

Universidad Autónoma de Nuevo León (UANL)  
Programa de Apoyo a la Investigación Científica y Tecnológica (PAICyT)

Propuesta para Proyecto de Investigación

Área: Ciencias Aplicadas / Ingeniería

## Métodos locales de agrupamiento de datos

Dra. Satu Elisa Schaeffer

División de Posgrado en Ingeniería de Sistemas (PISIS)  
Facultad de Ingeniería Mecánica y Eléctrica (FIME)  
Universidad Autónoma de Nuevo León (UANL)

Abril 2007

### Resumen

#### Importancia y Relevancia del Proyecto

En muchos problemas de los campos de bioinformática y minería de datos, es común que un instancia típica del problema está estructurado en una manera no uniforme, o sea, está compuesto de subproblemas más altamente restringidas que el problema completo. En muchos casos, los problemas de interés están relaxionados a uno o dos subproblemas densos y no es necesario utilizar los datos completos para ofrecer una solución buena a un subproblema.

En este trabajo planteamos el problema de la *identificación de subproblemas* “independientes” y “interesantes” de problemas del mundo real, es decir, el problema de *agrupamiento* (inglés: *clustering*). El trabajo inicia por definir buenas medidas estructurales para la indentificación de subproblemas. La teoría de grafos ofrece herramientas poderosas para tales definiciones. Existen varios métodos de agrupamiento *global* de grafos, que clasifiquen *el grafo entero* en subgrafos según ciertos criterios. Lo que interesa a nosotros es poder identificar *solamente* el subproblema o los subproblemas *relevantes* desde el punto de vista de una pregunta de naturaleza *local*. Ya existen algunas ideas iniciales sobre el agrupamiento local, pero hace falta trabajo riguroso en el desarrollo de algoritmos de agrupamiento local que sirvan para problemas del mundo real: en las aplicaciones, los datos suelen ser más complejos que grafos simples, y los algoritmos existentes no prestan mucha atención a las posibilidades de *ponderación* y *dirección* de la representación de los datos como un grafo.

Nuestro enfoque está en problemas del tipo de optimización combinatorial, donde por lo general se busca por configuraciones que cumplen con ciertos requisitos: las rutas de entrega más cortas o las ubicaciones más adecuadas para centros de distribución, etcétera. El problema general de agrupamiento, como muchos de optimización combinatorial, es NP-difícil. Sin entrar en detalles técnicos, no se cree que puedan existir algoritmos exactos de solución con tiempo de ejecución que aumente no más rápido que polinomialmente con el tamaño del problema. Con procesar solamente subproblemas, podemos reducir el tamaño del problema. Esto combinado con la aplicación de algoritmos de *aproximación* ya ofrece una solución viable para datos de tamaño masivo igual como para datos sujetos a cambios y actualizaciones frecuentes. El reto está en explotar efectivamente la estructura no uniforme de los datos por su representación de la forma de un grafo.

#### Metodología Propuesta

El desarrollo e implementación de algoritmos locales de agrupamiento de grafos que sean capaces de ofrecer soluciones de alta calidad rápidamente será el enfoque principal del trabajo propuesto. La hipótesis de la propuesta es que el problema en estudio puede ser resuelto de manera eficiente, es decir, desde un enfoque de la aplicación de métodos para optimizar *medidas locales* matemáticamente fundadas. Los detalles de la metodología se exponen en el protocolo del proyecto. En resumen, en este trabajo se pretende llevar a cabo investigación sobre el agrupamiento local, relevante en el campo de minería de datos y biotecnología. Aplicamos álgebra de grafos y técnicas avanzadas de optimización combinatorial para desarrollar métodos de solución eficientes - es aquí donde yace la contribución científica del trabajo propuesto. El trabajo involucra un componente computacional bastante significativo. Hasta donde se tiene conocimiento, el problema de agrupamiento local aquí planteado todavía es un campo abierto con muchas aplicaciones potenciales de gran impacto y muy poco trabajo existente.

