

Programación de Horarios Escolares basados en Ritmos Cognitivos usando un Algoritmo Genético de Clasificación No-dominada, NSGA-II

Víctor F. Suárez, Álvaro Guerrero y Omar D. Castrillón*

Universidad Nacional de Colombia, Facultad de Ingeniería y Arquitectura, Bloque Q,
Campus La Nubia, Manizales-Colombia.

(e-mail: vfsuarezc@unal.edu.co, aguerreroa@unal.edu.co, odcastrillong@unal.edu.co)

Recibido Jul. 13, 2012; Aceptado Ago. 22, 2012; Versión final recibida Sep. 04, 2012

Resumen

Se plantea la solución del problema de programación óptima de horarios escolares de una escuela pública Colombiana. Esto se hace considerando no solo la asignación adecuada de salones y docentes sino también los ritmos cognitivos que presentan los estudiantes como el factor más importante en el proceso de optimización. Se propone como método de solución el Algoritmo Genético de Clasificación No-dominada, NSGA-II. Los resultados muestran una mayor eficiencia del algoritmo en comparación con otros aplicados al mismo problema y evaluados en la misma forma. A nivel experimental, la metodología evidencia que los grupos programados mediante la orientación descrita presentan una reducción en los niveles de mortalidad académica en comparación con una programación horaria que no consideran los ritmos cognitivos de los estudiantes

Palabras claves: programación de horarios escolares, NSGA-II, ritmos de cognitivos, jornada escolar.

Scheduling of School Hours based on Cognitive Rhythms using a Non-dominated Sorting Genetic Algorithm, NSGA-II

Abstract

This paper proposes the implementation of an optimum academic agenda to solve the problem of scheduling a school day in Colombia. This is done by taking into account not only the appropriate assignment of rooms and professors but considering the cognitive rhythms of the students as the most relevant factor in the optimization. The Non-dominated Sorting Genetic Algorithm, NSGA-II is proposed as a solution method. The results show a higher efficiency of the algorithm when compared to other algorithms applied to the same problem and evaluated in the same form. At the experimental level the methodology shows a reduction in academic mortality rates in the groups where the methodology was implemented compared to groups that did not considered the cognitive rhythms of the students.

Keywords: school scheduling, NSGA-II, cognitive rhythms, constraints, school day

INTRODUCCION

Según Wren (1996), el problema de programación de horarios consiste en la disposición de objetos dentro de un patrón en el tiempo o el espacio, de tal manera que se logre el cumplimiento de unos objetivos determinados y que las restricciones sobre la forma en que se deben disponer dichos objetos se satisfagan. Es así como se identifican tres elementos esenciales dentro de cualquier proceso de programación: objetos, patrones y restricciones. Los objetos se refieren al recurso a organizar o disponer, los cuales pueden ser personas, máquinas, vehículos, clases, trabajos en una fábrica, etc.; el patrón es el orden que deben seguir los eventos y se crea como parte del proceso de programación; y las restricciones se definen como las relaciones tangibles o intangibles entre objetos y representan las reglas que se deben seguir para la construcción de un patrón. En el contexto educativo, este problema se enmarca en la disposición de asignaturas, estudiantes, profesores y salones, dentro de periodos de tiempo establecidos (jornadas, semanas, trimestres, semestres), los cuales están sujetos a unas condiciones instauradas de acuerdo a las características de la organización (colegios, universidades) y al tipo de evento que se desea programar (jornada, exámenes, uso de instalaciones).

La programación de horarios académicos generalmente se dirige a tres contextos: horarios universitarios, horarios escolares y horarios para exámenes. El presente artículo se centra en el problema de programación de horarios a nivel escolar. La asignación de horarios escolares tiene un enfoque en las clases y no en los estudiantes, por lo que la jornada diaria (mañana o tarde) se programa de forma continua, dividida en periodos de tiempos preestablecidos, para el desarrollo de actividades académicas en un número dado de asignaturas. Estas asignaturas se asocian a un nivel o grado respectivo, razón por la cual son cursadas por un mismo grupo de estudiantes, los cuales las toman a modo de curso obligatorio y no por demanda, como sucede con los cursos universitarios. En este sentido, la planta física también cumple con la orientación de clases, ya que los salones se asignan de forma permanente a un mismo grupo y se propicia de manera sistemática la rotación de los docentes por cada una de las aulas, según la distribución del horario y su asignación. No obstante, se deben tener en cuenta la existencia de aulas multipropósito o de uso especial, las cuales también pueden ser incluidas dentro de la programación para ser dispuestas de manera permanente en un horario dado para alguno de los grupos. En cuanto a los docentes, la asignación debe garantizar que un docente no esté en dos grupos diferentes al mismo tiempo.

En una proporción alta de instituciones, la solución a este problema es construida de manera manual, lo cual además de tomar días o semanas de trabajo, no permite establecer cierto tipo de restricciones que optimicen el uso del recurso humano y de infraestructura, al igual que da lugar a errores por traslapes en asignaturas y en la asignación de los docentes (Flores, 2011). Sin embargo, a nivel investigativo se han aplicado diversas técnicas analíticas y heurísticas a la solución del problema de programación de horarios escolares tales como: grafos coloreados (Čangalović y Schreuder, 1991), recocido simulado (Abramson, 1991), algoritmos genéticos (Colomi et al., 1998; Abramson y Abela, 1992; Beligianni et al., 2009), programación con restricciones (Valouxis y Housos, 2003), algoritmos de búsqueda local (Avella et al., 2007), programación lineal (Birbas et al., 2009), búsqueda tabú (Minh et al., 2010) y enjambres de partículas (Tassopoulos y Beligiannis, 2012). Igualmente, es importante resaltar que en la búsqueda de una metodología más integral ha sido necesario analizar la aplicación de estas técnicas inteligentes en una gran variedad de problemas como: Optimización de ambientes Job shop (Ruiz et al., 2012), Optimización de redes eléctricas por algoritmos genéticos (Anaut et al., 2009), tratamiento de emisiones atmosféricas (Pérez et al., 2004), procesos de simulación en degradación aeróbica (Rosa et al., 2004), etc.

