XXII Conferencia Latinoamericana de Informática - CLEI - 96

Partición de Sistemas de Ecuaciones para su Resolución Distribuida[1]

Benjamín Barán[+] *, Diana Benítez* y Rodrigo Ramos[+]

[+]Centro Nacional de Computación y Facultad de Ingeniería, Universidad Nacional de Asunción

Fax: (595) (21) 585619

*Facultad de Ciencias y Tecnología, Universidad Católica Nuestra Señora de la Asunción

Fax: (595) (21) 310587

E-mail: bbaran@una.py

Resumen

Este trabajo propone un método para descomponer sistemas de ecuaciones, de manera a facilitar su resolución utilizando ambientes distribuidos heterogéneos como computadores paralelos y redes de computadores. La presente propuesta asigna a los procesadores de un sistema distribuido heterogéneo distintos números de ecuaciones según sea la performance relativa de los procesadores a ser utilizados, de forma tal que su aplicación sea válida para un sistema general de ecuaciones. Así mismo, el método presentado permite utilizar la técnica de solapamiento parcial (*partial overlapping*) para mejorar la convergencia. Además, se introduce un parámetro de selección para escoger la mejor de un conjunto dado de particiones. Se presentan resultados experimentales que avalan la presente propuesta.

Palabras clave: Computación distribuída, convergencia de métodos iterativos, performance relativa.

[1]Este trabajo fue iniciado en el marco del proyecto UNESCO - ROSTLAC 885.610-4; con el auspicio de la Universidad Nacional de Asunción (UNA) y la Universidad Católica Nuestra Señora de la Asunción (UCA) y contó con el patrocinio de la Dirección de Investigaciones Postgrado y Relaciones Internacionales (DIPRI) de la UNA.

1. INTRODUCCIÓN

Las implementaciones paralelas de algoritmos iterativos son una opción importante dentro del contexto de la computación distribuida [1]. Para aprovechar sus ventajas en la resolución de sistemas de ecuaciones de gran porte es necesario descomponer el problema dado en subproblemas menores, los cuales pueden ser asignados a los distintos procesadores del sistema distribuido para su resolución paralela. Esta descomposición estará condicionada por dos factores:

a) la capacidad de procesamiento relativo de cada máquina del sistema distribuido, factor que determina las dimensiones de los subproblemas asignados a cada procesador;
b) el grado de acoplamiento que puedan tener los subproblemas entre sí, lo que determina la dependencia entre las variables y por consiguiente, influye en la convergencia del algoritmo [2].

Así, el sistema debe descomponerse de forma tal que a cada procesador se le asigne un subproblema de dimensión proporcional a su performance, y que la dependencia entre las variables actualizadas por cada procesador facilite la convergencia del algoritmo iterativo a ser utilizado.

Desde la década del 60 [3] fueron publicados diversos métodos de descomposición, en su mayoría diseñados para procesadores, problemas o algoritmos bien específicos [4-5]. Así, Vale et al. [6] formuló una propuesta para sistemas de ecuaciones lineales y simétricos utilizando computadores paralelos. Sin embargo, la Descomposición ε de Sezer y Šiljak [7] que desprecia ligaciones débiles, es la más utilizada en la actualidad.

Inspirada en el trabajo de Vale et al. [6], se desarrolla la presente propuesta que permite generar particiones en número y tamaño controlables por el operador, válida para un sistema general de ecuaciones. Además, se introduce un parámetro que permite escoger una buena partición de un conjunto de particiones y se sugiere la utilización de solapamiento parcial (*partial overlapping*) [8].

La sección 2 formula matemáticamente el problema, mientras la sección 3 presenta el método propuesto. Resultados experimentales se presentan en la sección 4 y las conclusiones en la sección 5.

2. FORMULACIÓN DEL PROBLEMA

Considerando un sistema de *n* ecuaciones algebraicas con n incógnitas dado por:

$$\Phi(x) = 0, \qquad x \in \Re^n, \qquad \Phi: \Re^n \to \Re^n \tag{2.1}$$

definido en un dominio $D \subset \Re^n$; se desea resolver (2.1) utilizando un sistema distribuido con *p* procesadores, $n \geq p$. Para esto, se debe descomponer el problema de forma tal que cada procesador *resuelva* solamente una parte del sistema, comunicando sus resultados parciales a los otros procesadores.

