



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

BELLEZA Y CALIDAD PERCIBIDA EN RUTEO DE VEHÍCULOS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE  
OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL  
MATEMÁTICO

LUCAS JOAQUÍN URIBE PESQUEIRA

PROFESOR GUÍA:  
FERNANDO ORDOÑEZ PIZARRO

MIEMBROS DE LA COMISIÓN:  
JORGE AMAYA ARRIAGADA  
ANDRÉS MUSALEM SAID  
ÁLVARO ECHEVERRÍA SOLÍS

SANTIAGO DE CHILE  
2020

**RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE:** Ingeniero Civil Matemático y grado de  
Magíster en Gestión de Operaciones  
**POR:** Lucas Joaquín Uribe Pesqueira  
**FECHA:** 27/10/2020  
**PROFESOR GUÍA:** Fernando Ordóñez Pizarro

## **BELLEZA Y CALIDAD PERCIBIDA EN RUTEO DE VEHÍCULOS**

Hay muchas razones por las cuales la aplicación de modelos y técnicas logísticas fallan al implementarlas en la vida real. Una de éstas es imponerlas como una solución final e inmodificable, sin considerar el feedback de los usuarios. Esto es particularmente relevante para SimpliRoute, empresa que provee un servicio de ruteo de vehículos, porque el atractivo de su producto, y muchas veces la decisión de compra por parte de sus clientes, dependen directamente de la percepción que tienen estos sobre la utilidad que entrega la plataforma.

En este contexto de integración de feedback, SimpliRoute detectó que muchos de sus clientes no estaban contentos con los resultados de su algoritmo, por razones del tipo: “las rutas planificadas son muy feas”. Por esto, se decidió investigar el fenómeno de manera formal, para entender por qué no les hacen sentido las soluciones entregadas, cómo son las que esperan, de qué formas se podría lograr entregar tales soluciones, y qué se perdería por hacerlo.

Se realizaron entrevistas con usuarios de la plataforma, en las que se profundizó sobre los criterios usados al editar las rutas entregadas por el software, lo que permitió definir variables relevantes a estudiar. Además, se revisó la literatura sobre el *atractivo visual* en el Problema de Ruteo de Vehículos (*VRP*), y se identificaron las distintas medidas utilizadas. Con la información recopilada se realizó una clasificación de las distintas formas de evaluar *atractivo visual*, y un estudio de su relación con las medidas tradicionales de optimalidad del *VRP*.

Por otro lado, para testear las hipótesis generadas a partir de la información recolectada hasta ese momento, y si las variables identificadas previamente tenían relevancia real en las decisiones de los usuarios, se realizaron dos estudios. En el primero se analizó la información almacenada en la base de datos de SimpliRoute, que guarda, para cada optimización que realiza algún cliente, el resultado que entrega el algoritmo y la versión final que decide utilizar luego de editar manualmente tal solución. El segundo consistió en una encuesta que fue respondida por varios clientes, en la que cada pregunta mostraba distintas soluciones a un *VRP*, y se les pidió que eligiesen la que consideraran mejor.

Los resultados indican que las medidas de belleza más importantes para los usuarios son la distancia de los puntos de cada ruta hacia su centro, y la cantidad de puntos dentro de las envolturas convexas de otras rutas; ambas medidas en conjunto son casi tan relevantes como la distancia total recorrida. Además, se identificó que éstas están altamente correlacionadas con el tiempo total de manejo en calle (0.91 y 0.55 respectivamente), lo cual indica que las soluciones que son percibidas como *visualmente atractivas* tienden a ser mejores soluciones que las que son calificadas como “más feas”. Por último, se observó que cuando se consideran solo las soluciones más cercanas al óptimo en la medida estándar, existe un pequeño *trade-off* (dependiente de la topología de la instancia) que no permite que mejoren al mismo tiempo las medidas de *atractivo visual* y las medidas de optimalidad estándar.

# Agradecimientos

A pesar de que nada de lo que escriba aquí puede llegar a reflejar mis reales agradecimientos a todas las personas que me acompañaron y ayudaron durante este camino, quiero intentarlo.

Primero, a Álvaro por confiar en mí para este proyecto y darme la oportunidad de ser parte de SimpliRoute. A Fernando por su disposición, guía, motivación, crítica y buena onda; 5 elementos que fueron clave para el desarrollo de este trabajo. A Andrés por su paciencia, su invaluable apoyo técnico y sus ganas de colaborar desde el momento en que le contamos sobre el proyecto. También a Linda y Fernanda por su disponibilidad absoluta y por no soltarme en todos estos años.

Al equipo completo de SimpliRoute, y en especial a mis compañeros y compañeras de Data Science, por hacer que mi día a día sea desafiante, entretenido y acogedor. A Víctor, por ser mi partner y complemento desde el principio de esta aventura.

A mis amigos y amigas, por todo, cualquier otra forma de decirlo no sería suficiente. A todas las personas con las que compartí en el grupo Scout Meli Witran Mapu, y en particular a las que trabajaron conmigo en el Staff; no imagino mi vida sin ustedes.

A Fernanda, por quererme y entenderme como nadie más, y por soportar mis largos días y noches de estudio.

A mi mamá, papá, hermanos y hermana. Por estar ahí siempre que lo necesité, y por motivarme y desafiarme con cariño, por ser un ejemplo a seguir para mí, por darme mi espacio sin dejar que me pierda. Muy especialmente, por su apoyo en el momento en que más dudé de mi carrera y de mí mismo.

Lucas Uribe Pesqueira.

# Tabla de contenido

|  |           |
|--|-----------|
| <b>Resumen</b>   | <b>I</b>  |
| <b>Agradecimientos</b>   | <b>II</b> |
| <b>1. Introducción</b>   | <b>1</b>  |
| 1.1. Motivación . . . . .  | 1         |
| 1.2. Definición del Problema . . . . .                                 | 5         |
| 1.3. Objetivos . . . . .   | 6         |
| 1.3.1. Objetivo General . . . . .                                      | 6         |
| 1.3.2. Objetivos Específicos . . . . .                                 | 6         |
| 1.4. Hipótesis . . . . .   | 6         |
| 1.5. Alcances . . . . .  | 7         |
| 1.6. Estructura de la tesis . . . . .                                  | 7         |
| <b>2. Marco Conceptual</b>   | <b>9</b>  |
| 2.1. Formulación genérica del problema de ruteo de vehículos . . . . . | 9         |
| 2.1.1. Métodos de resolución . . . . .                                 | 12        |
| 2.2. Ruteo de Vehículos Comercial . . . . .                            | 17        |

|   |           |
|---|-----------|
| 2.2.1. SimpliRoute . . . . .  | 19        |
| 2.3. Medidas de evaluación de Atractivo Visual . . . . .                        | 20        |
| 2.4. Trabajos Relacionados . . . . .  | 22        |
| <b>3. Entrevistas a Clientes</b>  | <b>31</b> |
| 3.1. Contexto y metodología . . . . .   | 31        |
| 3.2. Resumen de resultados . . . . .  | 33        |
| <b>4. Medidas de evaluación de Atractivo Visual</b>                             | <b>41</b> |
| 4.1. Clasificación de medidas usadas en la literatura . . . . .                 | 41        |
| 4.2. Nuevas medidas propuestas . . . . .  | 45        |
| 4.3. Relación entre las medidas en consideración . . . . .                      | 47        |
| 4.3.1. Metodología . . . . .  | 47        |
| 4.3.2. Resultados . . . . .   | 48        |
| <b>5. Confirmación de Relevancia</b>  | <b>60</b> |
| 5.1. Evaluación de cambios hechos por clientes en su operación diaria . . . . . | 60        |
| 5.2. Encuestas a Clientes . . . . .   | 71        |
| 5.2.1. Metodología . . . . .  | 71        |
| 5.2.2. Resultados . . . . .   | 77        |
| <b>6. Conclusiones y trabajos futuros</b>                                       | <b>83</b> |
| <b>Bibliografía</b>   | <b>86</b> |

# Índice de Tablas

|  |    |
|--|----|
| 3.1. Características de las empresas entrevistadas . . . . .   | 32 |
| 4.1. Resumen de la cantidad de instancias y soluciones generadas, antes y después de eliminar las soluciones que no alcanzaron el mínimo número de vehículos conocido . . . . .  | 49 |
| 4.2. Resumen descriptivo del <b>PercentageDT</b> en las soluciones generadas, para cada grupo de instancias . . . . .  | 49 |
| 4.3. Correlaciones entre las variables de <i>VA</i> y las tradicionales. . . . .   | 50 |
| 4.4. Promedios por instancia de correlaciones entre variables de <i>VA</i> y tradicionales. . . . .  | 52 |
| 5.1. Número de instancias incluidas en el análisis por cantidad de vehículos y puntos, y porcentaje de soluciones editadas en cada categoría. . . . .  | 62 |
| 5.2. Número de instancias incluidas en el análisis de cada grupo, por cantidad de vehículos y de puntos . . . . .  | 63 |
| 5.3. Promedio de valores alcanzados por las distintas variables consideradas, en las soluciones de las instancias que no fueron editadas ('Sin Cambio'), y en las que sí cambiaron, para los casos antes de editar ('Original'), después de editar ('Editada'), en cada grupo. . . . . | 64 |
| 5.4. Porcentaje de instancias que mejora, se mantiene igual, o empeora con la edición. . . . .   | 65 |

|  |    |
|--|----|
| 5.5. Número de soluciones que empeoran, se mantienen igual, o mejoran en <b>Avg-<br/>ForeignPointsHull</b> , según si su valor original era mínimo o no, según si la edición mejoró el <b>DT</b> o no, y según si la edición mantuvo la factibilidad o no. . . . .   | 66 |
| 5.6. Número de soluciones que empeoran, se mantienen igual, o mejoran en <b>Avg-<br/>Cross</b> , según si su valor original era mínimo o no, según si la edición mejoró el <b>DT</b> o no, y según si la edición mantuvo la factibilidad o no. . . . .   | 67 |
| 5.7. Número de soluciones que empeoran, se mantienen igual, o mejoran en <b>Avg-<br/>DistCenter</b> , según si la edición mejoró el <b>DT</b> o no, y según si la edición mantuvo la factibilidad o no. . . . .  | 68 |
| 5.8. Número de soluciones que empeoran, se mantienen igual, o mejoran en <b>Avg-<br/>DistBetween</b> , según si la edición mejoró el <b>DT</b> o no, y según si la edición mantuvo la factibilidad o no. . . . .   | 69 |
| 5.9. Número de soluciones que empeoran, se mantienen igual, o mejoran en <b>AvgRa-<br/>tioCenterV1</b> , según si su valor original era mínimo o no, según si la edición mejoró el <b>DT</b> o no, y según si la edición mantuvo la factibilidad o no. . . . .   | 69 |
| 5.10. Número de soluciones que empeoran, se mantienen igual, o mejoran en <b>Av-<br/>gRatioCenterV2</b> , según si la edición mejoró el <b>DT</b> o no, y según si la edición mantuvo la factibilidad o no. . . . .  | 70 |
| 5.11. Valores pertenecientes a cada nivel en los distintos atributos después de la discretización. Los porcentajes asociados a <b>PercentageDT</b> y <b>Percentage-<br/>Center</b> corresponden al desvío porcentual de <b>DT</b> y <b>AvgDistCenter</b> con respecto a la mejor solución encontrada para cada una. Los valores asociados a <b>RelativePointsHull</b> corresponden a <b>SumForeignPointsHull / (número de puntos insertados)</b> . . . . . | 73 |
| 5.12. Arreglo ortogonal utilizado como base para escoger las características de las soluciones a incluir. La última columna indica si la combinación de niveles fue posible de generar o no. . . . .   | 74 |

|   |    |
|---|----|
| 5.13. Promedio y percentiles 2,5% y 97,5% de las distribuciones a posteriori de la Media y la Desviación Estándar de cada parámetro del modelo. . . . . | 81 |
|---|----|



# Índice de Ilustraciones

|   |    |
|---|----|
| 1.1. Izquierda: Solución original entregada por el algoritmo. Derecha: Solución después de la edición manual de un cliente. . . . .   | 2  |
| 4.1. <b>SumDistCenter</b> vs <b>MaxPointsHull</b> por instancia . . . . .   | 51 |
| 4.2. <b>SumDistCenter</b> vs <b>PercentageDT</b> por instancia y grupo . . . . .  | 54 |
| 4.3. <b>SumForeignPointsHull</b> vs <b>PercentageDT</b> por instancia y grupo . . . . .   | 55 |
| 4.4. <b>SumForeignPointsHull</b> vs <b>PercentageDT</b> por instancia para grupo C1. Los puntos rojos muestran la frontera de Pareto y el punto verde corresponde a la MSC. . . . .   | 57 |
| 4.5. <b>SumForeignPointsHull</b> vs <b>PercentageDT</b> por instancia para grupo R1. Los puntos rojos muestran la frontera de Pareto y el punto verde corresponde a la MSC. . . . .   | 58 |
| 5.1. Izquierda: Solución original. Derecha: Solución después de ser editada por el planificador. . . . .  | 61 |
| 5.2. Puntos y envolturas convexas de cada ruta, en una de las soluciones escogidas inicialmente para aparecer en la encuesta. Si bien esta solución tiene nivel 1 para <b>RPH</b> , a pesar de tener pocos puntos dentro de las envolturas convexas de rutas que no son la suya propia, las superficies de éstas sí se intersectan considerablemente. . . . . | 75 |

|   |    |
|---|----|
| 5.3. <b>RPH vs PC, RPH vs PDT y PC vs PDT</b> para instancia elegida. Los puntos naranjos corresponden a las soluciones incluidas en las preguntas de la encuesta.  | 76 |
| 5.4. Valores de las medias ( $\mu_i$ ) de las variables aleatorias normales asociadas a la utilidad de cada nivel (sólo se muestran 3 niveles por atributo, pues el cuarto está fijado en 0) para los tres atributos en consideración, a través de la cadena.   | 78 |
| 5.5. Utilidades parciales ( <i>part-worths</i> ) de cada nivel para los tres atributos en consideración. . . . .  | 79 |
| 5.6. Distribución <i>a posteriori</i> de las medias ( $\mu_i$ ) para los niveles 1, 2 y 3 de cada atributo (el nivel 4 está fijo en 0 en las 3 variables), indicando con la media <i>a posteriori</i> (en rojo), el intervalo de $\pm 2$ errores estándar numéricos en torno a ésta (en amarillo), y los intervalos de credibilidad bayesiana al 95 % y 99 % (en verde y azul respectivamente). . . . . | 80 |
| 5.7. Importancia relativa de cada atributo. . . . .   | 81 |

# Capítulo 1

## Introducción

### 1.1. Motivación

El problema de ruteo de vehículos [15] (*Vehicle Routing Problem (VRP)*) es un ejemplo clásico de la Investigación de Operaciones (*IO*). Ha sido estudiado extensivamente y ha nutrido al sector productivo con técnicas y modelos que han sido aprovechados ampliamente en la práctica [29].

Una de las formas en las que las organizaciones pueden beneficiarse de los avances de esta área de investigación, sin tener que invertir en desarrollar modelos propios, es a través de empresas de *software* que ofrecen sistemas que les permiten planificar su operación logística de manera eficiente y organizada.

Hace algunos años pocas empresas podían utilizar este tipo de herramientas por el alto costo monetario que tenían. En los resultados de la encuesta del año 2008 sobre proveedores de *software* de ruteo de vehículos (de aquí en adelante *CVRS* por su nombre en inglés (*Commercial Vehicle Routing Software*)) que se publica bianualmente en la revista *OR/MS Today* de *INFORMS* [26] se menciona que es difícil encontrar proveedores que ofrezcan este tipo de productos por menos de 10.000 dólares. Sin embargo, en el último tiempo han surgido *CVRSs* que aprovechándose de los avances que ha tenido la tecnología (como la mayor disponibilidad y accesibilidad de servicios *cloud* y el aumento en la capacidad de cómputo), pueden entregar servicios de igual calidad a valores mucho menores. Desde el año 2010 en adelante, los resultados de esta misma encuesta [27] mencionan la existencia de alternativas

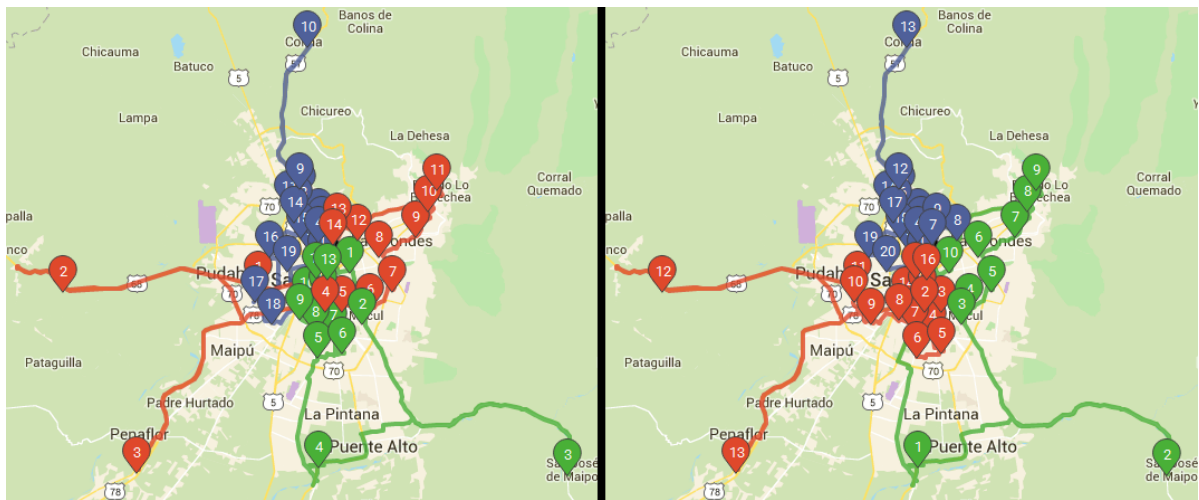


Figura 1.1: Izquierda: Solución original entregada por el algoritmo. Derecha: Solución después de la edición manual de un cliente.

*SaaS* (Software como servicio, por sus iniciales en inglés) con valores partiendo desde los 100 dólares mensuales. Este cambio ha permitido la masificación del uso de estas herramientas, ampliando y diversificando el tipo de usuarios de sus plataformas.

Uno de estos *CVRs* es SimpliRoute, *Startup* tecnológico chileno que desde el año 2015 ofrece un *SaaS* de ruteo de vehículos en la nube. Después de un año de trabajo, el equipo de SimpliRoute comenzó a notar que varios de sus clientes no estaban 100% satisfechos con los resultados que entregaba su plataforma. Algunos tomaban las soluciones sugeridas por el algoritmo y las editaban excesivamente de forma manual, para que éstas se ajustaran de mejor forma a sus necesidades y gustos particulares, y otros se aprovechaban de la libertad que brinda el *software* para definir algunos parámetros y restricciones que influyen en la ejecución del algoritmo, y en vez de utilizarlos con el fin para el que fueron diseñados, los usaban para forzar resultados con las características que ellos buscaban. En la figura 1.1 se muestra un ejemplo de una solución antes y después de que un cliente la edite; cada uno de los tres colores representa la ruta a visitar por cada vehículo, y los números dentro de los puntos muestran el orden en que éstos deben de ser visitados. Cabe mencionar que en este caso, las rutas de la solución editada recorren una mayor cantidad de kilómetros que las de la solución original.

Al mismo tiempo, los encargados de ventas de SimpliRoute pudieron detectar que muchos de sus potenciales clientes decidían no comprar su servicio porque las rutas que entregaba el algoritmo no cumplían con ciertos requisitos que ellos declaraban. Al preguntarles la razón

por la que no les gustaban las soluciones mencionaban cosas como “*aunque la solución que entregan usa menos camiones que los que nosotros utilizamos actualmente, las rutas en el mapa se ven muy desorganizadas, muy feas. No hacen sentido*”.

Este tema comenzó a tomar cada vez más relevancia para SimpliRoute, no sólo porque podían perder algunos posibles clientes, sino que también porque la promesa de ofrecer un servicio de ruteo inteligente que entrega automáticamente los resultados que los clientes esperan, no se cumple si es que sus usuarios deben de recurrir a otros procesos para llegar a esos resultados. Esto motivó la investigación en profundidad del problema de percepción de calidad de las rutas, con el objetivo de entender mejor el fenómeno, para poder entregar soluciones que se ajusten mejor a las necesidades de sus clientes y así brindar un mejor servicio. En este contexto se enmarca el trabajo este proyecto de tesis.

A pesar de que esta problemática surge de una dificultad inherentemente práctica, ha sido investigada parcialmente en el mundo académico. El hecho de que usuarios de diversos sistemas y plataformas prefieran soluciones más *bonitas* es un comportamiento que ha sido documentado en varios contextos distintos [16,40,49,73,82,85]. Aunque la comunidad científica aún no ha llegado a un acuerdo sobre los mecanismos que regulan este proceso, una de las posibles explicaciones, propuesta por Miller [58], dice que la relación entre la estética y la usabilidad de un sistema o producto depende de la Carga Cognitiva (concepto desarrollado por [80] que hace referencia al esfuerzo usado por la Memoria de Trabajo). El mecanismo plantea que la estética implica orden y claridad, y que estos dos conceptos ayudan a disminuir la Carga Cognitiva, lo cual afectaría positivamente la percepción de eficiencia y satisfacción.

En el contexto del *VRP* también se ha reportado la relevancia de considerar estos factores. Poot et al. [64] mencionan que para los planificadores, el *atractivo visual* de una solución (*VA* por sus siglas en inglés) juega un rol importante a la hora de decidir si aceptarla o no, y Sahoo et al. [72] comentan que para la empresa que estudian en su trabajo, el *atractivo visual* y la compacidad de las rutas de las soluciones son características muy importantes. Bräysy & Hasle [7] en el capítulo sobre *CVRSS* del libro [83], plantean que este tipo de requerimientos fue observado en bastantes usuarios, pero notan que la motivación para preferir la *belleza* no es clara. Además, listan la compacidad de las rutas como un tema importante para investigaciones futuras. Más recientemente, Vössing [86] realiza un estudio en el que

investiga la dificultad que existe para aplicar los métodos publicados en el mundo académico, en solucionar problemas de ruteo de vehículos del mundo real. En este trabajo se entrevistaron varias empresas, las que mencionaron que los factores blandos (como la *VA* de las rutas, la carga de trabajo equilibrada entre los empleados, la imparcialidad general, etc.) están siendo cada vez más importantes para éstas.

Por otro lado, en el campo de la Investigación de Operaciones Conductual, que estudia la relación entre los modelos de *IO* y los aspectos conductuales humanos de sus aplicaciones prácticas, se ha puesto relevancia en el hecho de que las intervenciones en organizaciones no logran los resultados esperados si es que los usuarios finales no las aplican y utilizan de la forma correcta. En este contexto, Becker [3] enfatiza la importancia de hacer investigación que lleve a técnicas de *IO* que no sólo se centren en la precisión analítica, sino que además se enfoquen en aspectos como la robustez, la conveniencia de uso, la implementabilidad, la transparencia para el usuario final y la aceptación tecnológica.

Un ejemplo de este enfoque aplicado al *VRP* es el trabajo de Hollis & Green [33], en el que se propone un algoritmo que además de minimizar el número de vehículos a utilizar, intenta entregar soluciones con buenos niveles de *belleza*. Ellos justifican este enfoque diciendo que las soluciones que son más atractivas desde el punto de vista visual son más robustas, y por lo tanto, es más probable que proporcionen los ahorros de costos que deberían derivarse del uso de *CVRSSs*. Esta consideración por la robustez va en la línea de las tendencias relevantes de investigación en ruteo de vehículos planteadas por Coelho et al. [12], pues se enfoca no sólo en el costo de las soluciones generadas, sino también en la calidad del servicio que se deriva de éstas. Bajo la misma idea, pero en el contexto de solucionar una variante del *VRP* llamada Problema de Recolección de Desechos, Parchitelli et al. [62] hacen hincapié en la necesidad de que exista un método que no sólo se enfoque en los objetivos tradicionales del *VRP*, sino que además intente maximizar la aceptación general de las soluciones.

Por último, uno de los pocos trabajos que hace referencia explícita a la relación entre las medidas de calidad tradicional y la *VA* es el de Lu & Dessouky [50]. Ellos comentan que, aunque los resultados de los experimentos realizados por MacGregor & Ormerod en [53] sugieren que el atractivo visual de una ruta única refleja implícitamente la longitud ésta, no hay investigaciones que apoyen que esa relación también existe entre la distribución geográfica

de diferentes rutas y la distancia total de la solución.

## 1.2. Definición del Problema

Para dar una definición precisa del problema a estudiar, es necesario aclarar algunos términos y supuestos que se utilizarán durante el desarrollo de este trabajo. Dado lo expuesto en la sección 1.1, se asumirá que los usuarios de *CVRSSs* evalúan las soluciones entregadas por éstos basándose en una medida de *calidad subjetiva* o *calidad percibida*, y que en base a ésta deciden elegir una solución por sobre otra. Más aun, para estudiar esta *calidad percibida* se mantendrá el foco en la relación que ésta tiene con las medidas tradicionales de calidad, y con las de *VA*.

Dicho esto, el problema a abordar en este trabajo de tesis consiste en responder las siguientes preguntas:

- ¿Es posible explicar con métricas cuantitativas la percepción de calidad que tienen los usuarios sobre las soluciones? ¿Qué variables se asocian de mejor forma a esta percepción?
- ¿Cómo se relaciona la *VA* de las soluciones con la calidad percibida por los usuarios?
- ¿Cuánto y cómo influyen las medidas tradicionales de calidad del VRP en la percepción de calidad de los usuarios?
- ¿Cuánto están dispuestos los usuarios a perder en términos de calidad tradicional, para aumentar los índices de calidad percibida?
- ¿Cuál es el *trade-off* que existe entre calidad tradicional y calidad percibida? ¿De qué depende que éste sea mayor o menor?

## 1.3. Objetivos

### 1.3.1. Objetivo General

Estudiar y comprender la percepción de calidad de las soluciones del *VRP* que tienen los clientes de SimpliRoute, y particularmente su relación con la *VA*, para generar conocimientos que permitan entregar resultados más cercanos a lo esperado por sus usuarios.

### 1.3.2. Objetivos Específicos

- Definir métricas cuantitativas que reflejen las características que los usuarios de SimpliRoute buscan en las soluciones.
- Identificar, categorizar y estudiar las propiedades de las métricas que se han utilizado para cuantificar la *VA* en la literatura (en el contexto del *VRP*).
- Caracterizar la relación que tienen las medidas de *VA* con las medidas tradicionales de optimalidad en el *VRP*.
- Evaluar la importancia que tienen las medidas de *VA* en la *calidad percibida* de los usuarios.
- Medir la sensibilidad de los clientes a cambios en las medidas de *VA*.

## 1.4. Hipótesis

Tomando como referencia el conocimiento basado en la experiencia de distintos empleados de SimpliRoute, se espera encontrar razones relevantes para preferir soluciones con mayor *VA*, pero también que una parte de esa preferencia esté asociada a costumbres o creencias equivocadas sobre cómo esperan que sean las “buenas soluciones”. Por lo mismo, se hipotetiza que las medidas tradicionales de calidad resultarán importantes para entender la *calidad percibida*, pero no suficientes como para explicarla totalmente.



Además, se espera que las variables que miden la intersección entre las rutas y la compacidad de éstas, estén asociadas positivamente con la *calidad percibida*, y que los usuarios estén dispuestos a perder un poco de calidad tradicional en pos de ganar en términos de *VA*. Más aun, se conjetura que el *trade-off* entre calidad tradicional y percibida será en general bajo, y que para instancias con distribuciones de puntos clusterizadas geográficamente será más bajo todavía.

## 1.5. Alcances

A pesar de su relevancia, no se encuentra dentro de los objetivos de este proyecto de Tesis, por la naturaleza de la pregunta, el entender el mecanismo que regula en las personas, la relación entre el *atractivo visual* de las soluciones y la *calidad percibida* de las mismas.

No se estudiará cómo afectan las distintas restricciones posibles de las instancias (como las ventanas de tiempo, la carga y demanda, el número y las características de vehículos disponibles, etc.) en la relación de las medidas de *belleza* entre ellas, ni con los objetivos tradicionales. Tampoco el cómo influye la localización de los *depots* en éstas.

Otro punto importante que no está dentro de los alcances del proyecto es investigar las posibles metodologías algorítmicas de generación de soluciones con mejores niveles de *VA*, ni la factibilidad y/o simpleza de implementación de las diversas medidas de *belleza* dentro de éstas.

## 1.6. Estructura de la tesis

El resto de esta tesis se estructura de la siguiente manera; en el capítulo 2 se define formalmente el problema a trabajar y se revisa la literatura existente sobre *VA* en el *VRP*. El capítulo 3 describe las entrevistas exploratorias realizadas a clientes de SimpliRoute y resume los resultados obtenidos en éstas, y el capítulo 4 estudia en detalle las propiedades de las medidas de *VA* y su relación con las funciones objetivo tradicionales del *VRP*. En el capítulo 5 se presentan los resultados obtenidos por dos estudios enfocados en evaluar la relevancia

que tienen las medidas de  $VA$  para los usuarios de la plataforma, y el trabajo termina con las conclusiones y una discusión sobre los trabajos futuros en el capítulo 6.

# Capítulo 2

## Marco Conceptual

El problema abordado en este trabajo tiene como base el *VRP*. Este problema consiste en, dado un conjunto de puntos en un plano y un conjunto de vehículos, definir rutas para estos de forma que cada punto sea visitado una y solo una vez, y que se minimice la distancia, el tiempo o el costo total de los viajes. El *VRP* junto a sus múltiples variantes, ha sido ampliamente estudiado desde que fue introducido por Dantzig y Ramser [15] en 1959. El interés en su estudio proviene tanto de la teoría como de la práctica, por su relevancia en la logística y en la optimización combinatorial. Como este problema puede ser visto como una generalización del problema del vendedor viajero (*TSP* por sus siglas en inglés) en la cual los puntos a visitar se pueden repartir entre varios vendedores, pertenece a la clase de complejidad NP-Difícil.

En la sección 2.1 se profundizará en su formulación matemática, en sus variantes y en sus métodos de resolución, y luego en la sección 2.2 se abordará el ruteo de vehículos desde el punto de vista comercial. En la sección 2.3 se introducirán las medidas de evaluación de *VA*, y se terminará el capítulo con una revisión de la literatura relacionada a éstas.

### 2.1. Formulación genérica del problema de ruteo de vehículos

Para definir formalmente, el *VRP* se considera un digrafo  $G = (N, A)$ , donde  $N = \{0, 1, \dots, n\}$  es el conjunto de nodos y  $A = \{(i, j) : \forall i, j \in N, i \neq j\}$  el conjunto de arcos. El

nodo 0 es el punto de partida y llegada de todos los vehículos y es llamado *depot* (palabra en inglés que significa bodega o depósito). Cada arco  $(i, j) \in A$  tiene asociado un costo  $c_{ij} \geq 0$  correspondiente al tiempo de viaje que toma a un vehículo recorrer el tramo entre los nodos  $i$  y  $j$ . Además, se considera un conjunto de vehículos  $K = \{1, \dots, |K|\}$ .

