

Optimización evolutiva multi-objetivo en la planificación de controles a clase en la educación superior cubana

Pavel Novoa-Hernández

Universidad de Holguín, Departamento de Licenciatura en Matemática,
Cuba

pnovoa@facinf.uho.edu.cu

Resumen. En Cuba, un control a clase es el proceso en el que se evalúa la calidad de la docencia impartida por el profesor universitario. La planificación de los controles a clase es realizada comúnmente por el jefe de departamento, el cual debe cumplir con varios criterios a la vez. Algunos de estos criterios son: la presencia en el tribunal de al menos un miembro del mismo colectivo y/o con categoría igual o superior que el profesor evaluado, la disponibilidad y nivel de utilización de los miembros del tribunal, entre otros. Esta tarea se torna compleja cuando se tienen en cuenta un gran número de profesores y controles al mismo tiempo. Con el objetivo de resolver esta problemática en el presente trabajo se propone un enfoque computacional basado en: 1) la modelación de este proceso como un problema de optimización multi-objetivo, y 2) su solución mediante una variante del conocido algoritmo evolutivo NSGA-II. Los resultados de los experimentos computacionales realizados muestran que nuestras propuestas ofrecen soluciones útiles y de calidad.

Palabras clave. Planificación de controles a clase, construcción de equipos, optimización evolutiva multi-objetivo, NSGA-II.

Evolutionary Multi-objective Optimization for Scheduling Professor Evaluations in Cuban Higher Education

Abstract. In Cuba, university professors are often evaluated during the development of a class. Scheduling such evaluations along the academic semester is a common task of the head of the department, who must satisfy several criteria at the same time. Examples of these criteria are the presence of at least one tribunal member of the same academic unit, the presence of at least one tribunal member with equal docent category, availability and the utilization

level of the tribunal members, among others. However, scheduling professor evaluations can become a complex task if several professors and evaluations are considered. Aiming at solving this problem, in the present work we propose a computational approach based on 1) modelling this task as a multi-objective optimization problem and 2) solving this problem by adapting a variant of a very well-known evolutionary algorithm, NSGA-II. The results of the performed computational experiments show that our proposal contributes to obtaining useful quality solutions.

Keywords. Professor evaluations scheduling, team building, evolutionary multi-objective optimization, NSGA-II.

1. Introducción

En Cuba, un control a clase es el proceso en el que se evalúa la calidad de la docencia impartida por un profesor. Esta evaluación tiene lugar durante la clase y es realizada por un tribunal, generalmente formado por colegas del propio departamento o institución. Comúnmente, la planificación de los controles a clase es realizada por el jefe de departamento en períodos académicos (ej. semestralmente), y teniendo en cuenta varios criterios relacionados con los profesores. Ejemplos de estos criterios son: la presencia de al menos un miembro del tribunal con categoría docente igual o superior que el profesor evaluado, la presencia de al menos un miembro del tribunal del mismo colectivo o disciplina del profesor, entre otros. Además, en relación a la conformación de los tribunales, el planificador debe considerar también la disponibilidad de los miembros el día del control a clase, así como distribuir el trabajo lo más

equitativamente posible entre los posibles evaluadores. Para los propósitos de este trabajo, planificar un control a clase a un profesor determinado consiste en elegir un turno de clase dentro del período académico considerado, y asignarle un tribunal de evaluación.

Desde el punto de vista administrativo, la complejidad de esta tarea pudiera minimizarse aplicando el principio “divide y vencerás”, esto es, dejando que los colectivos o disciplinas del departamento sean los que planifiquen los controles de sus respectivos profesores. Sin embargo, varias razones apuntan a que resulta más adecuado considerar a todos los profesores del departamento en la planificación de los controles. Un ejemplo sencillo es el siguiente: considere un colectivo en el que su jefe es el profesor con la mayor categoría docente. Entonces, éste nunca podrá ser evaluado dentro de su colectivo, por no existir ningún otro profesor con igual o mayor categoría. Como consecuencia, habría que contar con la colaboración de profesores de otros colectivos. Asimismo, una planificación global tendría como beneficio adicional, una distribución más equitativa de la carga de trabajo de los profesores.

No obstante lo anterior, esta planificación puede tornarse compleja si existe un gran número de profesores, y por tanto, de controles a clase. A esto se le suma el hecho de que cada profesor tiene una agenda de trabajo diferente al resto de sus compañeros, esto es, cantidad de turnos de clase y otras actividades en horarios diferentes. De manera que no todos los profesores estarían disponibles siempre para formar los tribunales. Adicionalmente, intentar cumplir al mismo tiempo con los criterios mencionados anteriormente resulta generalmente difícil para un solo ser humano. Teniendo en cuenta esta problemática, en el presente trabajo se propone un enfoque computacional basado en: 1) la modelación de este proceso como un problema de optimización multi-objetivo, y 2) su solución mediante la adaptación del algoritmo NSGA-II [1]. El objetivo de esta propuesta es automatizar el proceso de planificación de controles a clase, favoreciendo así la toma de decisiones del jefe de departamento en cuestión. En la literatura existen investigaciones similares, las cuales tratan de manera independiente, o la

construcción de equipos [2, 3, 4] o la planificación docente [5, 6, 7]. Sin embargo, hasta donde se conoce, no existen investigaciones que se planteen resolver ambos problemas al mismo tiempo, como es el caso del presente trabajo.

Para una mejor exposición de nuestra propuesta, el resto del artículo queda organizado de la forma siguiente: en la Sec. 2, se describe el modelo matemático del problema propuesto. Más adelante, la Sec. 3 está explicar el método de solución, así como las adaptaciones realizadas al mismo. Por su parte, la Sec. 4 expone los resultados obtenidos a partir de los experimentos computacionales realizados. Con el objetivo de validar nuestras propuestas se resuelven varios escenarios relacionados con departamentos docentes de la Universidad de Holguín (Cuba). Finalmente, en la Sec. 5 se brindan las conclusiones y trabajos futuros de la investigación.

2. El problema de la planificación de controles a clase

En esta sección se describe más detalladamente el problema a resolver, incluyendo el modelo matemático propuesto.

2.1. Descripción general del problema

El problema de la planificación de los controles a clase es el siguiente¹:

Sean conocidos por cada profesor del departamento, su categoría docente, cantidad de controles a clase a realizarle, pertenencia a un colectivo o disciplina, y agenda de trabajo durante el semestre. Planificar en las fechas disponibles de su agenda, los controles a clase a realizarle, y asignarle un tribunal de evaluación, teniendo en cuenta los siguientes requerimientos:

1. en el tribunal debe existir al menos un miembro con categoría igual o superior a la del profesor evaluado,

¹ Aunque la descripción de este problema puede variar en otras universidades cubanas, en esencia, las diferencias son despreciables.

2. en el tribunal debe haber al menos un miembro perteneciente al colectivo del profesor evaluado,
3. se debe minimizar la cantidad de miembros del tribunal con actividades el mismo día del control a clase, especialmente cuando estas sean actividades docentes (ej. clases),
4. distribuir equitativamente la participación de los profesores del departamento en los tribunales,
5. distribuir uniformemente los controles a clase a lo largo del período académico.

