

Algoritmos Genéticos Distribuidos: Heterogeneidad y Migración

German Dupuy¹, Natalia Stark¹, Fernando Sanz Troiani²,
Hugo Alfonso¹, Carlos Bermudez, Carolina Salto¹, Gabriela Minetti¹
Laboratorio de Investigación en Sistemas Inteligentes (LISI)
Facultad de Ingeniería - Universidad Nacional de La Pampa
Calle 110 Esq. 9 (6360) General Pico - La Pampa - Rep. Argentina
Te. / Fax: (02302) 422780/422372, Int. 6302
e-mail: ¹{saltoc, alfonsoh@ing.unlpam.edu.ar}

Resumen Esta línea de investigación se enfoca en las metaheurísticas distribuidas, en especial los algoritmos genéticos distribuidos, dado que permiten reducir significativamente la complejidad temporal de la resolución y mejorar la calidad de las soluciones obtenidas. Dos características sumamente importantes a la hora de resolver problemas NP-duros y NP-completos.

Por un lado, esta línea de investigación se dedica a la configuración y evaluación del desempeño de los algoritmos genéticos distribuidos ejecutados sobre plataformas heterogéneas. En este sentido surge una nueva metodología, denominada HAPA, cuyo fin es brindar una implementación eficiente y eficaz de este tipo de algoritmos.

Por otro lado, la investigación se enfoca en una nueva política migratoria de los algoritmos genéticos distribuidos, con el objetivo de mejorar su desempeño. Para ello se propone una estrategia centrada en la obligatoriedad de participación en el crossover de los individuos recibidos por medio de la migración.

Palabras claves: Metaheurísticas, algoritmos genéticos distribuidos, paralelismo, heterogeneidad, migración

Contexto

Esta línea de investigación se desarrolla en el marco de dos proyectos de investigación dirigidos por la Dra. Carolina Salto y llevados a cabo en el Laboratorio de Investigación de Sistemas Inteligentes (LISI), de la Facultad de Ingeniería

de la Universidad Nacional de La Pampa. Uno de ellos acreditado por dicha facultad y el otro por la Agencia Nacional de Promoción Científica y Tecnológica. Cabe destacar que desde hace varios años, los integrantes de estos proyectos mantienen una importante vinculación con investigadores de la Universidad Nacional de San Luis (Argentina) y de la Universidad de Málaga (España), con quienes se realizan publicaciones conjuntas.

Introducción

Los algoritmos genéticos distribuidos (dGA) son una clase de metaheurísticas, de algoritmos evolutivos en particular, cuyas soluciones se evalúan y evolucionan en paralelo. Si bien el uso de metaheurísticas logra reducir significativamente la complejidad temporal de la búsqueda, la exploración sigue consumiendo mucho tiempo a la hora de resolver problemas industriales. En consecuencia, el paralelismo es necesario no solo para reducir el tiempo de resolución, sino también para mejorar la calidad de las soluciones obtenidas [1], [2], [3], ya que cuando se usa un dGA la búsqueda puede seguir un progreso diferente al de un algoritmo genético tradicional. En tanto que, la mayoría de los resultados presentados sobre dGAs asumen que las plataformas computacionales subyacentes tienen características idénticas (ambiente homogéneo) con respecto al hardware (procesadores, memoria, red) y a los componentes de software (sistema operativo) [4], [5]. Pero la realidad es que a los laboratorios actuales les resulta muy difícil mantener hardware homogéneo y cobra importan-

cia la necesidad de configurar y evaluar el desempeño de los dGAs sobre plataformas heterogéneas.

En este sentido esta línea de investigación ataca dos frentes distintos relacionados con los dGAs. El primero estudia el desempeño y la adaptación del proceso de búsqueda de los dGAs en entornos computacionales heterogéneos. En tanto que, el segundo frente está directamente vinculado con los cambios en el proceso de búsqueda atendiendo a la topología y a la política migratoria.

