
UN ENFOQUE ESTOCASTICO Y MULTI-DEPOT PARA EL DISEÑO DE FLOTAS DE REPARTO: MODELACION Y ALGORITMO DE SOLUCION

Pablo A. Miranda, Rodrigo A. Garrido
Departamento de Ingeniería de Transporte
Pontificia Universidad Católica de Chile
Casilla 306, Santiago 22, cod. 105, Chile.
Tel: (56 – 2) 686 – 4270, Fax (56 – 2) 688 – 4818
e-mail: pmiranda@ing.puc.cl, rgarrido@ing.puc.cl

RESUMEN

En el presente paper se desarrolla un enfoque estocástico para el diseño de flotas de reparto, la cual es utilizada en la entrega de productos de una empresa distribuidora. Esta modelación se basa en la existencia de un conjunto conocido de centros de distribución, de modo que se debe agrupar los clientes de dicha empresa en conjuntos que serán servidos por un mismo camión. El modelo presentado permite decidir el número de camiones utilizados en el reparto de productos, y la asignación de clientes a cada uno de ellos, respetando la restricción de capacidad de cada vehículo. Lo anterior se hace considerando una demanda estocástica para cada uno de los clientes, modelada a través de su media y su varianza.

El modelo desarrollado se diferencia de los tradicionales modelos encontrados en la literatura, en que se asume simultáneamente la existencia de una restricción de capacidad estocástica, y la existencia de múltiples centros de distribución, los cuales servirán de origen para las respectivas rutas de reparto. La modelación de costos de ruteo es realizada a través de una estructura tipo Hub&Spokes. Adicionalmente se presenta una heurística de resolución, junto a un algoritmo basado en relajación lagrangeana que permite evaluar la calidad de la solución entregada por la heurística.

1. INTRODUCCION Y ANTECEDENTES GENERALES

Dentro de la gestión de la cadena de abastecimientos o gestión logística para empresas productoras-distribuidoras, los problemas se han clasificado dentro de tres niveles jerárquicos; el nivel estratégico (largo plazo), nivel táctico (mediano plazo) y nivel operacional (corto plazo). Dentro del primer nivel se puede mencionar el diseño de la red de distribución y la planificación agregada de la producción. Dentro del nivel operativo están las decisiones asociadas al ruteo diario de vehículos de reparto y las decisiones de ordenamiento y/o abastecimiento de cada centro de distribución. Entre estos dos niveles, es decir en el nivel táctico, están una serie de problemas, como el diseño de espacios de almacenamiento, definición de estrategias de distribución, elección de estrategias de inventario, etc., las cuales a su vez pueden estar más cerca del nivel operativo, o del nivel estratégico. Uno de estos últimos problemas, es el diseño de flota de distribución y agrupación de clientes en zonas de demanda o rutas de reparto. Este problema, si bien puede ser considerado un problema de tipo táctico, especialmente porque es revisado cada períodos de tiempo relativamente cortos, se acerca bastante al nivel estratégico, debido a la importancia dentro del diseño de la red de distribución, pudiendo incluso determinar el éxito o fracaso de una empresa. Para empresas de gran tamaño, donde los clientes pueden ser cientos o miles, la utilización de juicios de expertos puede tornarse impracticable, de modo que la utilización de modelos matemáticos para resolver este tipo de problemas se hace imprescindible.

En la literatura se ha encontrado que el problema de agrupación de clientes, conocido como Clustering Problem, es abordado mediante diferentes enfoques. Por un lado está la modelación discreta, como en Koskosidis and Powell (1992) y Yang et al. (1998) en que se debe decidir si un cliente pertenece o no a una ruta de reparto, versus una modelación continua, como en Novaes et al. (2000) y Newell and Daganzo (1986), en que las zonas pueden ser definidas mediante coordenadas polares. Por otro lado está la modelación determinística, en contraste con una modelación estocástica, especialmente para la demanda de clientes y la restricción de capacidad. Ver Novaes et al. (2000), Powell et al. (1995), Birge and Louveaux (1997) y Stewart and Golden (1983), en donde se muestra como incorporar programación estocástica a problemas de ruteo vehicular. La modelación estocástica de restricciones consiste en establecer una probabilidad o nivel de servicio que restringe las soluciones a no superar el consumo de un recurso, con una probabilidad igual a dicho nivel de servicio. Este enfoque es conocido como Chance Constrained Problem.

Este problema, Clustering Problem, es considerado NP-Completo, de modo que su resolución de manera exacta para instancias de gran tamaño es prácticamente imposible. De este modo en la literatura se puede encontrar una serie de algoritmos, heurísticas y meta-heurísticas, que permiten encontrar buenas aproximaciones para este tipo de problemas. Uno de ellos es el método de relajación lagrangeana, que permite encontrar buenas soluciones y buenas cotas duales para la función objetivo, entregando además una cota para el error cometido por la solución encontrada por las heurísticas (duality gap). En Miranda and Garrido (2003), Garrido (2000), Melkote and Daskin (2001), Nozick (2001), Koskosidis and Powell (1992), Bramel and Simchi-Levi (2000), y Daskin (1995), es posible encontrar heurísticas de solución y aplicaciones del método de relajación lagrangeana, a diferentes problemas asociados a la gestión logística.