Aunque los métodos mencionados contemplan las limitaciones generales del problema, se encuentran diferentes enfoques en los que se establecen restricciones especiales sobre la infraestructura, la inclusión de asignaturas electivas, el balanceo de la carga académica diaria de los docentes y otras condiciones particulares de acuerdo al tipo de modelo académico acogido experimentalmente. No obstante, a pesar de las consideraciones expuestas, ninguno de los métodos aborda de manera explícita la relación existente entre los horarios programados y el proceso pedagógico, dejando de lado el papel vital que este asunto tiene dentro de la gestión del tiempo dentro del aula, la cual está directamente relacionada con el grado de aprendizaje de un estudiante (Murillo, 2008; Cornejo y Redondo, 2007). En este sentido, es necesario aclarar que el proceso cognitivo a nivel escolar depende de cuatro factores relacionados con las necesidades del alumno: las preferencias emocionales, el ambiente, las necesidades sociales y las necesidades fisiológicas (Maciques, 2012). Este artículo se centra en el último factor, el cual además de vincular las necesidades alimenticias y las necesidades de movimiento del individuo, resalta la existencia de una hora óptima del día para desarrollar un trabajo. Por tanto, conocer dicha hora favorece la programación de horarios particulares para la consecución de rendimientos óptimos en la asimilación de contenidos (De La Cruz Nogales, 2009).

Surgen entonces dos preguntas: ¿Cuándo enseñar? y ¿qué enseñar? Algunas creencias a nivel escolar indican que en las primeras horas de la mañana son las más adecuadas para ubicar las asignaturas más complejas, pues suponen que los estudiantes tienen una mayor disposición y atención. No obstante, como señala Fernández (2000), las investigaciones sobre atención y fatiga sugieren que los estudiantes tienen en la mañana un nivel bajo de atención, el cual sube y se mantiene hasta antes del mediodía, momento en el que vuelve a descender. En la tarde el fenómeno se repite, sólo que el proceso de ascenso se ve reflejado sobre las cuatro de la tarde, lo cual implica que en esta parte de la jornada, el nivel más alto de atención se logra de manera más tardía. Es preciso señalar, que aunque existe un debate alrededor de la implementación de una jornada escolar continua o partida (Feito, 2007) y sobre la duración de la jornada (Banco Mundial, 2009), el presente trabajo se orienta en la programación de horarios de forma continua, ya sea en la jornada de la mañana o en la tarde, bajo los parámetros establecidos por el Ministerio de educación de Colombia para las instituciones de carácter público (MEN, 2012).

En concordancia con lo mencionado, Hederich et al. (2004), presentaron el trabajo titulado “Ritmos cognitivos en la escuela”, una investigación que expone las características cognitivas de los estudiantes y su relación directa con la planificación de la jornada escolar, dentro del modelo educativo Colombiano. De sus resultados se destaca: a) La primera hora de la jornada (mañana – tarde) presenta los niveles más bajos de activación fisiológica y de atención. b) Los momentos ideales para el trabajo, intelectualmente exigente, pertenecen al intermedio de la jornada, en los cuales los niveles de activación fisiológica son elevados con un nivel de atención ideal. c) Las últimas horas de la jornada tampoco favorecen el trabajo en asignaturas con un alto componente de atención o que requieran de una actividad intelectual sostenida, por lo que se recomienda ubicar asignaturas de un alto componente motor. d) Dada la dificultad existente con la cantidad de horas de clase semanales y la necesidad de jornadas continuas como parte de la optimización de la infraestructura, se propone que las asignaturas con mayor demanda cognitiva como matemática y lenguajes se ubiquen preferiblemente al intermedio de la jornada, o en su defecto al inicio de la misma, siempre y cuando se privilegien los momentos de presentación de datos durante las primeras horas y los momentos de interpretación y análisis en las segundas horas.

Con base en lo anterior, se propone un modelo de programación de horarios acogiendo no sólo las restricciones generales sobre docentes, asignaturas y salones, sino incluyendo también un acercamiento a la práctica pedagógica, al considerar los momentos de la jornada en los que es más propicio el aprendizaje para ciertas áreas del conocimiento. En este caso, se estipula una programación horaria semanal de 30 horas, repartida de lunes a viernes en 6 clases diarias de una hora. La jornada diaria se divide en tres segmentos de dos horas de acuerdo al nivel de atención previsto, determinando un periodo de inicio bajo, un periodo intermedio alto y un tercer periodo de descenso. A cada periodo se asocian un conjunto de asignaturas de acuerdo a las características de sus contenidos, por ejemplo al primer bloque, áreas de ciencias naturales y ciencias sociales; al segundo bloque, áreas de ciencias exactas, lenguaje e idiomas; y al tercer bloque, áreas de deportes, tecnología y manualidades. El objetivo de la programación es ubicar la mayor cantidad posible de asignaturas en el segmento horario que les corresponde, ya que esto permite el desarrollo de actividades propias al nivel cognitivo que el estudiante tiene en los diferentes momentos de la jornada.

La estructura de la metodología presentada en este documento emplea como estrategia de solución el algoritmo NSGA-II, un algoritmo genético multiobjetivo con elitismo (Carranza et al., 2008; Ruizet al., 2012), el cual ha sido adaptado para la solución de este problema. La ventaja de su uso está en que permite abordar el problema general a partir de la construcción de soluciones individuales para cada grupo, dando cumplimiento a las restricciones de carácter obligatorio y seleccionando aquellas que incumplen en menor proporción las restricciones blandas, de tal manera que se dé prioridad a las soluciones con mejor disposición de acuerdo al enfoque en los ritmos de aprendizaje. Para mostrar la alta eficiencia del algoritmo bajo el enfoque establecido, se realiza la comparación de los resultados de su función de evaluación, con los entregados por un algoritmo genético simple y un algoritmo de búsqueda aleatoria, en la solución de un mismo problema.

