

Adaptación de técnicas metaheurísticas a diferentes escenarios

German Dupuy¹, Natalia Stark¹, Fernando Sanz Troiani²,
Hugo Alfonso¹, Gabriela Minetti¹, Carolina Salto¹

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{minettig, saltoc, alfonsoh@ing.unlpam.edu.ar}, ²fstnando@gmail.com

Resumen La finalidad de esta línea de investigación es el estudio, desarrollo y adaptación de técnicas metaheurísticas a diferentes escenarios. Si bien las metaheurísticas han resuelto exitosamente un gran número de diversos problemas, todavía quedan muchas características de las mismas a mejorar. En particular, nosotros abordamos: el tratamiento de datos ruidosos en problemas de optimización combinatoria, el control adaptativo de parámetros y la representación eficientes de las soluciones. El objetivo de nuestro trabajo es analizar estas áreas de investigación poco exploradas y formular estrategias para mejorar la eficiencia de las metaheurísticas.

Palabras claves: Metaheurísticas, ruido, algoritmos proactivos, control adaptativo, representaciones, frecuencia migratoria, mutación, optimización combinatoria.

CONTEXTO

Los desarrollos de esta línea de investigación se enmarcan en los proyectos de investigación “Resolviendo problemas complejos con técnicas metaheurísticas avanzadas” y PICTO 2011-0278. Ambos proyectos son 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. Los integrantes de este laboratorio mantienen desde hace varios años una importante vinculación con investigadores de la Universidad Nacional de San Luis (Argentina) y

de la Universidad de Málaga (España), con quienes se han realizado varias publicaciones conjuntas.

INTRODUCCIÓN

A diferencia de los métodos exactos, las metaheurísticas permiten atacar problemas con espacios de búsqueda muy grandes, entregando soluciones satisfactorias en tiempos razonables [1]. La popularidad de estas técnicas ha crecido en los últimos 25 años, su uso en múltiples aplicaciones ha demostrado eficacia y eficiencia a la hora de resolver problemas complejos y de gran tamaño. Las metaheurísticas, comunmente, resuelven problemas de: diseño en general, telecomunicaciones, aprendizaje de máquina y minería de datos en bioinformática y biología computacional, modelado de sistemas, planificación de tareas en problemas de enrutamiento, problemas de programación y producción, logística y transporte, entre muchos otros [2], [3], [4], [1].

A pesar de su gran popularidad y exitosa utilización en una gran variedad de problemas, todavía existen aspectos importantes de estas herramientas que no han recibido la atención suficiente por parte de la comunidad científica. Entre estos aspectos podemos encontrar: el basamento teórico de estas herramientas, el análisis de robustez de dichos algoritmos ante datos del problema que contengan errores, el ajuste de los valores paramétricos (tamaño de población, tipos de operadores, probabilidad de aplicación de los operadores, política migratoria, cantidad de iteraciones, etc.) y el análisis y

desarrollo de representaciones y operadores más eficientes.

Con el objetivo de aportar un grano de arena en este sentido, los integrantes de esta línea de investigación, analizamos, diseñamos y desarrollamos distintas estrategias que permitan una mejor adaptación y mayor eficiencia de las metaheurísticas en los siguientes escenarios: el tratamiento de datos que puedan contener errores (ruido), el control adaptativo de parámetros claves cómo son la probabilidad de mutación y la frecuencia de migración en los Algoritmos Evolutivos (EAs) y, por último, la selección de operadores que permita un comportamiento más eficiente del GA.

DESARROLLO

En esta sección describimos los desarrollos que se llevan a cabo en esta línea de investigación:

- **Tratamiento de datos ruidosos.** Con frecuencia se publican técnicas metaheurísticas que resuelven con éxito casos no ruidosos de problemas. Sin embargo, sólo hay unos pocos estudios que se han ocupado del ruido en los datos de las instancias de los problemas. En consecuencia, el objetivo de nuestro trabajo es analizar la robustez de metaheurísticas basadas en trayectoria, tales como Simulated Annealing (SA) y Problem Aware Local Search (PALS), y basadas en población, tales como GAs y la optimización basada en colonia de hormigas (ACO).

En primer lugar, para llevar a cabo este estudio es necesario analizar las fuentes de ruido en los dos problemas seleccionados para la experimentación, a saber: el problema de ensamblado de fragmentos de ADN (FAP) y el popular problema del viajante de comercio (TSP).