En este documento se presenta una versión estocástica del problema de agrupación de clientes, considerando restricción de capacidad. Además se considera la existencia de múltiples depósitos

que servirán de origen para las respectivas rutas de reparto. Estos depósitos a su vez pueden o no ser servidos por una planta o proveedor. Esta estructura de la red de distribución es representada en la Figura 1. Se presenta además una heurística de solución que mejora iterativamente una solución factible, a través de mejoras locales. Por otra parte se presenta una metodología basada en relajación lagrangeana, que permite encontrar una cota inferior al valor óptimo del problema, permitiendo evaluar la calidad de las soluciones entregadas por la heurística señalada.

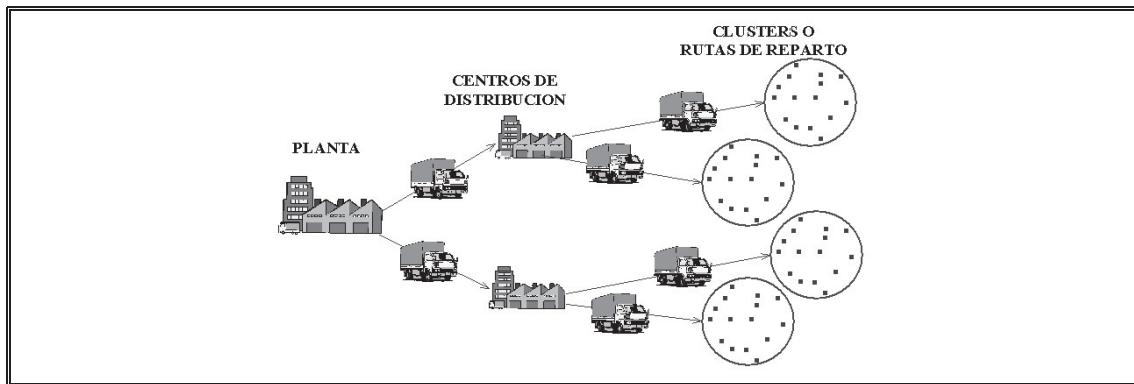


Figura 1: Representación Gráfica de una Red de Distribución de Productos

En el Capítulo 2 se presenta el modelo matemático que resuelve el problema descrito, Stochastic Capacitated Multi-Depot Clustering Problem (SCMDCP). El Capítulo 3 describe la heurística desarrollada para la resolución del SCMDCP. En el Capítulo 4 se describe la metodología utilizada para la obtención de una cota inferior al valor óptimo del problema. El Capítulo 5 presenta los resultados de aplicar las metodologías descritas en el Capítulo 3 y Capítulo 4 a un ejemplo numérico. Finalmente el Capítulo 6 presenta las principales conclusiones y una breve discusión de posibles extensiones y líneas de investigación.

2. UN MODELO ESTOCASTICO PARA LA AGRUPACION DE CLIENTES, CON RESTRICCION DE CAPACIDAD Y MULTIPLES DEPOSITOS

El modelo que se presenta, SCMDCP, permite diseñar la flota de reparto de una empresa con múltiples centros de distribución, los cuales a su vez sirven a un conjunto de clientes espacialmente distribuidos. Cada uno de estos clientes debe ser asignado a una ruta de entrega, la cual deberá ser servida por un camión de capacidad fija y conocida. Eventualmente cada camión deberá resolver un problema del vendedor viajero, a fin de obtener una ruta de mínimo costo para la entrega de productos a los clientes asignados. Sin embargo en la presente investigación, dado el nivel estratégico del problema, no se considera la elección de rutas específicas para cada cluster o grupo de clientes. Lo anterior se justifica además por el hecho de que la demanda de clientes es estocástica, de modo que cada cliente en un día cualquiera posee una probabilidad mayor a cero de no presentar demanda. De este modo la ruta de entrega debe diseñarse día a día, dependiendo de los clientes que participen en cada instancia, en contraste con el problema de diseño de flota que es resuelto a comienzos del período de planificación.

De este modo, en vez de estudiar los costos dentro de cada cluster basados en la búsqueda de rutas óptimas, se modelaron dichos costos con una estructura tipo Hub&Spokes, en la cual existe

un costo de entrega directo al centroide de cada grupo de clientes o cluster, y costos de transporte entre cada uno de los clientes y el respectivo centroide al cual ha sido asignado. Esta estructura de costos puede ser visualizada en la Figura 2. En adelante nos referiremos indistintamente a rutas de reparto o clusters. Por otro lado la modelación se basa en una estructura discreta, en la cual existe un conjunto fijo de centros de distribución y un conjunto fijo de clientes.