Desde el punto de vista matemático, a partir de esta descripción informal se puede obtener un problema de optimización, específicamente de tipo combinatorio multi-objetivo. Nótese que los requerimientos anteriores pueden ser tratados como funciones objetivo independientes, y/o restricciones.

2.2. Modelo matemático propuesto

El modelo matemático propuesto para el problema descrito anteriormente, se detalla a continuación.

Datos. Se asumirán conocidos los siguientes datos.

n : número de profesores. Cada profesor será identificado por un número natural i en el rango $[1, n]$.

m : número de turnos de clase en el semestre. Cada turno de clase estará identificado por un número natural en el rango $[1, m]$.

c_i : categoría docente del profesor i . En particular $c_i \in \{0, 1, 2, 3, 4\}$, donde los valores de c_i significa lo siguiente:

- 0, el profesor es *Adiestrado*,
- 1, el profesor es *Instructor*,
- 2, el profesor es *Asistente*,
- 3, el profesor es *Profesor Auxiliar*,
- 4, el profesor es *Profesor Titular*.

d_i : colectivo al que pertenece el profesor i . En este caso $d_i \in \mathbb{N}$ donde un valor natural está asociado a un colectivo determinado.

u_i : cantidad de controles a realizar al profesor i . Donde $u_i \in \{1, 2, 3\}$, teniendo como significado:

- $u_i = 1$, si el profesor es *Profesor Auxiliar* o *Profesor Titular*,
- $u_i = 2$, si el profesor es *Instructor* o *Asistente*,
- $u_i = 3$, si el profesor es *Adiestrado*.

a_{ij} : disponibilidad (agenda) del profesor i en el turno j ($j = 1, 2, \dots, m$). En particular, $a_{ij} \in \{0, 1, 2\}$ con el siguiente significado:

- 0 indica que el profesor está Disponible,
- 1 indica que el profesor tiene alguna Actividad profesional o personal (no está disponible),
- 2 indica que el profesor tiene clase.

Variables de decisión. Las variables de decisión se han definido mediante una tupla x_k formada por índices.

Formalmente, $x_k = \langle p_e, p_1, p_2, p_3, t \rangle$, que significa que el control a clase k , será realizado al profesor p_e en el turno t , en el que participan como tribunal, los profesores p_1, p_2, p_3 .

Es de notar que los elementos variables son p_1, p_2, p_3 y t , pues p_e se incluye solo por motivos de identificación, esto es, para determinar a qué profesor se le realizará el control a clase. Además:

- $k = 1, 2, \dots, D$, siendo $D = \sum_n u_i$ el total de controles a clase (dimensión del vector),
- $p_e, p_1, p_2, p_3 \in \{1, 2, \dots, n\}$ y $t \in \{1, 2, \dots, m\}$.

De manera que una solución vendría dada por un vector D -dimensional: $\mathbf{x} = \{x_1, x_2, \dots, x_D\}^T$.

Restricciones. Como la mayoría de los escenarios de decisión reales, el problema de los controles a clase posee varias restricciones. En ese sentido, es importante mencionar que las restricciones funcionales, como las relacionadas con los criterios 1 y 2 de la Sec. 2.1, serán consideradas como funciones objetivo. La idea es tratar de manera flexible el cumplimiento de estos objetivos, ya que en la práctica pudiera darse el caso de que algunos tribunales pudieran funcionar sin cumplir con los mismos. Sin embargo, existen otros tipos de restricciones que provocarían la infactibilidad de las soluciones en la realidad, tal es el caso de las siguientes:

- un profesor no puede ser al mismo tiempo tribunal y evaluado ($p_e^k \neq p_i^k$, con $i = 1, 2, 3$),

- un mismo profesor solo puede ser elegido una vez para un mismo tribunal ($p_1^k \neq p_2^k, p_1^k \neq p_3^k, p_2^k \neq p_3^k$),
- no se puede planificar un control a clase en un turno en el que el profesor a evaluar no tenga clase. Más formalmente, la disponibilidad del profesor p_e en el turno t^k , tiene que ser igual a 2 ($a_{p_e, t^k} = 2$).

Función objetivo. Teniendo en cuenta los criterios mencionados en la Sec. 2.1, se definen los siguientes objetivos:

1. *Coincidencia entre la categoría docente del tribunal y del profesor.*

$$\min f_1(\mathbf{x}) = \sum_{k=1}^D \max\{c_{p_e}^k - c_{max}^k, 0\}, \quad (1)$$

donde $c_{p_e}^k$ es la categoría docente del profesor p_e , mientras que c_{max}^k es la máxima categoría docente dentro del tribunal asignado al control k . Formalmente, $c_{max}^k = \max\{c_{p_1}^k, c_{p_2}^k, c_{p_3}^k\}$.

La interpretación de la expresión (1) es sencilla. Se trata de comparar la máxima categoría del tribunal con la categoría del profesor a evaluar. Mediante la resta de estas categorías, se obtiene tres posibles resultados: un valor negativo (el profesor tiene menor categoría), 0 (el profesor tiene la misma categoría), y un valor positivo (el profesor tiene mayor categoría). Nótese que es solo a partir de este último resultado que se obtendría un error distinto de 0, lo cual es lo deseable para caracterizar al criterio 1) de la Sec. 2.1.

2. *Coincidencia entre el colectivo del tribunal y del profesor*

$$\min f_2(\mathbf{x}) = \sum_{k=1}^D \Theta_{colect}(d_{p_e}^k, D^k), \quad (2)$$

donde $d_{p_e}^k$ es el colectivo del profesor p_e y, $D^k = \{d_{p_1}^k, d_{p_2}^k, d_{p_3}^k\}$ los de los miembros del tribunal p_1, p_2, p_3 , respectivamente. En particular, Θ_{colect} es una función binaria, que devuelve 0 si existe al menos un miembro del tribunal del mismo colectivo que el profesor, y 1 cuando ocurre lo contrario. Formalmente, está definida por la expresión (3):

$$\Theta_{colect}(d, D) = \begin{cases} 0 & \text{si } \exists d_e \in D: d_e = d, \\ 1 & \text{en caso contrario.} \end{cases} \quad (3)$$

3. *Disponibilidad del tribunal el día del control a clase.*

$$\min f_3(\mathbf{x}) = \sum_{k=1}^D \sum_{i=1}^3 a_{p_i, t^k}, \quad (4)$$

donde p_i es el miembro del tribunal i , y t^k es el turno de clase en el que se desarrollará el control k . De manera que a_{p_i, t^k} es la disponibilidad del profesor p_i en el turno t^k . En este caso, la idea es contabilizar la cantidad de actividades de los miembros del tribunal durante el turno t^k . Esta operación se facilita en gran medida por la definición de los valores de a_{ij} (véase más arriba el apartado correspondiente a *Datos*).

