UNIVERSIDAD DE LA LAGUNA

Optimización metaheurística para la planificación de redes WDM

Autor: Melián Batista, María Belén

Directores: Manuel Laguna y José A. Moreno Pérez

Departamento de Estadística, Investigación Operativa y Computación

Contents

Prólogo (español)

Durante las últimas cuatro décadas, la planificación en redes de telecomunicaciones ha constituido un problema muy fértil para desarrollar y aplicar modelos de optimización. Dos aspectos principales alientan estos esfuerzos: las enormes inversiones de capital en comunicaciones ofrecen oportunidades significativas para reducir costes, incluso con mejoras en los diseños y en la administración de las redes; y los rápidos cambios tecnológicos proporcionan constantes alternativas de diseño y entornos de operación.

Los diversos desarrollos llevados a cabo en las últimas décadas, así como el incremento del volumen de demanda han derivado en una nueva era en telecomunicaciones, con la sustitución de transmisión analógica por tecnología digital, disminuyendo costes e incrementando anchos de banda con equipo de transmisión de fibra óptica sustituto de los tradicionales cables de cobre.

La tecnología de fibra óptica se ha convertido rápidamente en una de las principales componentes de las redes de comunicación. Este medio de transmisión es de coste efectivo, fiable y proporciona capacidad casi ilimitada. Esta combinación permite establecer nuevos servicios que requieren grandes cantidades de ancho de banda. De forma paralela, las características únicas de esta tecnología implican la necesidad de nuevos métodos de planificación.

Se define *red óptica* como una red de telecomunicaciones con fibras ópticas como enlaces de transmisión, y con una arquitectura diseñada para explotar las características de las fibras. Estas arquitecturas involucran dispositivos ópticos y electrónicos. Por lo tanto, el término red óptica no implica necesariamente red puramente óptica, pero sí algo más que un conjunto de fibras que terminan en dispositivos electrónicos.

En teoría, la fibra óptica tiene un ancho de banda extremadamente alto. Sin embargo, dado que la tasa a la cual puede acceder un usuario final está limitada a velocidad electrónica, sólo pueden ser alcanzadas velocidades de unos pocos gigabits por segundo.

Por tanto, la conversión óptica-electrónica impide explotar el ancho de banda de una sóla fibra. La *multiplexación por división de longitudes de onda* (WDM - *wavelength division multiplexing*) y los amplificadores de fibra (EDFA - *erbium-doped fiber amplifier*) son dos desarrollos recientes que permiten superar estas limitaciones.

WDM es la transmisión de múltiples señales láser a diferentes longitudes de onda (colores) en la misma dirección, al mismo tiempo, y sobre el mismo hilo de fibra. Las tecnologías tradicionales tales como SDH (*Synchronous Digital Hierarchy*) o su equivalente americano SONET (*Synchronous Optical network*) están con frecuencia basadas en anillos interconectados, mientras que la tecnología WDM usualmente no restringe a ninguna arquitectura de red especial.

El sistema WDM está equipado con amplificadores que permiten la transmisión en un canal. El desarrollo reciente más importante ha sido la comercialización de los EDFA, que amplifican las señales en diferentes longitudes de onda simultáneamente.

Las implementaciones actuales de las redes de telecomunicaciones no tienen la capacidad suficiente para soportar las demandas inminentes de ancho de banda. La tecnología WDM permite incrementar la capacidad de las redes de fibra óptica existentes siendo compatible con el equipamiento instalado en las mismas. En esta memoria se aborda el estudio de un problema real que surge en este marco. El problema consiste en incrementar la capacidad, a mínimo coste, de una red de fibra óptica existente para satisfacer un conjunto de requerimientos de demanda. Por tanto, estudiamos el *Problema de Provisión y Conducción* (*Provisioning and Routing Problem*) en redes ópticas WDM.

La alta capacidad que proporcionan las redes WDM da lugar a diseños de red más dispersos que los asociados con la tecnología de cobre. Consecuentemente, el tráfico conducido a través de un único enlace, y por tanto también la interrupción de servicios si falla un enalce o nodo de red, es significativamente mayor. Todo ello implica que además de la reducción de costes, la supervivencia frente a fallos es uno de los aspectos más importantes en el diseño de redes ópticas. Se llama supervivencia a la habilidad de restaurar los servicios de red en caso de un fallo en la misma.

Los algoritmos y las formulaciones de modelos generalmente consideran que los datos de los problemas dados están perfectamente determinados. Sin embargo, en la mayoría de los problemas que surgen en el mundo real esto no es así. Por el contrario, los problemas reales suelen incluir datos inciertos debido a procedimientos de medida y errores. Este problema surge, por ejemplo, cuando los datos representan demands futuras de tráfico y costes de productos que no pueden ser determinados con certeza.

El capítulo 1 de esta memoria proporciona una descripción de las redes de telecomunicaciones ópticas, así como de los problemas que surgen al planificar redes ópticas. También se proporciona una descripción de la tecnología óptica utilizada en esta memoria.

El capítulo 2 presenta una introducción a las metaheurísticas centrada en el análisis de la integración de varias metaheurísticas en un procedimiento híbrido. Se describen los problemas de optimización en general y la programación lineal como una herramienta para aportar soluciones a un problema. Se plantea como alternativa el diseño de metaheurísticas, describiendo los elementos de las heurísticas constructivas y las mejoras locales. Se describen brevemente los elementos fundamentales de las metaheurísticas básicas. Se analizan las características más importantes de las metaheurísticas que se integran posteriormente para construir un híbrido que las aproveche. Se presta especial atención a las características más relevantes de las metaheurísticas de búsqueda tabú, de búsqueda dispersa y del arranque múltiple que se usan para construir el procedimiento híbrido presentado en el siguiente capítulo. Finalmente se consideran las propiedades que sería deseable que tuvieran las metaheurísticas.

El capítulo 3 realiza una descripción del problema general objeto del estudio sin restringir la posible tecnología a instalar, mostrando sus aplicaciones así como una revisión de la literatura de las mismas. La siguiente sección contiene una revisión de la literatura del problema de Provisión y Conducción en redes ópticas WDM. Se describe en detalle el problema, analizando las características del mismo que pueden ser importantes para su resolución. Se considera un modelo de Programación Matemática Lineal Entera basado en el propuesto por otros autores y se analizan las dimensiones del mismo. Se propone como método de resolución un procedimiento metaheurístico híbrido, que combina la búsqueda dispersa (*scatter search*), la búsqueda tabú (*tabu search*), y el multiarranque (*multistart*). Este híbrido se compara con otro metaheurístico, con dos variantes de un procedimiento basado en permutaciones y con un optimizador comercial. El modelo propuesto tiene un número muy elevado de variables y restricciones, lo cual impide que sea resuelto de forma óptima por procedimientos de solución exactos cuando se consideran problemas reales de tamaño moderado o grande. El metaheurístico híbrido propuesto es capaz de obtener diseños de red con costes inferiores a los proporcionados por otras metaheurísitcas y por optimizadores comerciales después de un cierto tiempo de ejecución.

En el capítulo 4 proponemos un modelo heurístico alternativo para el problema objeto de estudio que tiene un menor número número de varibles y restricciones. Los resultados obtenidos por un optimizador comercial, que porporciona soluciones óptimas, son comparados con los proporcionados por el metaheurístico híbrido desarrollado en el capítulo anterior.

El capítulo 5 se dedica a desarrollar métodos de resolución para el problema de supervivencia en redes ópticas WDM y proponemos modelos de programación lineal entera para diferentes esquemas de protección de redes. Los trabajos anteriores que aparecen en la literatura considerando estos esquemas de protección para el problema en estudio, obtienen en una primera fase la red de trabajo o red de servicio, que permite satisfacer los requerimientos de demanda, y en una segunda fase, comenzando con la mejor red de trabajo obtenida, la red de protección, que protege contra fallos únicos en los enlaces de red. Los modelos propuestos en este capítulo tienen como dato de entrada la mejor red de trabajo. Sin embargo, proponemos un método de resolución basada en el metaheurístico híbrido desarrollado que detemina la red de protección no solamente para la mejor red de trabajo alcanzada, sino para el conjunto de redes de trabajo con costes inferiores proporcionadas por el metaheurístico. Hemos comprobado que el mejor diseño total, es decir, el que incurre en menos costes de incremento de capacidad para obtener las redes de trabajo y protección, no es en la mayoría de los casos obtenido al considerar la mejor red de trabajo, sino alguna otra incluida en el conjunto de las "buenas" (conjunto de referencia).

