

Proyecto de curso

# **El Problema de Minimizar la Polarización presente en una Población (MinPol)**

Análisis de Algoritmos II

Escuela de Ingeniería de Sistemas y Computación



Profesor Juan Francisco Díaz Frias      Profesor Jesús Alexander Aranda  
Monitor Mauricio Muñoz

Octubre de 2024

## **1. Introducción**

El presente proyecto tiene por objeto verificar que los estudiantes han adquirido el siguiente resultado de aprendizaje:

Construye modelos de optimización en términos de parámetros, variables, restricciones y función objetivo, a partir de un problema determinado, para explorar soluciones prácticas utilizando herramientas computacionales de modelamiento y solvers existentes.

Para ello, los estudiantes deben demostrar que logran:

- Utilizar el método branch and bound para resolver problemas de programación binaria, entera y mixta.
- Usa técnicas de programación lineal para modelar/solucionar problemas de programación lineal en términos de parámetros, variables, restricciones y función objetivo.
- Usa técnicas de programación entera para modelar/solucionar problemas de programación entera en términos de parámetros, variables, restricciones y función objetivo.
- Usa técnicas de programación entera mixta para modelar/solucionar problemas de programación entera mixta en términos de parámetros, variables, restricciones y función objetivo.
- Usa un lenguaje de modelamiento para escribir y probar modelos de programación lineal, entera y entera mixta.

Para ello el estudiante:

- Desarrolla un programa utilizando tecnologías de programación adecuadas para resolver en grupo un proyecto de programación planteado por el profesor.

- Escribe un informe de proyecto, presentando los aspectos más relevantes del desarrollo realizado, para que un lector pueda evaluar el proyecto.
- Desarrolla una presentación digital, con los aspectos más relevantes del desarrollo realizado, para sustentar el trabajo ante los compañeros y el profesor.

## 2. El Problema de Minimizar la Polarización presente en una Población: MinPol

### 2.1. Contexto del problema

La polarización, así como el extremismo, es un fenómeno que se presenta cada vez de forma más frecuente en nuestra sociedad, este fenómeno se agudiza cuando las sociedades se tienden a dividir en dos grandes bandos, de tamaños similares, donde la opinión de cada bando es totalmente opuesta a la del otro.

El nivel de polarización presente en una sociedad se puede medir de diversas maneras, una forma de hacerlo es considerando el esfuerzo requerido para llevar a toda la población a un consenso. Donde el consenso se refiere a que todas las personas han alcanzando una opinión común con respecto a un tema.<sup>1</sup>

Entendiendo que un alto nivel de polarización, al menos en algunos casos, tiene un potencial efecto corrosivo y perjudicial en el funcionamiento de comunidades, sociedades y democracias, es interesante estudiar estrategias que permitan reducir la polarización. Sin embargo, los esfuerzos para reducir la polarización tienen un costo que puede ser muy alto, por ende la toma de decisiones que se tomen en pro de reducir la polarización constituye un problema bastante relevante y es el tema central de este proyecto.

En este proyecto, partiendo de las opiniones iniciales de una población, se determinará que esfuerzos se harán para cambiar la opinión de algunos con el fin de alcanzar el menor nivel de polarización, entendiendo que los esfuerzos tienen un costo y que este no puede ser más elevado de un umbral definido.

### 2.2. El problema

El problema de minimizar la polarización en una población consiste en decidir qué esfuerzos se harán para cambiar la opinión de un grupo de personas y hacia donde, teniendo en cuenta que cada esfuerzo cuesta y que hay recursos limitados, de tal forma que la población termine lo menos polarizada posible.

En un ejemplo simplificado, imagine que en una población de tamaño 10, donde tenemos tres posibles opiniones: opinión 1, opinión 2 y opinión 3 ; las opiniones iniciales de la población sobre una propuesta son las siguientes:

- 8 personas comparten la opinión 1.
- 0 personas comparten la opinión 2.
- 2 personas comparten la opinión 3.

