Proyecto de curso
El Problema de Minimizar el extremismo presente en una Población (MinExt)
Análisis y Diseño de Algoritmos II
Escuela de Ingeniería de Sistemas y Computación

Angie Joya Duarte - 2322609   Emily Nuñez Ordoñez - 2240156
Sheila Marcela Valencia Chito - 2243011   Victoria Andrea Volveras Parra - 2241874

Julio de 2025

1. Implementación

1.1. Parámetros

Los parámetros del modelo son los siguientes:

- $n \in \mathbb{N}$: Cantidad total de personas.

$$n = \text{cardinal del conjunto de personas}$$

- $m \in \mathbb{N}$: Número de opiniones posibles que pueden tener las personas.

$$m = \text{cardinal del conjunto de opiniones}$$

- $p \in 0..n$: Cantidad de personas que tienen como opinión inicial la opinión $i \in 1..m$.

$$p = (p_1, p_2, \ldots, p_m) \in \mathbb{N}^m, \quad \sum_{i=1}^{m} p_i = n$$

- $ext \in [0,1]^m$: Nivel de extremismo asociado a cada opinión.

$$ext = (e_1, e_2, \ldots, e_m)$$

- $ce \in \mathbb{R}^m$: Costo adicional de asignar una persona a la opinión i si inicialmente no había nadie allí.

$$ce = (c_1, c_2, \ldots, c_m)$$

- $c \in \mathbb{R}^{m \times m}$: Esfuerzo requerido para mover una persona de la opinión i a la opinión j.

$$c = (c_{ij}), \quad i, j \in \{1, \ldots, m\}$$

- $ct \in \mathbb{R}$: Costo total máximo permitido.

$$\sum (\text{costos de cambio realizados}) \leq ct$$

- $maxM \in \mathbb{N}$: Número máximo de movimientos permitidos.

$$\sum (\text{personas movidas}) \leq maxM$$

- $maxp = \max(p)$: Cantidad máxima de personas que comparten una misma opinión inicialmente.

$$maxp = \max\{p_1, p_2, \ldots, p_m\}$$

1.2. Variables

Las variables del modelo son:

- $x \in \mathbb{N}^{m \times m}$: Número de personas que cambian su opinión de la i a la j.

$$x = (x_{ij}) \in \{0, \ldots, \text{maxp}\}^{m \times m}$$

- $extremismo \in \mathbb{R}$: Valor del extremismo de la población.

$$extremismo = \sum_{i=1}^{m} p_i \cdot ext_i$$

- $sol \in \mathbb{N}^m$: Cantidad de personas que terminan con la opinión i en la solución final.

$$sol = (s_1, s_2, \ldots, s_m) \in \{0, \ldots, n\}^m$$

- $total\_cost \in \mathbb{R}$: Costo total del cambio de opiniones.

$$\text{total\_cost} = \sum_{i=1}^{m} \sum_{j=1}^{m} \left[x_{i,j} \cdot c_{i,j} \left(1 + \frac{p_j}{n}\right) + x_{i,j} \cdot ce_j \cdot \mathbf{1}_{\{p_j=0\}} \right]$$

1.3. Restricciones

Las restricciones principales del modelo son:

- No negatividad: Restringe los valores de x a ser mayores o iguales que 0, dado que representan cantidades de personas y, por tanto, no pueden ser negativas. Es una restricción entera, ya que la variable involucrada es entera.

$$\forall i, j \in \{1, \ldots, m\}, \quad x_{i,j} \geq 0$$

- Cantidad de movimientos: Limita la cantidad total de cambios de opinión al máximo permitido. Es una restricción entera, ya que solo involucra variables enteras.

$$\sum_{\substack{i=1 \\ j=1 \\ x_{i,j} \neq 0}}^{m} |j - i| \cdot x_{i,j} \leq \text{maxM}$$

- Cambios de una opinión a otras: La cantidad de personas que cambian de la opinión i a otras opiniones no puede superar el número inicial de personas que tienen la opinión i. Es una restricción entera, ya que solo involucra variables enteras.

$$\sum_{j=1}^{m} x_{i,j} \leq p_i \quad \forall i \in \{1, \ldots, m\}$$

- Movimientos a la misma opinión: La cantidad de personas que pasan de la opinión i a la misma opinión i es cero, ya que no se considera un cambio real de opinión. Es una restricción entera, ya que solo involucra variables enteras.

$$x_{i,j} = 0 \quad \forall i, j \in \{1, \ldots, m\} \text{ tal que } i = j$$

- Solución final: Número de personas en la opinión i después de los cambios, calculado como el valor inicial menos quienes salieron y más quienes llegaron. Es una restricción entera, ya que solo involucra variables enteras.

$$\text{sol}_i = p_i - \sum_{j=1}^{m} x_{i,j} + \sum_{k=1}^{m} x_{k,i} \quad \forall i \in \{1, \ldots, m\}$$

- Costo total: Cantidad de recursos utilizados para realizar los cambios de opinión. Es una restricción mixta, ya que participan tanto variables enteras como variables continuas.

$$\text{total\_cost} = \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^{m} \left[x_{i,j} \cdot c_{i,j} \cdot \left(1 + \frac{p_i}{n}\right) + \begin{cases} ce_j \cdot x_{i,j} & \text{si } p_j = 0 \\ 0 & \text{si } p_j \neq 0 \end{cases} \right]$$

- Limite del costo total: El costo total de los cambios realizados debe ser menor o igual al costo máximo permitido. Es una restricción lineal, ya que participan variables continuas.

$$\text{total\_cost} \leq c_t$$

- Valor del extremismo: Nivel total de extremismo de la población después de los cambios de opinión. Es una restricción mixta, ya que participan tanto variables enteras como variables continuas.

