



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

HEURÍSTICA PARA LA GENERACIÓN DE ZONAS DE REPARTO PARA LOS
CARTEROS DE CORREOS DE CHILE

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN GESTIÓN DE OPERACIONES
MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL

FRANCISCO JAVIER VEGA ELGUETA

PROFESOR GUÍA:
RAFAEL EPSTEIN NUMHAUSER
PROFESOR CO-GUÍA:
RODOLFO URRUTIA URIBE

MIEMBROS DE LA COMISIÓN:
EDUARDO ÁLVAREZ MIRANDA
FERNANDO ORDOÑEZ PIZARRO

SANTIAGO DE CHILE
2017

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
POR: FRANCISCO JAVIER VEGA ELGUETA
FECHA: 2017

PROF. GUÍA: RAFAEL EPSTEIN NUMHAUSER PROFESOR CO-GUÍA: RODOLFO URRUTIA URIBE

HEURÍSTICA PARA LA GENERACIÓN DE ZONAS DE REPARTO PARA LOS CARTEROS DE CORREOS DE CHILE

Resumen

Este trabajo se presenta una heurística basada en métodos de particionamiento para el diseño de zonas de reparto para los carteros de Correos de Chile. La heurística es una elaboración propia, basada en métodos de localización y clusterización con restricciones de equidad de carga de la jornada laboral.

La empresa Correos de Chile es una empresa que mueve más de 250 millones de cartas y paquetes anualmente a lo largo de todo el país, para lo cual cuenta con cerca de 2.300 carteros para los envíos postales. Cada uno de estos carteros es responsable de una zona de reparto denominada cuartel. Estas zonas han sido consecuencia de una lenta evolución histórica y de manera aislada en cada comuna, y constituye un problema complejo debido a su gran tamaño. Mas aún, los parámetros que caracterizan el problema son dinámicos: crecimientos de las zonas urbanas, variaciones de tráfico de correspondencia, etc.

El tráfico de correspondencia varía por código postal y está caracterizada según los diferentes tipos de productos ofrecidos por la empresa (Carta Normal, Carta Certificada, Carta+, etc.), y tiene componentes de cantidad de envíos y tiempos de distribución, además de contar con factores geográficos y topográficos en las zonas de reparto.

La heurística propuesta se hace cargo del diseño de las zonas de reparto, equilibrando la carga entre cuarteles proveniente de la suma de las cargas de los códigos postales (CPs) que lo conforman. Para lo anterior se desarrolla un algoritmo que genera un grafo que representa las calles y CPs de cada comuna. Luego se desarrolla la heurística que asigna todos los CPs a los cuarteles más cercanos, y luego los traspasa entre cuarteles para equilibrar la carga.

Esta heurística fue probada en la comuna de Viña del Mar y Concón, cuya instancia tiene 11.831 códigos postales y 63 cuarteles a generar. Utilizamos dos mediciones de carga para la generación de las zonas de reparto, una que depende de la demanda, tamaño, tipo de inbueble, entre otros, y la segunda que depende de la cantidad de puertas pertenecientes a cada código postal.

La heurística permite obtener distintas soluciones factibles que garantizan equidad, lo cual permite a los tomadores de decisiones flexibilidad para seleccionar el diseño de los cuarteles.

Agradecimientos

Para finalizar esta etapa universitaria quisiera agradecer a todas las personas que me han acompañado en este proceso.

En primer lugar agradecer a toda mi familia, que han participado de todo este proceso conmigo y que han sido un apoyo fundamental para poder sacar adelante todos mis proyectos. Un agradecimiento especial a mis papás, Beatriz y Francisco, que son un ejemplo para mí, que siempre han estado presente en los momentos que los he necesitado y que permitieron desarrollarme tanto personal como profesionalmente.

Quisiera agradecer a los profesores que me acompañaron y me dieron la oportunidad de integrarme en este proyecto, además de enseñarme y guiarme a lo largo de todo este proceso.

También quiero dar gracias a todos mis amigos de la universidad, en particular agradecer a Felipe, Cesar, Natalie, Francisca y Alberto, ya que ellos fueron un gran apoyo a lo largo de toda la carrera.

Finalmente quiero agradecer a María Jesús, quien ha sido la persona más importante a lo largo de toda esta etapa, me acompañó y apoyó en todos los momentos, confió en mí, con su cariño me animó todos los días, y quien ha sido una excelente compañera de vida.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	4
1.2.1. Objetivo General	4
1.2.2. Objetivos Específicos	5
1.3. Alcances y Limitaciones	5
1.4. Estructura de la Tesis	6
2. Antecedentes y Marco Teórico	7
2.1. Modelos de Localización	7
2.1.1. Modelo Discreto	7
2.2. Grafo y Árboles	8
2.2.1. Grafo	8
2.2.2. Árbol	8
2.3. Chinese Postman Problem	9
2.4. Clusterización	10
2.4.1. Métodos de particionamiento	11
3. Descripción de la Heurística	13
3.1. Árbol de Conexión de Códigos Postales	13
3.2. Clustering	17
3.2.1.	17
3.2.2. Generador de Cuarteles	19
3.2.3. Polishing	21
4. Resultados	26
4.1. Caso de estudio: Comuna Viña del Mar y Concón	27
4.1.1. Generación del grafo	27
4.1.2. Función de Carga	29
4.1.3. Generador de Cuarteles	30
4.2. K-Means	35
4.3. Carga según Número de Puertas	36
4.4. Análisis de sensibilidad	38
5. Conclusiones	42
Bibliografía	45

A. Set Covering Model	46
B. P-Median	47
C. Max Covering Model	48
D. Cuarteles con K-Means	49
E. Cuarteles con carga igual al número de puertas	50

Índice de Ilustraciones

1.1.	Se muestran los cuarteles del centro de Viña del Mar, con sus respectivas manzanas y códigos postales.	3
1.2.	En el eje X se muestran los 63 cuarteles de la comuna Viña del Mar y Concón, y en el eje Y se muestra la carga total del cuartel, proveniente de una función de carga de trabajo recibida. Fuente: Elaboración propia.	5
2.1.	Grafo con 5 nodos y 5 aristas.	8
3.1.	Se muestran los 5 pasos de construcción de un grafo que cumple criterios de distancia y que solo rompe planaridad dentro de las manzanas, ya que de esta manera quedan todos los códigos postales pertenecientes a la misma manzana conectados entre sí. Paso 1: los códigos postales iniciales, Paso 2: división de los códigos postales en cuadrantes, Paso 3: árbol de expansión en cada cuadrante, Paso 4: agregar aristas por cuadrante sin romper planaridad, Paso 5: conexión entre cuadrantes sin romper planaridad.	14
3.2.	Izquierda: cuatro códigos postales unidos con arcos no planares. Derecha: solución deseable de planaridad	15
3.3.	Ejemplo de solución no deseable de cuarteles disconexos. Cada celda corresponde a un código postal con su respectiva carga (número de la celda). El cuartel azul no cumple el requerimiento de conectividad deseado.	17
3.4.	Ejemplo de solución no deseable de cuarteles, cuyas manzanas pertenecen a dos zonas de reparto diferentes.	18
3.5.	Transformación de CPs a manzanas: En la imagen de la izquierda se encuentran las manzanas con sus códigos postales (círculos azules), y los arcos que los conectan. En la imagen de la derecha se encuentran las manzanas representadas por su centro de gravedad (círculos rojos), y los arcos que las conectan entre ellas.	19
3.6.	En la imagen se muestra un cuartel no deseable (cóncavo), ya que el cartero podría preferir ir del punto A al punto B a través del punto C, el cual pertenece a otro cuartel, y por lo tanto, a otro cartero.	20
3.7.	Se muestra el cuartel con mayor carga (0) y sus cuarteles vecinos con el indicador de la cercanía con él.	22
3.8.	Se muestra como se va propagando la carga, desde el cuartel de mayor carga hacia afuera.	23
4.1.	Fuente: Censo 2002. http://www.vinadelmarchile.cl	26
4.2.	Fuente: Censo 2002. http://www.vinadelmarchile.cl	27

4.3.	A la izquierda están los códigos postales de Viña y Concón, y a la derecha se encuentran los mismos códigos postales y el color representa el cuadrante al que pertenecen.	28
4.4.	Árbol de expansión en cada cuadrante (imagen de la izquierda), y conexiones entre ellos (imagen de la derecha).	29
4.5.	Distribución de carga (eje Y) de los cuarteles (eje X) de las comunas de Viña del Mar y Concón. Tiene un GAP del 131 %, un promedio de carga de trabajo equivalente a 188,7 y una desviación estándar de 62,3.	30
4.6.	En gris se encuentran los códigos postales de las comunas de Viña del Mar y Concón, y en verde están las semillas utilizadas en la heurística.	31
4.7.	Distribución de carga de los cuarteles generados con el algoritmo. GAP de 24 %, promedio de carga de trabajo igual a 212,3 y con una desviación estándar de 7,4.	32
4.8.	Cuarteles generados con algoritmo. Hay conexidad de cuarteles, y todos los CP de una manzana pertenecen al mismo cuartel.	32
4.9.	Evolución del GAP de la solución a través de las 33 iteraciones del <i>Polishing</i>	33
4.10.	Cuarteles generados en la comuna de Viña del Mar y Concón. Se pueden ver cuarteles de área muy grande y otros de área pequeña, esto ocurre debido al equilibrio de cargas de trabajo y no a la cantidad de CPs de cada cuartel, razones por la que K-Means no funciona, ya que no se busca minimizar la distancia entre cada CP al centroide de su cuartel.	34
4.11.	Distribución de carga (eje Y) de los cuarteles generados (eje X) en la comuna de Viña del Mar y Concón, mediante el método de particionamiento K-Means. Se obtuvo un GAP del 88,3 %, un promedio de 107,9 y una desviación estándar igual a 47,5	35
4.12.	Distribución de carga de los cuarteles actuales con la carga de cada código postal igual al número de puertas que contiene. Eje Y corresponde a la cantidad de puertas que contiene cada cuartel del eje X. GAP igual a 206,6 %, promedio de 1.928 y desviación de estándar 862,8.	37
4.13.	Distribución de carga de los cuarteles generados con la heurística y carga de cada código postal igual al número de puertas que contiene. GAP de 30 %, promedio de 1.928 puertas y desviación estándar de 119.	38
4.14.	Evolución del GAP de dos instancias con semillas aleatorias, a lo largo de las iteraciones del <i>Polishing</i>	39
4.15.	Distribución de las semillas aleatorias y los cuarteles obtenidos a partir de ellas. GAP del 23 % obtenido en 13,4 minutos.	40
4.16.	Distribución de las semillas aleatorias y los cuarteles obtenidos a partir de ellas. GAP del 22 % obtenido en 14,2 minutos.	40
4.17.	Distribución de carga de trabajo de los 63 cuarteles para ambas instancias de semillas aleatorias.	41
5.1.	Cuarteles del centro de Viña generados en las 3 instancias descrita, y con sus cargas respectivas. El de la izquierda corresponde al generado con la heurística y con función de carga continua entre 1 y 10. El de arriba a la derecha corresponde a los cuarteles generados con K-Means e igual carga que el anterior. El de abajo a la derecha corresponde a los cuarteles generados con la heurística y carga igual al número de puertas del CP.	43

D.1. Cuarteles generados en la comuna de Viña del Mar y Concón mediante el método de particionamiento K-Means.	49
E.1. Cuarteles generados en la comuna de Viña del Mar y Concón mediante con la carga de los código postales igual al número de puertas que contienen.	50

Capítulo 1

Introducción

1.1. Motivación

La presente tesis, busca desarrollar un algoritmo de generación de zonas de reparto (zonas geográficas en las cuales los carteros distribuyen las cartas) para las operaciones de distribución de empresas del mercado postal, que tienen un tamaño de 70 mil millones de pesos anuales en Chile, y empresas del mercado CEP¹ con 230 mil millones. Dentro de las principales funciones de este tipo de empresas se encuentra la etapa de distribución, la cual involucra una gran cantidad de recursos y operaciones complejas.

Si bien se busca un algoritmo aplicable a cualquier empresa del rubro, se trabajará con Correos de Chile, empresa líder en el mercado postal con una participación de 64 %.

Correos de Chile es una empresa pública fundada en 1736, dedicada al servicio de correspondencia, giros postales y al mercado de envíos y encomiendas nacionales e internacionales, cumpliendo con las funciones de Servicio Postal Universal. En sus casi 270 años de existencia, Correos de Chile se ha ido perfeccionando para llegar a ser reconocida dentro de las mejores empresas de correos a nivel internacional. Esto la llevó a alcanzar una cifra histórica en la movilización de paquetes internacionales el año 2015, ver de Chile [2].

Correos de Chile tenía, para el año 2013, un costo anual asociado a las operaciones de la línea postal y CEP del orden de los *MM*\$47,000, correspondientes a un 64 % de sus costos operacionales. Dada la importancia de los montos asociados, un incremento marginal en la eficiencia implicaría un gran impacto económico para la compañía. Por ejemplo, lograr una mejora del 1 % en la eficiencia de los costos operacionales de las líneas Postal y CEP representaría un ahorro de *MM*\$470 anuales.

Dentro de las operaciones de Correos de Chile, nos enfocaremos principalmente en el mercado postal, el cual corresponde al servicio de correspondencia a todo Chile y el mundo. Dentro de los principales productos que ofrece la empresa en este mercado, se encuentran:

¹Courier, Expreso y Paquetería

- Carta Normal: envíos de correspondencia a todo Chile y el mundo. Servicio diseñado para empresas que requieren despachar grandes volúmenes de documentos a clientes en forma periódica.
- Carta Certificada: envíos de correspondencia a todo Chile, única con certificación de entrega. El servicio permite dar fe pública respecto de la entrega al destinatario, la fecha y el domicilio en el que fue recibida dicha carta.
- Carta Registrada: envíos de correspondencia nacional con digitalización de la constancia de entrega vía web en todo Chile.
- Servicios Especiales: soluciones integrales de distribución postal a la medida del cliente.
- Carta +: envíos de correspondencia de última generación, con visualización del estado de entrega vía web a través de carteros equipados con dispositivos GPS.

Correos de Chile es una empresa que mueve más de 250 millones de cartas y paquetes anualmente a lo largo de todo el país, y dentro de sus canales de distribución, la compañía cuenta con 200 vehículos para el mercado CEP, y cerca de 2.300 carteros para los envíos postales.

La unidad geográfica de distribución para los envíos postales se denomina *cuartel*, el cual corresponde a la zona de reparto asociada a un cartero. Cada una de estas zonas de reparto cuenta con una superficie, cantidad de puntos de entrega, tráficos de envío de correspondencia, topologías y tipologías distintas. Por su parte la unidad básica de un cuartel, es el código postal (Figura 1.1), el cual corresponde a un frente de manzana², y que cuenta con tráfico, número de puertas, tipo de estructura, etc.

La configuración de sectores actuales es consecuencia de una lenta evolución histórica. En Correos de Chile, actualmente existe una instancia en donde se evalúa la reestructuración y reasignación de las zonas de reparto, a cargo de la Comisión Nacional de Distribución. En dicha instancia la Comisión Nacional de Distribución estudia modificaciones posibles a las zonas de reparto, pero con un foco muy local: se reestructura una zona de reparto a la vez. Esta tarea se lleva a cabo sin la ayuda de sistemas de apoyo a la toma de decisiones.

El problema de definición de las zonas de reparto es complejo por su gran tamaño. Más aún, los parámetros que caracterizan el problema son dinámicos: variaciones de tráfico de correspondencia, aumento o crecimiento de las zonas urbanas o pobladas, construcción de edificios, cambios en el plan regulador municipal, y otros.

Para alcanzar sus objetivos en el nivel de servicio, Correos de Chile mantiene una constante innovación tecnológica y una fuerte motivación a mejorar la productividad de los casi 2.300 carteros, por lo que aparece la necesidad de reconfigurar sectores territoriales de reparto, con el fin de equilibrar la jornada laboral de sus empleados, y de ofrecer un servicio de mayor calidad a sus clientes.

Esta reconfiguración de los sectores territoriales, dependerá del tiempo que un cartero demore en satisfacer cada uno de los códigos postales de su cuartel durante un día promedio. Para esto, es necesario establecer una función de carga, que considere todas las variables antes mencionadas, y entregue como resultado el tiempo de trabajo necesario para satisfacer

²Todas las direcciones de un frente de manzana, están asociadas al mismo código postal.

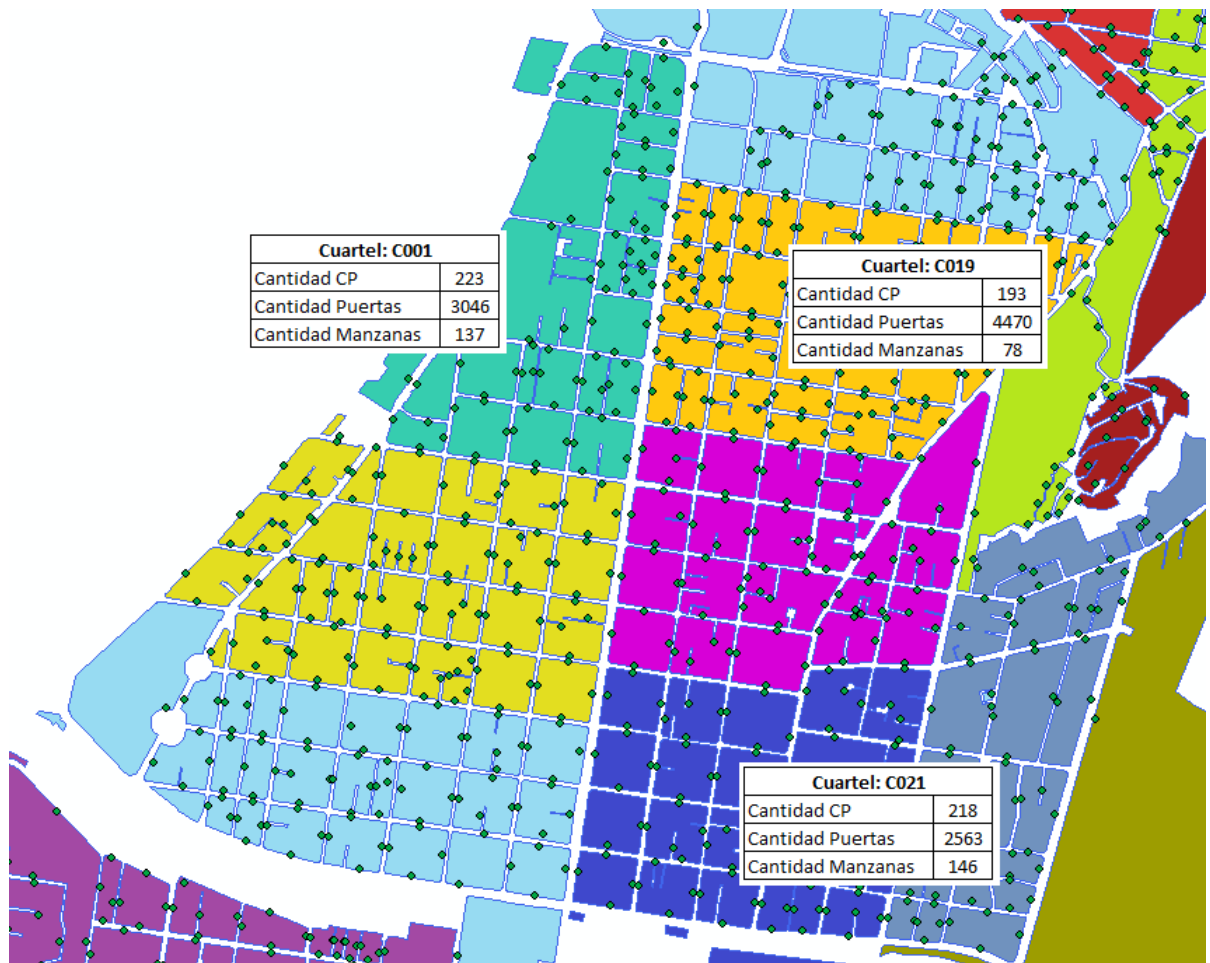


Figura 1.1: Se muestran los cuarteles del centro de Viña del Mar, con sus respectivas manzanas y códigos postales.

un cuartel, o más específicamente, cada uno de los códigos postales que lo componen.