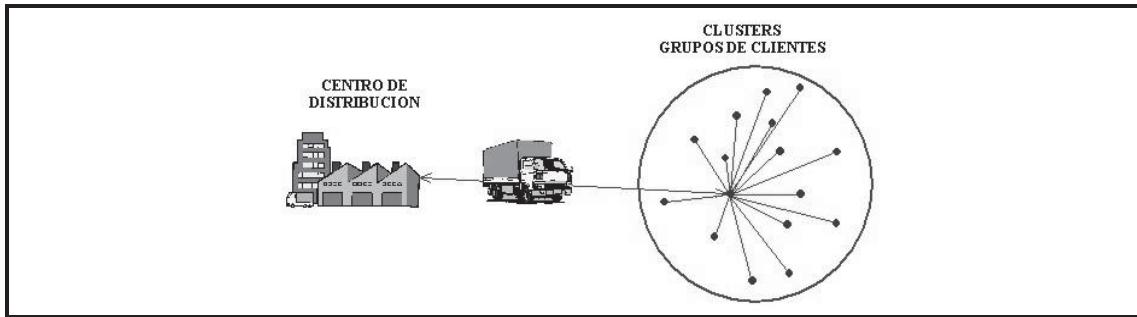


Figura 2: Estructura de Costos de Ruteo Dentro de Cada Cluster o Ruta de Entrega

A continuación se definen los parámetros, conjuntos y variables que se utilizan en la modelación:

- M : Conjunto de clientes.
- X_j : Variable binaria que indica si un cliente j , ha sido elegido como centroide de algún cluster o ruta de reparto.
- W_{jl} : Variable binaria que indica si el cliente l es asignado al cluster j , es decir cuyo centroide es el cliente j .
- D_j : Variable que indica la media de la demanda del cluster j .
- V_j : Variable que indica la varianza de la demanda del cluster j .
- $RCap$: Capacidad de los vehículos utilizados para servir los clientes de cada cluster.
- v_j : Demanda media diaria del cliente j .
- σ_j^2 : Varianza de la demanda diaria del cliente j .
- CT_j : Costo de transporte entre el cliente o centroide j y el centro de distribución más cercano a él.
- CR_{jl} : Costo de transporte entre el cliente l y el centroide j ¹.
- TGR : Gastos fijos por período (diarios) asociados a un vehículo de reparto.
- $Z_{1-\theta}$: Percentil de la distribución Normal (0,1) que acumula una probabilidad $1-\theta$.

Así el modelo que resuelve el problema descrito, SCMDCP, es el siguiente:

$$\text{Min} \quad \sum_{j=1}^M (TGR + CT_j) \cdot X_j + \sum_{j=1}^M \sum_{l=1}^M CR_{jl} \cdot W_{jl} \quad (1)$$

$$\text{Sujeto a:} \quad \sum_{j=1}^M W_{jl} = 1 \quad \forall l = 1, \dots, M \quad (2)$$

$$W_{jl} \leq X_j \quad \forall j, l = 1, \dots, M \quad (3)$$

$$D_j + Z_{1-\theta} \cdot \sqrt{V_j} \leq RCap \cdot X_j \quad \forall j = 1, \dots, M \quad (4)$$

¹ Debido a que potencialmente cualquier cliente puede ser elegido como centroide de algún cluster, este costo debe estar definido para todo par de clientes.

$$\sum_{l=1}^M W_{jl} \cdot v_l = D_j \quad \forall j = 1, \dots, M \quad (5)$$

$$\sum_{l=1}^M W_{jl} \cdot \sigma_l^2 = V_j \quad \forall j = 1, \dots, M \quad (6)$$

$$W_{jl}, X_j \in \{0,1\} \quad \forall j, l = 1, \dots, M \quad (7)$$

La ecuación (1) representa los costos totales de transporte del sistema, considerando la estructura presentada en la Figura 2. El primer término contempla los gastos fijos que se incurren en una ruta de entrega, como gastos de mantenimiento, licencias, permisos de circulación, sueldos, etc., representados por el parámetro TGR . Se contempla además el costo de transporte entre los centros de distribución y cada uno de los centroides j . En este caso se asumirá que cada cluster será asignado al centro de distribución más cercano al respectivo centroide. Este supuesto debe ser revisado y será discutido como una extensión del presente enfoque. El segundo término de (1) representa los costos de asignar clientes a cada uno de los cluster elegidos.

La restricción (2) obliga a que cada cliente sea asignado exactamente a un cluster o centroide. Esta restricción es válida tanto para clientes que no representan centroides como para aquellos que han sido elegido como tal. De hecho si un cliente j es elegido como centroide, es decir $X_j = 1$, automáticamente la solución implica asignar dicho cliente a sí mismo ($W_{jj} = 1$), pues obviamente el costo respectivo es cero ($CR_{jj} = 0$). La restricción (3) establece que sólo es posible asignar clientes a otros que hayan sido elegidos como centroides ($X_j = 1$). En caso contrario ($X_j = 0$), no debe ser posible asignar clientes a dicho centroide. La restricción (4) establece la restricción estocástica de capacidad. Esta restricción asegura que la demanda diaria generada en cada cluster no supere la capacidad $Rcap$ con una probabilidad $1 - \theta$. Las restricciones (5) y (6) establecen la relación existente entre la media y la varianza de la demanda de cada cluster y la media y la varianza de la demanda de cada cliente asignado a dicho cluster. Finalmente la restricción (7) establece la integralidad de las variables del modelo (W_{jl} y X_j).