En esta forma de modelar el problema, la variable de decisión binaria  $x_{ij}$  representa si es que en la solución se utiliza o no el arco  $(i, j)$  para viajar a través de él.

Considerando todo lo anterior, la formulación es:

$$\text{minimizar} \quad \sum_{(i,j) \in A} c_{ij} x_{ij} \quad (2.1a)$$

$$\text{s.a.} \quad \sum_{i \in N} x_{ij} = 1 \quad \forall j \in N \setminus \{0\}, \quad (2.1b)$$

$$\sum_{j \in N} x_{ij} = 1 \quad \forall i \in N \setminus \{0\}, \quad (2.1c)$$

$$\sum_{j \in N} x_{0j} = |K|, \quad (2.1d)$$

$$\sum_{i \in S} \sum_{j \in S} x_{ij} \leq |S| - 1 \quad \forall S \subseteq N \setminus \{0\}, S \neq \emptyset, \quad (2.1e)$$

$$x_{ij} \in \{0, 1\} \quad \forall (i, j) \in A. \quad (2.1f)$$

Así, (2.1a) indica que se busca minimizar la suma del tiempo de viaje total de todos los vehículos, (2.1b) y (2.1c) aseguran que sólo 1 vehículo entre y salga de cada nodo, mientras que (2.1d) fuerza a que salgan exactamente  $|K|$  vehículos del *depot*. Por último, (2.1e) corresponde a la restricción de eliminación de *subtours*, y (2.1f) indica la naturaleza de las variables.

Tomando esta formulación como base se pueden realizar diferentes extensiones. Drexler [17] revisa y clasifica las variantes más comunes desde la teoría y la práctica. Algunas de éstas son:

1. **VRP con Capacidades (CVRP)**: Es una de las variaciones más utilizadas, y considera que cada nodo tiene asociada una demanda  $d_i \geq 0$ ,  $\forall i \in N \setminus \{0\}$ , y que los vehículos tienen una capacidad  $q_i = q > 0$ ,  $\forall i \in K$ . Es posible extender esta variante

aún más, considerando que los nodos tienen más de una demanda que satisfacer, y los vehículos tienen capacidades específicas asociadas a cada una de estas demandas.

2. **VRP con Ventanas Horarias (VRPTW)**: Este tipo de problemas extiende el VRP clásico agregando restricciones para cada nodo, que sólo permiten la llegada de vehículos a tal punto en un rango horario  $[a_i, b_i]$ ,  $\forall i \in N \setminus \{0\}$ . Es posible asignar una ventana horaria al *depot* para definir una duración máxima de la jornada de trabajo de los conductores. Esta variante también puede extenderse más, considerando que los nodos tengan múltiples ventanas horarias disjuntas en las que sea posible visitarlos.
3. **VRP con Depots Múltiples (MDVRP)**: Esta extensión considera que los vehículos tienen un *depot* de inicio distinto al lugar donde deben de terminar. Una variante similar a esta es el *VRP con Finales Abiertos* en la cual los vehículos deben de partir todos desde el mismo *depot*, pero no hay restricciones asociadas a dónde deben de terminar al finalizar su ruta.
4. **VRP con Flota Heterogénea (HFVRP)**: Como su nombre lo indica, esta variante considera que existen distintos tipos de vehículos disponibles para resolver el problema. Las diferencias pueden estar en la capacidad de cada uno, en sus costos fijos o variables, en su velocidad de traslado, en los clientes a los que pueden visitar, y/o en los puntos de inicio y fin que éstos tienen.
5. **VRP con uso Múltiple de Vehículos**: En esta versión del *VRP*, en vez de restringir a que cada vehículo pueda salir sólo una vez del *depot*, se permite que una vez que éstos terminen una ruta y vuelvan al *depot*, puedan salir a servir clientes nuevamente (si es que todavía tienen tiempo disponible).
6. **VRP Estocástico (SVRP)**: Se le dice VRP estocástico a los problemas de ruteo de vehículos en los que uno o varios de los elementos del problema son aleatorios. Algunos ejemplos de características que podrían considerarse son: el tiempo de viaje entre los puntos, la demanda de cada punto, el tiempo de servicio de éstos, o incluso si los distintos puntos van a estar presentes o no.
7. **VRP Dinámico o Dependiente del Tiempo (TDVRP)**: La diferencia de este problema con la versión básica es que en este caso los costos o tiempos de viaje  $c_{ij}$

asociados a cada arco no son constantes, si no que varían según la hora a la que se comienza a usar cada arco.

Todas las variantes que agregan restricciones al problema también pueden ser consideradas en su versión *suave*, es decir, no requerir que éstas se cumplan estrictamente, si no que penalizar en la función objetivo las situaciones en las que esto no ocurre.

Como el caso que se estudia en este trabajo está motivado en problemas del *mundo real*, se hace necesario incluir simultáneamente la mayoría de las extensiones listadas a la hora de modelar el problema, para poder entregar soluciones que se ajusten a la operación particular de cada cliente. La decisión de incluir o no cada una de las restricciones influye fuertemente en la efectividad y eficiencia de los distintos algoritmos existentes para resolver el *VRP*. En la sección siguiente se revisan los métodos de resolución más comunes, junto con sus ventajas y desventajas.

### 2.1.1. Métodos de resolución

Si bien el *VRP* es una extensión del *TSP* conceptualmente sencilla, las adaptaciones de los exitosos métodos exactos de resolución existentes para este último problema no han alcanzado el mismo nivel de éxito en este nuevo campo [76]. Este escenario, junto con la alta relevancia práctica de este problema para la logística, han fomentado el estudio y desarrollo de un gran número de diversas metodologías de resolución, tanto exactas como heurísticas. A continuación se mencionan brevemente algunas de éstas. Para una revisión detallada, referirse a los capítulos 2, 3 y 4 del libro *Vehicle Routing: Problems, Methods and Applications* [83].

1. **Métodos Exactos:** Se basan en modelar el *VRP* como un problema de optimización lineal entera, y resolverlo utilizando técnicas de esta rama de las matemáticas, como lo son el *Branch and Bound*, la generación de *Cortes Válidos* y la *Generación de Columnas*. Como las distintas formulaciones posibles del problema tienen ya sea un número exponencial de restricciones o de variables, estos métodos sólo suelen ser capaces de resolver instancias de tamaño limitado.

- **Branch and Bound:** La mayoría de los acercamientos que utilizan esta técnica resultan de replicar las metodologías que resultaron exitosas para resolver el *TSP*. Dos de éstas (propuestas por Christofides y Eilon [10]) consisten en encontrar cotas para la poda resolviendo el problema simplificado, descartando las *Restricciones Generalizadas de Eliminación de Subtours (GSEC)* ya sea totalmente (transformando el problema en un *Problema de Transporte*) o parcialmente (de forma tal que sólo se garantice que la solución sea conexa, pero no necesariamente que los nodos tengan el grado correcto, transformando el problema en encontrar un *Árbol Recubridor Mínimo*). Una mejora a estos métodos, propuesta por Fisher [18] consiste en agregar a la función objetivo, de manera penalizada, algunas de las restricciones descartadas, para encontrar cotas más ajustadas.
- **Branch and Cut:** Esta metodología, al igual que la anterior, se basa en descartar algunas de las restricciones del problema, pero la diferencia está en que adicionalmente se descartan las restricciones de integralidad. Una vez resuelto el problema relajado, se evalúa si es que algunas de las restricciones que fueron descartadas son violadas por la solución encontrada, y en caso de que esto ocurra se agregan tales restricciones al problema y se vuelve a iterar. Cuando no se encuentran restricciones violadas, se ramifica el problema en los nodos con soluciones racionales. La primera aplicación de esta metodología al *VRP* fue realizada por Laporte, Nobert y Desrochers [45] el año 1985. Ellos proponen descartar las *Restricciones Generalizadas de Eliminación de Subtours (GSEC)*, y utilizar una heurística para encontrar las restricciones violadas, para luego agregarlas al problema junto con los cortes de Gomory identificados. Una versión más moderna de esta metodología es la propuesta por Lysgaard, Letchford y Eglese en [52] el 2004.
- **Generación de Columnas:** Los trabajos que utilizan esta metodología suelen usar una formulación alternativa del *VRP* basada en el *Problema del Conjunto de Cobertura (SCP)*, propuesta por Balinski y Quandt [1]. Esta formulación considera una variable de decisión binaria  $\lambda_r, r \in \Omega$ , donde  $\Omega$  es el conjunto de todas las rutas factibles y  $\lambda_r = 1$  si y solo si la ruta  $r$  es parte de la solución óptima. Como  $|\Omega|$  puede ser exponencialmente grande, la *Generación de Columnas* es particularmente útil para resolver el problema modelado de esta forma. Las ventajas de esta

formulación son que su relajación lineal suele entregar buenas cotas inferiores, y que la estructura de costos y restricciones que representan la factibilidad de cada ruta están implícitamente integradas en la definición del conjunto  $\Omega$  y del vector de costos  $(c_r)_{r \in \Omega}$  [76].

- **Branch and Cut and Price:** Propuesta por Fukasawa *et al.* [20], consiste en utilizar conjuntamente las versiones más exitosas de las tres técnicas descritas anteriormente. Variantes de esta metodología, son las que han logrado mejor desempeño a la fecha, logrando resolver consistentemente instancias con hasta 200 nodos [63].

2. **Heurísticas:** Las limitaciones de los métodos exactos, asociadas al tamaño máximo de instancias que pueden resolver, al largo y variable tiempo de ejecución que utilizan, y a la poca flexibilidad que tienen para introducir nuevas restricciones, los hacen poco factibles de utilizar para resolver problemas reales. Por estas razones es que en la práctica suelen utilizarse metodologías heurísticas, que para instancias de tamaño considerable entregan rápidamente una solución de calidad, aunque posiblemente subóptima.

- **Heurísticas Clásicas:** Este grupo de algoritmos consiste en metodologías desarrolladas específicamente para resolver el *VRP*. Dentro de éstas, se pueden identificar las *Heurísticas Constructivas*, cuyo objetivo es generar una buena solución factible, y las *Heurísticas de Mejora*, que reciben un solución factible generada previamente, y buscan mejorarla para encontrar soluciones más cercanas al óptimo.

Un ejemplo clásico de una *heurística constructiva* es la heurística de ahorro de Clarke y Wright [11]. Este algoritmo comienza generando rutas  $r_i = (0, i, 0)$ ,  $\forall i \in N \setminus \{0\}$  y las va uniendo secuencialmente según una regla de *ahorro*. Específicamente, dos rutas  $r_1 = (0, \dots, i, 0)$  y  $r_2 = (0, j, \dots, 0)$  se unen en una nueva ruta  $r_3 = (0, \dots, i, j, \dots, 0)$  si es que esta unión es factible y el valor del ahorro generado  $s_{ij} = c_{i0} + c_{0j} - c_{ij}$  es el mayor posible. Este proceso continúa hasta que no haya más pares de rutas que sea factible unir.

Las *heurísticas de mejora* consisten en realizar movimientos de *búsqueda local* que permiten explorar las vecindades de la solución actual en el espacio de soluciones. Algunos de éstos movimientos son los mismos utilizados para resolver el *TSP*, los



cuales se aplican a cada una de las rutas, por lo que son llamados movimientos *intra-ruta*. Un ejemplo de éstos es el *2-OPT* propuesto por Lin [48], en el que se remueven dos arcos de una ruta y se agregan dos nuevos arcos, de manera que la ruta siga siendo continua. Otros movimientos desarrollados específicamente para el *VRP* son los que involucran a más de una ruta al mismo tiempo, por lo que son llamados movimientos *inter-ruta*. Ejemplos de este tipo son el *RELOCATE* (remover un número de clientes consecutivos de una ruta e insertarlos en otra), *SWAP* (intercambiar un número de clientes consecutivos entre dos rutas) y *2-OPT\** (remover un arco de dos rutas e insertar dos nuevos arcos, de forma que las rutas queden cruzadas).

- **Metaheurísticas:** Las metaheurísticas son algoritmos generales de optimización heurística, que han sido aplicados al *VRP* con mucho éxito. Algunos de los más utilizados son:

- a) **Simulated Annealing (SA):** Esta metodología es similar a las *heurísticas de mejora*, puesto que recibe una solución factible y utiliza iterativamente movimientos de *búsqueda local* para encontrar mejores soluciones. La diferencia está en que usa una técnica probabilística que permite hacer algunos movimientos que empeoran la solución, para escapar de los óptimos locales y así producir soluciones de mayor calidad. Un ejemplo de este tipo de algoritmos aplicado al *VRP* es el propuesto por Osman [61].
- b) **Búsqueda Tabú (TS):** La *Búsqueda Tabú* también usa movimientos de *búsqueda local*, pero utiliza una técnica distinta para salir de los óptimos locales. En este caso se mantiene una lista de *soluciones prohibidas*, y en cada iteración del algoritmo se realiza el movimiento que lleva a la mejor solución que no está en la lista. De esta forma, si es que todos los movimientos que mejoran la solución actual llevan a soluciones que están en tal lista, el algoritmo elegirá un movimiento que empeore la solución, permitiendo así explorar una mayor parte del espacio de soluciones, llevando posiblemente a un mejor resultado. Cordeau, Laporte, y Mercier [14] proponen una implementación de esta metaheurística.
- c) **Optimización por Colonias de Hormigas (ACO):** Este tipo de algoritmos basa su funcionamiento en imitar el comportamiento de búsqueda de alimento

de las colonias de hormigas reales, simulando hormigas artificiales que recorren el espacio de soluciones buscando llegar a soluciones de mayor calidad. Cada hormiga elige hacia dónde moverse en el espacio de soluciones utilizando una regla probabilística que depende de la información generada por éstas mismas en iteraciones anteriores. Cuando una hormiga llega a una solución, se actualiza la información asociada al camino que recorrió para llegar a tal solución, incrementando o disminuyendo la probabilidad de ser escogido por otra hormiga en el futuro, dependiendo de la calidad de la solución encontrada. Un ejemplo de esta metodología aplicada al *VRP* se puede encontrar en Reimann, Doerner, y Hartl [67].

d) **Algoritmos Evolutivos (EA):** Estas metaheurísticas están inspiradas en la selección natural, y utilizan este concepto para generar iterativamente soluciones de mejor calidad. Durante la optimización se mantiene un conjunto de soluciones (*individuos*) llamado *población*, del cual se elige un subconjunto que será utilizado para *reproducirse* y así generar nuevas soluciones que formarán una nueva *generación*. Cada *individuo* de la *población* tiene asociado un puntaje asociado a la calidad de tal solución, que influye en la probabilidad de que éste sea elegido para *reproducirse* en la próxima iteración. Una de las implementaciones de este tipo de algoritmos que mejores resultados ha alcanzado es la de Nagata y Bräysy [59].

- **Metodologías Híbridas:** En el capítulo dedicado a heurísticas escrito por Laporte, Ropke y Vidal [46], sugieren que los algoritmos más modernos utilizados para resolver el *VRP* suelen utilizar técnicas provenientes de distintas metodologías, afirmando que *“a medida que la investigación avanza, hemos visto la emergencia de una amplia variedad de métodos híbridos que se basan en conceptos tomados de varios paradigmas algorítmicos como búsqueda local, vecindades de gran escala, inteligencia colectiva y búsquedas basadas en poblaciones, perturbaciones, programación entera, búsquedas en árboles, minería de datos, y computación paralela. La frontera entre los métodos se está volviendo cada vez más borrosa”*.

Dos ejemplos de propuestas híbridas interesantes son la de Prins [65] y la de Groër, Golden y Wasil [25]. La primera utiliza una estrategia evolutiva manteniendo una

*población* de soluciones en cada iteración, pero además aplica una técnica asociada a la metodología GRASP (Greedy Randomized Adaptive Search Procedure [68]), consistente en utilizar distintos puntos de partida para el algoritmo, junto con realizar una serie de *búsquedas locales* aleatorias en cada solución de la *población*. La segunda combina procedimientos de *heurísticas de mejora* y de *programación entera* (técnica llamada *mateheurística* [54]), generando un conjunto de rutas de alta calidad mediante la primera metodología y luego utilizando este conjunto para alimentar un programa de optimización entero utilizando la formulación basada en el *Problema del Conjunto de Cobertura (SCP)*.

## 2.2. Ruteo de Vehículos Comercial

El *VRP* es un problema importante tanto en la teoría como en la práctica, lo que se ve reflejado en los miles de artículos científicos publicados sobre el tema, y en la fundación de numerosas empresas que venden software de ruteo de vehículos comercial (*CVRSS*) en todo el mundo [17].

Una de las características que diferencia los problemas que se deben de resolver en el ámbito comercial de los que se abordan en el área académica, es que en estos últimos se suelen incorporar a los modelos sólo algunas de las extensiones mencionadas en la sección 2.1, mientras que en los primeros es común encontrar la presencia simultanea de múltiples y diversos tipos de restricciones y objetivos, pues cada organización tiene distintos modos de operar y de evaluar sus resultados. Drexl menciona en [17] que los resultados encontrados en su investigación apoyan la afirmación hecha por Sörensen et al. [79] que dice que *‘aunque hay un creciente foco científico en los llamados VRP “ricos” (que incorporan restricciones y objetivos más complejos), éstos no han alcanzado de ninguna manera toda la complejidad de los problemas de ruteo de la vida real’*.

Otra dificultad importante que enfrentan las empresas que ofrecen *CVRSSs* es que sus clientes suelen ser poco pacientes con los tiempos de ejecución de los algoritmos, esperando obtener soluciones para problemas de gran tamaño en pocos minutos. Este problema tampoco está resuelto totalmente, puesto que todavía hay una gran brecha entre los requisitos que

tienen los usuarios, y el rendimiento que alcanzan las aplicaciones actuales de optimización de rutas [7]. Hall y Partyka realizan bianualmente una encuesta sobre *CVRSs* y publican sus resultados en la revista *OR/MS Today* de *INFORMS*. En la versión del año 2014 [28] remarcan la importancia de brindar tiempos de ejecución cortos, y comentan que *“los desarrolladores y vendedores de software de ruteo de vehículos se encuentran bajo la inmensa presión de satisfacer las crecientes demandas de sus clientes, de obtener soluciones de ruteo de vehículos [...] en tiempo real”*.

Una última diferencia clave está en cómo se evalúan los resultados. En el área académica se suele medir la calidad de una solución como la distancia total recorrida o el tiempo total usado, pero en la práctica esto suele no ser suficiente, puesto que en las empresas que contratan *CVRSs* existen por lo menos tres actores que evalúan los resultados con distintos criterios: los altos mandos, los planificadores (que son los que utilizan directamente el *CVRS*) y los conductores. Los altos mandos suelen estar preocupados de que el uso de la herramienta que están adquiriendo traiga beneficios a la empresa, por lo que soluciones que minimizan la distancia total recorrida o el tiempo total usado, están dentro de lo que éstos prefieren. Por otro lado, los planificadores cuentan con conocimientos y con experiencias previas relacionadas al funcionamiento y a la logística interna de la empresa, que son difícilmente integrables en los modelos, y que los pueden hacer preferir soluciones que recorran mayor distancia, pero que sean mejores con respecto a otras condiciones relevantes para su operación. Cabe mencionar que hay veces que estos conocimientos previos son en realidad creencias equivocadas, pero incluso en estos casos es importante para los *CVRSs* tener herramientas que permitan abordar estos temas, pues favorecen el uso fluido de la plataforma, y además entregan la posibilidad de evaluar de forma más concreta cuáles los costos en los que incurre la empresa por decidir seguir estas creencias. Por último, hay que considerar que si los conductores no siguen las rutas que les entrega el planificador, los ahorros generados por el uso de la plataforma pueden verse fuertemente disminuidos.

Así, los algoritmos que necesitan los *CVRSs* no sólo deben de ser capaces de entregar soluciones para problemas de gran tamaño, que consideran múltiples restricciones simultáneamente, y de poder generarlas rápidamente, sino que además éstas deben de ser de buena calidad con respecto a la distancia total recorrida, deben de considerar las idiosincrasias

particulares asociadas a cada empresa y a cada planificador, y al mismo tiempo deben de incentivar a que cada conductor la cumpla de la manera más fielmente posible. Más aun, las metodologías que utilizan deben de ser lo suficientemente flexibles como para poder adaptarse a las cambiantes necesidades y expectativas de los clientes, en un contexto en el que a menudo a éstos les resulta difícil dar una definición clara de los objetivos y las limitaciones de su problema [17].

### 2.2.1. SimpliRoute

SimpliRoute es una empresa chilena fundada en enero de 2015, que ofrece un *CVR*S en el modelo de *SaaS*. Actualmente se encuentra presente con oficinas comerciales y de desarrollo en 6 países de Latinoamérica, y cuenta con más de 400 clientes en 23 países. Cada día, más de 8.000 vehículos usan su plataforma, realizando más de 120.000 entregas.

Como es común en los *CVR*Ss [17], SimpliRoute provee de una *interfaz a una base de datos* que permite leer la información relevante y guardar las soluciones generadas, un *sistema de información geográfica (GIS)* utilizado para geolocalizar datos de direcciones, calcular matrices de distancia y visualizar los datos y las soluciones en mapas, un *módulo de planificación* que es donde se realiza la optimización heurística y que al mismo tiempo permite editar manualmente los planes, un *módulo de telemática* que gestiona el intercambio de datos entre los vehículos y la oficina de despachos, así como el seguimiento de los vehículos, y por último un *módulo de estadísticas* que sirve para calcular indicadores de rendimiento clave y para crear informes.

SimpliRoute ha sido reconocido dentro del mundo del emprendimiento en distintas ocasiones, siendo ganadores del premio *Startup Chile*, recibiendo una mención honrosa en el *eCommerce Startup Competition 2015*, y siendo aceptados en la incubadora *500 Startups*. Además, tiene acuerdos comerciales con diversos *partners*, dentro de los que destacan *Telefónica* y *TomTom Telematics*.

## 2.3. Medidas de evaluación de Atractivo Visual

Dentro de las diferencias comentadas en la sección 2.2, la relacionada con la forma de evaluar la calidad de los resultados es la que más tiene relación este trabajo. Varios investigadores han notado que los usuarios de *CVR*Ss tienden a preferir soluciones *visualmente atractivas* [2, 13, 17, 24, 30, 31, 33, 34, 36–39, 41, 44, 50, 51, 55, 64, 71, 72, 81, 87], aunque algunos [7] comentan además que no es claro si es que esta preferencia está justificada y genera mejores resultados en la práctica, o si sólo está asociada a las falsas creencias de que este tipo de soluciones son más eficientes.

Poot et al. [64] fueron los primeros en utilizar este concepto en el contexto del *VRP*, y comentan que *‘Hemos notado que nuestros clientes a menudo usan medidas adicionales para decidir si un plan es aceptable o no. En particular, los planificadores operacionales tienden a preferir planes “visualmente atractivos”. Por ejemplo, les desagradan los planes en los que muchos viajes se cruzan. Los planes visualmente atractivos parecen ser más lógicos para ellos y más cercanos a la forma manual de trabajo. Por lo tanto, producir dichos planes ayuda a crear confianza entre los planificadores (y los controladores) en el sistema de planificación de rutas, lo que lleva a una rápida aceptación del sistema’*. Por otro lado, advierten que *“en una etapa posterior, cuando el sistema de ruteo sea totalmente aceptado, el gerente de logística de la empresa puede cambiar los parámetros del algoritmo para enfocarse más en la eficiencia en lugar del atractivo visual”*. Otras razones reportadas en la literatura como argumento para considerar este aspecto son:

- Facilita la implementación y el uso, pues genera confianza entre el software de ruteo y los planificadores, y también entre los conductores y las rutas generadas para éstos [2, 6, 33, 34, 41, 44, 50, 51, 71].
- Ayuda a aumentar la familiaridad del conductor con las calles y los clientes de la región particular, mejorando así la eficiencia de las operaciones [2, 24, 41, 51, 81].
- Produce soluciones más robustas y fáciles de modificar frente a cambios imprevistos sobre la marcha, como condiciones de tráfico no consideradas, problemas con alguna entrega o introducción de una nueva parada [13, 24, 33, 51].

- Permite modelar indirectamente restricciones y funciones objetivo que de otra forma sería difícil integrar en el modelo [2, 7, 13, 17, 36–39].
- Se hace cargo de la falsa creencia de que soluciones que no cumplen con estas características son poco eficientes, empeorando lo menos posible el valor de la función objetivo clásica [7, 71, 81].

A pesar de su relevancia, no existe un consenso sobre la definición formal del término *atractivo visual* en este contexto. Algunos de autores [2, 13, 51, 71] han notado que aunque el término es intuitivo, su subjetividad hace que sea difícil de definir cuantitativamente. A pesar de esto, la mayoría está de acuerdo en que hay al menos dos condiciones que son necesarias para que una solución sea considerada *visualmente atractiva*: que sus rutas se crucen entre ellas lo menos posible (*sectorización*), y que cada una de éstas sea lo más *compacta* que se pueda (es decir, que los puntos de cada ruta estén a distancias relativamente pequeñas entre sí).

Tampoco existe consenso sobre el nombre con el que referirse a este concepto; en estos trabajos se hace referencia a éste con diversos nombres, algunos de éstos son: atractivo o belleza visual, rutas bonitas, atractivo y calidad estética de las rutas y soluciones, medidas de calidad o restricciones no estándar, medidas cualitativas, objetivos subjetivos, métricas o restricciones de forma, forma de la solución, penalizaciones de forma y de compacidad, compacidad de las rutas, baja superposición de rutas, restricciones regionales, y soluciones operacionalmente robustas.

Otro aspecto relevante que se comenta en la literatura, y en el que hay algunos resultados contradictorios, es cómo se relaciona el *atractivo visual* con las medidas tradicionales de calidad del VRP. En algunos trabajos [50, 55, 64] se reporta que al encontrar soluciones más *atractivas visualmente* también se mejoran las medidas clásicas, y en otros casos [2, 13, 33, 34, 51, 72] se reporta lo contrario, aunque con distintos grados de *trade-off*.

Todos estos desacuerdos han dificultado el estudio sistemático de este tipo de características, lo cual se ve reflejado en que todavía no existen sets de instancias ni *benchmarks* con las cuales compararse consistentemente [30, 31].

Finalmente, dado el contexto de este trabajo, vale la pena notar el comentario realizado en las conclusiones del trabajo de Lum et al. [51]: “*Aunque hemos propuesto un método para generar rutas que se ajustan a las nociones de calidad en las métricas estéticas, no está claro que estas métricas capturen realmente las intuiciones para las que fueron diseñadas. Cada métrica tiene un principio de motivación central. Sin embargo, para cada característica visual dada, obviamente hay muchas maneras de intentar medirla. [...] Como el objetivo es crear rutas que un administrador de distribución quisiese implementar, sería útil ver qué tan bien estas métricas coinciden con las evaluaciones de los gerentes. Valdría la pena realizar una encuesta a los gerentes para investigar qué métricas corresponden a sus evaluaciones subjetivas de los diferentes planes de ruteo.*”.

## 2.4. Trabajos Relacionados

En esta sección se revisan cronológicamente los distintos trabajos que consideran dentro de sus modelos y/o algoritmos, directa o indirectamente, medidas asociadas a *atractivo visual* con el objetivo de generar rutas más *bonitas*. Cabe mencionar que existen metodologías para resolver problemas de ruteo de vehículos que generan soluciones que tienden a ser más *atractivas visualmente* por otras razones (como las llamadas *cluster-first route-second* [19,22]). En esta sección sólo se consideran los trabajos que declaran como uno de los objetivos de su trabajo el generar rutas más *bonitas*.

Como se comentó en la sección anterior, el primer trabajo que considera medidas de *VA* en el contexto del *VRP* es el de Poot, Kant y Wagelman [64] publicado el año 2002. Esta investigación fue realizada en conjunto con *ORTEC*, una consultora holandesa especializada en investigación de operaciones aplicada, y uno de los proveedores de *CVRSS* más grandes del mundo. Ellos proponen un algoritmo basado en la *heurística de ahorro*, que no considera directamente medidas de *VA* en su ejecución, pero al momento de comparar los resultados a los que llega este nuevo algoritmo con su heurística anterior, miden su calidad no sólo en términos tradicionales, si no que incorporan 6 medidas de calidad no estándar también. El nuevo algoritmo propuesto resulta ser superior al que utilizaban anteriormente, al ser testeado en un set de datos de la *vida real* presentado por ellos mismos, tanto en las medidas



tradicionales como en las no estándar.

El año 2004 Kim, Kim y Sahoo [43] presentan un trabajo en una conferencia, en el cual proponen un algoritmo para un solucionar un *VRP* en el contexto de un problema de recolección de basura, con restricciones de horario de almuerzo para sus conductores, y que además permite que los camiones hagan múltiples salidas durante el día, debiendo decidir a qué centro de descarga asistir cuando un vehículo se llena. Además de los objetivos tradicionales, incluye el de generar soluciones balanceadas, y “mejorar el atractivo visual” de la solución.

La heurística propuesta consiste en una fase inicial de construcción de rutas y una segunda fase de mejora. El grueso del trabajo se encuentra en la primera fase, que se basa en la estrategia *cluster-first route-second*. Se presentan dos alternativas para generar clusters balanceados: una ejecuta una versión del algoritmo *k-means* modificada y la otra se basa en un algoritmo para solucionar el problema de emparejamientos estables. En la segunda fase, los puntos que no están asignados a sus rutas más cercanas son reasignados a tales rutas, si es que ese movimiento mejora la calidad general de las rutas (usando como función objetivo una suma ponderada de dos medidas de *VA* previamente definidas), sin violar sus restricciones de capacidad. Para evaluar la calidad de sus algoritmos de clustering, estudian los resultados que éstos alcanzan en el set de datos del problema de clustering capacitado de Osman y Christofides [60], y muestran que logran encontrar mejores asignaciones en 9 de las 20 instancias.

Al año siguiente, los mismos autores (más Kraas y Popov Jr.) publican un trabajo [72] en el que extienden el trabajo realizado el año anterior. En esta ocasión sólo estudian la versión de la heurística que utiliza la versión modificada del *k-means*. Al compararse con las mejores soluciones conocidas de las instancias de Solomon [78], usando las medidas de evaluación tradicionales, su algoritmo no logra resultados tan favorables, pero los autores describen que éste sí suele generar soluciones visualmente más atractivas (aunque no reportan los resultados cuantitativamente).

El siguiente trabajo publicado fue el de Tang y Miller-Hooks [81] a principios del 2006. Ellos proponen una heurística interactiva que comienza construyendo rutas (posiblemente no factibles) con buena *forma*, basándose en puntos de referencia entregados por los planificadores.

Luego se realiza una fase de mejora, que considera simultáneamente el costo clásico y 2 medidas de *VA*, para hacer movimientos que logren factibilidad en la solución empeorando lo menos posible tales medidas. Para evaluar los resultados obtenidos, consideran dos instancias generadas con información de entregas de *FedEx* para una ciudad particular, y las evalúan en términos de las medidas utilizadas durante la ejecución. Los resultados de sus experimentos sugieren que considerar medidas de *atractivo visual* dentro de la heurística, puede lograr generar soluciones con mucho mejor *forma*, manteniendo resultados satisfactorios en términos de las medidas tradicionales del ruteo de vehículos.