En el capítulo 6 estudiamos el problema de provisión y conducción en redes ópticas WDM bajo incertidumbres en las demandas. Proponemos varios modelos matemáticos alternativos y métodos de resolución.

Agradecimientos (español)

Me gustaría agradecer a los profesores José A. Moreno Pérez y Manuel Laguna por su orientación en mi investigación en redes ópticas y optimización combinatoria, por haberme ayudado a convertirme en una investigadora más madura. También quiero agradecerles su contribución a los capítulos de esta tesis.

Ofrezco mi más sincero agradecimiento al profesor Manuel Laguna y a los compañeros del *Leeds School of Business* en la *University of Colorado at Boulder* por haberme acogido como una compañera más durante mis estancias en dicha universidad.

Me gustaría agradecer a los profesores Marcos Moreno Vega y Félix García López por su valiosa colaboración en diversos temas de investigación y por sus contribuciones a esta memoria.

También agradezco con toda sinceridad a mis compañeros del Departamento de Estadística, Investigación Operativa y Computación de la Universidad de La Laguna por contribuir al desarrollo de esta tesis proporcionando un muy buen ambiente de trabajo.

A la Universidad de La Laguna.

Finalmente, agradezco a mi familia y amigos su constante ayuda y estímulo a lo largo de toda mi vida.

Resumen (español)

CAPÍTULO 1.- Redes de Telecomunicaciones Ópticas.

Las redes y los sistemas de telecomunicaciones han evolucionado de forma explosiva en las últimas décadas, proporcionando nuevas oportunidades para la modelización y aplicación de procedimientos de optimización en problemas de planificación. Nuestro actual desarrollo social requiere tener acceso a la información *cuando la necesitemos* y *donde la necesitemos*. La información se proporciona a través de redes globales de comunicación, cuyas implementaciones actuales no disponen de capacidad suficiente para soportar las demandas de ancho de banda inminentes.

El mundo de las telecomunicaciones evoluciona vertiginosamente con la constante introducción de nuevas tecnologías en el mercado. Dado que la economía evoluciona de un pasado industrial a un futuro basado en la información, la demanda de más ancho de banda se está convirtiendo en un conductor dominante en la mayoría de las sociedades.

La mayoría de las redes de telecomunicaciones están divididas en tres niveles principales [3], [41], [93]: la red de larga distancia, la red metropolitana, y la red de acceso local. La red de larga distancia es la que típicamente conecta pares de ciudades a través de los nodos de paso. La red metropolitana interconecta oficinas centrales en diferentes grupos de clientes y proporciona acceso a los nodos de paso. Finalmente, la red de acceso local, conecta clientes individuales, pertenecientes a un grupo, a la correspondiente oficina central.

Estos tres niveles del sistema de comunicación se diferencian en diversos aspectos incluyendo los criterios de diseño particulares. Idealmente, el diseño de una red de telecomunicaciones debería considerar simultáneamente estos tres niveles. Sin embargo, debido a su complejidad, el problema de planificación completo se descompone considerando cada uno de los niveles independientemente. Se planifica la red de larga distancia con la información sobre las necesidades globales. En cada área metropolitana, se planifica la correspondiente red de comunicaciones. Finalmente, cada grupo de usuarios aborda el diseño de su propia red de acceso local.

Aunque la mayoría de las redes de telecomunicaciones actuales tienen fibra óptica, esto no las convierte en redes ópticas. En la mayoría de los casos en los que es usada la fibra, ésta sólo se despliega en los enlaces de transmisión para reemplazar a los cables de cobre. Sin embargo, una *red óptica* es una red de telecomunicaciones con fibras ópticas como enlaces de transmisión, y con una arquitectura diseñada para explotar las características ópticas de las fibras. Tales arquitecturas involucran tanto dispositivos ópticos como electrónicos. Por lo tanto, el término red óptica no implica exclusivamente red de conexiones puramente ópticas, sino algo más que un conjunto de fibras ópticas que terminan en dispositivos electrónicos.

La tecnología de fibra óptica se ha convertido rápidamente en una de las principales componentes de las redes de comunicación. Este medio de transmisión es de coste efectivo y fiable, y proporciona capacidad casi ilimitada. Esta combinación permite establecer nuevos servicios que requieren grandes cantidades de ancho de banda. De forma paralela, las características únicas de esta tecnología implican la necesidad de nuevos métodos de planificación.

Multiplexación por División de Longitudes de Onda

En teoría, la fibra óptica tiene un ancho de banda extremadamente alto. Sin embargo, dado que la tasa a la cual puede acceder un usuario final está limitada a velocidad electrónica, sólo pueden ser alcanzadas velocidades de unos pocos *gigabits* por segundo. Por tanto, es la conversión óptica-electrónica la que impide explotar el ancho de banda de una fibra. La multiplexación por división de longitudes de onda (WDM - *wavelength division multiplexing*) y los amplificadores de fibra (EDFA - *erbium-doped fiber amplifier*) son dos desarrollos recientes que permiten superar estas limitaciones.

La tecnología de multiplexación por división de longitudes de onda (WDM) es la transmisión de múltiples señales láser a diferentes longitudes de onda (colores) en la misma dirección, al mismo tiempo, y sobre el mismo hilo de fibra. Las tecnologías tradicionales tales como SDH (*Synchronous Digital Hierarchy*) o su equivalente americano SONET (*Synchronous Optical network*) están con frecuencia basadas en anillos interconectados, mientras que la tecnología WDM usualmente no restringe a ninguna arquitectura de red especial. A tales redes generales se les denomina redes de tipo malla.

Cada longitud de onda utilizada en la transmisión se convierte en un canal WDM. El sistema debe estar equipado con amplificadores que permiten la transmisión en un canal. El desarrollo reciente más importante ha sido la comercialización de los amplificadores EDFA, que amplifican las señales en muchas longitudes de onda diferentes simultáneamente. La mayor parte del coste de capacidad cuando se usan sistemas WDM está relacionado con las tarjetas de canal, que se añaden cuando se requieran. Esto significa que es posible instalar un sistema WDM con capacidad de hasta 96 canales donde sólo ocho canales están activos, y el diseño sólo implicaría el coste de equipar los ocho canales activos.

Para usar tecnología WDM, se debe contar con una unidad de equipo en ambos extremos de cada enlace de fibra. Además, para cada longitud de onda o canal en uso, se debe instalar el equipo de canal en ambos extremos del canal. Cada canal WDM es bidireccional y tiene la misma capacidad que un par de fibras.

La amplificación es el proceso de restaurar la señal óptica a su poder óptico original y sin distorsión después de que la señal haya perdido poder al pasar a través de un hilo de fibra. Este proceso es particularmente importante en entornos WDM. Los amplificadores estándares no tienen elementos electrónicos y consecuentemente no precisan la clásica conversión eléctrico-óptico y óptico-eléctrico, eliminando así la necesidad de ancho de banda adicional.

Nodos de Redes Ópticas

Los nodos de red pueden tener distintos tipos de funcionalidad. Se pueden clasificar en orden creciente de complejidad del siguiente modo:

- Nodos Estáticos. Son los acopladores direccionales y los conductores estáticos.

- Nodos Dinámicos. El nodo dinámico más simple es un interruptor de división de espacio, que es comúnmente llamado dispositivo de conexión óptica cruzada (OXC - *Optical Cross-Connect*). Las redes WDM pueden interconectar datos ópticamente mediante el uso de dispositivos OXC. Un sistema OXC puede incluir conversión óptico-electrónico o ser completamente óptico. Los sistemas que requieren conversión están equipados con transmisores y receptores. Estos sistemas convierten los datos del dominio óptico al electrónico, seguidamente los interconectan mediante el uso de un centro de cambio eléctrico, y finalmente convierten los datos nuevamente el dominio óptico. Los sistemas OXC completamente ópticos interconectan los datos enteramente en el dominio óptico.

