

Programa de Mejoramiento del Profesorado
PROMEPE
Área: Ingeniería y Tecnología
 Ciencias de la Computación
 Solicitud de Extensión al Proyecto de Investigación
Búsqueda adaptiva en redes dinámicas masivas

Dra. Satu Elisa Schaeffer

(1) Centro de Innovación, Investigación y Desarrollo en Ingeniería y Tecnología (CIIDIT)

(2) División de Posgrado en Ingeniería de Sistemas (PISIS),

Facultad de Ingeniería Mecánica y Eléctrica (FIME)

Universidad Autónoma de Nuevo León (UANL)

Período: del 1 de noviembre de 2008 al 31 de octubre de 2009

	1
1. El cumplimiento de las metas y objetivos planteados	2
2. Desglose financiero del primer año	4
3. Resultados obtenidos con la beca	6
4. Productos académicos obtenidos	8
5. Protocolo para el segundo año	10
5.1. Introducción y antecedentes científicos	10
5.1.1. Planteamiento del problema	12
5.1.2. Trabajo existente y áreas de oportunidad	14
5.2. Objetivos y metas	15
5.3. Metodología	17
5.4. Infraestructura disponible	17
5.4.1. Equipo de cómputo	18
5.4.2. Material bibliográfico	18
5.5. Incidencia del proyecto en el PIFI	20
5.6. Formación de recursos humanos	20
5.7. Vinculación y colaboración	21
5.7.1. Participantes en la UANL	21
5.7.2. Participantes externos a la UANL	22
Referencias bibliográficas	22
6. Cronograma de actividades	26
7. Solicitud de apoyos para el segundo año	27
8. Solicitud de beca de fomento a la permanencia	29
9. Productos académicos esperados	30
10. Documentos probatorios	
10.1. Formación de recursos humanos	30
10.2. Producción académica	30
10.3. Colaboración científica	30
10.4. Becario de licenciatura	30

Sección 1

El cumplimiento de las metas y objetivos planteados

- **Investigación:** Hemos efectuado investigación básica de problemas de búsqueda y problemas relacionados. Hemos desarrollado prototipos de técnicas de solución novedosas. Hemos avanzado el estado del arte en la área de búsqueda de recursos por publicaciones científicas en congresos internacionales.
- **Colaboración conjunta:** Hemos trabajado en colaboración entre entidades mexicanas y en nivel internacional con investigadores finlandeses. Hemos realizado las siguientes visitas de investigación entre las entidades participantes:
 1. Una visita de investigación de la Dra. Schaeffer en el Instituto Tecnológico de Ciudad Madero, visitando a la Dra. Laura Cruz Reyes, de 14 a 17 de noviembre del 2007, Ciudad Madero, Tamaulipas, México.
 2. Una visita de investigación de la Dra. Cruz Reyes y sus estudiantes del ITCM en la Universidad Autónoma de Nuevo León, de 10 a 15 de diciembre del 2007, visitando a la Dra. Schaeffer.
 3. Una visita de investigación de la Dra. Schaeffer y un estudiante suyo de la UANL en el Instituto Tecnológico de Ciudad Madero, visitando a la Dra. Cruz Reyes, de 30 de junio a 5 de julio del 2008, Ciudad Madero, Tamaulipas, México.
 4. Una estancia de investigación de la Dra. Schaeffer en Finlandia de 28 de julio a 12 de agosto del 2008, visitando al Dr. Orponen en la Universidad Politécnica de Helsinki TKK.
- **Formación de recursos humanos de alto nivel:** Hemos involucrado cuatro estudiantes de maestría y un becario de licenciatura de tesis en temas relacionadas con el proyecto bajo la dirección del investigador responsable, la Dra. Schaeffer. Los tesis de maestría están por graduar para el fin de este semestre de otoño del 2008.

También está involucrada un estudiante de doctorado de la Dra. Cruz, M.C. Claudia Guadalupe Gómez Santillán, quien ya lleva un avance significativo, y varios estudiantes de maestría de la ITCM en adición a Lic. Marco Aguirre, codirigido entre Dra. Cruz y Dra. Schaeffer. La Dra. Cruz pertenece al comité evaluador de Ing. Perla Cantú, tesista de Dra. Schaeffer.

En la dirección de la tesis de Ing. Ríos participa el Dr. Bustos de la UDP en Chile y el Dr. Berrones de la UANL.

- **Difusión:** Hemos publicado artículos y presentado trabajos en congresos nacionales e internacionales. Estamos finalizando artículos para su publicación en revistas indexadas internacionales.
- **Apoyo al programa:** Hemos apoyado la infraestructura del programa por proveer un servidor de experimentos (una Dell quadcore de 4 GB de RAM) y una computadora personal (una MacMini) para el uso de las tesis, junto con periféricos como impresoras (una láser blanco-negro económico y una inkjet multifuncional de color).

Sección 2

Desglose financiero del primer año

Cuadro 2.1: Monto ejercido por rubro en 2007-2008.

Rubro	Apoyado	Ejercido	Saldo	Justificación
Asistencia a reuniones académicas	16,000.00	15,050.53	949,47	Dra. Schaeffer asistió con los fondos del proyecto combinados con otros fondos de investigación a un <i>congreso nacional</i> CISCE 2008 (vuelos, septiembre) y un congreso internacional ILAS 2008 (viáticos y transporte local, junio).
Beca para estudiante	18,448.00	18,448.00	0.00	Se otorgó beca a Carlos Alberto Castillo Salazar. Realizó sus prácticas profesionales y está ahora escribiendo su tesis. El becario original, Gabriela Chamorro Sotelo, obtuvo un empleo de tiempo completo antes de la recepción de la beca.
Equipo	15,000.00	15,000.00	0.00	Solicitábamos fondos para dos <i>computadoras simples</i> tipo MacMini para el uso de los <i>estudiantes</i> del proyecto para compartir los recursos del servidor MacPro solicitado en sus simulaciones. Recibimos fondos suficientes para la compra de una acompañada con los periféricos necesarios, incluyendo un UPS no-break protector de voltaje, una memoria externa USB y una impresora multifuncional.

Continúa en la siguiente página...

Cuadro 2.1 – Continuoado

Rubro	Apoyado	Ejercido	Saldo	Justificación
Equipo para experimentación	30,000.00	29,025.00	975.00	De los 42,364 pesos solicitamos, fueron otorgados 30,000 pesos, que no alcanzó para la compra de un <i>servidor multiprocesador</i> MacPro, pero fueron suficientes para la compra de una con un procesador de cuatro núcleos marca Dell que hemos ocupado para realizar <i>experimentos</i> y <i>compartir</i> los recursos del servidor entre los investigadores y los estudiantes. También se compró una pantalla de 17 pulgadas, un disco duro externo y algunos periféricos necesarios.
Estancias cortas	17,000.00	17,000.00	0.00	Dra. Schaeffer realizó una estancia de investigación en la TKK en Finlandia en julio-agosto del 2008.
Gastos de publicación	0	0	0	No recibimos fondos en este rubro.
Materiales y consumibles	32,239.00	32,239.00	0.00	Se ha utilizado por la mayor parte software libre y otras herramientas sin costo, por lo cual no hubo altos gastos de software. El software comprado consiste de paquetes de herramientas de oficina y de sincronización para Mac para el uso en las computadoras de escritorio (iWork, iLife, MS Office y .Mac). Por corte en el monto otorgado en el rubro "Equipo para experimentación", no pudimos adquirir la máquina solicitada MacPro, sino uno de menor rendimiento de la marca Dell, por lo que no fue necesario una licencia del <i>sistema operativo</i> servidor. En libros se gastó los 15,000 pesos solicitados. Una lista completa de los títulos adquiridos está dada en sección 5.4.2. También se obtuvo acceso a las revistas SIAM, Science y dos revistas de la APS.

Continúa en la siguiente página...

