-atributo. Por lo tanto, existe una red de este tipo por cada relación definida en el archivo de parámetros. La red *NOMBRERED* se compone de todas estas relaciones.
- d) *NOMBRERED\_instituciones*. Red de nivel superior. En esta red cada nodo es una institución, específicamente es la institución donde se desempeña una persona.
- e) *NOMBRERED\_titulos*. Red de nivel superior. En esta red cada nodo es un título profesional de una persona. Se forma de la misma manera que *NOMBRERED\_instituciones*. Red de nivel superior. En esta red cada nodo es un par institución/título, se forma agrupando en un nodo todos los actores que comparten un mismo título e institución.
- f) *NOMBRERED\_disciplinas1*. Red de nivel superior. En esta red cada nodo es la disciplina del conocimiento en la que se desenvuelve una persona. Se forma de la misma manera que *NOMBRERED\_instituciones*.

La disciplina de una persona se identifica jerárquicamente en tres niveles: el nivel uno corresponde al área del conocimiento (Ej.: CIENCIAS EXACTAS Y NATURALES), el nivel dos es la disciplina general (Ej.: BIOLOGIA) y el nivel tres es una especialización dentro de esa disciplina (Ej.: INMUNOLOGIA). En esta red sólo considera el nivel uno.

- a) *NOMBRERED\_disciplinas2*. Red de nivel superior. Se forma igual que *NOMBRERED-disciplinas1*, sólo que aquí las disciplinas se consideran hasta el nivel dos. Por ejemplo: CIENCIAS EXACTAS Y NATURALES, BIOLOGIA.
- b) *NOMBRERED\_disciplinas3*. Red de nivel superior. Se forma igual que *NOMBRERED-disciplinas1*, sólo que aquí se las disciplinas se consideran hasta el nivel tres. Por ejemplo: CIENCIAS EXACTAS Y NATURALES, BIOLOGIA, INMUNOLOGIA.

- c) NOMBRE RED D seccion-atributo. Red de nivel superior. Por cada relación de similitud se genera una red derivada en que cada nodo se forma de los actores que comparten el mismo valor para la similitud seccion-atributo. Se crea una red por cada similitud. Se forma igual que NOMBRE RED\_instituciones.

#### 5.4. Diseño

Cada proceso mostrado en la figura 0.2, se implementó como un subsistema desacoplado unos de otros, sólo se comunican mediante los almacenes de datos definidos en el diagrama. Cada subsistema se compone de un conjunto de clases que lo implementan.

#### 5.5. Implementación

La implementación de cada clase del sistema se hizo en la plataforma Java. Específicamente se utilizó Java SDK 5.0 usando el editor NetBeans 4.0.

#### 5.6. Funcionamiento

Se muestra un caso con el objetivo de mostrar la interfaz grafica del programa, donde se ingresan los parámetros (Esto corresponde al proceso 3.0). En este caso se generó redes basadas en definiciones de filtros.

Se usó dos filtros:

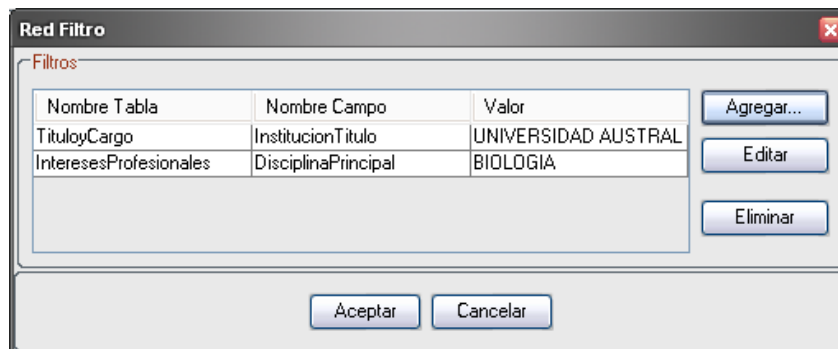
- Institución Titulo en Titulo y Cargo: UNIVERSIDAD AUSTRAL
- Disciplina Principal en Intereses Profesionales: BIOLOGIA

Se usó las siguientes relaciones de asociación:

- Co-autoría en publicación de libros, peso igual a 3
- Co-autoría en publicación de artículos en revistas o diarios, peso igual a 1.
- Co-autoría en publicación de apartados de libros, peso igual a 2.
- Co-autoría en publicación en actas de congresos, peso igual a 1

Los filtros se ingresan en la ventana “Red Filtro”:

Figura 5.6.1.  
Ventana Red Filtro



La ventana “Ingresar Parámetros” muestra todas las relaciones de asociación y similitud ingresadas, como se aprecia en la figura 5.6.2.