La restricción de *continuidad de longitud de onda* significa que la señal óptica debe tener la misma longitud de onda desde su origen hasta su destino. Es posible evitar esta restricción mediante la instalación de *convertidores* de longitudes de onda en los sistemas OXC completamente ópticos. Un convertidor de longitudes de onda es un dispositivo óptico capaz de trasladar la señal desde una longitud de onda entrante a una longitud de onda saliente posiblemente diferente entre las longitudes de onda disponibles en el sistema. Los convertidores de longitudes de onda relajan la restricción de continuidad de longitud de onda. Sólo es necesario que esté disponible alguna longitud de onda saliente entre las disponibles en el sistema. De esta manera se reduce el número de longitudes de onda necesarias para conducir un conjunto de demandas. Recíprocamente, serán necesarias menos longitudes de ondas distintas para conducir la demanda, resultando en un mejor aprovechamiento del ancho de banda.

Arquitecturas de Redes Ópticas WDM

Según el nivel de control en los nodos de red, se consideran tres clases de arquitecturas de redes ópticas WDM: redes de emisión y selección, redes de conducción de longitudes de onda, y redes de ondas luminosas lineales.

Trabajamos con redes de conducción de longitudes de onda, que pueden incluir selectividad estática o dinámica en los nodos de red. Estas redes no están restringidas a ninguna topología física particular. Una longitud de onda puede ser seleccionada en un nodo de red y ser conducida individualmente. Los nodos contienen transmisores y receptores que pueden seleccionar cualquier longitud de onda de un rango determinado. Entonces los caminos ópticos son punto-a-punto y para obtener conectividad multipunto es necesario el uso de múltiples conexiones ópticas punto-a-punto ópticas (WDM) y un par transmisor-receptor para casa conexión. Además es posible incluir conversión de longitudes de onda en la red.

Problemas en redes de conducción de longitudes de onda

Los problemas más relevantes que surgen en la conducción de longitudes de ondas son los siguientes.

Conducción y Asignación de Longitudes de Onda.

La conducción y asignación de longitudes de onda (RWA - Routing and Wavelength Assignment) es uno de los problemas fundamentales en redes de conducción de longitudes de onda. Este problema consiste en asignar una longitud de onda disponible a una conexión y establecer dicha longitud de onda en el transmisor y en el receptor. El problema está presente como subproblema en la mayoría de los problemas que surgen en la planificación de sistemas WDM. Un estado del arte muy completo sobre este problema aparece en [142].

Redes Convertibles.

La restricción de continuidad de longitudes de onda puede ser eliminada mediante la instalación de convertidores en los nodos de red. Las redes de conducción de longitudes de onda que poseen esta capacidad son llamadas redes convertibles. Un nodo con capacidad de conversión de longitudes de onda es llamado nodo convertidor de longitudes de onda. Las redes convertibles reducen la pérdida de ancho de banda, aunque pueden llegar a ser muy costosas. Los problemas de planificación surgen al estudiar la conveniencia de dotar a la red de convertidores y de determinar los nodos seleccionados para su instalación.

Diseño de la Topología Virtual.

Aparte de la topología física de la red determinada por los nodos y los enlaces físicos que los enlazan, las conexiones que se pueden realizar efectivamente condicionan el uso y gestión de la misma. La topología virtual de una red WDM está formada por un conjunto de caminos de luz establecidos entre un subconjunto de pares de nodos en la red. El problema de establecer la topología virtual de la red es similar al problema RWA de conducción y asignación de longitudes de onda, excepto que surgen en diferentes niveles de la red.

Reconfiguración de la Topología Virtual.

Dado que la topología virtual se diseña para satisfacer el tráfico estimado entre pares de nodos, esta topología puede no ser óptima para patrones de flujo diferentes. Por lo tanto, al cambiar las demandas de tráfico que debe soportar la red, es necesario realizar la reconfiguración de la topología virtual cuando mediante la eliminación de caminos de luz existentes y la adición de nuevos caminos. En este contexto se plantean dos objetivos, conseguir que la nueva configuración sea óptima y posteriormente determinar la forma más rentable de realizar los cambios (minimizando las interrupciones inevitables del servicio) o bien, cuando la configuración actual no permita soportar el tráfico demandado, determinar la forma más rentable de llegar a una nueva configuración que lo permita.

Redes con Supervivencia.

La alta capacidad que proporcionan las redes WDM da lugar a diseños de red más dispersos que los asociados con la tecnología de cobre. Consecuentemente, el tráfico conducido a través de un único enlace es significativamente mayor, y por lo tanto también lo es la interrupción de servicios si falla un enlace o nodo de red. Por tanto, además de la reducción de costes, la supervivencia contra fallos es uno de los aspectos más importantes durante el diseño de redes ópticas. La supervivencia de la red de comunicaciones es la posibilidad de restaurar los servicios de red en caso de un fallo catastrófico de la pérdida de conexión. Los proveedores de redes generalmente usan la supervivencia como un elemento competitivo con otros proveedores.

Los posibles fallos en una red óptica son producidos en los enlaces o en los dispositivos de interconexiones instalados en los nodos. Los fallos en enlaces son frecuentemente ocasionados por causas externas, mientras que los fallos de equipo en los nodos de red son debidos a causas internas.

Las técnicas de recuperación de las conexiones se clasifican en dos tipos: *restauración* y *protección*. Las técnicas de restauración deben restaurar el servicio reconduciendo dinámica y rápidamente el tráfico afectado usando la capacidad proporcionada en la red [66]. Las técnicas de protección usan rutas de protección para las demandas de tráfico que se determinan en el momento del establecimiento de las rutas de servicio [38], [116]. Dado que los servicios de red pueden ser generalmente restaurados en cuestión de horas o días en caso de que se produzca un fallo, los planificadores de redes no consideran la posibilidad de tener que afrontar más de un único fallo porque la probabilidad de otro fallo durante el período de reparación es pequeña. En los problemas de planificación de redes con sistemas WDM consideramos mecanismos de protección para fallos de un único enlace.

Los enfoques para proteger fallos de un único enlace están basados en dos mecanismos de supervivencia básicos: *protección de camino* y *protección de enlace*. La protección de camino asigna estáticamente un camino de protección entre los extremos de una conexión. La capacidad sobrante instalada en la red para proporcionar protección a un fallo puede ser o no compartida entre los distintos caminos de protección. Si no se permite compartir los recursos y se dedica un camino enlace-disjunto para el fallo, entonces se usa un *esquema de protección de camino dedicado* (también llamado protección 1+1). Si es posible compartir recursos, entonces se usa un *esquema de protección de camino compartido*.

En protección de enlace, el tráfico conducido a través del enlace defectuoso se reconduce alrededor de dicho enlace. Si durante el establecimiento de los caminos de servicio se reserva una ruta de protección entre los extremos del enlace que ha fallado, entonces se considera un *esquema de protección de enlace dedicado*. Si las rutas de protección pueden compartir la capacidad instalada sobre la red para proporcionar protección a la misma, entonces se considera un *esquema de protección de enlace compartido*.

Control y Gestión de Red

Un mecanismo de control debe ser capaz de determinar una ruta y asignar una longitud de onda a una conexión para responder dinámicamente a la demanda y configurar los centros de conexión a lo largo de la ruta. Los objetivos son maximizar el número de conexiones, minimizar los tiempos de establecimiento de las conexiones, y minimizar el ancho de banda usado por las señales de control.

Ventajas Económicas en Redes Ópticas de Múltiples Longitudes de Onda.