Cuadro 2.1 – Continuo

Rubro	Apoyado	Ejercido	Saldo	Justificación
				Los <i>materiales de oficina</i> adquiridos incluyen el papel y los cartuchos de una impresora blanco-negro láser y cartuchos de color y negros para inkjet. Como los estudiantes todavía no terminaron sus trabajos de tesis, no ha sido posible aprovechar de los fondos para la encuadernación de las tesis. También ha sido necesario comprar cables para los equipos que llegaron sin los adaptères y el cableado requerido para su instalación y uso.

Sección 3

Resultados obtenidos con la beca

El becario originalmente propuesto, Gabriela Chamorro Sotelo, obtuvo un empleo de tiempo completo en Gamesa antes de la recepción de la beca, por lo cual se tuvo que buscar otro becario entre los mejores alumnos de licenciatura de la facultad. Afortunadamente se encontró el candidato perfecto ya en noviembre de 2007 que nos permitió empezar a trabajar con el becario desde el inicio del proyecto.

Se otorgó la beca a Carlos Alberto Castillo Salazar, estudiante de licenciatura del noveno semestre de la carrera Ing. Mecánico Electricista en la Facultad de Ingeniería Mecánica y Eléctrica de la UANL. Él realizó sus prácticas profesionales en el proyecto y ya lleva un buen avance en su tesis. Oficialmente se le registrará como tesista en enero 2009 y continuará colaborando en el proyecto durante su segundo año.

Durante su estancia de trabajo en el 2008, Sr. Castillo ha estado trabajando en un proyecto de tesis acerca de redes sensoras inalámbricas. La propuesta consiste en implementar un algoritmo que tome en cuenta el costo de reemplazo de los nodos sensores que intervienen en la red sensora inalámbrica, para decidir la ruta de envío de la información.

Es decir, si un nodo sensor detecta un fenómeno, deberá decidir hacia que nodo sensor enviar esa información para hacerla llegar al usuario, además de tomar en cuenta la energía restante de dicho nodo, deberá tomar en cuenta el costo de reemplazo del mismo, ésto con el fin de incrementar la vida útil de los nodos sensores que tengan un alto costo de reemplazo, aunque esto implique reducir la vida útil de los nodos sensores con menor costo de reemplazo. La propuesta tiene como objetivo el reducir costos en el mantenimiento de la red.

Al inicio de la estancia de trabajo Sr. Castillo conocía en absoluto las redes sensoras inalámbricas. Debido a ésto, durante los primeros meses trabajó en recopilar información sobre el tema, con el objetivo de conocer las aplicaciones, problemas y puntos de oportunidad existentes. En la literatura revisada se encuentra una gran cantidad

de algoritmos conscientes del ahorro de energía en las redes, debido a que es un problema primordial, sin embargo, existen pocas o nulas aportaciones tomando en cuenta el costo de reemplazo de los nodos, es por eso que se decidió trabajar en éste tema.

En este momento se está trabajando en el modelo matemático, la parte experimental, simulación de terrenos y el despliegue de nodos sensores en ellos. Se busca en un futuro poder implementar el algoritmo, realizar simulaciones y evaluar si la propuesta es realmente viable y aplicable y de ser así, descubrir para qué condiciones de terrenos o topologías de redes se pudiese aplicar.

Sección 4

Productos académicos obtenidos

Los artículos, carteles y ponencias del proyecto durante su primer año incluyen los siguientes:

1. Bustos, Bersano, Schaeffer, Piquer, Iosup & Ciuffoletti: "Estimating the size of peer-to-peer networks using Lambert's W function". En *Grid Computing - Achievements and Prospects*, Berlin-Heidelberg, 2008. Springer Verlag. Presentado por Bustos en CoreGRID Integration Workshop 2008, Creta, Grecia. Abril 2008.
2. Avalos & Schaeffer: el cartel "Agrupamiento local de grafos dirigidos", en la XVIII Escuela de Optimización y Análisis (ENOAN 2008), Saltillo, Coahuila. Abril 2008. Ganador del concurso de mejor cartel.
3. Cantú & Schaeffer: el cartel "Optimización de métricas estructurales de grafos" en el ENOAN 2008, Saltillo, Coahuila. Abril 2008.
4. Orponen, Schaeffer & Avalos: "Locally computable approximations of absorption times for graph clustering", Congreso Internacional de la Sociedad de Álgebra Lineal ILAS 2008 en Cancún, Quintana Roo, México. Junio 2008.
5. Schaeffer. "Computación aleatorizada — probabilidad y algoritmos", conferencia magistral del Segundo Congreso Internacional de Sistemas Computacionales y Electrónicos CISCE 2008 en Ciudad de México, D.F., México. Septiembre 2008.
6. Schaeffer. "Retos del cómputo moderno - paralelo, distribuido e integrado", conferencia magistral del Cuarto Congreso Internacional de las Tendencias Tecnológicas en Computación CIDETEC 2008 en Ciudad de México, D.F., México. Programado para 15 de octubre del 2008.

7. Cruz, Gómez, Aguirre, Schaeffer, Turrubiates, Ortega & Fraire. **NAS Algorithm for Semantic Query Routing Systems in Complex Networks**. En *Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence (DECAI 2008)*. Aceptado para publicación. Springer.
8. Cruz, Gómez, Aguirre, Schaeffer, Turrubiates & Ortega. **Comparative Study of Search Processes in Algorithms for Semantic Query Routing Systems in Complex Networks**. En *Proceedings of the Fifteenth International Conference on Advanced Computer Systems, Artificial Intelligence, Software Technologies, Biometrics, and Information Technology Security; International Conference on Advanced Computer Systems (ACS-AISBIS 2008)*. Aceptado para publicación.
9. Avalos & Schaeffer, "Agrupamiento local de grafos dirigidos". Reporte de tesis aceptado para su presentación oral en el XLI Congreso Nacional de la Sociedad Matemática Mexicana SMM 2008.
10. Cantú & Schaeffer, "Análisis y optimización estructural de redes complejas", artículo aceptado a las memorias del Congreso Internacional sobre Innovación y Desarrollo Tecnológico CIINDET 2008.

Los trabajos de tesis del proyecto llevan el siguiente grado avance cada uno:

Tesista	Tema	Nivel	Avance
Marco Antonio Aguirre Lam	Algoritmos de reconocimiento de patrones que guían la búsqueda de información en redes complejas	Maestría	90%
David Juvencio Ríos Soria	Búsqueda dinámica de recursos para el balanceo de carga en redes P2P	Maestría	70%
Perla Elizabeth Cantú Cerda	Optimización estructural de redes complejas	Maestría	50%
Vanesa Avalos Gaytán	Agrupamiento local de grafos dirigidos	Maestría	80%
Carlos Alberto Castillo Salazar	Ruteo dinámico en redes sensoras inalámbricas para minimizar costo de reemplazo	Licenciatura	50%

Involucremos dos estudiantes de Verano Científico de nivel licenciatura, Mario Rivera Ramírez (de la Universidad de Guadalajara) y Jesús Manuel Lopéz Barrios (de la Universidad Autónoma de Sinaloa) quienes ambos están preparando un artículo científico para una revista nacional sobre sus trabajos. Rivera trabajó con Schaeffer y Avalos sobre un tema de caminatas aleatorias y su trabajo se enviará a la revista Ciencias UANL, mientras López trabajó con Schaeffer y Cantú sobre búsqueda local y propiedades estructurales; su trabajo se publicará en una revista de la Universidad de Sinaloa. También recibimos tres alumnos de licenciatura a través del programa PROVERICYT de la UANL.

Sección 5

Protocolo para el segundo año

Estudiamos redes dinámicas de tamaño masivo, como los sistemas entre pares y súpercomputadoras formadas por numerosos equipos pequeños que colaboren para lograr coordinación entre ellos. Nuestro enfoque está en los algoritmos de búsqueda distribuida de comportamiento escalable y dinámica que no sufren del crecimiento del sistema ni de cambios de topología. La importancia de tal estudio proviene del papel importante que ya han obtenido redes de información y comunicación en la sociedad moderna. La relevancia del proyecto está en diseñar soluciones capaces de reaccionar a patrones de uso del sistema y así evolucionar según los cambios de comportamiento y intereses de los usuarios. Queremos crear algoritmos adaptivos que aprendan del tráfico en la red y utilicen sus observaciones para mejorar la calidad de servicio a los búsquedas futuras. Utilizamos los descubrimientos de la última década sobre redes naturales y sistemas complejos para desarrollar métodos refinados en especial para instancias reales. El trabajo será analítico y experimental. El grupo de trabajo ya tiene colaboración establecida sobre el tema y fuertes conocimientos del diseño y análisis de algoritmos, experimentos computacionales y simulación.