Sea la Descomposición Cartesiana: $\Re^n = \Re^{n_1} \times \cdots \times \Re^{n_p}, \quad n_1 + \cdots + n_p = n$ (2.2)

con $D \subset \Re^n$ un dominio tal que $D = D_1 \times \cdots \times D_p, \quad D_i \subset \Re^{n_i}, \quad \forall i \in \{1,\ldots,p\}$; (2.3)

entonces: $x = \begin{bmatrix} x_1^T & \cdots & x_p^T \end{bmatrix}^T, \; x_i \in D_i, \; \forall\, i \in \{1,\ldots,p\}.$ (2.4)

Análogamente, $\Phi(x) = \begin{bmatrix} \Phi_1^T(x) & \cdots & \Phi_p^T(x) \end{bmatrix}^T, \quad \Phi_i : D \to \Re^{n_i}$ (2.5)

Consecuentemente, la ecuación a ser resuelta en cada procesador *i* (problema local), estará dada por:

$$\Phi_i(x) = 0, \quad \forall i \in \{1,\ldots,p\} \tag{2.6}$$

NOTA 1. *La descomposición adecuada de la función $\Phi(x)$ en p funciones $\Phi_i(x)$ es el objetivo central del presente trabajo.*

Para resolver el problema será utilizado un algoritmo iterativo de la forma: $x^{k+1} = G(x^k)$ (2.7)

donde *k* representa la iteración. Por lo tanto, en forma análoga a (2.4) y (2.5):

$$G(x) = \begin{bmatrix} G_1^T(x) & \cdots & G_p^T(x) \end{bmatrix}^T, \qquad G_i(x): D \to D_i, \qquad \forall i \in \{1,\ldots,p\} \tag{2.8}$$

La implementación síncrona de (2.7) para el procesador *i* será entonces: $x_i^{k+1} = G_i(x^k)$ (2.9)

Considerando un contexto asíncrono, se denota como $x_j(d_j^i(k))$ al 2valor de x_j enviado desde el procesador *j* y disponible en el procesador *i* al iniciar su iteración *k*. Denotamos así como $x^i(k)$ al vector

x disponible en el procesador i en el momento de su iteración k, esto es:

$$x^i(k) = \left[x_1^T(d_1^i(k)) \quad \cdots \quad x_p^T(d_p^i(k)) \right]^T, \qquad \forall\, i \in \{1,\ldots,p\} \tag{2.10}$$

Usando esta notación [9], podemos escribir el algoritmo asíncrono basado en (2.7) como:

$$x_i(k+1) = G_i(x^i(k)), \qquad \forall i \in \{1,\ldots,p\} \tag{2.11}$$

En [1] se deriva una condición suficiente de convergencia de (2.11), bajo las siguientes hipótesis:

Hipótesis 1 (Unicidad de la solución): *dado el sistema de ecuaciones (2.1) que puede ser reescrito como (2.6, definido en un conjunto cerrado* $D \subset \Re^n$ *que satisface (2.3), existe un operador G de la forma (2.8), tal3 que* $G(D) \subset D$ *y adicionalmente D presenta un y solamente un punto fijo*

$$x^* = \left[\left(x_1^*\right)^T \quad \cdots \quad \left(x_p^*\right)^T \right]^T, \qquad x_i^* \in D_i\,, \qquad \forall\, i \in \{1,\ldots,p\}$$

del operador G. El punto fijo x^* *es a su vez solución de (2.1).*

Hipótesis 2 (Asincronismo parcial): $\exists\, d \in \mathrm{N}\,,\ \forall k \in \mathrm{N}\,,$ *tal que 2 donde* $(k-d)$ *es el máximo 2atraso posible, determinado por la medida de asincronismo d.*

Hipótesis 3 (Bloque-Lipschitz): *cada operador* $G_i(x)$ *en (2.11) es Bloque-Lipschitz en D, esto es:*

$$\| G_i(x) - G_i(y) \| \le \sum l_{ij} \| x_i - y_j \|, \qquad \forall x,y \in D$$

Se demuestra así el siguiente teorema [1]:

<u>Teorema 1</u> (Condición suficiente de convergencia)

Bajo las hipótesis 1 a 3, el algoritmo asíncrono representado por (2.11) converge a un único punto fijo x en D si* $\rho(H) < 1$ *, donde* ρ *representa el radio espectral y la matriz de comparación H está dada por:*

$$H = (h_{ij}) \in \Re^{pxp}\,, \qquad con\ \ h_{ij} = l_{ij}$$

En resumen, el radio espectral de la matriz de comparación es un parámetro que nos permitiría asegurar a priori que el algoritmo converge, y por consiguiente podría ser utilizado para escoger buenas particiones.

