

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**ALGORITMO CULTURAL PARA EL PROBLEMA DE
LOCALIZACIÓN DE CLUSTERS DE ESTACIONES DE
SERVICIO DE BICICLETAS**

**BENJAMÍN EDGARDO APPEL GONZÁLEZ
CONIE ALEXANDRA UTRERAS BUSTOS**

INFORME FINAL PROYECTO DE TÍTULO
PARA OPTAR AL TÍTULO PROFESIONAL DE
INGENIERO CIVIL EN INFORMÁTICA

Abril 2018

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**ALGORITMO CULTURAL PARA EL PROBLEMA DE
LOCALIZACIÓN DE CLUSTERS DE ESTACIONES DE
SERVICIO DE BICICLETAS**

**BENJAMÍN EDGARDO APPEL GONZÁLEZ
CONIE ALEXANDRA UTRERAS BUSTOS**

Profesor Guía: **Guillermo Cabrera Guerrero**
Profesor Co-referente: **Pablo Maya Duque**

Abril 2018

Resumen

Los sistemas de bicicletas públicas ponen a disposición una serie de bicicletas a un grupo de usuarios, para que sean utilizadas temporalmente como medio de transporte. Comúnmente estos sistemas permiten recoger una bicicleta de una estación y devolverla en un punto diferente, de manera que el usuario tenga ésta en su posesión sólo durante el tiempo que demore en desplazarse de un punto a otro. Debido a la alta demanda de bicicletas, es necesario mantener al sistema balanceado a través de la localización de clusters o grupos de estaciones de bicicletas. Es por esto que se propone en el presente documento, la implementación de un algoritmo cultural para la resolución de este problema. Los algoritmos culturales son una rama de la computación evolutiva, que consisten en el material genético heredado por los padres a sus descendientes, donde el conocimiento adquirido por los individuos a través de las generaciones una vez codificado y almacenado, sirve para guiar el comportamiento de los individuos que pertenecen a una población. La cultura se presenta como un conjunto de fenómenos ideológicos que son compartidos por una población, donde un individuo puede interpretar sus experiencias y decidir su comportamiento.

Palabras Clave: Algoritmo Cultural, BOCA, Algoritmo Evolutivo, Matheurística, Optimización.

Abstract

Public bicycle systems give the users a number of bicycles, to be used temporarily as means of transportation. Usually, these systems allow for a bicycle to be picked up from a station and returned at a different location, so the user has it in possession only for the time elapsed to reach the destination. Due to the high demand for bicycles, it is necessary to keep the system under balance through the localization of clusters of terminals. This is why a cultural algorithm is proposed in this document to solve the problem. Cultural algorithms are a branch of evolutionary computing, consisting in genetic material being passed on from individuals to their descendants, in which the acquired knowledge throughout generations once coded and stored, is used to guide the behaviour of the population individuals. Culture is represented as a collection of ideological phenomena being shared by some population, where an individual can interpret its experiences and decide its behaviour.

Keywords: Cultural Algorithm, BOCA, Evolutionary Algorithm, Matheuristics, Optimization

Índice

1	Introducción	1
2	Descripción del Problema	2
3	Objetivos	3
3.1	Objetivo General	3
3.2	Objetivos Específicos	3
4	Marco Teórico	4
4.1	Estado del Arte	4
4.2	Algoritmos Culturales	5
4.2.1	Pseudocódigo	6
4.2.2	Tipos de Conocimiento	6
4.2.2.1	Conocimiento Normativo	6
4.2.2.2	Conocimiento Situacional	6
4.2.2.3	Conocimiento Temporal o Histórico	6
4.2.2.4	Conocimiento Topográfico	7
4.2.2.5	Conocimiento del Dominio	7
4.3	Algoritmos Exactos	7
4.3.1	Solver	7
5	Modelo Matemático	8
5.1	Conjuntos	8
5.2	Parámetros	8
5.3	Cálculo Conexión	9
5.4	Variables de Decisión	9
5.5	Función Objetivo	9
5.6	Restricciones	9
5.7	Restricciones de Dominio	10
6	Diseño e Implementación	11
6.1	Individuos	11
6.2	Población Inicial	11
6.3	Nueva generación	11
6.4	Crossover	12
6.5	Mutación	12
6.6	Espacio de creencias: Conocimiento Topográfico	12
6.7	Adyacencia	12
7	Resultados	13
7.1	Parámetros	13
7.2	Topologías	17
7.2.1	Topología 1	17
7.2.2	Topología 2	18
7.3	Resultados comparativos	20
8	Conclusión	21

Lista de Figuras

Figura 4.1	Espacios de un algoritmo cultural	5
Figura 6.1	Representación de Individuos	11
Figura 7.1	CrossRate VS Elitismo	13
Figura 7.2	CrossRate VS MutRate: Mejores resultados	15
Figura 7.3	CrossRate VS MutRate: Promedio mejores resultados	15
Figura 7.4	Topología 1	17
Figura 7.5	Topología 2	18

Lista de Tablas

Tabla 7.1	Tasa de Crossover VS Elitismo	13
Tabla 7.2	Tasa de Crossover VS Tasa de Mutación	14
Tabla 7.3	Mejores Resultados VS Tamaño de Población	16
Tabla 7.4	Resultados Topología 1	17
Tabla 7.5	Resultados Topología 2	19
Tabla 7.6	Resultados Algoritmo Exacto	20

1 Introducción

Durante las últimas décadas, el número de vehículos basados en combustible fósil ha aumentado considerablemente y con ello, la emisión de gases que contribuyen al calentamiento global. Es por esto que cada vez un mayor número de personas han decidido optar por modos de transporte sostenibles, siendo la bicicleta una de las mejores opciones. Dentro de las alternativas para promover el uso de este medio de transporte, se encuentran los Sistemas de Bicicletas Públicas (SBP) [1], los cuales ponen bicicletas a disposición de los usuarios, las que pueden ser retiradas, utilizadas por un periodo de tiempo y devueltas en un punto diferente.

Sin embargo, estos sistemas requieren que cada estación de servicio tenga disponibilidad de bicicletas en todo momento, lo que provoca que la distribución de estas en cada estación se vuelva un problema que se requiere modelar, involucrando distintas variables, como por ejemplo la cantidad de bicicletas que se encuentran en cada momento en una estación o los tiempos de alta y baja demanda, haciendo del modelado un tema complejo de abordar. Se pretende en el presente trabajo, utilizar una perspectiva táctica, la que será adoptada para resolver la problemática de localización de grupos de estaciones de bicicletas, de manera tal que se pueda llegar una solución aproximada basada en matheurística, la cual corresponde a algoritmos de optimización realizados por la interoperación de metaheurísticas y programación matemática. El presente proyecto se enfocará en modelar y resolver, específicamente el caso del sistema implementado en la ciudad de México, a través de la creación de clusters o conjuntos de estaciones, y la elección de un centro para dichos clusters.