5.1. Introducción y antecedentes científicos

Sistemas de comunicación ya forman una parte fundamental de la función de sociedad y industria moderna. Su tamaño está en crecimiento constante y la complejidad de tales sistemas posee desafíos grandes con respeto al mantenimiento, escalabilidad y desarrollo.

Nuestro interés está en redes masivas donde los *nodos* participantes, que típicamente son equipos de computación de capacidad diferente, comparten recursos — sea contenido o recursos computacionales — para fines colaborativos. Por ejemplo, se puede juntar una gran cantidad de computadoras personales para funcionar como una

súpercomputadora, coordinando el uso de sus recursos (es decir, procesadores y memoria) entre los nodos. Las *conexiones* que facilitan la comunicación del sistema son típicamente conexiones entre dos nodos así que para lograr comunicación entre nodos no directamente conectados uno al otro, hay que propagar mensajes a través de los nodos intermedios. Los nodos presentes y las conexiones entre ellos forman la *topología* del sistema, es decir, una red de comunicación.

Cuando el tamaño, en términos del número de nodos participantes, está grande, ya no es factible manejar la coordinación de la comunicación entre los nodos de una manera centralizada. La cantidad de información que debería ser guardada y procesada para poder tomar decisiones globalmente óptimas es demasiado grande para lograr funcionamiento eficiente. La reacción adecuada a la casi-imposibilidad de computación global centralizada es el uso de métodos *distribuidos* de computación local y algoritmos de *aproximación*.

Complejidad adicional surge de la naturaleza *dinámica* de tales sistemas: en cualquier momento, uno de los nodos puede ocuparse con un usuario local, sin poder continuar colaborando en la tarea distribuida. También nuevos nodos pueden entrar a la colaboración. Este tipo de cambio de la topología de las conexiones entre los nodos resulta en "invalidación" rápida de observaciones estructurales hechas de la red: lo que en un momento estaba comportamiento óptimo, puede que ya ni siquiera es factible en otro momento.

Además del uso de computadoras normales como una súpercomputadora, el mismo planteamiento repite en los sistemas *entre pares* (inglés: peer-to-peer networks) donde los nodos participantes contienen alguna información (por ejemplo, documentos de texto o música) y pueden realizar *búsquedas* para ubicar cierto contenido de interés, y en el caso que esté presente el contenido deseado, descargarlo de un nodo que ya lo posee [3, 8, 17, 19]. En general, una búsqueda en una red es el proceso de ubicar de una manera eficiente un nodo de *destino* desde un nodo de *origen*. Con eficiente se refiere a poder mantener las cantidades de tiempo de ejecución, memoria consumida y mensajes de comunicación entre los nodos en niveles moderados, preferiblemente minimizando por lo menos uno de estos tres factores.

Los nodos no necesariamente son computadoras conectadas a la Internet, sino también pueden ser terminales móviles (como teléfonos celulares) que forman una red *ad hoc* [10] a través de comunicación por radio (por ejemplo por Bluetooth). La necesidad de buscar por un nodo con cierto recurso, contenido o identidad es central también para la función de sistemas de *redes sensoras* donde una cantidad posiblemente enorme de elementos computacionales muy pequeños están en cargo de vigilar una área por un fenómeno específico para detectar y reportar eventos de interés.

5.1.1. Planteamiento del problema

Modelamos las redes de comunicación como *grafos*. Un grafo $G = (V, E)$ es una estructura de matemáticas discretas donde V es un conjunto de *vértices* y E es un conjunto de *aristas*. Los vértices representan los nodos — un vértice por nodo — y las aristas representan las conexiones — una arista (v, w) es un par de vértices, por lo cual si (v, w) está en el conjunto E , los nodos representados por los vértices v y w pueden comunicarse directamente uno al otro. El número de vértices $|V|$ se denota con n y el número de aristas $|E|$ con m .

El grafo está formada dinámicamente en tiempo así que vértices pueden entrar al grafo y partir del grafo. Cuando entra un vértice, puede formar aristas con los vértices ya presentes de una manera predeterminada que depende del sistema de mundo real estudiado. Cuando un vértice sale del grafo, desaparecen también todas sus aristas. Un vértice v es un *vecino* de otro vértice w si y sólo si la arista (v, w) pertenece al grafo. El número de vecinos que tiene un vértice se llama su *grado*. Un vértice v está conectado al otro vértice w , si en el grafo existe un *camino* de aristas por lo cual es posible "mover" desde v hasta w siguiendo las aristas. La *distancia* entre dos vértices es el número de aristas en el camino más corto que los conecta en el grafo. Si todos los vértices están conectados a todos los otros vértices, el grafo es *conexo*. La figura 5.1 ilustra estos conceptos.

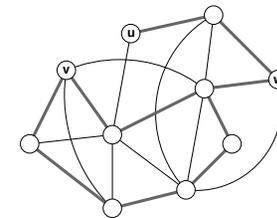


Figura 5.1: Un grafo de ejemplo: los vértices son los círculos y las aristas las líneas conectando un círculo al otro. Se muestra (en línea gruesa gris) dos posibles caminos desde el vértice v hasta el vértice w y un camino de w a un tercer vértice u . Las aristas que no forman parte de ninguno de los tres caminos están dibujadas en línea fina negra. En el ejemplo, $n = 8$ y $m = 20$ y el grafo es conexo. La distancia entre v y w es tres, entre v y u es dos y entre w y u es también dos.

Un vértice puede comunicar con cada vecino suyo por enviar un paquete por la arista que los conecta. Para comunicar a un vértice que no es su vecino, tienen que enviar el paquete primero a un vecino quien lo va a reenviar a otro vértice. Esto se repite hasta que el paquete llega a su destino. Para poder realizar el reenvío, es necesario que el paquete contenga una *identificación* del vértice de origen del mensaje

y también una identificación del vértice de destino. Esas identificaciones se llaman las *direcciones* de los vértices. El proceso de elegir a dónde un vértice debería reenviar un paquete destino de que no es el vértice mismo se llama *ruteo* y el camino recorrido por un paquete es una *ruta* de comunicación.

Una *tabla de ruteo* completa contiene información sobre todos los caminos de un grafo. Si cada vértice tiene acceso a una tabla de ruteo completa, tienen la información exacta para poder comunicar con cada otro vértice. Los vértices simplemente consultan en la tabla de ruteo por el camino más corto de ellos mismos al vértice destino del paquete. Si el grafo es muy grande o la capacidad de almacenar información de los vértices es limitada, no es posible manejar una tabla de ruteo completa. Si cada vértice solamente conoce una parte de la topología del grafo, es necesario incluir información (parcial) de la ruta en el paquete mismo, por si acaso es algún vértice intermedio no conoce la ruta. Además, cuando un vértice recibe un paquete con información de ruta incompleta y el vértice tampoco conoce la ruta, el vértice tiene que contar con un algoritmo de ruteo para poder elegir entre sus vecinos a cuál reenviar el paquete. Los algoritmos de ruteo juntos con los algoritmos de mantenimiento de una tabla de ruteo y el formato de los paquetes enviados forman un *protocolo de ruteo*.

Ruteo en grafos pequeños y estáticos es un problema ya resuelto. En este proyecto, queremos atacar el ruteo en grafos masivos y dinámicos. Con “masivo” queremos decir que los grafos son tan grandes que desde el punto de vista de un vértice son prácticamente infinitos: la fracción de vértices que son vecinos de un vértice es en promedio casi cero. Una restricción especial que planteamos se trata de la *capacidad computacional* de los vértices: cada vértice tiene un límite para cuánta información topológica puede “recordar” en su tabla de ruteo. Además el tamaño de los paquetes está limitado: no es posible incluir más que k unidades de información en un sólo paquete.