Jang et al. [34] realizan una consultoría para la *lotería de Missouri*, en la que desarrollan una *heurística de mejora*, para generar rutas que minimicen la distancia total de viaje manteniendo un balance en las cargas de trabajo. Una vez generadas las soluciones, por sugerencia los gerentes y de los representantes de ventas de lotería, modificaron varias de las rutas de éstas para disminuir las intersecciones y generar una solución más *clusterizada*, incluso cuando los cambios se produjeron a expensas de la eficiencia.

Los próximos en integrar medidas de *VA* en la ejecución de su heurística fueron Lu y Dessouky [50]. Ellos proponen una *heurística constructiva de inserción*, que además de considerar el costo tradicional de distancia recorrida a la hora de decidir insertar un punto en una ruta, también considera la reducción de la holgura disponible con respecto a las ventanas de tiempo relacionada con tal inserción. Además, incluyen una medida de *belleza intra-ruta* pero ninguna de *inter-ruta*, pues argumentan que para el primer caso existen trabajos [53] que sugieren que la *belleza* de una ruta única refleja implícitamente su largo total, pero que no hay estudios que muestren una relación de ese tipo para la *belleza inter-ruta*. Comparan los resultados obtenidos por su algoritmo en las instancias de Li y Lim [47] con los generados por otras dos *heurísticas de inserción* (una secuencial y una paralela), y además con los obtenidos por los autores que construyeron tales instancias. Concluyen que su algoritmo tiene mejor desempeño (en términos tradicionales y de *belleza*) que las primeras dos heurísticas, pero que el número de vehículos usados y el tiempo total de viaje obtenido fue ligeramente peor que el obtenido en [47] (lo que es esperable pues los tiempos de ejecución son mucho más bajos). No pudieron compararse en términos de *atractivo visual* con las soluciones de este último trabajo, pues ellos no reportaron las soluciones completas, sólo los valores de las medidas tradicionales

alcanzados por éstas.

Ese mismo año, Zhou et al. [87] publican un trabajo altamente similar al recién mencionado.

Finalizando el año 2006, Kim, Kim y Sahoo [44] publican un trabajo que extiende los que publicaron los años anteriores [43, 72]. En este caso, definen un set de instancias de prueba para su problema, y evalúan los resultados de su algoritmo con los alcanzados por una *heurística de inserción y mejora*. Miden los resultados considerando las medidas tradicionales (número de vehículos y distancia total recorrida), y también en términos de las dos medidas de *belleza* que proponen, y en términos de balanceo de tiempo total de trabajo. Aunque el número de vehículos obtenidos por el algoritmo basado en *clustering* a veces es mayor que el alcanzado por el de inserción, la solución del primero casi siempre es mejor en los otros aspectos considerados.

A principios del 2008, Kant et al. [41], nuevamente en el contexto de un trabajo realizado por *ORTEC*, reportan el proceso y las mejoras resultantes de implementar su *CVRS* en *Coca-Cola Enterprises*. Utilizan una heurística de construcción, que luego pasa por una fase de mejora utilizando movimientos de *búsqueda local*. Integran una medida de *VA* de forma penalizada en la función objetivo del algoritmo, y permiten que los usuarios ajusten el valor de la penalización, de forma que ellos mismos decidan cuánto están dispuestos a perder en términos de costos tradicionales, para generar soluciones más *bonitas*.

El siguiente artículo publicado fue el de Matis [55]. En este trabajo se propone un *Sistema de Soporte a Decisiones (DSS)* para ayudar a resolver el *Street Routing Problem (SRP)*. Además, se realiza una comparación entre nueve heurísticas para resolver este problema, penalizando la función objetivo con medidas de *VA* de diversas formas. Se presentan 3 medidas de *belleza*, y se propone una fórmula para integrarlas todas en un único indicador de *calidad subjetiva* de la solución. Los resultados obtenidos en un set de instancias creado por ellos mismos, muestran que aunque ninguna heurística domina totalmente a todas las otras, existe una tendencia que sugiere que las que logran mejores resultados en las medidas tradicionales también lo hacen en la medida de *calidad subjetiva*.

Luego, el año 2009 Juan et al. [39] presentan una heurística basada en simulación (después mejorada en [36], [38] y [37]), que agrega un componente aleatorio a la heurística de *ahorro*,

y de esta manera logran explorar el espacio de soluciones más ampliamente. Este algoritmo entrega como resultado una lista de “buenas soluciones”, para que el tomador de decisiones elija cuál prefiere. No consideran medidas de *atractivo visual* explícitamente, pero mencionan que una de las razones por las que es beneficioso entregar un conjunto de soluciones, en vez de sólo una, es que al elegir manualmente la solución, es posible considerar implícitamente este tipo de variables.

Hasle et al. [30,31] presentaron en dos congresos el año 2010, donde hablaron sobre cómo resuelven en *SINTEF* (organización noruega de investigación independiente) todo el proceso asociado al ruteo de vehículos para la distribución de periódicos y productos de medios en Noruega. Como es común en la investigación aplicada, consideran problemas de gran tamaño (miles de puntos), y varias restricciones (modos de transporte mixtos (vehículos y peatones), puntos de recogida variables, consideraciones de la topografía del lugar, etc.). Su antigua heurística consistía en una fase constructiva y luego una de mejora, e intentaron incluir una penalización en la función objetivo para mejorar el *atractivo visual* y el balanceo de las soluciones, pero los resultados no fueron satisfactorios. Por esto proponen un nuevo algoritmo similar al utilizado en [72], que comienza *clusterizando* los puntos usando una versión modificada del algoritmo *k-means* (resolviendo rápidamente un *TSP* dentro de cada *cluster* para estimar la duración de las rutas), y finalmente aplica movimientos de *búsqueda local* (*3-OPT* y posiblemente movimientos *inter-ruta* con restricciones en el deterioro permitido del *atractivo visual* y el balanceo). Mencionan que los resultados preliminares obtenidos en instancias industriales “difíciles” son prometedores, pero que todavía hay algunas instancias en las que no logran buenos resultados.

En el trabajo del año 2011 de Matis y Koháni [56] evalúan 10 métodos de agregación de clientes para reducir el tamaño de los problemas de *SRP* que resuelven (en el set de datos que consideran para este trabajo van desde 11.000 hasta 29.000 clientes), y para generar soluciones más *bonitas*. Modifican 2 de las 3 medidas propuestas en su trabajo anterior [55], y reportan los resultados de evaluar 5 de los métodos de *clusters* en términos de estas 3 medidas, y del número de *clusters* generados, y el promedio del valor aproximado de kilómetros recorridos por un *TSP* en cada *cluster*. La diferencia en el número de *clusters* generados por estos 5 métodos varía más de un 10% en el conjunto de instancias que utilizaron, siendo el algoritmo

de *clustering* difuso, presentado en [55], el que mejor resultados obtuvo.

El año 2012 Caccalano publica su tesis de magister [8] (y un paper asociado a ésta el 2015 [9]), en la cual resuelve un *VRP* para la entrega de material a líneas de montaje. Utiliza dentro de su heurística una de las métricas propuestas por [44] como función objetivo para una fase de mejora basada en *búsqueda local*. En particular, cada movimiento es realizado sólo si es que mejora el indicador de *VA* sin empeorar el costo tradicional en más de un 20 %.

El trabajo de Hollis y Green [33] del mismo año es uno de los más completos a la fecha. Ellos proponen una sofisticada heurística que, manteniendo buenos niveles de *belleza*, se enfoca en minimizar agresivamente el número de vehículos utilizados, característica que los trabajos anteriores en este contexto no habían logrado incorporar satisfactoriamente. Además, este trabajo es el primero que realiza una revisión comprensiva de las medidas de *atractivo visual* de las soluciones del *VRPs*. El algoritmo que proponen consta de dos fases. La primera es similar a la propuesta por Kant et al. en [41], pues modifican la heurística de *inserción secuencial* de Solomon [78] para generar soluciones visualmente más atractivas, y la segunda está basada en el algoritmo de *búsqueda local guiada* de Kilby et al. [42], donde utilizan una función objetivo penalizada con medidas de *VA*. Además, durante su ejecución, el algoritmo evita explícitamente rutas con baja utilización, puesto que éstas pueden ser un obstáculo importante cuando se intenta producir rutas con mejor *atractivo visual*. Cabe mencionar que, dado que el *VRP* que solucionan está motivado por un problema de la *vida real*, éste considera un alto número restricciones.

El algoritmo presentado por Hollis y Green [33] fue probado en un set de 40 instancias construido con información de las entregas diarias realizadas por *Schweppes Australia Pty. Ltd.*, y también en las instancias de Solomon [78]. Para la evaluación consideraron las medidas tradicionales (tiempo de ejecución, número de rutas generadas, distancia total recorrida) y 5 medidas de *VA*. Los resultados obtenidos en el conjunto de instancias propio fueron positivos, generando soluciones con buenos indicadores de *belleza* y bajo número de vehículos utilizados, incluso considerando el alto número de restricciones que tenían las instancias (ventanas horarias ajustadas, restricciones de compatibilidad vehículo-cliente, etc.). En las instancias de Solomon [78], al compararse con las mejores soluciones encontradas hasta ese momento, llegan a resultados que usan entre 11 % y 15 % más de rutas y entre 8 % y 31 %

más de kilómetros, pero que tienen buenos resultados en términos de las medidas de *VA* que consideran (aunque no las comparan con las alcanzadas por las mejores soluciones encontradas hasta ese momento).

El año 2013 Gretton y Kilby [24] publican un trabajo en el que integran dos medidas de *VA* al algoritmo de *ALNS* (*Adaptive Large Neighbourhood Search*) propuesto por Ropke y Pisinger [69] como penalización a la función objetivo. Una de las medidas que utilizan es computacionalmente costosa de calcular, por lo que proponen una técnica que permite, a costas de perder algo de exactitud, reducir el tiempo de cómputo de ésta. La otra de las medidas que consideran no había sido utilizada anteriormente en el contexto del *VRP*. Ellos evalúan los resultados obtenidos por su heurística en las instancias de Solomon [78], en la extensión de éstas realizada por Gehring y Homberger [21], y también en algunas instancias de la *vida real*, notando que logran buenos resultados en términos de *VA*, y empeorando en promedio sólo un 9% en términos de distancia total.

El trabajo de Battarra et al. [2] del año siguiente propone dos formulaciones de programación lineal entera para el *Clustered VRP* (*CluVRP*) (generalización del *VRP* en la que los puntos son agrupados en *clusters*, y donde la solución tiene que cumplir que si un vehículo visita algún cliente de un *cluster*, entonces tiene que visitar a todos los puntos de ese *cluster* antes de salir de él), mostrando algunas propiedades sobre estas formulaciones. Además proponen dos algoritmos exactos para resolver el problema, uno de *ramificación y cortes*, y otro de *branch-and-cut-and-price*. Los resultados alcanzados por estas dos metodologías al aplicarlas a varios sets de datos de la literatura, y también a instancias del *mundo real*, muestran que el algoritmo de *ramificación y cortes* logra mejores resultados. Los autores comentan que la posibilidad de resolver el *CluVRP* de manera exacta es una buena oportunidad para obtener soluciones para el *VRP* que sean al mismo tiempo eficientes en las medidas de costo clásicas, y que muestren un alto grado de *atractivo visual*.

Ese mismo año, Bosch publica su trabajo de tesis de Magister [6] realizada en conjunto con *ORTEC*. Los dos objetivos principales de su trabajo son encontrar la razón por la cual el plan generado usando su *Planificador de Rutas Táctico* (*TRP*) es menos atractivo visualmente que el generado por los planificadores después de realizar algunos ajustes manuales a la solución entregada, y también desarrollar una mejora al algoritmo actual usado por el *TRP*. Para

cuantificar la *belleza* de las soluciones, definen 4 indicadores, y comentan que los planificadores logran encontrar soluciones que mejoran los valores de esas 4 medidas simultáneamente, y al mismo tiempo alcanzan costos tradicionales más bajos. Realizan un extenso estudio comparativo de formas de construir la solución inicial, considerando factibilidad, optimización clásica y *VA*. Proponen dos formas de modificar la fase inicial de construcción de rutas del algoritmo actual que utilizan, una secuencial y una paralela, que toman como base el método de recubrimiento por círculos de Savelsbergh [74]. Las dos alternativas que proponen mejoran la solución actual, pero al compararse entre ellas, notan que la paralela genera soluciones más *clusterizadas*, pero entrega soluciones que utilizan un mayor número de vehículos, y con más infactibilidades.

El 2015 Constantino et al. [13] proponen un algoritmo exacto y una heurística para resolver una versión modificada de un *Mixed Capacitated Arc Routing Problem (MCARP)* en la que se agrega una restricción al modelo que solo permite soluciones con cierto nivel de *VA*. Realizan una revisión de la literatura sobre las metodologías y medidas utilizadas para obtener y evaluar las características asociadas a la sectorización de las soluciones, enfocados en los trabajos que utilizan técnicas de segmentación geográfica aplicadas a ruteo. Para evaluar el efecto de agregar esta nueva restricción en las soluciones, utilizan dos sets de instancias de la literatura del *MCARP*, y miden los resultados en términos del impacto en la función objetivo tradicional, el impacto en valores del *GAP* alcanzado y el tiempo de cómputo, y la calidad obtenida en 3 medidas de *VA*. Las soluciones que encuentran logran mejorar bastante en las medidas de *belleza* que consideran, generando un aumento ligero en el costo tradicional.

Al año siguiente, Rossit et al. [71] publican un trabajo en el que presentan resultados preliminares del desarrollo de un algoritmo cuyo objetivo es mejorar el *atractivo visual* de las soluciones de un *Capacitated VRP (CVRP)*. Primero realizan una revisión bibliográfica sobre las medidas de *belleza* utilizadas en el contexto de ruteo de vehículos, y luego describen el funcionamiento de su heurística. Ésta comienza construyendo una solución factible utilizando un método similar al propuesto por [44], y luego intenta mejorarla con movimientos de búsqueda local que consideran medidas de *VA* dentro de su objetivo y/o como restricción. Hacen pruebas en algunas de las instancias del set de Uchoa et al. [84] que muestran que su heurística puede encontrar buenas soluciones para las 4 medidas de *belleza* que consideran,

aunque comentan que es necesario mejorarla para que logre mejores resultados en términos de la distancia total recorrida y del tiempo total de cómputo.

Lum et al. [51] presentan el 2017 una heurística para el *Min-Max K Windy Rural Postman Problem (MMKWRPP)* que siendo algorítmicamente simple, produce soluciones que en promedio son sólo un 3% peor en términos de las medidas tradicionales, al compararse con las producidas por un método actual [4], y además tiene buen comportamiento con respecto a las 3 medidas de *VA* que consideran. El algoritmo propuesto consta de dos fases: primero construye las rutas utilizando la metodología *cluster-first route-second*, y luego aplica un procedimiento de mejora por *búsqueda local*, modificado para que no permita movimientos que empeoran una de las medidas de *belleza*. Realizan las pruebas de los algoritmos en dos sets de instancias: el primero está basado en 8 grafos construidos con información real de las calles de 8 ciudades del mundo, y el segundo en 10 grafos creados artificialmente simulando rectángulos de varios tamaños. Con cada una de estas 18 redes generan 6 instancias distintas en las que varían en número de vehículos disponibles y la ubicación del *depot*. Además, presentan una revisión bibliográfica sobre las medidas de *belleza* utilizadas en el contexto de ruteo de vehículos. Dos resultados interesantes a los que llegan son que, en términos del costo tradicional, la diferencia entre las dos heurísticas parece crecer linealmente en función del tamaño de la flota, y si el *depot* está ubicado al borde del problema, su heurística da resultados consistentemente peores que cuando éste está en el centro.



# Capítulo 3

## Entrevistas a Clientes

Luego de realizar la revisión de la literatura existente sobre *VA* en *VRPs*, y de entender las necesidades que otros investigadores han detectado, se realizaron entrevistas a 6 de los clientes de SimpliRoute que en algún momento habían declarado no estar 100 % conformes con las soluciones que entregaba el algoritmo, por temas relacionados con la *forma* de las soluciones.

El objetivo de estas entrevistas fue entender las razones por las que las necesidades de algunos clientes no son satisfechas a cabalidad por las soluciones que entrega el algoritmo, evaluar sistemáticamente la importancia que tiene este problema para ellos, y generar hipótesis sobre las variables que los usuarios de la plataforma consideran relevantes a la hora de evaluar la calidad y aplicabilidad de las soluciones. La información recabada de estas entrevistas, complementada con lo encontrado en la literatura, sirven de base para definir el conjunto de medidas de *VA* a estudiar.

En este capítulo se detalla el contexto y la metodología utilizada para realizar las entrevistas, y se entrega un resumen de los resultados obtenidos, junto con algunas conclusiones.

### 3.1. Contexto y metodología

Para elegir al grupo de clientes a entrevistar, se recurrió al conocimiento y la experiencia del equipo de *Customer Success* de SimpliRoute, junto a quienes se definieron los siguientes

criterios a considerar:

- Incluir sólo clientes que alguna vez hubiesen hecho comentarios sobre preferir rutas más *bonitas, compactas o sectorizadas*.
- Elegir empresas en las que el equipo de *Customer Success* creyese que los entrevistados pudiesen dar respuestas útiles (ya sea por tener una mejor relación con ellos, por las disponibilidades de tiempo, o por su juicio sobre las capacidades analíticas de los usuarios).
- Seleccionar empresas que fuesen similares en rubro (todas en el área de la elaboración de productos), y que utilizaran la plataforma para actividades del mismo tipo (distribución de sus productos a tiendas), pero que tuvieran operaciones de distintos tamaños.

Una descripción de las empresas seleccionadas que accedieron a ser entrevistadas se presenta en la tabla 3.1. Para cada una se indica el tipo de productos que elaboran, el tiempo que llevaban siendo clientes de SimpliRoute en el momento en que fueron entrevistadas, el número de vehículos que suelen utilizar cada día, y el número promedio de puntos que entregan diariamente.

| <b>Empresa</b> | <b>Rubro</b>                | <b>Antigüedad</b> | <b>Vehículos</b> | <b>Puntos</b> |
|----------------|-----------------------------|-------------------|------------------|---------------|
| A              | Alimentos de Origen Animal  | 12 meses          | 35               | 750           |
| B              | Alimentos de Origen Animal  | 8 meses           | 20               | 250           |
| C              | Alimentos de Origen Vegetal | 5 meses           | 12               | 50            |
| D              | Confites                    | 4 meses           | 80               | 800           |
| E              | Bebidas Envasadas           | 2 meses           | 7                | 100           |
| F              | Pinturas                    | 6 meses           | 33               | 180           |

Tabla 3.1: Características de las empresas entrevistadas

El desarrollo de las entrevistas tuvo la misma estructura en las 6 empresas. Todas fueron realizadas en conjunto con una o dos personas del equipo de *Customer Success*, que apoyaron activamente el proceso. Por parte de las empresas, siempre estuvieron presentes el o los usuarios directos de la plataforma (planificadores), y en 3 ocasiones también acudieron gerentes de área.

Las entrevistas comenzaron con una explicación de los objetivos de éstas a los participantes (*i.e.*, entender qué razones los motivan a realizar ediciones manuales a los resultados entregados

por la plataforma), y con una descripción de lo que realizaríamos en ese periodo. Se les pidió a los asistentes que cargasen en la plataforma los datos de su operación para ese día, y que antes de comenzar el proceso de edición manual de los resultados, comentasen las características positivas y negativas de la solución entregada por el algoritmo (ruta por ruta, y también para la solución completa). En los casos que se consideró pertinente se les pidió que profundizaran en alguna explicación. Luego, se realizó el mismo proceso con el resultado obtenido después de que editaran las rutas manualmente.

Se grabaron los audios de las 6 entrevistas, los que luego fueron analizados y codificados. Se identificaron todos los comentarios positivos y negativos realizados, junto con su explicación correspondiente, y con la solución propuesta para los problemas (estas últimas dos sólo cuando estaban presentes). Luego se clasificaron estos comentarios por categorías, las que se presentan a continuación junto con algunos ejemplos de las frases dichas en cada una.

## 3.2. Resumen de resultados

El tiempo promedio de edición manual de rutas fue de 26 minutos, con un rango de entre 15 y 55 minutos. A pesar de que los planificadores realizaron las modificaciones comentando en voz alta las razones por las que decidían hacer cada cambio, todos reportaron que el tiempo total que tomó el proceso fue cercano al habitual.

Antes de resumir las razones más comunes por las que los planificadores y gerentes deciden editar manualmente los resultados, se presentan algunas frases que mencionaron, relacionadas con las expectativas generales que tienen, y con la relevancia de considerar este tipo de variables en los *CVRSs*:

- “Una ruta está correcta cuando el vehículo asociado tiene un alto porcentaje de carga, va a un solo sector, y no pasa a dejar puntos de camino, va directo a su sector.”
- “Preferiríamos un resultado zonificado, aunque se pierda algo de optimalidad, por ejemplo, aunque use un vehículo más. Al zonificar, podíamos lograr dar un mejor servicio a nuestros clientes, por lo que el costo adicional de utilizar un camión más se compensa.”

- “Por ahora usamos las zonas fijas, porque la zonificación es un requisito para nosotros, pero no nos cerramos a probar planes en los que el algoritmo entregue resultados zonificados pero con zonas variables, para lograr una mejor eficiencia sin perder esa característica que necesitamos.”
- “Hoy prácticamente estamos ruteando manualmente, para obtener una solución ‘más lógica’ que la que nos da el algoritmo.”
- “Tener que hacer modificaciones manuales a lo que nos entrega el algoritmo nos hace perder muchas horas.”

Un punto importante que se puede extraer de estas frases es que a pesar de que los clientes pagan dinero por un *software* de optimización de rutas, pareciera que realizan una gran cantidad de ediciones manuales al resultado obtenido. Una de las hipótesis que se plantean para explicarlo es que estas ediciones se realizan en parte para lograr soluciones con mejor VA. Esta hipótesis se testea en el capítulo 5.

A continuación, se revisan las categorías que surgen más comúnmente cuando se les pregunta a los usuarios por qué prefieren soluciones diferentes a las entregadas por el algoritmo, y qué características buscan en éstas.

1. Creencias de que pueden mejorar el resultado: Como los algoritmos utilizados por los *CVRs* son heurísticos, existe la posibilidad de que el cliente mejore la solución manualmente. Sin embargo, hay muchas ideas y prácticas que ellos utilizaban al rutear manualmente, que creen que asegurándolas aprovechan de mejor forma los recursos (horario, carga, número de vehículos, etc.), y entonces piensan que son requisitos para que una solución sea buena. Un ejemplo de este tipo de creencias es el siguiente: “*Esa ruta (...) parte en Maipú y después va a San Bernardo, y hay otro vehículo que también va a San Bernardo y otro que también va a Maipú, utilizando 3 vehículos para servir estos dos sectores. Nosotros manualmente mandamos solo un camión a San Bernardo, (...) y otro solo a Maipú, y así nos reducimos 1 vehículo. Dicho esto, me llama mucho la atención que el resultado del algoritmo igual use 20 % menos de camiones de los que usamos nosotros manualmente*”.

Vale la pena mencionar en este punto, que realizando las entrevistas se logró identificar un caso en el que el algoritmo llega a un óptimo local que es fácilmente reconocible y mejorable por algunos usuarios. Esta situación también fue descrita por [6], y ocurre cuando los puntos de entrega están distribuidos en varias ciudades distintas, pero los *depots* están todos en la misma ciudad. En este caso, a veces el resultado generado envía más vehículos de los necesarios a una misma ciudad, lo que además de tener malos índices de *VA*, puede ser ineficiente dada la distancia adicional a recorrer.

2. Restricciones más flexibles de lo que especifican: Los usuarios a veces ingresan a la plataforma restricciones más estrictas de lo que realmente necesitan, para intentar generar soluciones más robustas. Por otro lado, algunos de estos tienen *preferencias* más que *restricciones*, pero la versión actual del algoritmo de SimpliRoute no tiene la opción de considerar *restricciones suaves*. Un ejemplo de esta categoría es el siguiente: “*Nosotros sabemos que algunas ventanas horarias se pueden romper, y en qué condiciones, pero no tenemos cómo decirle eso al sistema*”.
3. Costos asociados a zonas fijas: En algunos casos, las empresas definen estructuras de costos que son difícilmente integrables en los modelos tradicionales de ruteo de vehículos. Casos particulares de estas prácticas son cuando se les paga a los conductores dependiendo de los puntos particulares que visitan, o cuando se subcontrata la flota de vehículos y se definen por contrato estructuras de costos asociados a zonas. Uno de los entrevistados se refiere a este punto así: “*En nuestro modelo de costos, existen tarifas diferentes por ir a distintas zonas, por lo que si hago que un camión vaya a dos zonas innecesariamente, estoy gastando más plata de la necesaria*”.
4. Zonas en las que sólo pueden entrar ciertos vehículos: Ya sea por regulaciones legales o por políticas de la empresa (por ejemplo, preferencias de que algunos conductores y/o camiones particulares vayan a los barrios altos de la ciudad), los entrevistados mencionan que editan las soluciones para que algunos vehículos vayan o no vayan a ciertos sectores. Un ejemplo de estas preferencias es el siguiente: “*Al centro de Santiago no se puede entrar con camiones entre 8 y 10 am. Por otro lado, al sector oriente preferimos no entrar con vehículos pesados*”.
5. Balanceo (carga, kilómetros, puntos, tiempo): Esta categoría ha recibido en los últimos

años cada vez más atención desde la literatura académica (revisar Matl, Hartl & Vidal [57] para un estudio específico sobre el tema), pues mejora la aceptación de las soluciones por parte de los conductores y la calidad del servicio entregado. Los usuarios comentan cosas como: *“El problema es que [el algoritmo] me da, por ejemplo, 1 camión con 2 horas de utilización y que va cargado al 15 %, y otro con 5 horas al 90 %. Si los dos conductores ganan lo mismo, me es difícil darle una ruta corta a uno y otra larga a otro. O una ruta en que tiene que descargar muchísima carga a uno, y a otro muy poca”*.

6. Contingencias imprevistas: En la práctica pueden suceder muchas cosas durante la ejecución de un plan, que hacen que las entregas no se ajusten perfectamente a la solución original. Algunos eventos más regulares podrían modelarse con técnicas de optimización estocástica, pero otros son difícilmente predecibles. Para que la solución entregada por los *CVRs* sea de utilidad para sus usuarios, y en particular para que los ahorros que predicen se reflejen en la realidad, es importante generar planes que sean robustos a estos eventos. Uno de los entrevistados menciona que: *“Cualquier contingencia imprevista que pueda ocurrir (se cortó la luz y no se puede entrar o salir de la bodega, no está la persona que tiene que estar, hay que volver más tarde porque falta una factura, etc.) hacen que, si es que la ruta es muy extensa, (...) mi posibilidad de reacción para los últimos clientes de cada ruta sea casi nula”*.
7. Tiempos de servicio y ventanas de tiempo heterogéneas: Esta categoría habla directamente de una característica que el algoritmo de SimpliRoute no maneja (y que según Drexl [17], hasta el año 2012 pocos *CVRs* lo hacían), la posibilidad de definir distintos tiempos de servicio o ventanas de tiempo a los nodos, dependiendo del horario en el que son visitados o del conductor o vehículo que los atiende. Esta fue una de las razones que uno de los clientes dio cuando se le consultó sobre sus razones para editar manualmente la solución: *“Si va un conductor en particular, sabemos que él logra hacer una ruta con muchos clientes, pero si va otro, no. Entonces no podemos decirle al sistema que la ventana horaria es más grande, porque eso sólo es verdad si va un conductor en particular”*.
8. Tiempos de servicio y ventanas de tiempo desconocidas o poco predecibles: Si bien esta categoría es similar a la anterior, en este caso no se relaciona con las capacidades del

*CVRS*, si no que con la falta de conocimiento de los usuarios, y con la aleatoriedad del proceso. Un ejemplo de este caso es el siguiente: *“Los tiempos de servicio pueden ser muy variables. A veces no está el dueño, y aunque ‘viene llegando’, se demoran 1 hora más de lo previsto. Hay uno [de nuestros clientes] que tiene que hacerse quimioterapia, y si llegas cuando está haciéndosela, no te atiende hasta terminar, que es 3 horas después. Hay clientes que te dejan esperando 1 hora, entre que te entregan la factura, descargan el camión y ven donde van a dejar las cosas”*.

9. Familiaridad de conductor con la zona agiliza la operación: Varios de los entrevistados mencionaron razones que clasifican dentro de esta categoría. Ellos editan las rutas de la solución original para que éstas estén *zonificadas*, y luego asignan conductores a las distintas zonas según la experiencia que tienen ellos en cada sector. Cuando se les pregunta sobre esta práctica dicen cosas como: *“El desconocimiento de las condiciones particulares de un cliente (por ejemplo, no saber exactamente dónde queda, no saber dónde estacionar o si es que no se puede estacionar y el camión tiene que dar vueltas mientras el peoneta entrega, no saber quién y/o cómo y/o cuándo recibe, a veces se entrega en subterráneo o en el 3er piso, a veces la ventana de tiempo es de 10 a 14 pero el cheque llega después de las 12, etc.) (...) hacen que los puntos del final de la ruta queden muy expuestos a cualquier problema que se provoque en la primera mitad de ésta”*.
10. Supermercados: Esta categoría es en realidad es un ejemplo particular de las 3 categorías anteriores, pero dado que surgió repetidamente en todas las entrevistas, se consideró como una categoría en sí misma. Un ejemplo de ésta es el siguiente: *“Dado el poder que tiene el jefe de recepción, los supermercados son puntos de difícil atención. En general se respeta el orden de llegada, pero no siempre, y pueden tener ciertas preferencias hacia ciertos vehículos. Entonces si todos los días cambio el transportista que va a cierto supermercado, el recepcionista, como no conoce el transporte que viene, lo va dejando para el último porque no sabe cómo se va a desempeñar el transportista en el andén. Como los tiempos de andén son tan críticos, si llega un conductor que va todos los días y siempre descarga rápido, el jefe de recepción va a preferir que pase él antes que alguien que no conoce, que puede demorarse mucho y atrasarle todo el proceso”*.