Para esto, se diseñará e implementará un algoritmo cultural [17]. Dicho algoritmo expande la metáfora de los algoritmos genéticos de simular el proceso de evolución de individuos, agregando la abstracción del espacio social del ser humano y el conocimiento adquirido a través de varias generaciones, por medio del cual los mejores individuos seleccionados por la función de aptitud, no sólo influyen a la generación siguiente, sino que pueden influenciar las poblaciones futuras, a través de los denominados “espacios de creencias”. Además, con el objetivo de mejorar el rendimiento del algoritmo, se buscará agregar a la implementación el uso de un algoritmo exacto para la resolución de parte del modelo.

En este trabajo de título se expone la descripción del problema, a continuación los objetivos de este, tanto general como específicos. Luego se encuentra el marco teórico y se presenta una breve reseña del estado del arte, para seguir con el modelo matemático preliminar a utilizar en la solución de la problemática. Todo esto, como base necesaria para explicar el diseño utilizado para la implementación del algoritmo y utilización del modelo. Finalmente se presentan los resultados obtenidos a través de los distintos experimentos realizados y conclusiones al respecto.

2 Descripción del Problema

Los sistemas de bicicletas públicas, corresponden a sistemas de alquiler o préstamo de bicicletas en núcleos urbanos, donde se presta un servicio de movilidad práctico, rápido y pensado para el uso cotidiano. Estos sistemas están compuestos básicamente de estaciones y clusters. Siendo una estación, un punto definido en la ciudad donde se encuentran disponibles las bicicletas para su uso. Cada estación posee varios puntos de anclaje, estos puntos corresponden al espacio donde se retira o aparca una bicicleta. Mientras que un cluster corresponde a un conjunto de estaciones.

El problema de balanceo de bicicletas públicas se debe esencialmente a la necesidad de mantener el equilibrio entre la demanda y la oferta. Siendo la demanda la cantidad de veces que se requiere una bicicleta, y la oferta, la cantidad de estaciones y bicicletas que son puestas en servicio. Si la demanda es demasiado alta y la oferta es muy baja, se pueden llegar a producir desajustes en el sistema.

Si llega a existir una sobre-demanda en relación al número de bicicletas ofrecidas, hay un aumento en la circulación de éstas entre las estaciones, lo que podría provocar que algunas de estas se encuentren saturadas, llevando a los usuarios a tener que esperar en aquella estación por un punto de anclaje o dirigirse a otra para poder dejar la bicicleta. Caso contrario es que las estaciones queden desabastecidas, haciendo que el cliente deba caminar a la estación más cercana, esperando a que ésta tenga bicicletas disponibles. Ambos casos provocan insatisfacción de las personas y deteriora la imagen del sistema de transporte, pudiendo afectar su uso.

Se pretende, en el presente proyecto, resolver el problema de desequilibrio en el balanceo de las estaciones, a través de la creación de clusters, o grupos de estaciones. Estos son conformados por una estación centro y un grupo de estaciones asignadas a ese centro. Se deben tomar en cuenta restricciones de balanceo de estaciones, relacionadas a los requerimientos de bicicletas y los requerimientos de anclaje, es decir, cuántas bicicletas son sacadas de la estación y cuántas bicicletas son requeridas en ella, respectivamente. Estos grupos cumplen la función de centros de reposicionamiento, dentro de los cuales las bicicletas son trasladadas para mantener el balance requerido.

Además se deben considerar las prioridades de cada estación, donde cada cluster debe quedar conformado lo más homogéneo posible con estaciones de distintas prioridades. También se debe tener en cuenta que es necesario mantener un equilibrio entre el número de estaciones asignadas a cada uno de los clusters, para que no exista una sobrecarga en ellos. Finalmente el objetivo del problema mencionado y por el cual se está realizando este trabajo, es conformar los clusters, de manera tal que se minimice la distancia entre las estaciones centro de cluster y las estaciones pertenecientes a cada centro.

3 Objetivos

3.1 Objetivo General

Resolver el modelo matemático del problema de localización de clusters de estaciones de servicio de bicicletas públicas para la ciudad de México.

3.2 Objetivos Específicos

- Plantear un diseño de algoritmo cultural.
- Implementar el algoritmo cultural, utilizando herramientas open source, empleando el lenguaje de programación Julia.
- Complementar la implementación del algoritmo cultural, a través de la combinación de éste con algoritmos matemáticos exactos.
- Experimentar y evaluar los resultados para la optimización de la solución.

4 Marco Teórico

4.1 Estado del Arte

Para la implementación de un Sistema de Bicicletas Públicas es necesario tener en consideración los factores influyentes. En este apartado, se destacarán algunos trabajos realizados para resolver problemáticas basadas en factores endógenos, los cuales se relacionan con las características específicas del sistema. Según se menciona en [14], se pueden clasificar estas problemáticas en tres categorías: problemas de diseño, de gestión de recursos y de gestión de la demanda.

Para las problemáticas de diseño, se encuentran la asignación de recursos y planeación de localización, además de la interacción que existe con la infraestructura del transporte e intervenciones dirigidas al nivel de servicio, la capacidad, número de bicicletas que se deben mantener en cada una de las estaciones, puntos de anclaje y número óptimo a tener de bicicletas [14]. En cuanto a la solución de esta problemática algunos han optado por resolverlas enfocándose en el inventario de bicicletas que debe poseer cada estación [13]. También se ha resuelto esta problemática a través de determinaciones relacionadas con el nivel de servicio, tanto de manera estratégica [12] como a nivel táctico [20]. Por otro lado, existe un estudio en donde se determinan los requisitos de nivel de servicio en cada estación de uso compartido de bicicletas y diseño de rutas de vehículos casi óptimas para reequilibrar el inventario.[21]

Mientras que para el problema de gestión de la demanda, el sistema debe reequilibrarse adecuadamente para satisfacer la demanda de usuarios y para operar con éxito. Se ha utilizado para la solución de esta problemática la minería de datos, ya sea para entregar una clasificación de los sistemas públicos de bicicletas[15] o para identificar en las estaciones ciertos patrones de actividad[24]. Otros estudios se han dedicado a predecir la demanda potencial[23] y otros a identificar lugares en donde mejorar la infraestructura signifique captar más usuarios[11]. También se han recurrido a técnicas de series de tiempo para hacer una estimación de la demanda en las estaciones[22].

Finalmente en la gestión de recursos, se identificaron problemáticas en el mantenimiento de estaciones y bicicletas, además del reposicionamiento de estas teniendo en cuenta un balance en el sistema [14]. Se ha intentado resolver el problema de reposicionamiento mediante modelos matemáticos [5], a través de metaheurísticas [6], simulación[3] y programación por restricciones[9], para tratar de minimizar tiempo, distancias y demanda insatisfecha, siendo esta última obtenida a través de funciones lineales[2] y de funciones convexas posteriormente linealizadas[16].

