



Bases de datos académicas en Ciencias Computacionales

Seminario de Investigación II

M. Flores, P. Gómez-Gil, C. Feregrino

Versión: Julio 2011

La búsqueda de información en el área de interés es un paso fundamental en el desarrollo de cualquier investigación. Antes de desarrollar ideas propias es importante entender el estado del arte en el tema y familiarizarse con el trabajo previamente realizado, no sólo para aprender sino para evitar repetir el trabajo que otros han llevado a cabo [1, Cap. 13].

Dada la cantidad de material existente, realizar una revisión exhaustiva de la literatura en determinado tema es prácticamente imposible. Por lo tanto, para llevar a cabo una búsqueda de información efectiva es necesario identificar los trabajos *más importantes* que se relacionan con el problema específico de nuestro interés.

En la actualidad, las bases de datos académicas constituyen una de las herramientas de mayor utilidad para identificar publicaciones relevantes en determinada área del conocimiento. Estas bases de datos consisten en una colección de información, generalmente artículos de revistas y otras publicaciones científicas, organizada en un formato que permite acceder fácilmente a ella. Aunque es posible encontrar información sobre un tema usando algún motor de búsqueda, como Google, la principal ventaja al usar una base de datos académica radica en la calidad del material sobre el que se realiza la búsqueda. Mientras que un buscador encuentra todo tipo de trabajos sobre un tema, sin importar si son buenos o malos, en una base de datos académica se tiene acceso a material que ha sido previamente revisado por expertos y que, por lo tanto, tiene cierta autoridad.

El INAOE cuenta con acceso a algunas bases de datos académicas importantes en el ámbito científico. La lista completa de estas bases de datos se encuentra en la página <http://www.inaoep.mx/biblioteca/basesdatos/>.

En las siguientes secciones se describen brevemente cuatro de las bases de datos que pueden ser de mayor utilidad en el área de Ciencias Computacionales (CC), y algunas consideraciones básicas para utilizarlas de manera efectiva.

Es importante señalar que se debe estar dentro de las instalaciones del INAOE para utilizar las bases de datos descritas. Dos de ellas (*Web of Science* y *Computer & Applied Science Complete*) no proporcionan servicio gratuito. Las otras dos ofrecen solamente funcionalidad limitada para quienes no son clientes.

1. Bases de datos recomendadas en CC

1.1. Web of Science

Web of Science, que es parte de la *Web of Knowledge* de Thomson ISI [3], es una base de datos *de referencias* que permite al usuario explorar conexiones entre autores y/o artículos. La base de

datos contiene información bibliográfica sobre trabajos en distintas áreas del conocimiento, y no los documentos en sí.

Esta base de datos selecciona cuidadosamente las revistas que indexa, por lo que se tiene garantía de que la información que muestra es de calidad.

Rastrear referencias por medio de *Web of Science* permite, por ejemplo, averiguar quién citó el trabajo de determinado autor (esto ayuda a identificar grupos de trabajo en el tema de interés), cuántas veces ha sido citado un trabajo (esto da una idea del impacto que ha tenido dicho trabajo en el área) o qué trabajos, a su vez, fueron citados por éste. El sitio también ofrece información bibliográfica sobre los documentos de nuestro interés, lo que facilita la tarea de conseguirlos (ver Sección 3).

Si el usuario se registra en el sitio, se crea un historial de búsquedas al que es posible acceder en cualquier momento. Además, existe también la posibilidad de suscribirse al servicio de *RSS¹ feed* para recibir noticias sobre determinado artículo, autor o tema.

Un beneficio adicional para usuarios de *Web of Science* consiste en poder utilizar la herramienta *My EndNote*, que es una versión en línea del software *EndNote* (un gestor de referencias bibliográficas [2]). Con *MyEndNote* es posible organizar las referencias de interés en carpetas y crear archivos de bibliografía en varios formatos, por ejemplo, *.bib*.