A partir de lo anterior, hemos desarrollado un algoritmo heurístico que permita abordar el problema de la definición de zonas de reparto o cuarteles que enfrentan las empresas postales. Una heurística corresponde a un algoritmo que busca una buena solución al problema, sin embargo no es necesariamente óptima. Utilizaremos este mecanismo debido a la complejidad del problema y debido al tiempo de ejecución que involucraría encontrar el óptimo. Hacemos notar que los datos de medición de tiempos de reparto de algunos carteros contienen errores, dispersión y sesgo, por lo no es evidente que se encuentre el óptimo del problema.

La herramienta tendrá un gran requerimiento de información de entrada: información geográfica y geo-referenciada sobre las unidades operativas de distribución actuales, información sobre el tráfico de envíos y sus características, entre otras. La información viene caracterizada a nivel de código postal, es decir, para cada uno de los códigos postales de la comuna se tiene la ubicación geográfica, la cantidad de puertas, el tipo de inmueble, la cantidad de cartas que se entregan en dicho código postal (tráfico de cartas) e información topográfica. Por otra parte se tienen datos a nivel de carteros, que corresponde a mediciones realizadas por la empresa, que entregan el tiempo durante una jornada laboral que demora el cartero

en cada puerta del código postal entregando las cartas. También cuenta con el tiempo que demora desplazándose entre códigos postales y lo que demora dentro de un código postal, es decir en los traslados entre puertas pertenecientes al mismo código postal.

El objetivo de este trabajo es diseñar una herramienta que permita agrupar códigos postales de tal manera que exista carga homogénea entre cuarteles o zonas de reparto y conectividad³ entre los códigos postales de una misma zona.

El presente informe, si bien desarrolla una heurística aplicable para todas las empresas del rubro y para todas las comunas, se centrará en las comunas de Viña del Mar y Concón, por ser una zona acotada que reúne todos los conceptos requeridos para entender en profundidad el problema de definición y asignación de zonas de reparto.

La comuna de Viña del Mar pertenece a la V región, cuenta con una superficie de 121,6 km cuadrados, tiene una población de 290.000 habitantes y se encuentra a 120 kilómetros de Santiago. Por su parte, Concón cuenta con 76 km cuadrados de superficie y 63.000 habitantes. Entre ambas comunas tienen un total de 63 cuarteles y aparecen muchas de las características que influyen en el rendimiento de un cartero. En particular identificamos:

- Zonas urbanas y rurales.
- Edificios, Malls, Casas, Edificios corporativos, condominios, etc.
- Distintos estratos económicos.
- Zonas con distinta densidad poblacional y vehicular.

De la Figura 1.2 se puede ver la heterogeneidad en la carga de los cuarteles de la comuna de Viña del Mar. Esto genera complicaciones en la operación de la empresa, ya que la calidad de servicio, entendida como el tiempo de entrega de una carta, se verá favorecida o perjudicada según el cuartel a la cual pertenezca la dirección de destino. Este patrón se repite a lo largo de todas las comunas de Chile, lo que justifica la construcción de la herramienta antes descrita.

1.2. Objetivos

1.2.1. Objetivo General

El objetivo general es el desarrollo de una herramienta de agrupación de códigos postales, para la generación de zonas de reparto equilibradas en carga de trabajo para los carteros de Correos de Chile.

³Dentro de una zona de reparto, todos los códigos postales deben estar conectados. Es decir, se puede ir de un código postal a otro sin la necesidad de pasar a través de otro cuartel.

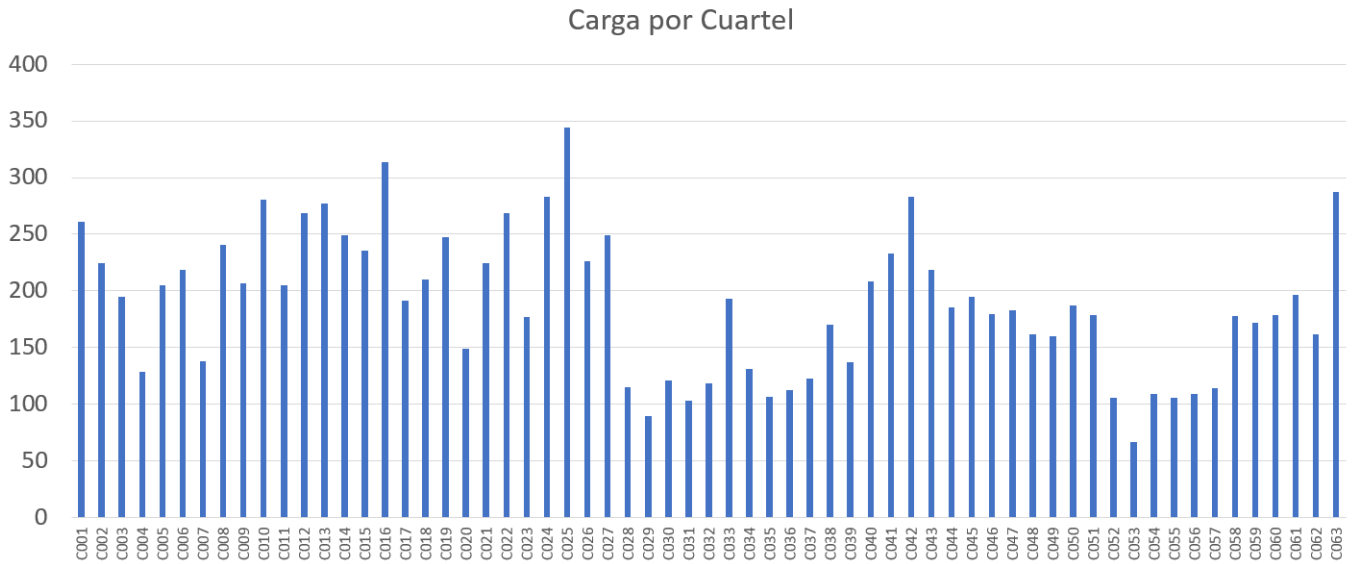


Figura 1.2: En el eje X se muestran los 63 cuarteles de la comuna Viña del Mar y Concón, y en el eje Y se muestra la carga total del cuartel, proveniente de una función de carga de trabajo recibida. Fuente: Elaboración propia.

1.2.2. Objetivos Específicos

1. Crear un algoritmo para la generación de un grafo que represente la adyacencia de los códigos postales de la comuna en cuestión.
2. Desarrollar y comparar heurísticas de clusterización de acuerdo a la carga de trabajo asociada de los cuarteles generados.
3. Realizar un análisis de sensibilidad de los parámetros del modelo para ver la robustez de éste.

1.3. Alcances y Limitaciones

Tal como fue descrito en las secciones 1.1 y 1.2 este trabajo desarrolla una metodología de clusterización para el diseño geográfico de los cuarteles de Correos de Chile. Este trabajo se hace cargo de las condiciones establecidas por la empresa, en particular de las restricciones de conectividad de los cuarteles y manzanas, aspectos de forma de cada cuadrante y la inclusión de la carga de los códigos postales.

Este trabajo tiene ciertas limitaciones y simplificaciones las cuales son nombradas a continuación:

- Se describe un algoritmo de generación de un grafo G que une códigos postales, el cual es necesario para la generación de las zonas de reparto, ya que con este se miden los tiempos de desplazamiento, además de ver vecindades entre códigos postales y manzanas. Este algoritmo se preocupa de generar conexiones planares y con restricciones de distancias.

Sin embargo no se evaluará dentro de este informe la posibilidad de la generación de otro grafo, ni se preocupará de entender si es o no el grafo óptimo del problema.

- No se encuentra dentro de los objetivos la generación de una función de carga de trabajo de carteros, ya que para efectos de esta tesis, se considerará como un input del problema, y la estimación de dicha carga de trabajo a nivel operacional no es un problema trivial por lo que es un tema en sí. Se cuenta con una función de carga, la cual recibe los diferentes parámetros asociados a cada código postal, número de puertas, tipo de inmueble, demanda y ubicación, y entrega un valor continuo entre 1 y 10 correspondiente a la carga que adiciona la incorporación de dicho código al cuartel. Esta función determinará la forma de los cuarteles, ya que dependiendo de la carga que adicione un código postal, se incorporará en un cuartel o en otro.
- Este problema tiene aleatoriedad en los parámetros, ya que a lo largo del tiempo, la demanda o tráfico de envío de correspondencia de cada código postal varía, además existen errores de medición en los tiempos de distribución por parte de Correos de Chile, debido a que no se alcanzan a medir todos los tiempos, y porque los carteros que pertenecieron a esta medición tienen un comportamiento distinto cuando los miden y cuando no. Por otra parte existen cambios topográficos de las comunas, que pueden generar cambios de los tipos de inmueble de un código postal, o incluso la incorporación de nuevas manzanas.
- Esta heurística genera una cantidad fija K de cuarteles, no entrega la cantidad óptima de zonas a asignar, si no que las equilibra. Sin embargo, se pueden construir y equilibrar las zonas de reparto variando la cantidad total de cuarteles K , buscando de esta manera la cantidad óptima de a generar.
- Al ser una heurística, no se llega a la asignación óptima de códigos postales. La medida de calidad de la empresa Correos de Chile es el *GAP* entre cuarteles. Dicho *GAP* consiste en un porcentaje equivalente al cociente entre la diferencia de carga de trabajo del cuartel de mayor carga y el de menor carga con una cota de carga establecida. Se definió como aceptable un *GAP* menor o igual al 25

1.4. Estructura de la Tesis

La estructura utilizada en esta tesis para representar el trabajo realizado, es la siguiente:

1. **Capítulo 2. Antecedentes y Marco Teórico:** Se divide en dos secciones, la primera asociada a modelos de localización. La segunda está asociada a la revisión bibliográfica asociada a métodos de clusterización, y literatura de aplicaciones similares en el ámbito de correos.
2. **Capítulo 3. Descripción de la Heurística:** Corresponde a la explicación de la heurística. Se explicarán los requerimientos necesarios y la implementación.
3. **Capítulo 4. Resultados:** Se presentan los resultados del algoritmo obtenidos para la comuna de Viña del Mar y Concón. Luego se presenta un análisis de sensibilidad y comparación con otros modelos de clustering.
4. **Capítulo 5. Conclusiones:** Se presentan las conclusiones relacionadas con el trabajo realizado, y propuestas para posibles trabajos futuros que nacen de la herramienta.

Capítulo 2

Antecedentes y Marco Teórico

En este capítulo se entrega la contextualización y conceptos relacionados con el problema de localización y de clusterización enfocados en la elaboración de cuarteles.

2.1. Modelos de Localización

Para comenzar a entender el diseño de territorios, es buena alternativa comenzar introduciendo los modelos de localización. Estos modelos buscan ubicar distintos puntos en el espacio con algún objetivo en particular, como por ejemplo, satisfacer la mayor demanda posible. Estos puntos a localizar serán denominados centros, y pueden ser interpretados como los centros de los cuarteles.

Según Daskin [1], existen distintos modelos de localización que toman diferentes supuestos, tales como asumir la demanda con distribución conocida, centros en una cantidad continua de lugares diferentes para ser localizados, entre otros. En particular interesan los modelos discretos de localización, ya que la demanda se encuentra en nodos, y los candidatos de centros están en un conjunto acotado de lugares.

2.1.1. Modelo Discreto

Dentro de los modelos de localización discreto existen varios tipos, que satisfacen distintos requerimientos. En particular se encuentran los modelos de *covering* (Apéndice A) y de *p-median* (Apéndice B).

El primero busca minimizar el tiempo de respuesta del centro a los nodos, buscando la mínima cantidad de nodos que logra satisfacer la demanda. Por otro lado el segundo busca localizar P centros de un conjunto de candidatos, para asignar la demanda a cada uno de ellos, de tal manera de minimizar el tiempo o distancia promedio desde los centros a los nodos.

Sin embargo, el modelo que más interesa es el de *Max Covering* (Apéndice C), que dado un número de centros, se busca cubrir la mayor demanda posible. Este, en particular, es el que más se asimila a lo buscado, ya que la cantidad de cuarteles a generar con el algoritmo, es un parámetro dado que depende de la comuna.

Si bien el modelo de Max Covering se aproxima a lo buscado, no se hace cargo de equilibrar la demanda de cada uno de los centros, o mejor dicho, no busca equilibrar la carga de cada uno de los cuarteles. Sin embargo genera la idea de arcos y nodos que permiten satisfacer criterios de conexidad y compacidad de cuarteles planteados por Correos de Chile.

2.2. Grafo y Árboles

2.2.1. Grafo

Un grafo se compone de vértices y aristas que los conectan. Formalmente un grafo es un par de conjuntos (V, E) , donde V corresponde al conjunto de vértices y E al conjunto de aristas formadas por pares de vértices.

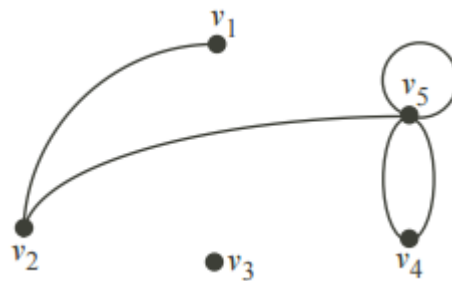


Figura 2.1: Grafo con 5 nodos y 5 aristas.

Tenemos $V = v_1, \dots, v_5$ y $E = (v_1, v_2), (v_2, v_5), (v_5, v_5), (v_5, v_4), (v_4, v_5)$

Si las aristas son unidireccionales, el conjunto de arcos se denomina con una letra D .

2.2.2. Árbol

Un árbol corresponde a un grafo que no tiene circuitos cerrados, es decir, que si avanzo a través de los nodos, jamás podré pasar nuevamente por uno que ya ha sido visitado, a menos que retroceda.

Por su parte, un árbol de expansión o *spanning tree*, es un árbol que conecta a todos los nodos de un grafo.

Teorema: Si el grafo G tiene n vértices y m arcos, entonces los siguientes son equivalentes:

1. G es un árbol
2. Hay exactamente una ruta que conecta 2 nodos cualquiera. G no tiene loops
3. G está totalmente conectado y $m = n - 1$
4. G no tiene circuitos cerrados y $m = n - 1$
5. G no tiene circuitos cerrados, pero al agregar cualquier arista al grafo, se genera uno

Teorema: *Un grafo que tiene todos sus nodos conectados, contiene al menos un spanning tree.*

Este último teorema es importante para más adelante, ya que permitirá que al generar el grafo para la realización de la heurística, éste tenga todos sus nodos conectados.

2.3. Chinese Postman Problem

El problema de los carteros ha sido estudiado anteriormente y se han modelado diversas variaciones de este problema. La idea de este problema es modelar el recorrido de un cartero para repartir todas las cartas.

Para definir el recorrido de los carteros, el problema se modela como un grafo, donde los arcos representan las calles de la comuna, y los vértices la intersección de ellas.

El Chinese Postman Problem (CPP) comenzó a ser estudiado en 1962 por Mei-Ko Kwan, quien estudió una metodología para rutear carteros. El objetivo del problema es encontrar la ruta mínima para repartir todas las cartas, es decir se busca encontrar la ruta de largo más corto, tal que el cartero pase a través de todos los arcos. Su principal innovación en el ruteo fue que un cartero podía pasar más de una vez por un arco a diferencia del *TSP* por ejemplo, haciendo del problema NP-Completo.

Más precisamente, un tour de un cartero, se define como un circuito del grafo G , que contiene todos los arcos e al menos una vez. Un tour Euler, es aquel circuito que contiene todos los arcos del grafo exactamente una vez.

A lo largo de los años han existido diversas variaciones del problema, tales como:

- CPP no direccionado: los arcos del grafo no son direccionados, es decir, si hay un arco que conecta los vértices A y B, entonces el cartero puede ir de igual forma de A a B, como de B a A.
- CPP direccionado: en este caso los arcos del grafo si tienen dirección única.
- CPP mixto: combina ambas metodologías mencionadas anteriormente.
- CPP con capacidad limitada: en este, los carteros tienen una capacidad acotada de distancia por recorrer.
- CPP con tiempo: en esta variación, las cartas deben ser entregadas con ventanas de tiempo, dentro de la jornada laboral del cartero.
- Entre otros.

Si bien el problema de los carteros ha sido bien estudiado en los últimos 50 años, ninguno de estos estudios se ha hecho cargo de generar las zonas geográficas de reparto de cada cartero y tampoco se hacen cargo de equilibrar la carga de trabajo entre carteros. Sin embargo, la idea del Chinese Postman Problem con capacidad limitada, entrega la idea base de la heurística que se quiere generar, donde la capacidad de cada cartero será definida como la cota de carga máxima que un cartero puede abastecer en su jornada laboral. Por lo tanto, la capacidad o cota entregada, determinará el tamaño y el diseño de los cuarteles generados. El problema de esto es que, si la cota no es lo suficientemente grande, puede haber códigos postales que no pertenecerán a ningún cuartel.