11. Estacionamientos: Este es un claro ejemplo de las diferencias que hay en los problemas que es necesario resolver para la investigación aplicada y la puramente académica. La decisión de cuándo y dónde estacionar influye fuertemente en la factibilidad de seguir la ruta planificada, pero esta variable es difícilmente integrable en los algoritmos de ruteo de vehículos, por la complejidad técnica de modelar el problema, y por la dificultad de tener todos los datos necesarios para que el resultado sea útil. Uno de los entrevistados se refiere a esta categoría diciendo que: *“Por ejemplo, en Recoleta (...), es muy difícil encontrar dónde estacionar. (...) También es difícil estacionar en Maipú. Entonces enviar varios camiones a una de estas zonas, complica todavía más ese proceso”*.
12. Repartos a pie: Hacer repartos a pie (para el conjunto particular de clientes entrevistados) es una técnica utilizada por los transportistas, que aprovechan la detención de un camión para realizar varias entregas cercanas (es decir, no es parte planificada de la operación, pero se hace cuando es posible). Esto puede ocurrir, por ejemplo, cuando el camión se logra detener en un sector de difícil estacionamiento, o si va a tener que estar detenido largo tiempo por una carga o descarga grande. Uno de los entrevistados se refiere a este último ejemplo diciendo que: *“Si el camión va con peoneta, mientras está estacionado descargando un punto grande, por ejemplo un supermercado, el peoneta puede ir caminando a entregar pedidos chicos cercanos. Eso (...) sólo pasa cuando la solución está zonificada”*.
13. No uso de autopistas: Cuando los vehículos y conductores que realizan los repartos no son de la empresa, en general prefieren no transitar por sectores que generen costos adicionales para ellos. Por esto, si las soluciones se generan considerando el uso de éstas, es poco probable que los conductores cumplan con el orden o los tiempos de llegada estimados. Una solución sencilla para este problema consiste en calcular las matrices de tiempos de viaje sin considerar autopistas, pero esa modalidad no estaba disponible en la plataforma de SimpliRoute al momento de realizar las entrevistas. Un ejemplo de esta categoría es el siguiente: *“Como la flota de vehículos es arrendada, los conductores no usan las autopistas para abaratar costos”*.
14. Reclamos de los conductores: Aunque esta categoría involucra a varias de las anteriores, tiene características particulares que hacen que merezca una clasificación independiente.



El éxito de la implementación de un *CVRS* en una organización, depende en última instancia de los transportistas, por lo que si la solución tiene características que estos perciben como negativas, todo el proceso puede verse afectado. En particular, los planificadores mencionaron que los transportistas se quejaban de temas como: rutas que “recorren todo Santiago” o que visitan muchas comunas<sup>1</sup> (en vez de que se mantengan en sólo un sector), rutas que se cruzan con otras, rutas más costosas o que generan menos ingresos para ellos, y rutas desbalanceadas. Uno de los entrevistados menciona que: *“Si yo le digo al transportista que me haga esa ruta, no le va a gustar. Los conductores reclaman, por ejemplo, por un tema de costos, y de que empiezan a las 6 am y terminan a las 6 pm”*.

15. Desconocimiento de sus propias condiciones: Esta categoría es similar a la de *Contingencias imprevistas*, pero en este caso en vez de considerar eventos externos, se refiere a la poca capacidad de la organización para predecir sus propias condiciones de trabajo. Para solucionar esta dificultad, uno de los entrevistados menciona que: *“Como no conocemos bien (...) la hora real en la que van a poder partir nuestros camiones, preferimos utilizar más vehículos con menos despachos cada uno, para que tengan más holgura para atender casos imprevistos”*.

Una conclusión importante que puede extraerse de las categorías recién expuestas es que si bien algunas de las motivaciones que reportan los usuarios al momento de editar no están asociadas con el foco de este trabajo (categorías 2, 7 y 13), la gran mayoría está directa o indirectamente relacionada con la inclusión de objetivos no estándar. Este hecho puede no ser extrapolable a otros clientes, pues uno de los criterios de inclusión para ser entrevistado fue haber mencionado disconformidad con la *VA* de las soluciones, pero sí indica que por lo menos para un grupo de clientes estas variables parecen ser relevantes.

Si nos enfocamos solo en las razones que caen dentro del marco de este trabajo, y analizamos las variables que los entrevistados mencionan que consideran a la hora de editar, y las posibles soluciones que ellos proponen para arreglar los problemas que detectaron, es posible organizarlas de la siguiente forma:

---

<sup>1</sup>División territorial administrativa más pequeña de Chile. Similar a lo que en otros países se llama municipio. [https://es.wikipedia.org/wiki/Comunas\\_de\\_Chile](https://es.wikipedia.org/wiki/Comunas_de_Chile)

- **Zonas Fijas:** Varios de los problemas identificados durante las entrevistas están relacionados con restricciones y/o costos asociados a sectores geográficos predefinidos. Los puntos 3, 4 y 9 de la enumeración anterior son ejemplos de casos que se verían beneficiados con la consideración de este tipo de variables en los modelos y algoritmos de ruteo de vehículos.
- **Zonas Variables:** Para el grueso de las categorías (puntos 1, 6, 8-15) los usuarios mencionan que solucionan los problemas modificando manualmente la solución para generar planes con rutas más sectorizadas (*“que las rutas se crucen lo menos posible entre ellas”*; *“que vaya cada vehículo a una zona distinta”*) y compactas (*“que todos los puntos que visite cada ruta estén cerca entre sí”*; *“que los vehículos no pasen a dejar pedidos ‘de camino’ al sector que tienen asignados”*). Es decir, les importa que las soluciones estén zonificadas, pero no quieren tener que definir previamente estas zonas ellos mismos.
- **Balanceo:** El punto 5 de la enumeración anterior es de naturaleza distinta a las dos clasificaciones anteriores. Los entrevistados mencionaron diferentes criterios a la hora de considerar elementos de balanceo. Algunos editaban con el objetivo de asegurar un mínimo de carga en cada camión, otros para equilibrar el número de puntos en cada ruta, y otros considerando la cantidad de kilómetros recorridos o el tiempo total de viaje.

Como las entrevistas se realizaron con un foco amplio sobre la edición manual de rutas, surgieron varios puntos que, aunque relevantes en la práctica, no están directamente relacionados con el *atractivo visual*, sino más bien con características que la plataforma de SimpliRoute no maneja, como *zonas fijas* y *balanceo*. Sin embargo, las conclusiones asociadas al grupo de *zonas variables* muestran resultados similares a los encontrados en la literatura, en términos de qué variables considerar (compacidad y sectorización) y también en la relevancia de éstas. Basándose en estos resultados, el siguiente capítulo aborda un análisis en profundidad de las distintas medidas propuestas en la literatura para evaluación de la VA en las soluciones del VRP.

# Capítulo 4

## Medidas de evaluación de Atractivo Visual

Este capítulo estudia en detalle las propiedades de las medidas de *VA*. Primero se describen una serie de características que clasifican las medidas encontradas en la literatura, y que permiten entender de mejor forma cómo se relacionan entre ellas. Luego se introducen dos medidas nuevas, con el objetivo de definir un indicador que represente de manera más completa la noción de *atractivo visual*, y finalmente se presentan los resultados de un estudio realizado para comprender en mayor profundidad la relación entre algunas medidas de *VA* y los objetivos tradicionales del *VRP*.

### 4.1. Clasificación de medidas usadas en la literatura

Como se mencionó en la sección 2.3, en los trabajos que abordan la *VA* en el *VRP* hay poco acuerdo sobre cómo definir las variables a utilizar. Sin embargo, aunque en cada uno de éstos se suelen considerar medidas diferentes, la mayoría comparte varias características comunes. Entender claramente cuáles son estos atributos y cómo varían entre los distintos enfoques es útil para guiar y ordenar la investigación en esta área, pues permite definir más claramente las medidas, evitar posibles ambigüedades, y en consecuencia, facilitar la comparación entre los resultados de distintos trabajos.

La característica implícita que se ha usado hasta el momento para categorizar las medidas

se centra en *qué* pretenden medir las variables en consideración, es decir, si intentan evaluar compacidad, sectorización u otra cualidad. Si bien esta clasificación es útil, donde hay mayor estructura es en el *cómo* se van a intentar cuantificar tales conceptos. En esta última línea, la mayoría de las medidas propuestas por los trabajos revisados pueden ser categorizadas basándose en las siguientes características:

- **Intra-ruta o Inter-ruta:** Hace referencia a si la medida cuantifica características relacionadas con cada ruta en sí misma o si evalúa cómo se relacionan las distintas rutas entre ellas. En general, suele ocurrir que las medidas Intra-ruta evalúan compacidad (*i.e.*, la densidad y distribución de los puntos de cada ruta) y las Inter-ruta evalúan sectorización (es decir, el nivel de intersección o cruce que hay entre las rutas).

Ejemplos de medidas Intra-ruta e Inter-ruta propuestas por Poot et al. [64] son, respectivamente, calcular para cada ruta la distancia promedio entre todos los pares de clientes visitados por ésta, y contar el número total de cruces entre arcos de diferentes rutas para la solución completa.

- **Cuantificación o Conteo:** Considera si la propiedad que se quiere evaluar se mide contando cuántos elementos cumplen cierto requerimiento, o si ésta se cuantifica directamente.

Las dos medidas mencionadas en el punto anterior también son buenos ejemplos para entender esta característica. La primera cuantifica el valor a evaluar, pues considera la distancia promedio entre los puntos, mientras que en el segundo caso simplemente se cuenta el número de veces que ocurre un fenómeno. Una versión similar a esta última medida, pero que es intra-ruta y cuantificada, es la llamada *Crossing Length Percentage (CLP)* propuesta por Lu y Dessouky [50]. El *CLP* de cada ruta se calcula asignando un valor (llamado *largo cruzado*) a cada punto en el que dos arcos de una misma ruta se intersectan, y luego el puntaje final se calcula sumando estos valores y dividiendo el resultado por el largo total de tal ruta.

- **Elemento geográfico de referencia:** Se refiere al elemento geográfico con respecto al cuál se realizan los cálculos; éstos pueden ser, entre otros, puntos, segmentos o áreas. Algunas de las elecciones más comunes en la literatura son definir un “centro” para

cada ruta, o considerar un área asociada a éstas.

Poot et al. [64] proponen dos medidas inter-ruta de conteo con elementos geográficos de referencia distintos. En la primera, se define el centro de cada ruta como el centro de gravedad de los puntos (en el plano) asociados a las visitas de tal ruta (es decir, el punto correspondiente al promedio por coordenadas); el indicador se calcula contando cuántos de los clientes de una ruta están más cerca del centro de otra ruta que del centro de la ruta a la que pertenecen. La segunda define como área asociada a cada ruta, a la envoltura convexa de los puntos pertenecientes a cada una; la medida se calcula contando la cantidad de clientes de una ruta, que están contenidos en el área asociada a alguna otra ruta.

Vale la pena mencionar que otros autores han propuesto diferentes definiciones de centro y área, para calcular las mismas u otras medidas. Por ejemplo, para los casos en los que se realizan los cálculos usando una matriz de distancias o tiempos de viaje en calle, puede ser preferible utilizar como centro de las rutas a alguno de los puntos visitados por la misma (en vez de un punto cualquiera en el plano). En este contexto, Tang y Miller-Hooks [81] sugieren usar el punto visitado por la ruta para el cual la distancia promedio entre éste y todos los otros clientes de la ruta es mínima. También proponen usar el punto visitado por la ruta para el cual la máxima distancia entre éste y todos los otros clientes de la ruta es mínima, pero comentan que prefieren la versión que utiliza el promedio, pues cuando se usa el máximo no se tiene en cuenta la distribución general de los puntos de la ruta en el espacio. Por otro lado, Kim et al. [44] proponen considerar el punto de la ruta que está más cerca del centro de gravedad de los puntos, y Kant et al. [41] usar el punto de la ruta que está en la posición  $n/2$  (para rutas con  $n$  puntos). Con respecto a las áreas, Matis y Kohani [56] proponen una medida que considera dos áreas distintas: la envoltura no-convexa más pequeña que contiene a los clientes de una ruta, y un círculo con el mismo perímetro que tiene tal envoltura no-convexa. Lum et al. [51] mencionan dos medidas usadas en la literatura de distritificación política, que consideran el área del círculo más pequeño que incluye totalmente al sector en consideración, y el perímetro del círculo con área igual a la que se está evaluando. En otra línea, Bosch [6] considera áreas fijas predefinidas (asociadas a distintas ciudades), y asigna estas áreas a las distintas rutas si es que la ruta tiene alguna parada dentro de

tales áreas.

- **Modo de Agregación:** Define la forma en la que la medida llega a un único valor final para la solución, a partir de varios valores que califican elementos parciales de ésta, por ejemplo, rutas o puntos en el mapa. Dado que estos elementos pueden estar anidados (en el caso del ejemplo, una ruta está formada por varios puntos, y una solución está compuesta por varias rutas), es posible que esta agregación sea por niveles (es decir, que se elija un Modo de Agregación para definir un valor para cada ruta a partir de los valores por punto, y luego otro para llegar al valor final por solución) o directa (es decir, no pasar por valores intermedios por ruta). También es posible que la medida esté definida inicialmente para la solución completa, y que por tanto no sea necesario definir un Modo de Agregación para ésta. Ejemplos comunes de Modos de Agregación son la suma, el promedio o el máximo de los valores.

Gretton y Kilby [24] hacen referencia a esta característica, notando que si bien utilizan una medida similar a una de las usadas por Tang y Miller-hooks en [81], ellos prefieren no agregar por ruta, si no que hacerlo directamente a nivel de solución, pues esa agregación intermedia podría ser problemática si los vehículos tienen diferentes capacidades, y también en las iteraciones iniciales de la búsqueda de soluciones, cuando las rutas que conforman la solución parcial están compuestas por un número significativamente distinto de visitas.

- **Distancia Utilizada:** Indica qué noción de distancia se usa para realizar los cálculos. Las opciones más comunes son la distancia euclidiana y la distancia “en calle” (correspondiente al largo del camino mínimo entre dos puntos a través de la red de caminos disponibles). Es importante mencionar que una medida puede no necesitar definir una noción de distancia, o puede incluir más de una definición de distancia distintas. Por ejemplo, en [43] se usa la distancia euclidiana para calcular centroides de las rutas, luego definen el centro de cada una como la visita más cercana al centroide (usando nuevamente la distancia euclidiana), y finalmente se asocia a cada punto la distancia “en calle” desde éste hacia el centro de su ruta, y se agregan estos valores sumándolos. Si bien algunos autores ([81], [44], [24] y [71]) han argumentado que es preferible usar la distancia “en calle”, pues permite aprovechar que la matriz de distancias entre puntos

está previamente calculada, Battarra et al. [2] comentan que prefieren utilizar la distancia euclidiana, porque la distancia “en calle” puede distorsionar fuertemente la forma de las rutas debido a restricciones de viraje y de calles de un sólo sentido.

- **Consideración del *Depot*:** Se refiere a si es que el *depot* es considerado o no como punto de la ruta para calcular la medida. Constantino et al. [13] mencionan la necesidad de estudiar las ventajas y desventajas que tiene esta característica. Además, notan que ninguna de las medidas presentadas en los trabajos que ellos citan (entre otros [64], [81], [50], [44], [55]) lo considera.

Un ejemplo de un indicador que sí lo hace es el propuesto por Lum et al. [51] llamado *Hull Overlap (HO)*. Esta medida considera la envoltura convexa de los puntos de cada ruta, y luego calcula, para cada ruta  $r_i$ , la fracción de su área que se intersecta con cada una de las envolturas convexas de las otras rutas, y suma todos esos valores. Es decir, para cada ruta calcula:

$$HO_{r_i} = \sum_{r \in R} \frac{\text{intersec}(\text{convex}(r), \text{convex}(r_i))}{\text{convex}(r_i)} \quad (4.1)$$

Finalmente, *HO* se calcula tomando el promedio de los  $HO_{r_i}$  entre todas las rutas. Así, ésta es una medida inter-ruta, de cuantificación, con la envoltura convexa de los puntos de la ruta como elemento geográfico de referencia, que usa el promedio como Modo de Agregación, que no utiliza nociones de distancia, y que incluye el *depot* como punto de la ruta para los cálculos.

## 4.2. Nuevas medidas propuestas

Las dos medidas propuestas intentan capturar, en un solo indicador, dos objetivos comúnmente encontrados en las medidas existentes en la literatura: que los puntos de cada ruta estén cerca de los centros de cada una de éstas, y que los puntos de cada ruta estén lejos de los centros de las otras rutas (así, se intenta evaluar simultáneamente compacidad y sectorización). Si llamamos  $R$  al conjunto de rutas de la solución, para cada  $r \in R$  se define:

$$RatioCenter1_r = \frac{1}{|r|} \sum_{p \in r} \frac{d(p, c_r)}{d(p, c^1(p))} \quad (4.2)$$

$$RatioCenter2_r = \frac{1}{|r|} \sum_{p \in r} \frac{d(p, c_r)}{d(p, c^1(p)) + d(p, c^2(p))} \quad (4.3)$$

Donde  $d(x, y)$  es alguna distancia entre los puntos  $x$  e  $y$ ,  $c_r$  es el centro de la ruta  $r$  y  $c^i(p)$  es el  $i$ -ésimo centro más cercano al punto  $p$ . Finalmente, las dos medidas para la solución completa se obtienen tomando el promedio por ruta de cada uno de estos valores.

Queda claro de la definición, que estas medidas son inter-ruta, de cuantificación, con “centros de ruta” como elemento geográfico de referencia, y con agregación en dos niveles (puntos a rutas, y rutas a solución) utilizando promedio en ambos casos. Con respecto a las otras características, para los resultados presentados más adelante, se consideró como centro de las rutas al centroide de la envoltura convexa de sus puntos, se utilizó la distancia euclidiana, y no se incluyó el *depot* como punto de las rutas.

Si bien la definición de la variable asociada a  $RatioCenter1_r$  es una consecuencia lógica del objetivo que se desea medir (que las rutas estén cerca de su propio centro y lejos de los otros), los experimentos iniciales mostraron ciertos problemas, relacionados por un lado con la falta de conciencia sobre la geometría de la instancia, y por otro lado, asociados a las rutas de sólo 1 punto. En este último caso, el centro más cercano al punto siempre corresponde al centro de su propia ruta, pero éste está exactamente en el mismo lugar que el punto, por lo que al calcular la medida el denominador se hace cero, lo que indetermina el valor de la variable. Por esta razón, se decide agregar la versión a  $RatioCenter2_r$ , que corrige parcialmente estos dos problemas.



## 4.3. Relación entre las medidas en consideración

### 4.3.1. Metodología

Para estudiar cómo se relacionan las distintas variables en consideración, en soluciones de las instancias de Solomon [78], se utiliza el algoritmo de Simulated Annealing de JSPIRT<sup>1</sup> [75] para generar la mayor cantidad posible de soluciones para cada una de estas instancias, y luego se evalúa cada una de éstas para las medidas tradicionales y las de *VA*.

Se usan las configuraciones por defecto de JSPRIT, pero se modifica la función objetivo para privilegiar la minimización de vehículos utilizados. También se agregan restricciones suaves relacionadas con las distintas medidas de *belleza*, y se ejecuta el algoritmo con diversas combinaciones de pesos (positivos y negativos) asociados a éstas, para favorecer la exploración del espacio de soluciones.

Todas las medidas de *VA* se calculan a nivel de ruta, y luego se utilizan 3 Modos de Agregación distintos: *suma*, *máximo* y *promedio*<sup>2</sup>. Las medidas consideradas son:

- **DT (DrivingTime)**: Es la función objetivo tradicional del ruteo de vehículos. Corresponde a la suma de los tiempos de viaje totales de cada ruta realizada.
- **PointsHull**: Para cada ruta, se cuenta el número total de puntos que están incluidos dentro de la envoltura convexa de ésta.
- **ForeignPointsHull**: Para cada ruta, se cuenta el número de puntos **de otras rutas** que están incluidos dentro de la envoltura convexa de ésta.
- **Cross**: Entre cada punto consecutivo de una ruta se traza la recta que los une, y se cuenta el número de veces que dos de estos segmentos se intersectan.
- **DistBetween**: Se define como el promedio de las distancias euclidianas entre los pares de clientes consecutivos de una ruta.

---

<sup>1</sup>Librería de código abierto en Java desarrollada y mantenida por Stefan Schröder, que permite resolver problemas de ruteo de vehículos.

<sup>2</sup>A pesar de que agrupar por suma y promedio en soluciones de la misma instancia y con el mismo número de vehículos utilizados es simplemente un cambio de escala, se decide considerar ambos modos pues esto no es así al evaluar conjuntamente soluciones con distintos números de vehículos usados.

- **DistCenter**: Se define como el promedio de las distancias euclidianas entre los clientes de una ruta y su propio centro. Se considera como centro al centroide de la envoltura convexa de los puntos de la ruta.
- **RatioCenterV1**: Definida en (4.2).
- **RatioCenterV2**: Definida en (4.3).

Si bien **Cross** no mide ni compacidad ni sectorización, se incluyó de igual forma en el análisis, por haber sido estudiada previamente por MacGregor & Ormerod en [53].

Cabe mencionar que no es posible incorporar estas medidas al modelo formulado en la sección 2.1 sin incluir información adicional. En particular, para la mayoría de las medidas de *VA* es necesario tener información sobre la ubicación en la que se encuentra cada uno de los puntos a visitar.

### 4.3.2. Resultados

Mediante este procedimiento se logra generar 9.466.375 soluciones distintas para las 56 instancias, sin embargo, para que la comparación entre soluciones de una misma instancia sea consistente, sólo se analizan las que alcanzan el mínimo número de vehículos conocido (según los valores reportados en [77]). Una vez eliminadas las que utilizaban vehículos adicionales, el total resultante fue de 1.234.362. La distribución de la cantidad de soluciones por tipo de instancia se presenta en la tabla 4.1. En ésta se aprecia que para las instancias de tipo *C*, el número de soluciones generadas fue bastante menor que para las de tipo *R* y *RC*, pero al mismo tiempo, estas últimas son las que tenían más soluciones con vehículos adicionales, por lo que al comparar después de la eliminación, en algunos casos quedan con cantidades menores que las de tipo *C*. Más aun, hubo 10 instancias de *R* y 7 instancias de *RC* para las que se descarta el total de las soluciones encontradas, por lo que tales instancias no se incluyen en el análisis posterior.

Para poder comparar resultados entre distintas instancias, es necesario utilizar medidas que no estén relacionadas con variables en las que las soluciones de éstas son distintas, como

| Grupo | Instancias pre | Instancias post | Soluciones pre | Soluciones post |
|-------|----------------|-----------------|----------------|-----------------|
| C1    | 9              | 9               | 617.430        | 251.385         |
| C2    | 8              | 8               | 469.425        | 254.265         |
| R1    | 12             | 7               | 2.554.892      | 30.077          |
| R2    | 11             | 6               | 2.371.790      | 412.577         |
| RC1   | 8              | 4               | 1.754.882      | 34.303          |
| RC2   | 8              | 5               | 1.697.956      | 251.755         |
| TODAS | 56             | 39              | 9.466.375      | 1.234.362       |

Tabla 4.1: Resumen de la cantidad de instancias y soluciones generadas, antes y después de eliminar las soluciones que no alcanzaron el mínimo número de vehículos conocido

por ejemplo el número de puntos de cada ruta, o el número de vehículos utilizados. Por esto, en vez de considerar directamente la medida **DT**, se usa **PercentageDT** que es el porcentaje de desvío del mínimo valor conocido de distancia total (según las Mejores Soluciones Conocidas (MSC) reportadas en [77]). La tabla 4.2 muestra un resumen de los resultados alcanzados para esta medida en las soluciones de los distintos grupos de instancias.

| Grupo | Promedio | Desv. Est. | Mínimo | Mediana | Máximo |
|-------|----------|------------|--------|---------|--------|
| C1    | 16,0 %   | 20,2 %     | 0 %    | 8,7 %   | 198 %  |
| C2    | 12,7 %   | 10,3 %     | 0 %    | 9,9 %   | 93 %   |
| R1    | 3,4 %    | 2,4 %      | 0,03 % | 2,8 %   | 22 %   |
| R2    | 13,6 %   | 9,2 %      | 0,07 % | 11,1 %  | 90 %   |
| RC1   | 5,8 %    | 3,2 %      | 0,01 % | 5,2 %   | 24 %   |
| RC2   | 19,2 %   | 14,4 %     | 0,08 % | 14,9 %  | 128 %  |
| TODAS | 14,6 %   | 13,7 %     | 0 %    | 10,7 %  | 198 %  |

Tabla 4.2: Resumen descriptivo del **PercentageDT** en las soluciones generadas, para cada grupo de instancias

Con respecto a las medidas de **VA**, no es claro en qué casos es deseable hacer esta normalización, porque no se conocen las formas en las que éstas se relacionan con las distintas características de las instancias ni de sus respectivas soluciones. Mas aún, incluso si se deseara hacerlo, no sería trivial pues no se conocen sus mínimos (incluso para algunas como **ForeignPointsHull** o **Cross** que pareciera que tienen un mínimo fácilmente distinguible, éste podría no ser 0 para las soluciones con el mínimo de vehículos, por restricciones de ventanas horarias u otras). Por estas razones, los análisis posteriores se realizan con los valores originales para estas variables.

## Relación entre las variables de Atractivo Visual

A modo de evaluación general inicial, se calculó la matriz de correlaciones (considerando las 1.234.362 soluciones) para las 21 variables asociadas al *atractivo visual* (3 por cada una de las 7 definidas previamente, por los distintos Modos de Agregación) más 3 relacionadas con el objetivo tradicional: el **DT**, la diferencia absoluta de éste con el valor de la MSC (**ZeroStartDT**), y el **PercentageDT**. El resultado se presenta en la tabla 4.3.

|                      |   | B    | C    | D    | E    | F    | G    | H    | I    | J     | K    | L     | M     | N    | O    | P     | Q     | R     | S     | T     | U     | V    | W     | X    |
|----------------------|---|------|------|------|------|------|------|------|------|-------|------|-------|-------|------|------|-------|-------|-------|-------|-------|-------|------|-------|------|
| SumPointsHull        | A | 0,74 | 0,66 | 1,00 | 0,95 | 0,95 | 0,08 | 0,07 | 0,08 | 0,35  | 0,50 | 0,49  | 0,17  | 0,56 | 0,61 | 0,12  | 0,63  | 0,82  | -0,04 | 0,09  | 0,52  | 0,51 | 0,46  | 0,37 |
| MaxPointsHull        | B |      | 0,96 | 0,74 | 0,82 | 0,84 | 0,19 | 0,24 | 0,27 | -0,12 | 0,74 | 0,85  | -0,34 | 0,36 | 0,67 | -0,39 | 0,23  | 0,75  | -0,55 | -0,07 | 0,22  | 0,15 | 0,27  | 0,24 |
| AvgPointsHull        | C |      |      | 0,66 | 0,72 | 0,79 | 0,17 | 0,23 | 0,27 | -0,22 | 0,73 | 0,89  | -0,45 | 0,27 | 0,63 | -0,49 | 0,11  | 0,70  | -0,64 | -0,11 | 0,14  | 0,04 | 0,21  | 0,20 |
| SumForeignPointsHull | D |      |      |      | 0,95 | 0,95 | 0,08 | 0,07 | 0,08 | 0,35  | 0,50 | 0,49  | 0,17  | 0,56 | 0,61 | 0,12  | 0,63  | 0,82  | -0,04 | 0,09  | 0,52  | 0,51 | 0,46  | 0,37 |
| MaxForeignPointsHull | E |      |      |      |      | 0,95 | 0,09 | 0,09 | 0,10 | 0,22  | 0,56 | 0,54  | 0,04  | 0,50 | 0,59 | 0,00  | 0,55  | 0,79  | -0,15 | 0,05  | 0,45  | 0,43 | 0,39  | 0,32 |
| AvgForeignPointsHull | F |      |      |      |      |      | 0,07 | 0,09 | 0,10 | 0,14  | 0,53 | 0,58  | -0,04 | 0,43 | 0,59 | -0,07 | 0,46  | 0,78  | -0,21 | 0,03  | 0,42  | 0,38 | 0,34  | 0,27 |
| SumCross             | G |      |      |      |      |      |      | 0,94 | 0,97 | -0,14 | 0,00 | 0,11  | 0,11  | 0,59 | 0,71 | -0,13 | 0,00  | 0,18  | -0,20 | -0,03 | 0,01  | 0,58 | 0,23  | 0,10 |
| MaxCross             | H |      |      |      |      |      |      |      | 0,96 | -0,18 | 0,05 | 0,17  | 0,02  | 0,55 | 0,69 | -0,20 | -0,03 | 0,18  | -0,26 | -0,05 | -0,02 | 0,50 | 0,26  | 0,15 |
| AvgCross             | I |      |      |      |      |      |      |      |      | -0,23 | 0,06 | 0,20  | -0,01 | 0,55 | 0,72 | -0,24 | -0,05 | 0,19  | -0,31 | -0,06 | -0,04 | 0,50 | 0,23  | 0,12 |
| SumDistBetween       | J |      |      |      |      |      |      |      |      |       | 0,19 | -0,03 | 0,85  | 0,46 | 0,19 | 0,87  | 0,81  | 0,51  | 0,66  | 0,28  | 0,53  | 0,60 | 0,38  | 0,33 |
| MaxDistBetween       | K |      |      |      |      |      |      |      |      |       |      | 0,89  | -0,24 | 0,34 | 0,51 | -0,24 | 0,37  | 0,71  | -0,43 | -0,02 | 0,22  | 0,05 | 0,36  | 0,37 |
| AvgDistBetween       | L |      |      |      |      |      |      |      |      |       |      |       | -0,39 | 0,27 | 0,60 | -0,42 | 0,15  | 0,71  | -0,61 | -0,09 | 0,13  | 0,01 | 0,27  | 0,28 |
| SumDistCenter        | M |      |      |      |      |      |      |      |      |       |      |       |       | 0,48 | 0,17 | 0,95  | 0,66  | 0,26  | 0,81  | 0,30  | 0,45  | 0,74 | 0,25  | 0,16 |
| MaxDistCenter        | N |      |      |      |      |      |      |      |      |       |      |       |       |      | 0,82 | 0,27  | 0,62  | 0,63  | 0,08  | 0,12  | 0,40  | 0,83 | 0,63  | 0,51 |
| AvgDistCenter        | O |      |      |      |      |      |      |      |      |       |      |       |       |      |      | -0,04 | 0,40  | 0,76  | -0,26 | 0,03  | 0,32  | 0,72 | 0,49  | 0,36 |
| SumRatioCenterV2     | P |      |      |      |      |      |      |      |      |       |      |       |       |      |      | 0,66  | 0,22  | 0,87  | 0,32  | 0,45  | 0,55  | 0,14 | 0,09  |      |
| MaxRatioCenterV2     | Q |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       | 0,71  | 0,46  | 0,30  | 0,64  | 0,63  | 0,46 | 0,39  |      |
| AvgRatioCenterV2     | R |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       | -0,07 | 0,13  | 0,55  | 0,54  | 0,43 | 0,35  |      |
| SumRatioCenterV1     | S |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       |       | 0,62  | 0,54  | 0,38  | 0,02 | -0,01 |      |
| MaxRatioCenterV1     | T |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       |       |       | 0,74  | 0,19  | 0,06 | 0,04  |      |
| AvgRatioCenterV1     | U |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       |       |       |       | 0,47  | 0,25 | 0,19  |      |
| DT                   | V |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       |       |       |       |       | 0,51 | 0,35  |      |
| ZeroStartDT          | W |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       |       |       |       |       |      |       | 0,97 |
| PercentageDT         | X |      |      |      |      |      |      |      |      |       |      |       |       |      |      |       |       |       |       |       |       |      |       |      |

En verde los valores mayores a 0,7 y en rojo los menores a -0,3.

Tabla 4.3: Correlaciones entre las variables de VA y las tradicionales.

El comportamiento general es el esperado, obteniendo altas correlaciones entre las medidas asociadas a las mismas variables pero con distintos Modos de Agregación (siendo la excepción más notoria la entre **SumDistBetween** y **AvgDistBetween** que es casi nula  $(-0,03)$ ), entre las variables asociadas a **PointsHull** y a **ForeignPointsHull**, y entre **SumDistCenter**, **SumRatioCenterV2** y **SumRatioCenterV1**.