## Protocolo de Investigación

### 1. Introducción y Antecedentes Científicos

**Introducción:** En muchos ambientes de minería de datos, tales como aquellos provenientes de la biotecnología o del mundo de negocios, es común que varias de las preguntas de interés son de carácter *local*. Con localidad queremos decir que una respuesta adecuada no cuenta con ninguna información *global* de los datos de entrada, ni necesita el procesamiento de los datos completos para su construcción. Preguntas de ese tipo son por ejemplo las siguientes:

- ✘ ¿Cuáles genes del paciente Hernández y del paciente Gonzáles son los más parecidos?
- ✘ ¿Cuáles proteínas tienen la estructura más similar?
- ✘ ¿Cuáles clientes compraron productos similares? ¿A quiénes deberíamos dirigir promociones?
- ✘ ¿Cuáles empresas tienen historia bursátil parecida?
- ✘ ¿Qué libros tiene la biblioteca de la misma tema que este libro interesante que ya leí?

En tal caso, cualquier método de buscar la respuesta que cuente con procesamiento completo del base de datos sería altamente ineficiente, aún más así si la cantidad de información es masiva y sujeto a modificación constante, por ejemplo a través de la introducción de datos nuevos. Por lo tanto, la pregunta de *cómo evaluar propiedades estructurales de naturaleza local* se convierte en crucial al encontrar respuestas a las preguntas de los campos de aplicación con eficacia y rapidez.

En este trabajo planteamos los problemas de definir medidas de calidad de la estructura local de datos con justificación matemática, por utilizar métodos de *la teoría y el álgebra de grafos* (Chung 1997). Pretendemos aplicar las medidas desarrolladas para identificar de subestructuras que corresponden directamente a la respuesta de una pregunta de naturaleza local. Tales medidas se basan en las relaciones entre los datos, fácilmente capturados en representaciones en forma de grafos. Un grafo consiste de un conjunto de nodos o *vértices* que están conectados por un conjunto de conexiones o *aristas*. Para capturar mejor los datos, típicamente se impone *pesos* en los vértices y aristas, así creando un *grafo ponderado*. Muchas veces las conexiones no son simétricas, y el grafo que resulta es *dirigido*.

El problema de identificación de subestructuras “relativamente independientes” y densamente conexas en grafos se llama *agrupamiento* (inglés: *clustering*). La mayoría de las formulaciones del problema de agrupamiento de grafos son sumamente difíciles de resolver: están clasificados como problemas NP-difíciles (Šima y Schaeffer 2005, Garey y Johnson 1979). Sin entrar en detalles técnicos, no se cree que sea posible construir un algoritmo que tuviera un tiempo de ejecución polinomial en términos de la cantidad de datos involucrados. En la práctica, se requiere soluciones rápidas y el tiempo de ejecución se convierte en un asunto crucial. Típicamente se evita la resolución de problemas NP-difíciles por construir soluciones *aproximadas* que no son las soluciones exactas óptimas. En el diseño de algoritmos de aproximación, el reto está en explotar efectivamente la estructura matemática del problema para poder encontrar rápidamente soluciones aproximadas de alta calidad.

Sin embargo, si el problema tiene tamaño masivo, incluso un método aproximado puede consumir demasiado tiempo. Igualmente en el caso que los datos están actualizados con frecuencia, resulta imposible construir ni una aproximación para la solución entera antes de que ya haya otro cambio en las entradas del algoritmo. Por eso, nuestro interés está en *algoritmos de cómputo local*. Aunque la solución exacta de los métodos locales confronta los mismos problemas de ser NP-difícil, la reducción en la cantidad de datos procesados puede resultar en ahorro inmenso con respecto al tiempo de ejecución y el uso de recursos de memoria. Esto facilita la solución de problemas de tamaño masivo de una manera distribuida y paralela, y además ofrece un método para contestar directamente preguntas interesantes de naturaleza local sin tener que resolver el problema entero. Con computación local, cambios frecuentes en los datos ya no poseen dificultades.