2.4. Clusterización

La clusterización es una herramienta fundamental en la minería de datos. El objetivo de los algoritmos de clusterización es encontrar conjuntos de datos, tal que dentro de cada conjunto los datos sean similares, pero los datos entre grupos sean diferentes.

Formalmente, la estructura de clusterización es representada como un conjunto de conjuntos $C = C_1, \dots, C_k$ de S , tal que $S = \cup_{i=1}^k C_i$ y $C_i \cap C_j = \emptyset \forall i \neq j$.

Dado que la clusterización es un agrupamiento de objetos/datos similares, es necesario alguna medida que determine si dos objetos son similares o distintos. De acuerdo a Maimon and Rokach [7], existen 2 tipos de medidas utilizadas para estimar esta relación: medidas de distancias, y medidas de similitud.

1. Medidas de distancias: Es el método más común de clusterización, y utiliza medidas de distancia para determinar la similitud entre pares de objetos x_i y x_j ($d(x_i, x_j)$).
2. Funciones de similitud: Es un método alternativo que compara dos vectores x_i y x_j . Esta función debe ser simétrica ($s(x_i, x_j) = s(x_j, x_i)$), tomar valores altos para objetos similares, y el valor más alto para objetos idénticos.

Dentro de los métodos de clusterización están los métodos jerárquicos, que construyen los clusters recursivamente con la metodología top-down o bottom-up, y los métodos de particionamiento.

1. Método jerárquico:
 - Clustering jerárquico aglomerativo: Cada objeto es una partición por si sola. Luego, sucesivamente estas particiones se van juntando hasta obtener la estructura deseada.
 - Clustering jerárquico divisivo: Todos los objetos pertenecen al mismo cluster, y recursivamente se va subdividiendo cada cluster en sub-clusters hasta obtener la estructura deseada.
2. Método de particionamiento: Este método consiste en traspasar objetos de un cluster a otro, partiendo de una partición inicial. Por lo general se requiere que la cantidad de particiones esté predefinida por el usuario.

A continuación profundizaremos más en los métodos de particionamiento.

2.4.1. Métodos de particionamiento

Este método traslada, iterativamente, objetos entre los k clusters.

Dentro de los tipos de métodos de particionamiento, se encuentran:

1. Algoritmos de minimización de errores: Estos algoritmos son los más intuitivos y más utilizados. La idea principal es encontrar estructura de clusters que minimicen cierto error de distancia. El criterio de error más conocido es el de Suma de Errores al Cuadrado (SSE), que puede ser optimizado si se entrega una solución aproximada del problema.

Dentro de los algoritmos que utilizan el criterio SSE, el más conocido es el K-Means, propuesto por MacQueen en 1967, el cual particiona los datos en K clusters (C_1, \dots, C_k), representados por sus centros.

Este método comienza con K centros de cluster, y todos los datos se asignan al centro del cluster más cercano. Luego se actualiza el centro del cluster, y se vuelve a comenzar. Este método iterativo termina cuando de una iteración a otra, los cluster ya no cambian.

Este algoritmo genera zonas convexas y muy bien distribuidas, sin embargo, no se preocupa de equilibrar cargas.

El centro de cada cluster es calculado como el promedio de los datos que pertenecen al cluster:

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q$$

Un criterio de parada de este algoritmo es cuando el error de la partición no se reduce si se traslada el centro del cluster

2. Clustering en grafos: Estos son métodos que generan cluster mediante grafos. Los arcos del grafo conectan los objetos llamados nodos. Uno de los algoritmos más conocidos es el Árbol de expansión mínima.

Dentro de los modelos de particionamiento de grafos, se encuentra el Problema de Clustering con Capacidades (CCP), el cual tiene como objetivo encontrar p clusters con capacidades. Este modelo nació con la idea del diseño de territorios, incluso ha sido utilizado en el diseño de rutas postales por Laporte et al. [6].

Una segunda versión del problema de clusterización involucra particionar los nodos de un grafo $G = (N, E)$, en p clusters tal que se minimice la suma de los pesos de los arcos dentro de él.

Una extensión de este problema es el propuesto por Jarrah and Bard [4]. Donde, dado un número n de clientes localizados en una cuadrícula de dos dimensiones, se busca particionarlos

en un número dado de clusters, que satisfagan restricciones geométricas y que la cantidad de clientes por cluster no excedan una cota máxima.

Esta extensión toma en cuenta la naturaleza aleatoria de la demanda de cada cliente, sin embargo no es capaz de resolver óptimamente estas consideraciones en conjunto con la forma de los clusters, y su medición de balanceo entre particiones, es la cantidad de puntos de retiro y puntos de despacho contenidos en cada partición.

Para la heurística desarrollada se complementan diversos métodos de clusterización. Se trabajó aglomerando códigos postales sobre un grafo, además, para comenzar la generación de los cuarteles, se utilizará un método aglomerativo que comienza con K semillas y va juntando CPs de acuerdo a criterios de distancia y de carga. Finalmente se utilizará un método de particionamiento, que traspasa códigos postales de un cuartel a otro, para satisfacer la cota de carga deseada y requerimientos operacionales impuestos por la empresa Correos de Chile.

Capítulo 3

Descripción de la Heurística

En este capítulo presentamos el algoritmo que determina un grafo de la comuna y la heurística de generación de zonas de reparto equilibradas para Correos de Chile. Cabe destacar que la heurística propuesta en esta tesis no es una asignación óptima, sino que busca encontrar los cuarteles más equilibrados en carga de acuerdo a semillas entregadas. Estas semillas corresponden a los centroides iniciales de cada cuartel, es decir, que a partir de estas semillas se generan las zonas de reparto.

Para efectos de esta tesis, la función de carga de trabajo será un input del problema. Esta función se deriva de una serie de variables correspondientes a cada código postal, tales como su número de puertas, el tipo de infraestructura, distribución de demanda, entre otros, y entrega un valor continuo entre 1 y 10 correspondiente a la carga de dicho código.

Por lo tanto los datos iniciales son el conjunto de códigos postales, con sus coordenadas geográficas, su carga de trabajo y la cantidad K de cuarteles o zonas de reparto a generar.

Se recuerda que la instancia a resolver corresponde a las comunas de Viña del Mar y Concón, que poseen 11.831 códigos postales y 63 cuarteles.

3.1. Árbol de Conexión de Códigos Postales

En primera instancia se debe generar un grafo, de tal manera que todos los códigos postales (desde ahora CPs) de la comuna queden conectados, con conexiones que no rompan planaridad¹ y que no superen cierto criterio de distancia.

Para lo anterior desarrollamos un algoritmo en distintas etapas:

1. **Cuadrantes:** La totalidad de códigos postales de la comuna (11.831), se divide arbitrariamente en 400 cuadrantes imaginarios. Para esto tomamos el código más al norte,

¹Planaridad en un grafo: que puede ser dibujado en un plano de tal manera que las aristas solo se intersecten en los nodos.

más al sur, más al este y más al oeste para dimensionar el tamaño que debería tener la grilla.

Luego el alto de la grilla se dividió en 20 partes iguales, al igual que el ancho de la misma. Luego cada cuadrante corresponde al conjunto de códigos postales que quedan en el interior de él (Imagen 2 de la Figura 3.1).

Esta división en cuadrantes, fue necesaria ya que de lo contrario, el problema de generación de grafo era muy grande para resolverlo. De esta manera se genera el grafo por cuadrante, para posteriormente conectar todos los cuadrantes y generar el grafo final de la comuna.

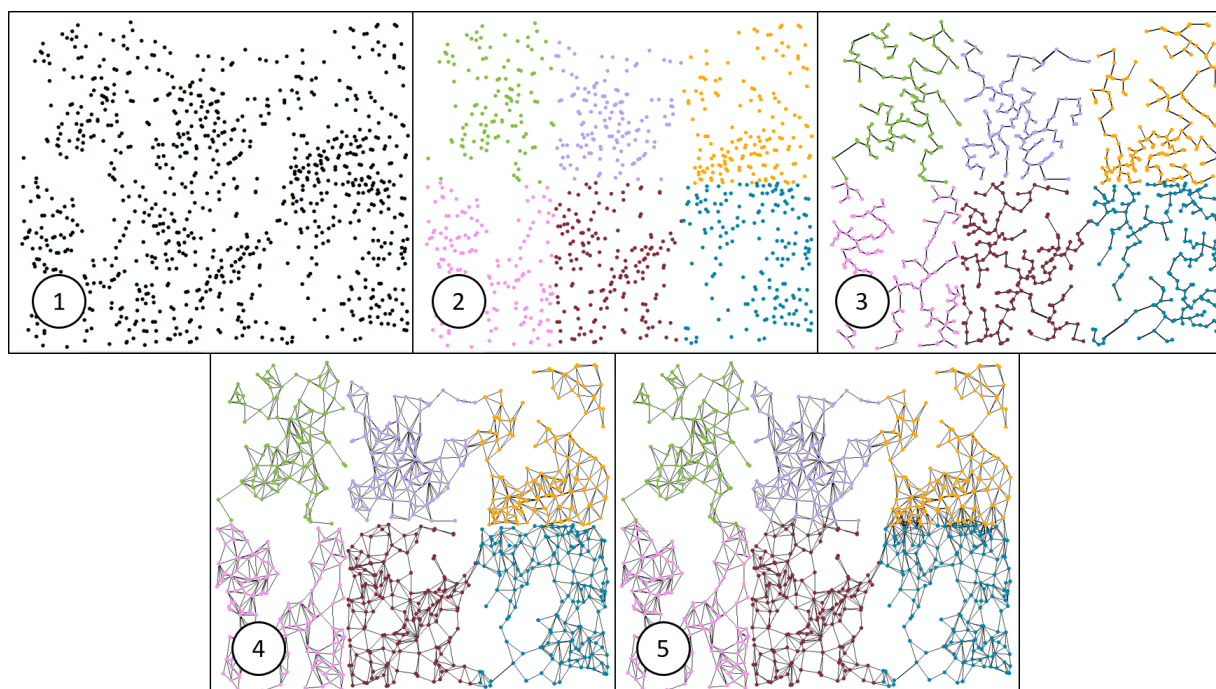


Figura 3.1: Se muestran los 5 pasos de construcción de un grafo que cumple criterios de distancia y que solo rompe planaridad dentro de las manzanas, ya que de esta manera quedan todos los códigos postales pertenecientes a la misma manzana conectados entre sí. Paso 1: los códigos postales iniciales, Paso 2: división de los códigos postales en cuadrantes, Paso 3: árbol de expansión en cada cuadrante, Paso 4: agregar aristas por cuadrante sin romper planaridad, Paso 5: conexión entre cuadrantes sin romper planaridad.

2. **Spanning Tree:** Cada uno de estos cuadrantes fue conectado mediante un árbol de expansión (Imagen 3 de la Figura 3.1), para asegurar que todos los códigos postales dentro del cuadrante estuvieran conectados.

El algoritmo es un método iterativo que logra conectar todos los CPs.

Algorithm 1 Spanning Tree

```
1: Conjuntos
2:  $A$  : conjunto de nodos que están en el árbol
3:  $X$  : conjunto de nodos que aún no se conectan al árbol ( $A \cap X = \emptyset$ )
4:  $G$  : conjunto total de arcos que conecta los nodos
5:  $E \subset G$  : lista de arcos que forman el árbol
6: Inicialización
7: while  $X \neq \emptyset$  do
8:    $E := E \cup \{z = (x, a) \in G \mid d_z \leq d_{(\bar{x}, \bar{a})} \forall \bar{x} \in X; \bar{a} \in A\}$ , se une al árbol el arco más
   corto que conecta un nodo libre con un nodo que pertenece al árbol
9:    $A := A \cup \{x\}$ , se agrega el nodo  $x$  al conjunto de nodos que están en el árbol
10:   $X := X \setminus \{x\}$ , se elimina el nodo  $x$  del conjunto de nodos libres
11: end while
12: return  $E$ 
```

3. **Planaridad en el cuadrante:** Para cada cuadrante se generan aristas que no superen una cota de distancia (Imagen 4 de la Figura 3.1), para generar arcos que representen la adyacencia real de los códigos postales. Dicha cota es igual a $\frac{\alpha}{\sqrt{n}}$, con $\alpha = 0,8$ y n equivalente al número de códigos postales (utilizada por Johnson et al. [5]), y que a la vez no rompan la planaridad del grafo (Figura 3.2).

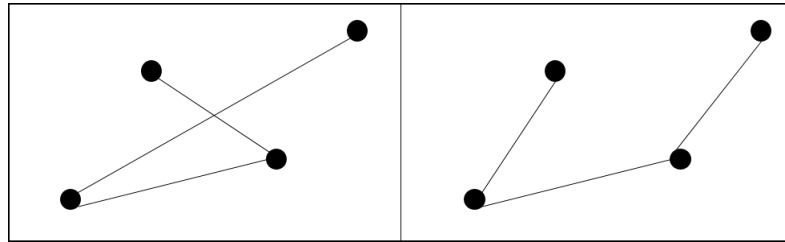


Figura 3.2: Izquierda: cuatro códigos postales unidos con arcos no planares. Derecha: solución deseable de planaridad

Se utiliza el siguiente criterio para ver si al conectar dos códigos postales (x_1, y_1) y (x_2, y_2) , la arista rompe planaridad:

Algorithm 2 Planaridad

```
1: Conjuntos
2:  $E$  conjunto de aristas
3:  $N$  conjunto de nodos
4: Inicialización
5: Sea  $n = (x_n, y_n)$ ,  $m = (x_m, y_m) \in N$ . Se busca ver si se agrega el arco  $(n, m)$  se rompe
   planaridad con los arcos de  $E$ 
6:  $s1_x := x_n - x_m$ 
7:  $s1_y := y_n - y_m$ 
8:  $intersecto = 0$ 
9: for  $\overline{AB} \in E$  do
10:   $s2_x := x_A - x_B$ 
11:   $s2_y := y_A - y_B$ 
12:  if  $s1_x * s2_y = s2_x * s1_y$  then
13:     $s := 0$  (no hay intersección)
14:     $t := 0$ 
15:  else
16:     $s = \frac{-s1_y * (x_2 - x_B) + s1_x * (y_2 - y_B)}{-s2_x * s1_y + s1_x * s2_y}$ 
17:     $t = \frac{s2_x * (y_2 - y_B) - s2_y * (x_2 - x_B)}{-s2_x * s1_y + s1_x * s2_y}$ 
18:  end if
19:  if  $0 < s < 1$  and  $0 < t < 1$  then
20:     $intersecto = 1$ 
21:    break
22:  end if
23: end for
24: return  $intersecto$ 
```

4. **Conexión entre cuadrantes:** En este paso se busca conectar todos los cuadrantes, para que así todos los CPs de la comuna queden conectados.

Para esto se conectó cada cuadrante con todos sus aledaños, agregando todas las aristas que no rompieran planaridad y que tampoco superen la cota de distancia establecida. Sin embargo se permitió romper planaridad dentro de las manzanas, ya que de esta manera todos los CP perteneciente a ella están conectados entre sí.

Finalmente si había cuadrantes que no quedaron conectados al resto de los cuadrantes (debido a que sus CPs superaban la cota de distancia), se conectaron mediante la arista más corta al cuadrante más cercano (Imagen 5 de la Figura 3.1).

El proceso completo de construcción del grafo que será utilizado para la implementación de la heurística que genera zonas de reparto equilibradas, se puede ver en la Figura 3.1. El grafo obtenido es estático para la comuna en cuestión. Este grafo dependerá de las cotas de distancia establecidos, el orden en que se conectan las aristas, y el nodo inicial con que se comienza la construcción del árbol. La construcción de este grafo determinará los cuarteles, ya que determina la adyacencia entre códigos postales.

3.2. Clustering

3.2.1.

Para el diseño de las zonas de reparto de los carteros de Correos de Chile, la empresa tuvo los siguientes requerimientos que debía satisfacer la heurtística a implementar:

1. Cuarteles conexos: todos los códigos postales pertenecientes al mismo cuartel, deben estar conectados de tal manera que, para ir de un CP a otro, no sea necesario atravesar otro cuartel (Figura 3.3).

8	10	9	1	4	8	6	5	3	1	1	9	9	4	6	2	10	6	5	1	5	1	3	10	6	1	7	6
9	8	7	7	6	10	6	2	3	10	8	1	2	5	8	1	2	8	10	7	5	2	4	7	10	7	3	10
3	6	5	2	4	7	10	4	1	2	8	7	6	8	3	7	9	1	3	10	10	1	8	4	2	8	8	6
4	6	8	10	6	2	2	6	6	9	8	6	7	5	1	9	1	1	4	4	7	5	8	4	3	10	6	
3	4	3	10	4	1	2	9	1	2	9	7	10	6	3	2	9	7	6	4	5	7	7	5	7	1	9	
1	1	6	10	1	6	6	2	3	8	2	7	1	9	8	8	5	6	10	10	2	3	10	3	5	5	1	2
3	1	4	3	8	7	9	8	6	3	9	3	9	6	1	1	7	8	10	6	5	2	10	8	2	3	5	5
7	5	6	8	2	1	5	9	8	1	4	6	10	6	9	6	9	9	10	5	1	2	9	10	4	4	6	1
10	5	1	7	3	7	5	4	6	2	4	5	4	7	10	6	4	10	2	8	3	5	4	1	10	10	3	6
4	4	2	1	1	10	6	7	2	10	9	1	5	3	7	6	2	4	5	2	4	4	5	2	5	1	7	4
8	7	1	10	10	3	8	3	7	4	3	4	4	4	2	6	9	6	5	9	6	9	3	4	2	6	2	8
2	9	10	3	3	9	7	2	5	2	5	8	8	6	8	2	8	1	7	8	6	7	10	9	9	2	8	2
1	8	5	4	4	6	5	10	1	9	5	5	10	9	7	5	7	4	3	7	2	5	2	4	8	9	6	5
5	4	7	1	2	5	7	2	4	6	5	4	5	4	5	1	8	1	8	10	3	6	2	8	2	6	10	9
3	3	2	3	9	9	9	4	3	3	10	7	9	8	2	3	4	6	8	7	8	4	4	5	2	2	7	9
10	6	2	10	2	4	10	2	3	6	7	5	6	3	1	5	3	9	3	4	9	10	1	4	4	1	10	3
8	5	4	10	3	4	10	10	2	8	1	5	9	5	6	8	9	3	9	10	1	7	4	1	4	9	2	5

Figura 3.3: Ejemplo de solución no deseable de cuarteles disconexos. Cada celda corresponde a un código postal con su respectiva carga (número de la celda). El cuartel azul no cumple el requerimiento de conectividad deseado.