Desarrollo

Actualmente, los entornos de computación distribuida se componen de muchas computadoras heterogéneas capaces de trabajar de forma cooperativa. A pesar de esto, la mayor parte del trabajo en metaheurísticas paralelas supone un hardware homogéneo como plataforma subyacente. Por un lado, nuestro trabajo se centra en la propuesta de una metodología que le permite a un algoritmo genético distribuido adaptarse para lograr una mayor eficiencia [6], usando recursos de hardware disponibles con diferente potencia de cálculo y todos ellos colaborando para resolver el mismo problema. Por otra parte, la estrategia de migración juega un rol importante en el diseño de un algoritmo genético distribuido. Motivo por el cual, también, hemos propuesto en [7] una nueva política migratoria, donde se impone que los individuos recibidos en una subpoblación por los efectos de la migración formen parte de las parejas para la recombinación.

Las investigaciones más recientes de algoritmos genéticos distribuidos sobre plataformas heterogéneas [8], [9], [10], [11], [12] se enfocan en la resolución de un determinado problema, y no en la metodología que podría aplicarse cuando es necesario usar hardware heterogéneo. En este sentido, nuestra propuesta consta de un nuevo procedimiento metodológico, y el consecuente diseño algorítmico, denominado metodología *Hardware Aware Parallel Algorithms* (HAPA) [6]. El objetivo de HAPA es facilitar una implementación eficiente y numéricamente precisa de algoritmos genéticos distribuidos sobre un conjunto de máquinas con diferentes procesadores, tamaños de memoria principal, y sistemas operativos. HAPA consta de tres

fases: (i) el cálculo de un ranking de los procesadores con el fin de conocer la plataforma (*offline*), (ii) el diseño de algoritmos derivados de la fase anterior, y (iii) la validación de un dGA (*online*). En particular se ha aplicado la metodología HAPA para regular las condiciones de parada del dGA y la frecuencia de la migración de dicho algoritmo.

En cuanto al estudio de la política migratoria de los dGA, nuestro enfoque define una estrategia relacionada con la incorporación de los individuos a una población debido al intercambio producto de la migración. La propuesta se centra en la obligatoriedad de elección de los nuevos individuos recibidos en una subpoblación como padre para la recombinación [7]. Esto marca un contraste con las reglas tradicionales de un dGA, donde el individuo recibido compite con el resto de los individuos de la subpoblación en el proceso de selección de los padres para participar en la recombinación. Esto puede ocasionar que estos individuos no se seleccionen como padres, con lo cual se estaría perdiendo el posible nuevo material genético que podrían aportar.

Resultados obtenidos

Los resultados obtenidos en los dos frentes atacados por esta línea de investigación son prometedores. Por un lado, hemos llevado a cabo una serie de pruebas con el fin de evaluar el desempeño de HAPA y compararlo con el de un dGA ejecutado en un entorno homogéneo. Al analizar la repercusión de la heterogeneidad en el rendimiento resultante de un dGA y, también, la de su eficiencia en el tiempo de ejecución se concluye que: la calidad de la solución es comparable con la obtenida usando una plataforma homogénea, pero lo más importante es que se logra una disminución del esfuerzo numérico y, consecuentemente, del tiempo de ejecución para hallar dichas soluciones cuando HAPA se aplica en ambientes heterogéneos. En resumen, este trabajo aporta una forma de evitar el diseño *ad-hoc* de algoritmos para plataformas heterogéneas, también de abre una nueva línea de investigación sobre la forma de inyectar conocimientos de hardware en los parámetros del software de los algoritmos. Además de ser una propuesta innovadora, se ha

demostrado que es numéricamente competitiva y que los tiempos de ejecución son reducidos.

Por otra parte el dGA, cuya estrategia de migración considera como padre para la recombinación a aquellos nuevos individuos ingresados por una operación de migración, ha sido evaluado en un amplio conjunto de instancias del problema de planificación de proyectos de software y comparado con un dGA tradicional. Los resultados indican que además de encontrar soluciones de buena calidad, el uso de esta estrategia permite encontrarla en una menor cantidad de evaluaciones que un dGA tradicional.