$$\text{extremismo} = \sum_{i=1}^{m} \text{sol}_i \cdot \text{ext}_i$$

1.4. Restricciones redundantes

Se añadieron las siguientes restricciones redundantes para mejorar la propagación:

- No negatividad: Tanto el extremismo como el costo total deben ser mayores o iguales que cero. Ambas son restricciones lineales, ya que participan variables continuas.

$$\text{total\_cost} \geq 0$$
$$\text{extremismo} \geq 0$$

- Número total de personas: La suma total de personas en todas las opiniones al final de los cambios debe ser igual al número total inicial de personas n. Es una restricción entera, ya que solo involucra variables enteras.

$$\sum_{i=1}^{m} \text{sol}_i = n$$

- Límite de extremismo: El extremismo resultante no puede superar el extremismo total posible con la distribución original de personas. Es una restricción mixta, ya que participan tanto variables enteras como variables continuas.

$$\text{extremismo} \leq \sum_{i=1}^{m} p_i \cdot \text{ext}_i$$

- Opiniones vacías: Si una opinión i no tiene personas al inicio ($p_i = 0$), entonces no puede haber movimientos desde esa opinión hacia otras. Es una restricción entera, ya que solo involucra variables enteras.

$$\forall i \in \{1, \ldots, m\}, \text{ si } p_i = 0 \Rightarrow \sum_{j=1}^{m} x_{i,j} = 0$$

- Mayor extremismo: Si el extremismo de la opinión j es mayor o igual al de la opinión i la cantidad de movimientos de i a j debe ser cero. Es una restricción mixta, ya que participan tanto variables enteras como variables continuas.

$$\forall i, j \in \{1, \ldots, m\}, \text{ si } \text{ext}_j \geq \text{ext}_i \Rightarrow x_{i,j} = 0$$

1.5. Función objetivo

El objetivo es alcanzar el mínimo valor de extremismo posible.

$$\text{mín } \text{extremismo}$$

1.6. Justificación

Las diferentes restricciones permiten garantizar una solución óptima, ya que consideran diversas limitaciones que aseguran que no se supere el costo total permitido ni la cantidad máxima de movimientos. Asimismo, se garantiza que la solución final sea coherente con la cantidad original de personas y opiniones. Se definieron las variables necesarias para aplicar todas las restricciones de manera adecuada. Además, se incluyeron restricciones redundantes que mejoran la propagación al reducir el espacio de búsqueda.

2. Casos de prueba

2.1. Propios

Se diseñaron 7 pruebas propias no robustas, las cuales facilitan el análisis de la técnica de branch and bound del modelo. Los valores fueron generados de manera aleatoria; sin embargo, se aseguró que el costo del esfuerzo de mover a una persona hacia una opinión con bajo nivel de extremismo fuera elevado. Esto se hizo con el objetivo de representar que alcanzar una opinión neutral implica mayor esfuerzo. Adicionalmente, se diseñaron 2 pruebas específicas (extremismo 8 y extremismo 9) para evaluar casos extremos: uno en el que el extremismo, el costo de cambiar una persona de opinión y el costo extra son elevados; y otro en el que dichos valores son bajos. A continuación, se presentan los valores óptimos de las soluciones de cada prueba:

Nombre prueba	Valor solución óptima
extremismo	1.7
extremismo1	0.53
extremismo2	1.52
extremismo3	3.16
extremismo4	15.146
extremismo5	10.599
extremismo6	4.432
extremismo7	94.948
extremismo8	13.614

Cuadro 1: Valores óptimos de las soluciones de las pruebas propias

3. Técnica branch and bound

Para evidenciar detalladamente la técnica de branch and bound se usara la prueba `Extremismo1` Para este se definieron las siguientes variables:

$$X[i,j] = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{bmatrix}$$

$$SOL[i] = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}$$

Considere el siguiente problema:

Problema PL0

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned}
C_1 :&\ x_1, \ldots, x_9 \geq 0 \\
C_2 :&\ x_1, x_2, x_5, x_7, x_8, x_9 = 0 \\
C_3 :&\ 2x_3 + x_4 + x_6 \leq 5 \\
C_4 :&\ x_3 \leq 2 \\
C_5 :&\ x_4 + x_6 \leq 2 \\
C_6 :&\ 2 - x_3 + x_4 = y_1 \\
C_7 :&\ 2 - x_4 - x_6 = y_2 \\
C_8 :&\ 1 + x_3 + x_6 = y_3 \\
C_9 :&\ (1 + 2/5)(5.201x_3 + 1.540x_2 + 6.721x_3) \leq 10.412 \\
C_{10} :&\ (1 + 2/5)(5.201x_3 + 1.540x_2 + 6.721x_3) \geq 0 \\
C_{11} :&\ y_1 + y_2 + y_3 = 5 \\
C_{12} :&\ z \geq 0 \\
C_{13} :&\ z \leq 1.33
\end{aligned}$$

Solución de PL0: $x_3 = 0,\ x_4 \approx 1.15,\ x_6 \approx 0.84,\ y_1 \approx 3.15,\ y_2 = 0,\ y_3 \approx 1.84,\ z \approx 0.47$

Esta solución no es entera, por lo tanto, se debe ramificar. Se puede evidenciar que hay cuatro variables no enteras, en este primer caso se denominara x_6 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $x_6 \leq 0$ o $x_6 \geq 1$.

Se generan dos problemas, agregando al problema anterior una de las restricciones anterior.

Problema PL1

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \leq 0 \end{gathered}$$

Problema PL2

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \geq 1 \end{gathered}$$

Problemas pendientes a resolver: { PL1, PL2 }. Primero se solucionara el problema **PL1**.