Entonces, el problema es qué información incluir en las tablas de ruteo junto con la pregunta qué información debería ser incluido en un paquete para poder lograr un tiempo ruteo casi óptimo. Estas preguntas no son fáciles de contestar en grafos dinámicos, porque los cambios constantes de topología causan partes de la información guardada volver inválida rápidamente. La idea es no enviar paquetes de control para averiguar el estado del grafo, sino adjuntar información en los paquetes que están relacionados a los pedidos y su servicio siempre y cuando los paquetes no están ocupando todas las k unidades de información disponibles. Así los vértices podrían propagar información que consideren esencial a los otros vértices que van a procesar el mismo paquete. Por permitir los vértices intermedios modificar esos contenidos adicionales de los paquetes podemos lograr comunicación más eficiente en el grafo.

5.1.2. Trabajo existente y áreas de oportunidad

En forma centralizada con información completa, el problema de *búsqueda* en grafos es simplemente la construcción de caminos de largo mínimo [9] que ya cuenta con varios algoritmos eficientes para diferentes variaciones del problema. Para grafos de tamaño masivo, ese tipo de procedimiento no es aplicable por el tamaño del grafo, ya que son algoritmos polinomiales en el tamaño del grafo — lo que nos interesa a nosotros es el diseño de métodos con complejidad que *no* dependa del grafo completo, sino del vecindario local de los vértices involucrados.

Cuando no toda la información de ruteo está disponible, uno típicamente tiene que hacer algún recorrido tipo *búsqueda en anchura* (inglés: breadth-first search) para encontrar un camino del origen al destino. En muchas aplicaciones ha resultado beneficioso permitir que la búsqueda progresa *en paralelo* ramificándose en las diferentes partes de la red, aunque esto puede resultar en congestión u otros problemas [5]. Ramificación a todas las direcciones se llama *inundación* (inglés: flooding), que en su forma más sencillo causa que cada mensaje propaga por cada arista. Unas variaciones probabilísticas y limitadas ofrecen resultados buenos [6, 13, 26]. Por lo general, el estudio y simulación de tales sistemas es más realístico en sistemas multiprocesador.

Las técnicas de *búsqueda local* están fuertemente relacionadas con búsqueda de recursos en redes masivos, porque son métodos para ubicar soluciones buenas en espacios enormes típicamente representadas a través de grafos. En nuestro problema, sin embargo, no existe una función objetivo natural para utilizar dirigir la búsqueda.

La comunidad científica que estudia redes ha comenzado en los últimos años a tomar cuenta de los efectos de *propiedades estructurales* de las redes mismas en el comportamiento de algoritmos. Tales propiedades incluyen por ejemplo los diámetros sorprendentemente pequeños en comparación con el tamaño del grafo [2]. Una medida propuesta por Sneppen *et al.* [20, 22] mide la facilidad de búsqueda de un grafo para clasificar y ajustar la topología con el fin de optimizar la realización de búsquedas en un grafo.

Se ha llegado a mostrar que la topología de una red puede tener un impacto significativo en el tiempo de búsqueda [11]. Adamic *et al.* [1] y Kim *et al.* [15] han estudiado búsquedas en grafos con *distribución de grado* especial, comúnmente visto en redes del mundo real. Ciertos tipos de topologías (en especial redes de *pequeño mundo* (inglés: small-world networks) minimizan el tiempo de búsqueda en unos escenarios típicos [11]. Kleinberg [16] muestra un modelo de construcción de redes de pequeño mundo que optimiza un cierto tipo de búsqueda utilizando *coordenadas* de los vértices. Búsqueda con información local en redes de pequeño mundo ha sido estudiado también por Zhu y Huang [28]. Métodos para “aumentar” grafos para que sean localmente navegables para búsquedas mejores fueron estudiado por Duchon *et al.* [12]. Un factor adicional importante es la tasa entre el costo de *seguir* una arista

y el costo de *averiguar* la existencia o largo de una arista [23].

En la ausencia total de información a priori, la problema de búsqueda es equivalente al comportamiento de un *camino aleatorio* (inglés: random walk) en el grafo. Los estudios de propiedades de caminos aleatorios en grafos de ciertas propiedades estructurales incluyen los de Annibaldi y Hopcraft [4] y Tadić [24, 25]. Propiedades de caminos aleatorios no cruzantes (inglés: self-avoiding random walk) fueron estudiados por Herrero [14] y Yang [27]. Típicamente el estudio de caminos aleatorios es fuertemente representado en el campo de física matemática.

Está establecido que un camino aleatorio con información de los vecinos *secundarios* (o sea, los vecinos de los vecinos) resulta en búsqueda ávida óptima de rutas [18]. En redes masivas del mundo real, sin embargo, el número de vecinos secundarios puede ser muy grande, especialmente cuando la topología es *libre de escala* (inglés: scale-free) — y casi siempre lo es en las redes del mundo real [7].

Lo que se propone es una continuación de algunas ideas presentadas en trabajo anterior [21] para estudiar *qué* información deberían guardar los nodos de una red localmente para que sea óptimo el comportamiento de una búsqueda en una red de tamaño masivo. Para evitar tráfico de control, queremos incluir información de ruteo adicional es cada paquete de datos que recorre en la red, llenando la capacidad del paquete según el estado actual de la “tabla local” de los nodos que procesan el paquete. Nuestro interés es en capacitar la red para cambios continuos de topología por llegadas y salidas de los nodos participantes y cambios en su contenido. También la generación de tráfico es dinámica y requiere reacciones proactivas del sistema.

5.2. Objetivos y metas

Este proyecto tiene objetivos generales y metas específicas tal en la área de desarrollo científico como en la área de formación de recursos humanos a través de dirección de estudiantes y de colaboración entre los científicos que toman parte en el proyecto, aprendiendo uno del otro sobre sus campos de especialidad y juntando los conocimientos diversos en crear algo de mayor aplicabilidad. Es muy importante para el grupo de trabajo además de producir resultados de alta calidad, divulgar los resultados obtenidos en foros nacionales e internacionales. Queremos enseñar a nuestros alumnos desde el comienzo a publicar en castellano y en inglés su trabajo, porque las publicaciones nos permiten encontrar aplicaciones nuevas a través de contactos nuevos y recibir retroalimentación de la comunidad científica que nos permite mejorar en el futuro.

- **Investigación:** Efectuar investigación básica de problemas de búsqueda en grafos masivos a fin de lograr un mayor entendimiento y técnicas de solución más

eficientes y matemáticamente justificadas. Avanzar significativamente el estado del arte en la área de búsqueda de recursos para resolver en paralelo o con aproximaciones las problemas de optimización combinatoria relacionados, permaneciendo a la vanguardia en dichas líneas de investigación.

- **Colaboración conjunta:** Realizar trabajo conjunto con investigadores en nivel nacional e internacional en áreas afines con la finalidad de fomentar los lazos de colaboración de nuestro programa con otros programas de reconocido nivel. Fortalecer las conexiones entre los participantes y crear conexiones directas entre sus contactos.
- **Formación de recursos humanos de alto nivel:** Involucrar a estudiantes, cuyos trabajos de tesis se enfocarán a tratar tareas específicas del proyecto y que serán dirigidos por Dra. Schaeffer y co-dirigidos por otros investigadores que participan 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.

Las metas específicas para el segundo año son las siguientes:

- **Investigación:** Tenemos un entendimiento de la estructura matemática del problema planteado que nos ha permitido poder desarrollar algunas técnicas de solución que exploten a ésta favorablemente. Estamos en la posición de estudiar variaciones aún más complejas con movilidad de nodos. Hemos desarrollado algoritmos de búsqueda para grafos masivos y dinámicos. Hemos desarrollado e implementado métodos de búsqueda en redes de tamaño masivo con propiedades estructurales de redes del mundo real.
- **Formación de recursos humanos:** Hemos dirigido a cuatro estudiantes de posgrado y uno de licenciatura. Hemos colaborado en la formación de un estudiante de doctorado de la Dra. Cruz, M.C. Claudia Guadalupe Gómez Santillán. En el año 2009, con la Dra. Schaeffer se incorporan Lic. Nidia L. Gómez Duarte como tesista de maestría y M.C. Tania Turrubiates López como tesista de doctorado. Sr. Carlos Alberto Castillo Salazar sigue como tesista de licenciatura. También colaborarán las doctoras Schaeffer y Cruz en la formación de maestros en el ITCM.
- **Publicación y diseminación de resultados:** Hemos expuesto resultados parciales y finales en un foro internacional y dos foros nacionales. Hemos publicado varios artículos en memorias de congresos. Estamos por enviar para publicación dos artículos a revistas arbitradas e indexadas. Material didáctico introductorio a los antecedentes ha sido publicado en línea.