3. HEURISTICA DE SOLUCION

El presente capítulo describe de manera general la heurística desarrollada para resolver el modelo SCMDCP. Koskosidis and Powell (1992), presentan un modelo determinístico para resolver el clustering, considerando una restricción de capacidad en cada cluster. Además se asume la existencia de un único depósito a partir del cual deben partir los camiones. Los autores proponen una heurística para resolver este problema, obteniendo resultados bastante eficientes. Sin embargo, dado que en el modelo establecido en el Capítulo 2, se asume la existencia de múltiples centros de distribución, y la restricción de capacidad se establece de manera estocástica, esta heurística no es aplicable a nuestro modelo. De este modo, se extendió dicha heurística con el fin de considerar las diferencias descritas. Este algoritmo se basa en, a partir de una solución inicial factible, mejorar dicha solución considerando un algoritmo de búsqueda local. Así el algoritmo siempre va disminuyendo el valor de la función objetivo del problema, de modo que, asumiendo la existencia de un óptimo global, la convergencia está garantizada (aunque no necesariamente a dicho óptimo global). Lo anterior se justifica por el hecho de que si en alguna iteración del algoritmo no se obtienen mejoras, el algoritmo se detiene.

La heurística consiste básicamente en cuatro pasos:

- Paso I** : Generación de solución inicial de centroides.
- Paso II** : Asignación tipo greedy de clientes a centroides.
- Paso III** : Reelección de centroides.
- Paso IV** : Intercambio de clientes.

El algoritmo itera entre los Pasos I y II, hasta encontrar una buena solución factible. El paso I consiste en la elección de un conjunto inicial de centroides en base algún criterio. En este caso se calcula el costo de establecer cada uno de los clientes como centroide, y se elige los de menor costo. La asignación tipo greedy del Paso II, consiste en asignar los clientes al centroide más cercano, en caso que sea factible; el paso además contempla la elección de un nuevo centroide, en caso que existan clientes que no se puedan asignar a ningún centroide. Luego, el Paso III, realiza una optimización local dentro de cada cluster (compuesto por el centroide y los clientes asignados a él), reeligiendo el centroide de modo de minimizar los costos dentro de éste. Finalmente, el Paso IV busca cambios de clientes asignados a diferentes centroides, de modo de reducir los costos totales de ambos clusters, respetando la restricción de capacidad de ambos. El algoritmo itera entre los pasos III y IV, hasta que no existan mejoras en la solución.

Las diferencias entre el algoritmo presentado en este artículo y el propuesto por Koskosidis and Powell (1992), están en el Paso I, donde la elección de centroides iniciales se hace sobre la base de los costos de cada cluster, los cuales dependen del centro de distribución asociado a cada uno de los clientes. Similarmente en el paso III, la reelección del centroide de cada cluster, también depende del centro de distribución asociado a cada uno de los clientes del cluster. Por otro lado en el Paso II, la asignación tipo greedy de clientes a centroides contempla la restricción de capacidad estocástica, dada por la expresión (4). Finalmente, en el Paso IV, el intercambio de clientes es factible si esta restricción de capacidad estocástica es respetada.

Cabe señalar que, como la mayoría de los algoritmos de búsqueda local tienden a encontrar soluciones “óptimos locales”, es interesante analizar si el algoritmo entrega mejores soluciones al considerar diferentes soluciones iniciales, representando una interesante línea de investigación. La selección de un conjunto de soluciones iniciales para comenzar el algoritmo puede ser realizada sobre la base de juicios de expertos, colapsando esta alternativa en caso de tratarse de una cantidad de clientes muy grande. Por otro lado es interesante generar métodos automáticos para la obtención de diferentes soluciones iniciales, que no requieran de intervención humana.

En relación con el problema de clustering, diversos autores se han dedicado a definir algoritmos de búsqueda de “una” solución inicial y que resulten en buenas soluciones finales (sin la necesidad de obtención de “varias” soluciones iniciales). Koskosidis and Powell (1992) presentan tres enfoques para la obtención una solución inicial tratando de minimizar la estructura de costos tipo Hub&Spokes, entregando resultados bastante eficientes. En el presente trabajo se considera una de estas tres metodologías, extendida a un caso estocástico y con la existencia de múltiples depósitos (dados por los Pasos I y II de la heurística).