Un resultado no esperado son las correlaciones negativas alcanzadas al considerar alguna de las siguientes variables **SumDistCenter**, **SumRatioCenterV2**, **SumRatioCenterV1**, junto con alguna de estas otras **MaxPointsHull**, **AvgPointsHull**, **AvgDistBetween**. Esto pareciera indicar que las soluciones con rutas más compactas tienden a intersectarse más y a tener arcos más largos. Sin embargo, un análisis más detallado de este caso muestra que el resultado es simplemente un caso de la paradoja de Simpson [5]. Esto puede verse en la figura 4.1 que muestra un gráfico de dispersión de las variables **SumDistCenter** y **MaxPointsHull** por instancia, donde se aprecia que aunque la correlación en el conjunto

total de datos es negativa, al evaluarla por instancia ese resultado se invierte o anula.

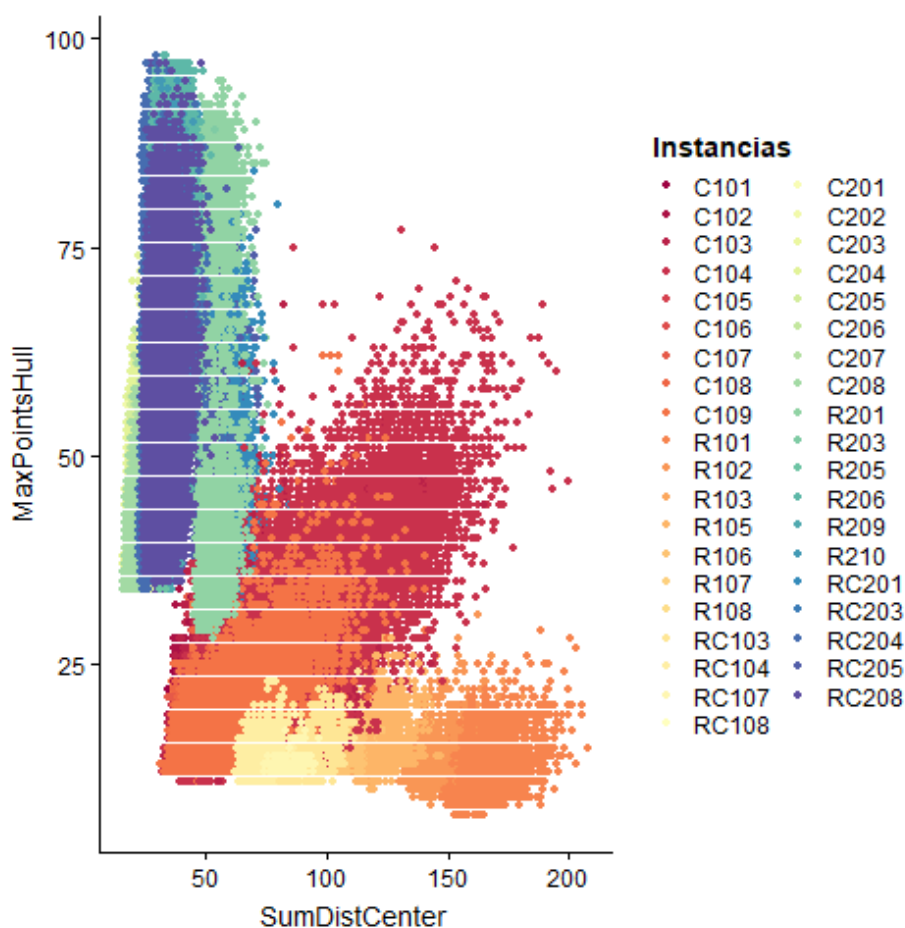


Figura 4.1: **SumDistCenter** vs **MaxPointsHull** por instancia

Ahora, este mismo fenómeno podría estar ocurriendo entre otros pares de variables, en mayor o menor grado. Incluso podría suceder en sentido contrario, es decir, que la correlación total parezca positiva o nula, pero que en realidad a nivel de instancias sea negativa. Para evaluar esta posibilidad, se calcula la matriz de correlaciones para cada una de las instancias, y luego se promedian estas 39 matrices<sup>3</sup>. Los resultados se presentan en la tabla 4.4.

Esta matriz contiene bastante información redundante, pues muchas de las variables incluidas en el análisis (al evaluarse para una misma instancia) son una transformación afín de alguna otra, y la correlación es invariante frente a cambios de escala y a traslaciones. Sin embargo, se incluyen igualmente, pues esto facilita la comparación con la tabla 4.3. En

<sup>3</sup>Es posible que el promedio de las 39 correlaciones asociadas a algún par de variables no sea un buen representante de cada una éstas, por lo que también se estudió la distribución de estos valores para todos los pares de variables. Las conclusiones a las que se llega mediante tal análisis son iguales a las que se obtienen con el estudio de los promedios, por lo que sólo se reporta este último.

|                      |   | B    | C    | D    | E    | F    | G     | H     | I     | J     | K     | L     | M    | N    | O    | P     | Q     | R     | S     | T     | U     | V    | W    | X    |
|----------------------|---|------|------|------|------|------|-------|-------|-------|-------|-------|-------|------|------|------|-------|-------|-------|-------|-------|-------|------|------|------|
| SumPointsHull        | A | 0,81 | 1,00 | 1,00 | 0,84 | 1,00 | -0,17 | -0,14 | -0,17 | 0,84  | 0,66  | 0,84  | 0,58 | 0,43 | 0,58 | 0,87  | 0,77  | 0,87  | 0,53  | 0,38  | 0,53  | 0,55 | 0,55 | 0,55 |
| MaxPointsHull        | B |      | 0,81 |      |      |      | -0,12 | -0,08 | -0,12 | 0,66  | 0,60  | 0,66  | 0,42 | 0,35 | 0,42 | 0,68  | 0,63  | 0,68  | 0,39  | 0,29  | 0,39  | 0,41 | 0,41 | 0,41 |
| AvgPointsHull        | C |      |      | 1,00 | 0,84 | 1,00 | -0,17 | -0,14 | -0,17 | 0,84  | 0,66  | 0,84  | 0,58 | 0,43 | 0,58 | 0,87  | 0,77  | 0,87  | 0,53  | 0,38  | 0,53  | 0,55 | 0,55 | 0,55 |
| SumForeignPointsHull | D |      |      |      | 0,84 | 1,00 | -0,17 | -0,14 | -0,17 | 0,84  | 0,66  | 0,84  | 0,58 | 0,43 | 0,58 | 0,87  | 0,77  | 0,87  | 0,53  | 0,38  | 0,53  | 0,55 | 0,55 | 0,55 |
| MaxForeignPointsHull | E |      |      |      |      | 0,84 | -0,12 | -0,08 | -0,12 | 0,70  | 0,65  | 0,70  | 0,46 | 0,42 | 0,46 | 0,71  | 0,67  | 0,71  | 0,42  | 0,31  | 0,42  | 0,44 | 0,44 | 0,44 |
| AvgForeignPointsHull | F |      |      |      |      |      | -0,17 | -0,14 | -0,17 | 0,84  | 0,66  | 0,84  | 0,58 | 0,43 | 0,58 | 0,87  | 0,77  | 0,87  | 0,53  | 0,38  | 0,53  | 0,55 | 0,55 | 0,55 |
| SumCross             | G |      |      |      |      |      | 0,84  | 1,00  |       | -0,20 | -0,18 | -0,20 | 0,28 | 0,22 | 0,28 | -0,19 | -0,18 | -0,19 | -0,14 | -0,11 | -0,14 | 0,30 | 0,30 | 0,30 |
| MaxCross             | H |      |      |      |      |      |       |       | 0,84  | -0,15 | -0,13 | -0,15 | 0,22 | 0,22 | 0,22 | -0,15 | -0,15 | -0,15 | -0,13 | -0,10 | -0,13 | 0,25 | 0,25 | 0,25 |
| AvgCross             | I |      |      |      |      |      |       |       |       | -0,20 | -0,18 | -0,20 | 0,28 | 0,22 | 0,28 | -0,19 | -0,18 | -0,19 | -0,14 | -0,11 | -0,14 | 0,30 | 0,30 | 0,30 |
| SumDistBetween       | J |      |      |      |      |      |       |       |       |       | 0,78  | 1,00  | 0,58 | 0,40 | 0,58 | 0,97  | 0,85  | 0,97  | 0,57  | 0,42  | 0,57  | 0,54 | 0,54 | 0,54 |
| MaxDistBetween       | K |      |      |      |      |      |       |       |       |       |       |       | 0,37 | 0,37 | 0,37 | 0,76  | 0,80  | 0,76  | 0,51  | 0,39  | 0,51  | 0,36 | 0,36 | 0,36 |
| AvgDistBetween       | L |      |      |      |      |      |       |       |       |       |       |       | 0,58 | 0,40 | 0,58 | 0,97  | 0,85  | 0,97  | 0,57  | 0,42  | 0,57  | 0,54 | 0,54 | 0,54 |
| SumDistCenter        | M |      |      |      |      |      |       |       |       |       |       |       | 0,75 | 1,00 |      | 0,57  | 0,48  | 0,57  | 0,30  | 0,19  | 0,30  | 0,91 | 0,91 | 0,91 |
| MaxDistCenter        | N |      |      |      |      |      |       |       |       |       |       |       |      |      | 0,75 | 0,40  | 0,40  | 0,40  | 0,23  | 0,16  | 0,23  | 0,69 | 0,69 | 0,69 |
| AvgDistCenter        | O |      |      |      |      |      |       |       |       |       |       |       |      |      |      | 0,57  | 0,48  | 0,57  | 0,30  | 0,19  | 0,30  | 0,91 | 0,91 | 0,91 |
| SumRatioCenterV2     | P |      |      |      |      |      |       |       |       |       |       |       |      |      |      | 0,88  | 1,00  |       | 0,63  | 0,45  | 0,63  | 0,54 | 0,54 | 0,54 |
| MaxRatioCenterV2     | Q |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       | 0,88  | 0,63  | 0,45  | 0,61  | 0,47 | 0,47 | 0,47 |
| AvgRatioCenterV2     | R |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       | 0,63  | 0,45  | 0,63  | 0,54 | 0,54 | 0,54 |
| SumRatioCenterV1     | S |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       | 0,94  | 1,00  |       | 0,28 | 0,28 | 0,28 |
| MaxRatioCenterV1     | T |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       |       | 0,94  |       | 0,17 | 0,17 | 0,17 |
| AvgRatioCenterV1     | U |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       |       |       |       | 0,28 | 0,28 | 0,28 |
| DT                   | V |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       |       |       |       | 1,00 | 1,00 | 1,00 |
| ZeroStartDT          | W |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       |       |       |       |      | 1,00 | 1,00 |
| PercentageDT         | X |      |      |      |      |      |       |       |       |       |       |       |      |      |      |       |       |       |       |       |       |      |      | 1,00 |

En verde los valores mayores a 0,7 y en rojo los menores a -0,3.

Tabla 4.4: Promedios por instancia de correlaciones entre variables de  $VA$  y tradicionales.

particular, la primera y tercera fila (y columna) de cada variable son iguales (pues usan los Modos de Agregación Suma y Promedio), y lo mismo ocurre con los valores correspondientes a **SumPointsHull** y **SumForeignPointsHull**, y con los asociados a las 3 variables relacionadas a **DT**.

Al comparar los valores de las tablas 4.3 y 4.4 se ve que las diferencias más grandes se dan en los pares que incluyen simultáneamente una de las siguientes variables: **SumDistBetween**, **SumDistCenter**, **SumRatioCenterV2**, y **SumRatioCenterV1**, y alguna de éstas otras: **MaxPointsHull**, **AvgPointsHull**, **AvgForeignPointsHull**, **MaxDistBetween** y **AvgDistBetween**. Estas combinaciones incluyen los pares para los que la tabla 4.3 entregaba correlaciones negativas, lo que confirma que el ejemplo presentado en la figura 4.1 también ocurre en los otros casos. En particular, la mayor diferencia se da para los pares (**SumRatioCenterV2**, **AvgDistBetween**) y (**SumRatioCenterV2**, **AvgPointsHull**) que pasaron de  $-0,42$  y  $-0,49$ , a  $0,97$  y  $0,87$  respectivamente. Por otro lado, los pares que disminuyeron lo hicieron en menor magnitud, siendo los casos más notables los de las variables asociadas a **Cross** con todas las otras (menos con **SumDistCenter** y **PercentageDT**), y los pares (**SumDistCenter**, **SumRatioCenterV2**) y (**SumDistCenter**, **SumRatioCenterV1**) que pasaron de  $0,81$  y  $0,95$ , a  $0,30$  y  $0,57$  respectivamente.

Con respecto a los resultados de la tabla 4.4 se ve que en general todas las correlaciones son positivas, menos las asociadas a las 3 variables de **Cross**, que son las únicas que consistentemente alcanzan valores negativos con otras variables (sólo obtienen números positivos al

compararse entre ellas, y al compararlas con las variables asociadas a **DT** y **DistCenter**). Por otro lado, al enfocarse en las correlaciones más altas de la tabla, además de los valores correspondientes a distintos Modos de Agregación de las mismas variables, sobresalen la fuerte correlación entre las variables asociadas a **DT** y **SumDistCenter**, y también los valores que alcanzan los pares que incluyen variables de **RatioCenterV2**, **DistBetween** y **PointsHull**. Por último, vale la pena notar que haber introducido **RatioCenterV2** parece haber logrado el objetivo de mitigar los problemas asociados a **RatioCenterV1**, pues ésta alcanza niveles de correlación mucho más altos con las otras variables de *VA*, y también con **PercentageDT**.

Estos resultados sugieren que las variables de *VA* en consideración están midiendo 4 elementos o *factores* ligeramente distintos. Uno de estos está relacionado con las variables de **PointsHull**, **DistBetween** y **RatioCenterV2**, otro con **DistCenter**, el tercero con **RatioCenterV1** y el último con **Cross**.

### Relación de las variables de Atractivo Visual con *DrivingTime*

Como se puede ver en la tabla 4.4, todas las medidas consideradas correlacionan positivamente con las variables asociadas a **DT**, siendo **SumDistCenter** y **MaxDistCenter** las que alcanzan los valores más altos (0,91 y 0,69 respectivamente), seguidas de **Sum(Foreign)-PointsHull**, **SumDistBetween** y **SumRatioCenterV2** (con 0,55, 0,54 y 0,54 respectivamente). Por otro lado, las que muestran una asociación más baja son **MaxRatioCenterV1**, **SumRatioCenterV1**, **MaxCross** y **SumCross** (0,17, 0,28, 0,25 y 0,30 respectivamente). Cabe mencionar que tener una alta correlación es positivo pues permitiría enfocarse sólo en una de las dos variables, y la otra mejoraría en consecuencia de esto, pero tener un grado de asociación bajo también podría ser provechoso pues indicaría la posibilidad de mejorar el nivel de *VA* de una solución, sin tener que perder en términos de **DT**.

Para tener una visión más amplia de este fenómeno se estudian los gráficos de dispersión de **PercentageDT** con las medidas de *VA*. Por razones de espacio, se restringe este estudio detallado a las variables **SumDistCenter** y **SumForeignPointsHull** (se escogen estas dos por sobre las otras pues el análisis previo sugiere que pertenecen a *factores* distintos, y que ambas presentan una alta correlación con **PercentageDT**).

Para poder evaluar las diferencias que existen en el comportamiento presentado en las distintas instancias y grupos de instancias, se grafican los datos correspondientes a las soluciones de cada grupo en gráficos distintos<sup>4</sup>, y dentro del gráfico de cada grupo se asocian colores diferentes a los puntos relacionados con las soluciones de cada instancia particular. Además, se calculan las medidas de *VA* para la MSC de cada instancia y se incluye tal dato en los gráficos, identificándolo con un círculo negro al rededor del punto (es importante notar que por definición, la MSC tiene **PercentageDT** igual a 0, por lo que estos puntos siempre se encuentran encima del eje X). Las figuras 4.2 y 4.3 muestran estos gráficos para las medidas **SumDistCenter** y **SumForeignPointsHull** respectivamente.

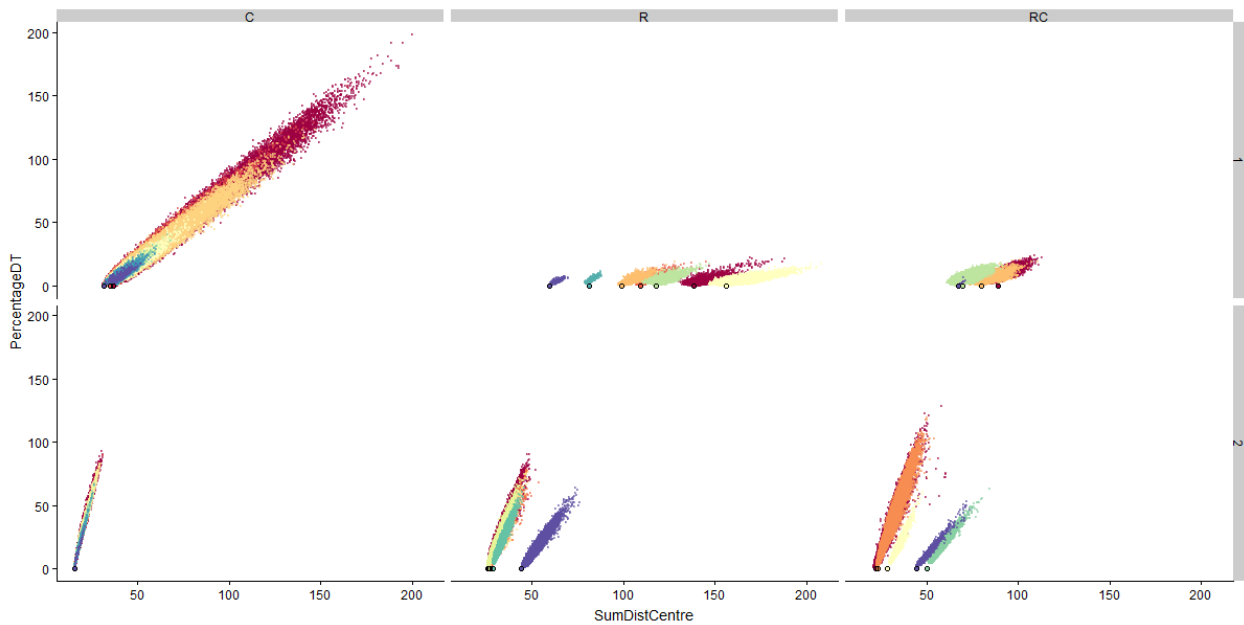


Figura 4.2: **SumDistCenter** vs **PercentageDT** por instancia y grupo

Se aprecia claramente que los datos correspondientes a **SumForeignPointsHull** resultan bastante más dispersos que los asociados a **SumDistCenter**, lo que refleja el hecho de que esta última alcanza una correlación con **PercentageDT** bastante más alta. Ahora, si comparamos la correlación alcanzada por las soluciones de las instancias entre los distintos grupos, se ve que las de tipo C logran una asociación más alta que las de tipo R y RC, lo cual resulta incluso más pronunciado para **SumForeignPointsHull**.

Otra diferencia importante que se puede ver en el comportamiento de los distintos grupos

<sup>4</sup>Pero se mantienen los límites de los ejes iguales en los 6 grupos de instancias, para facilitar la comparación entre éstos.



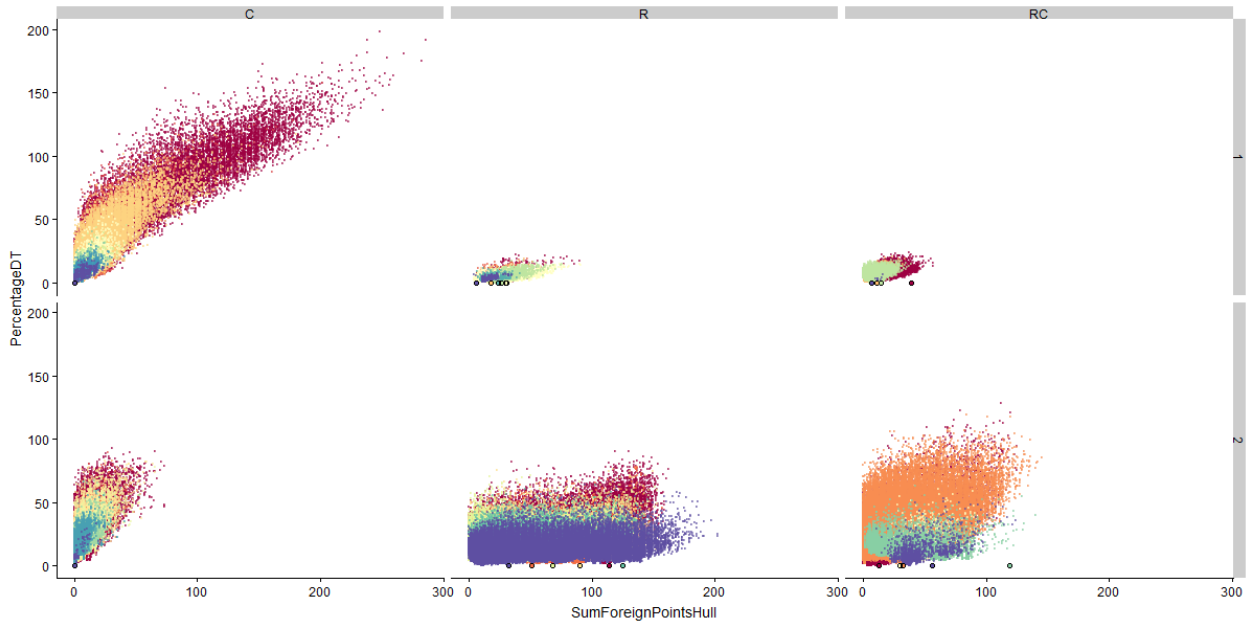


Figura 4.3: **SumForeignPointsHull** vs **PercentageDT** por instancia y grupo

de instancias se relaciona con la posición de la MSC en el eje X. Para las soluciones del grupo C, ésta suele ser simultáneamente la mejor solución para **DT** y para la medida de **VA**, pero en R y en RC (sobre todo para **SumForeignPointsHull**) hay algunas de éstas que alcanzan valores de **VA** bastante lejanos al óptimo (cabe mencionar que el hecho de que un punto esté lejos del eje Y no significa necesariamente que esté lejos del mínimo, pues los óptimos de **VA** pueden ser distintos para cada instancia. Una mejor forma de ver esto es evaluar cuánto más a la derecha está la MSC de una instancia, que el punto más a la izquierda de esa misma instancia). Esto indica que la distribución de los puntos en el plano afectan de manera importante la relación entre **VA** y **DT**. Por otro lado, en las instancias de tipo 1 para **SumDistCenter**, las pendientes son más bajas que en las de tipo 2; parte de esta diferencia se explica por el hecho de que las instancias de tipo 2 usan menos vehículos que las de tipo 1, pero el efecto se mantiene incluso después de corregir por esto.

El comportamiento de las soluciones de instancias del mismo grupo es bastante similar, aunque existen algunas diferencias. Por ejemplo, para **SumForeignPointsHull** en el grupo RC2, las instancias naranja y roja tuvieron muchas más soluciones y cubrieron una parte mayor del espacio que las instancias morada y verde. Algo similar ocurre en este mismo grupo pero para **SumDistCenter**, donde las pendientes asociadas a estas últimas dos instancias son distintas que las de las otras instancias del mismo grupo (nuevamente, esto se mantiene

después de corregir por el número de vehículos usados). Esto indica que incluso en instancias con distribuciones de puntos similares, existen otras variables que influyen en la relación entre  $VA$  y  $DT$ .

Con respecto al *trade-off* entre las variables tradicionales del  $VRP$  y la  $VA$ , se aprecia que para **SumDistCenter** en las instancias de tipo C y también en las de tipo 2, la correlación es tan alta que mejorar o empeorar en una variable obliga a hacer lo mismo en la otra. Esto no es así en las instancias de los grupos R1 y RC1, donde es posible aumentar o disminuir el valor de una variable manteniendo la otra constante, aunque esta libertad se reduce al acercarse al mínimo de cualquiera de las dos variables. Este comportamiento es interesante desde el punto de vista práctico, pues muestra que en una gran parte de los casos, mejorar el valor de **SumDistCenter** sin perder en términos de **PercentageDT**, es o bien obligatorio o al menos posible. Por otro lado, para **SumForeignPointsHull**, en los 6 grupos de instancias es posible tener distintos valores de  $VA$  para **PercentageDT** similares, aunque cerca del mínimo de **PercentageDT** esto sólo sigue siendo factible en los grupos R1, R2 y RC2. Este resultado muestra la relevancia de estudiar las características de las instancias para poder optimizar  $VA$  y  $DT$  simultáneamente.

Si nos enfocamos en la frontera de Pareto del conjunto de soluciones (es decir, en las soluciones no dominadas), el comportamiento también es fuertemente dependiente del grupo de instancias en evaluación. Para las de tipo C, en las dos medidas de  $VA$  en consideración, la frontera corresponde a un único punto. La figura 4.4 muestra este comportamiento para las 9 instancias de C1 con **SumForeignPointsHull**. En estos gráficos (y también en los de la figura 4.5) los ejes tienen límites distintos, para poder evaluar en detalle las instancias con valores más concentrados; la MSC se presenta en color verde y la frontera de Pareto en color rojo<sup>5</sup>. El comportamiento de la frontera en las instancias de R y RC es diferente en las distintas medidas de  $VA$ . Para **SumForeignPointsHull**, como se ve en la figura 4.5, la pendiente de la frontera tiende a ser más pronunciada cerca del mínimo de la variable de  $VA$  y a aplanarse al acercarse al mínimo de  $DT$ . Esto sugiere que para este tipo de instancias, es posible desviarse ligeramente de la MSC para mejorar sustancialmente en términos de  $VA$ ,

---

<sup>5</sup>Para calcular las fronteras de Pareto, sólo se consideraron las soluciones encontradas mediante la metodología presentada en la sección 4.3.1, por lo que hay casos en los que la MSC domina a algunas de las soluciones de ésta.

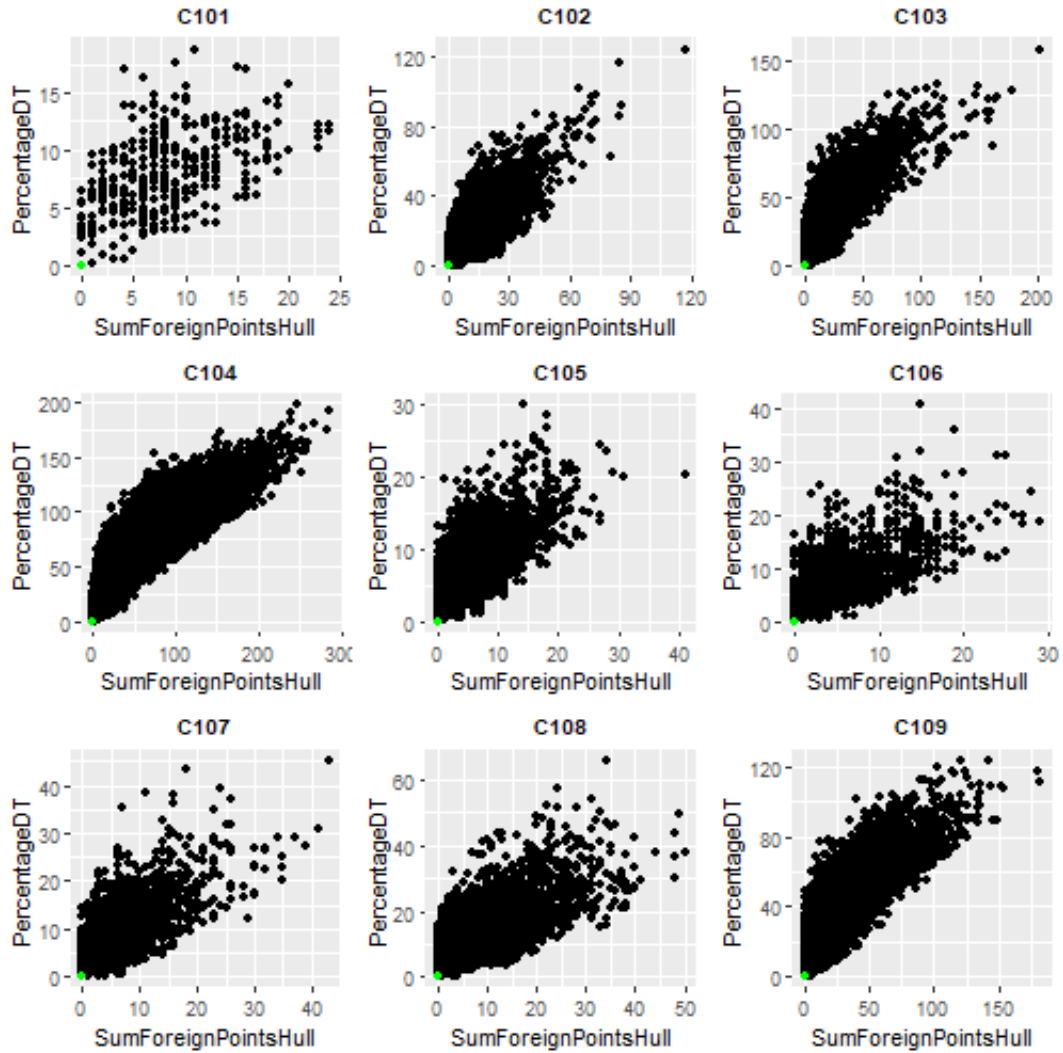


Figura 4.4: **SumForeignPointsHull** vs **PercentageDT** por instancia para grupo C1. Los puntos rojos muestran la frontera de Pareto y el punto verde corresponde a la MSC.

sin embargo para llegar al mínimo de ésta es necesario perder bastante más en términos de optimalidad clásica. Por otro lado, para **SumDistCenter** el comportamiento recién descrito sólo se ve en las instancias de tipo R1. Para las del grupo RC1 la pendiente es más bien constante, con una inclinación leve o intermedia, mientras que para las de R2 y RC2, al tener una correlación tan alta con **PercentageDT**, la frontera para algunas de las instancias es sólo 1 punto, y para otras consiste en valores para la VA muy similares, por lo que aunque la pendiente de ésta sea pronunciada o no, el cambio en **DT** es muy bajo.

