

2. ESTADO DEL ARTE

En este capítulo se desarrollarán las contribuciones más destacables relacionadas con la localización de terminales de consolidación de carga o *hubs* y las metodologías existentes para el envío y asignación de mercancía a las rutas de una red de transporte de muchos orígenes a muchos destinos. La producción científica para la resolución de este tipo de problemas se puede dividir en tres familias de metodologías de resolución:

- Aproximaciones continuas. Esta metodología de resolución analiza los sistemas logísticos a partir de las variables más relevantes, suponiendo que éstas no sufren variaciones importantes en la región de estudio. El método de las aproximaciones continuas (AC) se realiza reemplazando la información particular de cada elemento del sistema por funciones agregadas de la demanda y funciones sobre de la distribución espacial de terminales sobre la región de servicio. Con ello se logra reducir el tamaño del problema a un número acotado de variables que permiten resolverlo analíticamente. En el modelo se intenta manejar funciones continuas fácilmente resolubles mediante cálculo elemental. El desarrollo de estas técnicas de análisis aplicadas a sistemas logísticos son atribuibles al Prof. C. Daganzo y Prof. G. Newell y se recogen en Daganzo (2005) y Robusté (2005). Esta metodología se desarrollará en el subcapítulo 2.1.
- Programación matemática. Los métodos pertenecientes a este grupo tratan de desarrollar un modelo matemático para la optimización del sistema logístico. Las contribuciones en este campo empiezan en 1950 si bien su amplia difusión se alcanzó en el periodo 1980-2000. Mediante la definición de un conjunto de variables de decisión del problema, se trata de formular una función objetivo a minimizar (si se trata de costes) o a maximizar (si se habla de nivel de servicio) con una serie de restricciones sobre las variables. En función del objetivo, las características de las variables y su relación, se pueden encontrar un

amplio espectro problemas de programación matemática lineal o no lineal, problemas con variables reales o enteras, etc. El gran número de variables y ecuaciones empleadas produce que la solución exacta de algunos problemas de tamaño significativo (100 nodos o superior) no haya sido alcanzada por los altos requerimientos de memoria y tiempo de computación.

- Algoritmos heurísticos y meta-heurísticos. Debido a la alta complejidad del problema de distribución de muchos orígenes a muchos destinos y a la inexistencia de herramientas que puedan calcular la solución exacta o el óptimo global del problema en un tiempo razonable, se ha desarrollado un conjunto de métodos denominados algoritmos heurísticos. Éstos llegan a un razonable equilibrio entre el tiempo de resolución y la calidad del resultado obtenido. Sin embargo, estos procedimientos de cálculo aproximado han resultado ser inefectivos o ineficientes en algunos problemas *NP-Hard* por lo que ha sido necesario su perfeccionamiento. De este modo se han desarrollado los algoritmos metaheurísticos, procesos iterativos que guían la búsqueda de la solución con un heurístico subordinado. En el proceso de búsqueda, se aplican cambios, transiciones o perturbaciones a la solución actual y estrategias para su aceptación siguiendo unos criterios probabilísticos según se detalla en Golden y Assad (1998). El balance inteligente de estos procedimientos debe llevar a cubrir el espacio de soluciones y encontrar resultados o subdominios de soluciones cerca del óptimo global del sistema.

2.1. APROXIMACIONES CONTÍNUAS

La metodología de las aproximaciones continuas aplicada al problema de muchos orígenes a muchos destinos se desarrolla en Daganzo (2005). Esta contribución no sólo trata el problema de diseño topológico de la red sino que también fija la operativa temporal óptima de gestión de los envíos. Se parte de una región de servicio de área R_s con una serie de terminales que tienen asociados unos envíos de carga entre ellas. Adicionalmente, un subconjunto de estas terminales actúan también como centros de rotura y consolidación de carga (terminales *hub*). En una primera etapa, si bien existen distintas terminales de consolidación, se analizará el problema de transporte de envíos de muchos orígenes a muchos destinos permitiéndose que cada envío pase exclusivamente por una terminal de consolidación de carga.

En particular, el método de análisis propuesto se descompone en tres niveles de análisis diferenciados:

- Nivel operacional. El problema se centra en la configuración y organización de rutas de vehículos entre las terminales *hub* y el resto de delegaciones, determinando una cadena ordenada de nodos de visita asociadas a un vehículo particular. La localización de la terminal de consolidación *hub* así como la frecuencia de servicio entre el resto de delegaciones (entregas o recogidas) y las terminales *hub* están fijadas en otros niveles de análisis.
- Nivel táctico: Se parte de las localizaciones de las terminales como datos de entrada y se pretende coordinar las frecuencias de envío de todos los puntos servidos por el mismo *hub*.
- Nivel estratégico: El problema se basa en definir la localización óptima de las terminales de rotura de carga a partir de los flujos agregados entre orígenes y destinos de la región de servicio.

2.1.1. Problema operacional

Se parte de que cada terminal *hub* tiene asignado un conjunto de pares origen-destino a servir con una frecuencia de envío $1/P_f$ común para todos ellos, y con una demanda agregada de mercancía de valor λ en toda la zona de servicio R_s . Esta consideración determinará un intervalo de envíos en la región igual a P_f . De este modo, el número de paradas realizadas por los vehículos en las rutas debe ser ajustado por la localización relativa a cada terminal *hub* en relación con los ritmos de variación espacial de la demanda y de abastecimiento. Adicionalmente, se asume que:

- P_f es común para todas las terminales, para simplificar la operativa y los horarios
- los costes de inventario en ruta pueden suponerse negligibles
- los vehículos circulan llenos
- cada origen genera una carga a transportar inferior a la capacidad de un vehículo (el vehículo de recogida de una terminal únicamente realiza una parada en aquel origen).

Suponiendo un análisis uni-dimensional y bajo estos supuestos, el número de paradas de recogida efectuadas por distintos vehículos en un tiempo P_f en un origen, m^o , es igual al número de terminales *hub* a las que se encamina su carga. Es decir, se asigna un vehículo para cada terminal que visita los puntos a los que se debe recoger mercancía. En el mismo sentido, el número de paradas de entrega en un tiempo P_f en un destino, m^d , también equivale al número de *hubs* desde las que se sirve la mercancía a la terminal estándar. De este modo, el número de paradas realizadas en un tiempo P_f en una subregión de la zona de

servicio de extensión unitaria es: $\delta^o m^o + \delta^d m^d$, donde δ^o y δ^d son respectivamente las densidades de terminales origen y destino de los envíos en la subregión de servicio.

En este caso se puede definir una distancia de acceso a la subregión de servicio en la recogida (o entrega) como la distancia media entre la terminal desde (o hacia) la que se encamina la mercancía a todos los puntos de recogida (o entrega) multiplicada por el número de ciclos que empiezan en la terminal de consolidación. Si cada vehículo realizara únicamente una parada de reparto o de entrega, la distancia de acceso a la subregión sería igual a la distancia total recorrida (si se realiza el estudio en una dimensión).

De este modo, para comprobar la eficacia de distintas estrategias de asignación de puntos de servicio a terminales de consolidación basta con analizar el número de paradas y la distancia media de acceso a la subregión.

Una posible estrategia que se puede establecer para optimizar la distribución y reducir costes es definir una zona de influencia de recogida alrededor de la ubicación de una terminal (Figura 2.1), de forma que toda la mercancía de los puntos contenidos en esta área sea encaminada a través de la terminal *hub* en cuestión independientemente de su destino. En este sentido, los envíos que tengan como origen la delegación P_1 de la Figura 2.1a) serían gestionados por la terminal *hub* A, mientras que la mercancía con origen en la delegación P_2 sería consolidada en la terminal *hub* B.

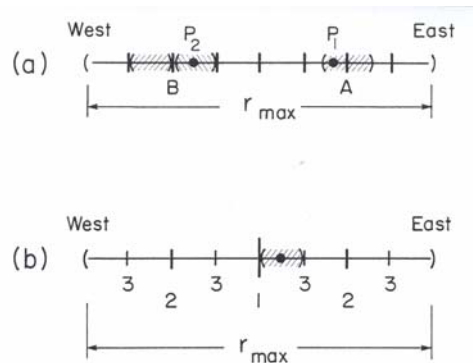


Fig. 2.1. Esquema del sistema de terminales con estructura estándar o jerarquizada en una dimensión. Fuente: Daganzo (2005)

Si todas las terminales de la zona de servicio operan de la misma forma, la determinación de las zonas de influencia se realiza a través de una partición de la zona de servicio. De este modo, si se considera h el número total de terminales de consolidación en servicio o terminales *hub*, se tendría que: $m^o = 1$, $m^d = h$; de forma que el número de paradas por cada franja de reparto y por unidad de longitud (o superficie en el caso bidimensional) es $\delta^o + \delta^d h$. Si el número de destinos es superior al de los puntos de origen, resulta mejor

definir zonas de influencia de reparto, de forma que el número de paradas resulte menor $\delta^o h + \delta^d$.

Sin embargo, esta estrategia presenta el inconveniente que en algunos pares de puntos origen-destino, la distancia media de acceso a la zona supera la distancia mínima que los une a través de una ruta directa sin pasar por la terminal. Este efecto se puede constatar en los envíos entre P_1 y P_2 . Si se realiza el envío a través de las terminales de consolidación A y B asociadas a las zonas de influencia de la Figura 2.1 se recorre más distancia que en un envío directo entre puntos.

Una forma alternativa de servicio que elimina este efecto es determinar distintas zonas de influencia asociadas a una terminal *hub*, como sucede con la terminal B de la Figura 2.1 a). La zona de influencia situada a la izquierda de B únicamente envía mercancía a través de B si sus destinos están situados al este de B así como los puntos de destino entre B y el origen dentro de la zona de influencia. Por otro lado, la zona de influencia de la izquierda únicamente envía mercancía a través de B para destinos situados al oeste.

Esta estrategia elimina el efecto de recorrer dos veces ciertos tramos de la ruta (*back-tracking*) para la mayoría de pares de orígenes y destinos, a excepción de los pares de puntos situados entre dos terminales de consolidación adyacentes. Estos ahorros en la distancia de acceso se consiguen a expensas de añadir una parada más en la recogida en cada origen, ya que $m^0 = 2$. Si r_{max} es la longitud del segmento de servicio de la Figura 2.1 a), el hecho de pasar de la primera estrategia a la estrategia alternativa supone ahorrarse $r_{max}/4h$ unidades de distancia al recorrer la distancia de acceso (en la primera estrategia la mitad del kilometraje desde los puntos origen a la terminal *hub* se realizan hacia atrás, mientras que en la segunda es prácticamente 0), pero se añaden δ^o paradas por unidad de longitud y por franja horaria de servicio.

2.1.2. Jerarquía de las terminales

Hasta el momento, se ha considerado que todas las terminales de consolidación *hub* tienen las mismas características de funcionamiento y de tamaño para poder servir cualquier punto del territorio, de modo que la asignación de los puntos a cada terminal *hub* se ha realizado por criterios de distancia. Sin embargo, se puede demostrar que incrementando en poca medida el número de paradas de recogida en delegaciones, se pueden reducir significativamente las paradas de entrega de mercancías.

Siguiendo con el análisis unidimensional, se puede establecer una jerarquía de terminales *hub* de forma que a la terminal *hub* más cercana al centro de la zona de servicio se la considere perteneciente al nivel $l=1$. De las dos subparticiones iguales de R_s que se generan en el segmento de servicio, se determinan las 2 terminales *hub* centrales de cada subpartición y se consideran pertenecientes al nivel $l=2$, procediendo de esta forma hasta que no queden terminales *hub* sin asignarles un nivel de jerarquía (Figura 2.1 b). Un sistema de distribución con un conjunto completo de terminales jerárquicas de niveles $l=1, \dots, L$ deberá tener un número de terminales igual a $(2^L - 1)$ y un total de 2^{l-1} terminales en cada nivel l .

De este modo, la estrategia de ruteo que puede ahorrar el recorrer dos veces un mismo tramo es la siguiente: se sirve cada par de delgaciones O-D a través de la terminal *hub* de mayor nivel jerárquico localizada entre el origen y el punto de destino. Si no se encuentra ninguna terminal entre los puntos se usa aquella situada más cerca de los puntos que pueda ser visitada desde los dos puntos con la menor distancia combinada. Con esta estrategia, cada punto recibe la mercancía de L terminales *hub* de la misma forma que encamina la mercancía a través de L terminales *hub*.

Con este procedimiento, sin incrementar la distancia de acceso en el reparto o recogida, el número de paradas puede ser reducido de $(2\delta^o + \delta^d h)$ a $(L\delta^o + L\delta^d)$. En Daganzo (2005) se estima que el ahorro en el número de paradas puede ser substancial si $\delta^o \approx \delta^d$, ya que se pasaría de $m^o + m^d \approx h$ a $m^o + m^d \approx 2L = 2\log_2(h+1)$. Por lo tanto, con la estrategia jerárquica el número de paradas crece de forma logarítmica con el número de terminales *hub*. Se debe constatar que los flujos a través de varias terminales son radicalmente diferentes incluso si los orígenes y los destinos son uniformemente distribuidos.

La forma de operación en terminales jerarquizadas es un sistema logístico muy común en distintas empresas de transporte, de forma que se constituye un *hub* de primer nivel y posteriormente se desarrollan terminales *hub* de menor jerarquía.

La extensión del análisis en dos dimensiones se puede resumir en la Figura 2.2, donde las terminales se deben localizar en la intersección de dos líneas de la misma jerarquía. En este caso, si $L=3$, el número máximo de terminales es $(2^L - 1)^2$ debido que hay $(2^L - 1)$ líneas de cada jerarquía.