4. *Nivel de utilización de los profesores en los tribunales.*

$$\min f_4(\mathbf{x}) = \sum_{i=1}^n |\mu_i^{ideal} - \mu_i^k|, \quad (5)$$

Aquí, μ_i^{ideal} y μ_i^k son el nivel de utilización ideal y real para el profesor i , respectivamente. En particular, μ_i^{ideal} se determina según la siguiente expresión:

$$\mu_i^{ideal} = 3 \left(\frac{D - u_i}{n} \right). \quad (6)$$

La expresión (6) es, en esencia, una aproximación del nivel de utilización ideal de un profesor determinado, teniendo en cuenta, tanto la cantidad de controles a realizar (D), así como aquellos que hay que realizarle a éste (u_i).

5. *Distribución equitativa de los controles a clase*

$$\min f_5(\mathbf{x}) = \delta_{ideal} - \min\{t_{l+1}^{ord} - t_l^{ord}\}, \quad (7)$$

$(l = 1, \dots, D - 1).$

Aquí, $\delta_{ideal} = \frac{m}{D}$ es el índice ideal de distribución de los controles a clase durante el período académico. Nótese que se trata de una simple división entre el total de turnos de clase y el número de controles a clase. Esto significa que lo "ideal" es tener un control a clase cada $\frac{m}{D}$ turnos. Por su parte, t_{l+1}^{ord} y t_l^{ord} representan

turnos de clase consecutivos de la solución \mathbf{x} . Note en este sentido que los turnos están marcados con el supra-índice *ord*, indicando que éstos han sido ordenados de manera ascendente. Este ordenamiento se hace con el objetivo de realizar la resta que aparece en la fórmula y obtener así, la distancia entre cada par de controles a clase. Obviamente, si se cuenta con D controles a clase, entonces se obtendrían $D - 1$ distancias (pares). Finalmente, el operador $\min\{\}$ indica que nos enfocaremos en la menor de estas diferencias, la cual al ser restada al índice ideal, nos daría una noción del error de \mathbf{x} con respecto a la distribución equitativa de los controles a lo largo del período académico.

A partir de las funciones objetivo anteriores se puede proseguir de varias formas. Por ejemplo, se pudieran agrupar todas en una sola función objetivo más general (ej. mediante una suma simple o ponderada). Asimismo, se pudieran dejar como tal, y que el método de solución intentara minimizar a todas las funciones al mismo tiempo (ej. aplicando los conceptos de dominancia² y optimalidad³ de Pareto). Sin embargo, en este caso particular, se trata de cinco objetivos, lo cual representa un número alto tanto en la teoría como en la práctica. Véase [8,9] para un análisis más detallado sobre el tema. Una variante intermedia sería crear objetivos más generales, esto es, a partir de la combinación de algunas funciones. Por ejemplo, pudieran obtenerse 3 objetivos a partir de la suma de los criterios 1) y 2), 3) y 4), y finalmente el 5). Esta es precisamente la variante que se seguirá en la presente investigación. Llamaremos $f_{categ-colect}$ al primero de estos objetivos, $f_{disp-util}$ al segundo, y f_{contr} al tercero. De manera que la función objetivo del modelo propuesto es la siguiente:

$$\begin{aligned} \min F(\mathbf{x}) \\ = [f_{categ-colect}(\mathbf{x}), f_{disp-util}(\mathbf{x}), f_{contr}(\mathbf{x})]. \end{aligned} \quad (8)$$

2.3. Sobre la complejidad del problema

Antes de finalizar esta sección, resulta importante hacer algunas consideraciones sobre la complejidad del problema. En sentido general, la complejidad computacional de los problemas combinatorios viene dada por el tamaño del espacio de búsqueda (soluciones) que, aunque finito, suele ser un número muy grande.

Como se pudo advertir en el modelo propuesto, las soluciones del problema son combinaciones de varios elementos: profesores (tribunal) y turnos de clase. Desde el punto de vista combinatorio, cada control a clase (variable) es independiente del resto. Formalmente, la cantidad de combinaciones (posibles valores) de una variable x_k es:

$$\omega_k = \frac{(n-1)!}{(n-4)!3!} \|T_{pe}^{dispo}\|, \quad (9)$$

donde n es la cantidad de profesores y T_{pe}^{dispo} los turnos de clase en los que el profesor a evaluar, tiene clase. Nótese que el primer factor de la expresión (9) es el número de combinaciones de 3 elementos a partir de $n - 1$, la cual es a su vez la cantidad de profesores disponibles para el tribunal k , sin contar al profesor a evaluar. También es importante notar que en estas combinaciones están algunas soluciones infactibles o al menos poco deseables, como es el caso de aquellos turnos de clase donde los miembros del tribunal no están disponibles.

De manera que la cantidad de soluciones del espacio de búsqueda de nuestro problema, vendría dada por:

$$\|\Omega\| = \prod_{k=1}^D \omega_k. \quad (10)$$

La expresión (10) muestra que la cantidad de posibles soluciones puede crecer rápidamente conforme el número de profesores y de controles a clase aumentan. Por ejemplo, considere el caso en el que a 10 profesores, con 80 turnos de clase cada uno, se le deben realizar 2 controles a

²Una solución A pareto-domina a otra B (formalmente $A > B$), cuando A es al menos tan buena como B en todos los objetivos, y superior a B en al menos uno de estos objetivos.

³ Una solución A es Pareto-óptima cuando no existe otra solución B que la domine. Al conjunto de soluciones Pareto-óptimas se le conoce como Conjunto de Pareto (PS), mientras que a la imagen de este conjunto se le conoce como Frente de Pareto (PF).

clase. Tendríamos entonces: $n = 10$, $\|T_{pe}^{dispo}\| = 80$ y $D = 2 \cdot n = 20$. Aplicando (9) y (10) el espacio de búsqueda correspondiente estaría formado por 3.5268×10^{76} soluciones. En el caso del mismo problema, pero con el doble de profesores ($n = 20$ y $D = 40$), tendríamos un espacio de búsqueda de 3.7718×10^{195} .

Para comprender mejor la dificultad de trabajar con estos espacios de búsqueda, supongamos que contamos con un algoritmo simple que genera todas las posibles soluciones, las evalúa y compara para obtener la mejor solución. Supongamos además que el tiempo de evaluar una solución es de un nanosegundo (1 ns), equivalente a 3.1689×10^{-17} años. Entonces, para el caso de 10 profesores, evaluar todas las posibles soluciones nos tomaría 1.1176×10^{60} años, mientras que en el segundo caso (20 profesores) unos 1.1952×10^{179} . Estos ejemplos muestran la intratabilidad de nuestro problema desde el punto de vista de algoritmos que consideren todo el espacio de búsqueda. Se infiere por tanto la aplicación de métodos de solución aproximados, con el objetivo de lograr un compromiso entre la calidad de las soluciones y el tiempo de ejecución.

3. Método de solución

La elección de modelar el problema de la planificación de controles a clase como un problema de optimización multi-objetivo, conlleva a la selección de un método que tenga en cuenta a dichos objetivos de manera independiente. De igual forma, la complejidad computacional del problema requiere de la aplicación de métodos aproximados, que ofrezcan soluciones cercanas a las óptimas en un tiempo razonablemente corto.

En este sentido, las denominadas metaheurísticas [10, 11] son métodos adecuados. Específicamente, los *Algoritmos Evolutivos Multi-objetivo* (MOEAs)⁴, considerados técnicas metaheurísticas, sobresalen por su alta efectividad y fácil implementación, lo cual los convierten en buenos candidatos para el problema en cuestión a diferencia de otras técnicas exactas [12].