La técnica de solapamiento parcial (*partial overlapping*) fue introducida para mejorar la convergencia de los métodos de resolución en bloque, y consiste en resolver una misma ecuación repetidamente en varios bloques (o procesadores, en una implementación paralela) [8].

3. PROPUESTA DEL TRABAJO

El método de descomposición propuesto utiliza una matriz cuadrada no negativa $M = \{m_{ij}\}$ ($m_{ij} \geq 0$) de dimensión n x n y elementos diagonales no nulos ($m_{ii} > 0$), cuyos elementos m_{ij} y m_{ji} $(i \neq j)$ representan el acoplamiento existente entre las variables x_i y x_j. En el caso lineal ($Ax=b$), los elementos m_{ij} de la matriz M están dados por $m_{ij} = | a_{ij} |$, donde A es la matriz de coeficientes del sistema o una versión reordenada de la misma. Si el sistema a ser resuelto fuera no lineal, la matriz M a ser utilizada dependerá del método de resolución.

Decimos que las incógnitas x_i y x_j no son adyacentes si $m_{ij} = m_{ji} = 0$; caso contrario x_i y x_j son adyacentes. En caso de ser adyacentes, se habla de variables débilmente acopladas (si m_{ij} y m_{ji} son pequeñas) o fuertemente acopladas (si m_{ij} ó m_{ji} son muy grandes).

Básicamente, el método consiste en la formación de p subproblemas a partir de p incógnitas iniciales, denominadas *semillas*, donde p representa también el número de procesadores. A estas incógnitas semillas se van agregando las demás incógnitas hasta que el sistema quede totalmente descompuesto. La determinación de las incógnitas a ser anexadas a una dada semilla depende de un ranking de incógnitas basado en pesos previamente asignados. A cada paso del proceso, la incógnita a ser agregada al subproblema es la adyacente más pesada. En caso de que dos o más subproblemas seleccionen a la misma incógnita como candidata, esta se asignará al subproblema con el cual posea un mayor acoplamiento. Si no es posible desempatar y la incógnita en disputa posee un peso significativo, se puede optar por solapamiento parcial. Para esto se asigna la incógnita en disputa a todos los subproblemas que la hayan seleccionado como candidata. A medida que los subproblemas van siendo formados, debe verificarse que el tamaño de los mismos se encuentre en la proporción deseada (performance relativa de los procesadores: w). Para esto, se anexan incógnitas sólo a aquellos subproblemas que aun no alcanzaron dicha proporción. El proceso de agrupamiento de incógnitas continua hasta que todas las incógnitas hayan sido asignadas a algún subproblema. En caso de obtenerse más de una descomposición, se utiliza un criterio de selección para elegir la partición más conveniente.

Basados en el principio descrito, se propone el siguiente método de cuatro etapas (ver detalles en [9]):

1) **Clasificación de incógnitas (Algoritmo 1):** Se hace un ranking de pesos que indica el grado de acoplamiento de cada incógnita al sistema. El peso P_i de una incógnita x_i esta dado por

$$P_i = \sum_{j=1 \;\; j \neq i}^{n} (m_{ij} \; / \; m_{ii})^{z_{ij}} \quad con \quad z_{ij} = m_{ij} \; / \; (m_{ii} \; / \; D) \qquad \forall i \in \{1,\ldots,n\} \tag{3.1}$$

donde $$D = \frac{1}{nc} \sum_{i=1}^{n} \sum_{j=1,\, j \neq i}^{n} m_{ij} \quad , \qquad nc = número\ de\ m_{ij}\ diferentes\ de\ cero$$

2) **Selección de semillas (Algoritmo 2):** Deben seleccionarse como semillas a las incógnitas que sean centro de aglutinamiento de incógnitas y que no se encuentren fuertemente acopladas unas con otras.

3) **Formación de particiones (Algoritmo 3):** Una vez determinadas las semillas para cada subsistema, es necesario adicionar las demás incógnitas a los subconjuntos adecuados.

4) **Selección de la mejor partición (Algoritmo 4):** Pueden ser determinados varios conjuntos de semillas que a su vez servirán de pie para generar diferentes particiones. Para identificar a la mejor partición de todas las obtenidas se utiliza un parámetro de selección. En [2] se proponen varios parámetros posibles de selección, recomendándose como óptimo al Parámetro de Acoplamiento "Par_A", dado por:

$$\text{Par\_A} = \sum m_{ij} \quad , \qquad \forall i, j \;\; tal\ que\ \ x_i\ y\ x_j\ pertenecen\ a\ subproblemas\ diferentes$$

En base a lo expuesto en la sección 2, este trabajo propone como parámetro de selección a $\rho(H)$, siendo elegida como óptima aquella partición que presente el mínimo valor del parámetro de selección. La efectividad del parámetro $\rho(H)$ frente a Par_A se verá en los resultados experimentales de la sección 4.