Luego, se estudia la robustez de una metaheurística entonces se analizan las diferencias entre las soluciones encontradas para las instancias sin y con ruido. Si no se detectan diferencias (estadísticamente significativas), la metaheurística muestra un comportamiento neutro (insensible, indistinto) a pequeñas variaciones en los datos de entrada. Consecuentemente, la metaheurística se considera robusta para resolver instancias ruidosas.

- **Control adaptativo de parámetros.** Tradicionalmente en los algoritmos evolutivos, los parámetros mantienen valores constantes durante la búsqueda. Sin embargo, dado un problema una dificultad importante es determinar a priori cuál es el valor más adecuado para cada parámetro. Además, existe una creciente demanda de software de optimización por parte de usuarios no especialista dentro de un entorno de desarrollo industrial.

En consecuencia, se propuso un algoritmo evolutivo de adaptación que incluye un mecanismo para modificar la probabilidad de mutación sin control externo. Este proceso de adaptación dinámica ocurre mientras que el algoritmo busca la solución al problema. De esta forma, se elimina una fase de cálculo muy costosa relacionada con el pre-ajuste de los parámetros algorítmicos.

Este nuevo algoritmo evita la convergencia prematura y la pérdida de la diversidad genética. Por otra parte, este método se implementa sin introducir cambios importantes en el diseño original del algoritmo. El algoritmo toma decisiones durante la búsqueda mediante el uso de una estrategia de control basada en información sobre la entropía de la población actual. El objetivo de nuestra estrategia adaptativa es aumentar la probabilidad de mutación si la diversidad genética se pierde gradualmente y, así, mantener una población distribuida en el espacio de búsqueda. De lo contrario, el valor de la probabilidad se reduce cuando se observa un aumento de la diversidad poblacional. Por lo tanto, estos cambios en el valor de probabilidad son también una fuente adicional para lograr un buen equilibrio entre la exploración y la explotación al monitorear la diversidad genotípica presente en la población. Es decir, cuando el valor de entropía se aproxima a 0, la población contiene individuos idénticos, de lo contrario la medida de entropía es positiva y obtiene el valor máximo cuando todos los individuos de la población son diferentes.

Por otra parte, propusimos una estrategia proactiva, denominada PROACT, para adaptar la política de migración de un Algoritmo Evo-

lutivo Distribuido (DEA). La decisión proactiva se hace a nivel local dentro de cada subpoblación, y está basada en la entropía de esa subpoblación. De esa manera, cada subpoblación puede pedir a sus vecinos que aumenten o disminuyan la frecuencia de migración con el fin de mantener la diversidad genética en un nivel deseado. Consecuentemente, se evita que las subpoblaciones queden atrapados en mínimos locales.

Finalmente, se combinaron estas dos estrategias en un DEA dando origen a Proact-PmAdap. El objetivo de Proact-PmAdap es permitir al algoritmo un control simple, sencillo y eficiente de la diversidad en la población. Proact-PmAdap actúa de forma proactiva cuando se espera una baja diversidad poblacional. La estrategia intenta incluir más diversidad genética en la isla con el propósito de la disposición de una población diversa de la que se pueden seleccionar soluciones en cada paso de migración.

- **Selección de operadores.** Con el objetivo de minimizar la duración y el costo de un proyecto de software, se propuso un AG tradicional usando codificación binaria para representar una solución al problema de planificación de proyectos de software (PSP). En particular, esta investigación se centró en la elección del operador de cruce y su probabilidad; para luego comparar el cambio en el rendimiento del GA al utilizar operadores genéticos tradicionales respecto de otros más específicos para el problema. Los experimentos mostraron que utilizar una recombinación tradicional es capaz de aumentar el rendimiento del algoritmo, manteniendo en niveles aceptables la velocidad de convergencia. El GA propuesto para resolver el PSP utiliza cromosomas binarios para representar las soluciones al problema. Estas soluciones son codificadas por medio de matrices que representan la dedicación de cada empleado para realizar una determinada tarea. Luego, para facilitar la manipulación de los cromosomas, estas matrices son traducidas a cadenas binarias. En consecuencia, pueden aplicarse distintos tipos de operadores de recombinación:

operadores de cruce tradicionales (un punto, dos puntos, uniforme, entre otros) que actuarán directamente sobre la cadena binaria que representa la matriz solución o bien aquellos diseñados para la recombinación de matrices. En otras palabras,

RESULTADOS OBTENIDOS/ESPERADOS

En esta sección presentamos los resultados obtenidos de nuestra investigación en el transcurso del año 2013 y los esperados en el 2014.