⁴Siglas en inglés de *Multiobjective Optimization Evolutionary Algorithms*.

Entre estos métodos, el algoritmo NSGA-II⁵ [1], basado en los conceptos de dominancia y optimalidad de Pareto, ha sido ampliamente aplicado con éxito en diversos contextos. Véase por ejemplo en el ámbito de la *Planificación*, los trabajos [13] y [14]; en el contexto de la *Administración* a [15], en *Bioinformática* a [16], y en la *Robótica y Diseño de Controladores* a [17, 18].

Específicamente, la variante de NSGA-II que se empleará es la denominada *elitista controlada*. En general, los algoritmos genéticos elitistas dan prioridad la obtención de individuos (soluciones) con mejores valores de la función objetivo (fitness) que, en el caso del NSGA-II, es una posición (o rango) obtenida a partir de los objetivos del problema. Asimismo, en estos algoritmos se favorece la creación de individuos que incrementen la diversidad de la población (conjunto de soluciones), incluso a pesar de que estos no posean un buen valor de fitness. En particular, mantener la diversidad de la población es un requisito necesario para obtener un frente de Pareto óptimo. En NSGA-II, la diversidad es mantenida mediante el control de los *individuos élite* de la población (ej. los mejores). En este sentido, el número de individuos élite es controlado por un parámetro predefinido, que limita el tamaño del frente de Pareto. Mientras que la diversidad es mantenida por una función de distancia que permite determinar los individuos que están relativamente distantes entre sí en el frente de Pareto. Los pasos generales del algoritmo NSGA-II se muestran en el pseudocódigo de la Figura 1. Nótese que el objetivo del algoritmo es determinar el mejor Conjunto de Pareto F_{mejor} , partiendo de una población inicial de soluciones, y aplicando de manera iterativa la generación de nuevos individuos a partir de operadores evolutivos (paso 8, Fig. 1).

Es importante resaltar que tanto la función de creación de la población inicial, como los operadores de Cruzamiento y Mutación han sido definidos teniendo en cuenta varios aspectos específicos del problema tratado. Un ejemplo bien conocido es la codificación (representación) de

⁵ Siglas en inglés de *Non-Dominated Sorting Genetic Algorithm*.

Algoritmo NSGA-II (variante elitista controlada)

-
- 1: $P \leftarrow \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\mu\}$ Generar aleatoriamente μ individuos (población inicial)
 - 2: Evaluar en la función objetivo, los individuos de P
 - 3: $R \leftarrow \{R_1, R_2, \dots\}$ Determinar los *conjuntos Pareto* a partir de los rangos (posición) de P
 - 4: Calcular la distancia de separación de los individuos por cada conjunto R_i
 - 5: $F_{mejor} \leftarrow$ conjunto de Pareto de P , esto es, R_1
 - 6: $G \leftarrow 0$ contador del número de generaciones (iteraciones)
 - 7: **mientras** $G \leq G_{max}$ **o** F_{mejor} no es el deseado **hacer**
 - 8: $Q \leftarrow$ generarHijos(P) mediante *Selección, Cruzamiento y Mutación*
 - 9: Evaluar en la función objetivo a los individuos de Q
 - 10: $Q \leftarrow Q \cup P$
 - 11: $P \leftarrow \emptyset$
 - 12: $R \leftarrow \{R_1, R_2, \dots\}$ Determinar los *conjuntos Pareto* a partir de los rangos (posición) de Q
 - 13: $F_{mejor} \leftarrow$ Conjunto de Pareto de Q
 - 14: **por cada** conjunto de Pareto $R_i \in R$ **hacer**
 - 15: Calcular la distancia de separación de los individuos de R_i
 - 16: **si** $\|P\| + \|R_i\| \geq \mu$ **entonces**
 - 17: Añadir a P , los $m - \|P\|$ individuos de R_i más diferentes (según distancia)
 - 18: **salir** del ciclo
 - 19: **si no**
 - 20: $P \leftarrow P \cup R_i$
 - 21: $G \leftarrow G + 1$
 - 22: **retornar** F_{mejor}
-

Fig. 1. Algoritmo NSGA-II [1], variante elitista controlada

las soluciones, la cual siempre afecta la definición de los operadores evolutivos. A continuación se detallan las principales adaptaciones realizadas:

Creación de la población inicial. El objetivo de este paso es crear soluciones aleatorias, esto es, a partir de un muestreo uniforme del espacio de búsqueda. En ese sentido, se implementó una función a nivel de componente (x_k , control a clase), la cual, dado un profesor a evaluar, selecciona de manera aleatoria, un día (turno de clase) y tres profesores para el tribunal asociado.

Con la intención de hacer lo más factible posible esta planificación, se tuvieron en cuenta las consideraciones descritas en la Sec. 2.1, específicamente en el apartado *Restricciones*. De manera que, para generar una solución (individuo de la población), esta función es llamada D veces (una vez por cada componente del vector solución).

Cruzamiento. El objetivo de este operador es crear nuevas soluciones a partir de otras existentes, esto es, mediante el intercambio o

combinación de determinadas variables. En este caso se implementó una función basada en el conocido cruzamiento disperso (*Scattered crossover*), el cual genera un nuevo individuo a partir de la mezcla aleatoria de las componentes de dos padres. Dado que las componentes (planificaciones de controles a clase) constituyen planificaciones factibles, este tipo de cruzamiento se puede realizar sin grandes consecuencias. Este operador depende de un parámetro, conocido como tasa de cruzamiento, que controla la cantidad de individuos a generar.

Mutación. Este operador tiene por objetivo introducir pequeñas y nuevas variaciones a soluciones existentes. Es una forma explícita de introducir diversidad a la población. Aquí se reutilizó la función descrita anteriormente para el caso de la *Creación de la población inicial*. Solo que esta vez, dada una solución, se selecciona aleatoriamente una de sus componentes y se reemplaza por otra generada de manera aleatoria.

4. Experimentos computacionales

Una vez descritas nuestras propuestas, este apartado tiene por objetivo mostrar los resultados obtenidos en la aplicación del método considerado. Los experimentos computacionales realizados están divididos en tres partes. La primera tiene por objetivo, identificar qué configuración de parámetros del algoritmo NSGA-II, resulta más efectiva en la solución del problema. A partir de la identificación de esta “mejor” configuración, en una segunda parte se mostrarán con un sentido más realista, un ejemplo de solución. Finalmente, en un tercer apartado, ilustraremos cómo se afecta la calidad de las soluciones cuando se resuelven escenarios más complejos.

Dado que el problema a resolver fue modelado en términos generales, resulta importante establecer instancias del mismo que puedan ser resolubles por el método elegido. En ese sentido, se consideró como caso de estudio, la planificación de controles a clase en el Departamento de Licenciatura en Matemática de la Universidad de Holguín. Específicamente durante el primer semestre de un curso escolar.