Solución de PL1: $x_3 \approx 0.83,\ x_4 = 2,\ x_6 = 0,\ y_1 \approx 3.16,\ y_2 = 0,\ y_3 \approx 1.83,\ z \approx 0.47$

Se denominara x_3 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $x_3 \leq 0$ o $x_3 \geq 1$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL3

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \leq 0 \\ x_3 \leq 0 \end{gathered}$$

Problema PL4

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \leq 0 \\ x_3 \geq 1 \end{gathered}$$

Problemas pendientes a resolver: { PL3, PL4, PL2 }. Primero se solucionara el problema **PL3**.

Solución de PL3: $x_3 = 0,\ x_4 = 2,\ x_6 = 0,\ y_1 = 4,\ y_2 = 0,\ y_3 = 1,\ z = 0.53$

Teniendo en cuenta que se obtuvo una solución entera entonces esta sera candidata para ser la óptima, por lo tanto, se debe de almacenar.

Problemas pendientes a resolver: { PL4, PL2 }. Primero se solucionara el problema **PL4**.

Solución de PL4: $x_3 = 1,\ x_4 \approx 1.45,\ x_6 = 0,\ y_1 \approx 2.45,\ y_2 \approx 0.54,\ y_3 = 2,\ z \approx 0.67$

Se denominara x_4 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $x_4 \leq 1$ o $x_3 \geq 2$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL5

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\leq 0 \\ x_3 &\geq 1 \\ x_4 &\leq 1 \end{aligned}$$

Problema PL6

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\leq 0 \\ x_3 &\geq 1 \\ x_4 &\geq 2 \end{aligned}$$

Problemas pendientes a resolver: { PL5, PL6, PL2 }. Primero se solucionara el problema **PL5**.

Solución de PL5: $x_3 \approx 1.13,\ x_4 = 1,\ x_6 = 0,\ y_1 \approx 1.86,\ y_2 = 1,\ y_3 \approx 2.13,\ z \approx 0.85$

Se denominara x_3 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $x_3 \leq 1$ o $x_3 \geq 2$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL7

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\leq 0 \\ x_3 &\geq 1 \\ x_4 &\leq 1 \\ x_3 &\leq 1 \end{aligned}$$

Problema PL8

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\leq 0 \\ x_3 &\geq 1 \\ x_4 &\leq 1 \\ x_3 &\geq 2 \end{aligned}$$

Problemas pendientes a resolver: { PL7, PL8, PL6, PL2 }. Primero se solucionara el problema **PL7**.

Solución de PL7: $x_3 = 1,\ x_4 = 1,\ x_6 = 0,\ y_1 = 2,\ y_2 = 1,\ y_3 = 0,\ z = 0.86$

Se obtuvo una solución entera, pero el **z** es mayor que el que se tenia almacenado, teniendo en cuenta que se busca minimizar esta solución se descarta.

Problemas pendientes a resolver: { PL8, PL6, PL2 }. Primero se solucionara el problema **PL8**, teniendo en cuenta las otras restricciones este no es satisfactible.

Problemas pendientes a resolver: { PL6, PL2 }. Primero se solucionara **PL6**, el cual no es satisfactible, por lo tanto, este problema esta agotado.

Se solucionara **PL2** siendo este el único problema por resolver.

Solución de PL2: $x_3 = 0,\ x_4 \approx 0.46,\ x_6 = 0,\ y_1 \approx 2.46,\ y_2 \approx 0.53,\ y_3 = 2,\ z \approx 0.67$

Se denominara x_4 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $x_4 \leq 0$ o $x_4 \geq 1$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL9

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \geq 1 \\ x_4 \leq 0 \end{gathered}$$

Problema PL10

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \geq 1 \\ x_4 \geq 1 \end{gathered}$$

Problemas pendientes a resolver: { PL9, PL10 }. Primero se solucionara el problema **PL9**.

Solución de PL9: $x_3 = 0,\ x_4 = 0,\ x_6 \approx 1.10,\ y_1 = 2,\ y_2 \approx 0.89,\ y_3 \approx 2.10,\ z \approx 0.80$

Se denominara y_3 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $y_3 \leq 2$ o $y_3 \geq 3$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL11

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \geq 1 \\ x_4 \leq 0 \\ y_3 \leq 2 \end{gathered}$$

Problema PL12

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \geq 1 \\ x_4 \leq 0 \\ y_3 \geq 3 \end{gathered}$$

Problemas pendientes a resolver: { PL11, P12, PL10 }. Primero se solucionara el problema **PL11**.

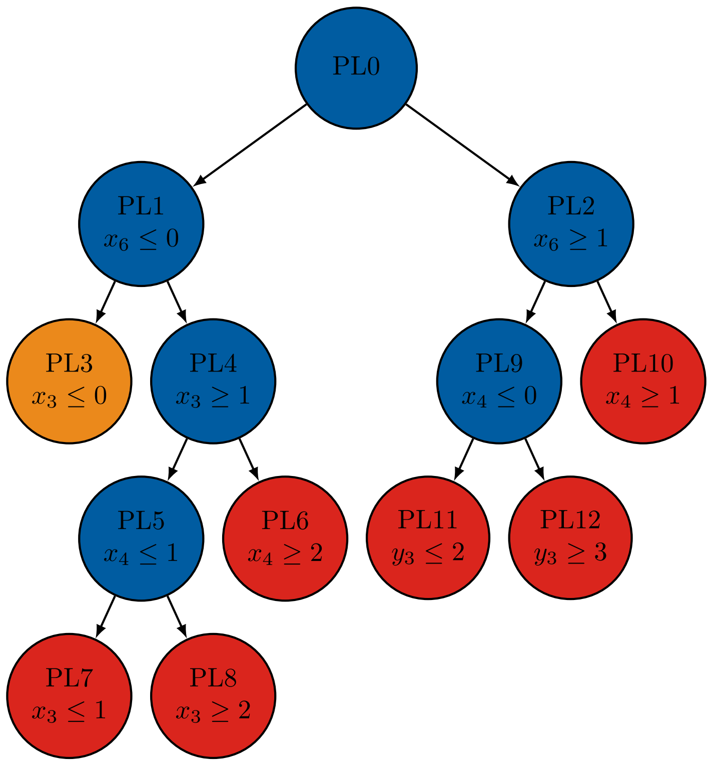
Solución de PL11: $x_3 = 0,\ x_4 = 0,\ x_6 = 1,\ y_1 = 2,\ y_2 = 1,\ y_3 = 2,\ z = 0.86$

Se obtuvo una solución entera, pero el **z** es mayor que el que se tenia almacenado, teniendo en cuenta que se busca minimizar esta solución se descarta.