**Planteamiento del Problema:** En un grafo  $G = (V, E)$ , donde  $V$  es un conjunto de los vértices y  $E$  un conjunto de aristas, ¿cuál es el subgrafo *relevante* de un dado vértice  $v$  en  $V$ ? Es decir, ¿cuáles otros vértices  $w$  en  $V$  son estructuralmente cercanos al vértice  $v$ . La Fig. 1 ilustra un ejemplo de un grafo con 55 vértices que forman siete “cuevas” de vértices cercanos, donde es fácil identificar la estructura de los grupos locales.

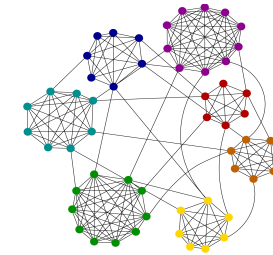


Figura 1: Un grafo de 55 vértices que forman siete subgrafos densos.

**Trabajo existente:** La investigación en el campo de agrupación de datos generales (inglés: *data clustering*) ha sido muy amplia (Jain y Dubes 1988) y las propuestas para agrupamiento *global* de grafos, es decir, la computación de todos los subgrupos simultáneamente utilizando toda la información de adyacencia del grafo, ya son varios (Newman 2004). Sin embargo, por lo general están basados en computaciones extensivas con un tiempo de ejecución larga. Métodos globales tienen una naturaleza muy distinta al enfoque que se presenta en este trabajo: la identificación de subestructuras relevantes por *computación local* y su solución rápida aproximada. Lograr tiempos de ejecución muy pequeños facilita la adaptación a cambios frecuentes en los datos de entrada, un escenario muy común en los campos de aplicación de minería de datos. Los métodos existentes para agrupamiento local son pocos: los primeros fueron los métodos Schaeffer (2005), Orponen y Schaeffer (2005) y Clauset (2005). Técnicamente una posibilidad para agrupamiento local fue sugerida por Wu y Huberman (2004), pero no contaba con utilización en computación local explícita, sino funcionaba con computación global.

**Áreas de oportunidad:** Hasta donde se tiene conocimiento, el problema de agrupación local no todavía tiene soluciones satisfactorias con justificación matemática y éxito práctica. Especialmente en el caso de **grafos ponderados y dirigidos** prácticamente todavía no existen resultados. Además, la aplicación de los métodos de agrupamiento local en problemas del mundo real ha sido bastante limitada por la naturaleza teórica de las pocas propuestas existentes.

En las aplicaciones, por lo general los grafos son por lo menos ponderados y en muchos casos también dirigidos, ambos aspectos dan al problema de agrupación local una estructura matemática diferente al problema de grafos simples no ponderados, para el cual es necesario su estudio y solución eficiente, lo cual constituye el componente científico básico de esta propuesta de investigación. En particular, por la definición de medidas con justificación matemática y su optimización por técnicas de búsqueda local pueden ofrecer métodos de solución innovativas, eficientes, y con resultados de alta calidad.

**Metodología propuesta:** El desarrollo e implementación de algoritmos eficientes de agrupación local de grafos ponderados y dirigidos, que todavía no existen, será el enfoque principal del trabajo propuesto. Dado su enfoque matemático en el diseño de las medidas de calidad de agrupamiento, el desarrollo y estudio de tales algoritmos pertenece a una área entre las matemáticas y ciencias computacionales. Las aplicaciones del mundo real provienen por de varios ramos de ciencia, actualmente en especial de la biotecnología. Para lograr a cabo investigación exitosa de agrupamiento local, pasamos por seis pasos básicos:

- (1) un **entendimiento matemático** de la estructura de subgrupos por álgebra de grafos y propiedades espectrales como los **eigenvectores** de las matrices de adyacencia del grafo de entrada,
- (2) las definiciones de **medidas de calidad** para determinar hasta qué grado un dado subgrafo forma un buen grupo para agrupamiento,
- (3) el diseño de **algoritmos exactos** para encontrar el mejor grupo para un dado vértice de interés, aunque serán por lo general de tiempo exponencial, por la naturaleza NP-dura del problema de agrupamiento,
- (4) el desarrollo de **algoritmos de aproximación** para agrupamiento local por el uso de las medidas como funciones objetivos de búsqueda local,
- (5) el **análisis teórica** de la eficiencia de los algoritmos desarrollados y de la exactitud de las aproximaciones en comparación con los algoritmos exactos para instancias pequeñas, y
- (6) la amplia **experimentación** aplicada con datos artificiales generados para probar los metodos y datos reales de problemas prácticas relevantes.