Un tutorial sobre el uso de esta base de datos está disponible en <http://science.thomsonreuters.com/tutorials/wos7/>

1.2. Computer & Applied Science Complete

Computer & Applied Science Complete (CASC) forma parte de un conjunto de bases de datos de investigación denominado *EBSCOhost*. CASC contiene una lista de más de 1300 revistas académicas orientadas a la teoría de computación y sistemas, el contexto social y profesional de nuevas tecnologías, y la ingeniería computacional.

Para acceder a esta base de datos se debe dar click en el link correspondiente (desde la página del INAOE), después dar click en la opción “*EBSCO search databases*” y luego “*Computer Science/Engineering Databases*”. Otra opción consiste en, después de ingresar a la página de EBSCO Host desde la página del INAOE, dar click en *EBSCO Discovery Service*. Esta segunda opción permite realizar búsquedas sin restringirse a la base de datos de computación e ingeniería.

En forma similar a *Web Science*, esta base de datos permite realizar una búsqueda de referencias y acceder al abstract e información bibliográfica de documentos. Adicionalmente, el INAOE cuenta con acceso a la mayoría de los trabajos en esta base de datos.

Es necesario registrarse como usuario para poder mantener un historial de las búsquedas realizadas, guardar referencias de interés y organizar las referencias en carpetas para facilitar su posterior consulta.

Los documentos que están disponibles en su totalidad puede ser identificados porque debajo del título que aparece en la lista de resultados es posible observar el botón *PDF full text*. Si este no es el caso, para obtener el artículo completo es necesario seguir los pasos descritos en la Sección 3 de este documento.

¹*Rich Site Summary*, un formato para informar de manera regular sobre contenido de la Red que cambia con frecuencia

1.3. SpringerLink

Esta base de datos funciona de manera similar a la base descrita en la sección anterior, y permite realizar búsquedas de artículos de revistas, capítulos de libros y referencias. De acuerdo con datos de la compañía, cuenta con casi 40,000 libros electrónicos y más de 2,250 revistas científicas [4].

Al realizar una búsqueda, los artículos que están disponibles para ser descargados directamente se identifican con un cuadrado verde que aparece a la izquierda del título del artículo; si el ícono es blanco, sólo se tiene acceso al abstract, información bibliográfica y referencias del documento (en este caso se debe proceder como se indica en la Sección 3 para obtener el artículo).

Aunque es posible realizar una consulta en la base de datos como “visitante”, registrarse permite recibir correos electrónicos que informen sobre algún tema, artículo o autor de interés, personalizar la página principal de la base de datos y establecer temas favoritos, guardar consultas para futura referencia y organizar datos de interés.

De este sitio es posible descargar libros que pertenecen a colecciones y cuya fecha de publicación sea anterior al año 2005.

1.4. IEEE Computer Society Digital Library

Esta base de datos permite realizar búsquedas de documentos (artículos de revistas y artículos de congreso) pertenecientes a la Sociedad Computacional del IEEE ².

Como en las bases de datos anteriores, es posible realizar búsquedas de referencias, obtener datos bibliográficos, leer abstracts y, si se trata de un usuario registrado, obtener actualizaciones sobre artículos de interés.

En el INAOE se tiene acceso a todo el material que aparece en la página de esta base de datos.

2. Cómo realizar búsquedas

2.1. Búsqueda básica

En esta sección se describen los pasos básicos para realizar una búsqueda en las bases de datos descritas en la sección anterior. Se asume que las búsquedas están llevándose a cabo en las instalaciones del INAOE.

Paso 1: Ingresar a la página de la base de datos en cuestión

Se recomienda ingresar a las bases de datos siguiendo las ligas que se encuentran en la página del INAOE para, de esta manera, asegurar que el sitio reconozca que la búsqueda se está realizando desde el Instituto.

Paso 2: Iniciar la búsqueda

- Especificar en el espacio correspondiente los términos (*keywords*) de interés. Se debe elegir las palabras que mejor describan el tema de investigación.