Cabe mencionar que para algunas medidas de VA (en particular para **SumDistCenter** y **SumForeignPointsHull**), hacer búsquedas locales intra-ruta en una solución dada no cambiaría sus valores, pero sí podría mejorar los de **PercentageDT**. Esto implica que en

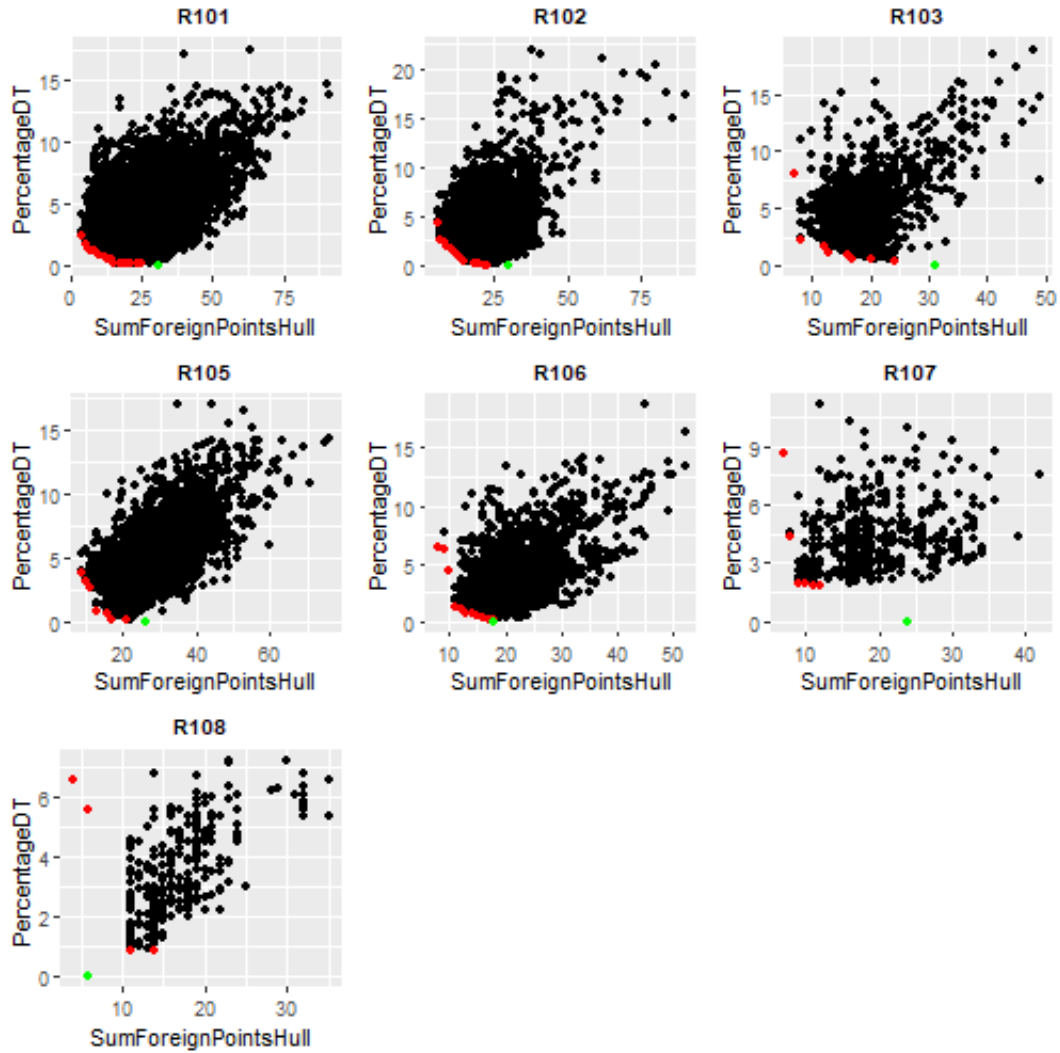


Figura 4.5: **SumForeignPointsHull** vs **PercentageDT** por instancia para grupo R1. Los puntos rojos muestran la frontera de Pareto y el punto verde corresponde a la MSC.

la práctica, muchas de las soluciones presentes en los datos utilizados podrían mejorar su **PercentageDT** sin empeorar en  $VA$ .

Es importante notar que no fue posible determinar si el conjunto de soluciones encontradas es una muestra representativa del espacio de soluciones completo, y tampoco si describe correctamente la forma de éste en los puntos cercanos a los mínimos locales. El hecho de que se haya encontrado un número tan distinto de soluciones para las diferentes instancias (con RC208 y R206 alcanzando 134.894 y 123.384 respectivamente, RC108 y C201 teniendo sólo 12 y 87, y con 17 instancias con 0 soluciones) entrega cierta confianza para las instancias con más soluciones, pero no se cuenta con información suficiente como para confirmar si en los casos con pocas soluciones esto es un indicador de que en estos datos falta una parte importante

del espacio de soluciones, o si esto se debe a que el número de soluciones factibles para el mínimo número de vehículos utilizable en tales instancias es simplemente más bajo.

Ahora, incluso si se tuviera acceso al espacio de soluciones completo y se conociera la frontera de Pareto de cada variable de  $VA$  con respecto a  $DT$ , esta información todavía no sería suficiente para poder tomar decisiones en la práctica, pues no se conoce cuánto cambio es necesario en las distintas variables de  $VA$  para que éste sea distinguible por los usuarios, y tampoco cuánto  $DT$  están dispuestos a perder por mejorar en las distintas medidas de  $VA$ . En el siguiente capítulo se proponen y aplican dos métodos distintos para abordar estas preguntas, y se presentan los resultados obtenidos.

# Capítulo 5

## Confirmación de Relevancia

Si en el capítulo anterior el foco fue analizar la interacción entre las variables de  $VA$  y el  $DT$  en el espacio de soluciones factibles, el de éste puede entenderse como el estudio de las soluciones *preferibles*. En concreto, se evalúa la influencia que tiene la  $VA$  a la hora de decidir qué solución es preferida por los usuarios de SimpliRoute, se mide cuánto cambio en  $VA$  es necesario para ser percibido por estos, y cuánta mejora en  $VA$  compensa un aumento de 1 kilómetro en  $DT$ .

Para esto se plantean dos estudios. En el primero, descrito en la sección 5.1, se utiliza la información guardada en la base de datos de SimpliRoute para analizar las ediciones manuales que hacen los usuarios de la plataforma a las soluciones entregadas por el algoritmo, con la idea de que los cambios que realizan podrían indicar qué variables son relevantes para ellos a la hora de evaluar la calidad de una solución. El segundo, presentado en la sección 5.2, consiste en una encuesta en la que se les pidió a distintos clientes de la empresa que respondan una serie de preguntas en las cuales se les mostraban dos soluciones y se les pedía que eligieran cuál consideraban más atractiva.

### 5.1. Evaluación de cambios hechos por clientes en su operación diaria

Se consideraron las 3.323 tuplas (instancia, solución entregada por el algoritmo, solución guardada después de la posible edición manual) almacenadas en la base de datos de SimpliRoute

correspondientes a optimizaciones realizadas por clientes en la plataforma entre el 20 de abril y el 15 de agosto del 2017. Se escogieron estas fechas porque el análisis se hizo el 16 de agosto de tal año, y pues no se pudo conseguir información más antigua debido a que el 19 de abril se realizó un cambio en la forma en la que se almacenaban los datos, que causó que no sea posible asociar confiablemente los resultados de la optimización entregados por el algoritmo con la solución guardada por el cliente después de editarla.

Para asegurar la consistencia y relevancia del análisis, se definieron los siguientes criterios para decidir qué tuplas incluir en el estudio:

- Que el número de vehículos usados en la solución editada sea el mismo que en la solución original, y que sea mayor a 1.
- Que los puntos visitados en la solución editada sean los mismos que los de la solución original, y que haya más de 4.
- Que los vehículos de ambas soluciones tengan como máximo 1 ruta asociada (es decir, que sólo salgan y vuelvan al *depot* una vez).

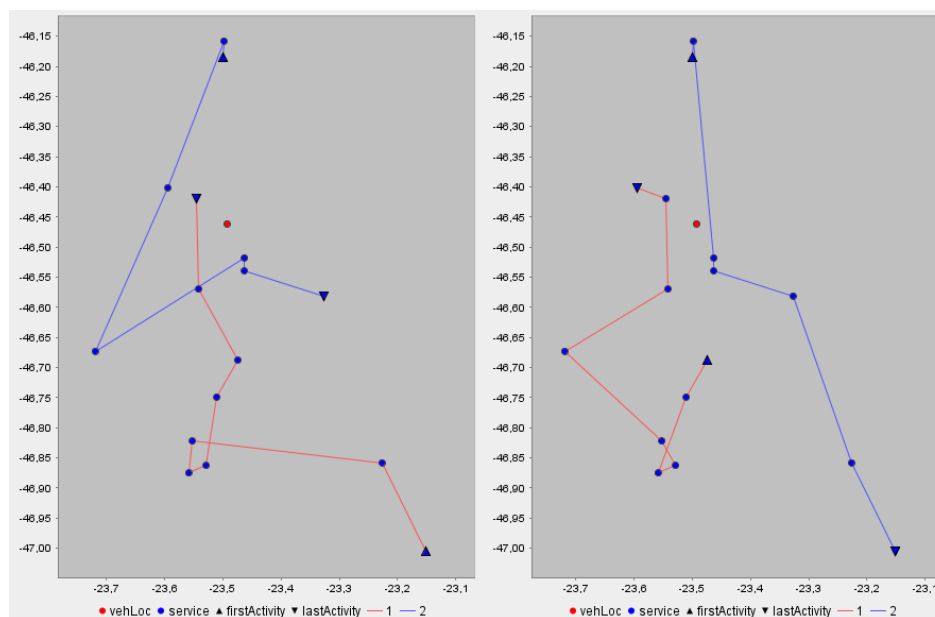


Figura 5.1: Izquierda: Solución original. Derecha: Solución después de ser editada por el planificador.

Después de aplicar estos criterios, quedaron 99 tuplas para analizar, que corresponden a instancias de optimización de cuentas de 31 clientes. La figura 5.1 muestra un ejemplo de una

de éstas. Los 3 clientes con más tuplas, en conjunto, abarcan más de la mitad de los datos, con 51 tuplas asociadas a sus cuentas. Por esto, en el análisis que se presenta a continuación, además de evaluar los resultados generales, se discute sobre las posibles diferencias que podría haber entre estos dos grupos (desde aquí en adelante se llamará ‘Grupo A’ al conjunto de las 51 instancias de las 3 cuentas más grandes, y ‘Grupo B’ a las 48 instancias restantes).

La tabla 5.1 muestra un resumen descriptivo de las 99 instancias incluidas en el análisis. Se puede ver que la gran mayoría de éstas (salvo 6) son pequeñas, con menos de 100 puntos y hasta 5 vehículos, y que en general el porcentaje de edición es independiente del tamaño de las instancias.

| <b>Vehículos</b> | <b>Número</b> | <b>Editadas</b> | <b>Puntos</b> | <b>Número</b> | <b>Editadas</b> |
|------------------|---------------|-----------------|---------------|---------------|-----------------|
| 2                | 77            | 78 %            | 5-10          | 24            | 75 %            |
| 3                | 10            | 60 %            | 11-20         | 47            | 74 %            |
| 4                | 5             | 40 %            | 21-30         | 9             | 67 %            |
| 5                | 5             | 100 %           | 31-50         | 11            | 82 %            |
| 10               | 1             | 0 %             | 51-100        | 2             | 50 %            |
| 23               | 1             | 0 %             | 101-150       | 4             | 75 %            |
|                  |               |                 | 151-250       | 2             | 50 %            |
| <b>TODAS</b>     | <b>99</b>     | <b>74 %</b>     | <b>TODAS</b>  | <b>99</b>     | <b>74 %</b>     |

Tabla 5.1: Número de instancias incluidas en el análisis por cantidad de vehículos y puntos, y porcentaje de soluciones editadas en cada categoría.

Dentro de las 26 instancias que no fueron editadas, sólo 4 pertenecen a al ‘Grupo A’. Esto implica que el porcentaje de soluciones editadas para este grupo (92,2 %) es mucho mayor que para el ‘Grupo B’ (54,2 %). A pesar de esta diferencia, incluso en el grupo con menor cantidad de ediciones, más de la mitad de las soluciones entregadas por la plataforma parecen no ser las esperadas por los clientes. Esto es más relevante todavía, si se considera que los criterios de inclusión dejaron fuera de este análisis muchas soluciones por editar ‘*demasiado*’.

También se encuentran diferencias significativas<sup>1</sup> entre estos dos grupos al comparar el número de vehículos usados, el número de puntos insertados, el **DT** y la mayoría de las variables de *VA* consideradas; la tabla 5.2 indica el número de instancias de cada grupo, según el número de vehículos y de puntos. Estas diferencias no son necesariamente un problema para el análisis que se presenta a continuación, pues en vez de considerar el cambio absoluto que

<sup>1</sup> $p < 0,01$  usando un t-test heterocedástico de dos colas.



| Vehículos    | Grupo A   | Grupo B   | Puntos       | Grupo A   | Grupo B   |
|--------------|-----------|-----------|--------------|-----------|-----------|
| 2            | 51        | 26        | 5-10         | 12        | 12        |
| 3            | 0         | 10        | 11-20        | 30        | 17        |
| 4            | 0         | 5         | 21-30        | 4         | 5         |
| 5            | 0         | 5         | 31-50        | 5         | 6         |
| 10           | 0         | 1         | 51-100       | 0         | 2         |
| 23           | 0         | 1         | 101-150      | 0         | 4         |
|              |           |           | 151-250      | 0         | 2         |
| <b>TODAS</b> | <b>51</b> | <b>48</b> | <b>TODAS</b> | <b>51</b> | <b>48</b> |

Tabla 5.2: Número de instancias incluidas en el análisis de cada grupo, por cantidad de vehículos y de puntos

causan las ediciones de las soluciones, se puede evaluar el cambio porcentual, para así evitar posibles distorsiones por las diferencias de escala. Si se considera tal medida, no se registran diferencias significativas entre estos dos grupos<sup>2</sup> (incluso si se incluyen sólo las instancias que fueron editadas). Así, los resultados presentados a continuación son válidos para toda la muestra en consideración.

Como se ve en la tabla 5.3 que muestra el promedio alcanzado en cada variable por las soluciones de cada grupo antes y después de editar, y también en las soluciones que no fueron editadas, todas las variables evaluadas tienden a mejorar en las soluciones editadas, siendo las asociadas a **ForeignPointsHull** las que mayor cambio muestran. La comparación entre los valores del grupo ‘Sin Cambio’ y los otros puede ser engañosa, pues al tratarse de instancias distintas, es posible que haya diferencias entre éstas que influyan en los valores absolutos de las distintas medidas. Dicho esto, resalta el hecho de que para la variable **ForeignPointsHull**, los valores de las instancias ‘Sin Cambio’ son incluso más bajos que en las soluciones editadas, incluso llegando al mínimo en las 4 del grupo A. Por otro lado, llama la atención que **DT** mejore al editar, pues eso podría indicar que la solución entregada por el algoritmo de SimpliRoute no es de muy buena calidad, pero en realidad la mayor parte de las soluciones que logran mejorar el **DT** lo hacen generando infactibilidades, generalmente rompiendo ventanas horarias de clientes o jornadas laborales de los vehículos. En concreto, de las 73 soluciones editadas, 42 mejoran en **DT** pero 24 de esas dejan de ser factibles (en comparación, entre las 31 que empeoran en **DT**, sólo 5 dejan de ser factibles). Más aún, las 18 soluciones que logran

<sup>2</sup> $p > 0,05$  usando un t-test heterocedástico de dos colas. Sólo una de las 21 variables de VA evaluadas tiene diferencias significativas, lo cual (independientemente de la alta correlación existente entre algunas variables) deja de ser significativo después de corregir por las múltiples comparaciones realizadas.

|                      | Grupo A  |         |            | Grupo B  |         |            |
|----------------------|----------|---------|------------|----------|---------|------------|
|                      | Original | Editada | Sin Cambio | Original | Editada | Sin Cambio |
| Número de Instancias | 47       | 47      | 4          | 26       | 26      | 22         |
| Clientes Insertados  | 15,34    | 15,34   | 14,00      | 38,77    | 38,77   | 29,14      |
| Vehículos Usados     | 2,00     | 2,00    | 2,00       | 2,96     | 2,96    | 3,77       |
| DT                   | 226,92   | 215,38  | 280,63     | 181,47   | 178,16  | 143,49     |
| SumPointsHull        | 16,11    | 15,64   | 14,00      | 53,38    | 46,08   | 35,36      |
| MaxPointsHull        | 10,19    | 9,11    | 8,50       | 22,54    | 18,04   | 12,32      |
| AvgPointsHull        | 8,05     | 7,82    | 7,00       | 14,66    | 12,75   | 9,09       |
| SumForeignPointsHull | 0,77     | 0,30    | 0,00       | 14,62    | 7,31    | 6,23       |
| MaxForeignPointsHull | 0,62     | 0,21    | 0,00       | 8,73     | 5,23    | 3,18       |
| AvgForeignPointsHull | 0,38     | 0,15    | 0,00       | 3,64     | 1,73    | 1,34       |
| SumCross             | 0,91     | 0,55    | 1,00       | 2,23     | 2,09    | 2,12       |
| MaxCross             | 0,83     | 0,55    | 0,75       | 1,46     | 1,24    | 1,14       |
| AvgCross             | 0,46     | 0,28    | 0,50       | 0,65     | 0,64    | 0,57       |
| SumDistCenter        | 24,77    | 22,81   | 26,85      | 15,71    | 15,01   | 14,38      |
| MaxDistCenter        | 16,13    | 13,69   | 15,67      | 7,98     | 7,33    | 6,21       |
| AvgDistCenter        | 12,38    | 11,41   | 13,42      | 5,41     | 5,17    | 4,04       |
| SumDistBetween       | 19,54    | 16,69   | 19,99      | 10,55    | 10,30   | 10,14      |
| MaxDistBetween       | 12,62    | 9,89    | 11,58      | 5,54     | 5,45    | 4,64       |
| AvgDistBetween       | 9,77     | 8,35    | 10,00      | 3,87     | 3,84    | 2,82       |
| SumRatioCenterV2     | 0,63     | 0,62    | 0,41       | 1,10     | 1,06    | 1,55       |
| MaxRatioCenterV2     | 0,40     | 0,37    | 0,24       | 0,48     | 0,45    | 0,51       |
| AvgRatioCenterV2     | 0,32     | 0,31    | 0,21       | 0,35     | 0,33    | 0,35       |
| SumRatioCenterV1     | 2,54     | 2,34    | 2,01       | 3,61     | 3,50    | 10,46      |
| MaxRatioCenterV1     | 1,51     | 1,34    | 1,01       | 1,45     | 1,35    | 7,09       |
| AvgRatioCenterV1     | 1,27     | 1,17    | 1,01       | 1,18     | 1,14    | 3,99       |

Tabla 5.3: Promedio de valores alcanzados por las distintas variables consideradas, en las soluciones de las instancias que no fueron editadas ('Sin Cambio'), y en las que sí cambiaron, para los casos antes de editar ('Original'), después de editar ('Editada'), en cada grupo.

reducir el **DT** manteniendo la factibilidad logran una disminución de un 5,4% en promedio, en comparación con el 11% que alcanzan las no factibles.

Con respecto a las variables de *VA*, se analizan sólo las con Modo de Agregación 'Promedio' para evitar resultados distorsionados por los cambios de escala entre las distintas soluciones. La tabla 5.4 muestra, para cada variable, el porcentaje de instancias que mejora, se mantiene igual, o empeora con la edición (con respecto a las 73 instancias editadas). En ésta se ve que las variables que menos mejoran son las relacionadas a **PointsHull**, pero que al mismo tiempo son las que menos empeoran. Esto llama la atención, pues la tabla anterior mostraba a **AvgForeignPointsHull** como una de las variables con mayor cambio, por lo que no queda claro si esta variable es relevante para los planificadores al momento de editar las soluciones; algo similar, pero en menor magnitud, ocurre con **AvgCross**. Por otro lado, las

|                      | <b>Bajó</b> | <b>Igual</b> | <b>Subió</b> |
|----------------------|-------------|--------------|--------------|
| AvgPointsHull        | 33 %        | 58 %         | 10 %         |
| AvgForeignPointsHull | 33 %        | 58 %         | 10 %         |
| AvgCross             | 38 %        | 36 %         | 26 %         |
| AvgDistCenter        | 42 %        | 14 %         | 44 %         |
| AvgDistBetween       | 62 %        | 0 %          | 38 %         |
| AvgRatioCenterV2     | 45 %        | 14 %         | 41 %         |
| AvgRatioCenterV1     | 47 %        | 21 %         | 33 %         |

Tabla 5.4: Porcentaje de instancias que mejora, se mantiene igual, o empeora con la edición.

variables **AvgDistCenter** y **AvgRatioCenterV2** muestran un número similar de instancias que mejoran y que empeoran, lo que sugiere que éstas podrían tener menor relevancia al momento de editar. Por último, llama la atención que la única variable que nunca se mantiene igual es **AvgDistBetween**, lo que podría sugerir que esta variable es altamente relevante al momento de editar, pero esto en realidad es consecuencia de la definición de la variable, pues cualquier edición realizada puede potencialmente cambiar el **AvgDistBetween**, mientras que las otras variables (como las asociadas a **PointsHull**, **DistCenter**, **RatioCenterV1** y **RatioCenterV2**) no cambian si es que se realizan ediciones que sólo alteran el orden de los puntos dentro de sus vehículos originales (sin mover puntos de un vehículo a otro).

Para entender de mejor forma cómo afecta la edición de las soluciones a las distintas variables, se estudia en detalle cada una de éstas por separado. Las tablas 5.5, 5.6, 5.7, 5.8, 5.9 y 5.10 muestran para cada medida en consideración, el número de soluciones que empeoran, se mantienen igual, o mejoran, segmentando según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no. Además, para **AvgForeignPointsHull**, **AvgCross**, y **AvgRatioCenterV1**, se segmenta adicionalmente según si el valor original de la medida era mínimo o no; esto se realiza sólo para estas 3 variables, pues el mínimo de éstas es fácilmente identificable.

- **AvgForeignPointsHull:** Si se analiza la tabla 5.5, se puede entender el hecho mencionado anteriormente, de que esta variable presenta un gran cambio pero en un bajo número de instancias. En ésta se ve que 45 de las 73 soluciones originales ya tienen el mínimo valor alcanzable por esta variable, por lo que en todos esos casos no es posible mejorar más. Al evaluar el cambio desde esta perspectiva, se ve que cuando el

| <b>AvgForeignPointsHull</b>        | <b>DT Empeoró</b> |                  | <b>DT Mejoró</b> |                  | <b>TOTAL</b> |
|------------------------------------|-------------------|------------------|------------------|------------------|--------------|
|                                    | <b>Mínimo</b>     | <b>Mejorable</b> | <b>Mínimo</b>    | <b>Mejorable</b> |              |
| Empeoró (Edición Factible)         | 3                 | 1                | 1                | 1                | 6            |
| Igual (Edición Factible)           | 18                | 1                | 8                | 0                | 27           |
| Mejoró (Edición Factible)          | 0                 | 3                | 0                | 8                | 11           |
| Empeoró (Edición NO Factible)      | 1                 | 0                | 0                | 0                | 1            |
| Igual (Edición NO Factible)        | 2                 | 0                | 12               | 1                | 15           |
| Mejoró (Edición NO Factible)       | 0                 | 2                | 0                | 11               | 13           |
| <b>TOTAL (Edición Factible)</b>    | <b>21</b>         | <b>5</b>         | <b>9</b>         | <b>9</b>         | <b>44</b>    |
| <b>TOTAL (Edición NO Factible)</b> | <b>3</b>          | <b>2</b>         | <b>12</b>        | <b>12</b>        | <b>29</b>    |

Tabla 5.5: Número de soluciones que empeoran, se mantienen igual, o mejoran en **AvgForeignPointsHull**, según si su valor original era mínimo o no, según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no.

**AvgForeignPointsHull** es mejorable, en la gran mayoría de los casos éste mejora, y cuando no lo es, en muy pocos casos se elige empeorarlo; todo esto, independientemente de si el **DT** mejora o no, y de si la solución deja de ser factible con la edición o no. Más aún, si se evalúa el porcentaje de soluciones que alcanza el mínimo, se ve que el 69,2% de las instancias que no fueron editadas llegaron a tal valor, mientras que para las que sí se modificaron este indicador fue de un 61,6% antes de editar y 78,1% después de hacerlo. Esto muestra dos puntos importantes: primero, que las instancias que no se editaron tenían un alto el porcentaje de soluciones con el mínimo, lo que podría indicar que esta es una característica que se busca en las soluciones (ya que cuando está presente se suele no modificar), y segundo, que aunque en las que sí se editaron este valor también era originalmente alto, luego de la edición éste aumentó aún más, dejando muy pocas instancias con valores mayores al mínimo posible.

Por otro lado, si se analiza la magnitud de las ediciones que se realizaron, se ve que dentro de las 14 soluciones que originalmente eran mejorables y terminaron siendo no factibles, hay sólo una que se mantiene igual, ninguna empeora, dos que mejoran en un 79% y 81% (lo que corresponde a un cambio absoluto de 27 y 7 respectivamente), y las otras nueve mejoran en un 100% (correspondiente a un 1,43 en promedio con valores entre  $\frac{1}{3}$  y 3). En comparación, entre las 14 soluciones que originalmente eran mejorables y siguieron siendo factibles después de la edición manual, también sólo una se mantiene igual, dos empeoran pasando de 0,5 a 1 y de 0,5 a 2, cinco mejoran un 42% en promedio (equivalente a un cambio absoluto de 1,56 en promedio con valores

entre 0,5 y 3,2), y seis mejoran en un 100% (correspondiente a un cambio absoluto de 1,2 en promedio con valores entre 0,5 y  $\frac{5}{3}$ ). Así, parece ser que esta variable es relevante para los planificadores al momento de editar las soluciones entregadas por el algoritmo, pues suelen mejorarla cuando es posible, muchas veces llegando al mínimo alcanzable, incluso si esto lleva a infactibilidades para el problema original.

| <b>AvgCross</b>                    | <b>DT Empeoró</b> |                  | <b>DT Mejoró</b> |                  | <b>TOTAL</b> |
|------------------------------------|-------------------|------------------|------------------|------------------|--------------|
|                                    | <b>Mínimo</b>     | <b>Mejorable</b> | <b>Mínimo</b>    | <b>Mejorable</b> |              |
| Empeoró (Edición Factible)         | 6                 | 3                | 5                | 0                | 14           |
| Igual (Edición Factible)           | 7                 | 1                | 6                | 3                | 17           |
| Mejoró (Edición Factible)          | 0                 | 9                | 0                | 4                | 13           |
| Empeoró (Edición NO Factible)      | 0                 | 1                | 3                | 1                | 5            |
| Igual (Edición NO Factible)        | 1                 | 0                | 7                | 1                | 9            |
| Mejoró (Edición NO Factible)       | 0                 | 3                | 0                | 12               | 15           |
| <b>TOTAL (Edición Factible)</b>    | <b>13</b>         | <b>13</b>        | <b>11</b>        | <b>7</b>         | <b>44</b>    |
| <b>TOTAL (Edición NO Factible)</b> | <b>1</b>          | <b>4</b>         | <b>10</b>        | <b>14</b>        | <b>29</b>    |

Tabla 5.6: Número de soluciones que empeoran, se mantienen igual, o mejoran en **AvgCross**, según si su valor original era mínimo o no, según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no.

- **AvgCross:** Las conclusiones a las que se puede llegar estudiando la tabla 5.6 son similares a las alcanzadas para **AvgForeignPointsHull**, pero menos pronunciadas, pues cuando esta variable es mejorable se suele mejorar, pero (a diferencia de **AvgForeignPointsHull**) cuando está en el mínimo, en cerca de la mitad de los casos se elige empeorarla. Esto último impacta en el porcentaje de soluciones que alcanzan el mínimo, que para las instancias no editadas fue de un 38,5%, mientras que en las que se modificaron, originalmente el valor era de un 47,9%, y después del cambio sólo sube a un 49,3%. Llama la atención el hecho de que la situación en la que el **DT** mejora y la nueva solución ya no es factible, es cuando la edición más favorece a esta variable. Todo esto sugiere que esta variable es relevante, pero en menor nivel que otras. Con respecto a la magnitud del cambio generado, cuando mejoran lo hacen en un promedio en un 77%, con 15 de las 28 instancias bajando un 100%, es decir, alcanzando el mínimo (esto equivale a un cambio absoluto de 0,73 en promedio, con valores entre 0,5 y 2). En comparación, cuando son mejorables, la 5 instancias que empeoran lo hacen en un -65% (equivalente a un cambio absoluto de -0,63 en promedio, con un rango de  $[-1, -\frac{1}{3}]$ ), y cuando no son mejorables lo hacen en una magnitud similar.

| <b>AvgDistCenter</b>               | <b>DT Empeoró</b> | <b>DT Mejoró</b> | <b>TOTAL</b> |
|------------------------------------|-------------------|------------------|--------------|
| Empeoró (Edición Factible)         | 13                | 8                | 21           |
| Igual (Edición Factible)           | 5                 | 2                | 7            |
| Mejoró (Edición Factible)          | 8                 | 8                | 16           |
| Empeoró (Edición NO Factible)      | 3                 | 8                | 11           |
| Igual (Edición NO Factible)        | 1                 | 2                | 3            |
| Mejoró (Edición NO Factible)       | 1                 | 14               | 15           |
| <b>TOTAL (Edición Factible)</b>    | <b>26</b>         | <b>18</b>        | <b>44</b>    |
| <b>TOTAL (Edición NO Factible)</b> | <b>5</b>          | <b>24</b>        | <b>29</b>    |

Tabla 5.7: Número de soluciones que empeoran, se mantienen igual, o mejoran en **AvgDistCenter**, según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no.