Tabla 1. Características del Departamento de Licenciatura en Matemática, Universidad de Holguín

Características	Valores
Número de profesores (n)	18
Profesores Titulares	5 (profesores 1,2,3,4,5)
Profesores Auxiliares	4 (profesores 6,7,8,9)
Profesores Asistentes	3 (profesores 10,11,12)
Profesores Instructores	1 (profesor 13)
Profesores Adiestrados	5 (profesores 14,15,16,17,18)
Cantidad de controles a clase a planificar (dimensión del vector solución) (D)	32
Cantidad de turnos de clase (m)	480
Miembros del colectivo 1 ($d_i = 1$)	10 (profesores 2,3,5,7,8,10,11,12,13,17)
Miembros del colectivo 2 ($d_i = 2$)	8 (profesores 1,4,6,9,14,15,16,18)
Promedio de turnos con clase ($\ T_{pe}^{dispo}\ $)	80
Tamaño del espacio de búsqueda ($\ \Omega\ $)	3.4607×10^{151}

Las características de este departamento se muestran en la Tabla 1. Nótese que el departamento cuenta con un total de 18 profesores, los cuales se identificarán por un número natural, en el rango $[1,18]$. Por tal motivo, en los apartados relacionados con las categorías docentes y los colectivos, no solo se especifica la cantidad de profesores, sino también cuales de estos pertenecen a cada grupo. Otras informaciones adicionales, como la agenda personal de cada profesor se omiten por razones de espacio.

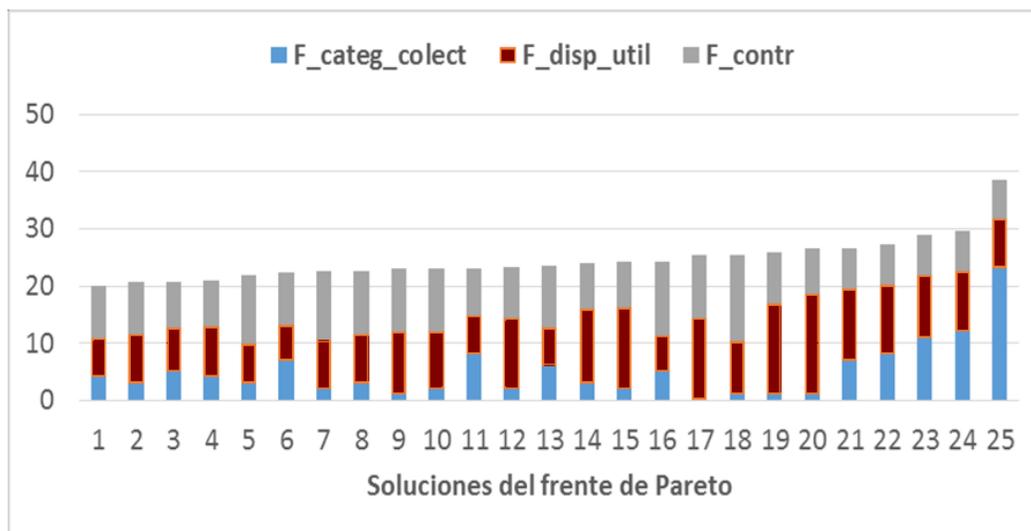


Fig. 2. Mejor aproximación al Frente óptimo de Pareto, según los objetivos definidos en la expresión (8). La aproximación es obtenida a partir de varias configuraciones del algoritmo NSGA-II y múltiples ejecuciones de las mismas

4.1. Análisis de las configuraciones del algoritmo NSGA-II

Como se advirtió anteriormente, la primera parte de los experimentos está relacionada con la determinación de la configuración más efectiva del algoritmo NSGA-II. En este caso se exploraron varias configuraciones de dos parámetros bien sensibles en este algoritmo: el tamaño de la población, y la tasa de cruzamiento. En relación al tamaño de la población, se consideraron los siguientes valores: 160, 320, y 480, que corresponden a 5, 10 y 15 veces el número de variables (i.e. 32). Por otra parte, las tasas de cruzamiento consideradas fueron: 0,4; 0,6 y 0,8. La combinación de estos factores genera un total de 9 variantes o configuraciones del algoritmo NSGA-II, los cuales fueron ejecutados un total de 20 veces, con semillas aleatorias diferentes.

Aunque la determinación de las medidas más idóneas para analizar y comparar algoritmos en la optimización evolutiva multi-objetivo es un tema aún en desarrollo, en la literatura se reportan ciertos avances. Véanse por ejemplo los trabajos [9, 11]. Sin embargo, la mayoría de las medidas asumen conocido el Frente óptimo de Pareto

(F_{OP}), lo cual suele ser difícil de determinar en problemas reales. Para solucionar esta dificultad existen varias alternativas, siendo una de ellas, encontrar una buena aproximación de este frente óptimo. Es precisamente este procedimiento el adoptado en la presente investigación. Así, se obtuvo la aproximación mostrada en la Figura 2. Dicha aproximación fue obtenida a partir de todos los frentes generados por las configuraciones del algoritmo NSGA-II, y todas las ejecuciones de las mismas. Para lograrlo, se compararon según el concepto de dominancia de Pareto todos los frentes de Pareto obtenidos. Nótese que este gráfico muestra 25 soluciones que no se dominan entre sí y que por tanto, poseen distintos valores de las funciones objetivo definidas en la expresión (8). Los valores de las funciones objetivo son representados por colores. Desde el punto de vista de la teoría de optimalidad de Pareto, todas estas soluciones son igual de buenas. La elección de cual implementar queda entonces por parte del decisor.

Una vez obtenida esta aproximación del frente óptimo de Pareto, se empleó como medida de convergencia el promedio de la distancia euclidiana al frente óptimo de Pareto. Esta medida, basada en la conocida *medida de distancia* [11] se define de la forma siguiente:

$$I_s = \frac{1}{\mu} \sum_{i=1}^{\mu} \min_{v \in S_{OP}} (\|F(v) - F(x_i)\|_2). \quad (9)$$

Aquí, S_{OP} es la mejor aproximación al conjunto óptimo de Pareto, y x_i es el i -ésimo individuo del conjunto de Pareto obtenido por la configuración s . Nótese que la idea es promediar, la menor distancia de las soluciones obtenidas por la configuración s y el frente F_{OP} . Por supuesto, un valor bajo de esta medida significa un buen nivel

de convergencia del algoritmo hacia al mejor frente de Pareto conocido. Los resultados en relación a esta medida de convergencia, para las distintas configuraciones estudiadas, se muestran en el gráfico de barras de la Figura 3. Como se aprecia, las mejores configuraciones son las que usan un tamaño de población mayor, esto es, 480 individuos. En particular, entre estas configuraciones sobresale la que emplea la tasa de cruzamiento igual a 0,6.