²Institute of Electrical and Electronic Engineers

- En el caso de *Web of Science* se debe elegir la opción *General Search* para realizar una búsqueda por tema, autor, revista, institución o área geográfica; se debe elegir *Cited Reference Search* para buscar artículos que citan a determinado autor o trabajo.
- Para delimitar o ampliar una búsqueda en las cuatro bases de datos mencionadas, se puede usar las siguientes herramientas:
 - Comillas (“ ”) para especificar una frase
 - Los operadores booleanos AND, OR y NOT para combinar términos
 - Un signo de interrogación (?) para reemplazar una letra en alguna palabra
 - Un asterisco (*) para buscar diferentes terminaciones de una palabra

La Tabla 1 muestra un ejemplo de términos de búsqueda y los resultados que se obtendrían usando las herramientas anteriores.

Tabla 1: *Ejemplo del uso de herramientas para delimitar o ampliar una búsqueda de términos en una base de datos académica.*

Expresión empleada	Resultado obtenido
classification	Trabajos que contienen la palabra classification
“random walk kernel”	Trabajos que contienen la frase (completa y exacta) random walk kernel
gender AND blog	Trabajos que contienen la palabra gender y la palabra blog
learning NOT clustering	Trabajos que contienen la palabra learning pero no la palabra clustering
document OR text	Trabajos que contienen la palabra document o la palabra text
wom?n	Trabajos que contienen las palabras woman o women
comput*	Trabajos que contienen palabras con el prefijo comput , por ejemplo compute , computing , computer

Paso 3: Revisar resultados obtenidos

- Revisar la lista de resultados obtenidos. Con base en el título, se puede determinar si resultan de interés para la investigación que se está llevando a cabo. La Figura 1a muestra un ejemplo de resultados obtenidos al realizar una búsqueda en *Web of Science*.
- Si se desea analizar con más detalle alguno de los trabajos enlistados en los resultados de búsqueda, se debe dar un click en el título de interés. Al dar el click aparecerá el abstract del trabajo y algunas acciones posibles, como descargar el trabajo (si está disponible), exportar la información bibliográfica, usar un marcador para examinar el trabajo en otra ocasión, etc. La Figura 1b muestra un ejemplo de la información que se obtiene al examinar uno de los trabajos en la lista de resultados.

Una vez que se ha examinado el artículo en cuestión, es posible hacer click en el botón *Back to Results* y seguir analizando la lista de resultados obtenidos.

ISI Web of KnowledgeSM

Web of Science | Additional Resources

Search | Cited Reference Search | Advanced Search | Search History | Marked List (0)

Web of Science® – with Conference Proceedings

<< Back to previous

Results Topic=(spatial AND graph)
 Refined by: Topic=(spatial AND graph AND mining)
 Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH.

Results: **53** Page 6 of 6

Print | E-mail | Add to Marked List | Save to EndNote® Web | Save to EndNote®, RefMan, ProCite | more options

51. Title: ImageMap: An Image indexing method based on spatial similarity
 Author(s): Petrakis EGM, Faloutsos C, Lin KI
 Source: IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING Volume: 14 Issue: 5 Pages: 979-987 Published: SE
 Times Cited: 26
 Full Text

52. Title: Multi-level clustering and its visualization for exploratory spatial analysis
 Author(s): Estivill-Castro V, Lee I
 Source: GEOINFORMATICA Volume: 6 Issue: 2 Pages: 123-152 Published: JUN 2002
 Times Cited: 4

Se tiene interés en este título

(a) Lista de resultados

ISI Web of KnowledgeSM

Web of Science | Additional Resources

Search | Cited Reference Search | Advanced Search | Search History | Marked List (0)

Web of Science® – with Conference Proceedings

< Back to results list | Record 52 of 53

Multi-level clustering and its visualization for exploratory spatial analysis

Print | E-mail | Add to Marked List | Save to EndNote® Web | Save to EndNote®, RefMan, ProCite | more options