2. Carga de trabajo equilibrada entre cuarteles: es decir que dada alguna función de carga, la carga entre cuarteles (corresponde a la suma total de carga de los códigos postales que lo componen) sea homogénea.
3. Una manzana pertenece a un solo cuartel: esto implica que todos los códigos postales que pertenecen a una misma manzana, deben pertenecer a la misma zona de reparto (Figura 3.4). Si bien esta restricción no debiese ser inviolable, en la realidad, operacionalmente es muy deseable.



Figura 3.4: Ejemplo de solución no deseable de cuarteles, cuyas manzanas pertenecen a dos zonas de reparto diferentes.

Por lo tanto, previo a la generación de los cuarteles, hubo un paso adicional que era transformar códigos postales en manzanas, es decir, que desde ahora no se trabajaría más con los CPs.

De esta manera se trabajará con un grafo en el cual los vértices son manzanas, y las aristas representan las vecindades entre manzanas.

Para esto se definieron las coordenadas de la manzana m como el centro de gravedad de los códigos postales que las conforman, y se consideró el peso de cada CP igual a 1.

$$x_m = \frac{\sum_{CP \in m} x_{cp}}{|CP \in m|}; \quad y_m = \frac{\sum_{CP \in m} y_{cp}}{|CP \in m|}$$

Para la generación de las aristas, se conectó la manzana m con todas las manzanas que tienen códigos postales conectados a algún CP de la manzana m (Figura 3.5).

De esta manera se tendrá inmediatamente el requerimiento número 3 de la empresa.

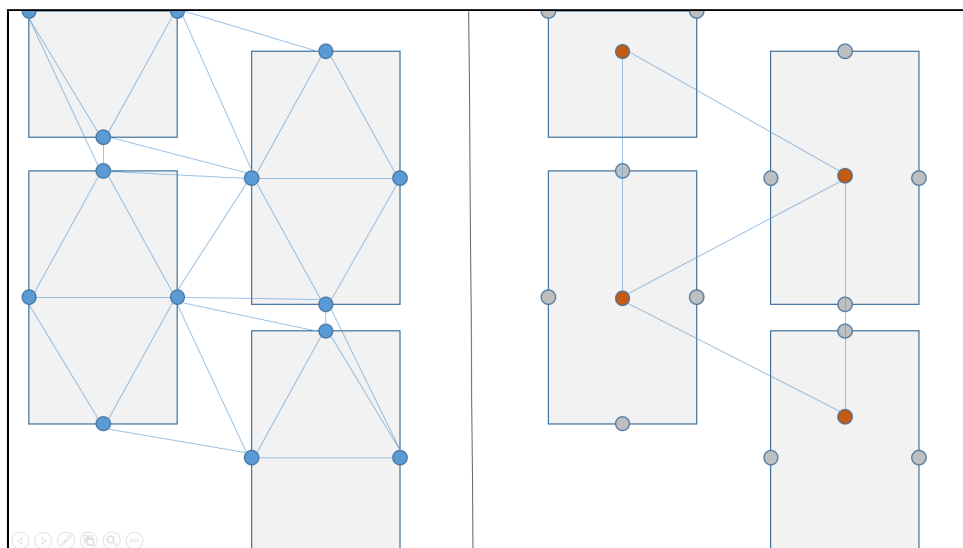


Figura 3.5: Transformación de CPs a manzanas: En la imagen de la izquierda se encuentran las manzanas con sus códigos postales (círculos azules), y los arcos que los conectan. En la imagen de la derecha se encuentran las manzanas representadas por su centro de gravedad (círculos rojos), y los arcos que las conectan entre ellas.

Con la transformación de los 11.831 códigos postales a las 6.798 manzanas, y el grafo que las conecta, se puede proceder a la construcción de los 63 cuarteles para las comunas de viña del Mar y Concón. Para estas instancias se utilizó 300 iteraciones y un GAP del 25%. Esta construcción consta de dos pasos:

1. Generador de cuarteles
2. Polishing

El primero paso consiste en asignar todas las manzanas a un cuartel, de tal manera que éste quede conexo y lo más convexo posible ². Esto es deseable, ya que, de esta manera el cartero no debe atravesar por otro cuartel para satisfacer el suyo (Figura 3.6).

El segundo paso de la generación de cuarteles consiste en una etapa de redistribución de algunas manzanas entre cuarteles, para que éstos queden con carga equilibrada.

3.2.2. Generador de Cuarteles

Este procedimiento es un algoritmo constructivo aleatorizado³, que comienza con un conjunto de K manzanas, al igual que el algoritmo de particionamiento K-means 2.4.1, que representan las "semillas" de cada uno de los cuarteles, y finalmente entrega una lista de

²Convexidad: Para cualquier par de puntos de una figura geométrica, la recta que los une está completamente contenida en esta.

³Algoritmo constructivo aleatorizado: algoritmo que comienza con una solución vacía, y va extendiendo la solución aleatoriamente, hasta que se obtiene una solución completa.

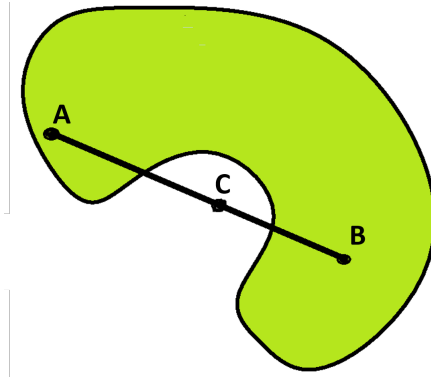


Figura 3.6: En la imagen se muestra un cuartel no deseable (cóncavo), ya que el cartero podría preferir ir del punto A al punto B a través del punto C, el cual pertenece a otro cuartel, y por lo tanto, a otro cartero.

cuarteles con sus respectivas manzanas asociadas y sin dejar manzanas sin un cuartel asignado.

Para efectos de esta tesis se utilizó un conjunto de semillas que, arbitrariamente, se encontraran dispersas en el plano. Estas semillas influyen en la forma y tamaño de los cuarteles finales.

La idea principal de este algoritmo es ir agregando secuencialmente, dado un orden aleatorio (en cada iteración se ordenan aleatoriamente los cuarteles), manzanas a cada cuartel, hasta que se llegue a una cota máxima de carga asociada a todo el cuartel. La cota viene dada por:

x : Código Postal

C : Comuna

Q_x : Carga de trabajo del CP

P : Cantidad de cuarteles a construir

Cota: Corresponde al cociente de la carga total de los códigos postales con la cantidad de cuarteles a generar, multiplicado por un margen de holgura, que para efectos de esta tesis se utilizó el 5%.

$$Cota = \frac{\sum_{x \in C} Q_x}{P} * 1,005$$

Para esto se recorre la lista aleatoria de cuarteles definida anteriormente, y se asigna la manzana más cercana al centro de gravedad de dicho cuartel y que no haya sido asignada. Luego de dicha asignación se actualiza el centro de gravedad del cuartel y la lista de manzanas vecinas del mismo, y se realiza el mismo procedimiento con el cuartel siguiente.

Este procedimiento se repite hasta que todos los cuarteles completaron la cota de carga máxima o no podían seguir creciendo (caso en que todas las manzanas vecinas ya estaban asignadas a otro cuartel), luego comienza un post proceso: se agregan todas las manzanas,

que aun no pertenecen a ninguna zona de reparto, al cuartel cuyo centro de gravedad sea el más cercano.

De esta manera todas las manzanas quedan seleccionadas en un cuartel, sin embargo el post-proceso permite que la carga de los cuarteles no quede equilibrada, por lo que surge la necesidad del siguiente paso de la heurística (Polishing).

Algorithm 3 Generador de cuarteles

```

1: Conjuntos
2:  $k$ : número de cuarteles
3:  $C_i = \{m_j\}$ : manzana  $m_j$  asignada al cuartel  $i$ ,  $C_i \cap C_l = \emptyset$ , todos los cuarteles comienzan con la manzana más cercana a la semilla
4:  $N = \cup_{i=1}^k C_i$ : conjunto de manzanas  $m_j$  asignadas a algún cuartel
5:  $M$ : conjunto de manzanas  $m_j$  sin ser asignadas
6:  $V_i$ : conjunto de manzanas  $m_j$  vecinas del cuartel  $i$  que no han sido asignadas
7:  $W_j$ : conjunto de manzanas vecinas de la manzana  $m_j$ 
8:  $Q_i$ : carga del cuartel  $i$  (suma de la carga de las manzanas pertenecientes al cuartel)
9:  $q_j$ : carga de la manzana  $j$ 
10: Inicialización
11: while True do
12:    $largo = |M|$ 
13:   for  $i = 1$  to  $k$  do
14:      $Q_i = \sum_{j|m_j \in C_i} q_j$ 
15:     if  $V_i \neq \emptyset$  and  $Q_i < cota$  then
16:        $C_i = C_i \cup \{m_j\}$ , con  $m_j \in V_i$  la manzana de menor distancia al centroide de  $C_i$ 
17:        $M = M \setminus \{m_j\}$ 
18:        $V_i = V_i \setminus \{m_j\} \cup (W_j \cap M)$ 
19:     end if
20:   end for
21:   if  $|M| = largo$  then
22:     Break
23:   end if
24: end while
25: for  $j = 1$  to  $|M|$  do
26:    $C_i = C_i \cup \{m_j\}$ , con  $C_i$  el cuartel con centroide más cercano a  $m_j$ 
27:    $M = M \setminus \{m_j\}$ 
28:    $V_i = V_i \setminus \{m_j\} \cup (W_j \cap M)$ 
29: end for
30: return  $N$ 

```

3.2.3. Polishing

Esta parte de la heurística tiene como objetivo la redistribución de las manzanas, de tal manera que se logre equilibrar la carga de los cuarteles.

La idea es que en cada iteración del *Polishing*, nos fijemos en el cuartel de mayor carga, generando una onda de propagación de manzanas, es decir, el cuartel de mayor carga entrega manzanas secuencialmente a todos sus vecinos hasta cumplir con la cota de carga definida anteriormente. Luego sus vecinos, de igual manera, entregan manzanas a sus respectivos vecinos que no son el cuartel inicial. Así sucesivamente hasta que todos los cuarteles han sido recorridos.

Esta etapa del *Polishing*, si bien no asegura homogeneidad en una pasada por todos los cuarteles, permite que la carga se vaya distribuyendo a lo largo de toda la comuna en cuestión. Por lo tanto es necesario generar varias veces esta onda de propagación, hasta que la carga quede lo más homogénea posible.

La heurística se detiene cuando se cumple una cantidad dada de iteraciones, o hasta que el gap entre el cuartel de mayor carga, con el de menor carga, sea menor a 25 %.

$$GAP = \frac{Q_M - Q_m}{Cota}$$

Con Q_M la carga del cuartel de mayor carga y Q_m la carga de trabajo del cuartel de menor carga.

Cada iteración del *Polishing* funciona de la siguiente manera:

1. Se comienza con el cuartel de mayor carga, y se genera una lista ordenada de todos los cuarteles según un indicador que dice el nivel de cercanía de dicho cuartel con el cuartel de mayor carga (Figura 3.7). El primer nivel (comenzando en el número 0) corresponde al cuartel de mayor carga, el segundo nivel de vecino corresponde a los cuarteles vecinos del cuartel principal (cuarteles con número 1). Luego el tercer nivel son los vecinos de los cuarteles perteneciente al segundo nivel, y así sucesivamente.



Figura 3.7: Se muestra el cuartel con mayor carga (0) y sus cuarteles vecinos con el indicador de la cercanía con él.

2. Nos paramos en el cuartel del nivel 0 (el de mayor carga), y entregamos secuencialmente, a todos sus cuarteles vecinos, la manzana más lejana al centroide del cuartel 0, que esté conectada al vecino correspondiente (1 manzana por cuartel vecino). Si se alcanza la cota de carga definida anteriormente pasamos al siguiente paso, si nose vuelve a recorrer los cuarteles de vecinos (sin haber actualizado la lista de vecinos).
3. Ahora nos paramos en el siguiente nivel de vecinos, es decir si estábamos en el nivel n vamos a el nivel $n + 1$, y uno por uno nos paramos en un cuartel de este nivel. De igual manera que con el cuartel inicial, este cuartel entrega manzanas a todos sus cuarteles vecinos hasta cumplir la cota de carga (Figura 3.8), sin embargo, no entrega manzanas a los vecinos que ya han repartido manzanas anteriormente. Acá puede ocurrir 3 casos:
 - (a) Entrega manzanas a un cuartel de un nivel vecino superior a él (si está en el nivel $n + 1$ y entrega a un cuartel del nivel $m \geq n + 2$). En este caso no hay problema, ya que aún no se recorren los niveles superiores.
 - (b) Entrega manzanas a un cuartel del mismo nivel de vecino (si está en el nivel $n + 1$ y entrega a un cuartel del nivel $m = n + 1$). Esto puede ocurrir porque aún no se recorre este segundo cuartel, o porque ya se pasó anteriormente por él, pero este no tenía manzanas a repartir ya que tenía menos carga que la cota definida. En este caso, una vez que termine con el cuartel que se está recorriendo, volvemos a recorrer todos los cuarteles del nivel de vecinos actual, para ver si alguno de los cuarteles quedó con mayor carga que la cota.
 - (c) Entrega manzanas a un cuartel de un nivel de vecinos anterior (si está en el nivel $n + 1$ y entrega manzanas a un cuartel del nivel $m \leq n$). Esto ocurre porque, aún cuando el cuartel receptor, recibió de todos sus vecinos anteriores o de su mismo nivel, nunca tuvo más carga que la cota definida, por lo que nunca repartió manzanas. En esta circunstancia, una vez que termine con el cuartel que se está recorriendo, se vuelve al nivel anterior y comenzamos a iterar nuevamente.

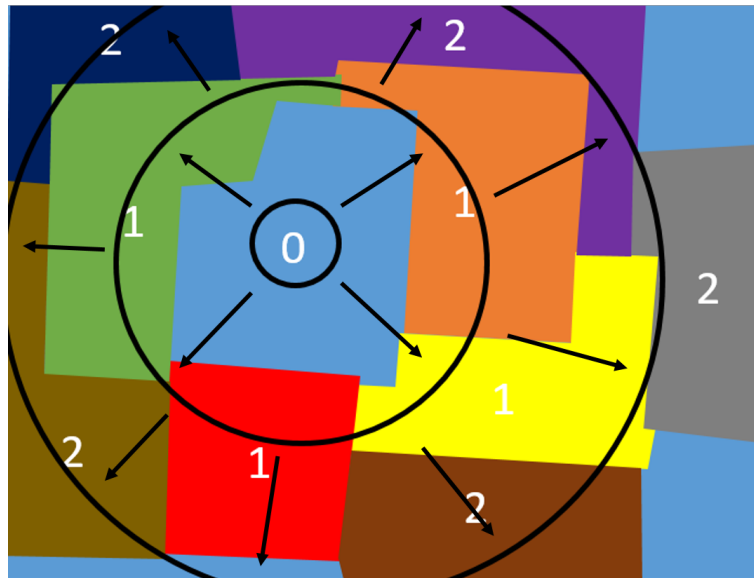


Figura 3.8: Se muestra como se va propagando la carga, desde el cuartel de mayor carga hacia afuera.

4. Una vez que se han recorrido todos los niveles de vecinos, se verifica si con el traspaso de manzanas entre cuarteles, se rompe la restricción de conectividad dentro del cuartel. Si no se rompe, entonces se verifica si se cumple el gap entre el cuartel de mayor carga con el de menor carga (cociente entre la carga del cuartel de menor carga de trabajo con el de mayor carga), o si se completaron las N iteraciones definidas. En cualquiera de los dos casos, se termina con la heurística, de lo contrario se comienza con la siguiente iteración y se vuelve al paso 1.

Sin embargo, si se rompió la conectividad de algún cuartel, entonces se procede a un proceso de entrega de manzanas entre cuarteles. Este funciona de la siguiente manera:

- (a) Para cada cuartel que rompe con la restricción, se enumeran las vecindades⁴ del cuartel.
- (b) Ver cual de las vecindades es la más grande (la que contiene mayor número de manzanas), y esa representará el cuartel. Para las otras vecindades, se toman cada una de las manzanas y se asignan a su cuartel vecino (cuartel que contenga alguna manzana vecina de la manzana en cuestión) cuyo centroide es el más cercano, sin importar la carga de dicho cuartel.
- (c) Una vez recorridas todas las vecindades de todos los cuarteles que rompieron la restricción, estamos seguros que se satisface la restricción de conectividad, sin embargo, la carga nuevamente se pudo haber desbalanceado. Por lo que se verifica si se cumple el gap entre el cuartel de mayor carga con el de menor carga, o si se completaron las N iteraciones definidas y se vuelve al paso 1 para comenzar una nueva iteración.

Es importante notar que existe un trade-off entre equilibrar cargas y generar cuarteles conectados, tal como se describió en el último paso del algoritmo. Este paso, perjudica el equilibrio para poder cumplir con los requerimientos de la empresa, sin embargo, a lo largo de las iteraciones del *Polishing* este desajuste es cada vez menor, pero no se garantiza que deje de ocurrir en algún momento, por lo que no garantiza el óptimo del problema.

Una vez terminadas las iteraciones o alcanzado el gap, se tienen los cuarteles con sus respectivas manzanas. Ahora es necesario volver a la transformación realizada al comienzo, donde se comenzó a trabajar con manzanas en vez de los CPs, y así tener los códigos postales pertenecientes a cada cuartel y el grafo original.

⁴Vecindad: Conjunto de manzanas pertenecientes a un mismo cuartel, y que se encuentran conectadas entre sí. Es decir, si para un cuartel existe una sola vecindad, entonces este cumple con la restricción de conectividad.