- **AvgDistCenter:** A pesar de que para esta variable no fue posible evaluar si es que las soluciones originales ya tenían el mínimo alcanzable o no, sólo 10 de las 73 instancias no cambiaron para esta medida, por lo que aunque es posible que algunas de estas soluciones no hayan mejorado porque no era factible hacerlo, el hecho de que sean pocas hace que esto no sea un problema importante para analizar los resultados. Como se vio en el capítulo 4, esta variable está fuertemente correlacionada con **DT**, y aunque algo de esto se ve en la tabla 5.7, pues cuando **DT** empeora, el número de instancias que empeora tiende a ser mayor que cuando **DT** mejora, también se ve que la cantidad de instancias que se mueve en la dirección opuesta es considerable. De hecho, cuando el **DT** y el **AvgDistCenter** empeoran, el **AvgDistCenter** sube en promedio un 29% cuando la solución se mantiene factible y un 68% cuando deja de serlo, en comparación con el aumento de un 14% y un 11% respectivamente cuando el **DT** mejora; un comportamiento similar se ve cuando el **AvgDistCenter** mejora, bajando un 6% y un 16% cuando el **DT** empeora (para los casos donde se mantiene la factibilidad o no, respectivamente), y un 10% adicional en cada caso cuando el **DT** mejora. Vale la pena mencionar que si comparamos el cambio de esta variable para las instancias en las que la edición mantiene la factibilidad en comparación con cuando ésta se pierde, se ve que en este último caso el cambio tiene a ser más positivo, es decir, la variable mejora más en los casos en los que mejora, y empeora menos en los casos en los que empeora.
- **AvgDistBetween:** De forma similar a como ocurrió con **AvgDistCenter**, el hecho de no poder evaluar si la solución original ya tenía el mínimo para esta variable sólo

| <b>AvgDistBetween</b>              | <b>DT Empeoró</b> | <b>DT Mejoró</b> | <b>TOTAL</b> |
|------------------------------------|-------------------|------------------|--------------|
| Empeoró (Edición Factible)         | 17                | 2                | 19           |
| Igual (Edición Factible)           | 0                 | 0                | 0            |
| Mejoró (Edición Factible)          | 9                 | 16               | 25           |
| Empeoró (Edición NO Factible)      | 3                 | 6                | 9            |
| Igual (Edición NO Factible)        | 0                 | 0                | 0            |
| Mejoró (Edición NO Factible)       | 2                 | 18               | 20           |
| <b>TOTAL (Edición Factible)</b>    | <b>26</b>         | <b>18</b>        | <b>44</b>    |
| <b>TOTAL (Edición NO Factible)</b> | <b>5</b>          | <b>24</b>        | <b>29</b>    |

Tabla 5.8: Número de soluciones que empeoran, se mantienen igual, o mejoran en **AvgDistBetween**, según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no.

afecta parcialmente el análisis, pues el 100 % de las instancias editadas cambiaron su valor para esta medida. Los resultados también son similares a los alcanzados por la variable recién mencionada, pero en este caso la asociación a **DT** parece ser más fuerte todavía, lo que hace que hipotetizar sobre el objetivo de la edición manual para esta variable sea menos confiable.

| <b>AvgRatioCenterV1</b>            | <b>DT Empeoró</b> |                  | <b>DT Mejoró</b> |                  | <b>TOTAL</b> |
|------------------------------------|-------------------|------------------|------------------|------------------|--------------|
|                                    | <b>Mínimo</b>     | <b>Mejorable</b> | <b>Mínimo</b>    | <b>Mejorable</b> |              |
| Empeoró (Edición Factible)         | 1                 | 11               | 2                | 2                | 17           |
| Igual (Edición Factible)           | 5                 | 2                | 3                | 1                | 11           |
| Mejoró (Edición Factible)          | 0                 | 6                | 0                | 10               | 16           |
| Empeoró (Edición NO Factible)      | 1                 | 0                | 1                | 5                | 7            |
| Igual (Edición NO Factible)        | 0                 | 1                | 2                | 1                | 4            |
| Mejoró (Edición NO Factible)       | 0                 | 3                | 0                | 15               | 18           |
| <b>TOTAL (Edición Factible)</b>    | <b>7</b>          | <b>19</b>        | <b>5</b>         | <b>13</b>        | <b>44</b>    |
| <b>TOTAL (Edición NO Factible)</b> | <b>1</b>          | <b>4</b>         | <b>3</b>         | <b>21</b>        | <b>29</b>    |

Tabla 5.9: Número de soluciones que empeoran, se mantienen igual, o mejoran en **AvgRatioCenterV1**, según si su valor original era mínimo o no, según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no.

- AvgRatioCenterV1:** Como esta medida sí tiene un mínimo fácilmente reconocible, la tabla 5.9 que resume los resultados alcanzados en ésta, también los segmenta según cuando la solución original era mejorable o no. Si bien en muchos casos la edición mejora el valor de esta medida, y aunque cuando este valor ya está en el mínimo suele no empeorarse, se ve también que cuando es mejorable (que es en la mayoría de los casos), el hecho de que mejore o empeore está fuertemente relacionado con si el **DT** mejora o no, así, nuevamente no es posible reconocer si es que la edición en este caso

está enfocada en mejorar esta variable o el **DT**.

| <b>AvgRatioCenterV2</b>            | <b>DT Empeoró</b> | <b>DT Mejoró</b> | <b>TOTAL</b> |
|------------------------------------|-------------------|------------------|--------------|
| Empeoró (Edición Factible)         | 14                | 9                | 23           |
| Igual (Edición Factible)           | 5                 | 2                | 7            |
| Mejoró (Edición Factible)          | 7                 | 7                | 14           |
| Empeoró (Edición NO Factible)      | 2                 | 5                | 7            |
| Igual (Edición NO Factible)        | 1                 | 2                | 3            |
| Mejoró (Edición NO Factible)       | 2                 | 17               | 19           |
| <b>TOTAL (Edición Factible)</b>    | <b>26</b>         | <b>18</b>        | <b>44</b>    |
| <b>TOTAL (Edición NO Factible)</b> | <b>5</b>          | <b>24</b>        | <b>29</b>    |

Tabla 5.10: Número de soluciones que empeoran, se mantienen igual, o mejoran en **AvgRatioCenterV2**, según si la edición mejoró el **DT** o no, y según si la edición mantuvo la factibilidad o no.

- AvgRatioCenterV2:** A pesar de que para esta medida existe algo de asociación en la dirección del cambio en **DT** y en sí misma, lo que más resalta en este caso es que la edición es mucho más favorable para ésta cuando el cambio hace que la solución deje de ser factible. De hecho, el porcentaje de instancias que mejora en este último caso es más del doble que el que lo hace cuando la edición mantiene la factibilidad (65,5% versus 31,8% respectivamente). Esto podría indicar que esta variable tiene cierta relevancia para los usuarios, pero que les es más difícil de mejorar manualmente.

Al evaluar en detalle el cambio de cada variable, se ve que **AvgForeignPointsHull** es la que más favorablemente se comporta, seguida de **AvgCross**. Las otras medidas muestran resultados mixtos, en general asociados a cambios en **DT**. Dicho esto, vale la pena mencionar que a pesar de que en el capítulo 4 se vio que la correlación entre **AvgDistCenter** y **DT** es más alta que la que alcanzan **AvgDistBetween** y **AvgRatioCenterV1** con **DT**, el cambio en las ediciones asocia más fuertemente a estas últimas dos variables que a la primera. Este desacople podría ser un indicador de que los usuarios están intentando mejorar el **AvgDistCenter**, independientemente del **DT**.

Es importante dejar claro que los diversos análisis realizados en esta sección sirven sólo como generadores de hipótesis, y no deben de ser interpretados como resultados comprobados. Para poder llegar a conclusiones más sólidas, en la sección 5.2 se utiliza la información de este análisis para guiar el diseño de un estudio que permita estimar la relevancia de la **VA**



para los usuarios de la plataforma.

## 5.2. Encuestas a Clientes

Para medir la importancia que tienen las distintas características de las soluciones para los planificadores a la hora de evaluar la calidad de éstas, se diseña una encuesta consistente en una serie de preguntas que muestran dos soluciones factibles de una misma instancia (junto con información sobre el número de kilómetros recorridos por los vehículos en cada una), y que pide a los encuestados elegir la que consideren mejor entre las dos.

Dado el enfoque confirmatorio de este estudio, y el alcance limitado en el número posible de respuestas a conseguir, se escoge evaluar sólo el objetivo tradicional del *VRP* (**DT**), y dos medidas de *VA*. Considerando esta restricción, es necesario elegir cuáles de las variables estudiadas anteriormente incluir. El trabajo realizado en el capítulo 4 propone que las variables podrían organizarse en 4 *factores* distintos, lo que implica que escoger medidas de *factores* diferentes probablemente entregue información más rica. Además, en ese capítulo se ve que las variables que más fuertemente correlacionan con **DT** son las asociadas a **DistCenter**. Por otro lado, el análisis realizado en la sección 5.1 sugiere que **AvgForeignPointsHull**, **AvgCross** y **AvgDistCenter** (en ese orden de importancia) son las variables que más probablemente tienen en mente los planificadores al momento de editar manualmente las soluciones. Por último, los resultados de las entrevistas presentados en el capítulo 3 indican que la compacidad y sectorización de las soluciones son características deseables. Así, considerando todos estos puntos, se escoge evaluar las medidas **AvgForeignPointsHull** y **AvgDistCenter**.

### 5.2.1. Metodología

Para estudiar la importancia de las 3 variables en consideración se planteó el problema bajo el marco de referencia del Análisis Conjunto [23], que es una técnica estadística que utiliza encuestas para evaluar la relevancia de los atributos de un conjunto de objetos basándose en las respuestas entregadas por los encuestados. En particular se utiliza el análisis conjunto basado en la elección (*choice-based conjoint (CBC)*), que consiste en preguntas en las que

las personas encuestadas deben de escoger la opción que más prefieran entre dos o más alternativas.

Para esta metodología se comienza definiendo un objeto a evaluar, que en este caso son soluciones a una instancia del *VRP*; se elige usar varias soluciones distintas de una misma instancia, pues de esta forma se reduce la varianza interna de las respuestas. La instancia por utilizar se escogió aleatoriamente dentro de la base de datos de SimpliRoute, considerando sólo las con puntos únicamente en Santiago (pues la encuesta sería respondida por clientes de esta ciudad), que tuviesen entre 30 y 80 puntos, y entre 4 y 6 vehículos, todos partiendo y terminando en el mismo *depot*. La instancia finalmente seleccionada tiene 6 vehículos (con jornada laboral de 4,5 horas) y 51 puntos, de los cuales 17 tienen ventana horaria (con largo promedio de 3,4 horas y un rango de entre 1 y 4 horas); dadas las características del problema, la capacidad de los vehículos no es una restricción activa. Se lograron generar 99.254 soluciones distintas (todas con 6 vehículos usados) con la misma metodología que se utilizó para generar soluciones en el capítulo 4.

Luego, se escoge un conjunto de atributos (del objeto) que se quieren estudiar, que en este caso corresponden a las 3 medidas seleccionadas previamente, y se definen ciertos niveles para cada una de estas características. Debido a que las variables que estamos considerando son continuas, para generar estos niveles es necesario discretizarlas, pero como no se conoce la capacidad visual que tienen las personas para distinguir diferencias en los atributos de las distintas soluciones, se escogió una discretización relativamente arbitraria, que segmentó cada variable en 4 niveles. Por esta razón, si los resultados del análisis no encuentran diferencias entre los niveles de alguna variable, podría deberse a que la segmentación realizada resultó muy gruesa. Para discretizar las variables **DT** y **AvgDistCenter**, se expresaron sus valores como el desvío porcentual de cada una con respecto a la mejor solución encontrada respectivamente; éstas variables reciben el nombre de **PDT = PercentageDT** y **PC = PercentageCenter**. Para **SumForeignPointsHull**<sup>3</sup>, dado que el mínimo es 0, se consideró **RPH = RelativePointsHull = SumForeignPointsHull / (número de puntos insertados)**. La tabla 5.11 muestra los valores de cada variable que se asociaron a

---

<sup>3</sup>Se utiliza el Modo de Agregación *Suma* en vez de *Promedio* pues, por un lado, como todas las soluciones en consideración usan el mismo número de vehículos, esta transformación es simplemente un cambio de escala, y por otro lado, esta medida más sencilla de interpretar.

cada nivel.

| Nivel | PercentageDT         | RelativePointsHull  | PercentageCenter     |
|-------|----------------------|---------------------|----------------------|
| 1     | [0 % – 2,5 %]        | [0 – 0,1]           | [0 % – 10 %]         |
| 2     | (2,5 % – 5 %]        | (0,1 – 0,2]         | (10 % – 20 %]        |
| 3     | (5 % – 10 %]         | (0,2 – 0,5]         | (20 % – 35 %]        |
| 4     | (10 % – <i>max</i> ] | (0,5 – <i>max</i> ] | (35 % – <i>max</i> ] |

Tabla 5.11: Valores pertenecientes a cada nivel en los distintos atributos después de la discretización. Los porcentajes asociados a **PercentageDT** y **PercentageCenter** corresponden al desvío porcentual de **DT** y **AvgDistCenter** con respecto a la mejor solución encontrada para cada una. Los valores asociados a **RelativePointsHull** corresponden a **SumForeignPointsHull** / (número de puntos insertados).

Por último, para generar la encuesta es necesario definir qué soluciones se van a mostrar en ésta (es decir, qué combinaciones de niveles en cada uno de los 3 atributos van a tener las soluciones que se van a enseñar), y en qué pares se van a asociar para definir cada una de las preguntas. Debido a que preguntar por todos los pares posibles de todas las combinaciones de los distintos niveles de cada atributo (es decir, utilizar un diseño factorial completo) sería prohibitivamente largo, se utiliza sólo un subconjunto de éstos basado en las llamadas tablas o arreglos ortogonales [32]. Estas tablas son matrices que tienen una columna por cada atributo que se está evaluando, cuyas entradas corresponden a los distintos niveles a considerar para cada atributo, y que cumplen la propiedad de que para cada submatriz con todas las filas y  $t$  columnas (con  $0 \leq t \leq k$ , donde  $k$  es el número de columnas), las filas contienen todas las  $t$ -tuplas posibles a formarse con los distintos niveles de cada atributo el mismo número de veces (la tabla 5.12 muestra el arreglo ortogonal finalmente utilizado). Si se interpreta cada fila de una éstas tablas como un objeto a ser incluido en la encuesta (es decir, la combinación de niveles de cada atributo que deben de tener las soluciones seleccionadas), entonces utilizar este tipo de tablas permite reducir considerablemente el número de preguntas de la encuesta, minimizando la cantidad de información perdida por dejar ciertas combinaciones fuera del análisis (se mantiene suficiente información como para estimar los efectos principales de cada atributo, pero no para estudiar las interacciones entre éstos).

Ahora, para este caso en particular, es difícil (e incluso podría no ser factible) generar algunas de las combinaciones de niveles de atributos que indican las tablas ortogonales, pues las 3 variables en consideración están fuertemente correlacionadas. De hecho, de todos los arreglos ortogonales posibles para 3 variables y 4 niveles por variable, el que menos combinaciones

| Solución | PDT | RPH | PC | Generada |
|----------|-----|-----|----|----------|
| 1        | 1   | 1   | 1  | Sí       |
| 2        | 1   | 2   | 4  | Sí       |
| 3        | 1   | 3   | 3  | Sí       |
| 4        | 1   | 4   | 2  | No       |
| 5        | 2   | 1   | 4  | Sí       |
| 6        | 2   | 2   | 1  | Sí       |
| 7        | 2   | 3   | 2  | Sí       |
| 8        | 2   | 4   | 3  | Sí       |
| 9        | 3   | 1   | 2  | Sí       |
| 10       | 3   | 2   | 3  | Sí       |
| 11       | 3   | 3   | 4  | Sí       |
| 12       | 3   | 4   | 1  | No       |
| 13       | 4   | 1   | 3  | Sí       |
| 14       | 4   | 2   | 2  | Sí       |
| 15       | 4   | 3   | 1  | Sí       |
| 16       | 4   | 4   | 4  | Sí       |

Tabla 5.12: Arreglo ortogonal utilizado como base para escoger las características de las soluciones a incluir. La última columna indica si la combinación de niveles fue posible de generar o no.

tiene que no están presentes en el conjunto de soluciones generadas es el presentado en la tabla 5.12, por lo que éste es el que se utiliza para construir la encuesta (las dos filas de la tabla que corresponden a soluciones que no fue posible construir, simplemente no se consideran). En los casos en los que había varias soluciones con los niveles de alguna de las filas de la tabla, simplemente se escogió una de éstas al azar. Una vez definido este conjunto de 14 soluciones, se ordenó aleatoriamente y se utilizó un diseño cíclico [35] para elegir los pares de soluciones de cada pregunta (es decir,  $\forall i \in [1, n - 1]$  la solución en la posición  $i$  de la lista ordenada aleatoriamente se empareja con la solución que quedó en la posición  $i + 1$ , y la última solución de la lista se empareja con la primera).

Una vez creada la encuesta, se les pidió a 10 empleados de SimpliRoute que la respondieran y que entregaran *feedback* sobre las instrucciones y el diseño general de ésta. Se utilizó esta información para mejorar el formato de las preguntas y de las instrucciones, y luego se analizaron los resultados obtenidos, para evaluar si la discretización y el hecho de no tener el 100% de las soluciones del arreglo ortogonal podía causar algún problema en el análisis. Dada la naturaleza de las variables en consideración, se esperaba que la utilidad de los niveles dentro de cada variable fuese monótonamente no creciente, lo cual ocurrió sólo para **PDT**,

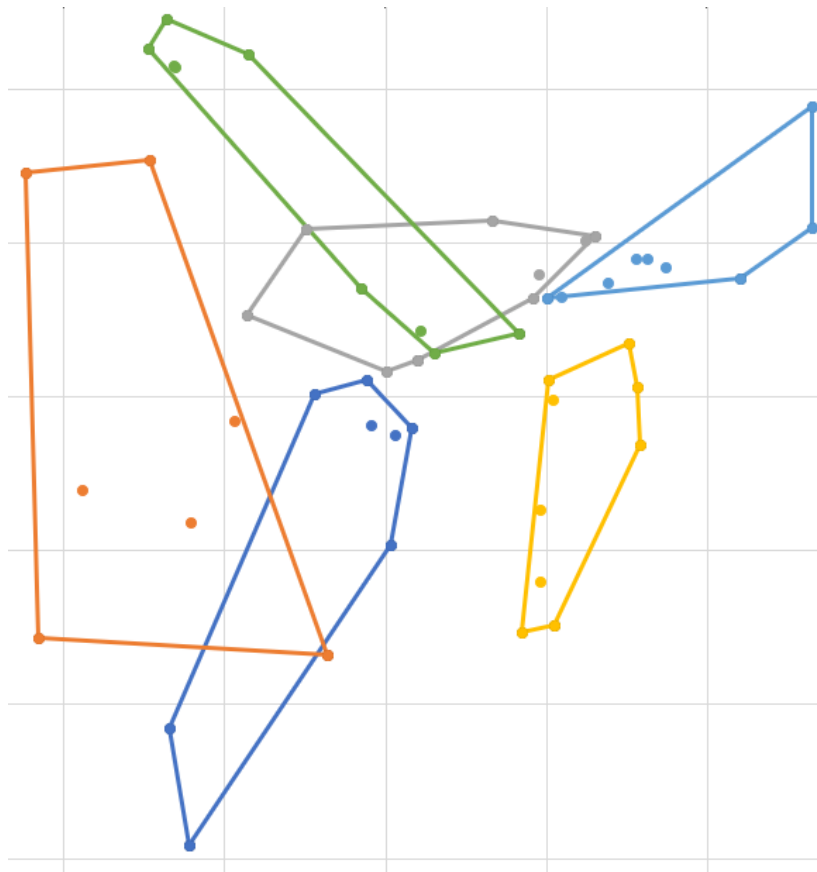


Figura 5.2: Puntos y envolturas convexas de cada ruta, en una de las soluciones escogidas inicialmente para aparecer en la encuesta. Si bien esta solución tiene nivel 1 para **RPH**, a pesar de tener pocos puntos dentro de las envolturas convexas de rutas que no son la suya propia, las superficies de éstas sí se intersectan considerablemente.

pero no para las otras dos medidas; en ambas el nivel 3 tuvo mayor valor que el nivel 2. Para entender de mejor forma este fenómeno, y evaluar si se debía a algún error o si ésta era la verdadera forma de la función de utilidad de los encuestados, se seleccionó a las 2 personas que tenían utilidades individuales menos monótonas, y se les pidió que volvieran a responder la encuesta, pero explicando en voz alta qué pensaban de cada solución, y por qué tomaban la decisión de elegir una por sobre otra en cada pregunta. Mediante estas entrevistas se pudo identificar que una de las personas respondía dándole más relevancia a los cruces intra-ruta que a los inter-ruta. A pesar de que este atributo podría ser relevante en las decisiones de los planificadores, como se están evaluando otras medidas, se tomó la decisión de controlar por este factor, y cambiar las soluciones de los pares que tenían cantidades diferentes de cruces intra-ruta. Por otro lado, las dos personas entrevistadas mencionaron en la misma pregunta, que ésta era difícil de responder, porque las soluciones enseñadas eran muy parecidas, pero

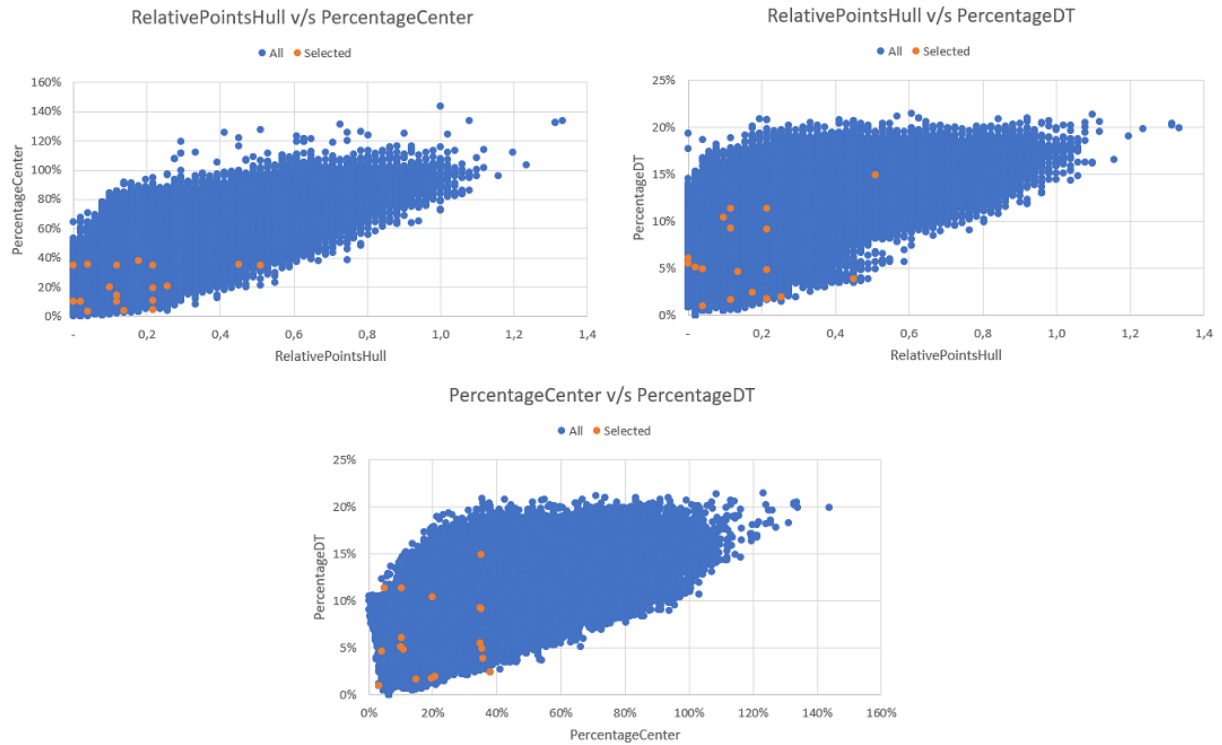


Figura 5.3: **RPH vs PC**, **RPH vs PDT** y **PC vs PDT** para instancia elegida. Los puntos naranjos corresponden a las soluciones incluidas en las preguntas de la encuesta.

esta percepción no representaba los niveles que tenían las dos soluciones en cada atributo, que eran bastante distintos ( $(\text{PDT}, \text{RPH}, \text{PC}) = (2, 4, 3)$  y  $(3, 1, 2)$ ); además, tal pregunta resultó ser la que mayor desacuerdo generaba entre todos los que respondieron la encuesta, con la mitad prefiriendo cada una de las dos soluciones. Por esto, se estudió el caso en detalle y se identificó que si bien las soluciones tenían nivel 4 y 1 para **RPH**, en ambos casos estaban muy cerca del punto de corte de la discretización (es decir, casi podrían haber sido 3 y 2 respectivamente). Además, estudiando visualmente ambas soluciones, se identificó una posible causa de confusión adicional: a pesar de que la solución con nivel 1 de **RPH**, tenía pocos puntos dentro de las envolturas convexas de rutas que no fuesen la suya propia, estas envolturas convexas sí se intersectaban bastante entre sí, como se puede ver en la figura 5.2. Por esta razón, se decidió cambiar esta solución por una que no presentara este problema. Todo este análisis llevó a sustituir las soluciones representantes de 4 de las 14 filas del arreglo ortogonal. Además, para asegurarse de que se está identificando correctamente la utilidad asociada a los niveles 2 y 3 de **RPH** y **PC**, se agregaron 2 preguntas adicionales formadas por 4 nuevas soluciones (con niveles  $(1,3,2)$  y  $(1,2,2)$ , y  $(3,1,2)$  y  $(3,1,3)$ ), específicamente

seleccionadas para diferenciar entre estos dos niveles de estos dos atributos. Luego de hacer estos cambios, se les pidió a las mismas 10 personas que respondieran la nueva versión de la encuesta, y en esta ocasión los resultados sí fueron satisfactorios.

Es importante mencionar que la correlación entre las 3 variables en consideración es mucho menor en el subconjunto de las 18 soluciones seleccionadas para la encuesta que en el conjunto total de soluciones generadas<sup>4</sup>, y que los valores en este último caso son similares a los reportados en el capítulo 4. Esto se puede ver gráficamente en las figuras en 5.3, que muestran gráficos de dispersión de los 3 pares de variables, distinguiendo especialmente las 18 soluciones incorporadas.

La encuesta se distribuyó a los planificadores de rutas de distintos clientes de SimpliRoute mediante un link que se les envió vía correo electrónico. Para incentivarlos a participar, se incluyó en el correo una sección que indicaba que todos los que respondieran la encuesta completamente participarían en un sorteo de 3 *gift cards* de una tienda de *retail* con un valor cercano a 15 dólares estadounidenses cada una.

## 5.2.2. Resultados

Del total de 100 correos enviados, 45 ingresaron al enlace de la encuesta, y 31 respondieron el 100 % de las preguntas. Los 14 casos que no terminaron de responder no fueron incluidos en el análisis.

El análisis de los datos se realizó usando el entorno de programación  $R$  [66] en su versión 3.5.1, y el paquete *bayesm* [70]. En particular, se utilizó la función *rHierMnlRwMixture* de dicho paquete, que hace uso de un algoritmo de Monte Carlo basado en Cadenas de Markov para estimar una regresión logística multinomial de coeficientes aleatorios bajo el paradigma jerárquico bayesiano. Se generaron  $R = 2.000.000$  muestras de la cadena, de las que se guardó sólo 1 de cada 5 de éstas (es decir, se definió *keep* = 5). Se utilizó el prior establecido por defecto, simplemente indicando el número de clases latentes a considerar. A pesar de que se evaluó la posibilidad de que existiese más de una clase latente en los encuestados, los

---

<sup>4</sup>En el conjunto total de soluciones la correlación para cada par es PDT-RPH = 0,62; PDT-PC = 0,65; PC-RPH = 0,76. Entre las 18 instancias elegidas los valores son PDT-RPH = 0,28; PDT-PC = 0,13; PC-RPH = 0,35.

resultados en estos casos generaban sólo una clase predominante y la(s) otra(s) mucho más pequeña(s), y además esto generaba que las utilidades volvieran a ser no monótonas, por lo que se decidió sólo considerar el caso de que todos los encuestados provienen de una misma clase latente.

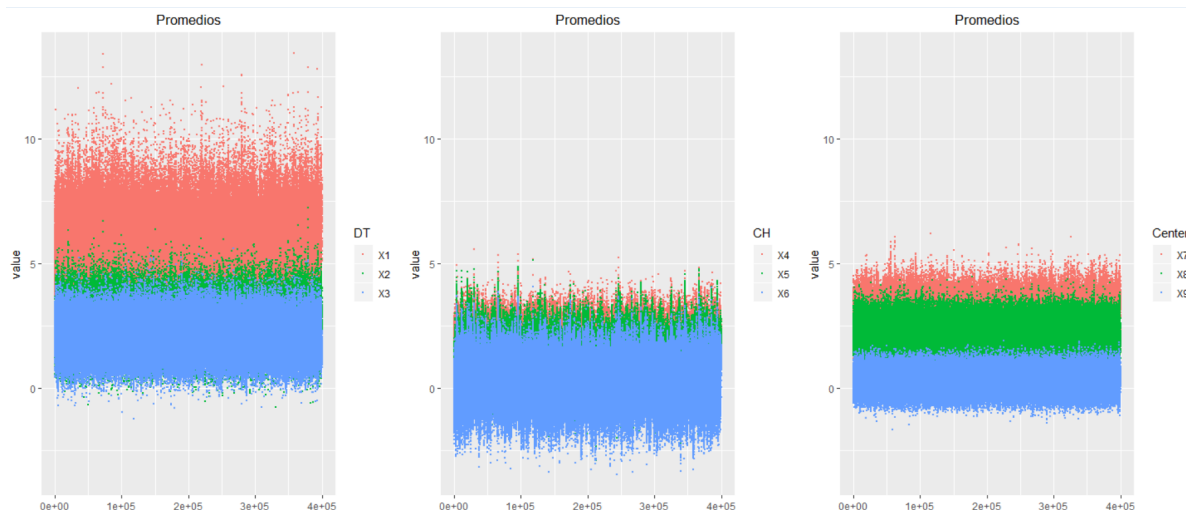


Figura 5.4: Valores de las medias ( $\mu_i$ ) de las variables aleatorias normales asociadas a la utilidad de cada nivel (sólo se muestran 3 niveles por atributo, pues el cuarto está fijado en 0) para los tres atributos en consideración, a través de la cadena.

Como se puede ver en la figura 5.4, que muestra los valores de las medias ( $\mu_i$ ) de las variables aleatorias normales asociadas a la utilidad de cada nivel para los tres atributos en consideración a través de la cadena, ésta converge bastante rápidamente, sin mostrar tendencias notorias en ninguna de las variables mostradas. Por otro lado, al estudiar los valores de la log-verosimilitud a través de la cadena tampoco se aprecian tendencias. Así, dada la rápida convergencia de ésta, se descarta sólo el primer 10% de los datos para calcular los distintos estimadores.

Para calcular las utilidades parciales (*part-worths*) de cada nivel de cada atributo, se considera la mediana de los valores de las medias ( $\mu_i$ ) a través de la cadena; estos valores se presentan en la figura 5.5. Se puede ver que las tres variables tienen utilidades decrecientes, con **PDT** siendo la que valores más altos alcanza. Pareciera no haber una diferencia importante entre los niveles 2 y 3 de **PDT**, entre los niveles 3 y 4 de **RPH**, ni entre los niveles 1 y 2, ni 3 y 4 de **PC**. Esto se confirma al estudiar la distribución *a posteriori* de estas variables, que se muestra en la figura 5.6 (y más en detalle en la tabla 5.13) junto con la media  $a$



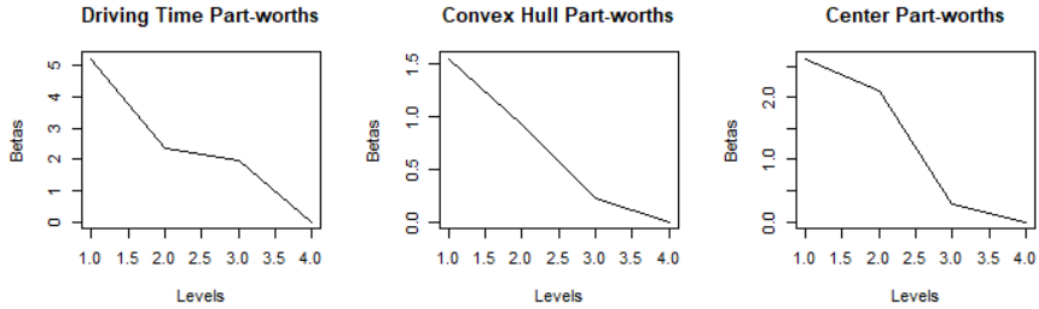


Figura 5.5: Utilidades parciales (*part-worths*) de cada nivel para los tres atributos en consideración.