La tecnología DWDM puede reducir el coste de añadir capacidad de fibra en redes de larga distancia en más de un 50%. Se denomina tecnología DWDM (*dense wavelength division multiplexing*) o tenconolgía WDM densa a la que permite utilizar más de 8 longitudes de onda distintas. Algunas empresas de transporte competitivas tales como Qwest y Comunicaciones IXC están creando nuevos mercados de capacidad óptica mediante el arrendamiento de longitudes de onda específicas a otras empresas. Algunas empresas fabricantes de equipos tales como Nortel, Lucent, Ciena y Alcatel, invirtieron casi $3B en equipo WDM en 1998. Como dato, el equipo WDM/DWDM desplegado en Norte América ha permitido aliviar la congestión en la red de larga distancia experimentada en 1997 y 1998 debida a internet y al tráfico de datos. Quizás más importante es que la tecnología WDM supone una mejora para los usuarios en las redes de telecomunicaciones ópticas, en las cuales el tráfico que fluye puede ser agregado y conducido más eficientemente, y restaurado más rápida y fiablemente tras fallos en la red.

El crecimiento de las demandas y los cambios en los patrones de tráfico conducen a que la red de acceso local deba soportar diferentes interfaces de servicio y, al mismo tiempo, ser escalable, permitir asignar ancho de banda a las demandas y ser fiable. La red metropolitana entre oficinas no es menos dinámica, pues los cambios que tienen lugar en la red de acceso tienen un impacto directo sobre la red entre oficinas. Consecuentemente, es crucial una respuesta rápida ante cambios impredecibles.

Las ventajas económicas de la implantación de los sistemas DWDM comparadas con el tendido de nueva fibra han sido menos claras para redes de acceso local y área metropolitana que para los transportes de larga distancia. Sin embargo, muchas compañías de acceso local están empezando a ver a las redes metro ópticas como una alternativa atractiva para los empujes de la nueva fibra. Mientras tanto, los vendedores trabajan para reducir los costes suficientemente tal que las uniones WDM/DWDM y equipo asociado tal como OXCs, que conducen y cambian longitudes de onda, sean claramente ventajosos para redes de corta distancia (área metropolitana y acceso local).

CAPÍTULO 2.- Metaheurísticas

En este capítulo se estudian los aspectos fundamentales de las metaheurísticas en su aplicación a los problemas de planificación de redes con tecnología WDM.

Introducción

Las metaheurísticas son estrategias maestras que guían y modifican otras heurísticas más allá de lo que lo hacen normalmente las encaminadas a la optimalidad local. La Teoría de la Optimización se refiere a los estudios de las soluciones óptimas de los problemas y los métodos para encontrarlas. La formulación estándar del modelo de un problema de optimización es:

$$\text{mininimizar } f(s), \text{ sujeto a } g(s) \geq 0.$$

En este modelo f representa la función objetivo a minimizar y g el conjunto de restricciones a que están sometidas las soluciones. En la programación lineal, ambas son funciones lineales. Para estos problemas existen potentes resultados para su resolución. Por tanto, se considera como una forma usual de abordar los problemas de optimización la representación mediante un modelo lineal y resolverlo mediante un optimizador comercial. Además de la posible inadecuación de las funciones lineales para representar situaciones reales, la cuestión determinante es el número de variables y restricciones del modelo resultante. Se plantea el dilema usual de obtener una solución exacta de un modelo aproximado o una solución aproximada de un modelo más ajustado como pueden aportar las metaheurísticas.

El problema de la conducción y asignación de longitudes de onda (RWA-*Routing and Wavelength assignment*) aparece frecuentemente como subproblema relevante en la planificación de redes ópticas, y permite introducir de manera clara los elementos básicos de las metaheurísticas constructivas y las mejoras locales. Este problema sirve para mostrar la importancia de los procedimientos basados en establecer el orden en que se conducen las demandas en la red y las dificultades a que puede dar lugar. Se plantea el esquema básico de una búsqueda local de mejora y la adopción de la estrategia voraz o *greedy* en su conducción.

Metaheurísticas Básicas

Las principales metaheurísticas básicas se describen brevemente aportando los elementos fundamentales. Las estrategia de arranque múltiple o *Multistart* consistente en aplicar una búsqueda de mejora local desde una serie de soluciones de arranque. Aparte de su gran simplicidad, la característica fundamental es el riesgo de que se produzca la denominada "catástrofe" del *Teorema Central de Límite*. La forma de abordar esta dificultad aparece en estudios sobre las características topológicas del espacio de soluciones como el que hemos realizado en [95].

La metaheurística de Búsqueda Tabú (TS *Tabu Search*) representa la forma elemental de introducir la memoria histórica de la búsqueda en su conducción. Básicamente, se trata de mantener información selectiva del proceso de búsqueda para reemplazar el entorno de la solución actual por un entorno modificado por la historia y orientar los movimientos aplicados en el recorrido. La modificación del entorno implica prohibir ciertos atributos o valores en las variables, pero también incluir soluciones candidatas no contempladas en principio. La metaheurística también contempla la modificación de la función objetivo que guíe la selección de las soluciones del entorno.

La búsqueda dispersa (SS, *Scatter Search*) es una metaheurística basada en poblaciones que usa un conjunto de referencia para combinar sus soluciones y construir otras. El método genera un conjunto de referencia desde una amplia población de soluciones. Se seleccionan subconjuntos del conjunto de referencia cuyas soluciones son combinadas para obtener soluciones de arranque para mejoras locales. El resultado de estas mejoras puede dar lugar a una actualización del conjunto de referencia e incluso de toda la población de la que volver a extraer el conjunto de referencia.

La estrategia de búsqueda dispersa incluye seis procedimientos y tres criterios de parada. Se incluyen los procedimientos de creación de la población inicial, de generación del conjunto de referencia, de selección del subconjunto a combinar, de combinación de soluciones seleccionadas, de mejora de la solución combinada y de actualización del conjunto de referencia.

La metaheurística de recocido simulado (SA, *Simulated Annealing*) se ha convertido desde su introducción a mediados de los 80 en una metaheurística importante en la resolución de problemas de optimización. Esta metaheurística aplica una búsqueda global, que escapa de los mínimos locales mediante un criterio probabilístico de aceptación de soluciones. El procedimiento está inspirado en un proceso físico de aplicación industrial. Se engloba dentro de los procedimientos que establecen un umbral para aceptar movimientos de no mejora. El recocido simulado utiliza un umbral probabilístico de tipo exponencial controlado por un parámetro denominado temperatura. La cuestión fundamental en su aplicación se encuentra en la selección del mecanismo apropiado de actualización de este parámetro que proporcione garantías de convergencia.

La metahuerística de entorno variable (VNS, *Variable Neighborhood Search*), por el contrario, es una metaheurística reciente basada en un principio simple: el cambio sistemático de la estructura de entorno sobre la que se articula la búsqueda. Desde su introducción avanzados los años 90 han crecido considerablemente las aplicaciones exitosas en la resolución de problemas de gran tamaño. Se emplea una serie finita de estructuras de entornos diferentes que pueden venir generadas por algún concepto de distancia entre soluciones o por la aplicación anidada de una estructura de entornos básica. La metaheurística se basa en tres hechos básicos: los mínimos locales con respecto a una estructura de entornos no lo son necesariamente con respecto a otra, el óptimo global es un óptimo local con respecto a todas las posibles estructuras de entornos, y finalmente, el hecho de que los mínimos locales, con respecto a la misma o distinta estructura de entornos, tienen características similares.

Aplicando el principio del cambio de entorno sistemático en una búsqueda local descendente se obtiene la búsqueda descendente de entorno variable (VND, *Variable Neighborhood Search*). La versión básica de la búsqueda de entorno variable (BVNS, *Basic Variable Neighborhood Search*) usa cambios determinísticos en la estructura de entornos para perturbar o agitar la solución actual y proporcionar nuevas soluciones de arranque para la aplicación de una búsqueda local de mejora. La versión que recientemente ha obtenido mayores éxitos es la búsqueda general de entorno variable que sustituye en la versión básica la búsqueda local por una búsqueda descendente de entorno variable (VND). Este procedimiento utiliza dos familias (posiblemente distintas) de estructuras de entornos para la agitación y para la mejora o descenso. Se han propuesto diversas extensiones para solventar las dificultades que se observan, destacando la búsqueda de entorno variable con descomposición (VNDS, *Variable Neighborhood Decomposition Search*) que aplica el cambio sistemático de la estructura de entorno en dos niveles: el del problema original y el de una descomposición basada en los entornos.