Algorithm 4 Polishings

```
1: Conjuntos
2:  $M = \{m : m \text{ es manzana}\}$  : conjunto de todas las manzanas
3:  $Q = \{C_i : C_i \text{ es cuartel}\}$  : conjunto de todos los cuarteles
4:  $E = \{(m, n) : m, n \in M\}$  : conjunto de todas las conexiones entre manzanas
5:  $C_i = \{m : m \in M\}$ : conjunto de manzanas pertenecientes al cuartel i
6:  $V_i = \{C_j : \exists m \in C_j \wedge \exists n \in C_i | (m, n) \in E\}$ : cuarteles vecinos del cuartel i
7:  $W_m = \{n : n \in M | (m, n) \in E\}$ : conjunto de manzanas vecinas de la manzana m
8:  $Nivel_0 = \{C_i : q_i = \max_{j \in Q} q_j\}$ , el cuartel de mayor carga
9:  $Nivel_n = \{C_i : \exists C_j \in Nivel_{n-1} | C_i \in V_j\}$ , conjunto de cuarteles tal que existe un cuartel
    vecino en el nivel anterior
10:  $T_i$  es el conjunto de vecindades del cuartel i
11: Inicialización
12:  $K$ : número de total cuarteles
13:  $P$ : número de iteraciones
14:  $Nivel$ : número de niveles de vecinos del cuartel de mayor carga
15:  $R = \{\}$ : conjunto de cuarteles recorridos en la presente iteración
16: anterior: 1 si el cuartel entrega manzanas a un cuartel del nivel anterior, 0 en caso
    contrario
17: anterior = 0
18: for  $p \in P$  do
19:   (*)
20:   for  $n \in Nivel$  do
21:     for  $C_i \in Nivel_n$  do
22:        $R = R \cup \{i\}$ 
23:       while  $q_i > cota$  do
24:         for  $C_j \in V_i$  do
25:           if  $C_j \in Nivel_{n-1}$  and  $j \notin R$  then
26:             anterior = 1
27:              $C_j = C_j \cup \{m\} : m \in C_i$  es la más lejana al centroide de  $C_i \wedge W_m \cap C_j \neq \emptyset$ 
28:              $C_i = C_i \setminus \{m\}$ 
29:             Actualizo  $V_i$  y  $V_j$ 
30:           else if  $C_j \in Nivel_{x:x \geq n}$  and  $j \notin R$  then
31:              $C_j = C_j \cup \{m\} : m \in C_i$  es la más lejana al centroide de  $C_i \wedge W_m \cap C_j \neq \emptyset$ 
32:              $C_i = C_i \setminus \{m\}$ 
33:             Actualizo  $V_i$  y  $V_j$ 
34:           end if
35:         end for
36:       end while
37:       if anterior = 1 then
38:         Volver a (*)
39:       end if
40:     end for
41:   end for
42:   for  $C_i \in Q$  do
43:     Actualizo  $T_i$ 
44:     if  $|T_i| > 1$  then
45:       Cedo las manzanas de cada vecindad al cuartel cuyo centroide es el más cercano
46:     end if
47:   end for
48: end for
```

Capítulo 4

Resultados

En este capítulo se mostrarán los resultados obtenidos con el algoritmo descrito en el capítulo anterior, para la comuna de Viña del Mar y Concón. Viña del Mar es una comuna de la V región, cuenta con una superficie de 121,6 km cuadrados, tiene una población de 290.000 habitantes y se encuentra a 120 kilómetros de Santiago.

La comuna cuenta con zonas urbanas y rurales, además de zonas con distintos estratos económicos. Por otra parte, la comuna tiene sectores con distinta densidad poblacional Fig.4.1 y distintos tipos de viviendas e inmuebles (Fig.4.2).

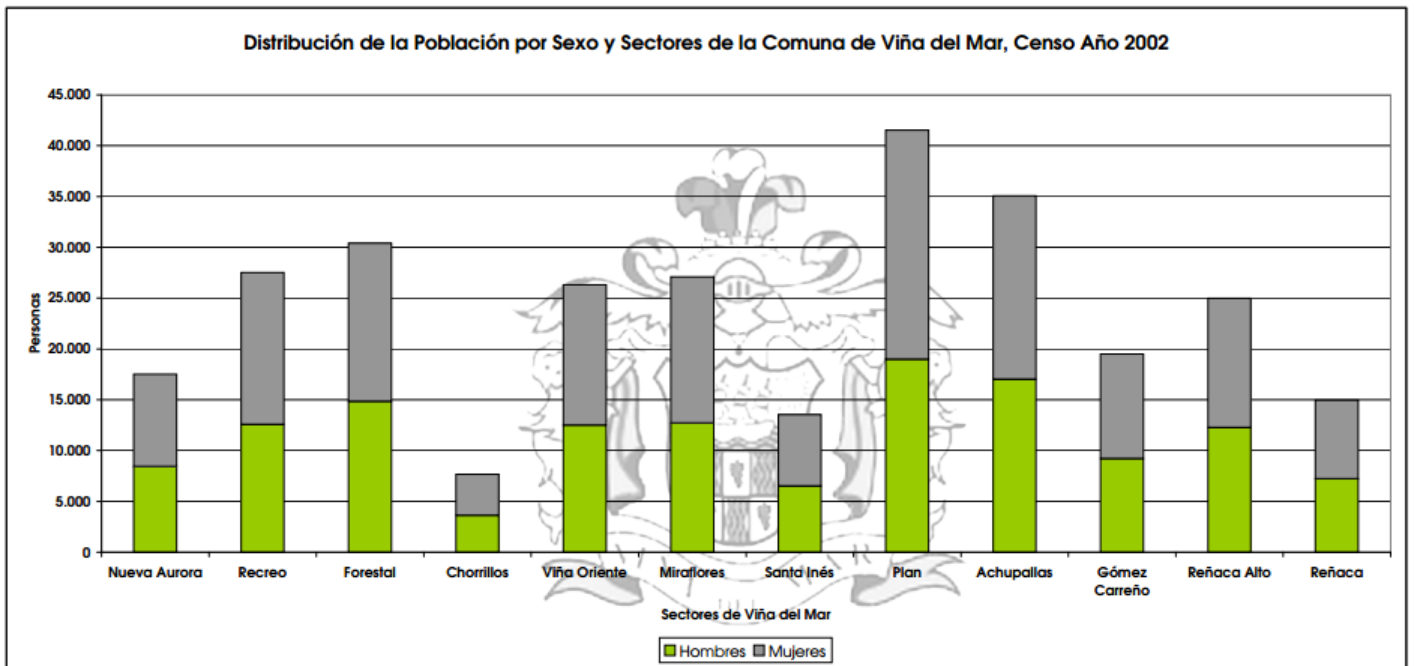


Figura 4.1: Fuente: Censo 2002. <http://www.vinadelmarchile.cl>

Esta variedad hace de la comuna una zona interesante para evaluar la heurística descrita anteriormente.

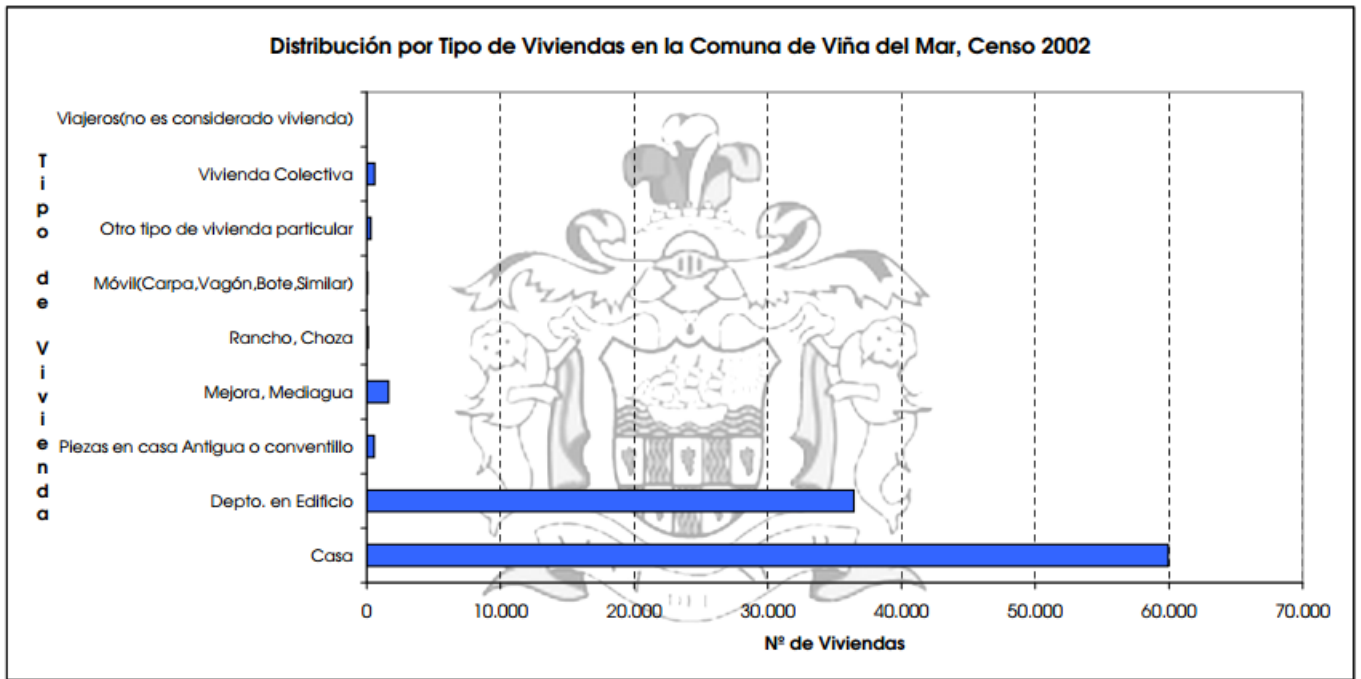


Figura 4.2: Fuente: Censo 2002. <http://www.vinadelmarchile.cl>

Los cuarteles actuales de estas comunas, han sido modificados a lo largo de su historia de manera parcial, nunca de manera integral, sin ningún apoyo de sistemas computacionales, lo que ha generado zonas de reparto poco eficientes, donde las cargas de trabajo entre zonas de reparto tiene gran varianza (Figura 1.2). Esta configuración merma la calidad de servicio de algunas zonas, ya que los tiempos promedios de distribución dependerán del cuartel a la cual correspondan las cartas a entregar.

Si bien esto ocurre en Viña del Mar, este patrón se repite a lo largo de todas las comunas de Chile.

4.1. Caso de estudio: Comuna Viña del Mar y Concón

Esta instancia cuenta con 11.831 códigos postales, los cuales deben particionarse en 63 cuarteles, correspondientes al número de cuarteles actuales de las comunas.

4.1.1. Generación del grafo

A partir de la posición geográfica de los códigos postales, se generaron los cuadrantes de la comuna de Viña del Mar, que en este caso se eligió arbitrariamente 400 (grilla de 20x20). Figura 4.3.

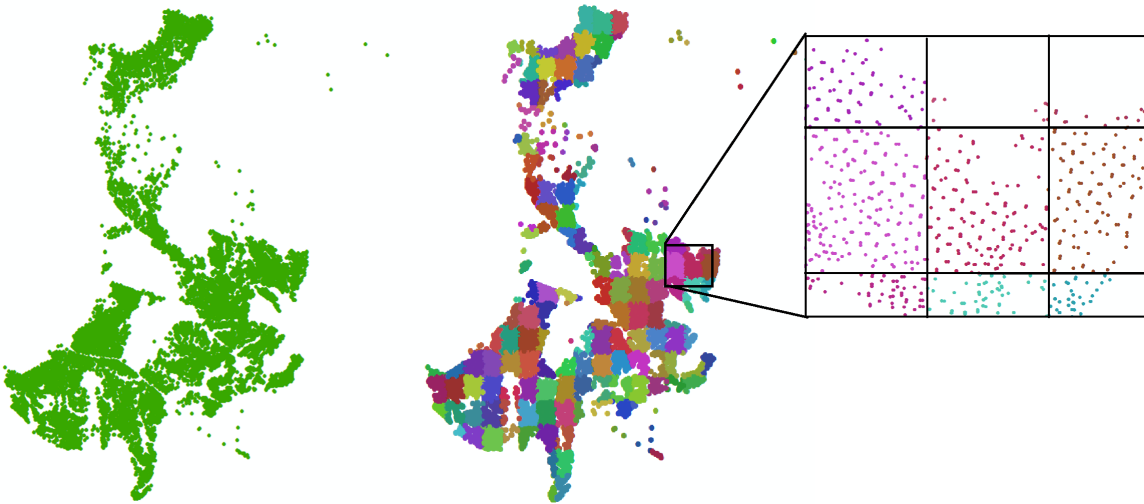


Figura 4.3: A la izquierda están los códigos postales de Viña y Concón, y a la derecha se encuentran los mismos códigos postales y el color representa el cuadrante al que pertenecen.

A partir de los 400 cuadrantes se generó el grafo para esta comuna, a partir del algoritmo descrito en la Sección 3.1:

1. Se genera árbol de expansión en cada cuadrante

2. Se agregan aristas a cada cuadrante, que cumpla con restricciones de planaridad y distancia

3. Se conectan cada cuadrante con todos sus cuadrantes vecinos adicionando todas las aristas que cumplan ambas restricciones mencionadas en el punto anterior

De esta manera se obtuvo el grafo que será utilizado para la generación de los cuarteles de la comuna de Viña del Mar. (Figura 4.4).

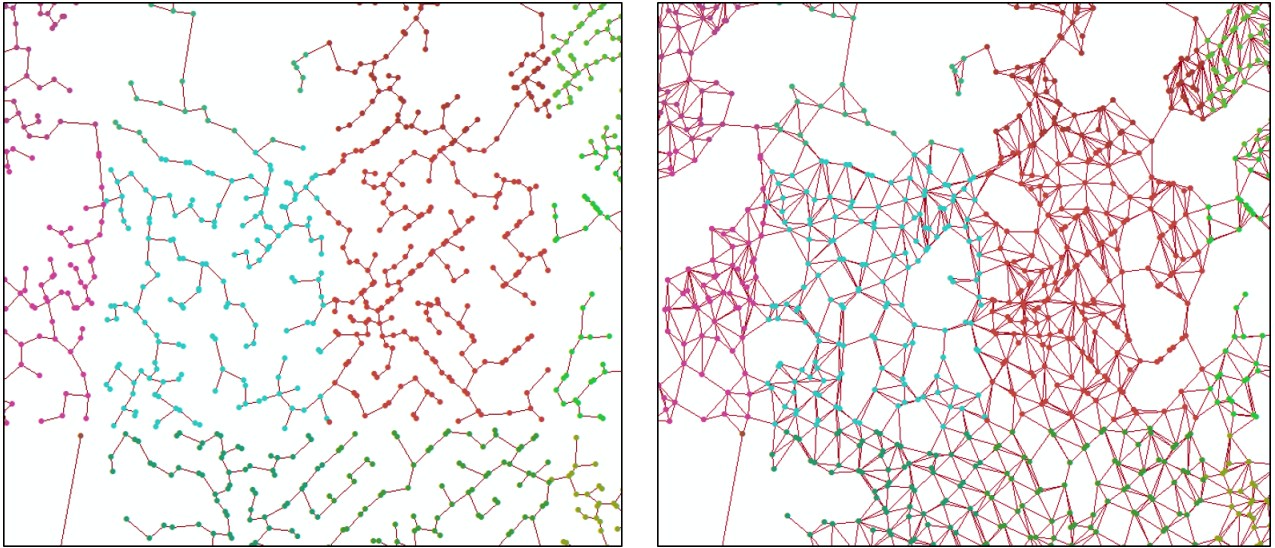


Figura 4.4: Árbol de expansión en cada cuadrante (imagen de la izquierda), y conexiones entre ellos (imagen de la derecha).

4.1.2. Función de Carga

Luego, para la carga de los códigos postales se utilizó una función que depende de características propias del CP:

- Tráfico de productos por cada puerta del CP
- Tiempo estimado de distribución

Como se mencionó anteriormente, la función de carga fue recibida como input. Para la elaboración de dicha función se utilizaron mediciones de tiempo, se realizaron en diferentes cuarteles, y se midió el tiempo que demora un cartero entregando las cartas en cada puerta perteneciente a su cuartel. También se midió los tiempos de desplazamiento del cartero dentro del código postal, es decir, el tiempo que demora recorriendo entre las puertas pertenecientes al mismo CP.

A partir de esto, se tomó el promedio del tráfico anual (del año 2016) de cartas entregadas en cada puerta, y se calculó la participación relativa de cada tipo de producto en el tráfico anual de la puerta en cuestión.

Por otra parte, con los datos de mediciones de tiempo, se tiene el tiempo que demora un cartero en entregar cada tipo de producto. Este número se multiplicó por la participación del producto en el tráfico de cartas. Luego se sumó el resultado con el de todos los tipos de cartas.

Posteriormente se sumó este número para todas las puertas de un código postal, para que finalmente se normalizara este valor para que quedaran valores entre 1 y 10, que corresponden a la carga del código postal.

Así se obtuvo una función que depende del tráfico de cartas, la cantidad de puertas de cada código postal, y características propias del CP, provenientes de las mediciones de tiempo entregadas por Correos de Chile.

Con esta caracterización de los códigos postales, los cuarteles quedan con la distribución de carga de la Figura 4.5.

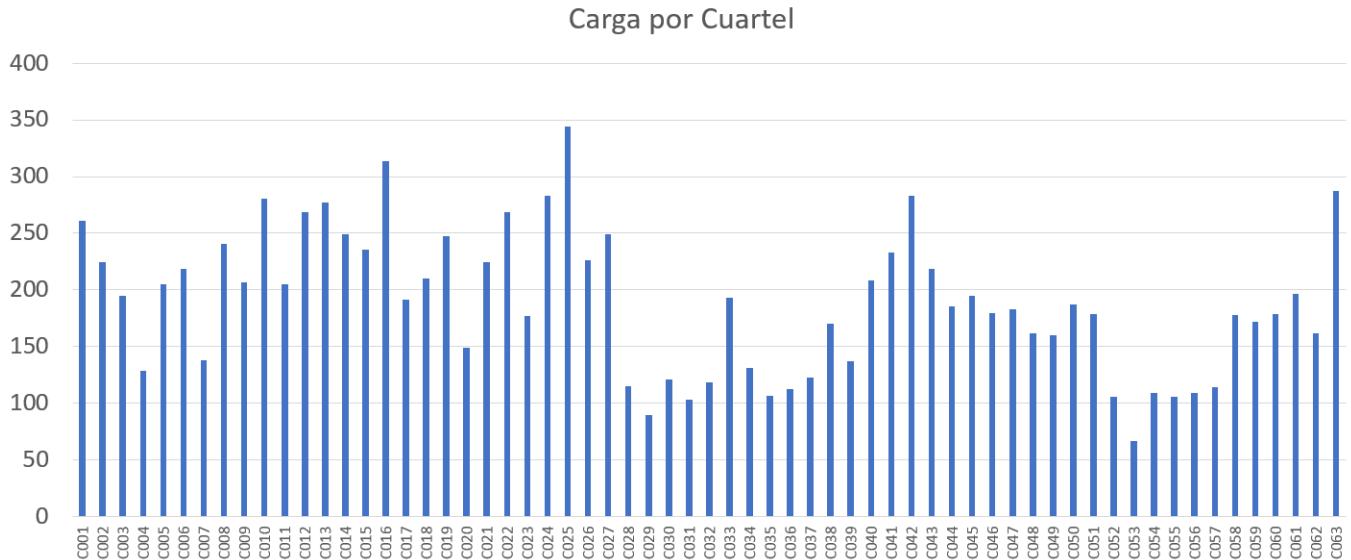


Figura 4.5: Distribución de carga (eje Y) de los cuarteles (eje X) de las comunas de Viña del Mar y Concón. Tiene un GAP del 131 %, un promedio de carga de trabajo equivalente a 188,7 y una desviación estándar de 62,3.