En cada uno de estos pasos, hay que cuidar especialmente la adaptación de los métodos para aplicar también con grafos ponderados y dirigidos. También hay que cuidar que explotamos en plena potencia la estructura no uniforme de las instancias típicas, en vez de desarrollar métodos para “casos promedios aleatorios”. Así podemos garantizar resultados nuevos y métodos útiles en uso práctico para los varios campos de aplicación.

## 2. Objetivos y Metas

### Objetivos Generales:

- × **Investigación:**
  - Efectuar investigación básica de problemas de agrupación de grafos a fin de lograr un mayor entendimiento y técnicas de solución más eficientes y matemáticamente justificadas. El trabajo estará basado en problemas de grafos de tamaño masivo para procesamiento en paralelo.
  - Efectuar investigación y desarrollo de medidas de calidad para subgrafos basada en las matemáticas de procesos estocásticos, aplicados a resolver el problema en cuestión.
  - Avanzar significativamente el estado del arte en la área de agrupación de datos para resolver en paralelo o con aproximaciones los problemas de optimización combinatoria, permaneciendo a la vanguardia en dichas líneas de investigación.
- × **Colaboración conjunta:** Realizar trabajo conjunto con investigadores de prestigio internacional en áreas afines con la finalidad de fomentar los lazos de colaboración de nuestro programa con otros programas de reconocido nivel.
- × **Formación de recursos humanos** de alto nivel: Involucrar a estudiantes de posgrado, cuyos trabajos de tesis se enfocarán a tratar tareas específicas del proyecto y que serán dirigidos por el investigador responsable y posiblemente co-dirigidos por algunos de los investigadores que colaboran en el proyecto.
- × **Difusión:** Publicar y diseminar resultados de la investigación realizada en revistas y congresos de prestigio internacional y nacional.
- × Apoyar a la infraestructura del programa.

### Metas Específicas:

- × Lograr un entendimiento profundo de la estructura matemática del problema planteado, con la finalidad de poder desarrollar técnicas de solución que exploten a ésta favorablemente y que permitan estudiar posteriormente otras variaciones más complejas de los problemas.
- × Desarrollar algoritmos de agrupamiento local para grafos dirigidos y ponderados, en paralelo a generalizar y adaptar los métodos ya desarrollados por el grupo de trabajo para grafos simples.
- × Desarrollar e implementar computacionalmente métodos de búsqueda local para resolver problemas de tamaño masivo con computación paralela y distribuida.
- × **Formación de recursos humanos:** Dirigir y formar al menos un estudiante de posgrado.
- × **Publicación y diseminación de resultados:** Se espera exponer resultados parciales y finales en al menos un foro internacional y un nacional y publicar al menos un artículo en revista internacional, uno en memoria de congreso internacional y uno en revista o memoria de congreso nacional.
- × **Adquisición de la infraestructura** necesaria para su exitosa realización:
  - Un equipo de cómputo tipo Dual Core con memoria suficiente para los experimentos.
  - Software especializado de matemáticas, simulación y programación.
  - Textos y revistas en áreas relacionadas al proyecto.

### 3. Metodología

- (1) La primera etapa del proyecto estará enfocada al estudio matemático del problema para poder establecer propiedades y caracterizaciones de nuevas y extensiones del trabajo existente para grafos dirigidos y ponderados.
- (2) La segunda etapa del proyecto estará enfocada a experimentar con las medidas y definiciones formuladas en la primera etapa y para diseñar conceptualmente cómo aplicar métodos de búsqueda local en la optimización de dichas medidas.
- (3) La tercera etapa del trabajo consistirá en desarrollar e implementar computacionalmente los algoritmos diseñados en la segunda etapa. El Dr. Pekka Orponen, experto reconocido en el área de teoría de computación, colaborará para darle un sólido soporte a nuestros algoritmos.
- (4) Finalmente, la cuarta etapa del trabajo consistirá en evaluar experimentalmente y por comparaciones los algoritmos desarrollados utilizando tanto instancias de datos reales como datos generados, así como la elaboración de los correspondientes artículos científicos.