Los algoritmos genéticos (GA, *Genetic Algorithm*) son la última de las metaherísticas básicas descritas. Se trata de una metaheurística de tipo evolutivo basada en una población de individuos que evoluciona de forma controlada sobre el espacio de búsqueda. Está inspirada en fenómenos naturales asociados a la evolución y se basa en una representación de las soluciones mediante ristras de símbolos denominados genes. Los elementos fundamentales de los algoritmos genéticos son las operaciones de cruzamiento y mutación. Estas operaciones utilizan la información de dos o una solución interpretada como individuo de una población para dar lugar a nuevos individuos. Estos individuos se seleccionan de acuerdo a la evaluación de su grado de ajuste en el que juega un papel fundamental la función objetivo del problema. Las generaciones de individuos hacen evolucionar la población mejorando sus individuos persiguiendo obtener soluciones de alta calidad.

Finalmente sólo se mencionan otras metaheurísticas importantes que han surgido en este campo.

Una cuestión clave para la aplicación de metaheurísticas en problemas reales es la posibilidad de explotar el paralelismo en su aplicación. Las metaheurísticas basadas en búsquedas locales pueden ser paralelizadas siguiendo, en principio tres tipos de estrategias. Una primera posibilidad consiste en paralelizar el tratamiento de los entornos distribuyendo diversas partes de los mismos entre los distintos procesadores. Esta sería la forma de abordar los procedimientos estándares de búsqueda local como la estrategia voraz. Otra posibilidad de aprovechar el paralelismo consiste simplemente en replicar la metaheurística en cada uno de los procesadores, que realizan búsquedas independientes. Ésta corresponde a la forma natural de paralelizar un híbrido entre la estrategia de arranque múltiple y la metaheurística de que se trate. La tercera vía alternativa intermedia consiste en involucrar en el paralelismo alguna de las componentes propias de la metaheurística. Estas tres vías han sido exploradas en [39] and [40] donde se consideran respectivamente paralelizaciones de la metaheurística de entorno variable (VNS) y la búsqueda dispersa (SS). Para la búsqueda de entorno variable (VNS) se considera la paralelización del proceso de agitación para generar soluciones de arranque para cada procesador que aplica la búsqueda con o sin actualización de la mejor solución encontrada. Para la búsqueda dispersa (SS) se considera la paralelización del proceso de combinación para proporcionar varios mínimos locales con los que actualizar el conjunto de referencia si fuera preciso.

El Papel de la Metaheurísticas

Las metaheurísticas proporcionan una forma de mejorar significativamente el rendimiento de los procedimientos heurísticos sencillos en la resolución de problemas difíciles. Se ha comprobado el éxito de las metaheurísticas al abordar problemas de complejidad y tamaños realistas en aplicaciones industriales y en el campo de las telecomunicaciones. Generalmente los mejores procedimientos alcanzan altas cotas de eficiencia aprovechando información del contexto.

El contraste entre los primeros procedimientos heurísticos orientados exclusivamente a la optimalidad local y las estrategias metaheurísticas de búsqueda global que escapan de este estancamiento hace cada vez más relevantes sus aplicaciones industriales. El impacto de la posibilidad de aportar soluciones de alta calidad con procedimientos metaheurísticos sencillos a problemas reales de gran complejidad ha sido importante.

La evolución actual del campo de las metaheurísticas está condicionada por las distintas interpretaciones de lo que constituye una búsqueda inteligente. Diversas características relevantes han contribuido al éxito de las distintas propuestas en mayor o menor medida. Por tanto, al contemplar la aplicación de una metaheurística híbrida que integre las características relevantes de diversos procedimientos procede realizar un análisis de la relevancia de estos aspectos en las distintas metehaurísticas. Entre tales características se encuentran el uso de memoria adaptativa para conducir la búsqueda, la elección del tipo de estructura de entorno sobre la que se articula la búsqueda o la interacción entre los atributos de las buenas soluciones para mejorarlas aún más.

Se analizan varios de estos aspectos como son, aparte de los ya mencionados, la necesidad de bordear la frontera de factibilidad para acercarse a los óptimos fronterizos, la estrategia de oscilación en los aspectos contrapuestos presentes en la búsqueda, la perturbación aleatoria de aspectos demasiado rígidos de las búsquedas o la inclusión de umbrales regulables dinámicamente. Algunos de los aspectos importantes que intervienen en el éxito del procedimiento híbrido son característicos de las metaheurísticas que se integran en el híbrido. Por ejemplo, en la búsqueda tabú son características la memoria adaptativa del proceso y la información que permite la exploración de regiones prometedoras. La memoria opera en cuatro dimensiones principales: la basada en frecuencia, la centrada en lo reciente, la de calidad del resultado y la de influencia en el proceso. Se contemplan dos tipos de memoria: la memoria explícita y la de atributos. Una componente importante en la búsqueda tabú, y en la práctica totalidad de las estrategias de búsqueda es la regulación entre la intensificación y diversificación de la búsqueda.

Entre las características relevantes de la búsqueda dispersa destaca la propuesta de realizar una combinación inteligente de las características de buenas soluciones en lugar de dejar esta responsabilidad al azar como ocurre con los algoritmos genéticos y otros algoritmos evolutivos. Es también relevante en la metaheurística de búsqueda dispersa el uso de un conjunto de referencia que contiene una representación de buenas soluciones. En el concepto de buenas soluciones intervienen no sólo cuestiones de calidad individual de las soluciones incluidas en el conjunto de referencia sino también la diversidad de las mismas. Sin embargo se trata de un conjunto de soluciones de un tamaño mucho menor que los de una población que incluso puede ser importante para obtener como resultado de la búsqueda, no sólo una única solución presumiblemente óptima, sino un conjunto disperso de tamaño moderado de soluciones de alta calidad. Estas soluciones pueden ser proporcionadas al decisor para que opte por una de ellas o utilizadas para fases posteriores de optimización. Esta última propiedad de la metaheurística de búsqueda dispersa se utiliza aquí para abordar el problema de supervivencia de la red WDM partiendo, no de una sóla red de servicio, sino de un conjunto de buenas redes de servicio.

Entre las características rentables de la metaheurística de arranque múltiple, aprovechadas en el diseño del procedimiento híbrido se encuentra la posibilidad de utilizar procedimientos constructivos inteligentes que aprovechen la información obtenida durante el proceso de búsqueda para realizar buenas elecciones. Estas características se apoyan en el principio de la persistencia atractiva y la validez marginal condicional.

Por último es importante realizar un análisis de las propiedades deseables de los procedimientos metaheurísticos. Una relación de tales propiedades es: sencillez, precisión, coherencia, eficiencia, efectividad, eficacia, generalidad, adaptabilidad, robustez, interactividad, multiplicidad, autonomía e innovación. Estas propiedades indican diversas direcciones en las que se pueden mejorar las prestaciones de las metaheurísticas. Algunas de ellas están estrechamente relacionadas o pueden incluso ser relativamente contrarias. Las propiedades de sencillez, precisión y coherencia posibilitan la rápida implementación de algoritmos rentables incluso para problemas poco estudiados. La eficiencia, efectividad y eficacia son tres aspectos en los que se puede analizar el rendimiento práctico de los algoritmos resultantes; hacen referencia al consumo de recursos computacionales, el grado de optimalidad de la solución aportada y la frecuencia con que se aporta la solución óptima. La generalidad, adaptabilidad y robustez son las que sustentan la amplia aplicabilidad de las metaheurísticas posibilitando la adaptación a distintas circunstancias, pero sin exigir un gran esfuerzo para las modificaciones pertinentes. La interactividad, multiplicidad y autonomía son las que propician su utilidad en los sistemas de ayuda a la decisión donde debe tener cabida el aprovechamiento del conocimiento experto del usuario por parte de la metaheurística, la posibilidad de proporcionar varias soluciones de alta calidad, pero diferentes y la conveniencia de que la metaheurística pueda funcionar sin agobiar al usuario con el ajuste de una gran cantidad de parámetros. Finalmente, las innovaciones asociadas a las metaheurísticas, tanto en el campo de aplicaciones exitosas como la incorporación de ideas novedosas facilita su popularización como ha ocurrido con las metaheurísticas inspiradas en fenómenos naturales estudiados por la física, la medicina, la economía o la biología.

