

## PROCESAMIENTO PARALELO DISTRIBUIDO HETEROGÉNEO EN APLICACIONES ESTRUCTURALES Y NUMÉRICAS

**Gustavo E. Vazquez y Nélica B. Brignole**

Grupo de Investigación y Desarrollo en Computación Científica GIDeCC  
Departamento de Ciencias e Ingeniería de la Computación  
Universidad Nacional del Sur, Avda. Alem 1253, 8000 Bahía Blanca, Argentina

Planta Piloto de Ingeniería Química PLAPIQUI (UNS-CONICET)  
Comp CRIBABB, km 7 Camino La Carrindanga, 8000 Bahía Blanca, Argentina  
Tel.: 0291 4861700 - Fax: 0291 4861600 - e-mail: [dybrigno@criba.edu.ar](mailto:dybrigno@criba.edu.ar)

### Resumen

El objetivo global de esta línea de investigación es diseñar e implementar nuevas estrategias de procesamiento paralelo en entornos de cómputo distribuido heterogéneo para facilitar la resolución de problemas tanto estructurales como numéricos cuya resolución mediante técnicas secuenciales resulta costosa en tiempos de ejecución debido a la magnitud o complejidad de las instancias que se desea resolver. Como resultado de estas investigaciones se ha logrado el desarrollo de técnicas robustas y eficientes aplicables a un amplio espectro de problemas de búsquedas en grafos y de optimización con función objetivo y restricciones no lineales.

En términos generales, es posible distinguir dos ramas de investigación para el desarrollo de algoritmos paralelos distribuidos: la paralelización de algoritmos secuenciales existentes y la creación de alternativas intrínsecamente paralelas. En el caso de problemas estructurales, se estudiaron los métodos secuenciales clásicos de búsqueda en grafos y se establecieron las limitaciones para su uso en redes de estaciones de trabajo. Sobre esta base se propuso un nuevo método de distribución semi-dinámica y se lo aplicó al algoritmo GS-FLCN para análisis de observabilidad. Por otra parte, en la línea de los algoritmos intrínsecamente paralelos se desarrolló un nuevo algoritmo de búsqueda totalmente distribuido con el objeto de aumentar la eficiencia de los recorridos para esta aplicación específica. En vez de la estrategia tipo Master-Worker de la propuesta anterior se diseñó otra mucho más eficiente que sigue un esquema descentralizado tipo Master-Supervisor-Worker. En cuanto a los problemas numéricos, se consideraron estrategias para aplicar el paralelismo a las secciones de cómputo intensivo de los algoritmos secuenciales existentes para optimización no lineal con restricciones conocidos como GRG y SQP. Asimismo se desarrolló una nueva técnica de descomposición de dominio con el objeto de ampliar el rango de aplicabilidad de un algoritmo intrínsecamente paralelo concebido originalmente para problemas sin restricciones de modo que se lo pudiera utilizar en forma eficiente para el tratamiento de los problemas de optimización no lineal con restricciones.

En cuanto a las verificaciones de desempeño, las implementaciones se testearon sobre redes de estaciones de trabajo homogéneas y heterogéneas pequeñas y se hicieron estudios de escalabilidad. Las métricas clásicas de speed-up debieron ser ajustadas con el objeto de tener en cuenta la heterogeneidad de los procesadores y así poder asegurar comparaciones justas. En tal sentido, todos los nuevos algoritmos propuestos lograron muy buen desempeño en cuanto al tiempo de ejecución en comparación con los algoritmos secuenciales correspondientes. Se analizaron casos de estudio académicos y problemas industriales reales de mediano y gran tamaño pertenecientes al área de ingeniería de procesos.

## Resultados Alcanzados

### 1. Procesamiento Paralelo Distribuido Aplicado al Desarrollo de Algoritmos para Reordenamiento Estructural de Matrices Ralas.

Los grafos poseen una estrecha relación con la teoría de matrices ralas y en particular con los métodos de reordenamiento y manipulación estructural de dichas matrices. De hecho, una de las formas más utilizadas para representar grafos es justamente a través de matrices. El reordenamiento estructural de matrices a patrones de raridad específicos puede efectuarse empleando búsquedas arbóreas. La permutación de matrices ralas a formas especiales permite ahorrar tiempos de cómputo y espacio de memoria. En particular, nosotros hemos trabajado en el reordenamiento estructural a forma triangular en bloques (FTB), el cual ha sido usado con éxito en la resolución de sistemas lineales, problemas de mínimos cuadrados, particionamiento de matrices para procesamiento paralelo, entre otras aplicaciones.