Con respecto al *análisis de robustez de ACO*. Luego de una intensa investigación sobre el origen de ruido en ambos problemas (FAP y TSP), se detectaron fehacientemente cuáles eran las fuentes de ruido en la información de los problemas. Esto permitió desarrollar distintos conjuntos de instancias con ruido¹ que posibilitaron el análisis de robustez de los algoritmos. En el caso de FAP, se estudió la robustez de SA, PALS, GA y un nuevo algoritmo que hibrida PALS con SA, denominado PH-PALS [5]. Este estudio arrojó que, a la hora de resolver instancias ruidosas de FAP, solamente tres de estos algoritmos son robustos: PALS, GA y PH-PALS. En tanto que para TSP, se espera encontrar evidencia empírica que demuestre la robustez de distintos algoritmos de la familia ACO.

Los resultados obtenidos por el AE que realiza un control adaptativo de la probabilidad de mutación son muy alentadores, ya que puede obtener soluciones de más alta calidad que variantes de EAs con probabilidades de mutación fija, y también exhibe velocidades de convergencia similares a la de estas variantes. Una observación importante es que nuestra propuesta no añade un tiempo computacional significativo al aplicar estrategia de control para adaptar los valores de probabilidad de mutación. Este año se espera poder contrastar los resultados de esta estrategias con las publicadas en la literatura.

En tanto que, **PROACT** y **Proact-PmAdap** permitieron mantener, eficientemente, la suficiente diversidad genética para evitar la convergencia prematura hacia mínimos locales. Esto es de especial interés en muchos escenarios de optimización donde es necesario reducir el número de

¹<http://mdk.ing.unlpam.edu.ar/lisi/principal.html>

evaluaciones de la función objetivo (por ejemplo, la optimización de la simulación) [6], [7].

En cuanto a la *selección de operadores para PSP*, el estudio consistió en analizar el comportamiento de un AG tradicional bajo dos tipos de cruce: el tradicional operador de un punto para representaciones binarias y un operador de cruce a nivel matricial ya que las soluciones al problema son matrices que representan la dedicación de cada empleado a las tareas. También se analizaron cuatro valores distintos de probabilidades de cruce ya que es un parámetro influyente en el rendimiento de un algoritmo. Los resultados muestran que para este problema es conveniente utilizar probabilidades de cruce altas. La variante algorítmica que utiliza el operador tradicional obtuvo el mejor rendimiento, obteniendo buena calidad de patrones de planificación. Los porcentajes de factibilidad de las soluciones obtenidas son altos para problemas de planificación no tan complejos, estos valores son similares a los obtenidos por otras variantes evolutivas de la literatura para resolver el problema en cuestión [8].

FORMACIÓN DE RECURSOS HUMANOS

Durante el año 2013, integrantes del proyecto han realizado diversos cursos de postgrado directamente relacionados con la temática del proyecto, con el objetivo de llevar a cabo carreras de postgrado en un futuro cercano.

En tanto que, en el LISI se trabaja con alumnos avanzados en la carrera Ingeniería en Sistemas 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.

REFERENCES

- [1] E.-G. Talbi, *Metaheuristics - From Design to Implementation*. Wiley, 2009.
- [2] E. Alba, *Parallel Metaheuristics: A New Class of Algorithms*, ser. WILEY Series on Parallel and Distributed Computing. Wiley, 2005.
- [3] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Comput. Surv.*, vol. 35, pp. 268–308, September 2003.
- [4] F. Glover and G. Kochenberger, *Handbook of Metaheuristics*. Norwell, MA: Kluwer Academic Publishers, 2002.
- [5] G. Minetti, G. Leguizamón, and E. Alba, "An improved trajectory-based hybrid metaheuristic applied to the noisy {DNA} Fragment Assembly Problem," *Information Sciences*, no. 0, pp.–, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0020025514001261>

- [6] C. Salto, F. Luna, and E. Alba, "Enhancing Distributed EAs Using Proactivity," in *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation*, ser. GECCO '13 Companion. New York, NY, USA: ACM, 2013, pp. 1747–1748. [Online]. Available: <http://doi.acm.org/10.1145/2464576.2480796>
- [7] C. Salto, F. Luna, and E. Alba, "Enhancing distributed EAs by a proactive strategy," *Cluster Computing*, pp. 1–11, 2013. [Online]. Available: <http://dx.doi.org/10.1007/s10586-013-0321-4>
- [8] G. Dupuy, N. Stark, and C. Salto, "Algoritmo evolutivo para el problema de planificación en proyectos de desarrollo de software," in *XIX Congreso Argentino de Ciencias de la Computación*, RedUNCI, Ed. Mar del Plata: Universidad CAECE, October 2013, pp. 110–119.