Problemas pendientes a resolver: { PL12, PL10 }. Primero se solucionara el problema **PL12**, teniendo en cuenta las otras restricciones este no es satisfactible.

El único problema pendiente por resolver es el **PL10**, el cual no es satisfactible, terminando así la búsqueda.

Solución optima: PL3



Se debe tener en cuenta que la elección de la variable de ramificación tiene un gran impacto en el espacio de búsqueda, ya que esta podría hacer que se expanda una mayor o menor cantidad de nodos. Por ejemplo:

Para este caso { **PL0,PL1,PL2,PL3,PL4,PL5,PL6,PL7,PL8,PL9,PL10** } se mantienen iguales que en el ejemplo anterior, para **PL9** se definirá como variable de ramificación x_6. Por consecuente como debe ser entera, se tiene que $x_6 \leq 1$ o $x_6 \geq 2$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL11

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{gathered} C_1 \\ \vdots \\ C_{13} \\ x_6 \geq 1 \\ x_4 \leq 0 \\ x_6 \leq 1 \end{gathered}$$

Problema PL12

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\geq 1 \\ x_4 &\leq 0 \\ x_6 &\geq 2 \end{aligned}$$

Problemas pendientes a resolver: { PL11, P12, PL10 }. Primero se solucionara el problema **PL11**.

Solución de PL11: $x_3 \approx 0.13,\ x_4 = 0,\ x_6 = 1,\ y_1 \approx 1.86,\ y_2 = 1,\ y_3 \approx 2.13,\ z \approx 0.85$

Se denominara x_3 como la variable de ramificación. Por consecuente como debe ser entera, se tiene que $x_3 \leq 0$ o $x_3 \geq 1$. Teniendo esto en cuenta se generan dos nuevos problemas.

Problema PL13

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\geq 1 \\ x_4 &\leq 0 \\ x_6 &\leq 1 \\ x_3 &\leq 0 \end{aligned}$$

Problema PL14

$$\text{mín } z = 0.12y_1 + 0.52y_2 + 0.05y_3$$

sujeto a

$$\begin{aligned} &C_1 \\ &\vdots \\ &C_{13} \\ x_6 &\geq 1 \\ x_4 &\leq 0 \\ x_6 &\leq 1 \\ x_3 &\geq 1 \end{aligned}$$

Problemas pendientes a resolver: { PL13, P14, P12, PL10 }. Primero se solucionara el problema **PL13**.

Solución de PL13: $x_3 = 0,\ x_4 = 0,\ x_6 = 1,\ y_1 = 2,\ y_2 = 1,\ y_3 = 2,\ z = 0.86$

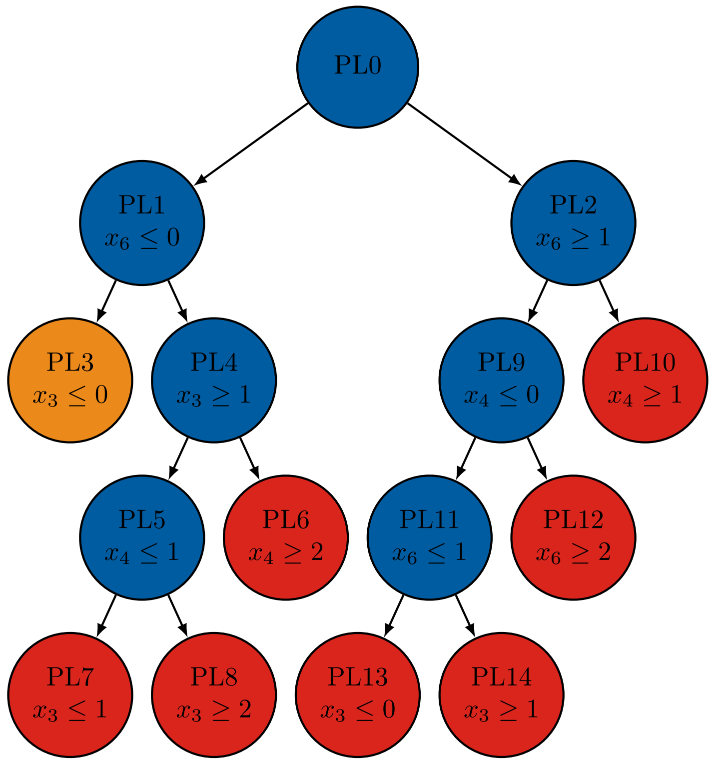
Se obtuvo una solución entera, pero el **z** es mayor que el que se tenia almacenado, teniendo en cuenta que se busca minimizar, esta solución se descarta.

Problemas pendientes a resolver: { P14, PL12, PL10 }. Primero se solucionara el problema **PL14**, teniendo en cuenta las otras restricciones este no es satisfactible.