El esfuerzo para mover una persona de una opinión a otra se muestra en la siguiente tabla:

Costo esfuerzo	opinión 1	opinión 2	opinión 3
opinión 1	0	3	5
opinión 2	2	0	3
opinión 3	4	3	0

---

<sup>1</sup>La polarización es un fenómeno diferente al extremismo, por ejemplo, un grupo de personas que comparta una opinión radical sobre un tema sería extremista, sin embargo, dicho grupo no estaría polarizado ya que todos piensan lo mismo.

Adicionalmente, el costo máximo permitido de los esfuerzos es 9.

¿Cuáles serían los esfuerzos o acciones que se harían para lograr minimizar la polarización, si las acciones permitidas sólo consisten en pasar  $x$  personas de una opinión  $i$  a una opinión  $j$ ?

En este caso, un esfuerzo podría ser el siguiente: mover dos personas de la opinión 3 a la opinión 1. Este esfuerzo, según la tabla, costaría en total 8 (ya que sería el costo individual de mover de  $Op_3$  a  $Op_1$  por dos ya que se mueven dos personas), siendo este costo menor al umbral de 9. Como en este caso se alcanza consenso en la opinión 1, el valor de la polarización alcanzado es 0. Por lo tanto mover las opiniones de las dos personas de la opinión 3 a la opinión 1 es una solución del problema.

En la realidad y en este ejercicio, el problema es un poco más complejo, porque el costo de mover  $x$  personas de una opinión  $i$  a una opinión  $j$  no es tan simple como en este ejemplo. Y el número de cambios de opinión pueden estar restringidos.

Por otro lado, ¿qué ocurre cuando no se alcanza consenso, porque el esfuerzo para ello es superior al umbral permitido? El objetivo sería llegar a una configuración donde la polarización sea mínima. Pero, ¿cómo se mide en ese caso la polarización? Para medir la polarización en general, incluyendo casos donde no se alcanza consenso, usaremos la siguiente fórmula:

$$Pol(p, v) = \sum_{i=1}^m p_i |v_i - mediana(p, v)| \quad \text{calcular polarización}$$

donde  $m$  es el número de posibles opiniones,  $p$  es el vector con la distribución de personas por opinión,  $p_i$  es el número de personas con opinión  $i$ ,  $v$  es el vector con los valores de todas las opiniones, es decir  $v_i$  es el valor de la opinión  $i$ ; y  $mediana(p, v)$  corresponde al valor de la mediana de los valores de las opiniones de todas las personas que forman la población de acuerdo a la distribución  $p$ .

### 2.3. Formalización

Sea  $n \in \mathbb{N}$  el número total de personas. Sea  $m \in \mathbb{N}$  el número de opiniones posibles que pueden tener las personas.

Sea  $p_i \in 0..n$  el número de personas que tienen como opinión inicial la opinión  $i \in 1..m$

Sea  $v_i \in [0, 1]$  el valor real correspondiente a la opinión  $i \in 1..m$

Sea  $c_{i,j} \in \mathbb{R}^+$  el costo del esfuerzo de mover una persona de opinión  $i$  a la opinión  $j$ , donde  $i, j \in 1..m$ .

Obviamente  $c_{i,i} = 0, i \in 1..m$ .

Sea  $ce_i \in \mathbb{R}^+$  el costo extra de mover una persona a la posición  $i$  si dicha posición estaba inicialmente sin personas  $i \in 1..m$ .

Sea  $ct \in \mathbb{R}^+$  el costo total máximo permitido de todos los esfuerzos.

Sea  $maxM \in \mathbb{N}^+$  el máximo número de movimientos permitidos.

El problema consiste en decidir qué esfuerzos (donde cada esfuerzo consiste en mover un determinado número de personas de una opinión  $i$  a una opinión  $j$ ) se harán tal que el costo total permitido para todos los esfuerzos no sea superado y que se minimice la polarización de la población.

Naturalmente, el número de personas que se mueven de una opinión a otras, no puede ser mayor que el número de personas que tenían inicialmente esa opinión.

Además, el número de movimientos total va a estar limitado por un número máximo de movimientos posibles. Mover una persona de la opinión  $i$  a la opinión  $j$  se contabilizará como  $|j - i|$  movimientos.