Todas las problemáticas aquí presentadas, como se ha mencionado en un principio, están enfocadas en resolver problemas del Sistema de Bicicletas Públicas asociado a los factores endógenos de este. Sin embargo, las soluciones han sido aplicadas, ya sea mediante heurísticas o algoritmos exactos. Lo que se pretende en el presente trabajo, es hacer una combinación de ambas técnicas, aplicándolas al problema presentado para resolver el desequilibrio en el balanceo de las estaciones del SPB, a través de la creación de clusters, o grupos de estaciones.

4.2 Algoritmos Culturales

En las sociedades humanas, la cultura es un medio de almacenamiento del conocimiento de manera independiente de los individuos que la generan, y es potencialmente accesible a todos los miembros de la sociedad. Como tal, la cultura es capaz de guiar las actividades de resolución de problemas e interacción social de los individuos de la población. Esto permite que información autoadaptativa, a la vez que otros tipos de conocimiento, sea almacenada y manipulada en un espacio separado de los individuos de la población social. Por lo tanto, los sistemas culturales comprenden un sistema de herencia doble donde, en cada paso de tiempo, el conocimiento tanto a nivel de la población como en el nivel de las creencias adquiridas sea transmitido a la siguiente generación [18]. El objetivo es incrementar las tasas de aprendizaje o convergencia, y de esta manera, que el sistema responda mejor a un gran número de problemas[8]

Los algoritmos culturales fueron desarrollados inicialmente por Robert Reynolds, como una extensión de los algoritmos evolutivos. Su funcionamiento, representado en la figura 4.1, es explicado en el siguiente modelo. El algoritmo opera en dos espacios, el espacio de la población y el espacio de creencias. En el primer espacio existe un número de individuos, donde cada uno tiene un conjunto de características independientes de los otros, cada individuo es evaluado según la función de aptitud, como en cualquier algoritmo evolutivo. A través del tiempo, dichos individuos pueden ser reemplazados por algunos de sus descendientes. El segundo espacio, es donde se almacenan los conocimientos adquiridos por los individuos a través de las generaciones. La información que se encuentra en este espacio, debe estar disponible a cualquier individuo. El espacio de creencias se actualiza después de cada iteración por los mejores individuos de la población, las categorías de conocimiento del espacio de creencias pueden afectar el componente de población a través de la influencia, la cual puede afectar a la población mediante la alteración del genoma o las acciones de los individuos. Para unir ambos espacios se establece un protocolo de comunicación que dicta las reglas del tipo de información que se debe intercambiar entre los espacios.

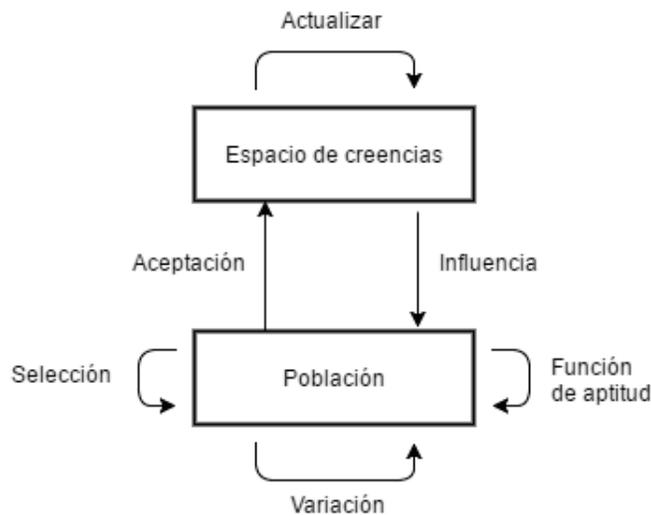


Figura 4.1: Espacios de un algoritmo cultural

4.2.1 Pseudocódigo

El pseudocódigo utilizado para algoritmos culturales es el siguiente:

Algorithm 1 Cultural Algorithm

- 1: *Inicializar el espacio de la población*
 - 2: *Inicializar espacio de creencias*
 - 3: *Repetir hasta cumplir condición de término*
 - a. *Realizar las acciones de los individuos en el espacio población*
 - b. *Evaluar cada individuo mediante el uso de la función de aptitud*
 - c. *Seleccionar los padres para reproducir una nueva generación de descendientes*
 - d. *Dejar que el espacio de creencias altere el genoma de la descendencia mediante el uso de la función de influencia*
 - e. *Actualizar el espacio de creencias mediante el uso de la función de aceptación, dejando los mejores individuos para afectar el espacio de creencias.*
-

4.2.2 Tipos de Conocimiento

Se han identificado cinco categorías básicas de conocimiento cultural: conocimiento normativo, conocimiento situacional, conocimiento temporal o histórico, conocimiento topográfico y conocimiento del dominio.[19]

4.2.2.1 Conocimiento Normativo

Es un conjunto de rangos variables prometedores, que proporcionan estándares para el comportamiento individual y las guías dentro de las cuales se pueden hacer ajustes individuales [4]. Este conocimiento lleva a los individuos de la población a "saltar al buen rango" si es que no están ya allí [19].

4.2.2.2 Conocimiento Situacional

Corresponde al conocimiento de los eventos importantes en el espacio de decisión, proporcionando un conjunto de casos ejemplares que son útiles para la interpretación de experiencias individuales específicas en contextos específicos. Este conocimiento lleva a los individuos a "moverse hacia los ejemplares-[19].

4.2.2.3 Conocimiento Temporal o Histórico

Se encarga de controlar el proceso de búsqueda y registrar eventos importantes en ella. Estos eventos importantes pueden ser un movimiento significativo en el espacio de búsqueda, o una detección de cambio de paisaje. Los individuos guiados bajo este conocimiento, pueden consultar aquellos eventos registrados para guiarlos en la selección de una dirección de movimiento [19].

4.2.2.4 Conocimiento Topográfico

Este conocimiento fue propuesto originalmente para razonar sobre patrones de paisajes funcionales basados en una región[25]. El paisaje funcional se divide inicialmente en celdas en un cierto nivel de granularidad, de acuerdo a las características espaciales y cada celda mantiene un seguimiento del mejor individuo encontrado hasta ahora en su región. Las celdas se pueden subdividir recursivamente en sub-celdas bajo ciertas circunstancias. Este conocimiento guía a los individuos hacia las celdas con mejor desempeño en el espacio de búsqueda [19].

4.2.2.5 Conocimiento del Dominio

Utiliza el conocimiento sobre el dominio del problema para guiar la búsqueda [19]. El conocimiento sobre las variables, parámetros, restricciones, etc., pueden ser durante el proceso de búsqueda.