5.3. Metodología

Nuestra metodología consiste en seis pasos básicos:

1. **un entendimiento matemático** de la estructura de las redes naturales masivas en las cuales se realiza búsquedas por experimentación y análisis,
2. **las definiciones** de medidas de utilidad de la información estructural local para poder determinar qué información preferir al momento de poder añadir información a un paquete de comunicación y para poder determinar qué información de un paquete recibido debería ser guardado en la memoria del vértice que lo procesa;
3. **el diseño de algoritmos** de ruteo dinámico para redes masivas;
4. **el desarrollo de protocolos** de comunicación realísticos para poder aplicar los algoritmos desarrollados en redes de comunicación alámbricas e inalámbricas;
5. **el análisis teórica** de las propiedades de los algoritmos y protocolos desarrollados;
6. **la amplia experimentación aplicada** del comportamiento práctico de los algoritmos y protocolos desarrollados en sistema multiprocesador con computación paralela.

En cada uno de estos pasos, cuidamos especialmente la escalabilidad de los métodos y medidas para grafos de tamaño prácticamente infinito. También buscamos asegurar que exploremos 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.

El comienzo el trabajo fue enfocado hacia un *estudio matemático* del problema para poder establecer propiedades y caracterizaciones nuevas y extensiones del trabajo existente. Después nos enfocamos a experimentar con las medidas y definiciones formuladas en la primera etapa y para diseñar conceptualmente los algoritmos de búsqueda en redes dinámicas. La tercera etapa consistió en desarrollar e implementar computacionalmente los algoritmos diseñados en nivel de prototipos. Hemos evaluado a través de experimentación los métodos desarrollados, realizando una serie de comparaciones sistemáticas con algoritmos de presentados en la literatura con anterioridad. Los experimentos se ha realizado utilizando tanto instancias de datos reales como datos artificiales.

5.4. Infraestructura 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. En esta sección se detalla la infraestructura actualmente disponible en el Programa Doctoral de la Facultad de Ingeniería Mecánica y Eléctrica para la realización del proyecto.

5.4.1. Equipo de cómputo

Las computadoras del laboratorio de estudiantes consiste en computadoras equipadas cada una con un procesador Pentium 2,66 GHz, 512 KB de memoria y un disco duro de 40 GB. El laboratorio de alto desempeño de PISIS cuenta con un servidor Sun-Fire V440, y unos veinte terminales gráficas que todos comparten los recursos del servidor. Todo el equipo de cómputo se encuentra conectado a la red de la UANL.

El grupo de trabajo ahora dispone de un servidor SSH de experimentos de cuatro núcleos y 4 GB de memoria RAM, dos computadoras de escritorio (iMac de Dra. Schaeffer y MacMini de sus estudiantes ambos con OS X) y una computadora portátil tablet PC para compartir entre los integrantes del grupo en la UANL. Adicionalmente tenemos en nuestro uso un servidor del Dr. Berrones que funciona como el servidor de web. Ambos servidores tienen el sistema operativo Ubuntu Linux 8.04.

Dos oficinas adicionales de Dra. Schaeffer, quien fue nombrado en mayo 2008 coordinador del área de Tecnología de Información & Software en el Centro de Innovación, Investigación y Desarrollo en Ingeniería y Tecnología (CIIDIT) de la UANL, en adición a su puesto principal de profesor investigador en el Posgrado de Ingeniería de Sistemas en la Facultad de Ingeniería Mecánica y Eléctrica de la misma, están equipados con computadoras desktop básicas.

5.4.2. 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 contempló la adquisición de textos en las áreas de optimización combinatoria, teoría de grafos y minería de datos. En el primer año, obtuvimos los siguientes títulos:

- The Elements of Style; Strunk Jr.
- Matrix Algebra Useful for Statistics; Searle
- Genetic Algorithms in Search, Optimization, and Machine Learning; Goldberg (2)
- Fundamentals of Wireless Communication; Tse & Viswanath
- C Programming Language; Kernighan & Ritchie

- UNIX for the Impatient; Abrahams & Larson
- Art of Computer Programming, Volumes 1-3; Knuth
- Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals; Batty
- Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence; Holland
- Art of Computer Programming, Volume 4, Fascicle 4: Generating All Trees — History of Combinatorial Generation; Knuth
- Evolutionary Dynamics: Exploring the Equations of Life; Nowak
- Introduction to Linear Optimization; Bertsimas & Tsitsiklis
- The Dynamics of Complex Urban Systems: An Interdisciplinary Approach; Albeverio et al.
- Weak Links: Stabilizers of Complex Systems from Proteins to Social Networks; Csermely
- Elements of Distributed Computing; Garg
- Protocols and Architectures for Wireless Sensor Networks; Karl & Willig
- Combinatorial Optimization: Networks and Matroids; Lawler
- An Introduction to Communication Network Analysis; Kesidis
- Evolution and Structure of the Internet: A Statistical Physics Approach; Pastor-Satorras & Vespignani
- Networks: Optimisation and Evolution; Whittle
- Random Graph Dynamics; Durrett
- Algorithmic Game Theory; Nisan, et al.
- Control Techniques for Complex Networks; Meyn
- Complex Graphs and Networks; Fan Chung & Lu
- How to Solve It: Modern Heuristics; Michalewicz & Fogel
- Handbook of Graphs and Networks: From the Genome to the Internet; Bornholdt & Schuster
- Ad Hoc & Sensor Networks: Theory And Applications; De Moraes & Agrawal
- Evolution of Networks: From Biological Nets to the Internet and WWW; Dorogovtsev & Mendes
- Random Geometric Graphs; Penrose

- Topological Structure and Analysis of Interconnection Networks; Xu
- Handbook of Discrete and Combinatorial Mathematics; Rosen
- Algebraic Graph Theory; Godsil & Royle
- Digraphs; Bang-Jensen & Gutin
- Intuitive Probability and Random Processes using MATLAB; Kay
- Teorías del Aprendizaje; Schunk
- Aprendizaje humano; Ormrod

Publicaciones periódicas reciente a las cuales obtuvimos acceso son todas las revistas de SIAM (en forma electrónica), la revista Science del AAAS (impresa y electrónica) y dos revistas del APS: Physical Review Y y Physical Review Letters (acceso electrónico). Como la gama de revistas disponibles vía suscripción institucional en la UANL está mejorando, se estima que los costos de acceso a revistas serán modestos.

5.5. Incidencia del proyecto en el PIFI

El Programa Integral de Fortalecimiento Institucional (PIFI) apoya al cuerpo académico del Programa de Posgrado de Ingeniería en Sistemas (PISIS). Por el recién ingreso de la Dra. Schaeffer a la UANL, todavía no está oficialmente miembro un cuerpo académico, pero ya colabora con la LGAC de *Modelaje, análisis y solución de sistemas determinísticos* y con el proyecto propuesto va a poder contribuir en mayor cantidad también a la otra línea de investigación del cuerpo académico sobre *sistemas estocásticos*. En la facultad se está preparando para la formación de un cuerpo académico nuevo en tecnologías de la información, el líder del cual será Dra. Schaeffer.

5.6. Formación de recursos humanos

Se busca involucrar más **estudiantes de licenciatura** en el proyecto como becarios para desarrollar sus tesis de licenciatura y para motivarles que continúen sus estudios en un programa de posgrado. Este meta coincide con el programa de honores que se implementa en PISIS en el nivel de la FIME para incluir estudiantes sobresalientes de licenciatura en la investigación. Aparte se involucra estudiantes de licenciatura a través de los programas de verano científico.