Hay que recordar que el objetivo de esta tesis no es determinar una buena función de carga, si no, que se considerará como un input del problema.

4.1.3. Generador de Cuarteles

Luego, con esta carga, se procedió a la generación de los cuarteles, mediante la heurística anteriormente descrita, donde se estableció que la cota de carga máxima que debía cumplir cada cuartel de la comuna, para esta instancia, corresponde a:

$$cota\ carga = \lfloor \frac{\sum_{CP} carga_{CP}}{63} * 1,005 \rfloor = 213$$

$$GAP = \frac{Q_M - Q_m}{213}$$

Para generar los cuarteles se utilizaron semillas distribuidas a lo largo de toda la comuna. Estas semillas fueron seleccionadas arbitrariamente, de tal manera que quedaran homogéneamente distribuidas. (Figura 4.6)

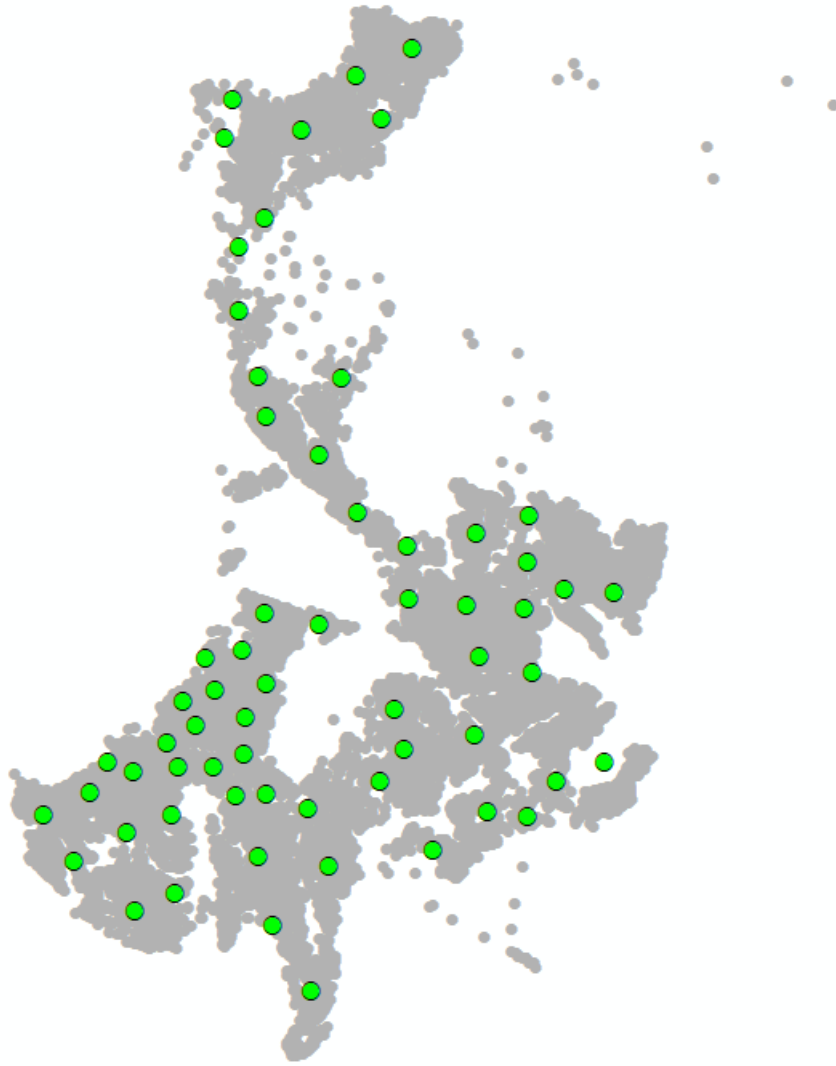


Figura 4.6: En gris se encuentran los códigos postales de las comunas de Viña del Mar y Concón, y en verde están las semillas utilizadas en la heurística.

El gap obtenido entre los cuarteles generados es menor al 25 % (Figura 4.7), logrado en 33 iteraciones del *Polishing*, y adicionalmente se observa que se cumplen los requerimientos de la empresa de generar cuarteles conexos, y donde todos los códigos postales de una manzana pertenezcan al mismo cuartel (Figura 4.8).

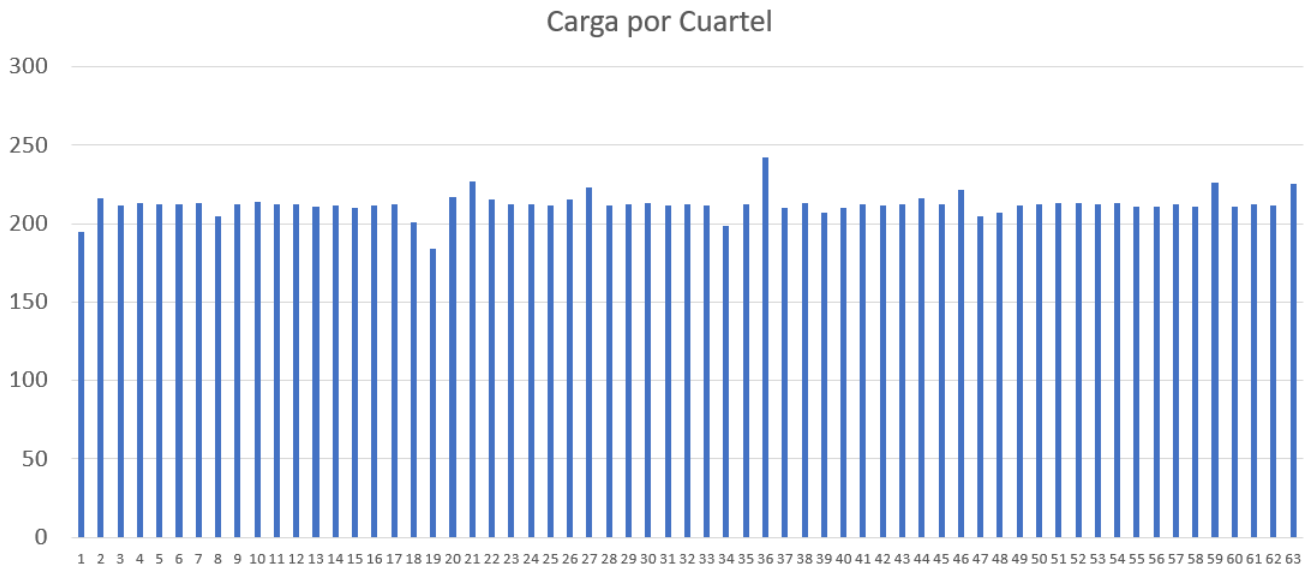


Figura 4.7: Distribución de carga de los cuarteles generados con el algoritmo. GAP de 24%, promedio de carga de trabajo igual a 212,3 y con una desviación estándar de 7,4.

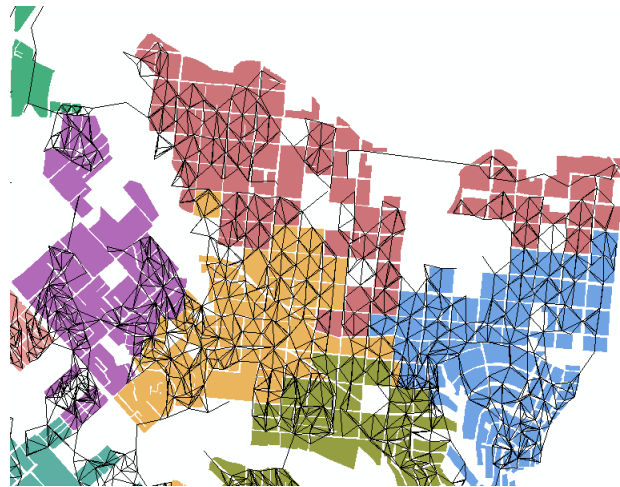


Figura 4.8: Cuarteles generados con algoritmo. Hay conexidad de cuarteles, y todos los CP de una manzana pertenecen al mismo cuartel.

La heurística de demoró 2,7 minutos en recorrer las 33 iteraciones, y en la Figura 4.9 se puede ver cómo evolucionó el GAP de la solución a medida que se avanzaba en ellas.

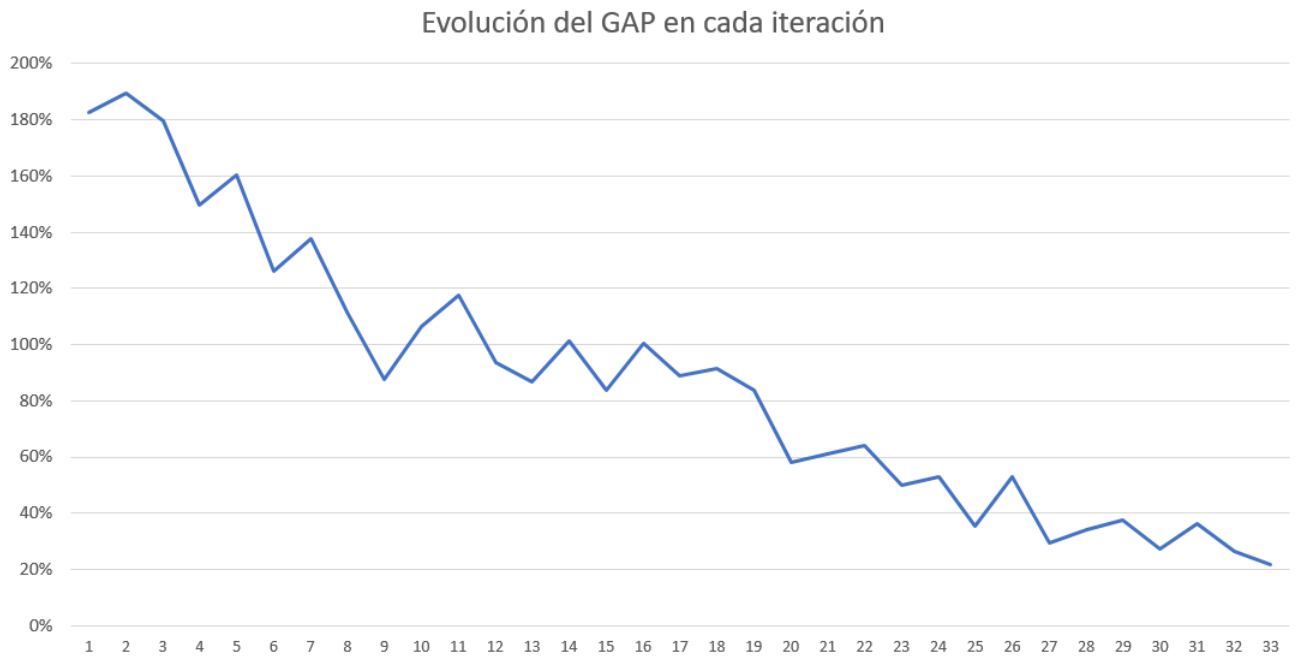


Figura 4.9: Evolución del GAP de la solución a través de las 33 iteraciones del *Polishing*.

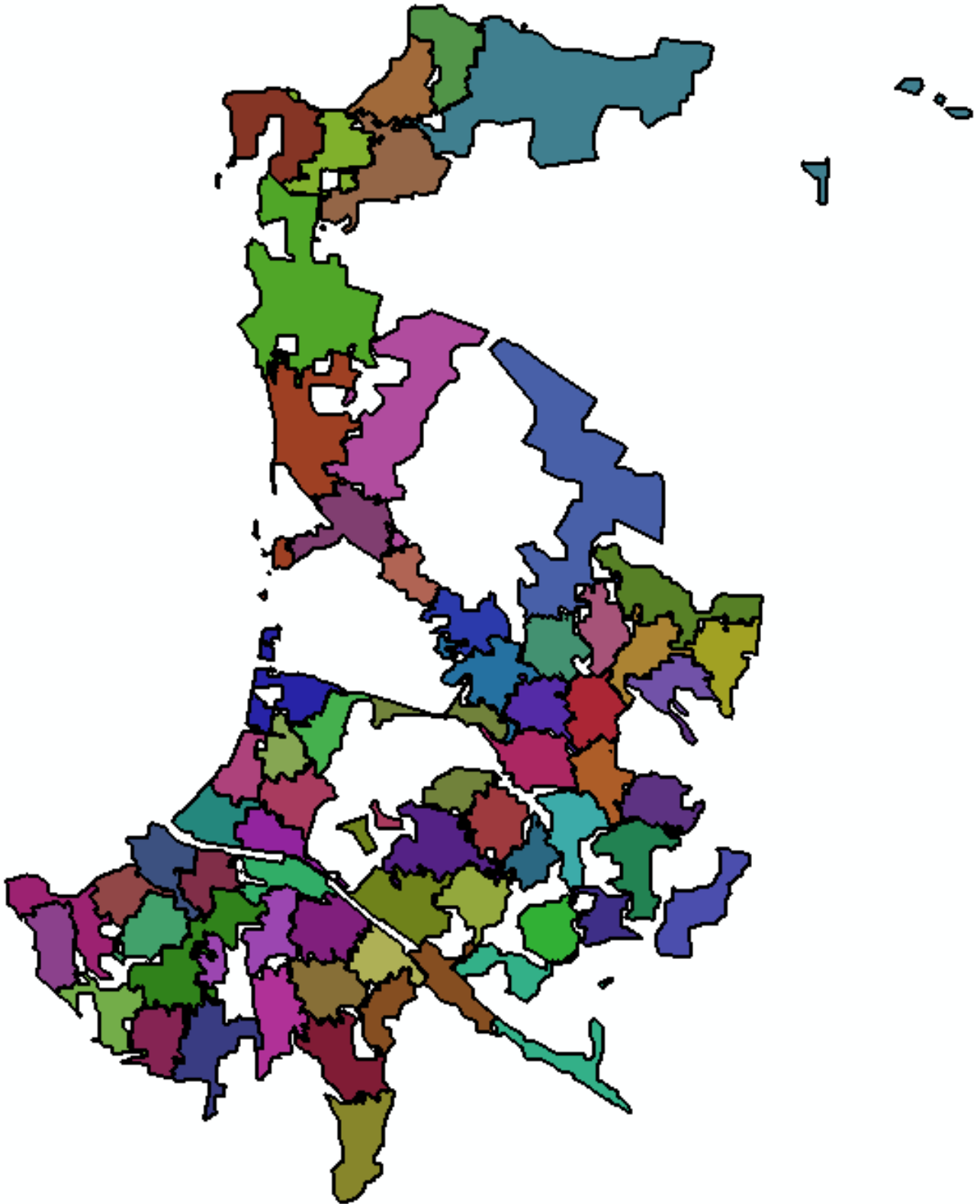


Figura 4.10: Cuarteles generados en la comuna de Viña del Mar y Concón. Se pueden ver cuarteles de área muy grande y otros de área pequeña, esto ocurre debido al equilibrio de cargas de trabajo y no a la cantidad de CPs de cada cuartel, razones por la que K-Means no funciona, ya que no se busca minimizar la distancia entre cada CP al centroide de su cuartel.

4.2. K-Means

Como mecanismo de comparación, se procedió a evaluar la heurística generada en esta tesis, comparándola con el mecanismo de clusterización más utilizado, K-Means 2.4.1.

Este método de particionamiento busca aglomerar los códigos postales en $K = 63$ cuarteles, de tal manera que se minimice la distancia de los códigos postales al centro de cada cuartel. La medida de distancia utilizada, fue la distancia euclidiana.

Adicionalmente, se le incorporó como restricción del problema, que todos los códigos postales de una misma manzana, deben pertenecer al mismo cuartel. Para este caso se utilizó las mismas semillas con las cuales se comenzó la iteración de generación de cuarteles. Finalmente se ejecutó el método de clusterización mencionado, y luego de 170 iteraciones, los centroides ya no se desplazaron, finalizando la ejecución.

Se puede ver que los cuarteles generados (Figura D.1) con este algoritmo, son zonas de reparto más compactas y convexas que los de la heurística presentada en este informe, sin embargo, la carga de los cuarteles, obtenidas utilizando la misma carga de los códigos postales utilizada anteriormente, tiene mayor varianza (Figura 4.11), alcanzando un gap entre el cuartel de mayor carga con el de menor carga, igual a 80 %.