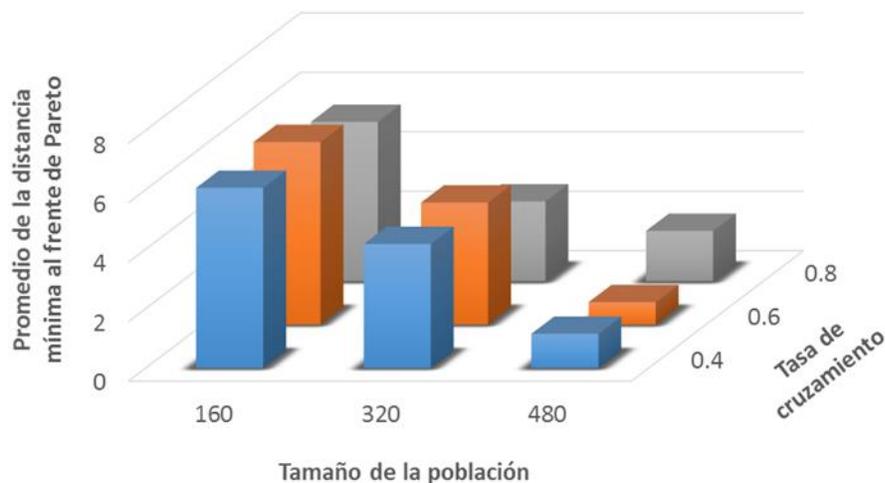


Fig. 3. Comparación entre las distintas configuraciones del algoritmo NSGA-II atendiendo al tamaño de la población y la tasa de cruzamiento

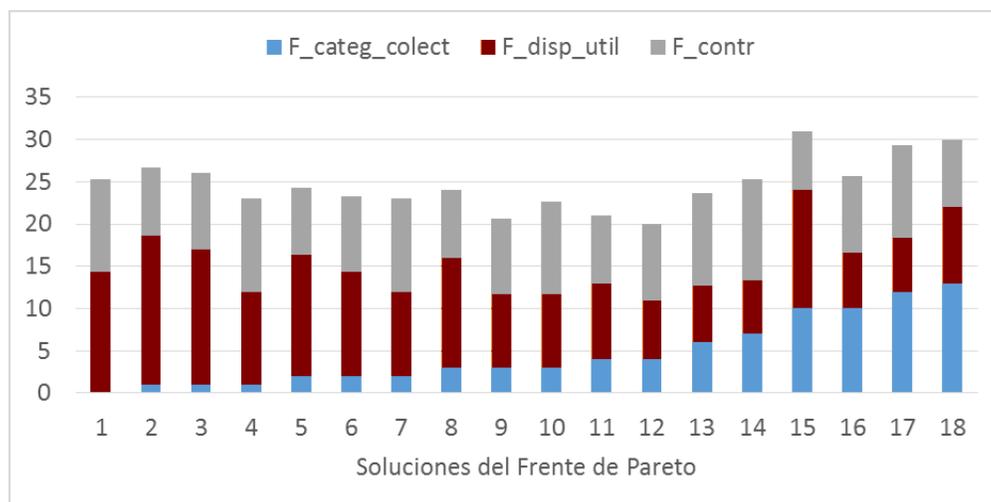


Fig. 4. Frente de Pareto obtenido por la variante NSGA-II (480-0,6) en 20 ejecuciones. Las soluciones fueron ordenadas ascendentemente según el objetivo $F_{categor-colect}$

Tabla 2. Solución obtenida por la variante NSGA-II (480-0,6). Los datos corresponden a la Solución 1 del frente de Pareto generado por esta variante

Tribunal del control a clase					
Control a clase	Profesor a evaluar	Miembro 1	Miembro 2	Miembro 3	Turno de clase
1	10	13	17	3	27
2	12	17	13	6	35
3	13	12	3	6	55
4	18	9	4	16	62
5	14	13	12	8	91
6	11	17	12	8	117
7	7	15	2	9	121
8	4	2	1	15	131
9	17	14	10	13	139
10	9	2	7	16	153
11	15	4	18	9	167
12	5	14	13	3	180
13	1	4	18	15	189
14	18	15	9	2	211
15	13	10	8	14	234
16	17	5	14	13	248
17	11	14	6	17	261
18	14	5	10	11	268
19	18	7	15	16	279
20	16	15	7	1	283
21	12	10	11	3	308
22	3	5	13	6	319
23	15	9	1	7	349
24	8	5	17	3	360
25	16	7	18	1	368
26	15	2	18	4	390
27	10	13	12	14	397
28	6	11	8	5	403
29	14	3	5	10	423
30	2	15	4	16	431
31	16	1	7	2	461
32	17	6	8	11	470

Precisamente, esta variante, que llamaremos NSGA-II(480-0,6), ha sido empleada para generar el frente de Pareto de la Figura 4. Véase que las 18 soluciones encontradas, se ordenaron según el criterio 1) Sec. 2.1, relacionado con la presencia en el tribunal de profesores con al menos igual categoría y del mismo colectivo, que el profesor a evaluar.

Obsérvese que la solución 1, tiene un error igual a 0 en este criterio, por lo que se cumple totalmente con este requerimiento. Igualmente, existen otras soluciones atractivas como la 9 y la 12 que muestran un equilibrio general en los tres objetivos.

4.2. Ejemplo de solución del problema

Para ilustrar más detalladamente la estructura real de las soluciones obtenidas, en la Tabla 2 se muestra la solución 1 correspondiente al frente de Pareto de la Figura 4. Es importante observar que los controles a clase han sido ordenados por la columna de los turnos de clase. Cada fila corresponde a un control a clase en específico.

Por ejemplo, la primera fila indica que el primer control a clase se le realizará al profesor 10, por un tribunal formado por los profesores 13, 17 y 3, en el turno de clase 27. Este turno de

Tabla 3. Configuración de parámetros correspondientes a escenarios más complejos del problema de la planificación a controles a clase

Características	Escenarios	
	Departamento grande	Dos departamentos
Número de profesores (n)	32	50
Profesores Titulares	1	6
Profesores Auxiliares	10	14
Profesores Asistentes	11	14
Profesores Instructores	7	8
Profesores Adiestrados	3	8
Cantidad de controles a clase a planificar (D)	56	88
Cantidad de turnos de clase (m)	480	480
Miembros del colectivo 1 ($d_i = 1$)	14	24
Miembros del colectivo 2 ($d_i = 2$)	18	26
Promedio de turnos con clase ($\ T_{pe}^{dispo}\ $)	80	80
Tamaño del espacio de búsqueda ($\ \Omega\ $)	$\approx 10^{311}$	$\approx 10^{543}$

clase, representado por un número natural, tiene asociado una hora y día únicos en el semestre.

Para comprobar la factibilidad de tal asignación desde el punto de vista del primer criterio, basta con comprobar en la Tabla 1, las categorías docentes y los colectivos de los profesores involucrados. En ese sentido, el profesor 10 es Asistente y pertenece al colectivo 1, requerimientos que cumple el tribunal asociado, debido a que el profesor 3 es: Profesor Titular y miembro del colectivo 1. De manera similar pueden comprobarse el resto de las filas. Es por ello que el valor de la solución en este criterio es $F_{categ-colect} = 0$. En relación al criterio 2, el valor de la solución en $F_{disp-util}$ es de 14,33. En cuyo caso, este valor proviene solamente del nivel de utilización, ya que el nivel de coincidencia (disponibilidad) igual a 0. Finalmente, nótese que los turnos de clase se encuentran bien espaciados entre sí, como se desea en el criterio 3. En particular, el promedio de distanciamiento entre turnos es de 14,29; muy cercano al ideal 15.