Unos **estudiantes de maestría** serán involucrados a través de los trabajos de tesis de maestría; actualmente ya trabajan en el tema cuatro estudiantes. Contemplamos la incorporación de 2 a 3 estudiantes nuevos de maestría para 2008–2009.

Por lo menos un *estudiante de doctorado* comenzará su trabajo en el primer semestre de 2009; al M.C. Arias Torres, postulante contemplado, se le atrasó su ingreso al doctorado por compromisos laborales y por asuntos familiares. Actualmente está preparándose la M.C. Tania Turrubiates López el anteproyecto (relacionado con este proyecto) para su admisión en la convocatoria actual del programa doctoral.

El becario originalmente contemplado, Gabriela Chamorro Sotelo, fue contratada a trabajar en una empresa y consecuentemente reemplazado por *Carlos Alberto Castillo Salazar*, un estudiante de **licenciatura** cursando el noveno semestre de la carrera IME en la Facultad de Ingeniería Mecánica y Eléctrica de la UANL.

5.7. Vinculación y colaboración

5.7.1. Participantes en la UANL

Una meta importante del PIFI del Programa Posgrado en Ingeniería de Sistemas (PISIS) es aumentar el liderazgo académico del grupo.

Ahora ya la Dra. Satu Elisa Schaeffer es coordinadora de un grupo de investigación y tiene proyectos de investigación. Está preparando propuestas de proyectos para CONACyT, entre otros.

Este proyecto propuesto ha sido uno de los primeros pasos de la formación del área de investigación en las tecnologías de información y software del CIIDIT, UANL. Se está formando en la FIME un cuerpo académico nuevo, el líder del cual será la Dra. Schaeffer, con dos LGACs relacionadas con la ingeniería de software.

Dra. Schaeffer tiene su doctorado en Teoría de Computación y sus últimas cinco publicaciones son

- Cruz, Gómez, Aguirre, Schaeffer, Turrubiates, Ortega y Fraire. **NAS Algorithm for Semantic Query Routing Systems in Complex Networks**. En *Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence (DECAI 2008)*. Aceptado para publicación. Springer.
- Cruz, Gómez, Aguirre, Schaeffer, Turrubiates y Ortega. **Comparative Study of Search Processes in Algorithms for Semantic Query Routing Systems in Complex Networks**. En *Proceedings of the Fifteenth International Conference on Advanced Computer Systems, Artificial Intelligence, Software Technologies, Biometrics, and Information Technology Security; International Conference on Advanced Computer Systems (ACS-AISBIS 2008)*. Aceptado para publicación.
- López, Litvinchev, Mata y Schaeffer. **R&D Portfolio Selection MILP Models for Large Instances in Public and Social Sectors**. *Computación y Sistemas*, Special Issue, aceptado

para publicación.

- Bustos, Bersano, Schaeffer, Piquer, Iosup y Ciuffoletti. **Estimating the size of peer-to-peer networks using Lambert's W function**. En Golatch, Fragopoulou, and Priol, editores, *Grid Computing - Achievements and Prospects*, Berlin-Heidelberg, 2008. Springer Verlag.
- Schaeffer. *Graph clustering*. *Computer Science Review* (1):1, pp. 27–64, 2007.

Del cuerpo académico consolidado de PISIS participa en el proyecto el *Dr. José Arturo Berrones Santos*. Su conocimiento de caminos aleatorios será muy importante para el proyecto y será miembro de todos los comités de tesis del proyecto en PISIS. Tiene su doctorado en Física y sus últimas cinco publicaciones son

- Berrones, *Stationary probability density of stochastic search processes in global optimization*, *J. Stat. Mech.* (2008) P01013.
- Berrones, Peña y Sánchez, **Stationary Probability Density of Stochastic Search Processes**, en *Encyclopedia of Artificial Intelligence*.
- Peña, Sánchez y Berrones, **Stationary Fokker–Planck Learning for the Optimization of Parameters in Nonlinear Models**, *Proceedings of MICAI*, Lecture Notes in Computer Science 4827, pp. 94–104, 2007.
- A. Berrones, **Generating Random Deviates Consistent with the Long Term Behavior of Stochastic Search Processes in Global Optimization**. En *Proceedings of IWANN 2007*, Lecture Notes in Computer Science, Vol. 4507, pp. 1–7. Springer, Berlín, Alemania 2007.
- A. Berrones, **Filtering by Sparsely Connected Networks Under the Presence of Strong Additive Noise**. En *Proceedings of the Seventh Mexican International Conference on Computer Science*, pp. 19–26. IEEE 2006.

5.7.2. Participantes externos a la UANL

Otra meta de importancia del PIFI es fomentar la vinculación con cuerpos académicos de otras instituciones. Dra. Schaeffer ya cuenta con buenas conexiones internacionales y colabora desde casi dos años sobre los temas del proyecto propuesto con un cuerpo académico nacional de otra institución en adición a colaborar fuertemente con el cuerpo académico de PISIS. Tiene en preparación la formación de un cuerpo académico en TI & Software en FIME para trabajos relacionados con las líneas de investigación que desarrolla su grupo en el CIIDIT de la UANL.

Helsinki University of Technology, Espoo, Finlandia En la Universidad Politécnica de Helsinki, Finlandia, trabaja un grupo de investigadores en el Laboratorio para Teoría de Computación. El líder del grupo *Dr. Olli Pekka Orponen* era el asesor del doctorado de la Dra. Schaeffer y su colaboración continua en varios temas de teoría de grafos y sus aplicaciones. Los otros investigadores del mismo grupo con quienes se colabora principalmente son *Dr. Harri Haanpää* y *M.C. André Schumacher*. El *Dr. Orponen* tiene su doctorado en Teoría de Computación y sus últimas cinco publicaciones son

- Schumacher, Orponen, Thaler y Haanpää, **Lifetime maximisation in wireless sensor networks by distributed binary search**. En *Proceedings of the 5th European Conference on Wireless Sensor Networks (EWSN'08)*, Lecture Notes in Computer Science, Berlin Heidelberg, to appear. Springer-Verlag.
- Haanpää, Schumacher, Thaler y Orponen, **Distributed computation of maximum lifetime spanning subgraphs in sensor networks**. En Zhang, et al., editores, *Proceedings of the 3rd International Conference on Mobile Ad-Hoc and Sensor Networks (MSN'07)*, volumen 4864 de Lecture Notes in Computer Science, pp. 445—456, Berlin Heidelberg, 2007. Springer-Verlag.
- Prasad, Schumacher, Haanpää y Orponen. **Balanced Multipath Source Routing**. En *Proceedings of the Twenty-first International Conference on Information Networking*, Lecture Notes in Computer Science. Springer, Berlín, Alemania, 2007.
- Schumacher, Haanpää, Schaeffer y Orponen. **Load Balancing by Distributed Optimization in Ad Hoc Networks**. En J. Cao et al. (editores), *Proceedings of the Second International Conference on Mobile Ad Hoc and Sensor Networks*, Lecture Notes in Computer Science, Vol. 4325, pp. 874—885. Springer, Berlín, Alemania, 2006.
- Florén, Kaski, Kohonen y Orponen. *Exact and Approximate Balanced Data Gathering in Energy-Constrained Sensor Networks*. Theoretical Computer Science, 344(1):30—46. Noviembre 2005.

Instituto Tecnológico de Ciudad Madero, Tamaulipas, México Dra. Schaeffer tiene un convenio de colaboración con el cuerpo académico de *Optimización Inteligente* del Instituto Tecnológico de Ciudad Madero. Es profesor visitante hasta el fin de 2008, con fines de aumentar la colaboración aún más en el futuro a través de dirección de tesis y proyectos conjuntos. El líder del cuerpo académico es la *Dra. Laura Cruz Reyes*, doctora de Ciencias de Computación. Sus últimas cinco publicaciones son

- Cruz, Gómez, Aguirre, Schaeffer, Turrubiates, Ortega y Fraire. **NAS Algorithm for Semantic Query Routing Systems in Complex Networks**. En *Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence (DECAI 2008)*. Aceptado para publicación. Springer.