CAPÍTULO 3.- El Problema de Provisión y Conducción (PRP)

Este capítulo presenta y describe las aplicaciones de un problema real que resulta de la necesidad de incrementar la capacidad de redes de telecomunicaciones ópticas. El contenido se organiza siguiendo el siguiente esquema:

- Realizar una introducción del problema de provisión y conducción o problema de síntesis y describir sus aplicaciones;
- Proporcionar una revisión bibliográfica del problema;
- Describir el problema;
- Desarrollar un modelo matemático y plantear algunas restricciones adicionales;
- Realizar una descripción detallada de un procedimiento metaheurístico híbrido basado en las estrategias multiarranque, búsqueda dispersa y búsqueda tabú; y
- Estudiar la efectividad del procedimiento propuesto anteriormente mediante comparación con otros métodos metaheurísticos y usando Cplex [28], que proporciona procedimientos exactos para resolver problemas de optimización lineal y entera.

En la literatura de las telecomunicaciones se han tratado principalmente dos problemas de flujo en redes: el problema de factibilidad y el problema de análisis. Dado un grafo y arcos con capacidades, el primer problema hace referencia a la factibilidad de un conjunto de flujos. Si además se considera un conjunto de requerimientos de demanda, el segundo problema hace referencia a la determinación de flujos factibles de tal manera que los requerimientos sean satisfechos. En esta memoria tratamos un tercer problema de flujo en redes: el problema de provisión (también referido como problema de síntesis o de dimensión), que consiste en minimizar el coste total de incrementar la capacidad de una red dada de manera que se satisfagan los requerimientos de demanda. En estos problemas, la topología física de la red y los requerimientos de demanda son conocidos y las variables de decisión están relacionadas con la capacidad que es necesario instalar en enlaces y nodos a mínimo coste.

Una topología de red existente, una estructura de costes y un conjunto de requerimientos de demanda caracterizan un ejemplo típico de un problema de provisión de red. La estructura de costes depende de cada situación, así como de la tecnología disponible. Cuando se resuelve este problema se asume que el sistema no incluye costes de conducción una vez que el equipo ha sido instalado.

Cuando el conjunto de requerimientos está formado por una única demanda, el problema de provisión y conducción se reduce a resolver un problema de camino más corto en un grafo con costes incrementales en los arcos. Los costes incrementales están asociados con el equipo requerido para conducir la demanda. Sin embargo, es más probable que el conjunto de requerimientos de demanda esté constituido por varios pares origen-destino. En este caso, las demandas son conducidas considerando la capacidad existente en la red actual. El problema de uso de la capacidad existente es un objetivo principal del estudio dado que el equipo instalado en enlaces y nodos para conducir una demanda bajo consideración, puede ser usada para conducir una demanda a ser considerada a continuación.

El problema de provisión tiene las siguientes aplicaciones adicionales en el mundo de las redes de telecomunicaciones: Problema Multi-hora, donde el conjunto de nodos en los que originan y finalizan las demandas depende de la hora del día; y Problema de Supervivencia, descrito anteriormente.

Problema de Provisión y Conducción en Redes WDM: Contribuciones Previas.

En la literatura hay varios trabajos relacionados con la expansión de capacidad en redes ópticas WDM. Estos trabajos pueden ser clasificados en dos categorías: el caso en el que se limita la fibra desplegada, donde el objetivo del problema de provisión es minimizar el número de longitudes de onda [4], [5], [108], [137]; y el caso en el que se limita el número de longitudes de onda por fibra, donde se trata de minimizar la cantidad de fibra requerida [109] o maximizar el tráfico establecido [114].

Caenegem, y otros [131] proponen una estrategia metaheurística basada en recocido simulado para incrementar la capacidad de redes WDM minimizando el coste total para una determinada demanda estática de tráfico. Consideran dos tipos diferentes de redes; redes que no usan conversión de longitudes de onda y redes que usan conversión en los nodos. Tratan además el problema de protección usando tres estrategias de reconducción en caso de fallo de un único enlace. Estos autores no estudian el problema con incertidumbre en las demandas de tráfico, que es uno de los objetivos de esta memoria.

Alanyali y Ayanoglu [2] se centran en métodos heurísticos para incrementar la capacidad sobre una red óptica WDM dado un conjunto estático de conexiones. Consideran que existe número fijo de longitudes de onda disponibles sobre cada fibra. Sin embargo, fijan un coste positivo para cada uno de los enlaces sin tener en cuenta los costes en los que se incurre al instalar las tarjetas de canal en cada canal WDM.

Kennington, y otros. [74] presentan un estudio empírico comparando soluciones a las que se les impone la restricción de continuidad con soluciones que permiten conversión para el problema de conducción y asignación de longitudes de onda en redes WDM. En [72] estos autores desarrollan un modelo de optimización y heurísticas para la versión sin conversión de longitudes de onda.

Kennington, y otros. en [73] estudian el problema de provisión y conducción en una red WDM considerando incertidumbre en las demandas. Para ello usan optimización robusta. Establecen el modelo para el PRP con demandas inciertas y fijan un presupuesto.

Cox, y otros. [27] proponen el problema de planificación en redes ópticas WDM que incluye simultáneamente los problemas de provisión, conducción y protección. Resuelven el problema usando un algoritmo genético, basado en la instalación de dispositivos ópticos en la red de forma incremental con el objetivo de minimizar el coste de conducción de cada demanda. Este algoritmo representa las soluciones mediante permutaciones de las demandas. Una permutación representa el orden según el cual se consideran las demandas, una a una, para propósitos de conducción. Dado que el equipo se instala para satisfacer la demanda actual, sin considerar las demandas que aún no han sido conducidas, cada permutación puede proporcionar un diseño de red diferente. Este enfoque no garantiza la existencia de una ordenación de las demandas que resulte en un diseño óptimo.

Descripción del Problema

El problema de optimización de expansión de capacidad trata con un conjunto de demandas que deben ser conducidas a través de una red óptica existente. Cada demanda está asociada con un nodo origen, un nodo destino y un tamaño, expresado en unidades OC-48. Cada demanda puede ser conducida enteramente sobre una o más fibras sin sistemas WDM instalados en ellas, sobre uno o más canales de un sistema WDM o puede ser cambiada de un WDM a otro a través de dispositivos OXC. El objetivo del problema es minimizar el coste total, que es la suma del coste de fibra adicional, de sistemas WDM y de equipos OXC.

Definimos el concepto de segmento como una secuencia de enlaces individuales que no pasan a través de ningún sistema OXC. En este caso, la fibra o sistema WDM pasa a través de cualquier nodo intermedio sin añadir o eliminar tráfico y sin requerimiento adicional de equipo en dichos nodos. Cada unidad OC-48 usa dos fibras sin WDM o un canal de un sistema WDM. Llamamos canal a la capacidad requerida por un OC-48. En un diseño de red óptimo, cada segmento debería seguir un camino de mínimo coste (con respecto al coste de fibra) desde su origen hasta su destino. Dado que el camino más corto entre cualesquiera dos nodos es un segmento potencial, la red de segmentos resulta en un grafo completo, que es intratable en la mayoría de los casos. Por lo tanto se genera un subconjunto de segmentos prometedores como una de las estrategias de la búsqueda.

Modelo Nodo-Segmento

En este capítulo presentamos un modelo de programación lineal entera para el PRP. Proporcionamos una formulación *nodo-segmento* para el problema con múltiples requerimientos de flujo sin incertidumbre en las demandas y sin protección. El modelo está basado en el propuesto por Cox, y otros en [27].

La topología de la red es representada por un grafo no dirigido $G = (N, E)$, donde N denota el conjunto de nodos y $E \subseteq N \times N$ denota el conjunto de segmentos. En esta formulación los enlaces y segmentos son equivalentes en el sentido de que representan una conexión entre dos puntos. El coste de usar un enlace individual o un segmento está incluido correctamente en la función objetivo. El conjunto de demandas está representado por $D = (o_1, d_1, R_1), (o_2, d_2, R_2), \cdots, (o_q, d_q, R_q)$, donde o_i representa el nodo origen, d_i el nodo destino, R_i el tamaño.