### 4. Infraestructura y apoyo técnico disponible

El Programa de Posgrado en Ingeniería de Sistemas es de reciente creación, motivo por el cual se encuentra en una etapa de desarrollo en materia de adquisición y reforzamiento de infraestructura. La actual infraestructura disponible en el Programa Doctoral de la Facultad de Ingeniería Mecánica y Eléctrica para la realización del proyecto es la siguiente:

- ✘ **Equipo de cómputo:**
  - Un laboratorio de cómputo con un servidor Sun Fire V440, una estación de trabajo Sun Ultra 10 y veinte terminales gráficas que comparten los recursos del servidor.
  - Una computadora del laboratorio de los estudiantes de PISIS con procesador de Pentium 2.66 Ghz, 512 KB de memoria y un disco duro de 40 GB en uso del investigador principal mientras no cuente con equipo de cómputo por proyectos.
  - **Software:**
    1. Compiladores de varios lenguajes de programación, como C, C++ y Java.
    2. Sistema de librerías para resolver problemas de optimización (CPLEX).
    3. Programas de modelación algebraica de problemas de optimización (AMPL y GAMS).
    4. Programa de modelación matemática (MATLAB y Octave).
  - **Red:** Todo el equipo de cómputo se encuentra conectado a la red de la UANL.
- ✘ **Material bibliográfico:** La biblioteca del Programa Doctoral cuenta con número limitado de textos y de publicaciones periódicas debido a la reciente creación del programa. Entre los requerimientos financieros del proyecto se contempla la adquisición de textos en las áreas de optimización combinatorial, teoría de grafos y minería de datos.
  - **Publicaciones periódicas recientes:** para el proyecto, las más relevantes son *INFORMS Journal on Computing*, *Journal of Combinatorial Optimization*, *Journal of Heuristics*, *Mathematical Programming* y *SIAM Review*.

### 5. Formación de Recursos Humanos

- ✘ Uno o dos estudiantes de **maestría** (dos estudiantes ya expresaron interés de trabajar en el tema).
- ✘ Uno o dos estudiantes de **tesis de licenciatura**.
- ✘ Dos estudiantes de licenciatura del Programa de **Verano Científico y Tecnológico** de la UANL.

### 6. Calendario de Actividades Trimestral

#### Trimestre 1/4

- ✘ Obtención de la literatura necesaria para la realización del proyecto.
- ✘ Obtención de equipo de cómputo, instalación del software necesario (por gran parte herramientas gratuitas o desarrolladas por los investigadores mismos).
- ✘ Estudio de la estructura matemática de grafos para definir medidas de calidad de agrupamiento.
- ✘ Derivación de propiedades y caracterización de soluciones de agrupación local.

#### Trimestre 2/4

- ✘ Derivación y diseño de métodos de solución por búsqueda local para optimizar las medidas definidas en el primer trimestre.
- ✘ Implementación de técnicas de optimización exacta para instancias de tamaño moderado.
- ✘ Pruebas computacionales preliminares con instancias de tamaño masivo.

#### Trimestre 3/4

- ✘ Estudio y derivación de métodos de búsqueda local para instancias de tamaño masivo.
- ✘ Refinación e implementación computacional de los algoritmos de agrupación a través de búsqueda local.
- ✘ Evaluación computacional preliminar de calidad de soluciones obtenidas por búsqueda local.

#### Trimestre 4/4

- ✘ Evaluación teórica y computacional de los métodos desarrollados con experimentación extensiva.
- ✘ Preparación de artículos científicos para publicar y divulgar los resultados teóricos y prácticos obtenidos durante los primeros tres cuatrimestres.
- ✘ Presentación de resultados en un foro internacional y uno nacional.