De forma general, en el estudio en el plano en métrica L_1 (métrica de red mallada, *grid* o Manhattan), únicamente los desplazamientos en los que el segmento que une los puntos origen y destino no cruza a ninguna línea de terminal comportan un cierto retroceso de distancia en la ruta.

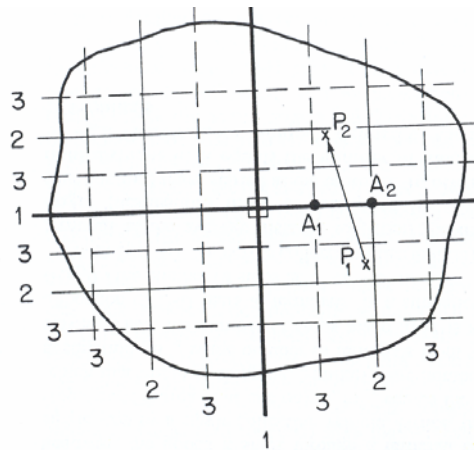


Fig. 2.2. Esquema del sistema de terminales con estructura jerarquizada en dos dimensiones. Fuente: Daganzo (2005)

Si los orígenes y destinos están independientemente distribuidos y $h \gg 1$, la probabilidad de que la ruta de un envío cometa un cierto retroceso en una dirección es del orden de $(1/h)^{1/2}$ y la distancia media adicional debida al retroceso es $1/3(|R_s|/h)^{1/2}$ (la tercera parte de la distancia entre terminales).

Debido que la probabilidad de retroceso en cada dirección es $(1/h)$, la distancia esperada adicional d_a entre todos los pares de puntos O-D se puede explicitar por la ecuación (2.1).

$$d_a = 2|R_s|^{1/2} / (3h) \tag{2.1}$$

Para la métrica Euclídea ó L_2 la expresión sería más compleja pero relativamente similar a la ecuación anterior. El número de paradas en la métrica L_2 utilizando para L la solución real de $h = (2^L - 1)^2$ y $L = \log_2(h^{1/2} + 1)$, se encuentra determinado por la ecuación (2.2).

$$m^0 = m^d \approx [\log_2(h^{1/2} + 1)]^2 \tag{2.2}$$

La ecuación anterior es exacta si el valor del término entre paréntesis es entero, en otros casos la ecuación tiende a infravalorar el valor medio alrededor de un 10%. Por esta razón se tiende a adoptar la ecuación (2.3) para compensar la infravaloración.

$$m^0 + m^d \approx [3h^{1/2}] \tag{2.3}$$

2.1.3. Problema táctico y estratégico

En Daganzo (2005), el problema táctico y estratégico se resuelve utilizando la solución óptima del problema operacional ideal. Se supone que la flota de vehículos es homogénea y de capacidad igual a C unidades de mercancía y un número de terminales *hub* igual a h . El coste unitario de recogida c_{PU} asociado al número de paradas está basado en la ecuación (2.3) y en la consideración que el número de paradas en cada origen es $m^0 = \frac{3}{2}h^{1/2}$, de modo que se puede expresar por la expresión (2.4).

$$c_{PU} = \alpha_2 \left[\frac{\delta_0}{\lambda P_f |R_s|} \right] \left(\frac{3}{2} h^{1/2} \right) \quad (2.4)$$

siendo $\alpha_2 = (c_d k \delta^{-1/2} + c_p)$

La cantidad entre paréntesis de la ecuación (2.4) corresponde al recíproco de la cantidad recogida en una zona origen en un intervalo. De este modo, el coste de recogida y reparto c_{PUD} asociado a la parada por unidad transportada se puede evaluar según la ecuación (2.5.a) si se asume que las densidades de reparto y recogida son similares e iguales a δ .

$$c_{PUD} = \left[\frac{3\alpha_2 \delta_0}{\lambda P_f |R_s|} \right] (h^{1/2}) \quad (2.5.a)$$

A partir de la ecuación (2.1) en la que se estima la distancia extra por vehículo debido a la visita del *hub*, se puede determinar el coste de circuito c_c (coste asociado al doble recorrido de un tramo de la ruta) por mercancía transportada en el envío sin paradas en el que se asume que los vehículos circulan llenos.

$$c_c = \left[\frac{c_d}{C} \right] \left(\frac{2|R_s|^{1/2}}{3h} \right) \quad (2.5.b)$$

El coste es proporcional a la distancia media entre los orígenes y destinos (coste básico aproximadamente igual a $c_d / C |R_s|^{1/2}$). Adicionalmente, se deberá considerar el coste de operación de la terminal c_{OT} por unidad de mercancía transportada, evaluado en la ecuación (2.5c); y el coste de inventario estacionario en los orígenes y destinos c_I determinado en (2.5.d).

$$c_{OT} = \alpha_5 + \alpha_6 \frac{h}{|R_s|} \quad (2.5.c)$$

$$\text{siendo } \alpha_5 = (c_t + c_i P_f + c_r^t (\tau^l + \tau^u))$$

$$\alpha_6 = (c_r^o + c_t^o) / \lambda$$

$$c_l = c_i P_f \quad (2.5.d)$$

La suma de las cuatro ecuaciones anteriores configura en este caso la función de coste logístico. Una vez establecida esta función de costes, pueden aparecer distintas decisiones estratégicas a solucionar. Por un lado, dado un intervalo de envío, se podría plantear el número óptimo de terminales *hub* con el que operar para una empresa de transporte que quiere empezar a servir un nicho de mercado.

Otra tipología de problemas podría aparecer cuando se pretende determinar el mejor intervalo de envío para un valor dado de h . En estos casos, los costes de circuito y de terminal son fijos y únicamente se deben balancear los costes locales de parada y de inventario, con lo que se obtiene un valor óptimo del intervalo de envío definido en la ecuación (2.6).

$$P_l^* \approx \left[\frac{3\alpha_2 \delta}{\lambda c_i |R_s|} h^{1/2} \right]^{1/2} \quad (2.6)$$

En este caso, el coste total c_T por unidad de mercancía transportada, sin incluir el coste de acceso a la zona sin realizar paradas, se puede expresar por la ecuación (2.7).

$$c_T = \alpha_5 + \alpha_6 \frac{h}{|R_s|} + \left(\frac{c_d}{C} \right) \left(\frac{2|R_s|^{1/2}}{3h} \right) + 2 \left(\frac{3\alpha_2 \delta c_i}{\lambda |R_s|} h^{1/2} \right)^{1/2} \quad (2.7)$$

Si se analiza el caso particular en que los costes fijos de terminal son despreciables y los costes de parada c_p son significativamente inferiores a la componente de distancia $kc_d \delta^{-1/2}$, el coste por objeto transportado se puede determinar mediante la expresión (2.8), donde N_o es el número de orígenes (o destinos), K una constante adimensional y asumiendo que $(3k^2)^{1/2} \approx 1$:

$$c_T \approx 2 \left[\frac{c_d |R|^{1/2}}{C} \right] \left\{ \frac{1}{3h} + \frac{K}{N_o} h^{1/4} \right\} \quad (2.8)$$

En este contexto, el valor mínimo del número de terminales *hub* (h^*) que optimiza la ecuación (2.8) se determina por la expresión (2.9):

$$h^* \approx \left(\frac{4 N_o}{3 K} \right)^{4/5} \quad (2.9)$$

Las formulaciones anteriores han sido desarrolladas bajo la perspectiva que todas las rutas no tienen ninguna restricción en su longitud ni en el número de paradas máximo. De forma adicional, las ecuaciones (2.5.a), (2.5.b), (2.5.c), (2.5.d) fueron desarrolladas para mercancías baratas y vehículos idénticos, pero se puede diseñar un sistema de configuración para situaciones diferentes, incluso con P_f variante según la jerarquía de la terminal de consolidación de cargas.

Si algunos de los puntos O-D de servicio tienen asociados unos volúmenes de carga muy superiores al resto de puntos, se pueden plantear estrategias de discriminación que consideren envíos directos entre estos puntos mientras que el resto de puntos sean servidos a través de la terminal *hub*. Los dos sistemas (puntos servidos por envíos directos y puntos pertenecientes a los envíos vía *hub*) tienen costes independientes de las acciones tomadas para controlar los envíos del otro sistema y por tanto pueden ser optimizados de forma separada.

2.1.4. Extensión a sistemas con varias terminales y dos transferencias

En los sistemas de distribución con dos transferencias en terminales *hub*, los problemas organizados de forma no jerarquizada son asimétricos con respecto a la recogida y distribución y por tanto requieren menores paradas en cada finalización del viaje. En este sentido, los sistemas de terminales jerarquizados no son utilizados para seleccionar la ruta de las mercancías a través de las terminales; cada par de puntos O-D es asignado al par de terminales *hub* con menor coste de circuito, considerando únicamente terminales en la inmediata vecindad de los orígenes y destinos. Si asumimos como terminales potenciales las situadas en las esquinas de la celda de terminales que contiene el punto origen y destino de la Figura 2.2., existen 16 posibilidades de determinación de las 2 terminales de encaminamiento del flujo, de la que se escogerá el par que añada la menor distancia.

Se asume que:

- Los vehículos que llegan y abandonan la terminal se desplazan a plena carga.

- El sistema está regulado con un intervalo de envío común P_f para todas las terminales.
- Cada par de terminales *hub* está comunicado por un vehículo que se desplaza sin realizar paradas (únicamente sirviendo estas dos terminales).

En relación a los costes, se debe considerar:

- El coste de inventario no es incrementado por la segunda transferencia y se mantiene en $c_i P_f$.
- Si se define el desplazamiento entre terminales *hub* como envío de larga distancia sin paradas (*line-haul*), el coste total de larga distancia expresado en veh-km todavía es igual al kilometraje total de los vehículos dividido por la capacidad del vehículo C . Para la estrategia de ruteo, el sobre coste de distancia asociado a la visita de una terminal *hub* (2.1) es igual a la distancia añadida por la visita de 2 terminales. De este modo, el coste de circuito por unidad transportada todavía está determinado por la ecuación (2.5. b).
- Los costes de transporte local son todavía proporcionales al número de paradas, con el mismo coeficiente de proporcionalidad. El número de paradas por terminal origen o destino es igual al número de terminales que sirven a la delegación en las dos direcciones. Sin embargo, el número determinado en la ecuación (2.3) para una transferencia se reduce en este caso a $4+4=8$. De este modo, el coste de paradas por unidad transportada, asumiendo que la densidad de recogidas es aproximadamente igual a la asociada a entregas es $8\alpha_2\delta/(\lambda P_f |R_s|)$.
- El coste de inventario en tránsito puede ser despreciado en relación a los costes de transporte.
- Los costes fijos de terminal no varían de la distribución con una terminal para un valor dado de h y demanda total λ , pero en este caso los costes de terminal proporcionales al flujo son aproximadamente el doble. De este modo, la mercancía se manipula dos veces y se requiere el doble de tiempo para moverla dentro de las terminales. De esta manera el coste de terminal por unidad de mercancía es $2\alpha_5 + \alpha_6 h / |R_s|$.

En este caso, el problema táctico y estratégico se desarrolla de forma análoga al problema de minimización de la ecuación (2.7) estableciendo una restricción de conservación de flujo. En la organización de los envíos mediante programación (sincronización) y bajo los supuestos de este apartado, el número de vehículos de recogida que llegan a la terminal es h . Se debe constatar que el número de vehículos que llegan a una terminal para realizar una transferencia debe ser igual al número de vehículos de salida hacia otras terminales, que es igual a $h-1$. Debido que una fracción igual a $1/h$ de toda la carga recogida es local, el número de paradas para recoger carga de cada vehículo es h (de forma similar se calcula el proceso de distribución). De esta forma se puede calcular el número de terminales por la expresión (2.10).

$$h = \frac{\lambda |R_s|^2 P_f}{hC} \quad (2.10)$$

Si el coste asociado a la terminal es despreciable y se adopta un valor de $K=0.5$ en la expresión para α_2 , el valor óptimo de terminales y el número de paradas puede ser expresado únicamente en función de N_o y K . Si se determina un valor de $N_o > 10^2 K$, se pueden adoptar los valores de las variables de decisión definidos en las expresiones (2.11a) y (2.11b).

$$h^* \approx (8N_o)^{1/4} \left(\frac{N_o}{K} \right)^{1/2} \quad (2.11.a)$$

$$\frac{n_s^*}{\sqrt{N_o}} \approx \frac{2^{1/2} K}{N_o} \quad (2.11.b)$$

En términos generales, la máxima diferencia entre un sistema con 1 y 2 transferencias es del orden del 7% del coste de transporte de una mercancía a lo largo de una región de servicio en vehículo a carga completa.

2.1.5. Generalizaciones. Situaciones alejadas de las hipótesis de partida

En el desarrollo de esta metodología se han realizado un conjunto de hipótesis que han permitido simplificar significativamente el problema de distribución y llegar a fórmulas compactas capaces de estimar el coste logístico total ligado a un problema de transporte. Sin embargo, los problemas de distribución reales presentan una serie de particularizaciones que se alejan relativamente de las consideraciones efectuadas en la metodología propuesta por Daganzo (2005) y por tanto, precisan la adaptación o reformulación de las funciones de coste. Los principales puntos en que los problemas reales se alejan de las hipótesis efectuadas se podrían resumir en:

- Distribución no uniforme de la demanda. La ubicación de los clientes potenciales en el territorio no es homogénea ya que existen múltiples criterios de localización de las actividades o empresas solicitantes de servicios de paquetería como el potencial económico de una zona, su dotación infraestructural, etc.

- Retornos en vacío. Las configuraciones de la red de distribución eficiente desarrolladas contemplan de forma segregada el reparto y la recogida de mercancías de una zona. Sin embargo, uno de los puntos de mayor capacidad de optimización del sistema es la reducción de los retornos hacia terminales en vacío (sin carga). En Daganzo (2005) y Hall (1990) se amplía el cuerpo metodológico desarrollado identificando las principales estrategias de optimización.