METODOLOGÍA

La metodología planteada para el desarrollo de esta investigación se describe de manera general en la figura 1. El proceso requiere del ingreso de la información básica sobre la cual se realiza la construcción del horario. Estos datos son: nombre de las materias y un código único de identificación sin importar el grupo, nombre del profesor que orienta cada materia y su respectivo código único de identificación, intensidad horaria semanal de cada materia, nombre del grupo en el que se orientará dicha materia, código de área a la que pertenece la materia (este código es un identificador del área del conocimiento en el que se ubica una asignatura, por ejemplo, materias como matemáticas, cálculo, geometría pertenecen al área de ciencias exactas), y por último, un número de identificación para el salón donde se orientará la asignatura. Se debe

verificar que la intensidad horaria semanal para cada grupo sea de 30 horas de clase y que el tiempo total de la asignación académica semanal de cada docente sea de 22 horas efectivas.

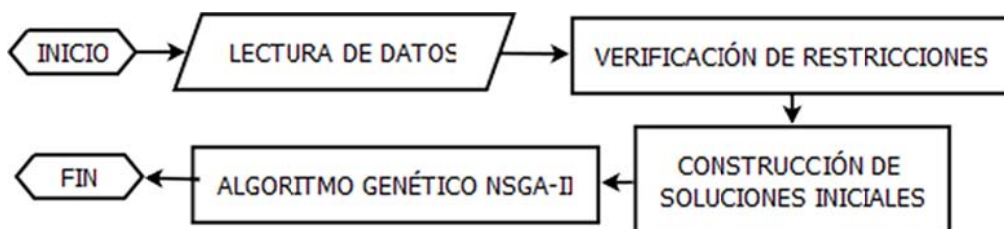


Fig. 1: Esquema general del programa

Una vez validados los datos se procede a la construcción de soluciones iniciales, las cuales corresponden a cuatro hipermatrices (matriz de tres dimensiones) denominadas: Materias, Profesores, Salones y Códigos de Área. En cada una de ellas se plasma de manera aleatoria la información correspondiente a cada grupo. La figura 2 muestra la disposición de dichas hipermatrices.

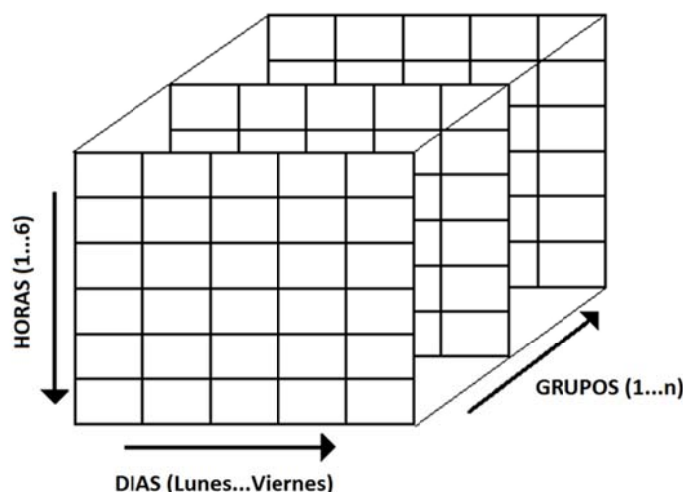


Fig. 2: Esquema de hipermatriz

Las soluciones iniciales constituyen la entrada a la sección que ejecuta el algoritmo NSGA-II, descrito en la figura 3. El algoritmo busca una solución óptima para cada grupo y a partir de estas construye la solución final del problema planteado.

Paso 1: Población Inicial.

Se toman todas las materias de un grupo y se conforman bloques de 2 horas de la misma materia. En caso de que la intensidad horaria de una materia sea impar, además de los bloques correspondientes quedara una hora suelta. Posteriormente se establece el parámetro L, éste representa la cantidad de bloques, en el cual las horas sueltas se incluyen como un bloque más. Además, es necesario determinar el tamaño de la población inicial, la cual es igual $2 \cdot L$. Esto último permitirá generar soluciones aleatorias a partir de una solución inicial, hasta alcanzar el tamaño de la población inicial.

Paso 2: Operadores Genéticos.

Los operadores genéticos de cruce y mutación son aplicados en cada individuo de la población inicial, con el fin de encontrar los nuevos individuos (Hijos). Para el cruce se aplica la técnica llamada Cruce PMX (Partially Matched Crossover) propuesta por Goldberg y Lingle (1985), el cual se realiza sobre las posiciones de los elementos en los individuos y no sobre sus valores. Para ello, se determinan dos posiciones dentro del vector que contiene el individuo; los elementos que se encuentran por fuera del segmento determinado por las posiciones se dejan igual en los hijos, luego los elementos que faltan en el

hijo 1 se colocan del padre 2 en el orden secuencial de aparición dentro del individuo, y de igual forma los elementos que faltan en el hijo 2 se toman del padre 1 en el orden secuencial en el que aparecen, tal como se puede observar en la tabla 1. En este ejemplo se tienen dos padres y se determinaron las posiciones 8 y 15 dentro del vector.