## 7. Referencias Bibliográficas

- F.R.K. CHUNG. Spectral Graph Theory. American Mathematical Society, Providence, RI, EUA, 1997.
- A. CLAUSET. Finding Local Community Structure in Networks. *Physical Review E*, 72:026132, 2005.
- M.R. GAREY y D.S. JOHNSON. Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, San Francisco, CA, EUA, 1975.
- M.E.J. NEWMAN. Detecting community structure in networks. *The European Physical Journal B*, 38:321-330, 2004.
- A.K. JAIN y R.C. DUBES. Algorithms for Clustering Data. Prentice-Hall, Englewood, NJ, EUA, 1988.
- P. ORPONEN y S.E. SCHAEFFER. Local Clustering of Large Graphs by Approximate Fiedler Vectors. En Sotiris Nikolettseas (editor), *Proceedings of the Fourth International Workshop on Efficient and Experimental Algorithms (WEA'05)*, Lecture Notes in Computer Science, Vol. 3593, pp. 524-533. Springer, Berlin, Alemania, 2005.
- J. ŠIMA y S.E. SCHAEFFER. On the NP-Completeness of Some Graph Cluster Measures. En J. Wiederman et al. (editores), *Proceedings of the Thirty-second International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM-06)*, Lecture Notes in Computer Science, Vol. 3831, pp. 350-357. Springer, Berlin, Alemania, 2006.
- S.E. SCHAEFFER. Stochastic Local Clustering of Massive Graphs. En T.B. Ho et al. (editores), *Proceedings of the Ninth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-05)*, Lecture Notes in Artificial Intelligence, Vol. 3518, pp. 354-360. Springer, Berlin, Alemania, 2005.
- F. WU y B.A. HUBERMAN. Finding communities in linear time: a physics approach. *The European Physical Journal B*, 38:331-338, 2004.

## 8. Participantes

### Responsable

**Nombre:** Satu Elisa Schaeffer

**Grado:** Doctor (D.Sc.), Teoría de Computación

**Nivel en el SNI:** solicitante de nuevo ingreso en la convocatoria de 2007 (primera oportunidad)

### Últimas cinco publicaciones:

- A. SCHUMACHER, H. HAANPÄÄ, S.E. SCHAEFFER y P. ORPONEN. Load Balancing by Distributed Optimisation in Ad Hoc Networks. En J. Cao et al. (editores), *Proceedings of the Second International Conference on Mobile Ad Hoc and Sensor Networks (MSN 2006)*, Lecture Notes in Computer Science, Vol. 4325, pp. 874-885. Springer, Berlin, Alemania, 2006.
- S.E. SCHAEFFER, S. MARINONI, M. SÄRELÄ y P. NIKANDER. Dynamic Local Clustering for Hierarchical Ad Hoc Networks. En *Proceedings of the Third Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON/IWWAN 2006)*, Vol. 2, pp. 67-672. IEEE Communications Society, Reston, VA, EUA, 2006.
- J. ŠIMA y S.E. SCHAEFFER. On the NP-Completeness of Some Graph Cluster Measures. En J. Wiederman et al. (editores), *Proceedings of the Thirty-second International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM-06)*, Lecture Notes in Computer Science, Vol. 3831, pp. 350-357. Springer, Berlin, Alemania, 2006.
- S.E. SCHAEFFER. Stochastic Local Clustering of Massive Graphs. En T.B. Ho et al. (editores), *Proceedings of the Ninth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-05)*, Lecture Notes in Artificial Intelligence, Vol. 3518, pp. 354-360. Springer, Berlin, Alemania, 2005.
- P. ORPONEN y S.E. SCHAEFFER. Local Clustering of Large Graphs by Approximate Fiedler Vectors. En S. Nikolettseas (editor), *Proceedings of the Fourth International Workshop on Efficient and Experimental Algorithms (WEA'05)*, Lecture Notes in Computer Science, Vol. 3593, pp. 524-533. Springer, Berlin, Alemania, 2005.