2.2. MODELOS BASADOS EN PROGRAMACIÓN MATEMÁTICA Y ALGORITMOS HEURÍSTICOS ASOCIADOS

Las aproximaciones continuas de las variables del problema de distribución física desarrolladas en el capítulo anterior han permitido identificar las variables de decisión sobre las que se debe actuar para optimizar la configuración logística actual.

Sin embargo, la mayoría de modelos de localización de almacenes o terminales de consolidación de la literatura científica se han basado en problemas con variables discretas en los que se conocía exactamente la demanda de cada cliente de la zona de servicio y su ubicación. En este sentido, a diferencia de las aproximaciones continuas, existen modelos en los que se plantea una función objetivo de costes logísticos a minimizar (principalmente en función de la distancia recorrida) en los que se debe tener como información de partida la localización exacta de los puntos de demanda de servicio (delegaciones-clientes-nodos), la red física que los comunica (arcos) y la identificación detallada de los flujos de carga o rutas logísticas de transporte. Si bien estos métodos proporcionan una solución más ajustada a la realidad, presentan como un punto débil la incapacidad de traslación de la solución a otros problemas parecidos así como la falta de identificación de las variables o subsistemas estratégicos del problema.

El desarrollo de las herramientas metodológicas de localización de centros de consolidación o terminales jerárquicas en problemas discretos se han formulado como un problema de programación matemática aunque ha sido incapaz de resolver configuraciones de tamaño relativamente pequeño debido a la complejidad del problema (Tabla 2.1). Dentro de la optimización combinatoria, esta tipología de problemas es *NP-Hard* por lo que en muchos casos ha sido necesario el desarrollo de algoritmos heurísticos de solución que aunque no proporcionen una solución óptima, tienen un tiempo computacional acotado.

Tabla 2.1. Resumen de las contribuciones científicas en la localización de hubs

Autor	Etapas del modelo	Hubs	Capacidad de hubs	Dominio de soluciones en Localización	Estrategias de envío contempladas	Método de solución	Costes considerados	Observaciones
Perl (1985)	Localización de hubs, determinación tamaño de hubs y ruteo	Número indeterminado, tamaño indeterminado	ilimitada	Discreto (dominio de soluciones mayor número de hubs determinado)	One-to-many	Localización y tamaño de hubs con heurísticos.	Costes unitarios de Transporte fijos y costes de inversión en almacenes	
Campbell (1986)	Localización de hubs	Número determinado	ilimitada	Continuo	Hub más cercano	Análítico	Costes de Transporte constantes	Variación temporal de la demanda
O'Kelly (1987)	Localización de hubs y asignación de flujos a la red	Número determinado	ilimitada	Discreto	Todos los envíos al través de hub (envío directo no permitido)	Localización: Programación entera cuadrática que se resuelve con algoritmos heurísticos. Asignación: 1 nodo al hub más cercano.	Costes iguales a la distancia recorrida	
Aykin (1990)	Asignación de flujos a la red	Número y localización determinada	ilimitada	Discreto	Todos los envíos al través de hub (envío directo no permitido)	Asignación: 1 nodo al hub de menor coste (no menor distancia)	Costes unitarios de Transporte fijos	
Kılınçcız (1991)	Asignación de flujos a la red	Número y localización determinada	ilimitada	Discreta		Heurísticos de asignación por distancia o multicriterio	Costes unitarios de Transporte fijos	Parte del modelo de O'Kelly 1987
Aykin y Brown (1992)	Localización de hubs y asignación de flujos a la red	Desconocidos	ilimitada	Continua	Todos los envíos al través de hub (envío directo no permitido)	Asignación: 1 nodo al hub de menor coste (no menor distancia)	Costes unitarios de Transporte fijos	Se considera un flujo simétrico
Aykin (1994)	Localización de hubs	Número determinado	ilimitada	Discreto (dominio de soluciones mayor número de hubs determinado)	Envío directo o al través de 1 o 2 hubs	Programación matemática entera y relajación Lagrangeana	Costes unitarios de Transporte fijos	
Aykin (1995a)	Localización de hubs y asignación de flujos a la red	Número determinado		Continua	Envío directo (fijados de antemano) o al través de 1 o 2 hubs	Programación matemática entera	Costes unitarios de Transporte fijos	
O'Kelly (1986)	Localización de hubs y asignación de flujos a la red	Número determinado	ilimitada	Discreto	Todos los envíos al través de 2 hubs	Asignación: 1 nodo a más de un hub	Costes unitarios de Transporte fijos y costes de inversión en almacenes	Comparativa de escenarios con un número menor de hubs con asignación múltiple a escenarios con más hubs sin asignación múltiple
O'Kelly (1998)	Localización de hubs y asignación de flujos a la red	Número determinado	ilimitada	Discreto	Todos los envíos al través de 2 hubs	Asignación: 1 nodo a más de un hub	Costes de Transporte variables con el flujo (no lineales)	
Ernst (1998)	Localización de hubs y asignación de flujos a la red	Número determinado	ilimitada	Discreto (dominio de soluciones mayor número de hubs determinado)	Envíos a través de un hub	Programación matemática no lineal. Método del branch & bound y heurísticos (SA y descenso aleatorio)	Costes unitarios de Transporte fijos	Comparativa de métodos para la distribución de servicios postales en Australia
Abdaimour-Helun (1999)	Localización de hubs y asignación de flujos a la red	Número indeterminado	ilimitada	Discreta (entre nodos)	Envíos a través de un hub	Metodología del Branch & Bound y algoritmos genéticos	Costes unitarios de Transporte fijos y costes de inversión en almacenes	
Wasner (2004)	Localización de hubs y ruteo	Número indeterminado, tamaño determinado	ilimitada	Continua	Envíos directos o al través de hubs	Localización y asignación mediante algoritmo heurístico	Costes unitarios de transporte entre hubs y del reparto fijos, costes de instalaciones no lineales con el flujo	Propuesta teórica más caso práctico de aplicación en Austria. Determina los límites de las zonas de reparto y entrega de cada ruta

Los conceptos básicos que permiten una comparación de las prestaciones diferenciales de cada modelo se basan en los siguientes puntos:

- **Etapas del modelo:** Los modelos desarrollados se basan en la resolución de distintas fases de los sistemas de distribución física: el proceso de localización de las terminales, la asignación a la red (entre terminales) de la mercancía asociada a cada par origen-destino y el proceso de determinación de rutas desde las terminales secundarias hasta los clientes. Los modelos presentados se focalizarán en una de estas fases o podrán dar solución a más de una fase a la vez.
- **Hubs:** El sistema logístico considerado puede partir de un número y localización de *hubs* prefijados a partir de los cuales se analiza el proceso de asignación de flujos a la red. Otra posibilidad es determinar de antemano el número de *hubs* que se utilizarán para realizar la distribución y calcular su localización óptima en el territorio.
- **Capacidad de hubs:** Un aspecto básico es la consideración de la capacidad de las terminales para dar servicio a toda la mercancía que se clasifica y se encamina a través de ella. La mayoría de modelos no consideran la capacidad de las instalaciones, y únicamente unos pocos llegan a definir un sistema capacitado.
- **Dominio de soluciones en localización:** Los modelos que incluyen la etapa de localización de las terminales en su desarrollo pueden permitir que éstas se efectúen en cualquier punto del territorio (continua) u obligar que la ubicación de las terminales se efectúe en un número de puntos singulares fijados de antemano (localización discreta en nodos especiales, delegaciones, puntos cercanos al mercado o a infraestructuras, etc.).
- **Estrategias de envío contempladas:** Los análisis científicos de los autores determinan distintas posibilidades de encaminamiento de la carga que se diferencian en envíos exclusivamente a través de *hubs* o la simultaneidad de envíos directos y a través de *hubs* (se suele fijar el número máximo de 2 paradas en *hubs*).
- **Métodos de solución:** Uno de los puntos más importantes de diferenciación de los distintos modelos que se han analizado ha sido las metodologías de resolución de los problemas. La mayor parte de los modelos han desarrollado la programación matemática del problema logístico expresada como una función objetivo de costes a minimizar sujeta a unas restricciones y limitaciones. Sin embargo, la resolución del problema en programación matemática únicamente se ha podido desarrollar en un tiempo computacional asumible en problemas de tamaño limitado. Este hecho ha obligado a aplicar una relajación de las variables del problema para poder determinar soluciones a

problemas de mayor tamaño o la adopción de técnicas heurísticas que si bien no aseguran que la solución sea óptima, realizan los cálculos en un tiempo acotado. Por último, las contribuciones científicas al problema más recientes han desarrollado algoritmos metaheurísticos o heurísticos combinados con técnicas de *fine-tuning*, que a partir de una configuración existente sub-óptima realizan un proceso de refinamiento y optimización de la solución.

- **Costes considerados:** La inclusión de los distintos elementos de transporte varía en los modelos analizados. En este sentido, existen unos modelos que únicamente consideran los costes de transporte proporcionales a la distancia y otro grupo que incluye los costes de las instalaciones fijas (manipulación de la mercancía y amortización del coste de la instalación). Sin embargo, existe otro grupo que considera que los costes de transporte varían linealmente con el flujo de la mercancía transportada en cada arco.

2.2.1. Programación matemática y heurísticos para la localización de terminales *hub*

En O’Kelly (1987) se plantea una programación entera cuadrática para la resolución de problemas de localización de *hubs* basado en los flujos entre orígenes y destinos. El problema viene definido por N_T puntos de servicio (nodos), el flujo W_{ij} entre cada par de nodo terminal (i,j) , el coste unitario de transporte c_{ij} específico para cada par de nodos y un número h de centros de consolidación que deben ser ubicados en el territorio ($h < N_T$). Por otro lado, el coste de transporte entre dos terminales que sean *hubs* se multiplica por un parámetro a ($0 < a < 1$) para reflejar las economías de escala existentes debido a los menores costes unitarios al utilizar vehículos de mayor capacidad. Paralelamente, se define una variable de decisión X_{ik} a la que se da valor igual a la unidad si el nodo i es servido a través del *hub* k y valor igual a 0 en caso contrario. Adicionalmente se considera $X_{ii}=1$ si el nodo i es un *hub* y 0 en caso contrario. Se puede comprobar que, en un caso general, todo punto debe estar servido por un *hub*, no permitiéndose el envío directo entre dos puntos de servicio. La solución del problema discreto recae en encontrar localizaciones de *hubs* y la asignación de los nodos al servicio de los *hub* que minimicen el coste total de transporte. La función objetivo del problema se detalla en la ecuación (2.12).

$$\text{Min } Z = \sum_i \sum_k \sum_m \sum_j W_{ij} [X_{ik} X_{jm} (c_{ik} + ac_{km} + c_{jm})] \quad (2.12)$$

sujeto a:

$$\begin{aligned} \sum_i X_{ik} &\leq (N_T - h + 1)X_{kk} \quad \forall k \\ \sum_k X_{ik} &= 1 \quad \forall i \\ \sum_k X_{kk} &= h \\ 0 \leq X_{ik} &\leq 1 \quad X_{ik} \text{ variable entera} \end{aligned}$$

La herramienta propuesta por O’Kelly (1987) para resolver la ecuación (2.12) son algoritmos heurísticos que deben ser utilizados sistemáticamente para determinar la asignación de los diferentes nodos al conjunto de *hubs*. Se plantean dos algoritmos heurísticos (HEUR1 y HEUR 2) que realizan una enumeración completa de todas las configuraciones posibles del problema. El problema considerado se caracteriza porque los *hubs* no tienen interacción entre sí y por que se adopta una asignación basada en que cada nodo es servido por el *hub* situado más próximo.

En Aykin (1990) y Aykin y Brown (1992) se analiza críticamente el método propuesto en O’Kelly. En primer lugar, O’Kelly plantea la evaluación de todas las posibles soluciones de asignar los N_T nodos (delegaciones) del problema a los h *hubs*. De forma particular, como considera en una primera aproximación dos *hubs*, existen $N_T(N_T-1)/2$ posibilidades de asignación a tener en cuenta. Este hecho incrementa significativamente el tiempo computacional en el caso de encontrarnos en problemas de gran tamaño y en problemas con un número significativo de *hubs*.

Adicionalmente, el método propuesto por O’Kelly propone una partición de los nodos en h grupos de forma que las subregiones no se solapen como criterio de optimalidad ya que cada nodo se asigna al *hub* más próximo. Sin embargo, puede haber problemas en que el método de O’Kelly de asignación de nodos terminales por criterios de proximidad a los *hubs* no contenga la configuración óptima, caracterizada por uno o más nodos no asignados a su *hub* más próximo.