4. OBTENCION DE UNA COTA INFERIOR, MEDIANTE EL USO DE RELAJACION LAGRANGEANA

En el presente capítulo se describe el enfoque utilizado para la obtención de una cota inferior para el valor óptimo del problema. De este modo la diferencia entre la cota superior, obtenida mediante la heurística desarrollada en el capítulo anterior, y la cota inferior obtenida mediante el presente enfoque, representa una cota superior para el error que se comete con la utilización de dicha heurística, o cota superior para el gap-dual. El enfoque se basa en relajación lagrangeana combinado con el método del subgradiente, para la actualización y optimización de las variables duales de las restricciones relajadas. La sección 4.1 presenta un fortalecimiento del problema SCMDCP, de modo de mejorar las cotas entregadas por el problema relajado. La sección 4.2 desarrolla la relajación de las restricciones seleccionadas, descomponiendo el problema en un subproblema para cada cliente. Se presenta además la manera de resolver cada uno de estos subproblemas. La sección 4.3 muestra como actualizar, a partir de la solución de los subproblemas, las variables duales de las restricciones relajadas.

4.1. Fortalecimiento del Problema SCMDCP

Se realizaron dos cambios a la formulación del modelo. El primero consiste en la consideración de las restricciones (5) y (6), como desigualdades, es decir son reemplazadas por:

$$\sum_{l=1}^M W_{jl} \cdot v_l \leq D_j \quad \forall j = 1, \dots, M \quad (8)$$

$$\sum_{l=1}^M W_{jl} \cdot \sigma_l^2 \leq V_j \quad \forall j = 1, \dots, M \quad (9)$$

Este cambio no altera la optimalidad del problema, debido a que para cualquier solución factible, considerando las restricciones (8) y (9), siempre será posible reducir los valores de D_j y V_j , alcanzándose la igualdad y manteniéndose el valor de la función objetivo inalterable. Además, con el objetivo de robustecer la formulación del problema, se incorporaron las siguientes restricciones o cortes:

$$0 \leq D_j \leq D_{\max} \quad \forall j = 1, \dots, M \quad (10)$$

$$0 \leq V_j \leq V_{\max} \quad \forall j = 1, \dots, M \quad (11)$$

Los parámetros D_{\max} y V_{\max} representan la máxima demanda media y varianza respectivamente, que es posible asignar a un cluster, respetando la capacidad disponible. Este tipo de restricciones, conocidas normalmente como restricciones de fortalecimiento, permiten mejorar las cotas entregadas por la relajación. Es fácil ver que se puede obtener buenos valores para D_{\max} y V_{\max} , buscando los clientes que presenten el mayor valor para v_l/σ_l^2 y σ_l^2/v_l , respectivamente. Luego, para el caso de D_{\max} , se determinará el máximo número de clientes de este tipo (es decir con una media y una varianza igual a la media y la varianza del cliente con el mayor valor para v_l/σ_l^2), de modo de alcanzar en igualdad la restricción de capacidad, dada por la ecuación (4). Se procede de manera análoga para la obtención de V_{\max} .

4.2. Descomposición y Resolución de Subproblemas

Las restricciones relajadas son las (2), (8) y (9), separando el problema en uno para cada cliente j . De este modo si ϕ^1 , ϕ^2 y ϕ^3 son los vectores de variables duales asociadas a las restricciones relajadas, la función dual-lagrangeana, consistente en la incorporación de manera penalizada de las restricciones relajadas en la función objetivo, se escribe como:

$$\begin{aligned} & \sum_{j=1}^M (TGR + CT_j) \cdot X_j + \sum_{j=1}^M \sum_{l=1}^M CR_{jl} \cdot W_{jl} + \sum_{l=1}^M \phi_l^1 \cdot \left(1 - \sum_{j=1}^M W_{jl} \right) \\ & \sum_{j=1}^M \phi_j^2 \cdot \left(\sum_{l=1}^M W_{jl} \cdot v_j - D_j \right) + \sum_{j=1}^M \phi_j^3 \cdot \left(\sum_{l=1}^M W_{jl} \cdot \sigma_l^2 - V_j \right) \end{aligned} \quad (12)$$

Este problema, para valores fijos de las variables duales asociadas a las restricciones relajadas, se descompone en un problema para cada cliente j , SP_j , el cual está dado por:

$$\text{Min} \quad (TGR + CT_j) \cdot X_j - \phi_j^2 \cdot D_j - \phi_j^3 \cdot V_j + \sum_{l=1}^M (CR_{jl} + \phi_j^2 \cdot v_l + \phi_j^3 \cdot \sigma_l^2 - \phi_l^1) \cdot W_{jl} \quad (13)$$

$$\text{sujeto a: } W_{jl} \leq X_j \quad \forall l = 1, \dots, M \quad (14)$$

$$D_j + Z_{1-\theta} \cdot \sqrt{V_j} \leq RCap \cdot X_j \quad (15)$$

$$0 \leq D_j \leq D_{\max} \quad (16)$$

$$0 \leq V_j \leq V_{\max} \quad (16)$$

$$W_{jl}, X_j \in \{0,1\} \quad \forall l = 1, \dots, M \quad (17)$$

Este subproblema se resuelve de la siguiente manera:

- Calcular:

$$\Delta_j = (TGR + CT_j) - \Pi_j + \sum_{l=1}^M \min \left[CR_{jl} + \phi_j^2 \cdot v_l + \phi_j^3 \cdot \sigma_l^2 - \phi_l^1, 0 \right] \quad (18)$$

donde:

$$\Pi_j = \begin{cases} 0 & \text{si } \phi_j^2 = \phi_j^3 = 0 \\ \phi_j^2 \cdot D_{\max} & \text{si } \phi_j^2 > 0, \phi_j^3 = 0 \\ \phi_j^3 \cdot V_{\max} & \text{si } \phi_j^2 = 0, \phi_j^3 > 0 \\ \phi_j^2 \cdot D_{\max} + \phi_j^3 \cdot \left(\frac{RCap - D_{\max}}{Z_{1-\theta}} \right)^2 & \text{si } \phi_j^2 \cdot D_{\max} > \phi_j^3 \cdot V_{\max}, \\ & \phi_j^2 > 0, \phi_j^3 > 0 \\ \phi_j^2 \cdot \left(RCap - Z_{1-\theta} \cdot \sqrt{V_{\max}} \right) + \phi_j^3 \cdot V_{\max} & \text{si } \phi_j^2 \cdot D_{\max} \leq \phi_j^3 \cdot V_{\max}, \\ & \phi_j^2 > 0, \phi_j^3 > 0 \end{cases} \quad (19)$$

- Luego si $\Delta_j < 0$, hacer $X_j = 1$,

$$W_{jl} = \begin{cases} 0 & \text{si } CR_{jl} + \phi_j^2 \cdot v_l + \phi_j^3 \cdot \sigma_l^2 - \phi_l^1 \geq 0 \\ 1 & \text{si } CR_{jl} + \phi_j^2 \cdot v_l + \phi_j^3 \cdot \sigma_l^2 - \phi_l^1 < 0 \end{cases} \quad \forall l = 1, \dots, M \quad (20)$$

y

$$(D_j, V_j) = \begin{cases} (0, 0) & \text{si } \phi_j^2 = \phi_j^3 = 0 \\ (D_{\max}, 0) & \text{si } \phi_j^2 > 0, \phi_j^3 = 0 \\ (0, V_{\max}) & \text{si } \phi_j^2 = 0, \phi_j^3 > 0 \\ \left(D_{\max}, \left(\frac{RCap - D_{\max}}{Z_{1-\theta}}\right)^2\right) & \text{si } \phi_j^2 \cdot D_{\max} > \phi_j^3 \cdot V_{\max}, \\ & \phi_j^2 > 0, \phi_j^3 > 0 \\ \left(RCap - Z_{1-\theta} \cdot \sqrt{V_{\max}}, V_{\max}\right) & \text{si } \phi_j^2 \cdot D_{\max} \leq \phi_j^3 \cdot V_{\max}, \\ & \phi_j^2 > 0, \phi_j^3 > 0 \end{cases} \quad (21)$$

- Si $\Delta_j \geq 0$, entonces $X_j = 0, D_j = V_j = 0, W_{jl} = 0, \forall l = 1, \dots, M$

De este modo en cada iteración k , dado un conjunto de variables duales asociadas a las restricciones relajadas, se dispondrá de las soluciones de los subproblemas, consistentes en $X_j^{(k)}$, $W_{jl}^{(k)}$, $D_j^{(k)}$ y $V_j^{(k)}$.

4.3. Actualización de Variables Duales y Cálculo de Cota Inferior

En cada iteración k , dado el valor de las variables duales, $\phi^{1(k)}$, $\phi^{2(k)}$ y $\phi^{3(k)}$, y dados los vectores de holguras de las restricciones relajadas (los cuales se calculan sobre la base de la solución de los subproblemas SP_j), dados por:

$$H1_l^{(k)} = 1 - \sum_{j=1}^M W_{jl}^{(k)} \quad (22)$$

$$H2_j^{(k)} = D_j^{(k)} - \sum_{l=1}^M W_{jl}^{(k)} \cdot v_l \quad (23)$$

$$H3_j^{(k)} = V_j^{(k)} - \sum_{l=1}^M W_{jl}^{(k)} \cdot \sigma_l^2 \quad (24)$$

la actualización de las variables duales está dada por:

$$\phi_l^{1(k+1)} = \phi_l^{1(k)} + \alpha^{(k)} \cdot H1_l^{(k)} \quad (25)$$

$$\phi_j^{2(k+1)} = \max \left\{ 0, \phi_j^{2(k)} - \alpha^{(k)} \cdot H2_j^{(k)} \right\} \quad (26)$$

$$\phi_j^{3(k+1)} = \max \left\{ 0, \phi_j^{3(k)} - \alpha^{(k)} \cdot H3_j^{(k)} \right\} \quad (27)$$