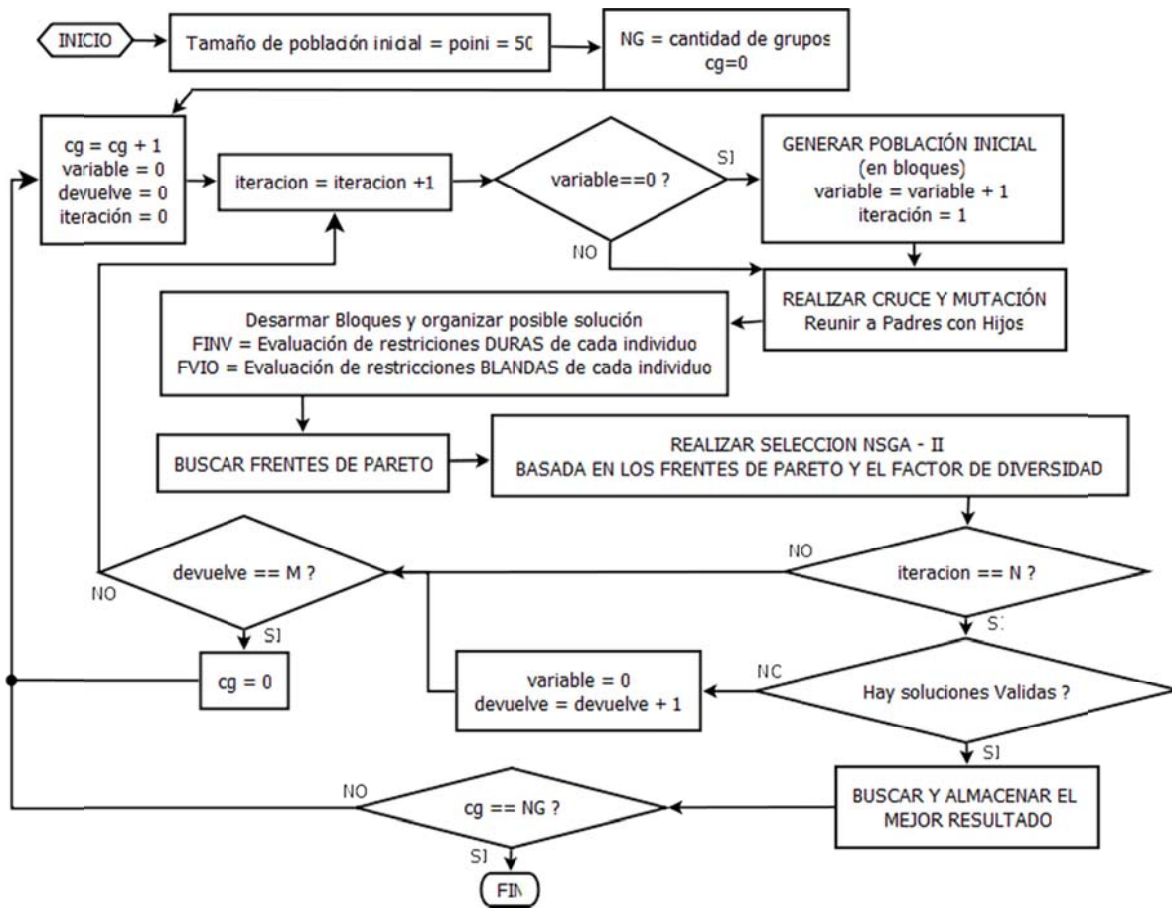


Fig 3: Esquema de ejecución del algoritmo genético

Tabla 1: Ejemplo Cruce PMX

PADRE 1	10	9	2	13	11	18	7	12	14	15	4	6	1	8	16	5	17	3
PADRE 2	8	12	14	9	15	3	13	11	1	2	4	6	7	16	10	18	5	17
HIJO 1	10	9	2	13	11	18	7									5	17	3
HIJO 2	8	12	14	9	15	3	13									18	5	17
HIJO 1	10	9	2	13	11	18	7	8	12	14	15	1	4	6	16	5	17	3
HIJO 2	8	12	14	9	15	3	13	10	2	11	7	4	6	1	16	18	5	17

La mutación se aplica a cada individuo de la población inicial determinando dos posiciones aleatorias en sus elementos. Aquellos que se encuentran por fuera de las posiciones se pasan al hijo como están en el padre y el sub-vector contenido dentro de las dos posiciones se invierte y se pasa al nuevo hijo.

Paso 3: Selección.

Se reúne la población de padres y la población de hijos, siendo necesario desarmar los bloques con el fin de formar cada una de las posibles soluciones. Estas soluciones son evaluadas mediante dos funciones objetivo, las cuales que describen las restricciones del problema y se deben cumplir en su totalidad para que la solución generada sea válida. Funciones duras (RD) y blandas (FITNESS), las cuales están definidas de

acuerdo a las condiciones específicas del planteamiento realizado en esta investigación y con base en algunos trabajos previos (Suarez y Castrillón, 2011). La función RD contempla las siguientes restricciones: a) En un mismo día no se deben dictar más de 2 horas de la misma materia. b) Sí en un día se dictan 2 horas de la misma materia estas deben ser consecutivas. c) Un profesor no puede dictar 2 materias a la misma hora del día. d) A una misma hora no pueden estar programadas dos clases en un mismo salón.

Las siguientes ecuaciones (1-5) representan las funciones de evaluación de estas restricciones, sobre las cuales operan las hipermatrices de Materias, Profesores y Salones:

$$RD = F1 + F2 \tag{1}$$

Los valores de F1 y F2 se describen de la siguiente manera.

$$F1 = \sum_{i=1}^{NG} \sum_{j=1}^5 F_{Rest1}(Materias) \tag{2}$$

$$F2 = \sum_{j=1}^5 \sum_{k=1}^6 F_{Rest2}(Profesores) + F_{Rest2}(Salones) \tag{3}$$

Donde:

- i = variable para recorrer la cantidad total de grupos a programar.
- j = variable para recorrer los días de la semana.
- k = variable para recorrer las horas del día.
- NG= cantidad total de grupos a programar

La evaluación F_{Rest1} y F_{Rest2} se realizan de la siguiente manera:

$$F_{Rest1} = \begin{cases} 1 & \text{Si algún código se repite más de 2 veces en la columna } j \\ 1 & \text{Si al repetirse algún código 2 veces en la columna } j, \text{ estos no están consecutivos} \\ 0 & \text{En otro caso} \end{cases} \tag{4}$$

$$F_{Rest2} = \begin{cases} 1 & \text{Si se repite algún elemento en el vector de profundidad determinado por } (j, k) \\ 0 & \text{En otro caso} \end{cases} \tag{5}$$

Es importante tener en cuenta que la solución no es válida en caso de que la función objetivo RD sea diferente de cero.