Figura 5.6.2.  
Ventana Ingresar Parámetros

**General**

Red Ego      Nombre red      

Red Filtro

---

**Asociaciones**

Nombre Tabla	Nombre Campo	Valor
Libro	AutoresLibro	3
ArticuloRevistasDiarios	AutoresArticulo	1
CapituloLibro	AutoresCapitulo	2
ActasCongr	Autores	1

---

**Similitudes**

Nombre Tabla	Nombre Campo	Valor	Sólo existentes
TituloyCargo	InstitucionEmpresa	1	No

En esta pantalla se puede elegir el tipo de red: Red Filtro o Red Ego (*Snowball*). En “Nombre Red” se le pone un nombre a la red que se quiere generar. Una vez ingresados todos los parámetros se puede guardar la red en un archivo (botón “Guardar”) la que se puede abrir mediante el botón “Abrir”, o simplemente presionar “Ejecutar” para generar las redes.

## 6. ANALISIS A REDES DE INVESTIGADORES

Como demostración del sistema se hizo un análisis a la red completa de currículos del SICTI, utilizando relaciones de asociación en publicaciones. Las similitudes sólo se usaran como ponderadores de las asociaciones (por lo que “Solo-existentes” tiene valor “SI”)

### 6.1. Red completa

Se incluyen todas las personas, en este no hay filtro, por lo que éste tiene el valor *null*.

Tabla 6.1.1.1.  
Red completa

Parámetro	Valor
Tipo de Red	Red Filtro
Nombre de la red	Red_completa

*Asociaciones:*

Tabla 6.1.1.2.  
Asociaciones

Sección	Campo	Peso
Libro	AutoresLibro	4
CapituloLibro	AutoresCapitulo	2
ActasCongr	Autores	1
ArticuloRevistasDiarios	AutoresArticulo	1
OtraProduccionBibliografica	AutoresOtraProduccion	1

*Similitudes:*

Tabla 6.1.1.3.  
Similitudes

Sección	Campo	Peso	Solo-Existente
TituloyCargo	InstitucionEmpresa	1	SI
InteresesProfesionales	disciplinaPrincipal	1	SI

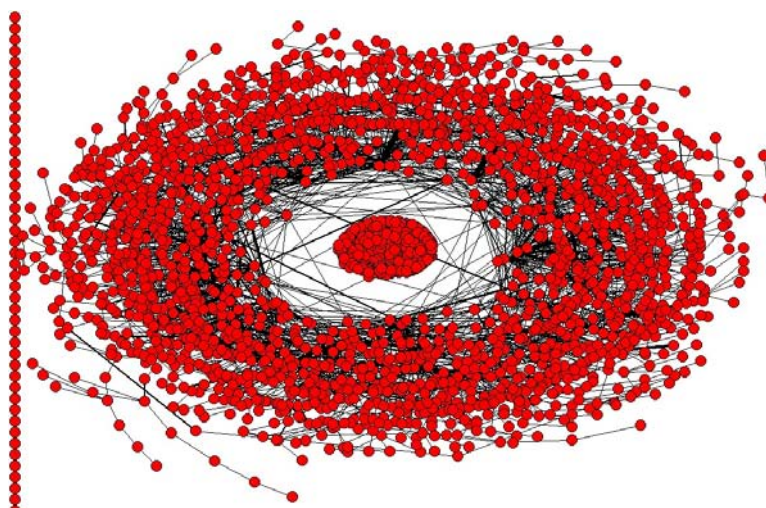
*Filtros:*

Tabla 6.1.1.4.  
Filtros

Nombre	Valor
null-null	Null

Esta gráfica muestra la red completa de actores. La red está compuesta por 7.433 actores. Debido a la gran cantidad de nodos es difícil apreciar la estructura a simple vista; pero basándose en los análisis, se puede observar que la red está constituida por un componente que agrupa a gran parte de los actores, más varios componentes aislados más pequeños.

Figura 6.1.1.  
Red Completa SICTI



En general se observa una red muy dispareja y con muchos componentes aislados. Esto es explicado, en parte, porque la red abarca la totalidad del universo de currículos, lo que produce gran heterogeneidad en los individuos. Es posible que otro factor de este aparente aislamiento sea la falta de información en algunos currículos lo que no permite mostrar todas las conexiones que podrían haber.

Como características positivas de la red, se puede mencionar la corta distancia promedio entre individuos conectados y la existencia de un componente de más de 2.000 individuos que indica la existencia de conexiones inter-disciplinas e inter-instituciones.

Como características negativas está la gran dispersión de individuos fuera del componente de 2000 individuos, y el poco tamaño de cada componente (el mayor es de tamaño 13). Otro punto importante es el de la concentración de las medidas de poder (grado de centralidad, cercanía, intermediación) que está en manos de pocos individuos; lo deseable sería que estuviera mejor distribuido.

Desde el punto de vista de cómo esto puede influir en la toma de decisiones del SICTI, esta la cuestión de la distribución de recursos; o sea dónde faltan y dónde sobran. Por ejemplo: es una buena idea darle más recursos a personas que no necesariamente son las más influyentes en la red, porque eso le daría el efecto de potenciarlo y que creciera como actor en la red. Visto esto mismo desde otro ángulo, se pueden reducir los recursos a los actores con más alto poder, para de este modo preferir a otros actores quizá cercanos a los de más alto poder, pero que no están relacionados con éstos, todo ello, con el objeto de equilibrar la red.