Y obviamente, cada persona tiene una y sólo una opinión, tanto en la distribución original, como en la resultante.

Ahora, ¿cuánto cuesta una solución?

El costo de una solución es la suma de los costos de cada movimiento.

El costo de un movimiento de  $x$  personas de la opinión  $i$  a la opinión  $j$  está dado por  $c_{i,j}(1 + p_i/n) * x$ , si  $p_j > 0$ . En el caso en que  $p_j = 0$ , a ese costo se le suma  $ce_j * x$ .

## El Problema de Minimizar la Polarización presente en una Población: MinPol

**Entrada:**  $n \in \mathbb{N}, m \in \mathbb{N}, p_i \in 0..n, ce_i \in \mathbb{R}^+, c_{i,j} \in \mathbb{R}^+, i, j \in 1..m, ct \in \mathbb{R}^+, maxM \in \mathbb{N}^+$

**Salida:**  $x_{i,j} \in \mathbb{N}, i \in 1..m, j \in 1..m$  son el número de personas que pasarán de tener una opinión  $i$  a una opinión  $j$ . Tal que se minimice la polarización, respetando las restricciones propias del problema.

### 2.4. ¿Entendimos el problema?

Una población formada por 20 personas, considerando 5 posibles diferentes opiniones, se distribuyen como se muestra en el cuadro 1.

Núm. personas por opinión	Opiniones
12	1
4	2
0	3
4	4
0	5

Cuadro 1: Distribución de población por opinión.

Los valores de las opiniones se aprecian en el cuadro 2.

Opiniones	Valor
1	0
2	0.25
3	0.5
4	0.75
5	1

Cuadro 2: Valores de las opiniones posibles.

El costo individual de ir de una opinión a otra se describe en el cuadro 3:

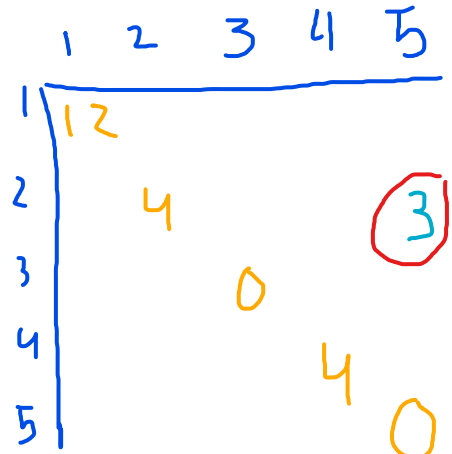
Costo	op. 1	op.2	op. 3	op. 4	op. 5
op. 1	0	2	4	5	7
op. 2	1	0	3	4	6
op. 3	3	2	0	2	4
op. 4	4	2	1	0	2
op. 5	8	5	3	2	0

Cuadro 3: Costos individuales de los cambios de opinión.

Adicionalmente en el cuadro 4 se muestra el costo extra del esfuerzo individual de pasar a una opinión de llegada cuando esta opinión inicialmente no era compartida por ninguna persona. Este costo extra sólo se aplicará bajo el cumplimiento de la condición anterior.

El costo total máximo permitido es 20, y el número máximo de movimientos es 18.

En este caso la entrada se describiría así:



Opiniones de llegada	Costo extra
1	1
2	2
3	1
4	1
5	3

Cuadro 4: Costos extras de desplazamientos a opiniones inicialmente no compartidas por ninguno.

**Entrada:**  $n = 20, m = 5, p_i \in \mathbb{N}, i \in 1..m$  según el cuadro 1,  $c_{i,j} \in \mathbb{R}^+, i, j \in 1..m$  según el cuadro 3,  $ce_i \in \mathbb{R}^+, i \in 1..m$  según el cuadro 4,  $ct = 20, MaxMovs = 18$ .

Describa al menos tres salidas diferentes para esta instancia, que cumpla todas las restricciones y calcule su respectiva polarización.

¿Es alguna de ellas una solución óptima? Si no, describa una solución óptima.