La función objetivo *FITNESS* hace referencia al enfoque en los ritmos de aprendizaje y su evaluación determina la calidad de las soluciones. El algoritmo busca que este valor se minimice. La evaluación de esta función se realiza con base a la segmentación de la jornada, con tres tipos de restricciones según la ubicación de cada asignatura durante el día:

- Restricción 1: Asignaturas donde los estudiantes tienen una mejor disposición para recibir su contenido en la primera y segunda hora de la jornada.
 - Restricción 2: Asignaturas donde los estudiantes tienen una mejor disposición para recibir su contenido en la tercera y cuarta hora de la jornada.
 - Restricción 3: Asignaturas donde los estudiantes tienen una mejor disposición para recibir su contenido en la quinta y sexta hora de la jornada.
- Estas restricciones se evalúan en la hipermatriz de Códigos de Área, a través de la función de evaluación presentada en la ecuación 6:

$$FITNESS = \min \left(\sum_{i=1}^{NG} \sum_{j=1}^5 \sum_{k=1}^6 F_{Rest3}(Códigos\ de\ Área) \right) \tag{6}$$

Donde F_{Rest3} evalúa una restricción por iteración, dependiendo del código de área. En la tabla 2, se pueden observar los valores de penalización que se tienen para cada restricción de acuerdo a la ubicación de cada asignatura según la hora del día.

Tabla 2: Valores de penalización restricciones blandas

Hora de la jornada	Valor Penalización Restricción 1	Valor Penalización Restricción 2	Valor Penalización Restricción 3
1	0	6	3
2	0	3	3
3	3	0	6
4	3	0	6
5	6	3	0
6	6	6	0

Una vez evaluadas todas las restricciones descritas, se procede a realizar la selección de los individuos para la nueva población. Dicha selección está basada en dos criterios: el rango de dominancia según los frentes óptimos de Pareto y el factor de diversidad de cada frente. Estos se calculan utilizando los resultados de las funciones objetivo, con el fin de encontrar grupos de soluciones en donde el primero es dominante (mejor calidad) sobre los demás. La obtención de los frentes de Pareto consiste en organizar los individuos de menor a mayor valor con respecto a la función objetivo *RD*, aplicar una división recursiva de la nueva población ordenada hasta que se puedan comparar sub-poblaciones de tamaño 1, en donde la sub-población derecha es dominada por la sub-población izquierda respecto a la función objetivo *RD*. El frente se conforma por la sub-población izquierda más los individuos de la sub-población derecha que no son dominados por ningún individuo de la sub-población izquierda con respecto a la función objetivo *FITNESS* (Peñuela y Granada, 2007).

Luego de encontrar un frente, sus individuos se descartan de la población ordenada y se repite el proceso hasta ubicar todos los individuos en frentes. La nueva población de individuos se construye agregando los individuos de cada frente hasta completar el tamaño de la población inicial. Sin embargo, puede ocurrir que al agregar un frente a la solución, la cantidad de individuos supere el tamaño de la población inicial. En este caso, si antes del último frente adicionado faltan 1 o 2 individuos, se adicionan el primer individuo o el primero y el último individuos del frente siguiente respectivamente. En caso de que falten más de 2 individuos, se agrega el primero y el último individuo más los escogidos mediante el factor de diversidad, hasta que el tamaño sea igual a la población inicial, teniendo prioridad aquellos individuos cuyo factor de diversidad sea más alto. Este factor permite medir la distancia entre los individuos del mismo frente y logra una mejor exploración del espacio de soluciones.

El proceso descrito desde el cruce y la mutación se repite hasta completar *N* iteraciones, después se verifica si en la última población seleccionada hay soluciones válidas, es decir, si el vector de la función objetivo *RD* tiene al menos un cero. En caso de que no se encuentre una solución válida se reinicia el algoritmo generando una nueva población inicial para el grupo actual. Si se debe generar más de *M* veces la población inicial de un grupo, el algoritmo vuelve a iniciar el proceso desde el grupo 1. Sino, se escoge la solución con el mejor valor de evaluación de la función *FITNESS* y se procede con el siguiente grupo.

Paso 4: Condición de parada.

El proceso debe continuar hasta completar la totalidad de grupos: En este momento el algoritmo entrega la solución final escogida.

EXPERIMENTACION

Paso 1: Pruebas.

Con el propósito de mostrar la eficiencia de la metodología propuesta, se establece la comparación de los resultados de su función de evaluación, con los obtenidos de la ejecución de un algoritmo genético simple y un algoritmo de búsqueda aleatoria, cuyo diseño y estructura también atienden a las condiciones específicas del planteamiento realizado en esta investigación. El algoritmo genético simple se diferencia del propuesto en la complejidad de sus operadores genéticos (no se manejan frentes) y el algoritmo de búsqueda aleatoria realiza la programación mediante una asignación recursiva en la que sólo examina el cumplimiento de las restricciones duras.

Se plantean 6 problemas con diferente número de profesores, asignaturas, salones y total de horas programadas (Tabla 3), los cuales son resueltos por cada algoritmo durante 10 repeticiones.