Formación de recursos humanos

Cada año se incorporan al proyecto alumnos avanzados en la carrera Ingeniería en Sistemas, quienes trabajan en temas relacionados a la resolución de problemas de optimización usando técnicas inteligentes, con el objeto de guiarlos en el desarrollo de sus tesis de grado y, también, de formar futuros investigadores científicos. Por otra parte, los docentes-investigadores que integran el proyecto realizan diversos cursos de posgrado relacionados con la temática del proyecto, con el objetivo de sumar los créditos necesarios para cursar carreras de posgrado.

REFERENCES

- [1] E. Alba and M. Tomassini, "Parallelism and evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 5, pp. 443–462, Oct 2002.
- [2] E. Cantu-Paz, *Efficient and Accurate Parallel Genetic Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers, 2000.
- [3] E. Alba, *Parallel Metaheuristics: A New Class of Algorithms*. Wiley, 2005, ISBN: 978-0-471-67806-9. [Online]. Available: <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471678066.html>
- [4] E. Alba, A. J. Nebro, and J. M. Troya, "Heterogeneous computing and parallel genetic algorithms," *Journal of Parallel and Distributed Computing*, vol. 62, no. 9, pp. 1362–1385, 2002.
- [5] G. Luque and E. Alba, *Parallel Genetic Algorithms: Theory and Real World Applications*. Springer, 2011, vol. 367.
- [6] C. Salto and E. Alba, "Adapting distributed evolutionary algorithms to heterogeneous hardware," *Trans. Computational Collective Intelligence*, vol. 19, pp. 103–125, 2015.

- [7] G. Dupuy and C. Salto, "Una estrategia para la selección de individuos recibidos en la migración en un algoritmo genético distribuido," in *Proceedings of ASAI 2015 Argentine Symposium on Artificial Intelligence, 44th Argentine Conference on Informatics*, 2015, pp. 96–104.
- [8] J. W. B. Jr. and S. V. Kumar, "Asynchronous genetic algorithms for heterogeneous networks using coarse-grained dataflow," in *GECCO*, ser. Lecture Notes in Computer Science, E. Cantu-Paz, J. A. Foster, K. Deb, L. Davis, R. Roy, U.-M. O'Reilly, H.-G. Beyer, R. K. Standish, G. Kendall, S. W. Wilson, M. Harman, J. Wegener, D. Dasgupta, M. A. Potter, A. C. Schultz, K. A. Dowsland, N. Jonoska, and J. F. Miller, Eds., vol. 2723. Springer, 2003, pp. 730–741. [Online]. Available: <http://dblp.uni-trier.de/db/conf/gecco/gecco2003-1.html#BaughK03>
- [9] M. García-Arenas, J. J. Merelo Guervós, P. Castillo, J. L. J. Laredo, G. Romero, and A. M. Mora, "Using free cloud storage services for distributed evolutionary algorithms," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '11. New York, NY, USA: ACM, 2011, pp. 1603–1610. [Online]. Available: <http://doi.acm.org/10.1145/2001576.2001792>
- [10] Y. Gong, M. Nakamura, and S. Tamaki, "Parallel genetic algorithms on line topology of heterogeneous computing resources," in *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '05. New York, NY, USA: ACM, 2005, pp. 1447–1454. [Online]. Available: <http://doi.acm.org/10.1145/1068009.1068239>
- [11] K. Meri, M. G. Arenas, A. M. Mora, J. J. Merelo, P. A. Castillo, P. García-Sánchez, and J. L. J. Laredo, "Cloud-based evolutionary algorithms: An algorithmic study," *Natural Computing*, vol. 12, no. 2, pp. 135–147, 2012. [Online]. Available: <http://dx.doi.org/10.1007/s11047-012-9358-1>
- [12] S. Mostaghim, J. Branke, A. Lewis, and H. Schmeck, "Parallel multi-objective optimization using master-slave model on heterogeneous resources," in *IEEE Congress on Evolutionary Computation*. IEEE, 2008, pp. 1981–1987. [Online]. Available: <http://dblp.uni-trier.de/db/conf/cec/cec2008.html#MostaghimBLS08>