De este modo, en Aykin (1990) se plantea una solución particular de la ecuación (2.12), definida por el conjunto de nodos delegaciones Z_r que son asignados al *hub* r : $Z_r = \{j: X_{jr} = 1\}$, $r = (N_T - h + 1), \dots, N_T$. El cambio de asignación de un nodo delegación cualquiera i que actualmente es servido por el *hub* t ($i \in Z_t$) a otro *hub* k produce una diferencia en el valor de la función objetivo que se cuantifica por la expresión (2.13).

$$2 \left[\sum_{j \in Z'_k} W_{ij} (C_{ik} + C_{jk}) + \sum_{h \in Z'_i} W_{ih} (C_{ik} + aC_{kt} + C_{ht}) + \sum_{r \neq k, \neq t, l \in Z_r} \sum W_{il} (C_{ik} + aC_{kr} + C_{lr}) \right] - \quad (2.13)$$

$$- 2 \left[\sum_{h \in Z_i} W_{ih} (C_{it} + C_{ht}) + \sum_{j \in Z_k} W_{ij} (C_{jk} + aC_{kt} + C_{it}) + \sum_{r \neq k, \neq t, l \in Z_r} \sum W_{il} (C_{it} + aC_{tr} + C_{lr}) \right]$$

donde

$$Z'_k = \{j : j \in Z_k \text{ o } j = i\} \quad \text{y} \quad Z'_i = \{j : j \in Z_i \text{ y } j \neq i\}$$

De este modo, el valor de la función objetivo puede ser mejorado si se reasigna el nodo a servir i al *hub* k , $k \neq t$, cuando se cumple la expresión (2.14).

$$\Delta S = \sum_r \sum_{j \in Z_k} W_{ij} (C_{ik} - C_{it}) + \left(\sum_{j \in Z_i} W_{ij} - \sum_{j \in Z_k} W_{ij} \right) aC_{kt} + \left(\sum_{r \neq k \neq t} \sum_{j \in Z_r} W_{ij} a(C_{kr} - C_{tr}) \right) \leq 0 \quad (2.14)$$

Si los nodos son asignados a los *hubs* en base a su proximidad, no se considera el flujo de objetos entre *hubs*. De este modo, si se define una constante positiva e , la condición anterior se puede reformular asignando el nodo i al *hub* k en lugar de t cuando se cumpla la expresión (2.15).

$$e(C_{ik} - C_{it}) \leq 0 \quad (2.15)$$

Para el caso de estudio de interacción entre *hubs*, en Aykin (1990) y Aykin y Brown (1992) se plantea una metodología de solución que determina la asignación de un nodo i a un *hub* k cuando se cumpla la ecuación (2.16) equivalente a que el valor A_{ik} calculado por la expresión (2.17) sea el mínimo global del sistema.

$$A_{ik} = \min_{1 \leq t \leq p} \{A_{it}\} \quad (2.16)$$

$$A_{it} = \left(\sum_r \sum_{j \in Z_r} W_{ij} \right) C_{it} + a \left(\sum_{r \neq t} \sum_{j \in Z_r} W_{ij} C_{tr} \right) \quad t = 1, \dots, h \quad (2.17)$$

La ecuación (2.17) resuelve el problema de asignación de nodos para una localización de *hubs* dada (problema discreto). Sin embargo, se pretende afrontar a la vez la determinación de una localización de *hubs* óptima con una asignación óptima de nodos terminales a estos *hubs*. Por lo tanto, en Aykin y Brown (1992) se analiza el problema inverso, en el que dada una asignación de los distintos nodos delegaciones a *hubs* (agrupación de nodos por zonas

asociadas a un *hub*) se pretende determinar la localización óptima de éstos. En este sentido, para determinar el problema de localización se reformula la función objetivo (2.12) y se establece la ecuación (2.18) a minimizar.

$$\min \left(\sum_i \sum_k \left(2X_{ik} \sum_j W_{ij} \right) C_{ik} + \sum_k \sum_t \left(a \sum_i \sum_j W_{ij} U_{ik} U_{jt} \right) C_{kt} \right) \quad (2.18)$$

Si se asume que el flujo entre nodos es simétrico, la ecuación (2.18) se puede simplificar y obtener la ecuación (2.19), donde $G_{ik} = 2U_{ik} \sum_j W_{ij}$, $L_{kt} = 2a \sum_i \sum_j W_{ij} U_{ik} U_{jt}$.

$$\min \left(\sum_i \sum_k (G_{ik} C_{ik}) + \sum_{1 \leq k < t \leq p} L_{kt} C_{kt} \right) \quad (2.19)$$

Se puede comprobar que los valores de G_{ik} y L_{kt} son conocidos si se dispone de la asignación de nodos a los distintos *hubs*. De esta forma se plantea un procedimiento iterativo para determinar el vector de coordenadas Q_k del *hub* k utilizado para resolver el problema de localización de multiplicidad de almacenes o terminales en distancia Euclídea, mediante la formulación de la expresión (2.20).

$$Q_k^{h+1} = \frac{\sum_{t \neq k} (L_{kt} Q_t^h / C_{kt}^h) + \sum_i (G_{ik} P_i / C_{ik}^h)}{\sum_{t \neq k} (L_{kt} / C_{kt}^h) + \sum_i (G_{ik} / C_{ik}^h)} \quad k = 1, \dots, h \quad (2.20)$$

Los superíndices de la ecuación (2.20) representan el número de la iteración en la que se encuentra el proceso iterativo.

Con todo, Aykin y Brown (1992) formulan un algoritmo para solucionar el problema global de localización de *hubs* y asignación de nodos terminales a los *hubs*. Este algoritmo resuelve el subproblema de localización de *hubs* y el de asignación de forma separada. La asignación a *hubs* se realiza mediante la ecuación (2.16) analizando únicamente la ubicación de los distintos nodos terminales del problema. De esta forma la interrelación entre *hubs* se mantiene constante y únicamente se necesita determinar el coste asociado a los nodos existentes. Las localizaciones de los *hubs* son sistemáticamente actualizadas mediante la ecuación (2.20). El algoritmo propuesto para resolver el problema de localización y asignación procede de la siguiente forma:

- Paso 0. Se designa una localización inicial de *hubs* con los vectores de posición Q_t^0 , $t=1, \dots, h$. Se fija $b=0$ y un número escalar de terminación del proceso iterativo $v \geq 0$. Por otro lado, se determina una asignación inicial X_{it}^0 , $i=1, \dots, N_T$ y $t=1, \dots, h$.
- Paso 1. Se reasignan los nodos terminales a los *hubs* utilizando la ecuación (2.16). Se repetirá la aplicación de (2.16) hasta que no se consigan mejores cambios de asignación en los X_{it}^{b+1} , $i=1, \dots, N_T$ y $t=1, \dots, h$.
- Paso 2. Se aplicará la ecuación (2.20) una vez para cada *hub* para actualizar la localización de los *hubs*, Q_t^{b+1} , $t=1, \dots, h$.
- Paso 3. Parar si $d(Q_t^{b+1}, Q_t^b) \leq v$ para todo *hub* t y $X_{it}^{b+1} = U_{it}^b$ para todo i y t . En caso contrario, fijar $b=b+1$ y dirigirse al paso 1.

Heurísticos para el problema de localización de *h-hubs*

En Klinecicz (1991) se analiza el problema de la localización de h *hubs* dentro de un conjunto de posibles soluciones ya definidas (identificación de las delegaciones candidatas a ser un *hub*). Se utilizan algoritmos heurísticos para su posible comparación sin la pretensión de encontrar la solución óptima dada la dificultad del problema.

El punto de partida del problema es el mismo definido en O’Kelly (1987). Si se evalúa la suma total del flujo generado o atraído por un nodo terminal genérico i ($O_i = \sum_j W_{i,j}$, $D_i = \sum_j W_{j,i}$), la función objetivo (2.12) se puede reformular por la ecuación (2.21).

$$\text{Min } Z = \sum_i \sum_k X_{ik} C_{ik} (O_i + D_i) + \sum_j \sum_m X_{jm} \sum_i \sum_k [X_{ik} a W_{ij} C_{km}] \quad (2.21)$$

En la referencia anterior se analizan distintos métodos de asignación de cada nodo a la región servida por cada *hub*:

a) Asignaciones basadas en la distancia

A nivel general, se asume que los costes por unidad de mercancía transportada son proporcionales a la distancia y por tanto se puede hablar indistintamente de asignación por distancia o por costes. Un ejemplo de este tipo de asignaciones es el propuesto en el heurístico propuesto en O’Kelly (1987) donde se enumeran todas las posibles configuraciones de *hubs*, asignando los nodos a un *hub* en base al criterio de mínima distancia. Sin embargo, este

procedimiento es computacionalmente inabordable para aquellos problemas con un número de nodos significativo.

b) Asignaciones multicriterio

Para reducir la cantidad de mercancía a transportar entre *hubs*, aquellos pares de nodos que tengan entre ellos un flujo importante de objetos, deben ser transportados por el mismo *hub*. Por lo tanto, se debe incluir un criterio de asignación que contemple esta estrategia en el proceso. En este sentido se define un criterio de asignación multicriterio que integre un criterio de tráfico entre nodos y el criterio de distancia o proximidad. Se debe tener en cuenta en este proceso que los h *hubs* ya están localizados en la zona de servicio y que únicamente el problema reside en asignar los nodos a los *hubs* ya conocidos. La asignación se realiza por una suma ponderada del tráfico del *hub* k (mercancía de las delegaciones ya asignadas a k) y el valor de la variable definida por el inverso de la distancia del nodo i al *hub* k , de forma que la delegación más próxima tenga el máximo valor de esta variable. En todos los casos, el valor de la variable de tráfico y de la distancia más próxima al *hub* se normalizará de forma que el valor máximo sea igual a 1.

La asignación por distancia tiende a minimizar el coste de transporte entre los nodos y el *hub* mientras que la medida de tráfico tiende a minimizar el coste de transporte de los desplazamientos entre *hubs*. La asignación sigue un proceso iterativo, en el que hay una asignación grosera en las primeras etapas que posteriormente se va reevaluando para reducir el valor de la función objetivo.

Heurísticos de intercambio

Las técnicas descritas anteriormente evalúan todas las posibles configuraciones de asignación de nodos a los h *hubs* determinados. Sin embargo, se pueden plantear algoritmos heurísticos que únicamente examinen configuraciones que potencialmente produzcan alguna mejora sobre la mejor solución actual encontrada en el proceso de optimización.

En este caso se permite el cambio de las delegaciones que actúan como *hub*. Las únicas configuraciones que se evalúan son aquellas que difieren en la localización de uno o dos *hubs* en relación a la mejor solución actual adoptada. La decisión de intercambiar nodos que actúen como *hubs* se basa en una variable de mejora local. Esta variable local determina el ahorro en la función objetivo si algún nodo $q \notin H$ reemplaza simplemente otro nodo $r \in H$ como *hub* (o $q_1, q_2 \notin H$ reemplazan r_1 y $r_2 \in H$), manteniendo la misma asignación para el resto de nodos. Se parte de la evaluación del tráfico total *interhub* para la solución actual. El tráfico desde el *hub* k al *hub* l queda determinado por la expresión (2.22).

$$I_{kl} = \sum_{\{i|X_{ik}=1\}} \sum_{\{j|X_{jl}=1\}} W_{ij} \quad (2.22)$$

Si se asume que H^0 es el conjunto de *hubs* de la configuración de la solución actual, los ahorros producidos por el reemplazo del nodo r por el nodo q se puede evaluar por la expresión (2.23).

$$R_{qr} = \sum_{\{i|X_{ir}=1\}} O_i(C_{ir} - C_{iq}) + \sum_{\{i|X_{ir}=1\}} D_i(C_{ri} - C_{qi}) + \sum_{\substack{k \in H^0 \\ k \neq r}} I_{rk} a(C_{rk} - C_{qk}) + \sum_{\substack{k \in K^0 \\ k \neq r}} I_{kr} a(C_{kr} - C_{kq}) \quad (2.23)$$

El reemplazo de un nodo que actúe como *hub* por otro implica que se debe aplicar un algoritmo de asignación de nodos a los *hubs* en cada iteración de cambio.

Heurísticos de zonificación-agrupación

Los heurísticos de intercambio y el de enumeración propuesto por O’Kelly (1987) determinan en primer lugar h puntos potenciales de funcionar como *hub* y posteriormente asignan los restantes nodos al servicio de cada *hub*. Sin embargo, los algoritmos de zonificación-agrupación en primer lugar establecen h grupos de nodos y posteriormente escogen la localización del nodo de cada grupo que actuará como *hub*.

Se destaca el estudio realizado por Monma y Sheng (1986) en el que se propone un método de asignación de los nodos a las h particiones del territorio (*clusters*) basado en la asignación multicriterio desarrollada anteriormente. En este caso, debido que las localizaciones de los *hubs* no se conocen de antemano, la medida del criterio de distancia del heurístico de zonificación se realiza en el centro de gravedad de cada partición (*cluster*). Conociendo las coordenadas de cada nodo i , el centro de gravedad de la partición l tiene las coordenadas calculadas por la expresión (2.24). El centro de gravedad minimiza las distancias al cuadrado pero se considera una aproximación suficiente del problema de las medianas.

$$(X_G, Y_G) = \left(\frac{\sum_{i \text{ en cluster } l} (O_i + D_i)x_i}{\sum_{i \text{ en cluster } l} (O_i + D_i)}, \frac{\sum_{i \text{ en cluster } l} (O_i + D_i)y_i}{\sum_{i \text{ en cluster } l} (O_i + D_i)} \right) \quad (2.24)$$

El algoritmo se basa en los siguientes pasos:

- Paso 0. Inicialización. Se empieza con p nodos llamados semillas para las h particiones realizadas, por ejemplo los nodos con mayor O_i+D_i , determinando unos pesos de la asignación por distancia w_1, w_2 tal que $w_1+w_2=1$.
- Paso 1. Asignación. Se repite para cada nodo i el siguiente proceso:
 - Para cada *cluster* k , se calcula el inverso de la distancia des del nodo i hasta el centro de gravedad de k . Se normalizan las distancias inversas que se expresan como D_{ik} para $k=1, \dots, h$.
 - Para cada *cluster* k , se calcula el tráfico total intercambiado con los nodos ya presentes en k ($\sum_{j \text{ en cluster } k} W_{ij} + W_{ji}$), normalizando la medida para que el máximo sea igual a la unidad y denominándolos T_{ik} , para $k=1, \dots, h$.
 - Para cada *cluster* k , calcular $G_{ik}=w_1D_{ik}+w_2T_{ik}$
 - Asignar i al *cluster* con un valor máximo de G_{ik} .
- Paso 2. Estudio de asignación. Para cada nodo i se debe realizar el siguiente cálculo: Para cada *cluster* k , se calculan los ahorros S_{ik} que resultarían si i es reasignado a k en lugar del *cluster* actual, asumiendo que el *hub* estaría localizado en el centro de gravedad.
- Paso 3. Adopción del mejor candidato para la reasignación. Se escoge el máximo S_{ik} de entre todos los nodos i y todos los *clusters* k .
- Paso 4. Reasignación. Si $S_{ik}>0$, se reasigna el nodo i al *cluster* k y se repite el paso 2. En caso contrario pasar al punto 5.
- Paso 5. Elección de *hubs*. Para cada *cluster* se escoge el nodo más cercano al centro de gravedad para ser el *hub* de la partición. De este modo, se asignarán todos los nodos de aquella partición a su *hub*.