Tabla 3: Parámetros de los problemas de prueba

Problema	Cantidad de grupos	Cantidad de docentes	Cantidad de asignaturas	Cantidad de salones	Total de horas programadas
1	4	14	40	5	120
2	5	19	50	6	150
3	6	20	60	7	180
4	7	23	73	8	210
5	8	23	86	9	240
6	9	23	99	10	270

En cada una de las repeticiones, el resultado obtenido respecto a la función *FITNESS* es tomado como referencia. Se obtienen los siguientes valores (Tabla 4):

Tabla 4: Resultados función *FITNESS* para comparación de eficiencia de algoritmos

Problema	Método	Valor Fitness total por repetición									
		1	2	3	4	5	6	7	8	9	10
1	Aleatorio	339	336	354	357	294	324	372	285	333	321
2	Aleatorio	477	429	444	492	444	432	426	408	387	423
3	Aleatorio	543	483	486	555	492	462	525	483	513	513
4	Aleatorio	630	573	600	507	567	573	501	543	645	627
5	Aleatorio	699	702	651	696	591	690	714	672	693	675
6	Aleatorio	867	702	771	771	768	762	753	765	789	744
1	AG –Simple	228	273	279	261	252	240	261	276	267	258
2	AG –Simple	327	303	321	279	309	315	324	294	288	303
3	AG –Simple	387	411	330	387	447	423	375	390	384	411
4	AG –Simple	501	474	426	459	408	471	507	444	456	444
5	AG –Simple	537	534	528	552	585	546	573	549	516	564
6	AG –Simple	606	627	642	648	600	627	645	609	648	600
1	NSGA-II	201	177	150	165	165	210	168	189	183	204
2	NSGA-II	183	186	198	168	201	255	222	219	189	189
3	NSGA-II	318	300	258	267	276	285	264	285	276	264
4	NSGA-II	279	321	324	303	363	342	399	357	315	300
5	NSGA-II	312	327	423	426	390	384	396	423	411	348
6	NSGA-II	483	450	456	534	471	432	438	480	534	468

Se puede observar que en cada uno de los problemas la mejor solución hallada fue encontrada con el algoritmo NSGA-II, mientras que la peor solución está asociada al método de búsqueda aleatoria. Estas soluciones fueron comparadas mediante un análisis Anova (Tabla 5), para determinar si existen diferencias estadísticamente significativas entre los resultados entregados por cada algoritmo. Posteriormente, se aplicaron pruebas post hoc para establecer entre qué métodos existen dichas diferencias (Tabla 6).

Tabla 5: Análisis de varianza entre métodos de solución

Problema		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Inter-grupos	113016,600	2	56508,300	122,720	,000
	Intra-grupos	12432,600	27	460,467		
	Total	125449,200	29			
2	Inter-grupos	277603,800	2	138801,900	228,882	,000
	Intra-grupos	16373,700	27	606,433		
	Total	293977,500	29			

Tabla 5: (continuación)

3	Inter-grupos	255861,600	2	127930,800	174,459	,000
3	Intra-grupos	19799,100	27	733,300	174,459	,000
	Total	275660,700	29			
4	Inter-grupos	303523,800	2	151761,900	96,859	,000
4	Intra-grupos	42304,500	27	1566,833	96,859	,000
	Total	345828,300	29			
5	Inter-grupos	435046,200	2	217523,100	190,670	,000
5	Intra-grupos	30802,500	27	1140,833	190,670	,000
	Total	465848,700	29			
6	Inter-grupos	434018,400	2	217009,200	192,817	,000
6	Intra-grupos	30387,600	27	1125,467	192,817	,000
	Total	464406,000	29			

Tabla 6: Prueba Post hoc. Test de Tukey

Problema	(I) Tipo de Algoritmo	(J) Tipo de Algoritmo	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
1	Aleatorio	AG-Simple	72,000*	9,597	,000	48,21	95,79
		NSGA-II	150,300*	9,597	,000	126,51	174,09
	AG-Simple	NSGA-II	78,300*	9,597	,000	54,51	102,09
2	Aleatorio	AG-Simple	129,900*	11,013	,000	102,59	157,21
		NSGA-II	235,200*	11,013	,000	207,89	262,51
	AG-Simple	NSGA-II	105,300*	11,013	,000	77,99	132,61
3	Aleatorio	AG-Simple	111,000*	12,110	,000	80,97	141,03
		NSGA-II	226,200*	12,110	,000	196,17	256,23
	AG-Simple	NSGA-II	115,200*	12,110	,000	85,17	145,23
4	Aleatorio	AG-Simple	117,600*	17,702	,000	73,71	161,49
		NSGA-II	246,300*	17,702	,000	202,41	290,19
	AG-Simple	NSGA-II	128,700*	17,702	,000	84,81	172,59
5	Aleatorio	AG-Simple	129,900*	15,105	,000	92,45	167,35
		NSGA-II	294,300*	15,105	,000	256,85	331,75
	AG-Simple	NSGA-II	164,400*	15,105	,000	126,95	201,85
6	Aleatorio	AG-Simple	144,000*	15,003	,000	106,80	181,20
		NSGA-II	294,600*	15,003	,000	257,40	331,80
	AG-Simple	NSGA-II	150,600*	15,003	,000	113,40	187,80

El análisis de varianza muestra que la significancia de la prueba es inferior al 5% en los 6 problemas, por tanto, existen diferencias estadísticas entre los resultados obtenidos por los diferentes algoritmos. El test de Tukey determinó que dicha diferencia se da entre todos los métodos (sig=0,000). En este sentido, el valor medio de la función de evaluación del método NSGA-II en todos los problemas, presenta una eficiencia 29,31% y 44,74% mayor con respecto al algoritmo genético simple y al algoritmo de búsqueda aleatoria respectivamente. De esta manera, se establece que el algoritmo NSGA-II presenta los mejores resultados.

Paso 2: Prueba piloto.