El parámetro $\alpha^{(k)}$ representa el paso de avance, para el cual típicamente se utiliza la expresión:

$$\alpha^{(k)} = \rho^{(k)} \cdot \frac{\left(Z_{Sup} - Z_{Inf}^{(k)}\right)}{\|H1^{(k)}\|^2 + \|H2^{(k)}\|^2 + \|H3^{(k)}\|^2} \quad (28)$$

Z_{Sup} es la cota superior en evaluación y que ha sido obtenida mediante la heurística descrita en el capítulo anterior; $Z_{Inf}^{(k)}$ es una cota inferior al valor óptimo del problema SCMDCP disponible en

la iteración k ; $\rho^{(k)}$ es un parámetro de control entre 0 y 2, que decrece conforme avanzan las iteraciones, especialmente cuando se observa un alto número de iteraciones en que no se obtiene mejora en la cota inferior. Esta cota inferior se actualiza en cada iteración según:

$$Z_{Inf}^{(k+1)} = \text{Max} \left\{ Z_{Inf}^{(k)}, \begin{array}{l} \sum_{j=1}^M (TGR + CT_j) \cdot X_j^{(k)} + \sum_{j=1}^M \sum_{l=1}^M CR_{jl} \cdot W_{jl}^{(k)} + \sum_{l=1}^M \phi_l^{1(k)} \cdot \left(1 - \sum_{j=1}^M W_{jl}^{(k)} \right) \\ + \sum_{j=1}^M \phi_j^{2(k)} \cdot \left(\sum_{l=1}^M W_{jl}^{(k)} \cdot v_l - D_j^{(k)} \right) + \sum_{j=1}^M \phi_j^{3(k)} \cdot \left(\sum_{l=1}^M W_{jl}^{(k)} \cdot \sigma_l^2 - V_j^{(k)} \right) \end{array} \right\} \quad (29)$$

Finalmente, el algoritmo se detiene en la medida que el valor de Z_{Inf} no mejore luego de un número determinado de iteraciones sucesivas, o bien cuando la diferencia porcentual entre Z_{Sup} y $Z_{Inf}^{(k)}$ sea inferior a un valor predeterminado, típicamente en torno a un 1% (entre otros criterios).

5. APLICACION NUMERICA

El procedimiento heurístico descrito en la sección 3 fue aplicado a un ejemplo numérico, el cual considera 20 centros de distribución y 200 clientes. Los clientes fueron aleatoriamente ubicados en un área cuadrada de 1000 km. de ancho, mientras que los centros de distribución fueron uniformemente distribuidos sobre dicha área. Para los gastos fijos diarios, TGR, se estableció un valor de \$10,000, mientras que los costos de transporte fueron estimados basados en un costo unitario de 56 \$/km. La media de las demandas de los clientes fueron aleatoriamente simuladas en torno a 14 unidades, mientras que las varianzas fueron generadas aleatoriamente considerando un coeficiente de variación cercano a 1. Para la capacidad de los vehículos se asumió un valor de 160 unidades, mientras que el nivel de servicio de la restricción de capacidad fue fijado en 90% ($Z_{0.9} = 1.28$).

La primera solución factible encontrada por la heurística consideró 27 centroides. Luego, para el mismo número de centroides, se encontró una segunda solución la cual entregó un ahorro de 13.8%. Finalmente la mejor solución encontrada por el algoritmo entregó un ahorro de 33.4% respecto a la primera solución factible, considerando la utilización de 40 centroides o clusters. Se resolvió además otras instancias del problema, considerando diferentes valores de capacidad. Por ejemplo, para un aumento de la capacidad de un 6%, se obtuvo un aumento de los costos de un 1%, mientras que reduciendo la capacidad en un 6% se obtuvo un aumento de costos de un 1.2%. Cada instancia fue resuelta en unos pocos segundos, en procesador Pentium IV (1.7GHz) con 256MB de memoria ram.

Finalmente se aplicó la metodología propuesta en el Capítulo 4, para estimar el máximo error cometido por la heurística, la cual entregó un error máximo de un 2.92%, para la instancia cuya capacidad es de 160. Para un aumento de un 6% en la capacidad, la heurística cometió un error máximo de un 3.9%. Finalmente para la instancia con una reducción de un 6% en la capacidad, la heurística presentó un error máximo de un 4.04%. Se debe notar que estos valores son cotas superiores para el valor desconocido del gap-dual, el cual es típicamente mayor que cero, para problemas enteros, lo cual indica que las soluciones encontradas por la heurística están cercanas al óptimo del problema.

6. CONCLUSIONES E INVESTIGACIONES FUTURAS

En el presente paper se ha obtenido un modelo que permite diseñar la flota de reparto en un sistema de distribución con múltiples centros de almacenamiento y distribución, considerando una situación estocástica, especialmente para efectos de la restricción de capacidad para los clientes asignados a cada uno de los camiones de la flota. Se desarrolló un algoritmo heurístico para resolver este problema y un algoritmo para la búsqueda de una cota inferior al costo mínimo de la flota de reparto. Se aplicó ambos algoritmos a una instancia con 20 centros de distribución y 200 clientes, obteniéndose en pocos segundos tanto la solución heurística, como la cota inferior del costo total óptimo del sistema. Se debe destacar que el tamaño de este problema puede ser bastante realista, manifestando la aplicabilidad del presente enfoque, considerando que en situaciones en que los clientes sean miles, estos pueden ser agregados geográficamente reduciendo el tamaño del problema.