En Klincewicz (1991) se realiza adicionalmente un análisis comparativo de los resultados de aplicación de los heurísticos detallados a un conjunto de problemas. Una parte de estos problemas derivan de configuraciones de aeropuertos de EE.UU. y otras de ejemplos hipotéticos de distribuciones logísticas. A nivel general, los mejores resultados se obtienen con los heurísticos de enumeración, sin embargo el gran requerimiento de tiempo computacional para llegar a la solución lo desaconsejan para problemas de gran tamaño. En particular, para configuraciones superiores a los 25 nodos y 4 *hubs* ya no se opta por este cálculo. Para casos que deben ser solucionados repetidamente o para problemas de gran tamaño se propone el heurístico de doble intercambio como técnica de solución para la localización de *h-hubs*. La calidad de la solución es buena y el tiempo computacional se mantiene en valores asumibles.

Adicionalmente se ha realizado un análisis de sensibilidad de los pesos w_1 y w_2 utilizados en el algoritmo de intercambio de asignación multicriterio y el de asignación de *clusters*. Se

constata que no hay un dominio de aplicación general de valores de w_1 y w_2 que produzca soluciones casi-óptimas, si bien los valores $(w_1;w_2)=(0,7;0,3)$ producen soluciones aceptables.

2.2.2. Programación matemática para la localización de terminales *hub* y asignación de envíos a rutas

En Aykin (1995a) se analiza el problema de la localización de *hubs* y de rutas en el plano \mathcal{R}^2 permitiendo que los flujos de carga a transportar entre dos puntos se consoliden en uno o dos *hubs* o que sean entregados directamente sin realizar ninguna parada para la consolidación de carga. En este problema se considera que las rutas en las que se permite el envío directo (sin pasar por ningún *hub*) son conocidas de antemano.

En este contexto el problema se basa en determinar el tipo de servicio (directo, con parada en un *hub* o en dos) entre los puntos de demanda y la localización de cada *hub*, de forma que se minimicen los costes de transporte. Se propone un algoritmo de resolución basado en 4 métodos de determinación de la solución inicial. La localización de los *hubs* se permite en todo el espacio de soluciones.

Se consideran las siguientes variables y parámetros:

$P_i \in \mathcal{R}^2$, coordenadas de localización del punto de demanda i ($i=1, \dots, N_T$),

c_{ik} , coste de transporte directo de i a k por unidad de flujo y distancia ($\text{€/m}^3\text{-km}$)

W_{ij} , flujo total entre i y j

$Q_k \in \mathcal{R}^2$, coordenadas de localización del *hub* k ($k=1, \dots, h$)

$a_1, a, a_2 \in [0,1]$, constantes de proporcionalidad que determinan respectivamente las economías de escala para los segmentos de la cadena de transporte origen de ruta-*hub*, conexión *hub-hub* y *hub-final* de ruta respectivamente ($a \leq a_1, a_2$)

F_D , conjunto de rutas (i,j) en las que es permitido un servicio directo sin paradas en *hub*. En este sentido, se define la variable discreta O_{ij} que toma valores iguales a 1 si $(i,j) \in F_D$ y 0 en caso contrario.

Las variables de decisión que determinan si un envío es realizado a través de uno o dos *hubs* se detallan en las expresiones (2.25) y (2.26).

$$X_{ij} = \begin{cases} 1 & \text{si el flujo de } i \text{ a } j \text{ es enviado directamente} \\ 0 & \text{en caso contrario} \end{cases} \quad (2.25)$$

$$X_{iktj} = \begin{cases} 1 & \text{si el flujo de } i \text{ a } j \text{ es enviado de } i \rightarrow \text{hub } k \rightarrow \text{hub } t \rightarrow j \\ 0 & \text{en caso contrario} \end{cases} \quad (2.26)$$

La variable de decisión X_{iktj} es definida para los puntos de demanda $i, j=1, \dots, N_T$, $i \neq j$, y para los *hubs* $k, t=1, \dots, h$. El caso que $t=k$, representa que el envío únicamente realiza una parada en el *hub* k . En esta situación, el problema de transporte se puede formular de la forma siguiente:

Problema P1

$$\begin{aligned} \text{Min } Z = & \sum_i \sum_j O_{ij} W_{ij} c_{ij} d(P_i, P_j) X_{ij} + \\ & + \sum_i \sum_k \sum_t \sum_j W_{ij} (a_1 c_{ik} d(P_i, Q_k) + a c_{kt} d(Q_k, Q_t) + a_2 c_{tj} d(Q_t, P_j)) X_{iktj} \end{aligned} \quad (2.27)$$

sujeto a:

$$O_{ij} X_{ij} + \sum_k \sum_t X_{iktj} = 1 \quad \forall i, j \quad (2.28)$$

$$Q_k \in \mathfrak{R}^2 \quad \forall k, X_{ij}, X_{iktj} \in [0, 1] \quad (2.29)$$

El primer término de la función objetivo (2.27) representa el coste de los envíos directos, mientras que el segundo identifica el coste de los envíos realizados a través de uno o dos *hubs*. Si se define el número total de pares de puntos en el que se permite un envío directo por $|F_D| = d^*$, el número total de variables con valor 0-1 utilizadas en el problema es de $h^2 N_T (N_T - 1) + d^*$ y el número de variables de decisión continuas es de $2h$. Este hecho implica que para valores razonablemente altos de puntos de servicio y *hubs* la resolución del problema implica un tiempo computacional prohibitivo.

Sin embargo, Aykin (1995a) plantea que el problema se puede descomponer en un número de subproblemas de caminos mínimos que implican decisiones de servicio si se permite fijar la localización del *hub*. En este sentido, si la localización de los *hubs* es conocida, el problema se reduce a resolver con programación matemática entera las variables de decisión sobre tipo de servicio (envío directo o a través de *hub*):

Problema P2

$$Min Z = \sum_i \sum_j O_{ij} W_{ij} C_{ij} X_{ij} + \sum_i \sum_k \sum_t \sum_j W_{ij} C_{iktj} X_{iktj} \quad (2.30)$$

sujeto a:

$$O_{ij} X_{ij} + \sum_k \sum_t X_{iktj} = 1 \quad \forall i, j \quad (2.31)$$

$$X_{ij}, X_{iktj} \in [0,1] \text{ variables enteras } \quad \forall i, j, k, t \quad (2.32)$$

donde $C_{ij} = c_{ij}d(P_i, P_j)$ y $C_{iktj} = (a_1 c_{ik}d(P_i, Q_k) + a c_{kt}d(Q_k, Q_t) + a_2 c_{tj}d(Q_t, P_j))$. Conociendo la posición de los *hubs*, la determinación de la distancia entre puntos de demanda y *hubs* es directa. La solución para el problema P2 se puede obtener descomponiéndolo en problemas de caminos mínimos de menor tamaño, uno para cada ruta y posteriormente escoger el servicio de menor coste según las variables discretas definidas en (2.33a) y (2.33b).

$$X_{ij} = \begin{cases} 1 & \text{si } (i, j) \in F_D \text{ y } C_{ij} = \min_{u,v} \{C_{ij}, C_{iuvj}\} \\ 0 & \text{en caso contrario} \end{cases} \quad (2.33a)$$

$$X_{iktj} = \begin{cases} 1 & \text{si } (i, j) \in F_D \text{ y } C_{iktj} = \min_{u,v} \{C_{ij}, C_{iuvj}\} \\ 1 & \text{si } (i, j) \notin F_D \text{ y } C_{iktj} = \min_{u,v} \{C_{iuvj}\} \\ 0 & \text{en caso contrario} \end{cases} \quad (2.33b)$$

De esta forma, dada una solución del problema P2, con un conjunto de rutas directas $S_{oo} = \{(i, j) : X_{ij} = 1\}$ y un conjunto de rutas entre *hubs* $S_{kt} = \{(i, j) : X_{iktj} = 1\}$, la función objetivo del Problema P1 puede describirse mediante la ecuación (2.34).

$$Min Z = \sum_{(i,j) \in S_{oo}} W_{ij} c_{ij} d(P_i, P_j) + \sum_k \sum_t \sum_{(i,j) \in S_{kt}} W_{ij} (a_1 c_{ik} d(P_i, Q_k) + a c_{kt} d(Q_k, Q_t) + a_2 c_{tj} d(Q_t, P_j)) \quad (2.34)$$

Analizando la ecuación anterior se deduce que el primer término es constante ya que únicamente representa las rutas directas. De este modo, si se define μ_w como el primer término, y se reescribe el segundo término, el problema P1 se puede reescribir mediante las expresiones (2.35) y (2.36).

Problema P3

$$\text{Min } Z = \sum_k \sum_t V_{kt} d(Q_k, Q_t) + \sum_i \sum_k U_{ik} d(P_i, Q_k) + \mu_W \quad (2.35)$$

donde:

$$V_{kt} = a \sum_{(i,j) \in S_{kt}} c_{kt} W_{ij} \quad \text{y} \quad U_{ik} = a_1 \sum_t \sum_{(i,j) \in S_{kt}} c_{ik} W_{ij} + a_2 \sum_t \sum_{(j,i) \in S_{ik}} c_{ki} W_{ji} \quad (2.36)$$

El problema P3 es del tipo de localización de diversas instalaciones con localizaciones desconocidas de los *hubs* $Q_k \in E^2, k=1, \dots, h$. Para solucionar el problema P3 en métrica Euclídea, se propone un método iterativo basado en el procedimiento de Weiszfeld detallado en Ballou (1991) que utiliza la ecuación (2.37) para determinar las coordenadas Q_k^{b+1} de la terminal k en la iteración $b+1$.

$$Q_k^{b+1} = \frac{\sum_{t \neq k} \frac{(V_{kt} + V_{tk})}{\|Q_k^b - Q_t^b\|} Q_t^b + \sum_i \frac{U_{ik}}{\|Q_k^b - P_i\|} P_i}{\sum_{t \neq k} \frac{(V_{kt} + V_{tk})}{\|Q_k^b - Q_t^b\|} + \sum_i \frac{U_{ik}}{\|Q_k^b - P_i\|}} \quad k = 1, \dots, h \quad (2.37)$$

En este sentido, en Aykin (1995a) se propone el algoritmo *Location-Routing* basado en un procedimiento iterativo a partir de una solución inicial que se puede resumir en los siguientes pasos:

- **Paso 0.** Se determina un conjunto de localizaciones iniciales de *hubs* $\{Q_1^0, \dots, Q_p^0\}$ y unas tolerancias $\nu, \varepsilon > 0$ y $b = 0$.
- **Paso 1.** Dadas las localizaciones de *hubs* $\{Q_1^b, \dots, Q_p^b\}$, resolver el problema P2 utilizando (2.33a) y (2.33b) para determinar S_{00}^b y $S_{kt}^b \forall k, t$. Si $b=0$, ir al paso 2. Si $b>0$ y $S_{00}^b = S_{00}^{b-1} \forall k, t$, entonces parar.
- **Paso 2.** Utilizando S_{00}^b y S_{kt}^b del paso 1, reducir el problema al Problema 3. Establecer $r=1$ y $Q_k^r = Q_k^b, k=1, \dots, h$.
- **Paso 3.** Calcular $Q_k^{(r+1)}, k=1, \dots, h$, utilizando (2.37).
- **Paso 4.** Si $\|Q_k^{(r+1)} - Q_k^r\| > \nu$ para algún k , establecer $r=r+1$ e ir al paso 3. En caso contrario $Q_k^{(b+1)} = Q_k^{(r+1)}, k=1, \dots, h, b=b+1$, e ir al paso 1.

Los procedimientos mencionados para los subproblemas P2 y P3 garantizan una solución óptima. Sin embargo, la solución encontrada con el algoritmo *Location-Routing* no garantiza la solución óptima del problema P1 ya que las localizaciones de *hubs* y la estructura de la ruta no se determinan conjuntamente, por lo que el resultado final depende de la solución inicial de partida.

2.2.3. Extensión de los modelos de programación matemática

Capacidad limitada

En Aykin (1994) se plantea el problema de localización de servicios e infraestructuras (*hubs*) de capacidad limitada con interacción entre *hubs*, además de admitir una política de envío directo entre puntos de servicio y políticas de parada múltiple en uno o más *hubs*. Los *hubs* se asumen que tienen una capacidad limitada de encaminar los flujos hacia su destino. De este modo, el problema se basa en determinar para cada par de delegaciones el tipo de servicio (directo, vía *hub*) y la localización de los *hubs* de forma que se minimice el coste total de la distribución. Si Y_i es una variable entera (0-1) que toma el valor de 1 cuando la terminal i es *hub* y 0 en caso contrario, se define la variable $r_k(Y_i, Y_t, Y_j)$ como la capacidad necesaria en el *hub* k si el flujo de i a j es enviado desde i al *hub* k y posteriormente al *hub* t (siendo k el primer *hub* en la ruta). De la misma forma, $s_k(Y_i, Y_t, Y_j)$ es la capacidad necesaria en k si el flujo de i a j es enviado desde i al *hub* t y posteriormente al *hub* k .