En esta sección se realizó una prueba piloto para verificar el efecto de la programación de horario dentro del proceso académico de la institución educativa Colegio San Jorge, en la ciudad de Manizales, institución de

carácter público, que atiende aproximadamente 640 estudiantes repartidos en 17 grupos de sexto a once. Del grado sexto a décimo existen tres grupos por nivel, mientras que en undécimo sólo son dos grupos. Para lograr una comparación objetiva, se tomaron dos grupos por nivel, realizando la programación de estos con los algoritmos NSGA-II y Búsqueda Aleatoria. El propósito es verificar la mortalidad académica de todas las asignaturas (Cantidad de asignaturas perdidas) en cada uno de los grupos y determinar si existe alguna relación con el método de programación horaria utilizada. El problema aplicado a los dos algoritmos tiene los siguientes parámetros: 6 grupos, 23 profesores, 8 salones y 66 asignaturas. Se da la condición de uso de dos salones adicionales a la cantidad de grupos, los cuales corresponden a la sala de cómputo y al patio de deportes, los cuales se deben asignar a cada grupo sin que haya dos o más grupos asignados a estos salones al mismo tiempo.

RESULTADOS Y DISCUSIÓN

Con base en la metodología propuesta se solucionó el problema planteado en el paso 2 de la experimentación. La programación de horarios generada bajo los dos métodos, fue implementada en el colegio mencionado. Una vez aplicada la jornada propuesta bajo cada método (NSGA-II y Búsqueda Aleatoria), se recopiló la información de un periodo lectivo respecto al rendimiento académico en los grupos evaluados la cual se encuentra resumida en la tabla 7. Allí se especifica la cantidad de alumnos por grupo y la cantidad de asignaturas orientadas en cada nivel. La cantidad de asignaturas perdidas corresponde a la sumatoria de las asignaturas perdidas por todos los estudiantes de un grupo durante el periodo evaluado. El valor del porcentaje de pérdida intragrupo representa la proporción de estas asignaturas respecto a la cantidad total de asignaturas orientadas en un grupo.

Tabla 7. Resultados experimentación

Nivel y grupo	Método y valor Fitness	Cantidad alumnos	Cantidad asignaturas orientadas	Total asignaturas orientadas	Cantidad asignaturas perdidas	% Perdida intragrupo	% Diferencia entre grupos
6	1 aleatorio 108	49	10	490	89	18,2%	1,9%
	2 NSGA-II 24	48	10	480	78	16,3%	
7	1 aleatorio 96	42	10	420	85	20,2%	8,5%
	2 NSGA-II 54	41	10	410	48	11,7%	
8	1 aleatorio 84	41	10	410	68	16,6%	4,4%
	2 NSGA-II 51	42	10	420	51	12,1%	
9	1 aleatorio 84	41	10	410	106	25,9%	6,1%
	2 NSGA-II 60	42	10	420	83	19,8%	
10	1 aleatorio 99	42	13	546	129	23,6%	0,7%
	2 NSGA-II 48	39	13	507	116	22,9%	
11	1 aleatorio 72	38	13	494	77	15,6%	2,3%
	2 NSGA-II 54	37	13	481	64	13,3%	
Promedio						18,0%	4,0%

Los resultados muestran que los valores de la función *FITNESS* son menores en todos los grupos programados mediante el algoritmo NSGA-II. Adicional a esto, el valor promedio de asignaturas perdidas por grupo es del 18% y la diferencia entre grupos muestra un promedio de 4 puntos porcentuales por encima en los grupos programados con el algoritmo de búsqueda aleatoria. Esto indica que los grupos programados con la metodología propuesta basada, en el método NSGA-II, presentan una reducción del 22,22 % en la cantidad de materias perdidas por grupo.

Finalmente, se resalta que en la revisión literaria no se encuentra una metodología para la programación de horarios escolares que tenga como objetivo reducir la mortalidad académica o que contemple un conjunto de restricciones asociadas a los ritmos cognitivos de los estudiantes o alguna otra variable relacionada con los procesos de aprendizaje. Por lo general, las investigaciones existentes centran el problema en la infraestructura y la carga académica docente, desvinculando el proceso de programación horaria del proceso pedagógico. Por tanto, el modelo propuesto además de permitir la búsqueda de resultados con una alta eficiencia a través del algoritmo NSGA-II, aborda restricciones para obtener un mayor beneficio de las horas de del día en las que los estudiantes son más productivos.

CONCLUSIONES

Los resultados permiten obtener las siguientes conclusiones: 1) El enfoque en ritmos cognitivos busca que el problema de programación de horarios escolares se vincule a los procesos pedagógicos relacionados con las necesidades fisiológicas de los estudiantes, en los que determinados momentos de la jornada subrayan un mayor o menor nivel de atención para el desarrollo de cierto tipo de actividades. En este sentido, el uso del algoritmo NSGA-II para la resolución de un problema de esta complejidad, permite obtener un conjunto de soluciones que acogen las múltiples restricciones impuestas y que en comparación con otros algoritmos utilizados (Algoritmo genético simple, Algoritmo de búsqueda aleatoria), alcanza un grado de optimización de hasta un 44,74% superior, en la ubicación de las asignaturas en la jornada, en función de la capacidad intelectual que requieren. 2) La aplicación de la metodología a la programación horaria de una institución educativa, muestra que los grupos, en los que su distribución horaria tiene un enfoque basado en los ritmos de aprendizaje, presentaron un descenso en la mortalidad académica hasta de un 22,22%, en relación con a los demás grupos. 3) El algoritmo permite tener en cuenta características como el uso de salones adicionales a los correspondientes para cada grupo, conservando las limitaciones de traslape; además de habilitar la asignación de múltiples materias para un docente dentro de un mismo grupo. 4) La metodología propuesta está diseñada de acuerdo a las características de la escuela pública en Colombia, lo cual permite que pueda ser aplicada en diferentes instituciones educativas cuya jornada sea continua (mañana, tarde), empleando diversas técnicas inteligentes, con el fin de obtener un mejor análisis sobre su verdadero impacto en la disminución de la mortalidad académica.