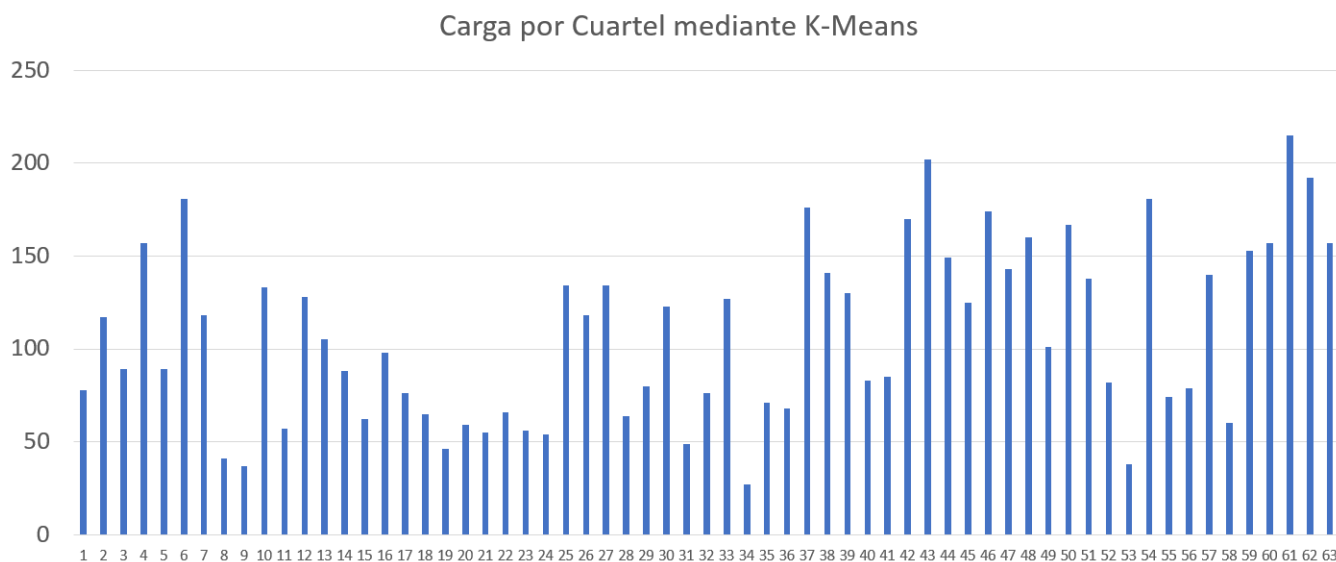


Figura 4.11: Distribución de carga (eje Y) de los cuarteles generados (eje X) en la comuna de Viña del Mar y Concón, mediante el método de particionamiento K-Means. Se obtuvo un GAP del 88,3 %, un promedio de 107,9 y una desviación estándar igual a 47,5

Si bien la forma geográfica de los cuarteles generados, son muy deseables por la empresa, la distribución de carga generada provoca diversas complejidades operacionales, por ejemplo, jornadas laborales variadas. A la vez, se puede ver que los cuarteles no son conexos, lo que se debe a que este mecanismo no utiliza el grafo y las vecindades para generar los clusters.

Luego se considera el desempeño de este método muy inferior para el problema abordado en esta tesis, por lo que se descarta la utilización de este método de particionamiento.

4.3. Carga según Número de Puertas

En esta sección se modificó la carga de los códigos postales, para verificar el funcionamiento de la heurística frente a un escenario diferente.

Se definió la carga de cada código postal como el número de puertas que éste tiene, así se busca generar cuarteles con igual número de puertas entre sí. Sin embargo, se estableció que si el número de puertas es mayor a 70, se fijó en 70, ya que de acuerdo a los estudios de tiempo realizados por la empresa Correos de Chile, donde se midió el tiempo que un cartero demoraba en cada código postal, no es significativa la diferencia cuando el número de puertas es más que 70.

Esta modificación se realizó pues es muy significativa la diferencia entre un código postal y otro, es decir, si un código postal tiene 300 puertas cambia radicalmente la carga total del cuartel al que pertenezca, generando cuarteles muy desbalanceados, sin embargo, en la realidad (según las mediciones de tiempo) el número de puertas deja de incrementar el tiempo de distribución (por ende la carga) pasadas las 70 puertas. Luego las cargas se desbalanceaban sin un mayor efecto real en los tiempos de distribución.

En la Figura 4.12 se puede ver la distribución de carga de los cuarteles actuales, en donde la diferencia entre el cuartel de mayor carga con el de menor carga, es mayor al 90 %.

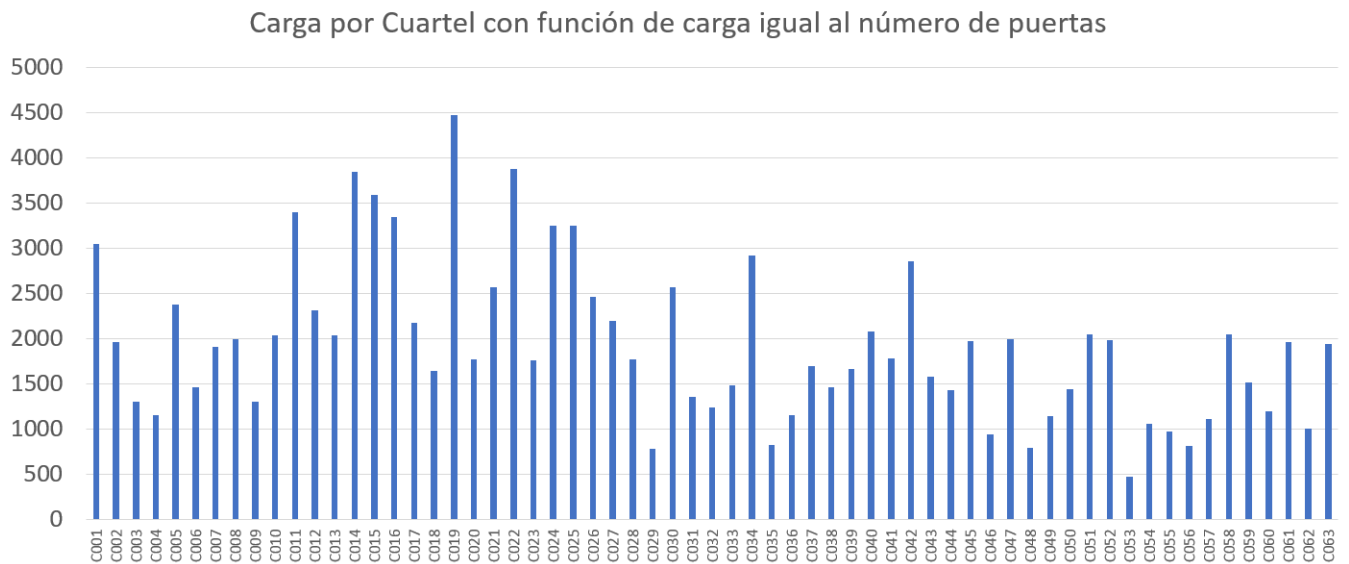


Figura 4.12: Distribución de carga de los cuarteles actuales con la carga de cada código postal igual al número de puertas que contiene. Eje Y corresponde a la cantidad de puertas que contiene cada cuartel del eje X. GAP igual a 206,6 %, promedio de 1.928 y desviación de estándar 862,8.

Para la generación de las zonas de reparto, se utilizó una cota de carga equivalente a:

$$cotacarga = \left\lfloor \frac{\sum_{CP} carga_{CP}}{63} * 1,005 \right\rfloor = 1938$$

Con esto en mente, se procedió a la generación de los cuarteles equilibrados, con el número de puertas como la carga del código postal. Al igual que en la instancia anterior, las semillas utilizadas fueron las mismas que para el caso original. Se corrió el algoritmo con 100 iteraciones, logrando un gap entre el cuartel de mayor carga con el de menor carga, igual al 30 %. (Figura 4.13).

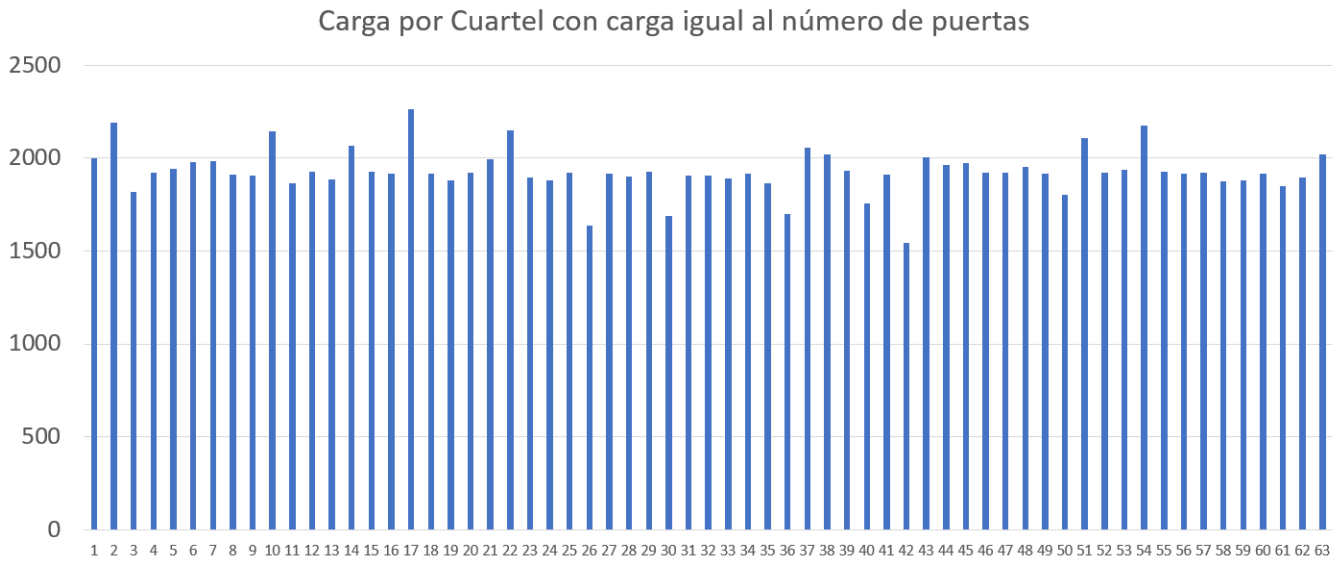


Figura 4.13: Distribución de carga de los cuarteles generados con la heurística y carga de cada código postal igual al número de puertas que contiene. GAP de 30 %, promedio de 1.928 puertas y desviación estándar de 119.

Se puede ver de ambos gráficos, que la carga se distribuyó entre los cuarteles, sin embargo, no se llegó al gap del 25 % en las iteraciones establecidas. Si bien, se podría haber realizado más de 100 iteraciones, se probó con 300 y el gap no mejoró. Esto ocurre principalmente porque la función de carga es discreta y con diferencias entre códigos postales muy grandes, por lo que al traspasar un CP de un cuartel a otro, la carga total de ambos cuarteles, cambia de manera importante. Por ejemplo, si traspaso un código postal con carga igual a 70, de un cuartel a otro, inmediatamente la diferencia de carga entre ambos cambió en 140.

De igual manera, los cuarteles generados, cumplen con las restricciones de las manzanas, y la conectividad dentro del cuartel (Figura E.1).

4.4. Análisis de sensibilidad

En esta sección se realizó un análisis de sensibilidad con respecto a las semillas iniciales de la heurística.

Anteriormente se utilizó semillas arbitrariamente seleccionadas, de tal manera que quedarán distribuidas a lo largo de toda la comuna, e intentando que sean equidistantes. En esta ocasión se utilizó dos alternativas de semillas, donde ambas fueron seleccionadas aleatoriamente en el plano.

En estas instancias la heurística tardó 13,4 y 14,2 minutos en llegar a los GAP de 23 % y 22 % respectivamente (Figura 4.14).

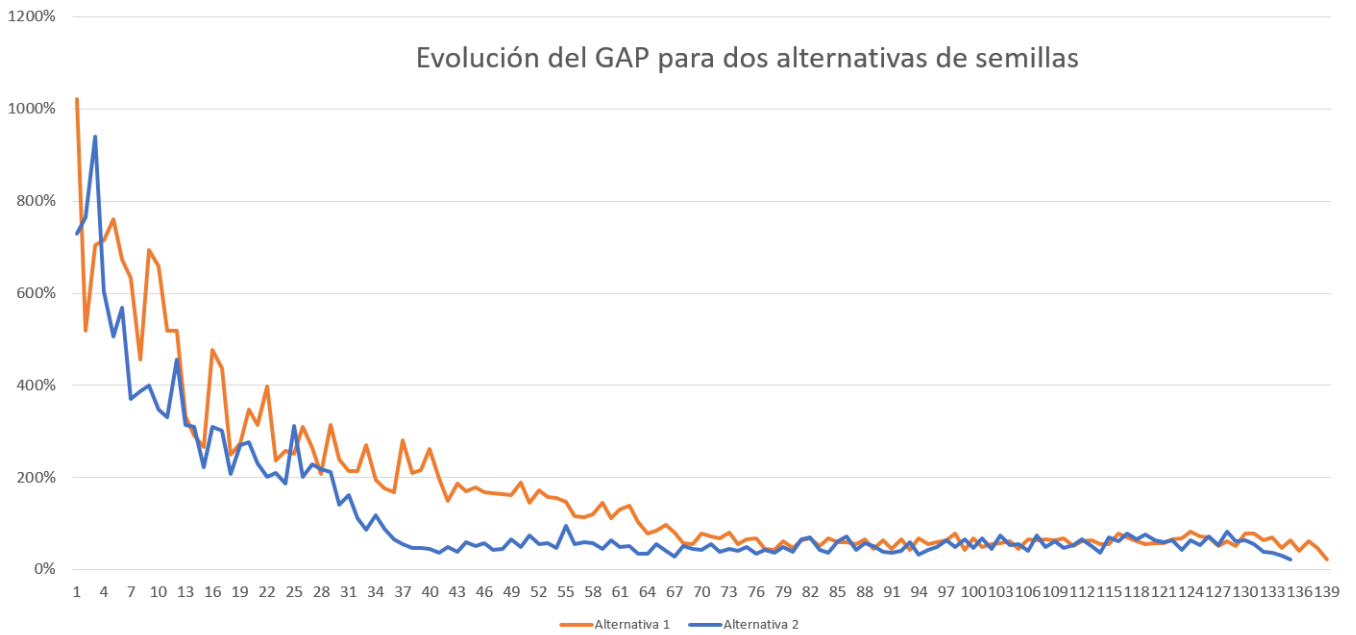


Figura 4.14: Evolución del GAP de dos instancias con semillas aleatorias, a lo largo de las iteraciones del *Polishing*.

En ambas instancias se alcanzó el GAP establecido y las restricciones establecidas por la empresa. Sin embargo, en ambas instancias la heurística demoró más que en el caso anterior, iterando más de 130 veces en el *Polishing*. Esto se debe a que en el primer paso de la heurística, donde se asignan los códigos postales a un cuartel, los cuarteles quedan muy desbalanceados, alcanzando un GAP hasta 10 veces más grande que antes. En las figuras 4.15 y 4.16 se puede ver las semillas iniciales y los cuarteles obtenidos en cada instancia.

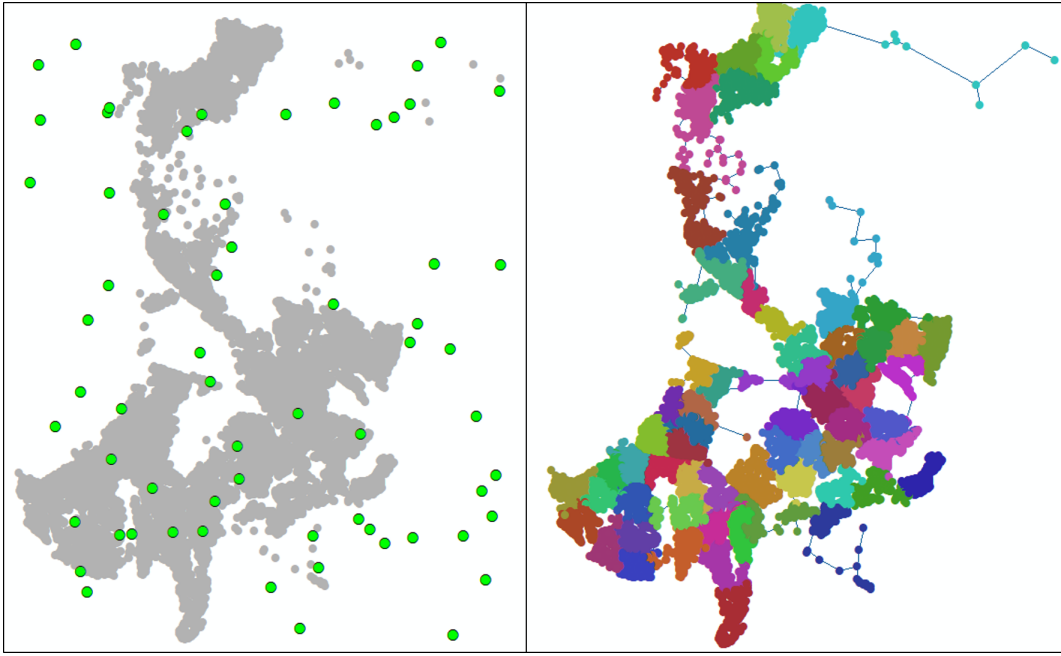


Figura 4.15: Distribución de las semillas aleatorias y los cuarteles obtenidos a partir de ellas. GAP del 23 % obtenido en 13,4 minutos.

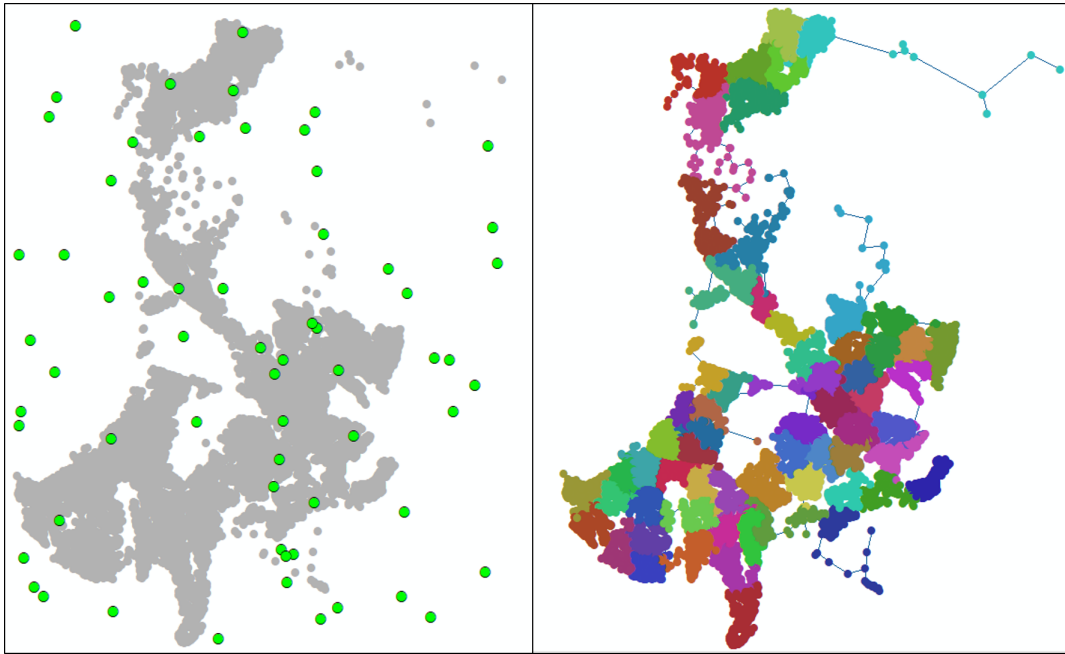


Figura 4.16: Distribución de las semillas aleatorias y los cuarteles obtenidos a partir de ellas. GAP del 22 % obtenido en 14,2 minutos.