### 3. El proyecto: Modelamiento e Implementación

Usted como ingeniero ha sido contratado para resolver el problema y debe:

- Proponer un modelo genérico para solucionar el problema. El modelo debe ser incluido en formato pdf y debe contener: parámetros, variables, restricciones, función objetivo. El modelo debe utilizar notación formal para que soporte cualquier instancia con la entrada definida en la Sección 3.1.
- Generar 5 instancias para retar a otros proyectos. Para cada instancia debe incluir la entrada y la salida esperada (el valor del óptimo, o por lo menos el valor de la mejor solución que su grupo haya encontrado)
- Implementar el modelo genérico en MiniZinc (**Proyecto.mzn**).
- Incluir una tabla con pruebas realizadas sobre las instancias que se proveen con el proyecto y las 5 instancias creadas por su grupo de trabajo. Realice un análisis sobre los resultados obtenidos (incluya el análisis en el informe con el modelo).
- Desarrollar una interfaz gráfica con la tecnología de su predilección que permita configurar o leer una entrada para el problema (la entrada deberá convertirse a formato dzn para poder ser ejecutada por el modelo cumpliendo con las características de la entrada definida en la Sección 3.1) y visualizar la salida. Esta interfaz junto con el modelo sería el entregable para el cliente y será utilizada por algún operario. La interfaz debe incluir un botón que al presionarlo:
  - Cree un archivo **DatosProyecto.dzn** con los datos proporcionados en la interfaz
  - Ejecute el modelo genérico **Proyecto.mzn** sobre los datos proporcionados
  - Despliegue los resultados de la solución
- Incluya los archivos fuente de su implementación gráfica en un directorio llamado **ProyectoGUI-Fuentes**

Para mayor información sobre la forma de ejecutar un modelo MiniZinc a través de línea de comandos visite:

- Modelamiento básico en MiniZinc: <https://www.minizinc.org/doc-2.2.3/en/modelling.html>

- Modelos más complejos: <https://www.minizinc.org/doc-2.2.3/en/modelling2.html>
- Hacer un vídeo de máximo 4 minutos donde muestre su aplicación funcionando. Debe explicar brevemente los componentes y realizar por lo menos dos pruebas con distintas configuraciones donde se muestren las soluciones en la interfaz gráfica. **Se debe incluir un enlace al video en el archivo pdf del informe.**

### 3.1. Entrada

La entrada se leerá de un archivo de texto \*.mpl con la siguiente información:

1. La primera línea contiene un entero indicando el número de personas (es decir, conteniendo  $n$ ).
2. La segunda línea contiene un entero indicando el número de posibles opiniones (es decir, conteniendo  $m$ ).
3. La siguiente línea contiene, una lista de  $m$  valores enteros correspondientes a la distribución de las personas según su opinión inicial, separados por comas,  $p_i, i \in 1..m$ . La suma de esos valores es  $n$ .
4. La siguiente línea contiene, una lista de  $m$  valores reales correspondientes a los valores de las opiniones posibles separados por comas, correspondientes, en su orden, a  $v_i, i \in 1..m$ .
5. La siguiente línea contiene, una lista de  $m$  valores reales separados por comas, correspondientes, en su orden, a los valores de los costos extras de las opiniones posibles, es decir a  $ce_i, i \in 1..m$ .
6. Las siguientes  $m$  líneas, contienen, cada una, una lista de  $m$  valores separados por comas, correspondientes, en su orden, a los costos del desplazamiento de cada opinión a las  $m$  opiniones posibles; es decir, cada línea contiene  $c_{i,j}, i, j \in 1..m$ .
7. La siguiente línea contiene, un valor real correspondiente al costo total máximo permitido (es decir, conteniendo  $ct$ ).
8. La siguiente línea contiene, un valor correspondiente al costo total máximo permitido (es decir, conteniendo  $MaxMvs$ ).