Author(s): Estivill-Castro V, Lee I

Source: GEOINFORMATICA Volume: 6 Issue: 2 Pages: 123-152 Published: JUN 2002

Times Cited: 4 | References: 37 | Citation Map

Abstract: Exploratory spatial analysis is increasingly necessary as larger spatial data is managed in electro-magnetic media. We propose an exploratory method that reveals a robust clustering hierarchy from 2-D point data. Our approach uses the Delaunay diagram to incorporate spatial proximity. It does not require prior knowledge about the data set, nor does it require preconditions. Multi-level clusters are successfully discovered by this new method in only $O(n \log n)$ time, where n is the size of the data set. The efficiency of our method allows us to construct and display a new type of tree graph that facilitates understanding of the complex hierarchy of clusters. We show that clustering methods adopting a raster-like or vector-like representation of proximity are not appropriate for spatial clustering. We conduct an experimental evaluation with synthetic data sets as well as real data sets to illustrate the robustness of our method.

Document Type: Article
 Language: English
 Author Keywords: clustering; data mining; exploratory spatial analysis; Delaunay diagram; cluster visualization

Guardar en MyEndNote

Abstract

(b) Detalles de un artículo

Figura 1: Ejemplo de resultados obtenidos al realizar una búsqueda en Web of Science. Los términos empleados fueron spatial AND graph AND mining.

Paso 4: Descargar artículos o guardar sus datos bibliográficos

- Las listas que se obtienen al realizar una búsqueda contienen un ícono cerca de cada título, que indica si es posible descargar el documento completo. La Figura 2 muestra el ícono para cada base de datos considerada en este documento.
- Si se desea descargar un artículo (y éste está disponible), se debe dar click en el botón correspondiente.
- Si sólo se desea guardar la referencia del trabajo en cuestión (y el usuario está registrado), se debe dar click en el botón *Add to folder*. Algunas bases de datos (por ejemplo *Web of Science*) permiten crear un archivo bibliográfico a partir de los documentos guardados y seleccionar el formato en el que se exportará la información (BibTex, ProCite, etc.). Para ver información previamente almacenada se debe hacer click en el botón *Filled folder*, *My SpringerLink* o *MyEndNote* que aparece en la base de datos, generalmente en la parte superior de la página, cerca del nombre de usuario.
- Si se desea descargar un artículo que no se encuentra disponible, es necesario guardar la información bibliográfica para después poder solicitarlo.

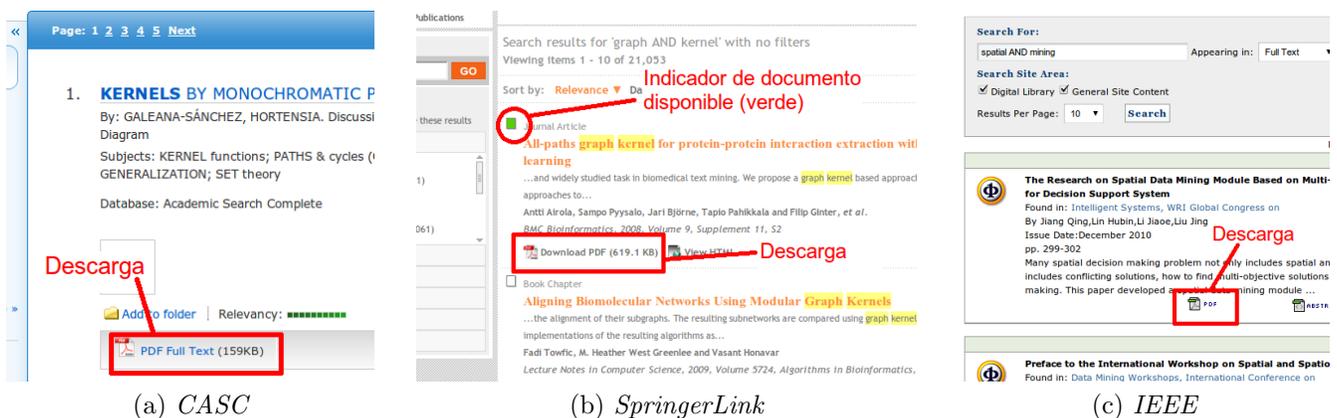


Figura 2: Botones para descargar un documento en las diferentes bases de datos.