El problema específico seleccionado para efectuar nuestras investigaciones consiste en el desarrollo de algoritmos para el reordenamiento estructural de las matrices de ocurrencia ralas asociadas a los sistemas de ecuaciones que modelan procesos industriales con vistas a lograr la FTB requerida para resolver problemas de diseño de instrumentación. En este ámbito, las técnicas combinatoriales poseen un elevado costo computacional, el cual puede ser disminuido mediante la aplicación criteriosa del concepto de paralelismo. Se trabajó sobre la base del mejor algoritmo secuencial de particionamiento existente, denominado GS-FLCN [1] y se desarrolló una versión paralela que se basa en búsquedas arbóreas tipo depth-first search (DFS). En cuanto a la metodología de trabajo la tarea se realizó en dos etapas, paralelizándose primero la técnica DFS genérica [2] para luego encarar específicamente la paralelización de GS-FLCN [3]. La etapa preliminar permitió fijar pautas vinculadas a aspectos esenciales tales como el estudio de los umbrales a partir de donde empezar a paralelizar y la definición de una métrica apropiada para cálculos de speed-up en arquitecturas heterogéneas que condujeran a comparaciones justas. Se obtuvo un software eficiente, flexible y escalable sobre clusters de estaciones de trabajo heterogéneas interconectadas por redes locales de comunicación.

En cuanto a la verificación de desempeño de la versión paralela, se testeó el algoritmo DFS paralelo sobre un cluster de tres estaciones de trabajo que consta de una computadora Digital Alpha, una PC Pentium II de 400 MHz y una PC Pentium de 133 MHz. Luego, el algoritmo GS-FLCN paralelo se probó con dos casos de estudio industriales en un cluster homogéneo de computadoras Pentium de 200 MHz disponibles en el Centro de Cómputos del Departamento de Ciencias de la Computación de la UNS. Se comenzó trabajando con dos máquinas, para luego efectuar un escalado incremental por agregado de nuevos nodos de procesamiento al sistema hasta totalizar doce procesadores.

Si bien este algoritmo paralelo distribuido permite ganancias significativas en tiempo, el enfoque no explota todo el potencial de paralelización del problema pues hay subgrafos que son recorridos más de una vez. El principal inconveniente es que el método no aprovecha la posibilidad de factorizar la detección de subcamino repetidos. Cada vez que la tarea Master llega a una misma hoja en su recorrido a través de un camino distinto, se genera una exploración repetida a lo largo del subgrafo cuya raíz es ese nodo. Esta deficiencia es más notoria a medida que crece el tamaño del grafo, tanto en cantidad de nodos como en densidad de arcos. Para evitar la reiteración de exploraciones fue necesario replantear el algoritmo desde su concepción. La nueva propuesta [4], que se basa en una arquitectura Master-Supervisor-Worker descentralizada especialmente diseñada para búsquedas en profundidad, evita la exploración repetida de subcamino. El trabajo se efectuó en varias etapas. En primer lugar se desarrollaron los pseudoalgoritmos para cada una de las tareas que componen el sistema. Se determinaron las estructuras de datos más adecuadas para lograr la

funcionalidad esperada, así como también los momentos y las formas de cada una de las interacciones entre las tareas. Se trabajó, además, en un protocolo de terminación distribuido del algoritmo. El nuevo algoritmo demostró ser más eficiente que la estrategia paralela centralizada, lográndose sustanciales mejoras en los tiempos de cómputo.

## **2. Paralelización en Entornos de Procesamiento Distribuidos de Algoritmos Secuenciales Clásicos para la Resolución de Problemas de Optimización con Función Objetivo y Restricciones No Lineales.**

En esta línea se desarrollaron algoritmos paralelos basados en métodos tipo gradiente, de probada robustez. Implementaciones secuenciales de estos métodos se han empleado con éxito en numerosos problemas de ingeniería. Las aplicaciones más interesantes desde el punto de vista de las ventajas que se derivarían de su paralelización son aquellas cuya función objetivo y restricciones son no lineales y costosas de evaluar. Dado que no existen antecedentes en la literatura sobre implementaciones de estas metodologías en ambientes distribuidos, este trabajo constituye un interesante aporte para la resolución eficiente de una amplia variedad de problemas de interés.

Dentro de la familia de los métodos tipo gradiente existen dos grandes técnicas: Programación Cuadrática Sucesiva (SQP) y Gradiente Reducido Generalizado (GRG). SQP resuelve internamente un problema con función objetivo cuadrática y restricciones no lineales para generar cada dirección de búsqueda. GRG, en cambio, resuelve un sistema lineal de ecuaciones. Por otra parte, en el SQP la función objetivo del problema de búsqueda en línea es un Lagrangiano aumentado mientras que para GRG la evaluación de dicha función objetivo exige resolver otro sistema de ecuaciones lineales. La tarea consistió básicamente en introducir paralelismo en aquellas etapas de cada algoritmo base que resultan críticas desde el punto de vista de los tiempos de cómputo. Los métodos tipo gradiente se basan en la generación de direcciones de búsqueda, para lo cual se emplea información sobre las derivadas primeras y segundas de las funciones involucradas. La tarea de evaluación de gradientes de funciones objetivo y las distintas técnicas alternativas para búsqueda en línea son los dos aspectos que concentran la mayor carga de cómputo de estos métodos. Por lo tanto, se consideró la paralelización del cálculo de gradientes y/o Hessianos, así como también del proceso de búsqueda en línea. Cada etapa de la búsqueda en línea consiste en encontrar un óptimo dentro de una dirección definida previamente. Por lo tanto, dado el vector dirección, se paralelizó el proceso de selección del tamaño de paso adecuado. De este modo se generaron versiones paralelas de los métodos SQP [5] y GRG [6]. En ambos casos se paralelizó la evaluación de todas las derivadas parciales y el proceso de búsqueda en línea. La única diferencia conceptual entre la filosofía con que se paralelizaron ambas estrategias está centrada en la búsqueda en línea. Para SQP, se evaluó la función objetivo simultáneamente en varios puntos, mientras que GRG involucró la evaluación paralela de las restricciones requeridas por la rutina de resolución de sistemas no lineales.