En la imagen que se encuentra en la Figura 4.17 se puede ver cómo distribuye la carga de trabajo en los 63 cuarteles de Viña del Mar y Concón, para ambas instancias de semillas aleatorias.

Distribución de la carga de trabajo por cuartel con dos instancias de semillas aleatorias

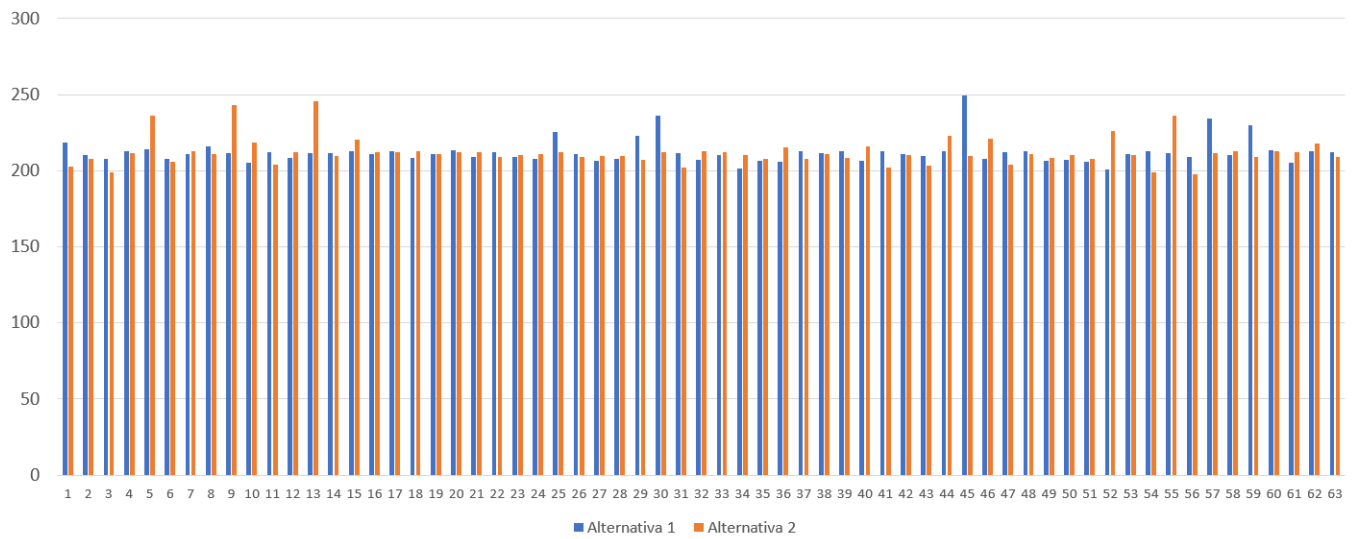


Figura 4.17: Distribución de carga de trabajo de los 63 cuarteles para ambas instancias de semillas aleatorias.

Capítulo 5

Conclusiones

Es claro que la empresa Correos de Chile, tiene un problema con el diseño de las zonas de reparto de sus casi 2300 carteros, por lo que se hizo indispensable mejorar los cuarteles para que se equilibren las cargas de trabajo, entendidas como la jornada laboral del cartero, y así mejorar la calidad de servicio de la compañía.

En este informe se presentó una herramienta computacional que permite la generación de zonas de reparto equilibradas para las empresas del mercado postal, a través de la minimización de la diferencia de carga entre el cuartel de mayor carga, con el de menor carga.

La heurística desarrollada fue basada en métodos de particionamiento jerárquicos y aglomerativos enfocados en grafos, donde los vértices corresponden a los códigos postales, y las aristas corresponden a las calles de la comuna. El principal aporte de esta heurística aglomerativa, corresponde a la incorporación de los vértices con carga.

La complejidad de la generación del grafo fue determinada por la cantidad de códigos postales. Para instancias con más códigos postales, se recomienda generar una grilla con más divisiones para disminuir la complejidad de la generación del árbol en cada cuadrante. Por su parte, el parámetro α definido para la cota de distancia, modifica las vecindades del grafo. Para comunas con códigos postales más lejanos, se hace necesario agrandar el parámetro para poder representar las calles de la comuna, y para garantizar la conectividad de todo el grafo.

Si bien, en esta tesis se entregó un algoritmo para la generación del grafo, este proceso no es necesario si ya se cuenta con un grafo que conecta todos los códigos postales, ya que la heurística creada, genera las zonas de reparto sin importar el grafo sobre el cual se trabaje.

Los cuarteles obtenidos por el modelo lograron una mayor equidad que los cuarteles actuales generados a través de una evolución histórica, pasando de un gap del 80 % a uno del 25 % en 33 iteraciones, y se pudo ver que el modelo satisface los requerimientos operacionales de la empresa.

Se realizó una comparación con el método de particionamiento más utilizado, K-Means, el cual tuvo un tiempo de procesamiento menor, y generó cuarteles convexos y más compactos. Sin embargo la desviación de las cargas entre cuarteles es mayor al 80 %, y no se respetó

la restricción de conectividad del cuartel, principalmente, porque el método K-Means no se trabajó sobre el grafo generado en este informe.

Adicionalmente se probó la heurística con una función de carga distinta a la utilizada anteriormente, para ver la robustez del algoritmo. En esta instancia se obtuvo una mejor equidad de la carga de los cuarteles en comparación con los cuarteles actuales, sin embargo, debido a la gran varianza de la carga entre los códigos postales, el gap obtenido entre el cuartel de mayor carga con el de menor carga, fue cercano al 30 % (Figura 5.1). El gap de carga no mejoró al incrementar el número de iteraciones de 100 a 300, concluyendo que la función de carga determinada, es de vital importancia para generar cuarteles equilibrados.

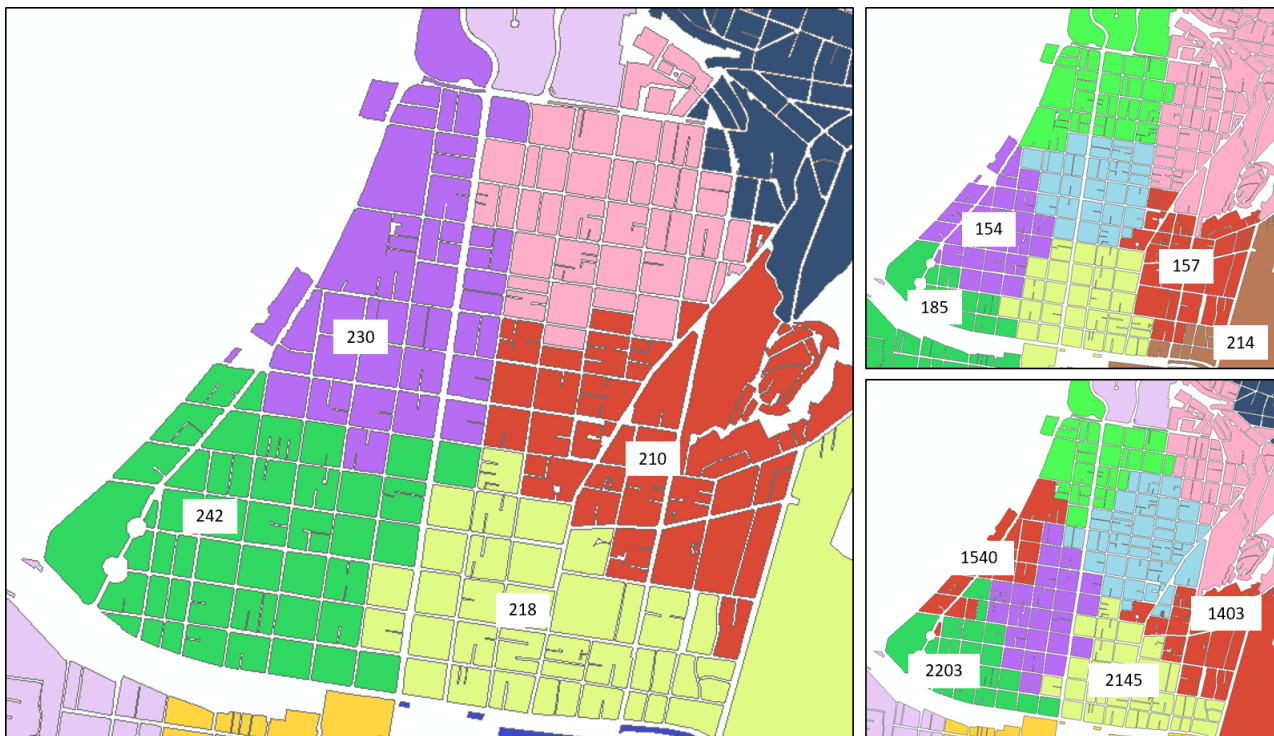


Figura 5.1: Cuarteles del centro de Viña generados en las 3 instancias descrita, y con sus cargas respectivas. El de la izquierda corresponde al generado con la heurística y con función de carga continua entre 1 y 10. El de arriba a la derecha corresponde a los cuarteles generados con K-Means e igual carga que el anterior. El de abajo a la derecha corresponde a los cuarteles generados con la heurística y carga igual al número de puertas del CP.

Limitaciones y trabajos propuestos

Si bien la heurística descrita a lo largo de este trabajo, cumplió con los requerimientos y objetivos planteados, depende de factores que pueden generar resultados distintos, entre ellos se encuentran:

- Función de carga: la carga que tiene cada código postal genera cuarteles diferentes dependiendo la varianza que ellos tengan, tal como se vio en la sección anterior.
- Semillas: las semillas iniciales afectan el diseño de los cuarteles. Se generó los cuarteles

con semillas concentradas en el centro, y las zonas de reparto generadas son menos compactas que cuando se generaron con semillas aleatorizadas.

- Cotas: las cotas de carga y distancia modifican el tamaño del cuartel, y las vecindades del grafo respectivamente.
- Entre otros.

A partir de lo anterior surge la necesidad de incorporar dentro de la función de carga, una variable que represente el tiempo de desplazamiento entre códigos postales. Esto es necesario para incluir de forma dinámica el tiempo adicional que genera en el ruteo del cartero, la incorporación de un determinado código postal al cuartel. Para esto puede ser necesario la utilización de un mecanismo de aproximación del TSP, que entregue una estimación del tiempo que adiciona un CP en la jornada laboral de un cartero si se incluye en un cuartel o en otro.

También se propone la incorporación de un mecanismo de generación de semillas eficientes, en donde se minimice el movimiento de códigos postales entre el proceso de Generación de Cuarteles con el de Polishing, o se minimice la perturbación entre los cuarteles finales con los cuarteles actuales de la comuna.

Para concluir, la heurística de generación de cuarteles presentada en este informe, permite a la empresa Correos de Chile, a través de una herramienta computacional, generar zonas de reparto más eficientes, de tal manera que se equilibre la jornada laboral de los carteros, y por lo tanto, la calidad de servicio entregada a sus clientes.

Bibliografía

- [1] Mark S Daskin. What you should know about location modeling. *Naval Research Logistics (NRL)*, 55(4):283–294, 2008.
- [2] Correos de Chile. *Memoria Anual*. 2015.
- [3] Jack Edmonds and Ellis L Johnson. Matching, euler tours and the chinese postman. *Mathematical programming*, 5(1):88–124, 1973.
- [4] Ahmad I Jarrah and Jonathan F Bard. Large-scale pickup and delivery work area design. *Computers & Operations Research*, 39(12):3102–3118, 2012.
- [5] David S Johnson, Maria Minkoff, and Steven Phillips. The prize collecting steiner tree problem: theory and practice. In *SODA*, volume 1, page 4, 2000.
- [6] Gilbert Laporte, Suzanne Chapleau, Philippe-Eric Landry, and H el ene Mercure. An algorithm for the design of mailbox collection routes in urban areas. *Transportation Research Part B: Methodological*, 23(4):271–280, 1989.
- [7] Oded Maimon and Lior Rokach. Data mining and knowledge discovery handbook. secaucus, 2005.
- [8] Keijo Ruohonen. *Graph Theory*. Tampere University of Technology, 2008.

Apéndice A

Set Covering Model

- d_{ij} : distancia entre el nodo i y el nodo j
- a_{ij} : 1 si $d_{ij} \leq D$, donde D corresponde a la máxima distancia entre un centro y un nodo de demanda que puede ser servido.
- x_j : 1 si un centro es ubicado en el nodo $j \in \mathcal{J}$, 0 si no

$$\begin{aligned} & \text{minimize} && \sum_{j \in \mathcal{J}} x_j \\ & \text{S.T.} && \sum_{j \in \mathcal{J}} a_{ij} x_j \geq 1, \forall i \in \mathcal{I} \\ & && x_j \in \{0, 1\}, \forall j \in \mathcal{J} \end{aligned}$$

Apéndice B

P-Median

- d_j : demanda del nodo j
- c_{ij} : costo unitario de satisfacer el nodo j desde el centro i
- p : número de centros a localizar
- y_i : 1 si el nodo i es centro, 0 si no
- x_{ij} : fracción de demanda del nodo j que es cubierta por el centro i

$$\begin{aligned} & \text{minimize} && \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} d_j c_{ij} x_{ij} \\ & \text{S.T.} && \sum_{i \in \mathcal{I}} x_{ij} = 1, \forall j \in \mathcal{J} \\ & && \sum_{j \in \mathcal{J}} y_i = p \\ & && x_{ij} - y_i \leq 0, \forall i \in \mathcal{I}; j \in \mathcal{J} \\ & && y_i \in \{0, 1\}, \forall i \in \mathcal{I} \\ & && x_{ij} \geq 0, \forall i \in \mathcal{I}; j \in \mathcal{J} \end{aligned}$$

Apéndice C

Max Covering Model

- d_{ij} : distancia entre el nodo i y el nodo j
- a_i : demanda del nodo i
- $N_i = \{j \in \mathcal{J} \mid d_{ij} \leq S\}$ con S la máxima distancia entre un centro y un nodo de demanda que puede ser servido
- y_i : 1 si algún centro puede abastecer al nodo i
- x_j : 1 si un centro es ubicado en el nodo j , 0 si no

$$\begin{aligned} & \text{maximize} && \sum_{i \in \mathcal{I}} a_i y_i \\ & \text{S.T.} && \sum_{j \in N_i} x_j \geq y_i, \quad \forall i \in \mathcal{I} \\ & && \sum_{j \in \mathcal{J}} x_j = p \\ & && y_i \in \{0, 1\}, \quad \forall i \in \mathcal{I} \\ & && x_j \in \{0, 1\}, \quad \forall j \in \mathcal{J} \end{aligned}$$

Apéndice D

Cuarteles con K-Means

Cuarteles generados mediante el método K-Means. Se muestra cada código postal con un color que representa el cuartel al que pertenece.

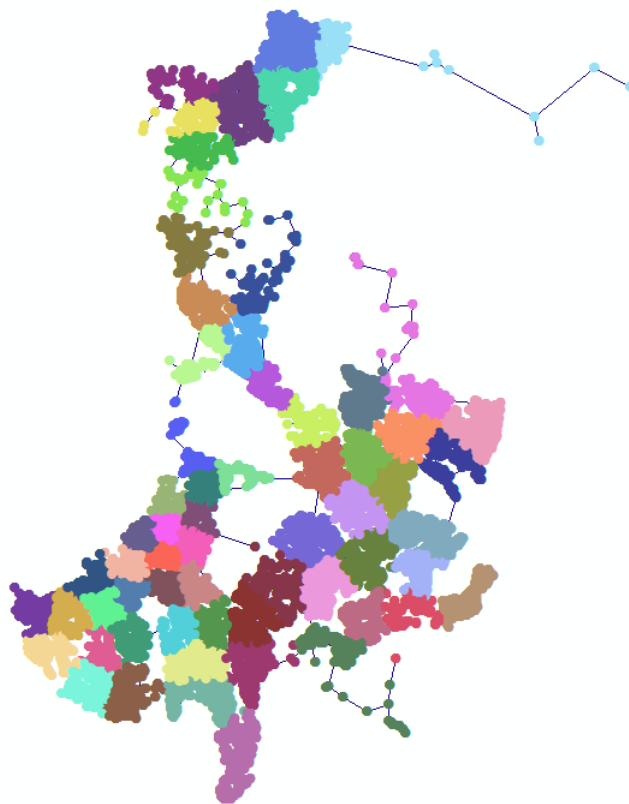


Figura D.1: Cuarteles generados en la comuna de Viña del Mar y Concón mediante el método de particionamiento K-Means.

Apéndice E

Cuarteles con carga igual al número de puertas

Cuarteles generados mediante la heurística descrita en este informe, y con función de carga de trabajo del código postal, equivalente al número de puertas que contiene. Se muestra cada código postal con un color que representa el cuartel al que pertenece.

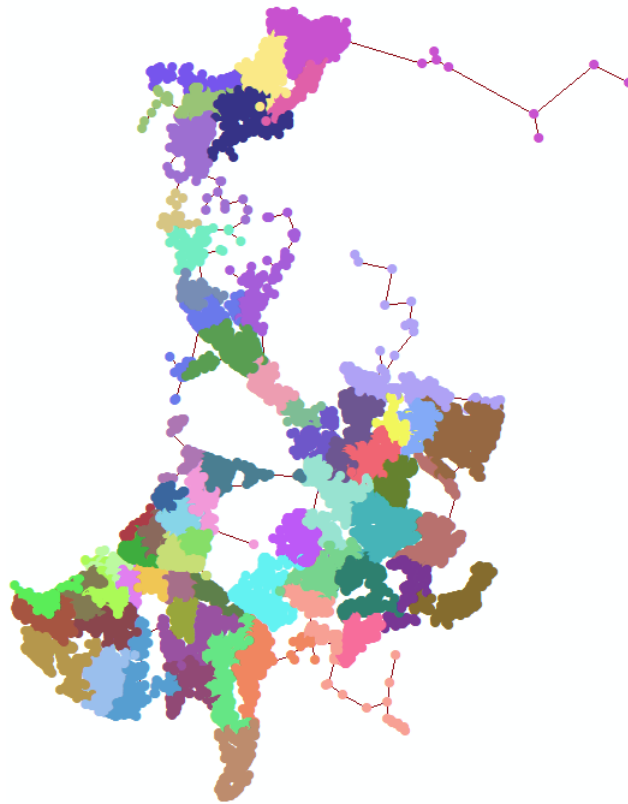


Figura E.1: Cuarteles generados en la comuna de Viña del Mar y Concón mediante con la carga de los código postales igual al número de puertas que contienen.