AGRADECIMIENTOS

Este artículo es parte del proyecto denominado "*Disminución de la mortalidad académica en instituciones de educación básica y media mediante el empleo de Técnicas Inteligentes en la asignación de horarios*" (código Hermes 13744), Financiado por la Universidad Nacional de Colombia.

REFERENCIAS

- Abramson, D, Constructing school timetables using simulated annealing: sequential and parallel algorithms. *Management Science*: 37(1) 98-113 (1991).
- Abramson, D., y J. Abela, A Parallel Genetic Algorithm for Solving the School Timetabling Problem. 15 Australian Computer Science Conference, 1-9 hobart, australia feb (1992).
- Anaut, D. y otros tres autores, Optimización de redes eléctricas mediante la aplicación de algoritmos genéticos, *Información Tecnológica*, 20 (4), 137-148 (2009).
- Avella, P. y otros tres autores, A computational study of local search algorithms for Italian high-school timetabling. *Journal of Heuristics*, 13(6) 543–556 (2007).
- Banco Mundial, La calidad de la educación en Colombia: Un análisis y algunas opciones para un programa de política. Misión residente en Colombia, Bogota, Colombia (2009).
- Beligiannis, G., C. Moschopoulos. y S. Likothanassis, A genetic algorithm approach to school timetabling. *Journal of the Operational Research Society*, 60(1) 23-42 (2009.)
- Birbas, T., S. Daskalaki. y E. Housos, School timetabling for quality student and teacher schedules. *Journal of Scheduling*, 12(2) 177–197 (2009).
- Čangalović, M. y J. A. Schreuder, Exact colouring algorithm for weighted graphs applied to timetabling problems with lectures of different lengths. *European Journal of Operational Research*, 51(2) 248-258 (1991).
- Carranza, H. E. y otros tres autores, Un enfoque multiobjetivo para la asignación de canales en sistemas celulares. *Información Tecnológica*: 19(1) 87-96 (2008).
- Colomi, A., M. Dorigo. y V. Maniezzo. Metaheuristics for high school timetabling. *Computational Optimization and Applications*: 9 (3) 275–298 (1998).
- Cornejo R., y J.M. Redondo, Variables y factores asociados al aprendizaje escolar. Una discusión desde la investigación actual. *Estudios Pedagógicos*, 33(2) 155-175 (2007).
- De La Cruz F. M, Las estrategias de aprendizaje: Técnicas para el estudio. *Revista Digita Innovación y Experiencias Educativa*, 1(15) 1-10 (2009).

- Feito A. R, Tiempos escolares: El debate sobre la jornada escolar continua y partida. Cuadernos de pedagogía, 1(365) 74-79 (2007).
- Fernández E. M, La hora de la escuela. Análisis y valoración de los procesos, los efectos y las opciones de implantación de la jornada continua. Universidad de Salamanca, Salamanca, España (2000).
- Flórez P, Revisión de algoritmos genéticos aplicados al problema de la programación de cursos universitarios. Programación Matemática y Software:3(1) 49-65 (2011).
- Goldberg, D. E. y J. R. Lingle, Loci and the traveling salesman problem. Proceedings of the First International Conference on Genetic Algorithms and Their Applications, 154-159, (1985).
- Hederich M, C., A. Camargo. y C. Reyes, Ritmos cognitivos en la escuela. Universidad Pedagógica Nacional, Bogota, Colombia (2004).
- Maciques R, Trastornos del aprendizaje. Estilos de aprendizaje y el diagnóstico psicopedagógico. (2004), http://www.sld.cu/galerias/pdf/sitios/rehabilitacion-equino/trastornos_del_aprendizaje_y_estilos_de_aprendizaje_1.pdf. Acceso: 8 de julio de 2012.
- MEN, Decreto 1850. Bogotá: Ministerio de Educación Nacional. (2002). <http://www.mineducacion.gov.co/1621/article-103274.html>. Acceso: 8 de julio de 2012.
- Minh, K. y otros tres autores, Using tabu search for solving a high school timetabling problem. Studies in Computational Intelligence: 1(283) 305-313 (2010).
- Murillo, F. J, Enfoque, situación y desafíos de la investigación sobre eficacia escolar en América Latina y el Caribe. En OREALC/UNESCO, y LLECE, Eficacia escolar y factores asociados en América Latina y el Caribe (págs. 17-48). Santiago: Oficina Regional de Educación de la UNESCO para América Latina y el Caribe y el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. (2008).
- Pérez. S., Z. Niño. y M. Llobregat. Desarrollo de un programa de simulación de procesos para el tratamiento de emisiones atmosféricas, Información Tecnológica, 15 (2), 55-62 (2004).
- Rosa, M. y otros tres autores, Selección mediante simulación de una estructura de modelo cinético para la degradación aeróbica de efluentes de una industria láctea, Información tecnológica, 15 (2), 75-78 (2004).
- Ruiz, S., O. D. Castrillón. y W. Sarache, Una metodología multiobjetivo para optimizar un ambiente job shop. Información Tecnológica, 35-46.(2012).
- Peñuela M. y M. Granada, Optimización multiobjetivo usando un algoritmo genético y un operador elitista basado en un ordenamiento no-dominado (NSGA-II). Scientia et Technica:1(35) 175-180. (2007).
- Suarez V. y Castrillón O, Diseño de una metodología basada en técnicas inteligentes para la distribución de procesos académicos en ambientes de trabajo Job Shop. XV Congreso de Ingeniería de la Organización, 285 – 289, Cartagena, España 7 -9 Septiembre de 2011.
- Tassopoulos, I. X., y G. Beligiannis, Using particle swarm optimization to solve effectively the school timetabling problem. Soft Computing: 16(7) 1229-1252 (2012).
- Valouxis, C. y E. Housos, Constraint programming approach for school timetabling. Computers and Operations Research:30(10) 1555–1572 (2003).
- Wren, A, Scheduling, Timetabling and Rostering – A Special Relationship. Practice and theory of automated timetabling: 1153(1) 46-75 (1996).