4. RESULTADOS EXPERIMENTALES

La eficiencia del método descrito fue verificada experimentalmente utilizando sistemas de ecuaciones lineales y no lineales. Las diferentes particiones generadas por el método propuesto fueron comparadas con otras particiones determinadas aleatoriamente y con particiones generadas por la Descomposición ε [7]. Además se realizó un estudio comparativo entre los parámetros de selección $\rho(H)$ y Par_A.

4.1. Sistema lineal de ecuaciones.

Se presentan resultados experimentales de un sistema lineal de 10 ecuaciones y 10 incógnitas [Ejemplo A, 10]. Se hizo una descomposición exhaustiva, resolviéndose las 45# particiones posibles utilizando el método bloque-Jacobi, considerando una performance relativa w = $[4\ 1]^T$. Se hicieron mediciones del número de iteraciones hasta la convergencia y se calcularon los parámetros $\rho(H)$ y Par_A.

En la Tabla 4, se observa que las particiones generadas por el método propuesto están dentro de las mejores particiones posibles y que son superiores al promedio. Como puede notarse, la partición obtenida por el método propuesto (nº 7) es mejor que la partición sugerida por la Descomposición ε (nº 19).

nº de partición	Iteraciones+	Posición en el ranking		$\rho(H)$	Par_A
		del nº de iteraciones*	de particiones#		
7	4.33	2º	2º	0.0551	10
23	4.83	4º	6º	0.0709	8.6
44	4.67	3º	4º	0.0794	12.7
2	4.83	4º	7º	0.1008	10.5
39	5	5º	12º	0.1014	10.6
18	6	9º	21º	0.1352	11.3
ε 19	5.5	6º	16º	0.1183	7.9

+ El nº de iteraciones es un promedio de pruebas realizadas con seis valores posibles de tolerancia.
* Existen 16 posibles valores del nº de iteraciones que van desde 4.17 hasta 7.33; siendo 5.78 el promedio.
# Existen $C_2^{10} = \dfrac{10\,!}{(10-2)\,!\;2\,!} = 45$ posibles particiones.   ε Partición generada por Descomposición ε.

Tabla 4: Resultados experimentales del Ejemplo A

Se puede observar además, que si se considera a Par_A como parámetro de selección, entonces sería elegida la partición nº 23 a pesar de ser inferior a la seleccionada por el parámetro $\rho(H)$.

En otros ejemplos de sistemas lineales [10], las particiones generadas por el método propuesto son las óptimas, al igual que las sugeridas por la Descomposición ε.

Finalmente, se destaca que el coeficiente de correlación entre el número de iteraciones y $\rho(H)$ fue de 0.96 frente a una correlación de solo 0.15 entre el número de iteraciones y Par_A, lo que enfatiza las bondades de $\rho(H)$ (frente a Par_A) como parámetro de selección.

4.2. Sistema no lineal de ecuaciones

Se aplicó la metodología propuesta a la descomposición del sistema IEEE de 14 barras para la resolución del problema de flujo de potencia. Dicho problema puede ser formulado como un sistema cuasi-lineal de ecuaciones [11] de la forma: $Yx = I(x)$, $Y \in C^{n \times n}$, $x \in C^n$, $I(x) \in C^n$, donde *Y* es la matriz de admitancias de la red, *x* el vector de voltajes de las barras e $I(x)$ el vector de corrientes inyectadas.

La matriz de acoplamientos *M* utilizada será la misma matriz *Y*, teniendo en cuenta que ésta representa en forma efectiva al acoplamiento entre los voltajes de las barras.

Este sistema fue resuelto en forma síncrona (2.9) y asíncrona (2.11), utilizando un sistema distribuido constituido por dos workstations: una SUN SPARC-Station 5 y una DEC 3000, con $w = [4\ 1]^T$.

La Tabla 5 muestra las posiciones de las 2 particiones generadas por el método propuesto en diversos tipos de ranking, tanto para el caso síncrono como para el asíncrono. La partición 59 es la seleccionada por el método propuesto, resultando para el caso asíncrono la partición óptima.

Número de Partición	Posición+ en el Ranking de						$\rho(H)$	Par_A
	Iteraciones.		Tiempo Real		Tiempo de CPU			
	sinc.	asinc.	sinc.	asinc.	sinc.	asinc.		
59	3°	7°	4°	1°	3°	1°	0.902	30.93
56	47°	82°	34°	45°	43°	56°	0.965	24

+ Existen 286 posibles particiones.