4.3 Algoritmos Exactos

Dentro de los algoritmos de optimización, se les denomina algoritmos exactos a aquellos que, a través de métodos exhaustivos, logran siempre llegar a la solución óptima del problema. Sin embargo, debido a que su funcionamiento es en base a una búsqueda exhaustiva a través de todas las posibles soluciones, estos algoritmos poseen el inconveniente de ser lentos para llegar a tal solución, lo que los hace inviables para problemas de alta complejidad [7]. Por esta razón es que se suele preferir los algoritmos heurísticos y metaheurísticos por sobre los algoritmos exactos.

Sin embargo, los algoritmos exactos pueden ser utilizados en conjunto con una metaheurística para mejorar el rendimiento de ésta, a través de la resolución de pequeños sub-problemas dentro del modelo general. A través de esta técnica, se logra llegar a soluciones óptimas dentro de tales sub-problemas, mejorando la solución general, sin sacrificar demasiado el desempeño del algoritmo.

4.3.1 Solver

Un *solver* es un programa matemático utilizado para la resolver problemas matemáticos. Como tal, los algoritmos usados por estos programas caen dentro de la categoría de algoritmos exactos, ya que encuentran la solución óptima a un problema matemático. En este proyecto se utilizó *Gurobi*, un *solver* libre para uso académico.

5 Modelo Matemático

Encontrar la manera óptima de localizar los clusters para balancear estaciones de bicicletas públicas es el principal objetivo del problema, en este caso utilizando datos de prueba de las estaciones de Ciudad de México. Para esto se presenta el modelo matemático preliminar de la solución, en el que se exponen los conjuntos y parámetros a utilizar, además del cálculo de la conexión entre dos estaciones, determinada por un rango máximo de distancia en la cual dos estaciones se pueden considerar unidas. Luego se nombran las variables de decisión, las cuales son incógnitas que deben ser determinadas a partir de la solución del modelo, a continuación se encuentra la función objetivo, que corresponde a la ecuación que será optimizada dadas las restricciones señaladas al final del apartado.

5.1 Conjuntos

El modelo considera tres conjuntos, primero el conjunto “E”, corresponde a las estaciones que deben ser asignadas a una zona, luego el conjunto “C”, que son las estaciones entre las cuales se escogerán los centros de los clusters, y por último el conjunto “P”, el cual indica la prioridad de reposicionamiento de bicicletas para cada estación.

- Estaciones “E”: $\{1, 2, 3, 4, \dots, I\}$
- Estaciones Candidatas “C”: $\{1, 2, 3, 4, \dots, J\}$
- Prioridades “P”: $\{1, 2, \dots, L\}$

5.2 Parámetros

A continuación se presentan los parámetros a considerar:

d_{ij} : distancia entre la estación $i \in E$ y la estación $j \in C$

I_{ij}^d : indicador de cobertura entre la estación $i \in E$ y la estación $j \in C$

I_{ij}^a : indicador de accesibilidad entre la estación $i \in E$ y la estación $j \in C$

d_{max} : radio de cobertura

c_{ij} : si existe conexión entre la estación $i \in E$ y la estación $j \in C$

r_i^+ : requerimiento de bicicletas en la estación $i \in E$

r_i^- : requerimiento de puntos de anclaje en la estación $i \in E$

p_{il} : si la estación $i \in E$ tiene la prioridad de reposicionamiento $l \in P$

cl : número de clusters a conformar

balance: Porcentaje de desbalance neto permitido en los requerimientos de cada cluster

prioridad: desbalance permitido en número de estaciones entre clúster de acuerdo al tipo de prioridad de las estaciones

5.3 Cálculo Conexión

En seguida se ilustra el cálculo de la conexión entre estaciones. Para indicar si dos estaciones están conectadas, se debe determinar si la distancia entre la estación i y la estación candidata j es inferior al radio de cobertura, de ser así se determina que hay conexión entre las estaciones, en caso contrario, no habrá conexión entre ambas.

```

 $\forall(i \in E, j \in C)$ 
if ( $d_{ij} < d_{max}$ )
     $I_{ij}^d = 1$ 
     $c_{ij} = I_{ij}^d * I_{ij}^a$ 
else
     $c_{ij} = 0$ 
end if

```

5.4 Variables de Decisión

$$x_{ij} = \begin{cases} 1 & \text{Si la estación } i \in E \text{ pertenece al cluster con centro en la estación } j \in C \\ 0 & \text{En caso contrario} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{Si la estación candidata } j \in C \text{ es el centro de un cluster} \\ 0 & \text{En caso contrario} \end{cases}$$

5.5 Función Objetivo

La función objetivo (1) del modelo es minimizar la distancia entre las estaciones centro de cluster y las estaciones incluidas dentro del cluster correspondiente.

$$\text{Min} \sum_{i \in E} \sum_{j \in C} d_{ij} x_{ij} \quad (1)$$

5.6 Restricciones

La función objetivo está sujeta a cinco restricciones, las cuales son descritas a continuación.

- Toda estación debe quedar asignada exactamente una vez.

$$\sum_{j \in C} x_{ij} = 1, \quad \forall i \in E \quad (2)$$

- Una estación solo se puede asignar a un cluster abierto y que exista conexión entre la estación y la estación elegida como centro de un cluster.

$$x_{ij} \leq c_{ij}y_j, \quad \forall i \in E, j \in C \quad (3)$$

- Hay tantas estaciones elegidas como centro de cluster como número de cluster a conformar.

$$\sum_{j \in C} y_j = cl \quad (4)$$

- El desbalance (en porcentaje o proporción) en los requerimientos netos de cada cluster debe ser inferior o igual a un valor fijado

$$\frac{|\sum_{i \in E} r_i^+ x_{ij} - \sum_{i \in E} r_i^- x_{ij}|}{\sum_{i \in E} r_i^+ x_{ij} + \sum_{i \in E} r_i^- x_{ij}} \leq balance, \quad \forall j \in C \quad (5)$$

Linealizando,

$$\frac{\sum_{i \in E} r_i^+ x_{ij} - \sum_{i \in E} r_i^- x_{ij}}{\sum_{i \in E} r_i^+ x_{ij} + \sum_{i \in E} r_i^- x_{ij}} \leq balance, \quad \forall j \in C \quad (6)$$

$$\frac{-\sum_{i \in E} r_i^+ x_{ij} + \sum_{i \in E} r_i^- x_{ij}}{\sum_{i \in E} r_i^+ x_{ij} + \sum_{i \in E} r_i^- x_{ij}} \leq balance, \quad \forall j \in C \quad (7)$$

- Mantener el equilibrio entre el número de estaciones asignadas a cada uno de los clusters frente a un valor teórico de referencia que debería tener cada cluster.

$$|\sum_{i \in E} p_{il} x_{ij} - \lfloor \frac{\sum_{i \in E} p_{il}}{cl} \rfloor y_j| \leq prioridad, \quad \forall j \in C, l \in P \quad (8)$$

Linealizando,

$$\sum_{i \in E} p_{il} x_{ij} - \lfloor \frac{\sum_{i \in E} p_{il}}{cl} \rfloor y_j \leq prioridad, \quad \forall j \in C, l \in P \quad (9)$$

$$\lfloor \frac{\sum_{i \in E} p_{il}}{cl} \rfloor y_j - \sum_{i \in E} p_{il} x_{ij} \leq prioridad, \quad \forall j \in C, l \in P \quad (10)$$

5.7 Restricciones de Dominio

Las variables están sujetas a las siguientes restricciones de dominio.

$$x_{ij} \in \{0, 1\}$$

$$y_j \in \{0, 1\}$$

6 Diseño e Implementación

A continuación se presenta el diseño e implementación de la solución al modelo expuesto anteriormente. La implementación del algoritmo fue realizada en el lenguaje de programación Julia versión 0.5.2. en sistema operativo GNU/LINUX.