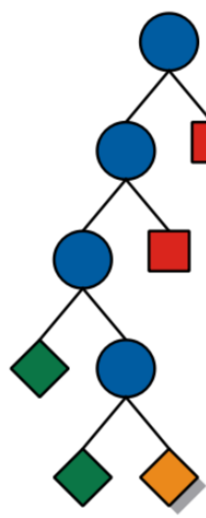
Problemas pendientes a resolver: { PL12, PL10 }. Primero se solucionara **PL12**, el cual no es satisfactible, por lo tanto, este problema esta agotado.

El único problema pendiente por resolver es el **PL10**, siendo este no satisfactible, la búsqueda termina.

Solución optima: PL3

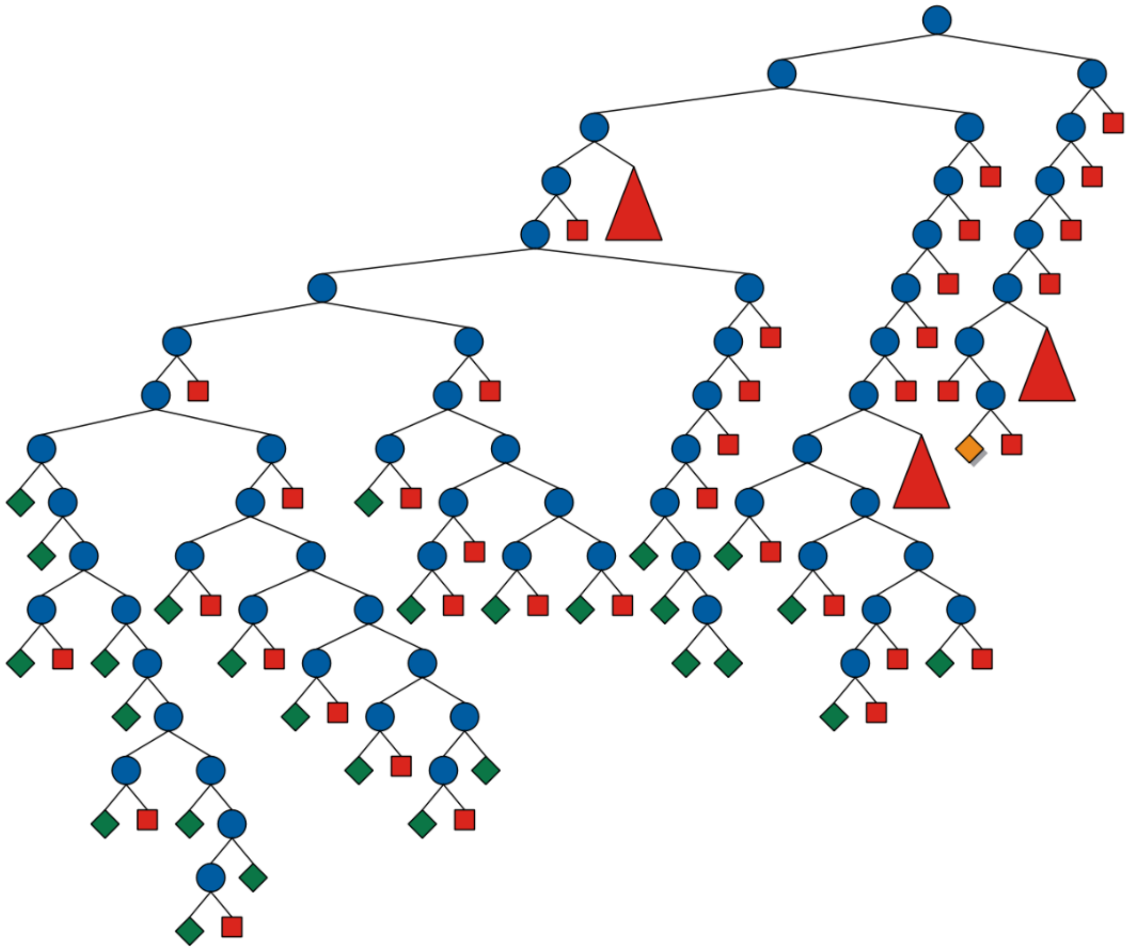


Se puede evidenciar que en el segundo caso se expande un nodo de mas y se tienen 2 nodos fallidos adicionales, demostrando la importancia de la elección de la variable de ramificación y como afecta el espacio de búsqueda. Al ejecutar la prueba con el solver Gecode Gist se obtuvo el siguiente árbol de búsqueda:



Se puede observar que con el solver se expande una menor cantidad de nodos, a su vez hay menos nodos fallidos y encuentra una mayor cantidad de soluciones, esto es debido a la estrategia de búsqueda que usa Gecode Gist, es decir su combinación de heurísticas, donde elige cual variable ramificar primero y cual valor de su dominio tomar, teniendo esto en cuenta, el solver real aplica branch and bound, las heurísticas y exploración completa para garantizar optimalidad, explicando la diferencia en la cantidad respectiva de nodos.

Se ejecuto la `Prueba7`, y se obtuvo el siguiente árbol de búsqueda:



Se tiene las siguientes variables:

$$X[i,j] = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ x_6 & x_7 & x_8 & x_9 & x_{10} \\ x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{16} & x_{17} & x_{18} & x_{19} & x_{20} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \end{bmatrix}$$

$$SOL[i] = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 & y_5 \end{bmatrix}$$

Teniendo en cuenta las restricciones se le asigna el valor a las siguientes variables:

$$x_1, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}, x_{19}, x_{21}, x_{22}, x_{23}, x_{24}, x_{25} = 0$$

Por lo tanto se tiene que hallar el valor que tienen que tomar las siguientes 9 variables:

$$x_2, x_3, x_4, x_5, x_8, x_9, x_{10}, x_{18}, x_{20}$$

La primera variable que ramifica es x_{20}, por lo tanto, para el nodo hijo de la rama izquierda explora $x_{20} \leq 0$, en esta se tiene como variable de ramificación x_{18}, nuevamente en la rama izquierda se explora $x_{18} \leq 0$, este mismo proceso se repite para las otras variables, explorándolas en orden descendente, al llegar al nodo donde se explora $x_2 \leq 0$ encuentra la primera solución, la cual es el problema inicial, donde ninguna persona cambia de opinión, igualmente esta será candidata para ser la óptima, por lo tanto, es almacenada.