El archivo correspondiente a la entrada del ejemplo de la sección 2.4 sería:

```
20
5
12,4,0,4,0
0,0.25,0.5,0.75,1
1,2,1,1,3
0,2,4,5,7
1,0,3,4,6
3,2,0,2,4
4,2,1,0,2
8,5,3,2,0
20
18
```

En el campus se compartirán ejemplos de entradas con sus respectivas salidas (Nota: es posible que en algunos casos haya varias soluciones que tengan el mismo valor para la función objetivo)

### 3.2. Sobre el Informe...

El grupo deberá entregar un informe del proyecto, en formato pdf, que contenga, al menos, los siguientes aspectos:

- El modelo: una descripción del modelo y una justificación de su adecuación al problema planteado.
- Detalles importantes de implementación: lo más relevante de la implementación, sin incluir código.
- El análisis de los árboles generados por su modelo para el ejemplo, y explicar sobre él cómo funcionó el mecanismo de *Branch and Bound*.
- Pruebas: descripción de las pruebas realizadas a su implementación.
- Análisis: de los resultados de las pruebas realizadas, buscando responder a los diferentes criterios de evaluación definidos en la rúbrica. Desarrolle y soporte su análisis utilizando los métodos apropiados (tablas, gráficos, indicadores estadísticos), donde puedan apreciarse las variaciones de acuerdo al tamaño y naturaleza de los datos de entrada. Explique claramente el significado de sus datos y cómo se analizaron.
- Un enlace al video explicatorio de la interfaz gráfica
- Conclusiones: Esta es una de las partes más interesantes del trabajo (pero no por ello la que más vale). En ella se espera que usted analice los resultados obtenidos y **justifique** claramente sus afirmaciones.

### 3.3. Grupos de trabajo

El proyecto puede ser desarrollado por grupos de máximo 4 personas.

## 4. Entrega, sustentación y evaluación

### 4.1. Entrega

La entrega se debe realizar vía el campus virtual en las fechas previstas para ello, por uno sólo de los integrantes del grupo. **La fecha de entrega límite es el 12 de noviembre de 2024 a las 23:59. La sustentación será en las sesiones del 28 de noviembre y 3 de diciembre (o en otras si hay un cambio de última hora).** Se debe subir al campus virtual en el enlace correspondiente a este proyecto un archivo comprimido **.zip** que siga la convención *CódigodeEstudiante1-CódigodeEstudiante2-CódigodeEstudiante3-CódigodeEstudiante4-Proyecto2-AdaII.2024-I.zip*. El comprimido deberá contener:

1. Archivo **Readme.txt** que describa todos los archivos entregados y las instrucciones para ejecutar la aplicación.
2. Archivo **Informe.pdf** acorde a la Sección 3.2. Recuerde incluir el link al video explicatorio.
3. Archivo **Proyecto.mzn** con la implementación del modelo
4. Directorio **DatosProyecto** con los datos con que fue probado su modelo.
5. Directorio **ProyectoGUIFuentes** con los archivos fuente de la implementación de la interfaz gráfica
6. Directorio **MisInstancias** con las 5 instancias generadas por su equipo de trabajo para retar a otros proyectos que resuelvan el mismo problema.

## 4.2. Evaluación

La evaluación de cada proyecto se hará de acuerdo a la rúbrica publicada en el campus virtual, diseñada para observar los indicadores de logro asociados a este proyecto.

## 4.3. Sustentación

El trabajo debe ser sustentado por los autores en día y hora señalados arriba. En todos los casos la sustentación será pilar fundamental de la nota individual asignada a cada integrante del grupo. En la sustentación se demuestra la capacidad del grupo de navegar en el código y realizar cambios rápidamente en él, así como la capacidad de responder con solvencia a las preguntas que se le realicen.

Cada persona de cada grupo, después de la sustentación, tendrá asignado un número real (el factor de multiplicación) entre 0 y 1, correspondiente al grado de calidad de su sustentación. **Su nota definitiva será la calificación del proyecto para el grupo, multiplicada por ese valor.** Si su asignación es 1, su nota será la del proyecto. Pero si su asignación es 0.9, su nota será 0.9 por la nota del proyecto. La no asistencia a la sustentación tendrá como resultado una asignación de un factor de 0.

La idea es que lo que no sea debidamente sustentado no vale así funcione muy bien!!! Y que, del trabajo en grupo, es importante que todos aprendan, no sólo algunos.

Éxitos!!!