Las capacidades r_k y s_k están expresadas en función de Y_i ya que dependen de si los nodos i, j ya son propiamente un *hub* y si coinciden con los *hubs* de análisis ($i=k, i=t, j=k$ o $j=t$). De este modo, se añade a la formulación de Aykin (1995a) una restricción adicional en la que la valoración de $r_k(Y_i, Y_t, Y_j)$ y $s_k(Y_i, Y_t, Y_j)$ para $\forall i, j \in N; t \in H$ no debe superar la capacidad nominal del *hub* k .

En el artículo en cuestión se plantean distintas técnicas de resolución que simplifican el modelo y pueden evaluar una solución óptima. La técnica de Relajación Lagrangeana del problema permite articular en dos partes su resolución: la determinación de las localizaciones de los *hubs* y los tipos de servicio o envío de cada par de nodos. En este sentido, si la localización de los *hubs* es conocida, el problema se reduce a resolver con programación matemática entera el tipo de servicio (envío directo o a través de *hub*).

Asignación múltiple

En O’Kelly *et al* (1996) se analizan las propiedades de las redes entre terminales y, mediante un conjunto de ejemplos numéricos, difundir una serie de propiedades de estas redes. Se parte del problema propuesto en O’Kelly (1987) asumiendo que todos los *hubs* están interconectados, que no se permite el envío directo entre clientes y que se permite tanto la asignación simple de un punto a un sólo *hub* o la asignación múltiple en la que un punto puede ser servido por 2 o más *hubs* (Figura 2.3). El estudio teórico analítico de asignación múltiple está extensamente desarrollado en Campbell (1997).

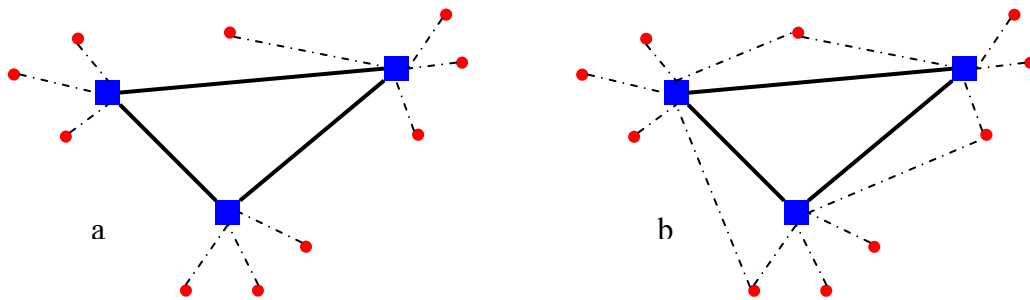


Fig. 2.3. Ejemplos de redes hub & spoke. a) Asignación simple de los nodos a un solo hub. b) Asignación múltiple de un nodo a más de 1 hub. Fuente: O’Kelly *et al* (1996)

Costes no lineales

En O’Kelly y Bryan (1998) se propone una modelización analítica de los costes en la red con *hubs* que permita considerar costes unitarios variables. En este sentido, los coeficientes de coste se multiplican por un factor Ω (ecuación 2.38) que decrece a medida que el flujo en el arco es mayor. La base de la metodología es hacer depender los costes de los arcos entre *hubs* con el flujo que circula a través de ellos, y reproducir la reducción de costes unitarios por efecto de la consolidación de carga en los arcos *interhub* tal como muestra la Figura 2.4.

$$\Omega = \left[1 - \theta \left(\frac{\sum_i \sum_j W_{ij} X_{ijkm}}{\sum_i \sum_j W_{ij}} \right)^\beta \right] \quad (2.38)$$

donde

W_{qr}	Flujo total que se transporta entre los nodos q y r
X_{qrkm}	Variable que toma el valor de 1 si los flujos entre el nodo q y r son enviados vía el <i>hub</i> k y m , 0 en caso contrario
θ, β	Parámetros; $\theta > 0, \beta > 0$

$$\sum_i \sum_j W_{ij} X_{ijkm} \quad \text{Flujo total transportado a través del arco } interhub (k,m)$$

$$\sum_i \sum_j W_{ij} \quad \text{Flujo total de la red}$$

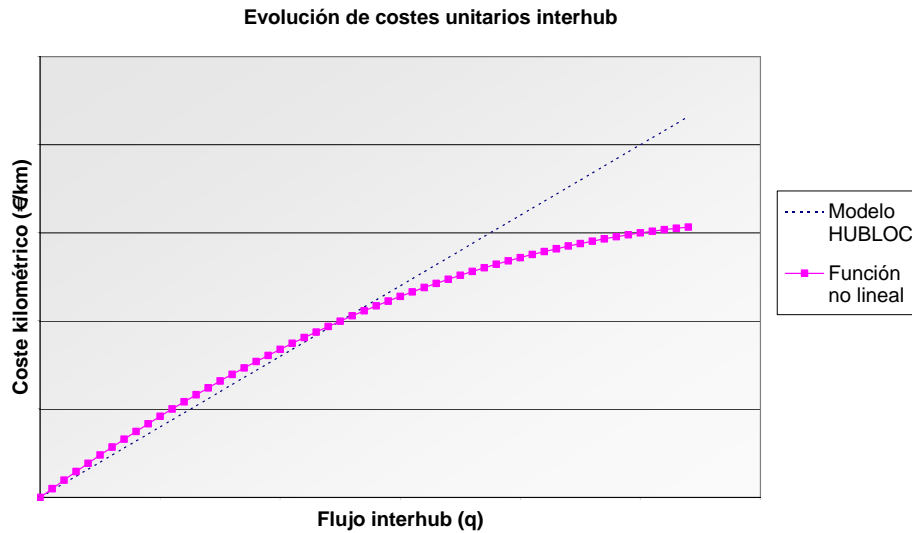


Fig. 2.4 Función de coste por distancia no lineal y modelo HUBLOC. Fuente: O'Kelly y Bryan (1998)

2.2.4. Algoritmos metaheurísticos para el problema de localización y asignación a rutas

Los algoritmos metaheurísticos son unos métodos matemáticos de resolución de problemas de optimización combinatoria de amplia utilización en problemas de ruteo (como el *Vehicle Routing Problem*, VRP, o el *Travelling Salesman Problem*, TSP) o en la asignación y gestión de flotas. Partiendo de una función objetivo a optimizar, los algoritmos metaheurísticos realizan un procedimiento de búsqueda de soluciones o estados. Al conjunto de todos los estados o soluciones candidatas se les llama espacio de búsqueda.

Este grupo de algoritmos se caracteriza por trabajar en un instante de ejecución con un estado actual que es variado por uno nuevo por medio de una transición de estado, movimiento o perturbación. La transición de estado o movimiento tiene como objetivo explorar nuevas configuraciones que aporten ahorros significativos en la función de coste en el subdominio de soluciones próximo al estado actual. Los algoritmos metaheurísticos varían y aceptan las configuraciones asociadas a los cambios de estado según procedimientos probabilísticos. En la mayoría de casos, los algoritmos permiten un incremento positivo en la función de coste a

costa de poder explorar nuevos dominios de la solución, superar óptimos locales y de este modo alcanzar el óptimo global de la solución (búsqueda colina arriba). Los más conocidos son: el Recocido Simulado (*Simulated Annealing*), los Algoritmos Genéticos (*Genetic Algorithms*), la Búsqueda Tabú (*Tabu Search*), y la Colonia de Hormigas (*Ant Colony Optimization*).

Recocido simulado (SA)

El Recocido Simulado (*Simulated Annealing*, SA, también traducido a veces como *Cristalización Simulada*) es un método iterativo que genera aleatoriamente modificaciones de la solución actual y las somete a reglas de probabilidad para su aceptación, evitando de esta manera la caída en óptimos locales. Este método pertenece a la familia de probabilísticos de *hill-climbing* (colina arriba), porque permite la aceptación de movimientos o cambios de la solución que impliquen un sobre coste en la métrica de evaluación o función de costes para evitar limitar la exploración a los extremos locales.

El SA desarrollado en Kirkpatrick *et al.* (1983) se basa en la analogía con el recocido de los metales y consiste en “calentar” a alta temperatura el sistema a optimizar, para luego disminuir la temperatura muy lentamente para permitir la cristalización y ordenamiento óptimo de la estructura del sistema con mínimo coste. La variable de control del proceso y, en analogía, del modelo de optimización es la temperatura.

La reducción de temperatura en el proceso físico se reproduce de forma discreta en el modelo (a saltos). De este modo, la probabilidad que marca el criterio de aceptación de la solución depende del estado del sistema, esto es, de la variable temperatura. En estados iniciales (alta temperatura) el criterio de aceptación permite incluir movimientos que incrementan el coste del sistema; mientras que en estados avanzados del proceso (baja temperatura) las soluciones con coste superior a la mejor solución tienen una probabilidad de aceptación casi nula. El criterio de finalización se basa en la convergencia de la solución a un valor determinado o al cumplir un cierto número de iteraciones. Se ha demostrado que bajo ciertas condiciones este proceso converge asintóticamente al óptimo global con probabilidad uno (Mitra *et al.*, 1986). La demostración de este resultado se basa en el hecho de que el SA puede considerarse como un algoritmo que continuamente intenta transformar la configuración actual en una de sus configuraciones vecinas. Este mecanismo puede ser fácilmente descrito por medio de una cadena no estacionaria de Markov o una secuencia de cadenas homogéneas de Markov.

Una de las primeras aplicaciones del SA para la resolución global y afinado de soluciones manuales de VRP fue la de Robusté *et al.* (1990). Más centrado en el problema objeto de esta tesis, Aykin (1995b) propone una herramienta computacional para la localización óptima de h

hubs y la asignación de N_T nodos o terminales a estos *hubs* basada en el algoritmo del SA desarrollado en Kirkpatrick *et al.*(1983). El “Simulated Annealing Based Greedy-Interchange Heuristic” (SAGI) propuesto por Aykin se articula en dos fases diferenciadas: la construcción de una solución inicial y el proceso de refinamiento de esta solución inicial.

a) Solución inicial

El algoritmo parte de una configuración inicial en la que se considera que cada nodo terminal puede actuar como *hub* de forma que $h=N_T$. En esta configuración, se determina para cada *hub* el coste incremental que se produciría si éste no actuara como tal (no existiera) y la carga se consolidara en el resto de *hubs* del sistema de transporte.

En un segundo paso, el algoritmo elimina o “cierra”el *hub* que tiene asociado un menor coste incremental en la función de costes del sistema. Este procedimiento se va repitiendo iterativamente reasignando los nodos que estaban asignados al *hub* clausurado al nuevo conjunto de *hubs* hasta quedar una configuración con h *hubs* en el territorio. Para proceder a la asignación de cada nodo terminal i , se determina la variable $A_{i,r}$ de forma similar a la que se hacía en la formulación (2.16) y metodología de asignación propuesta en Klincewicz (1991). De este forma se asigna el nodo delegación i al *hub* $r \in H$ que presente un valor inferior de la variable $A_{i,r}$ en cada iteración determinada según la expresión (2.39).

$$A_{i,r} = \min_{k \in H} \{A_{i,k}\} = \min \left\{ \sum_{t \in H} \sum_{j \in K_t} (a_1 W_{i,j} + a_2 W_{j,i}) C_{i,k} + \sum_{\substack{t \in H \\ t \neq k}} \sum_{l \neq i} a (W_{i,l} + W_{l,i}) C_{k,t} \right\} \quad \forall i \in K_r, r \in H \quad (2.39)$$

donde:

- $A_{i,k}$ Coste de transporte de la carga con origen o destino en el nodo delegación i si éste está asignado al *hub* k .
- H Conjunto de *hubs*
- K_t Conjunto de nodos delegación asignados al *hub* t

En este caso, se generan los pesos a_1, a_2, a asociados a los costes de transporte para computar de forma diferencial los envíos desde la terminal origen i al *hub* k , envíos desde el *hub* t a la terminal destino j y el coste del transporte entre *hubs* respectivamente.

b) Proceso de mejora

Esta etapa del algoritmo es propiamente la aplicación del SA y trata de realizar transiciones de estado o movimientos en la solución actual entre el conjunto de *hubs* y el conjunto de nodos terminales que no son *hubs*. El movimiento en si trata de analizar el incremento de coste provocado al considerar como *hub* una delegación $r \notin H$ en substitución de un *hub* $k \in H$. Para determinar cuando es rentable el cambio de un nodo $k \in H$ con el nodo $r \notin H$ en el movimiento de cambio de *hub*, se analiza aquel *hub* $t \in H$ que presente un valor máximo del coeficiente $R_{t,r}$ según la expresión (2.40).

$$U_{t,r} = \max_{k \in H} \{R_{k,r}\} = \max_{k \in H} \left\{ \begin{aligned} & \sum_{\{i/Z_{i,r}=1\}} O_i (C_{i,r} - C_{i,k}) + \sum_{\{i/Z_{i,r}=1\}} D_i (C_{r,i} - C_{k,i}) + \\ & + \sum_{\substack{n \in N \\ n \neq r}} I_{r,n} a(C_{r,n} - C_{k,n}) + \sum_{\substack{n \in N \\ n \neq r}} I_{n,r} a(C_{n,r} - C_{n,k}) \end{aligned} \right\} \quad (2.40)$$

donde:

$I_{n,l} = \sum_{\{i/Z_{i,n}=1\}} \sum_{\{j/Z_{j,l}=1\}} (W_{i,j})$ es la carga que va desde el *hub* n al *hub* l

$W_{i,j}$ es el volumen de mercancía que se envía desde i hasta j

O_n es la suma de la mercancía que se envía desde el nodo terminal n

D_n es la suma de la carga que tiene como destino final el nodo terminal n

$Z_{i,n} = 1$ si el nodo terminal i está asignado al *hub* n

$= 0$ en los demás casos

Si $U_{t,r} \geq 0$, el *hub* t se elimina de la solución y es substituido por el *hub* r ya que el movimiento o transición de estado produce una reducción de costes del sistema. En el caso que la variable sea igual a 0 los costes podrían aumentar, disminuir o mantenerse.