6.1 Individuos

Como punto de partida, se presenta el modelo de individuo a utilizar para generar las poblaciones, el cual constará de dos vectores de números enteros. El primer vector X , de largo cl (la cantidad de clusters a conformar), representará las estaciones candidatas seleccionadas para el rol de centro de cluster. El segundo vector Y , de largo I (la cantidad de estaciones del problema), contendrá en cada la posición del vector X donde está almacenada la estación centro de cluster, de manera que para cada estación i , si $X_i = n$, Y_n representará el cluster al cual dicha estación pertenece. Cabe mencionar que según esta manera de expresar los individuos, para cada estación centro de cluster $Y_j = m$, $X_m = j$.

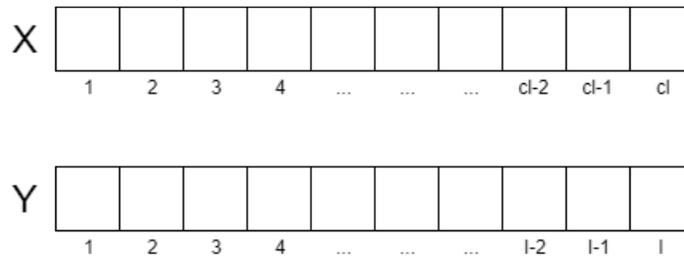


Figura 6.1: Representación de Individuos

6.2 Población Inicial

Para la generación de la población inicial, cada individuo será creado de la siguiente manera:

(1) Se llena el vector X , seleccionando aleatoriamente una estación de cada cuadrante.

(2) Para el vector Y , se utiliza un solver, en este caso Gurobi, para encontrar el óptimo para el vector X dado.

6.3 Nueva generación

Para la creación de cada nueva generación, se utilizan tres parámetros: elitismo, probabilidad de crossover, y probabilidad de mutación. En primera instancia, se seleccionarán los mejores individuos de cada generación, en una cantidad especificada por el elitismo, para que formen parte de la nueva generación sin ser modificados. Luego, el resto de los individuos serán creados por crossover o por copia. Finalmente, se aplica la mutación a cada nuevo individuo.

6.4 Crossover

El proceso de *crossover* se detalla a continuación:

(1) Se seleccionan los individuos padres a través de un torneo, es decir, se selecciona un número predefinido de individuos, los cuales a través de una llave de torneo van progresando si tienen mejor *fitness* que su contrincante, hasta solamente quedar dos individuos.

(2) Se selecciona un punto de cruce para la mezcla del vector X de los individuos seleccionados, generando dos hijos que serán agregados a la población después de terminar todo el proceso, usando el solver para generar el vector Y .

6.5 Mutación

El proceso de mutación se llevará a cabo como se expone a continuación:

(1) Para cada nuevo individuo, se recorre su vector de centros, eligiendo los que van a ser mutados según el parámetro probabilidad de mutación.

(2) Si el centro es seleccionado para ser mutado, se selecciona un nuevo centro al azar, dentro de los centros del cuadrante al que el centro original pertenece.

(3) Para el nuevo vector X , se genera el Y óptimo usando el solver.

6.6 Espacio de creencias: Conocimiento Topográfico

Para mejorar el rendimiento del algoritmo, se ha decidido utilizar el conocimiento topográfico, el cual fue considerado adecuado debido al componente geográfico del problema real. Para esto, se han dividido las estaciones en tantas celdas como clusters deba tener la solución, utilizando la ubicación geográfica de cada una. Existiendo múltiples opciones de cómo hacer tal división, se buscarán las que obtengan los mejores resultados.

6.7 Adyacencia

Con el fin de mejorar el rendimiento del algoritmo, se decidió implementar en conjunto con el conocimiento topográfico el concepto de adyacencia. Con este concepto, se reduce el número de estaciones candidatas a formar parte de un cluster, únicamente a aquellas que se encuentren en el cuadrante donde se halla la estación centro, o los cuadrantes directamente adyacentes a éste.

7 Resultados

En este apartado se presentan los resultados obtenidos mediante una serie de pruebas realizadas, con el fin de presentar un análisis de estos. En primera instancia se ilustran tablas con las pruebas efectuadas con el fin de decidir los parámetros más óptimos de elitismo, tasa de crossover y tasa de mutación que serán utilizadas para el resto de las pruebas. Esto ha sido acotado debido a la inmensa cantidad de pruebas que podrían ser realizadas, que por efectos de tiempo no fueron posibles de realizar. Luego se presentan las topologías utilizadas y los resultados obtenidos para cada una de ellas.

7.1 Parámetros

Las siguientes pruebas fueron realizadas con un tamaño de población de 100 individuos, para 50 generaciones, con un balance de 0,7 y prioridad 15. La tabla 7.3 denota los resultados obtenidos utilizando los parámetros mencionados y con una tasa de mutación de 0,075. La figura 7.1 muestra una comparación entre los parámetros de elitismo con la tasa de crossover. A través de esto se ha concluido que un elitismo de 5 es apropiado para realizar las pruebas.