Tabla 5: Ranking y parámetros de selección

Se nota además que si se utilizara Par_A como parámetro de selección, la partición 56 sería elegida, siendo que ésta es en realidad inferior a la partición 59 seleccionada por el parámetro $\rho(H)$.

La Tabla 6 indica las correlaciones entre los valores medidos y los parámetros de selección comparados, tanto en el caso síncrono como en el asíncrono. Se verifica en ambos casos que el radio espectral $\rho(H)$ presenta una mayor correlación que el parámetro Par_A con respecto a los tiempos de procesamiento y número de iteraciones.

	Iteraciones		Tiempo Real		Tiempo de CPU	
	sinc.	asinc.	sinc.	asinc.	sinc.	asinc.
$\rho(H)$	0.414	0.87	0.378	0.343	0.41	0.392
Par_A	0.0377	-0.05	0.0104	0.089	0.0137	0.0137

Tabla 6: Correlaciones

5. CONCLUSIONES

Dado el número de procesadores, los métodos de descomposición tradicionalmente utilizados no poseen dominio sobre el tamaño relativo de los subproblemas, o consideran sistemas muy particulares. Esta carencia dificulta el óptimo aprovechamiento de los sistemas distribuidos, generalmente heterogéneos. El presente trabajo presentó un método que permite descomponer cualquier sistema de ecuaciones con el desbalanceamiento deseado. Para esto, se introdujo además un parámetro de selección de particiones que mostró mejor comportamiento experimental que los presentados en trabajos anteriores.

Los resultados experimentales permiten concluir que:

- El método propuesto genera mejores particiones que una selección aleatoria, inclusive en varios casos la óptima.
- Las particiones generadas por el método propuesto son mejores ó iguales que las generadas por la Descomposición-ε, con la ventaja adicional de ejercer control sobre el número y tamaño de los subproblemas.
- $\rho(H)$ es un mejor parámetro de selección que el parámetro Par_A.

En resumen, la presente propuesta nos brinda una solución superior a las existentes para la descomposición de problemas con el objeto de resolverlos en un sistema distribuido heterogéneo.

BIBLIOGRAFÍA

[1] Barán B., Kaszkurewicz E. y Bhaya A., "Parallel Asynchronous Team Algorithms: Convergence and Performance Analysis". A ser publicado en la *IEEE Trans. on Parallel & Distributed Systems*, 1996.

[2] Vale M.H., *Descomposição de Redes Eléctricas para Procesamiento Paralelo*. Tesis Doctoral COPPE/UFRJ. Río de Janeiro, Brasil, 1995.

[3] Carré B.A., "Solution of Load-Flow by Partitioning Systems into Trees", *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-88, pp. 1931-1938, noviembre 1968.

[4] Irving M.R. y Sterling M.J.H., "Optical Network Tearing Using Simulated Annealing", *IEEE Proceedings,* vol. 137, no. 1, pp. 69-72, enero 1990.

[5] Gupta M. y Banerjee P., "Demostration of Automatic Data Partitioning Techniques for Parallelizing Compilers o8n Multicomputers", *IEEE Transactions on Parallel and Distributed Systems*, vol. 3, no.2, pp. 179-193, marzo 1992.

[6] Vale M.H., Falcão D.M. y Kaszkurewicz E., "Electrical Power Network Decomposition for Parallel Computations". *IEEE International Symposium on Circuits and Systems*-ISCAS 92. San Diego, California, 1992.

[7] Sezer M. y Šiljak D.D., "Nested epsilon decompositions of complex systems", IFAC" 9th *World Congress*, Budapest, Hungría, julio 1984.

[8] Ikeda M. y Šiljak D.D., "Overlapping decomposition, expansions and contractions of dynamic systems" *Large Scale System 1*, North-Holland Publishing Co., pp.29-38, 1980

[9] Barán B., *Estudo de Algoritmos Combinados Paralelos Assíncronos*. Tesis Doctoral COPPE/UFRJ. Río de Janeiro, Brasil. Octubre 1993.

[10] Barán B., Benítez D., Ramos R., "Partición de Sistemas de Ecuaciones para su Resolución en un Ambiente Computacional Distribuído", *Reporte Técnico del CNC-UNA*, TR-001/96, enero 1996.

[11] Monticelli A., *Fluxo de Carga em Redes de Energia Eléctrica*, Editora Edgar Blücher Ltda, 1983.