En términos algorítmicos, es interesante estudiar la eficiencia de los algoritmos presentados, para diferentes configuraciones de clientes (cantidad, volumen y variabilidad de las demandas, distribución geográfica, etc.). Luego, se puede evaluar la posibilidad de mejorar la heurística presentada, comenzando a partir de diferentes soluciones iniciales encontradas sobre la base de diferentes criterios. Por otro lado, es posible incorporar metodologías conocidas como meta heurísticas, como los algoritmos genéticos, el annealing simulado y la búsqueda tabú entre otros, que permitirían hacer búsquedas locales más inteligentes y sofisticadas. Así, se puede estudiar si las mejoras obtenidas con este tipo de metodologías, justifican la mayor complejidad de implementación.

Una de las extensiones básicas que puede ser desarrollada, es la optimización del nivel de servicio asociado a la restricción de capacidad, lo cual requeriría el establecimiento de un equilibrio entre los costos marginales incurridos en el sistema de distribución, al aumentar el nivel de servicio, y el beneficio marginal de reducción de costos por la reducción de demanda eventualmente no satisfecha (o satisfecha con mayores costos de distribución) en caso que se supere la restricción de capacidad.

Por otro lado, se debe recordar que en el Capítulo 2 se supuso que cada cluster es asignado al centro de distribución más cercano al centroide respectivo, lo cual es claramente discutible ya que correspondería a un enfoque tipo greedy. Sin embargo, se pretende hacer interactuar este modelo, a modo de subproblema, con un modelo de localización de instalaciones, que permita asignar de manera óptima estos centroides o clusters a los diferentes centros de distribución, decidiendo además la instalación de ellos (pudiendo incluso elegirse la capacidad de dichos centros).

Se destaca la posibilidad de evaluar de manera más precisa, mediante simulación u otro enfoque, los costos de ruteo dentro de los clusters, a fin de validar el enfoque establecido. Así es posible considerar estructuras de demanda más sofisticadas, que consideren correlaciones entre demandas de diferentes clientes o correlaciones temporales de demanda.

Finalmente es posible extender nuestro enfoque, considerando múltiples productos y múltiples períodos, en que los clientes presenten diferentes demandas en cada período.

AGRADECIMIENTOS

Se agradece al Fondo Nacional de Desarrollo Científico y Tecnológico, a través del proyecto FONDECYT 1030608, y a la empresa Nestlé Chile S.A., por financiar y fomentar tanto ésta como otras investigaciones.

REFERENCIAS

- Birge, J. R. and F. Louveaux (1997) **Introduction to Stochastic Programming**. New York, Springer-Verlag.
- Bramel, J. and D. Simchi-Levi (2000) **The Logic of Logistic**. New York, Springer-Verlag.
- Daskin, M. S. (1995) **Network and Discrete Location: Models, Algorithms, and Applications**. New York, Wiley-Interscience.
- Garrido, R. (2000) **Modelación de Sistemas de Distribución de Carga**. Santiago, Ediciones Universidad Católica de Chile.
- Koskosidis, Y. A. and W. B. Powell (1992) Clustering Algorithms for Consolidation of Customer Into Vehicle Shipments. **Transportation Research 26B (5): 365-379**.
- Melkote, S. and M. S. Daskin (2001) An Integrated Model of Facility Location and Transportation Network Design. **Transportation Research 35A (6): 515-538**.
- Miranda, P. A. and R. A. Garrido (2003) Incorporating Inventory Control Decisions into a Strategic Distribution Network Design Model with Stochastic Demand. **En proceso de revisión**.
- Newell, G. F. and C. F. Daganzo (1986) Design of Multiple-Vehicle Delivery Tours-I: A Ring Radial Networks. **Transportation Research 20B (5): 345-363**.
- Novaes, A. G. N., et al. (2000) A Continuous Approach to the Design of Physical Distribution Systems. **Computers and Operations Research 27 (9): 877-893**.
- Nozick, L. K. (2001) The Fixed Charge Facility Location Problem With Coverage Restriction. **Transportation Research 37E (4): 281-296**.
- Powell, W. B., et al. (1995) Stochastic and Dynamic Networks and Routing. **Handbooks in Operations Research & Management Science 8 (3)**.
- Stewart, W. R. and B. L. Golden (1983) Stochastic Vehicle Routing: A Comprehensive Approach. **European Journal of Operational Research 12: 371-385**.
- Yang, J., et al. (1998) On-Line Algorithm for Truck Fleet Assignment and Scheduling under Real-Time Information. **Working Paper**.