### Colaborador

**Nombre:** Pekka Orponen, Helsinki University of Technology (TKK), Espoo, Finlandia

**Grado:** Doctor (Ph.D.), Teoría de Computación

**Nivel en el SNI:** N/A

### Últimas cinco publicaciones:

- S. PRASAD, A. SCHUMACHER, H. HAANPÄÄ y P. ORPONEN. Balanced Multipath Source Routing. En *Proceedings of the Twenty-first International Conference on Information Networking (ICOIN'07)*, Lecture Notes in Computer Science. Springer, Berlin, Alemania, 2007.
- A. SCHUMACHER, H. HAANPÄÄ, S.E. SCHAEFFER y P. ORPONEN. Load Balancing by Distributed Optimisation in Ad Hoc Networks. En J. Cao et al. (editores), *Proceedings of the Second International Conference on Mobile Ad Hoc and Sensor Networks (MSN 2006)*, Lecture Notes in Computer Science, Vol. 4325, pp. 874-885. Springer, Berlin, Alemania, 2006.
- P. FLORÉEN, P. KASKI, J. KOHONEN y P. ORPONEN. Exact and Approximate Balanced Data Gathering in Energy-Constrained Sensor Networks. *Theoretical Computer Science*, 344(1):30-46. Noviembre 2005.
- S. SEITZ, M. ALAVA y P. ORPONEN. Focused Local Search for Random 3-Satisfiability. *Journal of Statistical Mechanics: Theory and Experiment*, P06006:1-27. Junio 2005.
- E. GRIFFITHS y P. ORPONEN. Optimization, Block Designs and No Free Lunch Theorems. *Information Processing Letters*, 94(2):55-66. Abril 2005.

## 9. Justificación Financiera

Rubro: 100 Gasto corriente y 200 Gasto de inversión

Clave	Partida	Descripción	Importe (MN)
101	Viáticos		0
102	Pasajes	Boleto de avión a congreso internacional: 590 USD X 11 MN/USD = \$6,490.	6,490
103	Trabajo de campo	N/A	0
104	Ediciones e impresiones	Impresión de artículos y encuadernación de tesis.	1,900
105	Servicios externos	Servicios de copiado	500
106	Cuotas de inscripción	Un congreso nacional \$2,000 MN y un congreso internacional 350 USD; las suscripciones a Institute for Operations Research and Management Science (130 USD), Mathematical Programming Society (80 USD), Society for Industrial and Applied Mathematics (120 USD). 210 USD x 11 MN/USD = \$7,480.	9,480
107	Artículos diversos	Consumibles, software, mobiliario para equipo	4,500
108	Libros y revistas	Suscripción a revistas (\$2,000); textos relevantes en las áreas de matemáticas, computación y optimización (\$3,000).	5,000
109	Animales	N/A	0
110	Becas	Dos (2) estudiantes de verano científico.	2,000
		<b>Total de gasto corriente</b>	<b>29,870</b>
201	Equipo de apoyo		0
202	Equipo de cómputo	Una computadora desktop Mac Pro con sistema operativo UNIX/MacOS con las siguientes especificaciones: <ul style="list-style-type: none"> <li>✘ 2 procesadores con doblenúcleo a 2,66 Ghz</li> <li>✘ 4 MB de caché de nivel 2 compartida</li> <li>✘ Dos buses frontales a 1,33 Ghz</li> <li>✘ 1 GB de memoria a 667 Mhz</li> <li>✘ Tarjeta gráfica con 256 MB de memoria</li> <li>✘ Disco duro SATA de 250 GB a 7.200 rpm</li> <li>✘ SuperDrive de DVD±RW y CD-RW</li> </ul> Adjuntado están la cotización utilizada de Mac Digital S.A. y otra cotización recibida (con precio de 35,977 pesos) de Mac S.A. Los otros proveedores contactados no enviaron cotizaciones.	35,116
		<b>Total de gasto de inversión</b>	<b>35,116</b>
		<b>Suma Total</b>	<b>64,986</b>