Tabla 7.1: Tasa de Crossover VS Elitismo

Elitismo	Tasa de Crossover	Mejor Resultado	Promedio Mejores
1	0.8	277569	282625
3	0.8	281454	282718
5	0.8	273484	277455
1	0.9	278637	280518
3	0.9	277189	281511
5	0.9	275562	275862
1	0.7	280003	281423
3	0.7	278718	281341
5	0.7	276663	277609

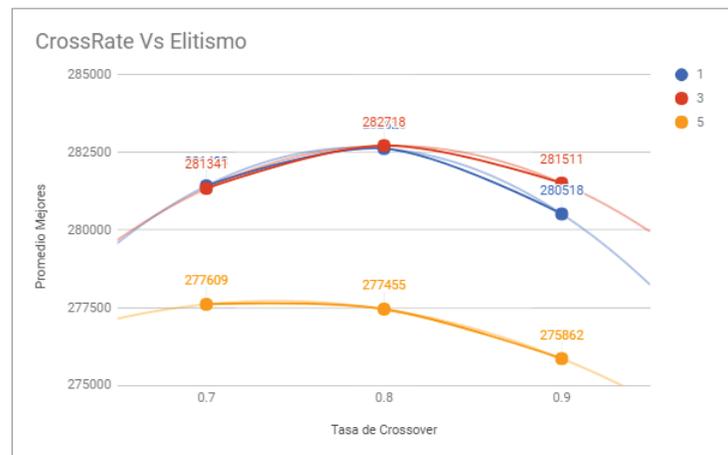


Figura 7.1: CrossRate VS Elitismo

La tabla 7.2 hace referencia a los resultados obtenidos para los parámetros anteriormente mencionados, esta vez variando su tasa de crossover y mutación, logrando con ello, hacer una comparación gráfica de la tabla en las figuras 7.2 y 7.3 y concluir que los parámetros de tasa de crossover y mutación a utilizar para las pruebas serán de 0,90 y 0,075 respectivamente.

Tabla 7.2: Tasa de Crossover VS Tasa de Mutación

Tasa de Crossover	Tasa de Mutación	Mejor Resultado	Promedio Mejores
0.75	0.025	273886	281184
0.80	0.025	276694	280342
0.85	0.025	274520	279081
0.90	0.025	275069	279696
0.95	0.025	277553	281032
0.75	0.050	277533	280513
0.80	0.050	276368	279569
0.85	0.050	277816	280292
0.90	0.050	275495	276654
0.95	0.050	275779	278487
0.75	0.075	274835	280592
0.80	0.075	274699	280806
0.85	0.075	270731	276724
0.90	0.075	274666	278510
0.95	0.075	274003	276666
0.75	0.100	275288	279915
0.80	0.100	276187	278175
0.85	0.100	274874	277732
0.90	0.100	274414	278538
0.95	0.100	272415	276238
0.75	0.125	278536	281670
0.80	0.125	274519	280574
0.85	0.125	276412	280057
0.90	0.125	275089	278466
0.95	0.125	277812	280383

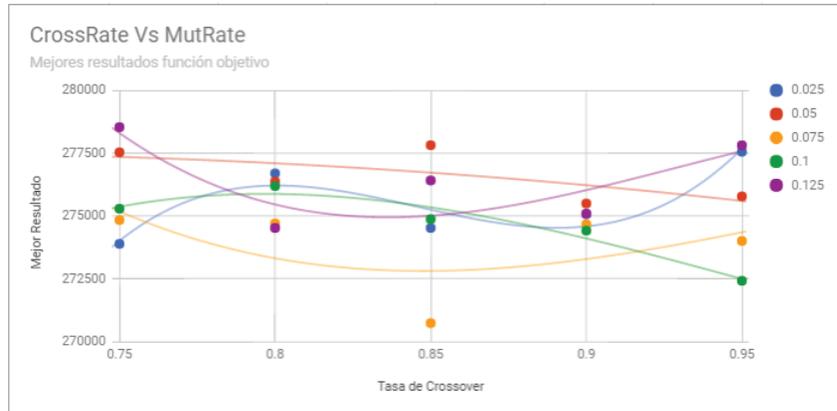


Figura 7.2: CrossRate VS MutRate: Mejores resultados

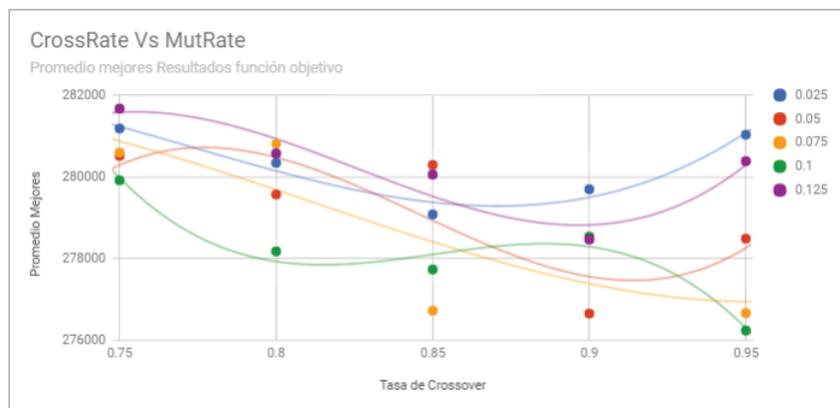


Figura 7.3: CrossRate VS MutRate: Promedio mejores resultados

En cuanto al tamaño de la población se realizaron distintas pruebas variando su tamaño, la tabla 7.3 muestra los mejores resultados obtenidos para 50 generaciones, en donde se observa que existe una mejora en los resultados para un tamaño de población desde 180 individuos. También se observa que aproximadamente desde la generación 40 no existen muchas variaciones en cuanto a resultados. Por lo que en vista de lo anterior, se ha decidido reducir las generaciones a 40 y aumentar la población a 180 individuos para efecto de las pruebas siguientes.