## 6.2. Red de instituciones

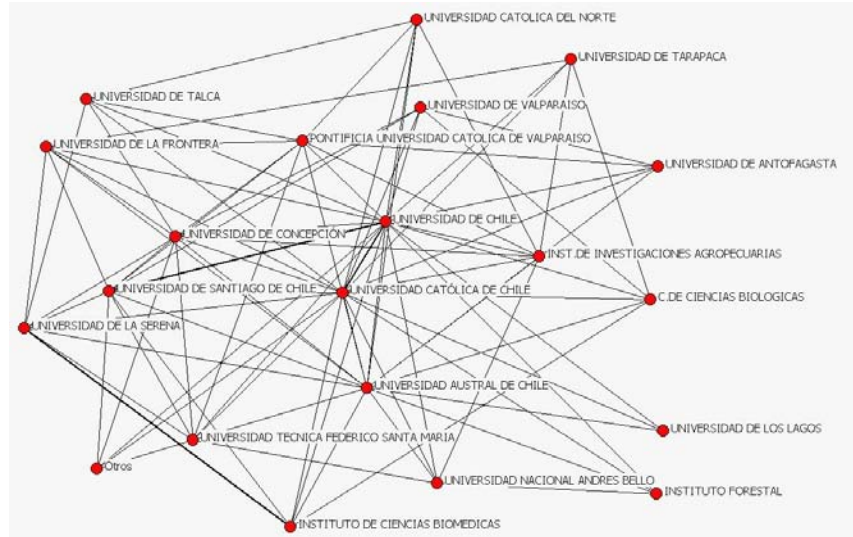
Como es de esperar, debido a que esta red se deriva de la red completa, se observa un componente grande, más una serie de componentes pequeños. En el componente grande se observan en promedio distancias cortas lo que es un indicador positivo para éste.

Las métricas de poder indican que hay 5 instituciones que tienen un importante rol en el ámbito de las publicaciones científicas. Estas son: Universidad de Chile, Universidad Austral de Chile, Universidad de Concepción, Pontificia Universidad Católica de Chile y Universidad de Santiago de Chile.



A continuación se muestra un acercamiento a la red de instituciones, aquellas con mayor puntuación en las medidas de poder.

Figura 6.2.1.  
Acercamiento Red de instituciones



## 7. CONCLUSIONES

### 7.1 Conclusiones generales sobre los resultados obtenidos

El objetivo general de esta investigación fue utilizar el análisis de redes sociales para descubrir cómo están estructuradas las redes de investigadores que utilizan el sistema SICTI y servir de apoyo a la toma de decisiones. La comprensión de la estructura se logró mediante la creación de un sistema completo de extracción y procesamiento de la información contenida en los currículos de investigadores en ciencia y tecnología y la utilización de software de análisis de redes sociales. El hecho de que su implementación haya permitido extraer información útil del SICTI demuestra que puede ser usado efectivamente para la toma de decisiones.

### 7.2. Posibilidades para el futuro

Finalmente, se prevé una serie de cambios que de hacerse efectivos podrían incidir de manera directa o indirecta sobre el SICTI, tales como:

1. *Mejoras en el software:* relacionados con la capacidad de definir más parámetros lo que, a su vez, permitiría un control más fino en la generación de las redes; mejoras en la aplicación para que detecte nombres mal escritos de instituciones, disciplinas, etc.; optimizaciones en los algoritmos de generación de red para disminuir su tiempo de ejecución.

2. *Aumento de información*: se espera una mayor cantidad de currículos en SICTI y de información en currículos ya contenidos, aumentando su calidad.
3. *Uso de otras Métricas y Fuentes de información independientes o complementarias*: el uso de otras Métricas no consideradas en esta investigación, como asimismo, de otras fuentes distintas a la del SICTI para ser utilizadas de manera independiente o complementaria permitiría robustecer el sistema.

## REFERENCIAS

- [1] *Social Network Analysis: Introduction and Resources*: <http://lrs.ed.uiuc.edu/tse-portal/analysis/social-network-analysis/>
- [2] *An Introduction to Social Network Analysis*:  
<http://www.orgnet.com/sna.html>
- [3] D. Krackhardt (1993): “Informal Networks”:  
<http://www.andrew.cmu.edu/~krack/papers/1993%20Informal%20Networks.pdf>
- [4] R. A. Hanneman (2001): “Introduction to Social Network Methods”:  
<http://faculty.ucr.edu/~hanneman/nettext/>
- [5] V. Batagelj y A. Mrvar: “Software Pajek”:  
<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>
- [6] S. Borgatti: “Software NetDraw”:  
<http://www.analytictech.com/downloadnd.htm>