Sin embargo, en el caso que $U_{t,r} < 0$, la función de costes se verá incrementada de modo que el algoritmo determina una probabilidad de aceptación p del intercambio entre *hubs* ($0 < p < 1$) basada en la distribución de densidad de probabilidad de Boltzmann. El criterio de aceptación de la solución queda determinado, pues, por la ecuación (2.41).

$$p = \begin{cases} \exp\left(\frac{U_{t,r}/Z_{\min}}{T}\right) & \text{si } U_{t,r} < 0 \\ 1 & \text{si } U_{t,r} \geq 0 \end{cases} \quad (2.41)$$

donde

Z_{min} : valor mínimo obtenido de la función de costes de transporte

T : parámetro de control (temperatura)

El proceso empieza actualizando la asignación de los *hubs* con la ecuación (2.39) de forma que si el costo de la solución obtenida es mejor que los encontrados hasta el momento, el nodo r substituye al *hub* t y se actualiza la solución. Si sucede lo contrario, la solución se mantiene igual. El parámetro inicial de control (T) se obtiene especificando la probabilidad de no examinar un intercambio para aquellos casos en que la función de ahorro indica un incremento del $100\delta\%$ en el valor de la función objetivo (siendo $\delta = U_{t,r}^*/Z_{min}$). Para calcular el parámetro temperatura T en cada iteración $t+1$ del proceso, se evalúa la condición de rechazo $R(\delta)$ en la iteración t mediante la ecuación (2.42), considerando adicionalmente el parámetro ν ($\nu > 1$) como una constante de enfriamiento del proceso.

$$P(R(\delta))_{t+1} = \nu P(R(\delta))_t \quad (2.42)$$

En este proceso, si se determina una solución mejor con $U_{t,r} < 0$, el parámetro de control T se actualiza en la iteración $t+1$ por medio de la expresión (2.43), de forma que la probabilidad de rechazo se incremente a medida que el proceso iterativo avanza.

$$T = \frac{-\delta}{\ln(1 - P(R(\delta))_{t+1})} \quad (2.43)$$

En Ernst y Krishnamoorthy (1999) también se desarrolla un primer análisis sobre la localización de *hubs* con capacidad limitada con clientes asignados a un solo *hub* (asignación simple) con un algoritmo basado en el Simulated Annealing y en el descenso aleatorio.

En este caso, se parte de una solución inicial escogiendo de forma arbitraria entre el 10 y el 90% de los nodos como *hubs*. El resto de nodos son asignados aleatoriamente a estos *hubs* iniciales. Posteriormente, para determinar y explorar el dominio más próximo de la solución, se definen un conjunto más completo de movimientos o cambios de estado que en Aykin (1995a). De esta forma, se consideran 6 tipos de transiciones de la solución actual siempre que no violen las restricciones de capacidad:

- Reubicación del *hub*: Cambia la localización de un *hub* dentro de un *cluster* (región de servicio) a otro nodo de forma aleatoria dentro del *cluster*.

- Reasignación de nodo: Cambia la asignación de un nodo que no es *hub* (escogido aleatoriamente) a otro *cluster*. Si el *cluster* contiene únicamente aquel nodo, se permite asignar el nodo a otro *cluster* y eliminar el *cluster* actual.
- Nuevo *hub*: Crea un nuevo *cluster* con un solo nodo haciéndolo a su vez *hub*.
- Fusión de *clusters*: Se unen 2 *clusters* aleatoriamente escogidos por medio de la asignación de todos los nodos en un *cluster* al *hub* del otro *cluster*.
- Partición de *clusters*: Se divide un *cluster* en dos *clusters*, por medio de la asignación de algunos de los nodos en el *cluster* original a un nodo que no era *hub* pero que en el nuevo *cluster* deviene *hub*.
- Intercambio de nodos: Intercambia las asignaciones de dos nodos que no son *hubs* entre distintos *clusters*. Cada nodo es asignado al *hub* del otro nodo.

El criterio de aceptación de la solución actual en cada iteración es el mismo marcado por la ecuación (2.41). En este caso, se constata que el método del Recocido Simulado no explora todo el dominio de soluciones y por lo tanto se repite 10 veces el procedimiento sobre soluciones iniciales diferentes.

El método del Descenso Aleatorio (RDH), a diferencia del Recocido Simulado que puede llegar a aceptar soluciones que comporten un empeoramiento local de la solución, siempre exige que una nueva solución comporte un beneficio de la solución. El método RDH, a partir de una solución inicial arbitraria, aplica los mismos 6 movimientos. En aplicar uno de los movimientos, se acepta la configuración realizada como nueva solución si comporta una mejora en la función objetivo. Cabe destacar que este procedimiento es diferente del heurístico del ogro (Greedy), ya que el primero no busca la mejor solución posible en cada iteración sino una solución mejor arbitrariamente.

Los dos heurísticos planteados por Ernst, A.T., M.Krishnamoorthy (1999) se aplicaron 10 veces a un subconjunto de varios nodos ($10 < N_7 < 50$) derivado de un problema base de reparto de correos con 200 nodos. Como conclusión, el algoritmo de RDH produce la mejor solución en problemas de tamaño pequeño o mediano (hasta 50 puntos de servicio), a pesar de un tiempo computacional más elevado. Sin embargo, para problemas habituales que afecten a 200 nodos o más, el algoritmo del descenso aleatorio no produce buenos resultados ni se resuelve con un tiempo computacional adecuado.

Algoritmos genéticos (GA)

Los algoritmos genéticos hacen evolucionar un estado actual compuesto por varios individuos (soluciones) sometidos a acciones aleatorias semejantes a las que actúan en la evolución biológica (mutaciones y recombinaciones genéticas), a la vez que se establece un criterio de

selección por el cual se decide qué individuos sobreviven y qué individuos son descartados. En este sentido, el GA se distingue claramente de todos los metaheurísticos básicamente por el hecho de que en cada iteración se tiene un conjunto de soluciones, o población de individuos y no una única solución en curso. Las soluciones sucesoras se obtienen a partir de la combinación de atributos (cadenas de nodos) de los elementos de la población de soluciones y no mediante la transformación de la solución en curso. La información es encriptada o codificada en forma de cadena, generalmente binaria, llamada cromosoma. La globalidad de símbolos presentes en la cadena (nodos) se denominan genes.

La selección de las diversas soluciones que se utilizan en la generación de una nueva solución sucesora se realiza mediante la definición de una función que evalúa la bondad y ajuste de cada solución de la población, *la función fitness* f_f . En este caso, la función fitness será una función decreciente con el coste de cada solución. De este modo, la probabilidad de selección de una solución de la población en curso es proporcional al valor de la función fitness, mediante la Regla de la Ruleta. Si S es el conjunto o la población de soluciones en curso, se asigna a cada solución $s^* \in S$ un sector circular de ángulo α_{s^*} determinado por la ecuación (2.44).

$$\alpha_{s^*} = (2\pi) \frac{f_f(s^*)}{\sum_{\forall s \in S} f_f(s)} \quad (2.44)$$

De este modo, en cada momento que se quiera seleccionar una solución de la población, se “tirará la ruleta”, es decir, se generará un número aleatorio entre 0 y 2π radianes y se escogerá la solución que tenga el sector circular asociado a este valor.

La transición de estado para crear las siguientes generaciones se realizan a partir de dos movimientos o operadores: el operador de cruzamiento (se intercambian subcadenas de 2 cromosomas) y el operador de mutación (cambia uno o varios genes de un cromosoma).

Una vez se ha incrementado la población, ésta se debe reducir a un tamaño conveniente, eligiendo un cierto número de soluciones mediante procesos idénticos a los de selección (proceso de regeneración).

El esquema general de operación se detalla a continuación:

- Generar aleatoriamente una población o conjunto de soluciones de un problema
- Evaluar cada una de las soluciones por medio de la función *fitness*

- Escoger de la población de soluciones un subconjunto que tenga un valor mayor de la función *fitness*.
- Mutar (cambiar) y entrecruzar (combinar) las diferentes soluciones de esa parte escogida, para reconstruir la población.
- Repetir un número máximo de iteraciones o hasta que se cumpla la convergencia a un valor fijado.

En Abdinnour-Helm y Venkataramanan (1995) y en Abdinnour-Helm y Venkataramanan (1996) se ha desarrollado un procedimiento basado en la aplicación del metaheurístico de GA aplicado a problemas de *many-to-many*. Su análisis y valoración de resultados se realizará comparándolos con los resultados encontrados con el método de la Búsqueda Tabú.

Búsqueda Tabú

El algoritmo de Búsqueda Tabú o *Tabu Search* (TS) ideado por Glover (1986) es un procedimiento iterativo de búsqueda y exploración en el dominio de soluciones que trata de evitar los óptimos locales y el comportamiento cíclico del algoritmo mediante la prohibición de unos movimientos de la solución actual (considerados tabú) y una estructuración de la memoria de soluciones pasadas.

El método de TS es inicialmente determinista, ya que no existe un criterio probabilístico de aceptación de la solución. Éste se basa en sustituir la solución generada o modificada sin ninguna condición de coste, es decir, tanto si es mejor que la solución en curso como si no. Esta única condición podría presentar el inconveniente de que aparezca como solución en curso una configuración que lo haya sido en una iteración anterior y que, por lo tanto, el algoritmo tenga un comportamiento cíclico. Para evitar este fenómeno, se realiza una lista de soluciones o atributos de solución prohibidos o tabús. Esto es, se define un vecindario (conjunto de soluciones derivadas de la aplicación de un movimiento a la solución actual) constituido exclusivamente por soluciones que no contengan atributos “tabús”. Sin embargo, en algunos casos se puede aceptar una solución con elementos tabús siempre que ésta sea la solución de mínimo coste durante el proceso de búsqueda (criterio de aspiración). La lista tabú no es estática, ya que los movimientos incluidos únicamente mantienen la prohibición durante un número de iteraciones acotado (memoria a corto plazo).

Según Glover y Laguna (2002) el método de TS se basa en tres aspectos fundamentales:

1. El uso de estructuras flexibles de memoria basadas en atributos para una mejor explotación y evaluación de los dominios de soluciones.
2. Mecanismo de control de las estructuras de memoria basado en la interacción entre las condiciones que limitan y hacen más flexible el proceso de búsqueda.

3. La incorporación de memorias de diferente duración (de corto a largo plazo), para implementar estrategias que intensifiquen y diversifiquen la búsqueda.

El método de TS integra los conceptos de intensificación y diversificación en la memoria a largo plazo. El primero consiste en intensificar la búsqueda a partir de soluciones con determinadas características o que hayan demostrado (históricamente) ser buenas, y el segundo, interesa si el algoritmo se estanca en un cierto tipo de soluciones que se repiten con mucha frecuencia en la solución actual. Este algoritmo, al igual que el SA, no tiene una condición intrínseca de finalización, por lo tanto, se ha de definir una como por ejemplo, alcanzar un cierto número de iteraciones o el hecho de no progresar.

En Glover y Laguna (2002) se sugiere como nombre alternativo al TS el de algoritmo de "inhibición débil", ya que los movimientos tabús son generalmente una pequeña fracción del total de movimientos disponibles y tienen un período de afectación relativamente corto. En este sentido, este método puede contrastarse con la técnica del *branch and bound* que también prohíbe ciertos movimientos para evitar ciclos, pero lo hace de una manera más rígida, por lo que se le considera como una forma de búsqueda con "inhibición fuerte".

En Abdinnour-Helm y Venkataramanan (1995) y en Abdinnour-Helm y Venkataramanan (1996) se proponen dos métodos basados en los metaheurísticos GA y TS para solucionar el problema de localización de *hubs* sin restricciones de capacidad y con asignación simple y única de cada cliente a su *hub* asociado. El modelo definido en la ecuación (2.45) parte de la ecuación de costes de O'Kelly (1992) donde el número de *hubs* es una variable de decisión y el coste fijo de establecer un *hub* se incluye en la formulación.

$$\min \sum_i \sum_k X_{ik} c_{ik} (O_i + D_i) + \sum_j \sum_m X_{jm} \sum_i \sum_k X_{ik} (\alpha T_{ij} c_{km}) + \sum_j X_{jj} f_j \quad (2.45)$$

sujeeto a

$$X_{jj} - \sum_i X_{ij} \geq 0 \quad \forall j$$

$$\sum_j X_{ij} = 1 \quad \forall i$$

$$X_{ij} \in \{0,1\} \quad \forall i, j$$

La variable de decisión X_{ik} determina que el nodo i es asignado al *hub* k cuando $X_{ik}=1$, mientras que en caso contrario $X_{ik}=0$, de la misma forma que si j es un nodo terminal *hub* $X_{jj}=1$. Las variables W_{ij} y c_{ij} representan respectivamente el flujo entre las terminales (i,j) y el coste por unidad de mercancía entre (i, j) mientras que f_i representa el coste fijo de instalación si el nodo terminal i es un *hub*.