- Cruz, Gómez, Aguirre, Schaeffer, Turrubiates y Ortega. **Comparative Study of Search Processes in Algorithms for Semantic Query Routing Systems in Complex Networks**. En *Proceedings of the Fifteenth International Conference on Advanced Computer Systems, Artificial Intelligence, Software Technologies, Biometrics, and Information Technology Security; International Conference on Advanced Computer Systems (ACS-AISBIS 2008)*. Aceptado para publicación.
- Cruz, Meza, Turrubiates, Gómez y Ortega. **Statistical Selection of Relevant Features to Classify Random, Scale Free and Exponential Networks**, *Innovations in Hybrid Intelligent Systems, Advances in Soft Computing*, Vol. 44. Springer-Verlag, Berlin Heidelberg New York (2007) pp. 454—461.
- Ortega, Conde, Gómez, Cruz y Turrubiates. **Studying the Impact of Growing Dynamics in the Internet Topology**, *Polish Journal of Environmental Studies* 16(5B): 117—120, 2007.
- Cruz, Meza, Turrubiates, Gómez y Ortega. **Experimental design for selection of characterization functions that allow discriminate among random, scale free and exponential networks**. *Polish Journal of Environmental Studies* 16(5B): 67—71, 2007.

Universidad Diego Portales, Santiago, Chile En el año 2005 Dra. Schaeffer realizó una estancia de investigación de 10 meses en la Universidad de Chile, Santiago, colaborando con el *Dr. Ricardo Baeza Yates*. Durante esa estancia, estableció colaboración con el *Dr. Javier Bustos Jiménez*, actualmente profesor de la UDP.

Dra. Schaeffer fue invitada a la UDP en abril 2007 para fortalecer la colaboración existente y para formar conexiones con los nuevos colegas del *Dr. Bustos*. El motivo de la visita era relacionado con la tecnología de *redes sensoras*, sobre la cual impartió seminarios la Dra. Schaeffer y cual es un campo activo de aplicación del conocimiento que se generará en el proyecto propuesto. La colaboración con los chilenos fortalece otra meta del PIFI de PISIS: incrementar la vinculación con el sector empresarial.

Dr. Bustos tiene su doctorado en Ciencias de Computación y sus últimas cinco publicaciones son

- Bustos, Bersano, Schaeffer, Piquer, Iosup y Ciuffoletti. **Estimating the size of peer-to-peer networks using Lambert's W function**. En Golatch, Fragopoulou, and Priol, editores, *Grid Computing - Achievements and Prospects*, Berlin-Heidelberg, 2008. Springer Verlag.
- J. Bustos Jiménez, C. Varas y J.M. Piquer. **Sub-contracts: Delegating Contracts for Resource Discovery**. En Talia, Yahyapour, Ziegler, editores, *Grid Middleware and Services: Challenges and Solutions*, Springer, New York, USA, pp. 95—104, 2008.
- J. Bustos Jiménez, D. Caromel, M. Leyton y J.M. Piquer. **Load Information Sharing Policies in Communication-Intensive Parallel Applications**. En *Proceedings of Advanced Distributed Systems: 6th International School and Symposium*, Lecture Notes in Computer Science. Springer, Berlín, Alemania, 2006.

- J. Bustos Jiménez, D. Caromel, M. Leyton, A. Iosup y J.M. Piquer. **The Rockchair and the Grid: Balancing load across Project Grids**. En *Proceedings of CoreGRID Integration Workshop*, 2006.
- J. Bustos Jiménez, D. Caromel, M. Leyton y J.M. Piquer. **Coupling Contracts for Deployment on Alien Grids**. En *Proceedings of the International Euro-Par Workshops*, Lecture Notes in Computer Science. Springer, Berlín, Alemania, 2006.

Sección 6

Cronograma de actividades

■ Nov/Dic 2008:

- Adaptación de los métodos de búsqueda para grafos dinámicos.
- Aprovechamiento mayor de estructurales en los métodos búsqueda.
- Finalización de una publicación en una revista indexada de prestigio.
- Finalización y presentación de los resultados de las tesis de cuatro estudiantes de maestría.

■ Primavera de 2009:

- Evaluación teórica de los métodos desarrollados.
- Experimentación extensiva computacional con instancias de tamaño masivo.
- Finalización y presentación de los resultados de la tesis de licenciatura del becario del proyecto.
- Incorporación de un estudiante de doctorado.

■ Verano de 2009:

- Segunda estancia corta de investigación en el extranjero; particularmente en Chile.
- Preparación de publicaciones que incorporan los resultados teóricos en los resultados prácticos ya obtenidos.
- Participación a un congreso internacional.

■ Otoño de 2009:

- Preparación de los informes, reportes y las últimas publicaciones del proyecto.
- Solicitación de fondos de proyectos de continuación de CONACyT.
- Establecimiento de convenios de colaboración con empresas.

Sección 7

Solicitud de apoyos para el segundo año

Cuadro 7.1: Montos solicitados por rubro para 2008-2009.

Rubro	Monto Solicitado	Justificación
Equipo para experimentación	0	No se requiere equipo mayor adicional. Estamos trabajando con la máquina ya obtenida, aunque es algo inferior a la especificación original y esperamos obtener fondos internos para obtener un servidor más potente en el CIIDIT en los próximos meses.
Equipo menor o complementario	14,000	En el primer año, recibimos fondos para una computadora simple tipo escritorio MacMini de las dos solicitadas. Con la incorporación de un estudiante de doctorado nos urge tener otra computadora dedicada para el uso de estudiantes, por lo cual solicitamos nuevamente fondos para una MacMini. El modelo apto para trabajo científico cuesta 10,489 pesos ya con IVA (http://store.apple.com/mx/ ; viene sin periférico ninguno. Adicionalmente se necesita una pantalla suficientemente grande para visualizar simulaciones. El Samsung 920LM cuesta 2,459 pesos más IVA (http://www.pcdomino.com/). El resto del monto solicitado se destina a teclado y ratón.

Continúa en la siguiente página...

Cuadro 7.1 – Continuado

Rubro	Monto Solicitado	Justificación
Asistencia a reuniones académicas	33,500	Esperamos asistir a más congresos en el segundo año. El primer año queríamos incluir un estimado para la asistencia en el presupuesto del segundo año, pero el sistema en línea no nos permitió modificar los campos relacionados. Se solicita 12,000 pesos para el pasaje a un congreso internacional organizado en Europa, 3,000 pesos para la inscripción y 8,000 pesos para los viáticos. Para un congreso nacional, estimamos 3,500 de pasaje, 2,000 de inscripción y 5,000 para viáticos.
Estancias cortas	18,000	Solicitamos 18,000 pesos para el pasaje de una estancia corta de Dra. Schaeffer en Chile. Se utilizará fondos de los colaboradores para cubrir algunos gastos de las estancias; en especial se espera poder cubrir los viáticos de parte de ellos.
Beca para estudiante	19,185	El becario Carlos Castillo continuará laborando como becario en el proyecto durante el segundo año. Ha terminado la revisión de literatura y la redacción de los antecedentes de su tesis, igual como el diseño y la documentación del modelo. Estamos en la fase de implementación de la simulación. Durante la primavera del 2009 realizaremos los experimentos y la tesis estará concluida durante el verano de 2009.
Materiales y consumibles	26,800	Se solicita 1,500 pesos para actualizaciones de software. La situación de los recursos bibliográficos ya ha mejorado. El estado de arte avanza rápidamente lo que crea la necesidad de complementar la colección de libros que tenemos. Solicitamos 10,000 pesos para libros en el segundo año. Los costos de inscripción a revistas internacionales fueron mayores a lo que esperamos en el primer año y tuvimos que buscar fondos alternativos para muchas de las revistas que usamos. Se estima un gasto de 10,000 pesos en revistas en el segundo año. Los <i>materiales de oficina</i> necesarios incluyen el papel y los cartuchos de las impresoras blanco-negro tipo láser y de color y negro tipo inkjet, en uso compartida entre el profesor y los estudiantes del proyecto; se estima un costo de 2,500 pesos. También se necesitará imprimir <i>carteles</i> para presentar en congresos u otros eventos de investigación y posiblemente impresión de mayor calidad para las tesis (las impresoras conseguidas son de clase económica); se estimó un gasto de 2,000 pesos en impresiones. Para el segundo año contemplamos la <i>encuadernación</i> de tesis de un alumno (800 pesos).