El coste de los segmentos se divide en dos componentes: el coste relacionado con la fibra y el coste relacionado con los canales si algún sistema WDM ha sido instalado en la fibra. Además, el coste de fibra se subdivide en tres cantidades diferentes: el coste de los sistemas WDM instalados sobre la fibra, que es una catidad fija; el coste de la fibra, que depende de la longitud del segmento; y el coste de los amplificadores, que depende del número de ellos instalados sobre la fibra. Estos tres costes diferentes se suman al coste total para cada fibra usada sobre cada segmento. Por cada canal WDM activo, deben sumarse al coste total el coste de las tarjetas de canal y del convertidor de longitud de onda. Finalmente, los costes de los nodos incluyen el coste de instalar el OXC y los puertos requeridos como origen y destino de cada canal, que puede ser un canal WDM o un par de fibras.

La formulación presentada usa las siguientes definiciones:

Datos de Costes

- C_e^F = coste de una fibra sobre el segmento e.
- $C_e^{W_j}$ = coste de un WDM tipo $j \in J$ en el segmento e.
- C^{O_l} = coste de un OXC tipo $l \in L$.
- C^{c_j} = coste de un canal de un WDM tipo j.
- C^{p_l} = coste de un puerto de un OXC tipo l.

Datos de Capacidad

- M^{w_j} = capacidad de una unidad WDM tipo j.
- M^{o_l} = capacidad de una unidad OXC tipo l.

Infraestructura existente

- g_e^j = canales WDM disponibles sobre sistemas WDM tipo j en el segmento e.
- h_n^l = puertos OXC disponibles en sistemas OXC de tipo l en el nodo n.

Variables de Decisión

- x_{ie} = cantidad de demanda i conducida sobre el segmento e.
- x_{ie}^F = cantidad de demanda i conducida sobre el segmento e en dirección directa.

- x_{ie}^R = cantidad de demanda i conducida sobre el segmento e en dirección inversa.
- f_e = número de pares de fibra sin WDM sobre el segmento e.
- w_e^j = número de unidades WDM tipo j en el segmento e.
- v_e^j = número de canales WDM tipo j en el segmento e.
- y_n^l = número de unidades OXC tipo l instaladas en el nodo n.
- u_n^l = número de puertos OXC tipo l instalados en el nodo n.

Función Objetivo

La función objetivo a ser minimizada es la suma de los costes de fibra (primer término), costes de WDM (segundo término) y costes de OXC (tercer término).

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \sum_{j \in J} \left(\left(C_e^F + C_e^{W_j} \right) w_e^j + C^{c_j} v_e^j \right) +$$
$$\sum_{n \in N} \sum_{l \in L} \left(C^{O_l} y_n^l + C^{p_l} u_n^l \right)$$

Restricciones

Las siguientes restricciones imponen que todas las demandas sean conducidas a través de la red, que no sea asignada a un enlace más demanda de la que su capacidad permita conducir y que los elementos de cambio tengan suficiente capacidad.

<u>Conservación de Flujo</u>

$$\sum_{e \in E, o_i = start(e)} x_{ie}^F + \sum_{e \in E, o_i = end(e)} x_{ie}^R - \sum_{e \in E, o_i = end(e)} x_{ie}^F -$$

$$\sum_{e\in E, o_i=start(e)} x_{ie}^R = R_i, \forall (o_i, d_i, R_i) \in D$$

$$\sum_{e\in E, d_i=end(e)} x_{ie}^F + \sum_{e\in E, d_i=start(e)} x_{ie}^R - \sum_{e\in E, d_i=start(e)} x_{ie}^F -$$

$$\sum_{e\in E, d_i=end(e)} x_{ie}^R = R_i, \forall (o_i, d_i, R_i) \in D$$

$$\sum_{e\in E, j=end(e)} x_{ie}^F + \sum_{e\in E, j=start(e)} x_{ie}^R - \sum_{e\in E, j=start(e)} x_{ie}^F -$$

$$\sum_{e\in E, j=end(e)} x_{ie}^R = R_i, \forall (o_i, d_i, R_i) \in D, \forall j \in N, j \neq o_i, d_i$$

$$x_{ie}^F + x_{ie}^R = x_{ie}, \forall (o_i, d_i, R_i) \in D, \forall e \in E$$

Capacidad de Segmento

$$\sum_{(o_i, d_i, R_i)\in D} x_{ie} \leq f_e + \sum_{j\in J} v_e^j, \forall e \in E$$

$$v_e^j \leq w_e^j M^{w_j} + g_e^j, \forall e \in E, \forall j \in J$$

Capacidad de OXCs

$$\sum_{n=end(e)} \left(f_e + \sum_{j\in J} v_e^j \right) \leq \sum_{l\in L} u_n^l, \forall n \in N$$

$$u_n^l \leq y_n^l M^{O_l} + h_n^l, \forall n \in N, \forall l \in L$$

Todas las variables son enteras no negativas.

La formulación estudiada tiene el siguiente número de variables:

$$|E| + 2|E||J| + 2|N||L| + 3|D||E|,$$

y el siguiente número de restricciones:

$$2|D| + |D|(|N| - 2) + |D||E| + |E| + |E||J| + |N| + |N||L|.$$

El número de variables y restricciones escala con el número de enlaces y nodos en la red y con el número de demandas.

La función objetivo precedente, las variables de decisión, y las restricciones especifican una versión formal del problema de expansión de capacidad. En la práctica, sin embargo, las demandas y los costes son típicamente inciertos, mientras que las opciones de tecnología disponible, tales como capacidades de los sistemas WDM y OXC, cambian frecuentemente al introducirse nuevos productos en el mercado. Por lo tanto, esta versión formal del problema es una aproximación de un problema de expansión más complejo, que incluye incertidumbre en los datos. Este problema es uno de los problemas en estudio de esta memoria.

Enfoque de Solución Metaheurístico

El objetivo de esta sección es presentar un procedimiento metaheurístico desarrollado para resolver el problema de provisión y conducción en redes WDM. Para problemas de planificación de pequeño tamaño, la formulación presentada anteriormente puede ser resuelta en una cantidad razonable de tiempo computacional. Sin embargo, la solución exacta de este modelo proporciona sólo una cota inferior del problema real pues no limita el número máximo de segmentos entre los puntos origen y destino de una demanda, lo cual es una restricción impuesta por las limitaciones de la tecnología actual.

Nuestro enfoque de solución usa la noción de *red base*, que inicialmente está formada por el diseño de red actual. Una red base es un diseño de red incompleto que no satisface el conjunto de requerimientos de demanda que debería ser capaz de satisfacer un diseño de red completo. A medida que el proceso itera, la red base evoluciona y el coste estimado de conducir una demanda a través de la red es más exacto. Una red base que ha sido evolucionada incluye equipo adicional, que ha sido añadido a la red base original. Cuando se considera una demanda que debe ser conducida por la red que ha evolucionado, esta demanda puede compartir la capacidad adicional con otros requerimientos de demanda, derivando en una estimación de costes más exacta.

El enfoque de solución propuesto genera una lista de caminos para cada demanda haciendo uso de una implementación eficiente del algoritmo de k-caminos más cortos. Los caminos obtenidos para cada demanda se obtienen calculando el coste incremental de conducir la demanda entera a través de la red base.

Los cuatro elementos básicos comunes a cualquier búsqueda heurística, independientemente de la estrategia considerada, son: la representación de la solución, el objetivo, la función de evaluación, y el mecanismo de movimiento.

- *Representación de la solución.* La construcción de una solución comienza con la selección de un camino para cada demanda. Tras la asignación de una demanda a un camino, se calcula el coste del diseño resultante. El coste se asocia al equipo requerido para satisfacer las demandas usando los caminos elegidos. Una solución está perfectamente determinada mediante una estructura de datos que almacena las asignaciones demandas a caminos y el equipo requerido en cada elemento de la red original.

- *Objetivo.* El objetivo del problema es minimizar la suma de los costes de fibra adicional, de equipo WDM, y sus equipos OXC terminales correspondientes.