Tabla 7.3: Mejores Resultados VS Tamaño de Población

Generación	Tamaño de Población					
	50	100	140	180	220	260
1	320455.6	315784.2	311035.8	311677.4	312079.4	308149.6
2	312243.8	313730.0	309286.0	308956.2	311073.4	305719.8
3	311202.2	310826.8	307938.4	307151.4	307877.2	301794.0
4	308767.2	309851.8	306102.8	306512.0	304687.0	300632.4
5	308348.0	307117.8	302694.0	306004.4	303097.0	297940.4
6	305430.0	303411.8	301396.4	303439.6	301968.4	296965.4
7	304235.4	300234.4	300866.2	302309.6	300686.6	295819.0
8	302382.0	299839.8	298633.6	300867.0	299297.2	293914.8
9	299796.2	298049.0	297941.0	299919.4	295585.6	293532.4
10	298669.6	297905.8	296172.4	296339.0	294650.0	293532.4
11	298227.0	297232.0	295516.8	296339.0	293078.2	292554.0
12	297402.2	296367.2	295159.4	295777.0	292504.2	291772.2
13	295643.6	296174.0	294453.4	295552.6	291205.2	290765.2
14	295208.2	293531.8	293488.2	294771.4	290760.2	289924.6
15	292889.0	292926.6	292283.8	293154.0	290380.8	288141.2
16	292294.4	292010.6	292239.4	292489.4	290055.6	287420.8
17	291711.2	291582.4	290960.2	290705.8	289337.0	287040.8
18	290783.4	290762.4	290593.4	290705.8	288656.8	286470.2
19	289393.6	290533.6	290297.2	289603.6	288115.6	286060.8
20	288763.8	289834.4	289643.8	288896.0	287924.4	285611.0
21	288400.2	289052.8	289327.6	288599.8	287845.8	284741.2
22	287958.8	28870.09	288184.2	288071.6	287202.6	284697.4
23	287958.8	287877.4	288050.8	287288.4	286710.6	284694.0
24	287154.8	287427.4	286694.0	286908.8	285778.4	284017.4
25	286659.8	287007.6	285979.2	286133.2	284601.4	283491.4
26	286246.4	286539.2	285163.4	285917.8	283826.8	282868.0
27	286246.4	286079.4	284765.2	284901.4	283826.8	282760.8
28	286110.0	285938.4	284765.2	284549.4	283052.6	282217.8
29	285890.6	285398.2	284479.0	283446.4	283052.6	282175.2
30	285109.6	285343.2	283892.6	283134.0	282582.4	281645.6
31	284317.0	284816.2	283285.2	282365.8	282012.6	281546.4
32	283752.6	284805.6	283089.6	280507.4	281962.0	281189.6
33	283530.4	283790.6	283079.6	280442.2	281688.4	281179.8
34	282800.6	283363.0	282847.8	280442.2	281602.0	280381.4
35	282691.4	282432.6	282432.2	280209.0	281179.0	280095.0
36	282691.4	282432.6	281921.4	279634.2	280663.4	279872.0
37	282630.8	282432.6	281590.0	279173.2	280630.4	279475.6
38	282362.4	282104.0	280909.4	278857.6	280620.8	279441.6
39	282362.4	282104.0	280504.4	278704.4	280340.6	279158.2
40	282154.8	281847.2	280316.6	278633.6	280039.6	278944.6
41	282022.2	281847.2	280028.8	278485.0	279751.0	278380.6
42	281898.8	281581.4	280002.8	278394.4	279573.4	278376.6
43	281889.6	281486.0	279667.2	278091.6	279509.2	277697.4
44	281788.8	281421.8	279619.8	277790.0	279161.6	277644.4
45	281313.0	280866.6	279498.6	277451.6	278980.2	277347.2
46	281050.6	280866.6	279313.8	277367.2	278669.2	277295.6
47	281024.8	280430.8	279140.2	277206.2	278450.2	276982.4
48	280835.6	280224.6	278904.4	277073.4	278075.2	276896.4
49	280564.6	280188.6	278652.2	277073.4	277948.8	276676.8
50	280564.6	280039.6	278556.2	276623.6	277620.8	276676.8

7.2 Topologías

A continuación se presentan las dos topologías que fueron utilizadas para realizar las pruebas. Cabe destacar que la topología 1 se ha ejecutado con los parámetros distintos a los que se presentaron en los dos puntos anteriores, mientras que para la topología 2 se utilizaron estos.

7.2.1 Topología 1

La figura 7.4 ilustra la primera topología utilizada, donde las divisiones se hacen en base a la cantidad de centros de cluster a conformar. La Tabla 7.4 muestra los resultados obtenidos para la topología.



Figura 7.4: Topología 1

Tabla 7.4: Resultados Topología 1

Balance	Prioridad	Función Objetivo
100 %	15	296918
100 %	12	299720
100 %	8	369838
70 %	15	294075
70 %	12	300802
70 %	8	379503

7.2.2 Topología 2

Una segunda topología implementada se realizó por medio de divisiones por cuadrantes, donde la cantidad de cuadrantes a conformar se rige por la cantidad de centros de cluster existentes. Cabe mencionar que los límites de los cuadrantes se delimitan por la ubicación geográfica de las estaciones, tal como lo ilustra la figura 7.5. La Tabla 7.5 muestra los resultados obtenidos para la topología expresada, donde para todos aquellos experimentos en el que la prioridad es de 5 o menos no se encontraron resultados optimos, mientras que aquellos con un tiempo de ejecución mayor a 6 horas fueron finalizados antes de alcanzar el número total generaciones.

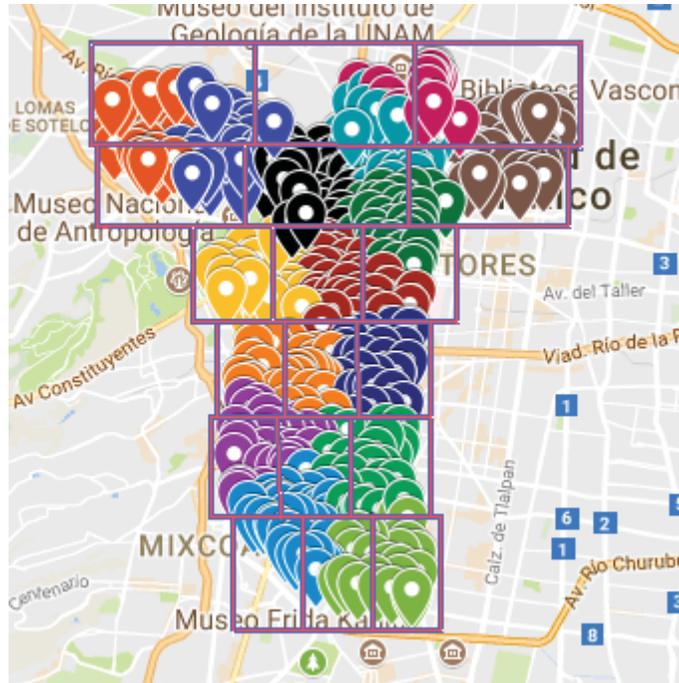


Figura 7.5: Topología 2

Tabla 7.5: Resultados Topología 2

Balance	Prioridad	Tiempo	Función Objetivo
100 %	15	9 min.	273511
100 %	12	10 min.	278460
100 %	8	12 min.	333105
70 %	15	10 min.	280417
70 %	12	11 min.	281172
70 %	8	28 min.	343093
35 %	15	3 hrs. 5 min.	291195
35 %	12	3 hrs. 12 min.	294395
35 %	8	3 hrs. 21 min.	363227
20 %	15	>6hrs.	387690
20 %	12	>6hrs.	391224
20 %	8	>6hrs.	468901
18 %	15	>6hrs.	405368
18 %	12	>6hrs.	413201
18 %	8	>6hrs.	479702
15 %	15	>6hrs.	418398
15 %	12	>6hrs.	436810
-	5	-	Resultados Inviabiles

7.3 Resultados comparativos

A continuación, en la Tabla 7.6, se muestran por motivos comparativos los resultados obtenidos para el mismo problema con un algoritmo exacto[10].