*posteriori*, el intervalo de  $\pm 2$  errores estándar numéricos en torno a ésta, y los Intervalos de Credibilidad Bayesiana (ICB) al 95 % y 99 %. Además, se puede ver que para el nivel 2 de **RPH**, el 0 está dentro del ICB al 95 %, por lo que tampoco es posible asegurar que este nivel sea preferible sobre los niveles 3 o 4. Así, aunque el nivel 1 de **RPH** tenga una utilidad parcial más baja que las otras dos variables, pareciera que los encuestados son altamente sensibles soluciones peores que las de nivel 1 para esta variable (es decir, soluciones para las que el número de puntos dentro de envolturas convexas de otras rutas es mayor que el 10 % del total de puntos de la instancia). Por otro lado, el hecho de que no se vean diferencias entre los niveles 2 y 3 de **PDT** indica que cuando no es posible estar cerca del mínimo (es decir, a menos de 2,5 % de éste), entonces los planificadores no percibirían mucha diferencia entre estar a 5 % del mínimo que estar a 10 % de éste (aunque a mayor distancia que esa sí se valoraría negativamente el impacto de tal alza). Por último, con respecto a **PC** pareciera que esta variable sólo comienza a influir negativamente cuando sus valores están a más del 20 % del mínimo.

Si nos enfocamos en la diferencia entre la utilidad máxima y mínima asociada a los distintos niveles de cada atributo, se puede obtener una medida de la importancia relativa que tiene cada uno de éstos. En la figura 5.7 se ve que aunque la función objetivo tradicional del *VRP* es la que más relevancia tiene, las dos medidas de *VA* combinadas tienen casi el mismo peso que ésta al momento de decidir qué solución prefieren los encuestados.

Para poder llegar a un resultado que permita establecer una equivalencia para cambios en **DT**, **ForeignPointsHull** y **DistCenter**, se buscará la cantidad de cambio en **PDT**, **RPH** y **PC** que genera un aumento de 1 punto de utilidad, y luego se retornaran los valores de

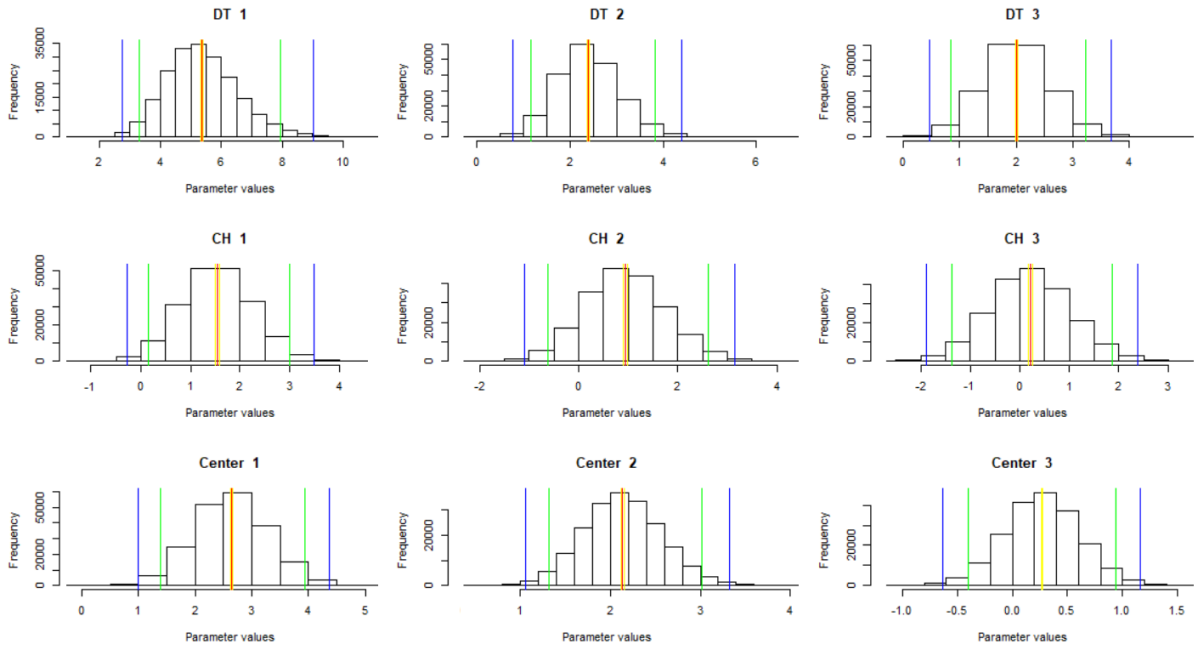


Figura 5.6: Distribución *a posteriori* de las medias ( $\mu_i$ ) para los niveles 1, 2 y 3 de cada atributo (el nivel 4 está fijo en 0 en las 3 variables), indicando con la media *a posteriori* (en rojo), el intervalo de  $\pm 2$  errores estándar numéricos en torno a ésta (en amarillo), y los intervalos de credibilidad bayesiana al 95% y 99% (en verde y azul respectivamente).

estas variables a sus unidades originales. Para esto, primero se debe expresar la relación entre las variables y la utilidad en términos que no dependan de la discretización realizada. A pesar de que la función de utilidad para cada atributo no es lineal en los niveles de éstos, y de que el valor que toman las variables en los distintos niveles tampoco lo es, las aproximaciones lineales de éstas incurren en poco error, y permiten entregar un valor estándar que no depende del nivel en el que se encuentra la solución. Así, para cada atributo se considera la línea que atraviesa los puntos  $(1, u_1)$  y  $(4, u_4) = (4, 0)$ , es decir, la que pasa por los puntos de utilidad más alta y baja; las pendientes de éstas son  $(m_{u_{PDT}}, m_{u_{RPH}}, m_{u_{PC}}) = (-1.74, -0.51, -0.87)$ , o sea, subir un nivel para **RPH** hace disminuir la utilidad cerca de medio punto. Por otro lado, como los niveles están definidos como intervalos, para definir un valor representante de cada nivel se considera el promedio de esta variable dentro de las soluciones (que se mostraron en la encuesta) que pertenecen a ese nivel para tal variable. Luego se calcula la pendiente de esta curva de igual forma que para las utilidades, y se obtienen los siguientes valores  $(m_{PDT}, m_{RPH}, m_{PC}) = (3.4\%, 0.16, 10.6\%)$ , es decir, subir un nivel en **PDT** eleva el valor de tal variable cerca de 3.4%. Ahora, para obtener el resultado buscado basta con dividir los valores de estas pendientes, es decir calcular  $\frac{m_x}{m_{u_x}}$  para las tres variables,

|          | Media    |       |        | Desv. Est. |       |        |
|----------|----------|-------|--------|------------|-------|--------|
|          | Promedio | 2,5 % | 97,5 % | Promedio   | 2,5 % | 97,5 % |
| DT 1     | 5,40     | 3,34  | 7,94   | 4,23       | 2,45  | 6,69   |
| DT 2     | 2,41     | 1,17  | 3,83   | 2,70       | 1,68  | 4,10   |
| DT 3     | 2,00     | 0,84  | 3,21   | 1,87       | 1,07  | 3,10   |
| CH 1     | 1,56     | 0,15  | 3,01   | 1,86       | 1,05  | 3,10   |
| CH 2     | 0,97     | -0,61 | 2,63   | 1,38       | 0,86  | 2,32   |
| CH 3     | 0,24     | -1,37 | 1,88   | 1,55       | 0,92  | 2,65   |
| Center 1 | 2,65     | 1,42  | 3,94   | 1,63       | 0,97  | 2,64   |
| Center 2 | 2,14     | 1,31  | 3,02   | 1,32       | 0,87  | 1,97   |
| Center 3 | 0,27     | -0,40 | 0,94   | 1,10       | 0,77  | 1,57   |

Tabla 5.13: Promedio y percentiles 2,5 % y 97,5 % de las distribuciones a posteriori de la Media y la Desviación Estándar de cada parámetro del modelo.

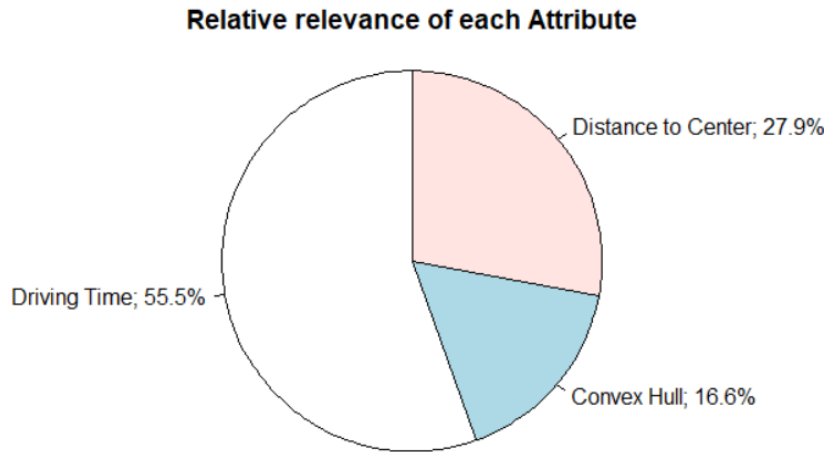


Figura 5.7: Importancia relativa de cada atributo.

lo cual muestra que para elevar la utilidad 1 punto es necesario disminuir cada variable en  $(\mathbf{PDT}, \mathbf{RPH}, \mathbf{PC}) = (2.0\%, 0.31, 12.1\%)$ . Finalmente, para expresar este resultado en términos de los valores originales de las tres medidas consideradas, basta con multiplicar los valores asociados a  $\mathbf{PDT}$  y  $\mathbf{PC}$  por sus respectivos mínimos y luego sumarles tal valor, y al asociado a  $\mathbf{RPH}$  multiplicarlo por el número de puntos de la instancia; esto resulta en  $(\mathbf{DT}, \mathbf{SumForeignPointsHull}, \mathbf{AvgDistCenter}) = (13.7, 15.8, 2.32)$ . Así, como para cada variable una baja en tales magnitudes genera el mismo aumento en términos de utilidad, esto quiere decir que mejorar el objetivo tradicional de la solución en 1 kilómetro, es percibido como equivalente a bajar en  $\frac{15.8}{13.7} = 1.15$  el número de puntos dentro de envolturas convexas de rutas que no son la suya propia, y a disminuir en  $\frac{2.32}{13.7} = 0.17$  kilómetros la distancia promedio de los puntos de cada ruta hacia su centro.

Es importante notar que debido a que sólo se evaluaron soluciones de la región más cercana a los mínimos de cada variable, es posible que en soluciones con valores más grandes, la percepción de alguna de estas medidas sea más negativa aun.

# Capítulo 6

## Conclusiones y trabajos futuros

El presente trabajo describe un estudio de las medidas de *atractivo visual* en la evaluación de la calidad de las soluciones del *VRP*, específicamente en el contexto del *CVRS SimpliRoute*. Se realizaron entrevistas con clientes de la empresa para desarrollar hipótesis sobre las variables a considerar, y luego se estudiaron y clasificaron las distintas medidas utilizadas en la literatura para cuantificar este fenómeno. Se evaluó la relación que éstas tienen entre sí mismas, y con los objetivos tradicionales del *VRP*, y finalmente se implementaron dos métodos para medir la importancia que tiene el incluir estos factores para los usuarios de la plataforma.

Las entrevistas descritas en el capítulo 3 levantaron temas que generaban insatisfacción, asociados en su mayoría a la preferencia por soluciones más compactas y zonificadas. Estas preferencias son similares a las descritas en la literatura, y su relevancia fue confirmada empíricamente en el capítulo 5, en el cual se encontró que las medidas de *belleza* de “distancia de los puntos de cada ruta hacia su centro” y “cantidad de puntos dentro de las envolturas convexas de otras rutas” en conjunto, son casi igual de importantes para los usuarios que la medida de “distancia total recorrida”. Este resultado es un avance importante en el entendimiento del rol que juega el *atractivo visual* en la percepción de calidad de las soluciones del *VRP*, y en la posibilidad de cuantificarlo con variables objetivas. Además, responde a la inquietud planteada por Lum et al. [51] sobre la necesidad de confirmar que las métricas definidas para evaluar este fenómeno realmente se correspondan con las evaluaciones subjetivas. La metodología utilizada la sección 5.2 también permitió estimar que la disminución en 1 kilómetro de la distancia total recorrida en una solución, es percibida como equivalente a una

reducción de 1.15 en el número de puntos dentro de envolturas convexas de rutas que no son la suya propia, y a una baja de 0.17 kilómetros en la distancia promedio de los puntos de cada ruta hacia su centro. Estos valores deben de ser interpretados cuidadosamente, pues provienen de la evaluación de soluciones de una sola instancia por un grupo reducido de personas, por lo que podrían ser distintos si se evalúan en condiciones diferentes.

Por otro lado, el análisis realizado en el capítulo 4 mostró que las medidas de *VA* suelen tener correlaciones entre medianas y altas con la distancia total recorrida, alcanzando valores de hasta 0.91, lo cual indica que las soluciones que son percibidas como “visualmente atractivas” tienden a ser mejores soluciones que las que son calificadas como “más feas”. Este resultado llena el vacío reportado por Lu y Dessouky [50] sobre la no existencia de trabajos que evalúen la relación entre las medidas de *belleza inter-ruta* y la distancia total de las soluciones, mostrando que éstas alcanzan correlaciones aún más altas que las previamente consideradas. Al estudiar la frontera de Pareto entre las medidas de *VA* y los objetivos tradicionales, se encontró que ésta tiene comportamientos distintos para las instancias de tipo C que para las de tipo R y RC (exhibiendo un mayor *trade-off* en las instancias de estos dos últimos tipos). A pesar de estas diferencias, se pudo ver que en la mayoría de los casos evaluados, o bien la mejor solución conocida en términos de distancia total recorrida también tenía buenos niveles de *belleza*, o bien era posible encontrar soluciones con resultados bastante mejores en términos de *VA* perdiendo poco en el objetivo tradicional.

Estos resultados en conjunto muestran que por un lado, mantener el foco del estudio del *VRP* sólo en el tiempo o distancia total recorridos implica que se están desarrollando modelos matemáticos y algoritmos computacionales complejos para encontrar soluciones que sólo capturan la mitad de lo que los planificadores consideran a la hora de decidir aceptar o no una solución, y que por otro lado, entregar soluciones con mejores niveles de *belleza* sería en general simultáneamente posible y beneficioso en términos de la mejora en la calidad percibida por los usuarios.

Una extensión natural al trabajo realizado en esta tesis sería el desarrollo de un algoritmo que incluya dentro de sus objetivos alguna(s) medida(s) de *VA*. También sería interesante evaluar formalmente la posible estructura factorial de las distintas medidas de *belleza* comentada en el capítulo 4, con métodos como Análisis de Componentes Principales o Análisis Factorial

Exploratorio, para entender en más profundidad el fenómeno, y para reducir el número de variables relevantes. Por otro lado, un punto importante a considerar para los *CVR*S es el cómo influyen la gran cantidad de restricciones adicionales que estos suelen manejar (como ventanas de tiempo, asociaciones vehículo-visita, heterogeneidad de los vehículos, etc.) en los resultados de los capítulos 4 y 5.

La investigación de operaciones, y en particular el ruteo de vehículos, son ejemplos ilustres de cómo la investigación en modelos matemáticos y métodos computacionales pueden beneficiar a las empresas y organizaciones. Sin embargo, para esta relación continúe siendo beneficiosa, será necesario que la academia considere integralmente dentro de sus soluciones, el hecho de que los problemas que aborda involucran inevitablemente a seres humanos, ya sea porque sus resultados influyen en el quehacer de estos, o porque dependen de estos para su implementación.

# Bibliografía

- [1] M. Balinski and R. Quandt. On an integer program for a delivery problem. *Mathematical Programming*, 12(2):300–304, 1964.
- [2] M. Battarra, G. Erdoğan, and D. Vigo. Exact algorithms for the clustered vehicle routing problem. *Operations Research*, 62(1):58–/71, 2014.
- [3] Kai Helge Becker. An outlook on behavioural or – three tasks, three pitfalls, one definition. *European Journal of Operational Research*, 249(3):806–815, 2016.
- [4] E. Benavent, Á. Corberán, and J.M. Sanchis. A metaheuristic for the min–max windy rural postman problem with k vehicles. *Computational Management Science*, 7(3):269–287, Jul 2010.
- [5] C. Blyth. On simpson’s paradox and the sure-thing principle. *Journal of the American Statistical Association*, 67(338):364–366, 1972.
- [6] A. Bosch. Seed selection in a multi-period planning with time windows. Master’s thesis, University of Twente, 7522 NB Enschede, Netherlands, 2 2014.
- [7] O. Bräysy and G. Hasle. Chapter 12: Software tools and emerging technologies for vehicle routing and intermodal transportation. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 351–380. SIAM, 2014.
- [8] L. Caccalano. Roteirização de veículos para o abastecimento de linhas de produção. Master’s thesis, Escola Politécnica, Universidad de São Paulo, São Paulo, Brasil, 5 2012.
- [9] L. Caccalano and C. Barbieri da Cunha. Roteirização de veículos para o abastecimento de linhas de produção. *Gestão & Produção*, 22:846 – 860, 12 2015.



- [10] N. Christofides and S. Eilon. An algorithm for the vehicle dispatching problem. *Operational Research Quarterly*, 20(3):309–318, 1969.
- [11] G. Clarke and J.W. Wright. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12:568–581, 1964.
- [12] L. C. Coelho, J. Renaud, and G. Laporte. Road-based goods transportation: a survey of real-world logistics applications from 2000 to 2015. *INFOR: Information Systems and Operational Research*, 54(2):79–96, 2016.
- [13] M. Constantino, L. Gouveia, M.C. Mourão, and A.C. Nunes. The mixed capacitated arc routing problem with non-overlapping routes. *European Journal of Operational Research*, 244(2):445–456, 2015.
- [14] J.-F. Cordeau, G. Laporte, and A. Mercier. A unified tabu search heuristic for vehicle routing problems with time windows. *Journal of the Operational Research Society*, 52:928–936, 2001.
- [15] G.B. Dantzig and R.H. Ramser. The truck dispatching problem. *Management Science*, 6(1):80–91, 1959.
- [16] A. David and P.R. Glore. The impact of design and aesthetics on usability, credibility, and learning in an online environment. *Online Journal of Distance Learning Administration*, 13(4), 2010.
- [17] M. Drexler. Rich vehicle routing in theory and practice. *Logistics Research*, 5:47–63, 2012.
- [18] M. Fisher. Optimal solution of vehicle routing problems using minimum k-trees. *Operations Research*, 42(4):626–642, 1994.
- [19] M.L. Fisher and R. Jaikumar. A generalized assignment heuristic for vehicle routing. *Networks*, 11(2):109–124, 1981.
- [20] R. Fukasawa, H. Longo, J. Lysgaard, M. Poggi de Aragão, M. Reis, E. Uchoa, and R. F. Werneck. Robust branch-and-cut-and-price for the capacitated vehicle routing problem. *Mathematical Programming*, 106(3):491–511, 2006.

- [21] H. Gehring and J. Homberger. A parallel hybrid evolutionary metaheuristic for the vehicle routing problem with time windows. In K. Miettinen, M. Makela, and J. Toivanen, editors, *Proceeding of EUROGEN99 - Short Course on Evolutionary Algorithms in Engineering and Computer Science*, pages 57–64. University of Jyväskylä, Enero 1999.
- [22] B.E. Gillett and L.R. Miller. A heuristic algorithm for the vehicle-dispatch problem. *Operations Research*, 22(2):340–349, 1974.
- [23] P. Green and V. Rao. Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, 8(3):355–363, 1971.
- [24] C. Gretton and P. Kilby. A study of shape penalties in vehicle routing. In *Proceedings of Eighth Triennial Symposium on Transportation Analysis*, Jun 2013.
- [25] C. Groër, B.L. Golden, and E.A. Wasil. A parallel algorithm for the vehicle routing problem. *INFORMS Journal on Computing*, 23:315–330, 2011.
- [26] R. Hall and J. Partyka. On the road to mobility. *OR/MS Today*, 35(1), Febrero 2008.
- [27] R. Hall and J. Partyka. On the road to connectivity. *OR/MS Today*, 37(1), Febrero 2010.
- [28] R. Hall and J. Partyka. Vehicle routing software survey: VR delivers the goods. *OR/MS Today*, 41(1):40–46, Febrero 2014.
- [29] G. Hasle and O. Kloster. *Industrial Vehicle Routing*, pages 397–435. 01 2007.
- [30] G. Hasle, O. Kloster, and M. Smedsrud. Aspects of routing problems in media product distribution. Invited talk at the 24th European Conference on Operational Research (EURO XXIV), Lisbon, Portugal, Jul 2010.
- [31] G. Hasle, O. Kloster, and M. Smedsrud. Vehicle routing problems in media product distribution. Talk at the 4th Nordic Optimization Symposium (4th NOS), Aarhus, Denmark, Oct 2010.
- [32] A. S. Hedayat, N. J. A. Sloane, and J. Stufken. *Orthogonal Arrays: Theory and Applications*. Springer Series in Statistics. Springer New York, 1999.

- [33] B. L. Hollis and P. J. Green. Real-life vehicle routing with time windows for visual attractiveness and operational robustness. *Asia-Pacific Journal of Operational Research*, 29(04):1250017, 2012.
- [34] W. Jang, H. H. Lim, T. J. Crowe, G. Raskin, and T. E. Perkins. The missouri lottery optimizes its scheduling and routing to improve efficiency and balance. *Interfaces*, 36(4):302–313, July 2006.
- [35] J. A. John. Factorial experiments in cyclic designs. *The Annals of Statistics*, 1(1):188–194, 1973.
- [36] A.A. Juan, F. Adelantado, S.E. Grasman, J. Faulin, and J.R. Montoya-Torres. Solving the capacitated vehicle routing problem with maximum traveling distance and service time requirements: An approach based on monte carlo simulation. In *Winter Simulation Conference*, WSC '09, pages 2467–2475. Winter Simulation Conference, 2009.
- [37] A.A. Juan, J. Faulin, J. Jorba, D. Riera, D. Masip, and B. Barrios. On the use of monte carlo simulation, cache and splitting techniques to improve the clarke and wright savings heuristics. *Journal of the Operational Research Society*, 62(6):1085–1097, 2011.
- [38] A.A. Juan, J. Faulin, R. Ruiz, B. Barrios, and S. Caballé. The SR-GCWS hybrid algorithm for solving the capacitated vehicle routing problem. *Applied Soft Computing*, 10(1):215–224, 2010.
- [39] A.A. Juan, J. Faulin, R. Ruiz, B. Barrios, M. Gilibert, and X. Vilajosana. Using oriented random search to provide a set of alternative solutions to the capacitated vehicle routing problem. In J.W. Chinneck, B. Kristjansson, and M.J. Saltzman, editors, *Operations Research and Cyber-Infrastructure*, pages 331–345, Boston, MA, 2009. Springer US.
- [40] T. Kallio. Why we choose the more attractive looking objects: Somatic markers and somaesthetics in user experience. In *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces*, DPPI '03, pages 142–143, New York, NY, USA, 2003. ACM.
- [41] G. Kant, M. Jacks, and C. Aantjes. Coca-cola enterprises optimizes vehicle routes for efficient product delivery. *Interfaces*, 38(1):40–50, 2008.

- [42] P. Kilby, P. Prosser, and P. Shaw. Guided local search for the vehicle routing problem with time windows. In S. Voß, S. Martello, I.H. Osman, and C. Roucairol, editors, *Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization*, pages 473–486. Springer US, Boston, MA, 1999.
- [43] B. Kim, S. Kim, and S. Sahoo. Balanced clustering algorithms for improving shapes on vehicle routing problems. In *IIE Annual Conference and Exhibition 2004*, May 2004.
- [44] B. Kim, S. Kim, and S. Sahoo. Waste collection vehicle routing problem with time windows. *Computers & Operations Research*, 224:3624–3642, 2006.
- [45] G. Laporte, Y. Nobert, and M. Desrochers. Optimal routing under capacity and distance restrictions. *Operations Research*, 33(5):1050–1073, 1985.
- [46] G. Laporte, F. Ropke, and T. Vidal. Chapter 4: Heuristics for the vehicle routing problem. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 87–116. SIAM, 2014.
- [47] H. Li and A. Lim. A metaheuristic for the pickup and delivery problem with time windows. In *Proceedings 13th IEEE International Conference on Tools with Artificial Intelligence. ICTAI 2001*, pages 160–167, Nov 2001.
- [48] S. Lin. Computer solutions of the traveling salesman problem. *Bell System Technical Journal*, 44:2245–2269, 1965.
- [49] Y. Lin and W.J. Zhang. Integrated design of function, usability, and aesthetics for automobile interiors: State of the art, challenges, and solutions. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 220(8):697–708, 2006.
- [50] Q. Lu and M.M. Dessouky. A new insertion-based construction heuristic for solving the pickup and delivery problem with time windows. *European Journal of Operational Research*, 175(2):672–687, 2006.
- [51] O. Lum, C. Cerrone, B. Golden, and E. Wasil. Partitioning a street network into compact, balanced, and visually appealing routes. *Networks*, 69(3):290–303, May 2017.

- [52] J. Lysgaard, A.N. Letchford, and R.W. Eglese. A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming*, 100(2):423–445, 2004.
- [53] J. N. MacGregor and T. Ormerod. Human performance on the traveling salesman problem. *Perception & Psychophysics*, 58(4):527–539, Jun 1996.
- [54] V. Maniezzo, T. Stützle, and S. Voss. Matheuristics: Hybridizing metaheuristics and mathematical programming. 2009.
- [55] P. Matis. Decision support system for solving the street routing problem. *Transport*, 23(3):230–235, 2008.
- [56] P. Matis and M. Koháni. Very large street routing problem with mixed transportation mode. *Central European Journal of Operations Research*, 19(3):359–369, Sep 2011.
- [57] P. Matl, R. F. Hartl, and T. Vidal. Workload equity in vehicle routing problems: A survey and analysis. *Transportation Science*, 52(2):239–260, 2018.
- [58] C. Miller. Aesthetics and e-assessment: the interplay of emotional design and learner performance. *Distance Education*, 32(3):307–337, 2011.
- [59] Y. Nagata and O. Bräysy. Edge assembly-based memetic algorithm for the capacitated vehicle routing problem. *Networks*, 54:205–215, 2009.
- [60] Ibrahim H. Osman and Nicos Christofides. Capacitated clustering problems by hybrid simulated annealing and tabu search. *International Transactions in Operational Research*, 1(3):317 – 336, 1994.
- [61] I.H. Osman. Metastrategy simulated annealing and tabu search algorithms for the vehicle routing problem. *Annals of Operations Research*, 41:421–451, 1993.
- [62] A. Parchitelli, F. Nocera, G. Iacobellis, M. Mongiello, T. Di Noia, and E. Di Sciascio. A pre-process clustering methods for the waste collection problem. In *2017 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, pages 242–247, Sept 2017.

- [63] M. Poggi and E. Uchoa. Chapter 3: New exact algorithms for the capacitated vehicle routing problem. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 59–86. SIAM, 2014.
- [64] A. Poot, G. Kant, and A. P. M. Wagelmans. A savings based method for real-life vehicle routing problems. *Journal of the Operational Research Society*, 53(1):57–68, 2002.
- [65] C. Prins. A GRASP  $\times$  evolutionary local search hybrid for the vehicle routing problem. *Bio-Inspired Algorithms for the Vehicle Routing Problem*, F. Pereira and J. Tavares, eds., Springer, Berlin, Heidelberg, 161:35–53, 2009.
- [66] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [67] K.F. Reimann, M. Doerner and R.F. Hartl. D-ants: Savings based ants divide and conquer the vehicle routing problem. *Computers & Operations Research*, 31:563–591, 2004.
- [68] M.G.C. Resende and C.C. Ribeiro. Greedy randomized adaptive search procedures: Advances, hybridizations, and applications. *Handbook of Metaheuristics*, M. Gendreau and J.-Y. Potvin, eds., Springer, New York, pages 283–319, 2010.
- [69] S. Ropke and D. Pisinger. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, 40(4):455–472, 11 2006.
- [70] Peter Rossi. *bayesm: Bayesian Inference for Marketing/Micro-Econometrics*, 2017. R package version 3.1-0.1.
- [71] D.G Rossit, D. Vigo, F. Tohmé, and M. Frutos. Improving visual attractiveness in capacitated vehicle routing problems: a heuristic algorithm. In *XVIII Latin-Iberoamerican Conference on Operations Research*, pages 748–755, 2016.
- [72] S. Sahoo, S. Kim, B. Kim, B. Kraas, and A Popov. Routing optimization for waste management. *Interfaces*, 35(1):24–36, 2005.

- [73] J. Sauer and A. Sonderegger. The influence of product aesthetics and user state in usability testing. *Behaviour & Information Technology*, 30(6):787–796, 2011.
- [74] M.W.P. Savelsbergh. A parallel insertion heuristic for vehicle routing with side constraints. *Statistica Neerlandica*, 44(3):139–148, Sept 1990.
- [75] S. Schröder. Jsprit. <https://github.com/graphhopper/jsprit>, Accessed in May 2017.
- [76] F. Semet, P. Toth, and D. Vigo. Chapter 2: Classical exact algorithms for the capacitated vehicle routing problem. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 37–57. SIAM, 2014.
- [77] SINTEF. Solomons 100 customers best known solution values. <https://www.sintef.no/projectweb/top/vrptw/solomon-benchmark/100-customers/>, Accessed in July 2017.
- [78] M. Solomon. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, 35(2):254–265, 1987.
- [79] K. Sörensen, M. Sevaux, and P. Schittekat. “Multiple neighbourhood” search in commercial VRP packages: evolving towards self-adaptive methods. *Cotta C, Sevaux M, Sörensen K (eds) Adaptive and multilevel metaheuristics, volume 136 of studies in computational intelligence. Springer, Berlin*, pages 239–253, 2008.
- [80] J Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):257–285, 1988.
- [81] H. Tang and E. Miller-Hooks. Interactive heuristic for practical vehicle routing problem with solution shape constraints. *Transportation Research Record*, 1964(1):9–18, 2006.
- [82] M. Thüring and S. Mahlke. Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, 42(4):253–264.
- [83] P. Toth and D. Vigo. *Vehicle Routing: Problems, Methods and Applications*. Mathematical Optimization Society and the Society for Industrial and Applied Mathematics, Philadelphia, 2014.

- [84] E. Uchoa, D. Pecin, A. Pessoa, M. Poggi, T. Vidal, and A. Subramanian. New benchmark instances for the capacitated vehicle routing problem. *European Journal of Operational Research*, 257(3):845–858, 2017.
- [85] P. van Schaik and J. Ling. The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human Computer Studies*, 67(1):79–89, 2009.
- [86] M. Vössing. Towards managing complexity and uncertainty in field service technician planning. In *2017 IEEE 19th Conference on Business Informatics (CBI)*, volume 01, pages 312–319, July 2017.
- [87] C. Zhou, Y. Tan, L. Liao, and Y. Liu. Solving the multi-vehicle pick-up and delivery problem with time widows by new construction heuristic. In *Sixth International Conference on Intelligent Systems Design and Applications*, volume 2, pages 1035–1042, Oct 2006.