Al explorar $x_2 \geq 1$ se hace un recorrido parecido al mencionado anteriormente, es decir, primero se explora el menor valor posible del dominio que puede tomar x_2 ($x_2 \leq 1$) este es el hijo izquierdo del nodo que se esta expandiendo ($x_2 \geq 1$), posteriormente se explora $x_2 \geq 2$, y así sucesivamente, este proceso se repite para todas las variables. Se debe tener en cuenta que para la búsqueda se aplica backtracking y branch and bound.

En el árbol de búsqueda los nodos azules representan los nodos que se están expandiendo, es decir, donde se esta aplicando la ramificación. Los nodos verdes representan valores que cumplen todas las restricciones, adicionalmente estos son candidatos para ser la solución óptima, por lo tanto, a medida que va avanzando la búsqueda el valor del extremismo de estos nodos va disminuyendo, teniendo en cuenta que se esta minimizando. Los nodos rojos representan valores donde no se encuentra una solución, es decir, cuando es no satisfactible, también puede haber casos donde se encuentre una solución que cumpla con todas las restricciones, pero el valor de esta sea mayor o igual al extremismo de la solución almacenada hasta ese momento. El triángulo rojo representa la exploración de uno o más nodos donde no se encontró ninguna solución candidata para ser la óptima. Finalmente el nodo naranja representa la solución óptima del problema.

4. Análisis

4.1. Tiempo de ejecución

Debido a que se utilizó el solver Gurobi 12.0.3, el tiempo de ejecución de todas las pruebas es corto, variando entre 0,6 y 1,6 segundos. Se puede apreciar que este tiende a aumentar a la medida que aumenta el tamaño de n, pero sigue siendo un valor considerablemente bajo, por lo que es fácilmente influenciado por los procesos de scheduling del sistema operativo y no es prudente concluir al respecto.

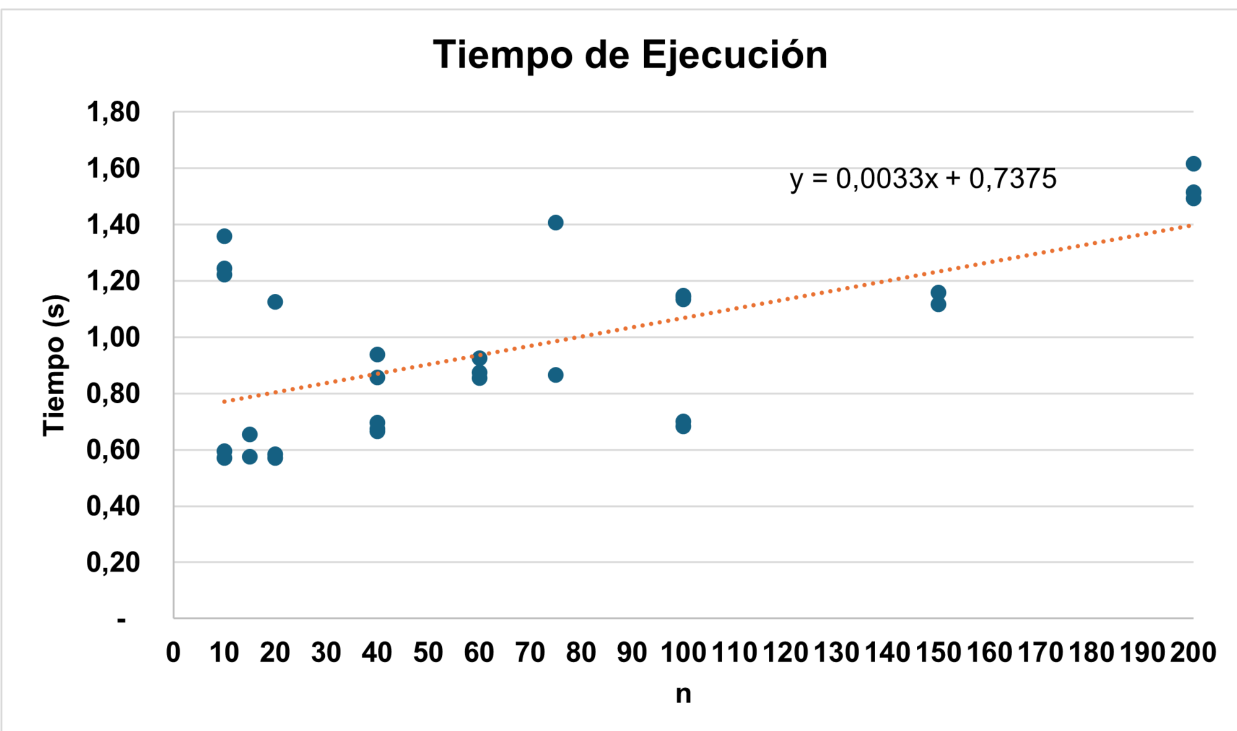


Figura 1: Tiempo de ejecución del algoritmo para poblaciones de 10, 25, 20, 40 y 200 personas.

4.2. Árboles de ramificación

4.2.1. Baterías de pruebas

Como se mencionó anteriormente, el modelo no presenta una estrategia de búsqueda que permita obtener los árboles de ramificación eficientemente. En consecuencia, no se pudo realizar el análisis de los árboles de ramificación de las 30 pruebas.