Tabla 7.6: Resultados Algoritmo Exacto

Balance	Prioridad	Función Objetivo
100 %	15	258446
100 %	12	258726
100 %	8	263696
70 %	15	264665
70 %	12	264019
70 %	8	278952
35 %	15	285160
35 %	12	285092
35 %	8	287102
20 %	15	315387
20 %	12	315893
20 %	8	323337
18 %	15	328831
18 %	12	329092
18 %	8	333642
15 %	15	364927
15 %	12	351250

8 Conclusión

Ante una postura global donde la conciencia por el medio ambiente comienza a ser una pieza esencial en la evolución, se desenvuelve la idea de sostenibilidad, una nueva táctica empresarial apoyada en producir beneficios y considerar al mismo tiempo el medioambiente. En lugares de todo el mundo se han comenzado nuevas aspiraciones como campañas de concienciación, la generación de energía limpia, o proyectos sostenibles, en donde la opción de sistemas de transporte público es una de las más emprendidas.

A través de la investigación realizada fue posible implementar un algoritmo cultural con el propósito de resolver la problemática identificada en un comienzo. El desarrollo de éste fue incremental, implementando en un principio un algoritmo evolutivo simple, para luego ir agregándole funcionalidades en busca de un mejor desempeño. El resultado final es un algoritmo cultural, el cual posee conocimiento topológico del problema a través de la separación de las estaciones en cuadrantes según su distribución geográfica.

Finalmente, a pesar de que el algoritmo implementado, no logró alcanzar los resultados obtenidos por el algoritmo exacto, si se pudo demostrar que la elección de la topología de los cuadrantes tienen efecto en el desempeño del algoritmo, por lo que queda por delante el analizar y encontrar mejores distribuciones de las estaciones. La ventaja del algoritmo yace en el tiempo que demora en encontrar las soluciones. Se espera que esta investigación sirva de punto de inicio para desarrollos futuros, los cuales puedan construir sobre esta base para poder encontrar mejores soluciones, de manera más efectiva y eficaz. Un posible camino para esto podría ser el expandir el algoritmo para incorporar la optimización multi-objetivo, lo cual se presenta como una opción natural dadas las características del problema.

Referencias

- [1] Miguel Bea Alonso. Los sistemas de bicicletas públicas urbanas. Barcelona, España, 2009. Documentos de investigación del Programa de Doctorado de Economía Aplicada de la Universidad Autónoma de Barcelona.
- [2] C. Morency C. Contardo and L. Rousseau. Balancing a dynamic public bike sharing system. Canada, 2012. CIRRELT.
- [3] Leonardo Caggiani and Michele Ottomanelli. A dynamic simulation based model for optimal fleet repositioning in bike-sharing systems. *Procedia - Social and Behavioral Sciences*, 87:203–210, 2013.
- [4] C. Chung and CAEP Reynolds, G. R. An evolution-based tool for real-valued function optimization using cultural algorithm. *International Journal on Artificial Intelligence Tools*, 7:239–291, 1998.
- [5] Mauro Dell’Amico, Eleni Hadjicostantinou, Manuel Iori, and Stefano Novellani. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, 45:7–19, 2014.
- [6] Mauro Dell’Amico, Manuel Iori, Stefano Novellani, and Thomas Stützle. A destroy and repair algorithm for the bike sharing rebalancing problem. *Computers Operations Research*, 71:149 – 162, 2016.
- [7] Aitana Vidal Esmorís. Algoritmos heurísticos en optimización. *Universidad de Santiago de Compostela*, 2013.
- [8] Benjamin Franklin and Marcel Bergerman. Cultural algorithms: Concepts and experiments. In *Proceedings of the 2000 Congress on Evolutionary Computation.*, volume 2, pages 1245–1251, 2000.
- [9] Luca Di Gaspero, Andrea Rendl, and Tommaso Urli. Balancing bike sharing systems with constraint programming. *Constraints*, 21(2):318–348, 2016.
- [10] Miguel Angel Arroyave Guerrero. Diseño de estrategias para el reposicionamiento de unidades en sistemas públicos de bicicletas. *Universidad de Antioquia*, 2017.
- [11] Hartwig H. Hochmair. Assessment of latent bicycle demand in street networks. University of Florida, 2009. Research and Education Center.
- [12] Jenn-Rong Lin and Ta-Hui Yang. Strategic design of public bicycle sharing systems with service level constraints. *Transportation Research Part E: Logistics and Transportation Review*, 47(2):284–294, 2011.
- [13] Jenn-Rong Lin, Ta-Hui Yang, and Yu-Chung Chang. A hub location inventory model for bicycle sharing system design: Formulation and solution. *Computers Industrial Engineering*, 65(1):77–86, 2013.
- [14] D. Perez M. Arroyave, P. Maya. Gestión de operaciones y logística oportunidades en movilidad sostenible: puntos de intervención en sistemas públicos de bicicletas. *The 8th International Conference on Production Research*, 2016.

- [15] Oliver O'Brien, James Cheshire, and Michael Batty. Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography*, 34:262–273, 2014.
- [16] Tal Raviv, Michal Tzur, and Iris A. Forma. Static repositioning in a bike-sharing system: models and solution approaches. *EURO Journal on Transportation and Logistics*, 2:187–229, 2013.
- [17] Robert G. Reynolds. An introduction to cultural algorithms. In Anthony V. Sebald and Lawrence J. Fogel, editors, *Evolutionary Programming — Proceedings of the Third Annual Conference*, pages 131–139, San Diego, CA, USA, 1994. World Scientific Press.
- [18] Robert G. Reynolds. New ideas in optimization. chapter Cultural Algorithms: Theory and Applications, pages 367–378. McGraw-Hill Ltd., UK, Maidenhead, UK, England, 1999.
- [19] Robert G. Reynolds and Bin Peng. Cultural algorithms: Computational modeling of how cultures learn to solve problems: An engineering example. *Cybernetics and Systems*, 36(8):753–771, 2005.
- [20] J. Schuijbroek, R.C. Hampshire, and W.-J. van Hoes. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992–1004, 2017.
- [21] J. Schuijbroek, R.C. Hampshire, and W.-J. van Hoes. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992–1004, 2017.
- [22] Aaron Hottenstein Shawn Turner and Gordon Shunk. Bicycle and pedestrian travel demand forecasting: Literature review. College Station, Texas, 1997. Texas Transportation Institute, Texas Department of Transportation.
- [23] Divya Singhvi, Somya Singhvi, Peter I. Frazier, Shane G. Henderson, Eoin O'Mahony, David B. Shmoys, and Dawn B. Woodard. Predicting bike usage for new york city's bike sharing system. In *AAAI Workshop: Computational Sustainability*, 2015.
- [24] Patrick Vogel, Torsten Greiser, and Dirk Christian Mattfeld. Understanding bike-sharing systems using data mining: Exploring activity patterns. *Procedia - Social and Behavioral Sciences*, 20:514 – 523, 2011.
- [25] Jin Xidong and Robert Reynolds. Using knowledge-based evolutionary computation to solve nonlinear constraint optimization problems: A cultural algorithm approach. In *Proceedings of the 1999 Congress on Evolutionary Computation*, volume 3, page 1678 Vol. 3, 1999.