La implementación y chequeo de los algoritmos adecuadamente paralelizados permitió definir la ganancia que puede esperarse gracias al empleo de esta metodología. Con este propósito se implementó un conjunto de casos de prueba que corresponden a problemas clásicos empleados para la evaluación de desempeño de algoritmos de optimización [7]. Se seleccionaron problemas representativos de diferentes tipos de dificultades que pueden presentarse para el cálculo del óptimo de modo que sirvieran para evaluar la calidad del algoritmo desarrollado.

En cuanto al ambiente de cómputo utilizado, se trabajó sobre una red Ethernet heterogénea pequeña, con una estación de trabajo Digital Alpha, una PC Pentium II de 400 MHz y una PC Pentium de 133 MHz. Se estudió en detalle el desempeño en cuanto a la cantidad de nodos de procesamiento usados, speed-up y eficiencia. Desde un punto de vista más general, el diseño y programación de estas estrategias permitió extraer conclusiones acerca de la utilización óptima de CPUs y de los efectos provocados por el uso de redes de comunicación de datos.

### 3. Desarrollo e Implementación de Algoritmos Intrínsecamente Paralelos para la Resolución de Problemas de Optimización con Función Objetivo y Restricciones No Lineales.

En este campo se propuso un nuevo algoritmo de optimización paralelo que se basa en el algoritmo de distribución paralela de variables (PVD) [8]. Este enfoque está basado en la descomposición del dominio del problema, lo cual básicamente se efectúa particionando el vector de variables de optimización. La técnica original sólo contemplaba problemas de optimización sin restricciones. Esto es claramente insuficiente para el tratamiento de la mayoría de los problemas prácticos complejos. Por lo tanto, se decidió modificar la técnica para poder considerar problemas restringidos. El principal aporte de la nueva metodología es la política incorporada para manejar restricciones no lineales y la propuesta de una nueva política de particionamiento de dominio.

El algoritmo paralelo se implementó en un entorno de procesamiento distribuido y el código se empleó en primer término para resolver varios modelos no lineales pequeños asociados a problemas clásicos de ingeniería [7]. A continuación, se encaró un problema industrial más grande y más complejo, correspondiente al modelo riguroso de una planta de expansión existente. Se lograron valores satisfactorios de speed-up y eficiencia. Una descripción detallada del algoritmo y el análisis de desempeño sobre todos estos ejemplos puede consultarse en [9].

#### Referencias

[1] **Ponzoni I., M.C. Sanchez, N.B. Brignole** "A New Structural Algorithm for Observability Classification" *Ind. Eng. Chem. Res.* (ISSN: 0888-5885), 38, 8, 1999, 3027-3035.

[2] **Vazquez G.E., I. Ponzoni, N.B. Brignole** "Parallel Depth-First Search on Clusters of Workstations", *PS SIAM Annual Meeting 1999*, Atlanta, USA, 12-15 Mayo 1999, 198

[3] **Ponzoni I., G.E. Vazquez, M.C. Sánchez, N.B. Brignole.** "Parallel Observability Analysis on Networks of Workstations", *Comp. & Chem. Eng.* (ISSN: 0098-1354), 25/7-8, 997-1002, 2001

[4] **Fapitalle, F., G.E. Vazquez, I. Ponzoni, N.B. Brignole** "A Decentralized Parallel-Distributed Observability Algorithm", aceptado *V WCCM (Fifth World Congress on Computational Mechanics)*, Viena, Austria, Julio 7-12, 2002

[5] **Vazquez G.E., Brignole N.B.** "Parallel NLP Strategies using PVM on Heterogeneous Distributed Environments", *Lecture Notes in Computer Science* (ISSN: 3-540-66549-8), 1697: *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, 533-540, Springer-Verlag, Berlin, 1999.

[6] **Vazquez G.E., R. Rainoldi, N.B. Brignole** "Non-Linear Constrained GRG Optimization under Heterogeneous Parallel-Distributed Computing Environments", *Computer-Aided Chemical Engineering* (ISBN: 0-444-82409-X), Vol. 8, Elsevier, 2000, 127-132.

[7] **Hock W., Schittkowsky K.**, "Test Examples for Nonlinear Programming Codes", Springer-Verlag, New York, 1981.

[8] **M.C. Ferris, O.L. Mangasarian** "Parallel Variable Distribution", *SIAM Journal of Optimization*, 4, 1994, 815-832.

[9] **Vazquez G.E., S. Diaz, N.B. Brignole, J.A. Bandoni** "Optimization of Industrial Problems using Parallel Processing under Distributed Environments", *Chemical Engineering Communications* (ISSN: 0098-4645), 189, 6, 2002.