Se toma una muestra de 13 árboles, que corresponden a las pruebas en las que el *solver* Geocode Gist 6.3.0 generó el árbol en un tiempo inferior a 10 horas.

Es importante mencionar que, en las primeras ejecuciones del modelo, la cantidad de nodos expandidos era demasiado grande, por lo que se incluyó la restricción *Mayor extremismo* (explicada en la Sección 2.4) para acotar dicho conjunto. Después de esta implementación, la cantidad de nodos expandidos se redujo notablemente, permitiendo obtener los árboles en menor tiempo de ejecución.

A continuación, se presenta una gráfica de los nodos explorados:

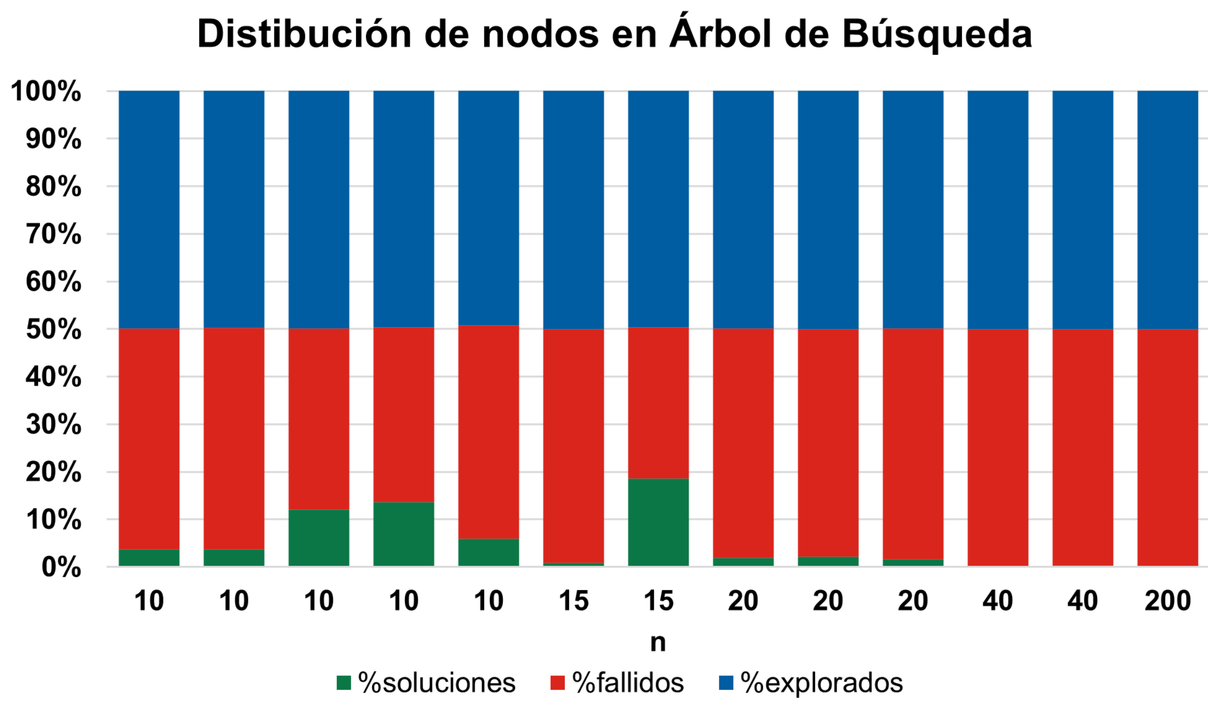


Figura 2: Distribución de nodos fallidos, explorados y solución para poblaciones de 10, 25, 20, 40 y 200 personas.

Geocode Gist 6.3.0 realiza una poda del 50 % de los nodos explorados, evidenciando que la estrategia de búsqueda empleada en esta muestra es eficiente, reduciendo el consumo de recursos computacionales en la búsqueda de la solución óptima.

Por otra parte, los nodos solución no superan el 20 %. Este porcentaje puede ser reflejo de las restricciones del modelo, ya que, acotan el conjunto de soluciones, lo que asegura la solución optima.

Se puede observar que en la mayoría de los casos a medida que va creciendo la cantidad de opiniones, también aumentan los nodos expandidos, esto es debido a que el tamaño de la matriz de la variable x depende de este valor. Basándonos en el modelo, se debe de tener muy en cuenta el extremismo de cada una de las opiniones, ya que luego de que se aplique la restricción mayor extremismo algunas variables de x se les asignara el valor de 0, dejando solamente las variables donde no se cumple que $x[j] \geq x[i]$, estos son las valores que varían en el espacio de búsqueda para encontrar la solución optima, impactando así la cantidad de nodos expandidos.

4.3. Estadisticas

En el siguiente enlace: Análisis, se encuentran dos hojas de cálculo. En la primera hoja se presentan los resultados obtenidos con las baterías de prueba, y en la segunda, los resultados de las pruebas propias.

Para ambos casos se detalla:

- El tiempo de ejecución.
- El valor de la solución.
- El costo de la solución.
- Cuántas personas cambiaron de opinión (suma x).
- El porcentaje de costo que empleó el modelo del costo total disponible ($ct\,\%$).
- El porcentaje de personas que cambiaron de opinión ($n\,\%$).

Además, se incluyen las estadísticas de los árboles obtenidos. Las casillas que se encuentran en gris corresponden a las pruebas para las que el *solver* no pudo encontrar una solución en un tiempo menor a 10 horas.

4.3.1. Baterías de pruebas

La gráfica ilustra el porcentaje de mejora, el costo empleado y la cantidad de personas moderadas; por lo tanto, no se puede evidenciar una correlación positiva ni negativa entre estas variables. Esto también es visible en las columnas ($ct\,\%$) y ($n\,\%$) de las hojas de Excel.