El algoritmo-GA es un procedimiento basado exclusivamente en los algoritmos genéticos mientras que el algoritmo-GATS es un algoritmo híbrido, mezcla de los algoritmos genéticos y de los algoritmos de la búsqueda tabú. En el algoritmo de GATS, los algoritmos genéticos se utilizan para cuantificar el número óptimo de *hubs*, mientras que el TS se aplica para identificar la asignación óptima de los puntos a cada *hub* a la vez que se hace en Skorin-Kapov et al. (1994) para el *h-hub* median problem.

Tabla 2.2. Resultados de aplicación del GA y GATS y solución exacta en problemas de localización de aeropuertos en EE.UU. Fuente: Abdinnour-Helm (1999)

	Solución de flujo global		Solución de flujo local	
E.E.U.U. Territorio Continental	Coste fijo	125	Coste fijo	125
	Óptimo	1.563,7 (1,11,15)	Óptimo	1.162,7 (5,7,10,14)
	GA	1.563,7 (1,11,15)	GA	1.162,7 (5,7,10,14)
	GATS	1.563,7 (1,11,15)	GATS	1.162,7 (5,7,10,14)
E.E.U.U. Costa Este	Coste fijo	20	Coste fijo	20
	Óptimo	282,1 (10)	Óptimo	230 (2,16)
	GA	282,1 (10)	GA	235,7 (8,18)
	GATS	282,1 (10)	GATS	231,5 (16,17)
E.E.U.U. Costa Suroeste	Coste fijo	40	Coste fijo	60
	Óptimo	618,3 (7,8)	Óptimo	593,8 (4,5,7,11)
	GA	629,6 (7)	GA	619,9 (4,7,8)
	GATS	618,3 (7,8)	GATS	595,9 (7,11)

Tabla 2.3. Resultados de aplicación del GA y GATS y solución exacta en problemas de localización de aeropuertos en Canadá y Europa.. Fuente: Abdinnour-Helm (1999)

	Solución de flujo global		Solución de flujo local	
Canadá	Coste fijo	180	Coste fijo	180
	Óptimo	1.595,9 (6,13)	Óptimo	1.189,7 (5,15,16)
	GA	1.601,8 (6,13)	GA	1.189,7 (5,15,16)
	GATS	1.595,9 (6,13)	GATS	1.189,7 (5,15,16)
Europa Occidental	Coste fijo	75	Coste fijo	75
	Óptimo	1.072,7 (2,12)	Óptimo	938,8 (7,11,12)
	GA	1.073,4 (2,12)	GA	938,8 (7,11,12)
	GATS	1.072,7 (2,12)	GATS	938,8 (7,11,12)

Los algoritmos GA y GATS se han aplicado a una serie de problemas conocida como CAB-data, derivada de la problemática de *hubs* en EE.UU. y que se recoge en O'Kelly (1987). En Abdinnour-Helm (1999) se han comparado los resultados de solución con la aplicación únicamente de algoritmos genéticos (GA) de búsqueda tabú (TS) y con la solución exacta en la Tabla 2.2 y 2.3. Con todo, la fusión de la metodología del TS y del GA en un algoritmo común mejora significativamente las soluciones encontradas por cada uno de éstos procedimientos. En todos los casos, la metodología combinada mejora o iguala la mejor solución encontrada con GA. Cabe destacar que los movimientos de cambio de un nodo son mucho más eficaces que los movimientos de intercambio entre nodos.

Colonia de Hormigas (ACO)

Los algoritmos ACO (*Ant Colony Optimization*) son modelos que reproducen el comportamiento de colonias de hormigas en la naturaleza. La capacidad de transmitirse información entre la colonia por medio de un rastro de feromona en su camino de ida y vuelta hacia un destino, les permite determinar la ruta más corta. En una etapa inicial, las rutas de hormigas entre dos puntos pueden generarse de forma aleatoria, pero cuando el sistema ha superado un tiempo necesario, la mayor concentración de feromonas se encontrará en el camino de menor longitud y podrá guiar a la colonia en su búsqueda de minimización de la distancia recorrida. La operativa irá marcando con más claridad (concentración) la ruta de menor coste a medida que el tiempo o iteraciones se incrementen.

En definitiva, la probabilidad con la que una hormiga escoge un camino aumenta con el número de hormigas que previamente hayan elegido el mismo camino (Barcos, 2002), aunque la feromona también se va evaporando con el paso del tiempo provocando que el rastro sufra cierto debilitamiento.

La base de la modelización de rutas de distribución con la aplicación de ACO se puede ver en Dorigo et al. (1996). Se caracteriza por ser un proceso iterativo en que se introduce una colonia de m hormigas en cada iteración que construye una solución al problema. En cada iteración, las soluciones construidas por cada hormiga se basan en criterios probabilísticos en función de la concentración de feromona artificial y una información inicial (solución de partida). Cuando todas las hormigas han construido una solución se debe actualizar la feromona en cada arco y tras ésta, puede comenzarse una nueva iteración.

En Barcos (2002) se aborda el problema del transporte de muchos orígenes a muchos destinos con varios *hubs* con estrategias de envío directo, a través de uno o dos *hubs*, mediante una ruta con paradas múltiples en origen (*peddling en origen*), y mediante una ruta con paradas múltiples en destino (*peddling en destino*). Los resultados obtenidos han sido contrastados,

para pequeños ejemplos, con los resultados exactos encontrados utilizando la programación entera. Finalmente, se aplica el algoritmo a una empresa que ofrece servicio de entrega en 24h en España peninsular y Baleares. La solución aportada por el heurístico (338,12m€) genera un coste por unidad de carga transportada un 8,1% inferior al coste actual de la empresa (368,06m€). Se ha comprobado adicionalmente que los costes son muy sensibles a los datos temporales y del nivel de servicio como son la velocidad de recorrido, el margen de tiempo de servicio o el tiempo de organización de la carga en terminales.

2.2.5. Metodologías integradas de planificación de envíos en red troncal y red capilar

En Wasner y Zäpfel (2004) se propone un método para la localización generalizada de *hubs* conjuntamente con un modelo de rutas de vehículos en la red capilar para la distribución de mercancías aplicado en Austria. En particular, el documento define la configuración de una red de distribución que gestione el número, la localización, las áreas de servicio y las rutas desde los clientes o puntos de servicio hasta los almacenes y viceversa.

El artículo expone que los principales análisis científicos sobre la localización de *hubs* se han basado en el problema de transporte de larga distancia directo entre un punto de visita y el *hub* asociado (O’Kelly 1986, Kliencewicz 1991, O’Kelly and Millar 1994, Ernst and Krishnamoorthy 1998, Ebery *et al.* 2000). Sin embargo, estos modelos propuestos no incorporan la visión de reparto del problema, en especial la fase de diseño de la recogida y reparto a los puntos de servicio de forma coordinada. Los modelos de localización de *hubs* suponen que la unidad de envío desde un *hub* a un cliente es independiente de la ruta escogida para visitar el cliente; de modo que los costes de entrega únicamente dependen de la suma del producto de los costes de envío desde el *hub* al cliente y el número de unidades repartidas al cliente.

No obstante, los costes de entrega de los clientes para sistemas con vehículos con carga menor a la capacidad (*LTL*) dependen de la ruta de entrega de los vehículos. Por lo tanto, es necesario integrar en los modelos de localización el problema de diseño de rutas reparto y recogida para proponer una red de distribución física que integre los costes del envío directo al *hub* (*line-haul*), considerando a la vez los costes de reparto y recogida, costes de terminal y de terminal *hub*. Los costes de larga distancia (*line-haul*) ocurrirían entre *hubs*, entre terminales o entre *hubs* y terminales, mientras que los costes de recogida y entrega ocurrirían entre puntos de servicio (clientes) y terminales.

Debido a la capacidad de empaquetamiento de los flujos correspondientes a una zona de transporte dada, así como a la posibilidad de la utilización de vehículos de mayor carga, los costes de envío entre puntos de servicio y *hubs* (*line-haul*) son inferiores a los costes de entrega y reparto. Este hecho se manifiesta en Salhi and Rand (1989), artículo en el que se recogen las siguientes representaciones de costes:

- Costes de *line-haul* 15-25%
- Costes de entrega y recogida 35-60%
- Costes de operación en terminales y *hubs* 25-45%.

De este modo, queda patente la necesidad de tener en cuenta de forma simultánea y coordinada la consideración del problema de localización de *hubs* y terminales con el de rutas de reparto o entrega.

En primer lugar se fija el número de almacenes y posteriormente su localización, de forma que posteriormente se determina la estructura de transporte a larga distancia. Si no se dispone de información aproximada acerca del valor óptimo del número de almacenes, es necesaria una fase de inicialización para determinar su convergencia a un valor. El procedimiento de resolución de la Figura 2.5 se repite con distintos valores de número de almacenes (4, 10, 40 por ejemplo) y se determina el coste total del sistema. La idea que debe prevalecer es que el coste total del sistema sigue una función de costes convexa que depende del número de almacenes, N_T . Una vez se haya identificado el valor subóptimo (por ejemplo 10) se repite el proceso para un intervalo más restringido de N_T (valores 6, 10, 15) para determinar la configuración total de costes, con lo que se termina la fase de inicialización. La determinación de las localizaciones de los almacenes se realiza especialmente durante la inicialización, de forma que cada zona postal de servicio puede tener localizada un almacén. Debido que el número de almacenes está fijado, los costes fijos de todas las localizaciones pueden considerarse constantes.

A partir de un método aproximado, las zonas postales donde se localizarán los almacenes serán determinadas. Los métodos disponibles son un algoritmo de *Añadida*, en el que se parte de 0 localizaciones y se identifica la localización que reduce al máximo la función objetivo, y el algoritmo de *Supresión*, en el que empezando con todas las localizaciones, en cada iteración se selecciona el almacén cuya eliminación comporta un incremento menor en la función objetivo.

Las nuevas localizaciones de los almacenes se pueden determinar alternativamente de tres formas diferentes:

- Localizaciones completamente nuevas, independientes de las localizaciones previas
- Movimiento de un solo almacén, manteniendo la localización de los demás
- Se mueven simultáneamente múltiples localizaciones de almacenes.

La estructura de larga distancia se desarrolla en tres etapas:

a) **PZ- asignación de almacenes.** Se tiene en cuenta la distancia y el tiempo entre la zona postal y la localización de almacenes, hecho que puede comportar soluciones substancialmente diferentes según el criterio sea distancia o tiempo. Se asignan probabilidades a cada zona postal (dependiente de la distancia) para realizar diferentes asignaciones en las etapas de iteración.

b) **Determinación de la estructura de larga distancia.** La decisión recae únicamente en la definición de la red óptima que determine la localización de *hubs*, su número total y tipo de envío (directo o con transferencia) entre pares de almacenes, ya que el volumen total de mercancía a transportar queda fijado por las fases anteriores. Sin embargo, una solución exacta a este tipo de problemas en condiciones reales parece como improbable y por tanto se deben realizar simplificaciones.

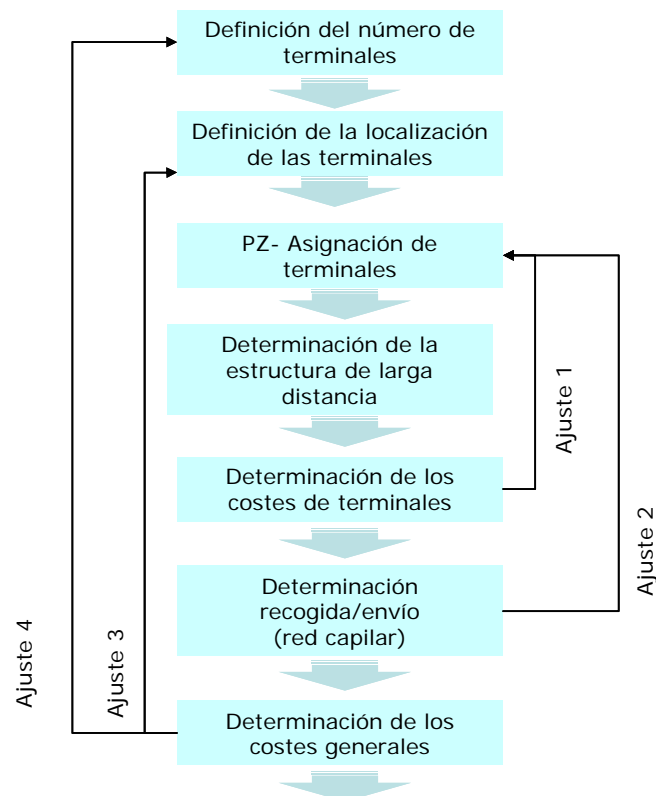


Fig.2.5 Esquema de aplicación del algoritmo del procedimiento propuesto en Wasner y Zäpfel (2004)

c) **Determinación de los costes de almacenes.** La asignación de las zonas postales a los almacenes determina el volumen total de tráfico a ser transferido en los almacenes individuales, de forma que los costes pueden ser calculados directamente de las funciones de costes modeladas por el autor. Se realiza un estudio particular de reasignación de almacenes especialmente en las zonas postales de recogida que están en la frontera del área asignada a un almacén.

La siguiente fase es la definición de los transportes de reparto y recogida, que al depender de las asignaciones de almacenes se permite un segundo *feedback* para minimizar el coste total de los transportes de larga distancia, de reparto y recogida. En esta fase, la solución del problema se puede realizar con la aplicación de alguna técnica heurística de la extensa bibliografía y producción científica sobre el tema.

Después de la primera iteración los costes son computados en la solución actual en el paso *Determinación de los costes totales*, de modo que se determina una primera solución para posteriores iteraciones.

Este método se ha aplicado para la re-planificación de una compañía de transporte en Austria con 10 almacenes y 1 *hub*. Los resultados derivados del estudio han logrado reducir los costes de distribución anteriores (1.000.000 unidades económicas) a 853.000 unidades económicas con la aplicación de la metodología.