Sección 8

Solicitud de beca de fomento a la permanencia

Les solicito la continuación de la beca de fomento a la permanencia por un segundo año. Durante el primer año he recibido una beca mensual de 6,000 pesos.

Sección 9

Productos académicos esperados

En el segundo año del proyecto se espera a llegar a tener los siguientes productos, aparte de los ya incluidos en este reporte. Se presenta un estimado de la cantidad mínima por cada tipo de producto.

- Artículos científicos publicados en
 - revistas internacionales (2),
 - revistas nacionales (3),
 - congresos internacionales (2) y
 - congresos nacionales (2).
- Ponencias y carteles en congresos (3).
- Conclusión de tesis de maestría (4).
- Conclusión de tesis de licenciatura (1).
- Presentación de la propuesta de proyecto doctoral (1).
- Incorporación de practicantes de verano científico (2).

Sección 10

Documentos probatorios

10.1. Formación de recursos humanos

- Asignación de tesista: David Ríos Soria
- Asignación de tesista: Perla Elizabeth Cantú Cerda
- Asignación de tesista: Vanesa Avalos Gaytán
- Asignación de tesista: Marco Antonio Aguirre Lam
- Asignación de practicante Delfín: Mario Rivera Ramírez
- Asignación de practicante PROVERICYT (2)
- Constancia PROVERICYT 2008

10.2. Producción académica

- Resumen del artículo en DCAI 2008
- Tabla de contenidos del libro *Grid Middleware and Services: Challenges and Solutions*, Springer.
- Constancia de ponencia ILAS 2008
- Constancia de ponencia CISCE 2008
- Carta de aceptación ILAS 2008
- Carta de aceptación CIINDET 2008
- Carta de aceptación SMM 2008
- Carta de aceptación ACS 2008
- Constancias de participación, carteles y premio de ENOAN 2008
- Invitación de ponencia CIDETEC 2008

10.3. Colaboración científica

- Estancia de Ing. Cantú del ITCM
- Estancia de Dra. Schaeffer del ITCM
- Estancia en la UANL de M.C. Turrubiates
- Estancia de Dra. Schaeffer en Finlandia

10.4. Becario de licenciatura

- Kárdex del becario
- Aceptación a prácticas profesionales
- Terminación exitosa de prácticas profesionales

Referencias bibliográficas

- [1] L.A. Adamic, R.M. Lukose, A.R. Puniyani y B.A. Huberman. Search in power-law networks. *Physical Review E*, 64(4):046135, 2001.
- [2] R. Albert, H. Jeong y A.-L. Barabási. Diameter of the World Wide Web. *Nature*, 401:130–131, 1999.
- [3] S. Androutsellis-Theotokis and D. Spinellis. A survey of peer-to-peer content distribution technologies. *ACM Computing Surveys*, 36(4):335–371, 2004.
- [4] S.V. Annibaldi and K.I. Hopcraft. Random walks with power-law fluctuations in the number of steps. *Journal of Physics A*, 35(41):8635–8645, 2002.
- [5] A. Arenas, A. Cabrales, A. Díaz-Guilera, R. Guimerà, and F. Vega-Redondo. Search and congestion in complex networks. En R. Pastor-Satorras, M. Rubi y A. Diaz-Guilera, editores, *Statistical Mechanics of Complex Networks*, volumen 625 de *Lecture Notes in Physics*, páginas 175–194, Berlín, Alemania, 2003. Springer, Berlín, Alemania..
- [6] F. Banaei-Kashani and C. Shahabi. Criticality-based analysis and design of unstructured peer-to-peer networks as “complex systems”. En *Third International Workshop on Global and Peer-to-Peer Computing*, 2003.
- [7] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [8] B.F. Cooper. Quickly routing searches without having to move content. En *Proceedings of the Fourth International Workshop on Peer-to-Peer Systems*, 2005.
- [9] T.H. Cormen, C.E. Leiserson, R.L. Rivest y C. Stein. *Introduction to Algorithms*. McGraw-Hill Book Co., Boston, MA, EUA, edición segunda, 2001.
- [10] M.S. Corson and A. Ephremides. A distributed routing algorithm for mobile wireless networks. *Wireless Networks*, 1(1):61–81, 1995.
- [11] A.P.S. de Moura, A.E. Motter y C. Grebogi. Searching in small-world networks. *Physical Review E*, 68(3):036106, 2003.
- [12] P. Duchon, N. Hanusse, E. Lebhar y N. Schabanel. Could any graph be turned into a small-world? *Theoretical Computer Science*, 355(1):96–103, 2006.

- [13] C. Gkantsidis, M. Mihail y A. Saberi. Hybrid search schemes for unstructured peer-to-peer networks. En *Proceedings of the Twenty-fourth Annual Joint Conference of the IEEE Computer and Communications Societies*, volumen 3, páginas 1526–1537, Los Alamitos, CA, EUA, 2005. IEEE Computer Society Press.
- [14] C.P. Herrero. Self-avoiding walks on scale-free networks. *Physical Review E*, 71(1):016103, 2005.
- [15] B.J. Kim, C.N. Yoon, S.K. Han y H. Jeong. Path finding strategies in scale-free networks. *Physical Review E*, 65(2):027103, 2002.
- [16] J.M. Kleinberg. The small-world phenomenon: an algorithmic perspective. En *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*, páginas 163–170, New York, NY, EUA, 2000. ACM Press.
- [17] G. Lv, P. Cao, E. Cohan, K. Li y S. Shenker. Search and replication in unstructured peer-to-peer networks. En *Proceedings of the Sixteenth Annual ACM International Conference on Supercomputing*, páginas 84–95, New York, NY, EUA, 2002. ACM Press.
- [18] G.S. Manku, M. Naor y U. Wieder. Know thy neighbor's neighbor: The power of lookahead in randomized P2P networks. En *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, páginas 54–63, New York, NY, EUA, 2004. ACM Press.
- [19] J. Risson and T. Moors. Survey of research towards robust peer-to-peer networks: search methods. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 50(17):3485–3521, 2006.
- [20] M. Rosvall, A. Trusina, P. Minnhagen y K. Sneppen. Networks and cities: An information perspective. *Physical Review Letters*, 94(2):028701, 2005.
- [21] S.E. Schaeffer. Algorithms for nonuniform networks. Informe de Investigación A102, Helsinki University of Technology, Laboratory for Theoretical Computer Science, Espoo, Finlandia, 2006.
- [22] K. Sneppen, A. Trusina y M. Rosvall. Hide-and-seek on complex networks. *Europhysics Letters*, 69(5):853–859, 2005.
- [23] C. Szepesvári. Shortest path discovery problems: A framework, algorithms and experimental results. En D.L. McGuinness and G. Ferguson, editores, *Proceedings of the Nineteenth National Conference on Artificial Intelligence and Sixteenth Conference on Innovative Applications of Artificial Intelligence*, páginas 550–555, Menlo Park, CA, EUA, 2004. AAAI Press/The MIT Press.
- [24] B. Tadić. Adaptive random walks on the class of Web graphs. *European Physical Journal B*, 23(2):221–228, 2001.
- [25] B. Tadić. Growth and structure of the World-Wide Web: Towards realistic modeling. *Computer Physics Communications*, 147(1–2):586–589, 2002.
- [26] H. Wang, T. Lin, C.H. Chen y Y. Shen. Dynamic search in peer-to-peer networks. En *Proceedings of the Thirteenth International World-Wide Web Conference*, New York, NY, EUA, 2004. ACM Press.

- [27] S.-J. Yang. Exploring complex networks by walking on them. *Physical Review E*, 71(1):016107, 2005.
- [28] H. Zhu and Z.-X. Huang. Navigation in a small world with local information. *Physical Review E*, 70(3):036117, 2004.