Estos resultados, aunque derivados de una sola instancia de problema, evidencian que el método empleado, bajo las adaptaciones

propuestas, permite obtener soluciones útiles y con una calidad apreciable.

4.3. Resultados en escenarios más complejos

En este apartado, nos centraremos en escenarios más complejos que el analizado anteriormente. En este sentido hemos considerado los escenarios mostrados en la Tabla 3, los cuales corresponden igualmente a la Universidad de Holguín. Estos son:

1. *Departamento grande*: Departamento de Matemática (Prestación de servicios).
2. *Dos departamentos*: Licenciatura en Matemática (Carrera), y Matemática (Prestación de servicios).

Como se aprecia en la Tabla 3, estos escenarios poseen más profesores y por tanto, más controles a realizar. Como consecuencia del incremento de estos parámetros, el tamaño del espacio de búsqueda aumenta significativamente. Véase la última fila de la Tabla 3.

Con el objetivo de verificar si las configuraciones del algoritmo NSGA-II de los apartados anteriores, mantienen el mismo comportamiento en estos dos escenarios más

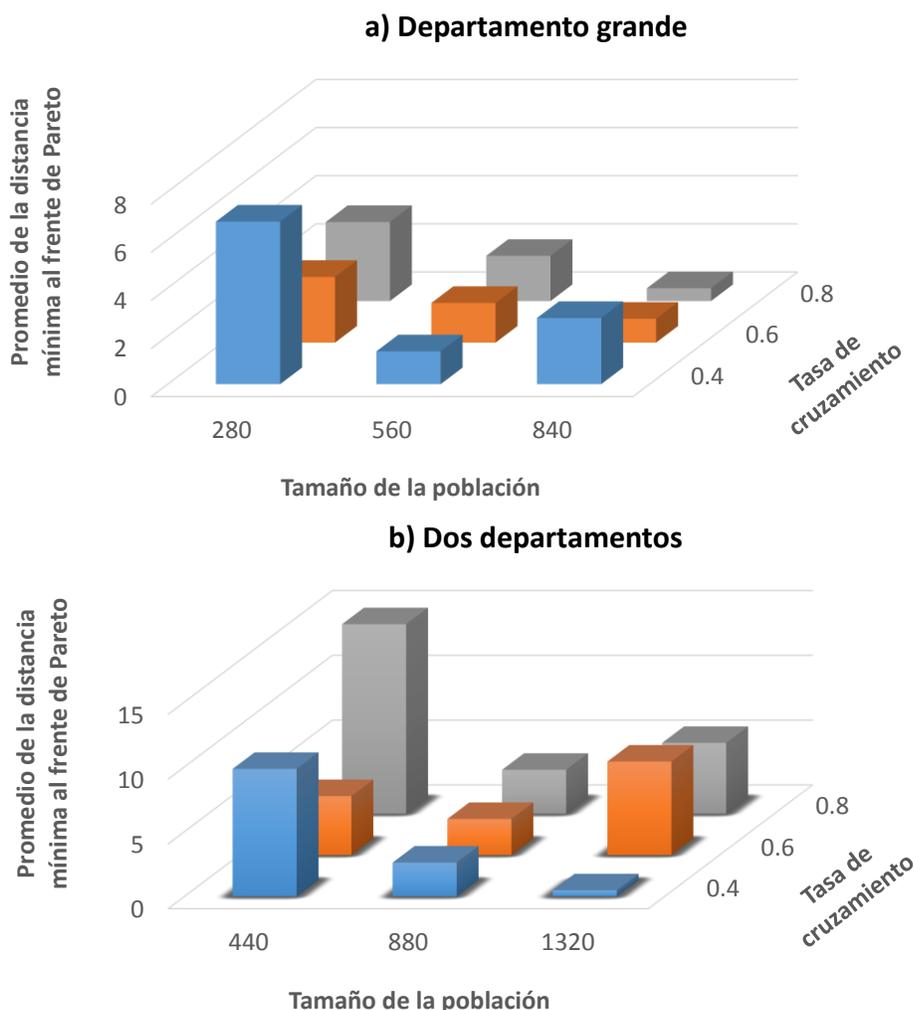


Fig. 6. Resultados de distintas configuraciones del algoritmo NSGA-II atendiendo al tamaño de la población y la tasa de cruzamiento en escenarios más complejos

complejos, hemos realizado experimentos computacionales adicionales. Los resultados, en términos del promedio de la distancia mínima al frente de Pareto, se muestran en la Figura 5. Nótese que en ambos escenarios se obtienen resultados diferentes.

Por ejemplo, en el caso del escenario *Departamento grande* (Figura 5-a) la mejor configuración es NSGA-II (840-0,8), mientras que en el escenario *Dos departamentos* es NSGA-II (1320-0,4). Lo común para ambas configuraciones es el empleo del mayor tamaño

de la población. Sin embargo, en cuanto a la tasa de cruzamiento, no queda claro que valor brinda los mejores resultados en los dos escenarios. Esto último sugiere que este parámetro pudiera ser objeto de variación durante la ejecución del algoritmo (ej. mediante una técnica de control de parámetros [19]).

Adicionalmente, con la intención de apreciar el efecto de la complejidad de estos escenarios sobre la calidad de las soluciones obtenidas, en la Figura 6 se muestra una comparación entre los tres escenarios abordados. Nótese que hemos

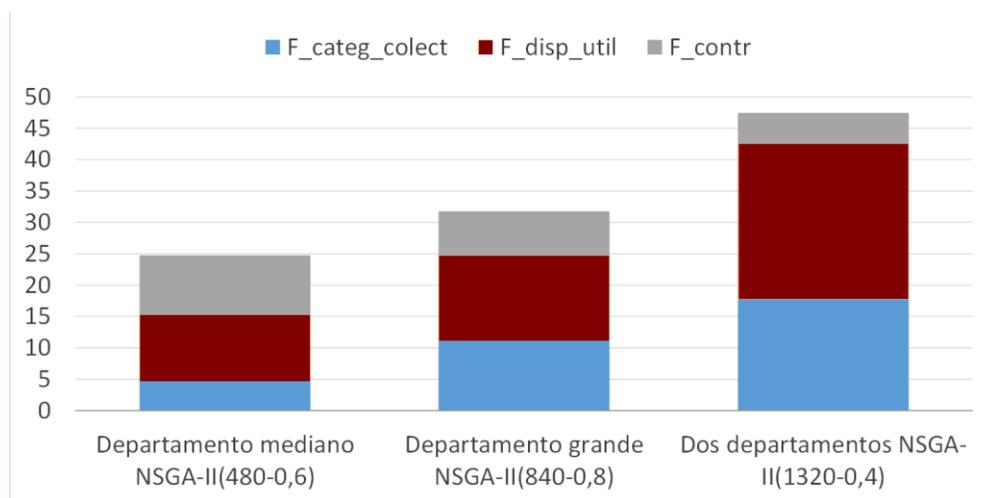


Fig. 5. Efecto de la complejidad del problema en la calidad de las mejores soluciones obtenidas

denominado al escenario de los apartados anteriores como *Departamento mediano*. Además, las gráficas corresponden al promedio de los frentes de Pareto obtenidos a partir de las soluciones encontradas por las mejores configuraciones de NSGA-II. A partir de estas gráficas es fácil ver que, de manera general, la calidad de las soluciones obtenidas decrece conforme la complejidad del problema aumenta.