Por ejemplo, en la prueba 4, el costo es bajo y el porcentaje de cambio de opinión de personas es alto; mientras que, en la prueba 24, tanto el costo como el porcentaje de personas moderadas son altos.

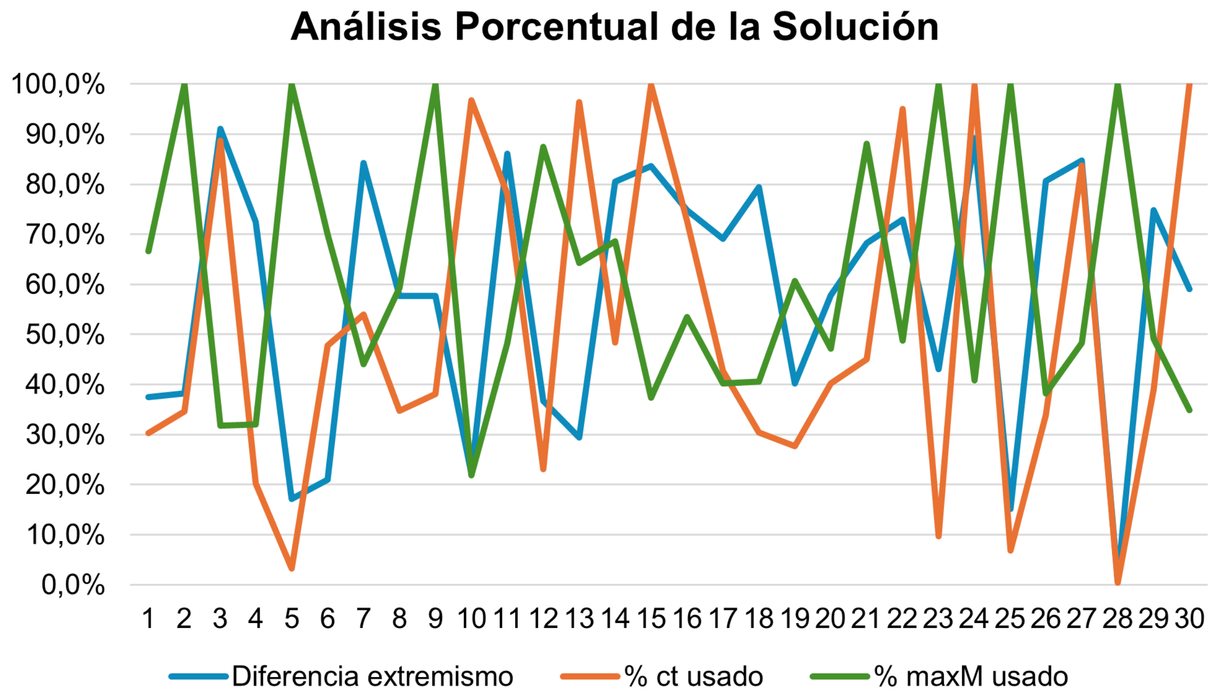


Figura 3: Análisis porcentual de la solución para las 30 pruebas.

Por consiguiente, se puede afirmar que el comportamiento de dichos valores no guardar relación directa entre sí. Sin embargo, los parámetros de entrada y las restricciones afecta su moderación.

Por otra parte, el 63,33 % de las pruebas cambiaron a más del 50 % de personas de opinión, mientras que el 36,66 % gastó más del 50 % del costo total disponible.

También, el 63,33 % de las pruebas redujeron el extremismo en más del 50 %, demostrando la eficacia del modelo desarrollado.

4.3.2. Pruebas propias

Como se mencionó anteriormente, no se puede evidenciar una correlación positiva ni negativa entre el porcentaje de costo, las personas moderadas y la mejora.

Por otra parte, el 33,33 % de las pruebas cambiaron a más del 50 % de personas de opinión, mientras que el 66,66 % gastó más del 50 % del costo total disponible.

Se puede observar que, en comparación con la batería de pruebas, las pruebas propias requieren un mayor costo para moderar personas. Este comportamiento, se debe en parte a los criterios en su diseño, los cuales se explicaron en la sección 3.

Además, el 44,44 % de las pruebas redujeron el extremismo en más del 50 %, demostrando que las pruebas implementadas representan poblaciones que necesitan de esfuerzos elevados para su alcanzar opiniones neutrales.

En el caso de la prueba *extremismo7*, donde se representó una población con extremismo elevado, se evidenció que eran necesarios valores altos de ct para reducir dicho extremismo. Por ejemplo, cuando su ct fue de 1000, la reducción del extremismo fue del 97.5 %.

Video

El video se puede encontrar en el siguiente link: Ver carpeta de Drive

5. Conclusiones

Se logró desarrollar un modelo que garantiza la solución óptima para el problema de minimizar el extremismo presente en una población. El 63,33 % de las pruebas redujo el extremismo en más del 50 %. Se emplearon técnicas de programación lineal y entera mixta para modelar las restricciones, parámetros, variables y función objetivo del modelo.

Por otra parte, la rapidez en la ejecución del método *branch and bound* está sujeta a la heurística del *solver* que realice las pruebas, ya que el modelo no tiene una estrategia de búsqueda propia, siendo la elección de las variables de ramificación un factor relevante en la eficiencia del programa.

Por consiguiente, se logró obtener los árboles de ramificación de 19 pruebas, donde en todos los casos el 50 % de los nodos fueron fallidos, y los nodos solución no sobrepasan el 19 %.

Los tiempos de ejecución del modelo en las 39 pruebas testeadas, empleando el *solver* Gurobi, varían desde los 584 ms hasta los 1615,72 ms, encontrando la solución óptima en el 100 % de los casos. Las pruebas evidenciaron que el tiempo de ejecución no depende necesariamente del tamaño de la población, sino que varía según la complejidad de la entrada.