2.2. Búsqueda avanzada

Las bases de datos que se han mencionado en este documento cuentan con la opción de realizar una “Búsqueda avanzada”. Para ello es necesario dar click en el botón *Advanced search* que aparece en la página de la base de datos en cuestión; generalmente el botón se localiza en la parte superior de la página. La Figura 3a muestra el botón para el caso de la base de datos de *IEEE*.

Cuando se selecciona esta opción es posible realizar una búsqueda que combine varios criterios usando operadores booleanos y especificar otros parámetros para delimitar la búsqueda de información, como fechas, revistas específicas, categorías u orden en que se desean los resultados. Por *default*, una búsqueda se hace en todas las fuentes (revistas, libros, memorias de congreso) y todos los años cubiertos por la base de datos en cuestión. La Figura 3b muestra un ejemplo de parámetros especificados para realizar una búsqueda avanzada en *IEEE*.



Figura 3: Ejemplo de una búsqueda avanzada en la base de datos IEEE.

3. Cómo obtener material

El INAOE tiene acceso a parte del material indexado en las bases de datos descritas anteriormente. Para ver los íconos que indican si un artículo está disponible y, en su caso, descargar el material, es necesario estar en las instalaciones del INAOE.

De no estar disponible el material, todavía es posible obtener el artículo de interés gracias al convenio que tiene el INAOE con las bibliotecas de algunos de los centros de investigación más importantes del país.

En este caso es necesario contactar al encargado del servicio de *Préstamo interbibliotecario*, ya sea en su oficina (ubicada en el Centro de Información del INAOE) o por medio de un mensaje electrónico que contenga todos los datos bibliográficos necesarios (título completo, autores, revista, año, páginas) para identificar los artículos que se están solicitando.

El encargado del servicio se contacta entonces con otras bibliotecas del país, gestiona el intercambio de material y envía los artículos en cuestión a la persona que los solicitó en, aproximadamente, tres días hábiles.

Actualmente la persona encargada de préstamo interbibliotecario es Victoria Ramírez Saucedo (vramírez@inaoep.mx, extensión telefónica 7052); también es posible contactar a Lilliana Perea Centeno (lperea@inaoep.mx). Para aclarar cualquier otra duda se debe acudir con el personal de la Biblioteca.

4. Conclusiones

En este documento se describieron cuatro de las bases de datos a las que se tiene acceso en el INAOE y que resultan particularmente útiles en Ciencias Computacionales. Asimismo, se mostró un conjunto de pasos básicos para realizar una búsqueda en estas bases de datos y para solicitar algún documento de interés que no pueda descargarse directamente.

El material que se provee en este tutorial no es, de ninguna manera, exhaustivo y se recomienda consultar otros tutoriales disponibles en la Red (especialmente en las páginas mismas de las bases de datos) para explotar al máximo los servicios y funcionalidades que cada una de estas bases de datos ofrece.

Agradecimientos

Se agradece al personal de la biblioteca del INAOE por su ayuda en la realización de estas notas, especialmente a la Srita. Liliana Perea.

Referencias

- [1] M. Berndtsson. *Thesis projects: a guide for students in computer science and information systems*. Springer, 2008.
- [2] Thomson Reuters. EndNote. Sitio Web <http://www.endnote.com>. Consultado el 12 de julio de 2011.
- [3] Thomson Reuters. Web of knowledge. Sitio Web <http://wokinfo.com/>. Consultado el 12 de julio de 2011.
- [4] Springer. Springerlink. Sitio Web <http://www.springerlink.com/>. Consultado el 12 de julio de 2011.