5. Conclusiones

En la presente investigación se aborda el problema de la planificación de controles a clase en la educación superior cubana. En ese sentido, se propone una solución computacional que consistió en la modelación de este proceso como un problema de optimización multi-objetivo. En consecuencia, se aplica como método de solución el conocido algoritmo NSGA-II, el cual se basa en los conceptos de dominancia y optimalidad de Pareto. En particular, se realizaron algunas adaptaciones a los operadores evolutivos de dicho método, debido a las características del modelo propuesto.

Con la intención de validar tanto el modelo y método considerados, se analizó como caso de estudio, la planificación de los controles a clase en el Departamento de Licenciatura en Matemática de la Universidad de Holguín. En ese

sentido, se realizaron varios experimentos computacionales que tuvieron como objetivo determinar la mejor configuración del algoritmo NSGA-II en la solución del caso de estudio. Como principal conclusión se puede mencionar que tanto el modelo como el método aplicado, resultan adecuados para tratar el caso de estudio, obteniéndose soluciones útiles y de buena calidad. No obstante los resultados alcanzados hasta aquí, creemos que estos son solo el punto de partida para el desarrollo de un sistema informático inteligente, que facilite al decisor, la interacción con dicho proceso y la toma de decisiones. Nuestras futuras investigaciones estarán orientadas a tratar este tema.

Agradecimientos

El autor desea agradecer al colectivo de profesores del Departamento de Licenciatura en Matemática de la Universidad de Holguín, por su colaboración e intercambios en el desarrollo de la presente investigación.

Además, el autor cuenta con el apoyo de la beca postdoctoral otorgada por el proyecto Eureka SD (Erasmus Mundus Action 2), coordinado por la Universidad de Oldenburg (Alemania).

Referencias

1. **Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002).** A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, pp. 182–197. DOI: 10.1109/4235.996017
2. **Wegener, I. (2005).** *Complexity Theory. Exploring the Limits of Efficient Algorithms*. Springer-Verlag, Berlin Heidelberg, 308 p.
3. **Novoa-Hernández, P., Novoa-Hernández, M.A., & Rivero Peña, Y. (2013).** Propuesta de técnicas evolutivas para la confección automática de tribunales de trabajos de diploma. *Revista Cubana de Ciencias Informáticas*, Vol. 7, No. 4, pp. 90–99.
4. **Escalera Fariñas, K., Infante Abreu, A.L., André Ampuero, M., & Rosete Suárez, A. (2014).** Uso de estrategias de paralelización en algoritmos metaheurísticos para la conformación de equipos de software. *Revista Cubana de Ciencias Informáticas*, Vol. 8, No.3, pp. 90–99.
5. **Schaerf, A. (1999).** A Survey of Automated Timetabling. *Artificial Intelligence Review*, Vol. 13, No. 2, pp. 87–127. DOI: 10.1023/A:1006576209967
6. **Burke, E.K. & Petrovic, S. (2002).** Recent Research Directions in Automated Timetabling. *European Journal of Operational Research*, Vol. 140, No. 2, pp. 266–280. DOI: 10.1016/S0377-2217(02)00069-3
7. **Petrovic, S. & Burke, E.K. (2004).** University Timetabling. In: **Leung J. (ed.) Handbook of Scheduling: Algorithms, Models, and Performance Analysis**, Chp. 45, CRC Press.
8. **Coello Coello, C.A., Van Veldhuizen, D.A., & Lamont, G. (2005).** *Evolutionary Algorithms for Solving Multi-Objective Problems*. 2nd Edition. Springer Science+Business Media, LLC, 800 p.
9. **Zhou, A., Qu, B.Y., Li, H., Zhao, S.Z., Suganthan, P.N. & Zhang, Q. (2011).** Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, Vol. 1, pp. 32–49. DOI: 10.1016/j.swevo.2011.03.001
10. **Melián, B., Moreno Pérez, J., & Moreno Vega, J. (2003).** Metaheurísticas: Una visión global. *Revista Iberoamericana de Inteligencia Artificial*, Vol. 19, pp. 7–28.
11. **Talbi, E.G. (2009).** *Metaheuristics: from design to implementation*. John Wiley and Sons, 593 p.
12. **Ehrgott, M. (2005).** *Multicriteria Optimization* (2nd edition). Springer Berlin – Heidelberg.
13. **Sarker, R. & Ray, T. (2009).** An improved evolutionary algorithm for solving multiobjective crop planning models. *Computers and Electronics in Agriculture*, Vol. 68, pp. 191–199. DOI: 10.1016/j.compag.2009.06.002
14. **Saadatseresht, M., Mansourian, A. & Taleai, M. (2009).** Evacuation planning using multiobjective evolutionary optimization approach. *European Journal of Operational Research*, Vol. 198, No.1, pp. 305–314. DOI: 10.1016/j.ejor.2008.07.032
15. **Ting, C.K., Lee, C.N., Chang, H.C. & Wu, J.S. (2009).** Wireless heterogeneous transmitter placement using multiobjective variable-length genetic algorithm. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, Vol. 39, No. 4, pp. 945–958. DOI: 10.1109/TSMCB.2008.2010951
16. **Shin, S.Y., Lee, I.H., Kim, D. & Zhang, B.T. (2005).** Multiobjective evolutionary optimization of DNA sequences for reliable DNA computing. *IEEE Transactions on Evolutionary Computation*, Vol. 9, No. 2, pp. 143–158.
17. **Saravanan, R., Ramabalan, S., Ebenezer, N.G.R., & Dharmaraja, C. (2009).** Evolutionary multi criteria design optimization of robot grippers. *Applied Soft Computing*, Vol. 9, No. 1, pp. 159–172. DOI: 10.1016/j.asoc.2008.04.001
18. **Wozniak, P. (2011).** Preferences in multi-objective evolutionary optimisation of electric motor speed control with hardware in the loop. *Applied Soft Computing*, Vol. 11, No. 1, pp. 49–55
19. **Novoa-Hernández, P., Cruz Corona, C. & Pelta, D.A. (2015).** Self-adaptation in dynamic environments – a survey and open issues. *International Journal of Bio-Inspired Computation* (in press).

Pavel Novoa-Hernández es Ingeniero Informático por la Universidad de Holguín (Cuba) desde el 2007. Es Máster en Ciencias de la Computación e Inteligencia Artificial por la Universidad Central de Las Villas (Cuba), desde el 2010. En el 2013 obtuvo el grado de Doctor en Informática por la Universidad de Granada (España). Actualmente es Profesor Auxiliar del Departamento de Licenciatura en Matemática de la Universidad de Holguín. Sus intereses investigativos, relacionados con el área de la Soft Computing, incluyen: problemas dinámicos de optimización, meta-heurísticas, problemas de optimización multi-objetivo.

*Artículo recibido el 26/08/2014; aceptado el 17/04/2015.
Autor de correspondencia es Pavel Novoa Hernández.*