- *Evaluación.* Cuando cada demanda ha sido asignada a un camino en su lista potencial, la evaluación de la solución consiste en calcular el incremento de capacidad requerido en los elementos de la red para conducir las demandas a través de los caminos asignados.

- *Mecanismo de Movimiento.* Cada solución tiene asociado un entorno, constituido por todas las soluciones factibles que pueden ser alcanzadas cambiando una demanda de un camino a otro.

La estrategia de solución propuesta es un método metaheurístico híbrido que combina ideas de la búsqueda dispersa [82], búsqueda multiarranque [51], y la búsqueda tabú [50]. La búsqueda tabú contribuye con un componente de memoria diseñada para evitar el ciclado de soluciones. La búsqueda dispersa incluye un mecanismo de generación de nuevas soluciones mediante la combinación de soluciones almacenadas en un conjunto de referencia. Finalmente, el componente multiarranque usa una memoria a largo plazo que permite construir nuevas soluciones en un rango amplio del espacio de soluciones.

El procedimiento híbrido desarrollado comienza con la generación de un conjunto de segmentos prometedores usando el algoritmo de camino más corto con distancias como pesos en las aristas. El procedimiento usa estos segmentos para ejecutar el algoritmo de generación de los k-caminos más cortos para cada demanda (usando costes incrementales como pesos en las aristas). Una vez determinada la red base original formada por segmentos, se obtiene la capacidad existente sobre la red, para asignar costes incrementales de conducción de cada demanda. Se genera un conjunto de referencia inicial haciendo uso de un método constructivo, que asigna las demandas a aquellos caminos que utilicen más eficientemente la capacidad existente en la base. Las soluciones del conjunto de referencia se ordenan según su coste total de menor a mayor. La solución actual, que al comienzo es la primera solución del conjunto de referencia, es usada para determinar el orden de las demandas según su coste por unidad. Se ejecuta una búsqueda local en el entorno de la solución actual usando la ordenación de las demandas determinada con anterioridad. Esto es, el primer movimiento candidato es reasignar la primera demanda dada por la ordenación. Si la nueva solución es mejor que la peor solución en el conjunto de referencia, entonces se actualiza el mismo. Si no existe ningún movimiento de mejora que involucre la reasignación de la primera demanda, entonces se considera la segunda demanda. Se itera el proceso hasta que la reasignación de alguna de las demandas produzca un movimiento de mejora. Si se consideran todas las demandas y no se produce ninguna mejora, se finaliza la búsqueda local. Una vez finalizada la búsqueda se analiza el conjunto de referencia. Si no se ha introducido ninguna nueva solución en el mismo, se reconstruye.

Numerosos estudios demuestran que los procedimientos metaheurísticos efectivos mantienen un balance entre diversificación e intensificación de la búsqueda. Para proporcionar este balance, se evoluciona la red base original usando la información proporcionada por el conjunto de referencia. Uno de los principales criterios usados en el procedimiento de evolución de la base está relacionado con el número de veces que aparece un segmento en los caminos asignados a las demandas en las soluciones del conjunto de referencia. El procedimiento usa información global (referida a todo el proceso de búsqueda) y local (referida al conjunto de referencia actual) mediante contadores que almacenan el número de canales usados en cada segmento con el objetivo de determinar dónde añadir equipo en la base actual. La diferencia entre el máximo número global y local de canales usados en cada segmento muestra su importancia. Cuanto menor sea esta diferencia, más importante es el segmento en el diseño de red final.

Cuando la red base ha sido evolucionada, la capacidad existente cambia y por tanto los costes incrementales de conducir las demandas a través de la red. Por lo tanto, se obtienen nuevamente los k-caminos más cortos con costes incrementales como pesos en las aristas para cada una de las demandas. Se genera una nueva solución actual, que usa de forma eficiente la capacidad añadida en la red base, y se ejecuta el procedimiento de búsqueda local desarrollado. El proceso finaliza después de un número especificado de iteraciones, donde una iteración es una búsqueda local.

Resultados Computacionales

Los experimentos realizados con datos reales y datos generados aleatoriamente muestran el mérito del procedimiento metaheurístico híbrido cuando se compara con un enfoque basado en permutaciones y con las cotas inferiores generadas al resolver el modelo matemático con Cplex. Usamos un test estadístico no paramétrico para comparar nuestro procedimiento y dos variantes del enfoque basado en permutaciones. El test muestra la efectividad de nuestra búsqueda local, la cual mejora las soluciones construidas con el enfoque basado en permutaciones hasta un punto en el que el método resultante es estadísticamente comparable con el metaheurístico híbrido propuesto.

CAPÍTULO 4.- El Problema de Provisión y Conducción: Un Modelo Alternativo

El propósito de este capítulo es proponer un modelo de programación lineal entera para resolver el problema de incremento de capacidad de una red existente a mínimo coste. Este modelo es referenciado como modelo *segmento-camino*, y tiene significativamente menos variables y restricciones que el modelo presentado anteriormente y propuesto por Cox, y otros. Este modelo usa las siguientes definiciones adicionales:

- A_n = conjunto de segmentos adyacentes al nodo n.
- J_{od} = conjunto de caminos posibles desde el origen o hasta el destino d, que pueden

ser usados para conducir la demanda (o, d).

- $x_p^{od} = 1$ si la demanda (o, d) es conducida a través del camino p, y 0 en otro caso.

El modelo es establecido como sigue:

Función Objetivo

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) +$$
$$\sum_{n \in N} \left(C^O y_n + C^p u_n \right)$$

Restricciones

$$\sum_{p \in J_{od}} x_p^{od} = 1, \forall (o, d) \in D$$

$$\sum_{(o,d) \in D} R_{od} \sum_{p \in J_{od}, e \in p} x_p^{od} \leq v_e + f_e, \forall e \in E$$

$$v_e \leq M^W w_e + g_e, \forall e \in E$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N$$

$$u_n \leq M^O y_n + h_n, \forall n \in N$$

Todas las variables de decisión son enteras no negativas.

El objetivo es minimizar la suma de los costes de fibra, sistemas WDM y equipos OXC. El primer conjunto de restricciones impone la satisfacción de las demandas. El siguiente convierte la capacidad sobre los caminos a capacidad sobre los segmentos y ésta a fibras y canales. Los tres conjuntos de restricciones siguientes convierten la capacidad de los segmentos en unidades WDM, acumulan los canales sobre los enlaces para añadir el número de puertos requeridos en cada nodo, y convierten la capacidad de los nodos en unidades OXC, respectivamente.

CAPÍTULO 5.- El Problema de Supervivencia

El propósito de este capítulo es resolver el problema real de incremento de capacidad de una red óptica a mínimo coste incorporando protección ante posibles fallos de enlaces. Desarrollamos varios modelos con el objetivo de resolver el problema usando un esquema basado en enlace compartido y un esquema basado en camino compartido. Cuando se usa el primer esquema, se reserva un camino de protección alrededor de cada enlace, permitiendo que las rutas de protección de distintos enlaces compartan la capacidad reservada. Cuando se usa el segundo esquema de protección, se determina una ruta de protección para cada demanda cuyo camino de trabajo contenga un enlace defectuoso.

Esquema de protección de enlace compartido

Presentamos un modelo segmento-camino que obtiene las rutas de protección para cada enlace defectuoso comenzando con una solución al modelo sin protección. El modelo usa las definiciones siguientes:

- NC_e = número de canales sobre el segmento e requeridos para obtener los caminos de trabajo. Entonces, si falla el segmento e, el tráfico NC_e debe ser reconducido entre los nodos origen y destino de e.
- J_e = conjunto de rutas de protección alternativas entre los nodos extremo del segmento e.
- $y_q^e = 1$ si el tráfico sobre el segmento e es reconducido sobre el camino $q \in J_e$, y 0

en otro caso.

El modelo asume que el tráfico sobre un segmento no puede ser dividido al ser reconducido. El modelo se establece de la siguiente forma:

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right)$$

Sujeto a:

$$\sum_{q \in J_e} y_q^e = 1, \forall e \in E$$

$$\sum_{q \in J_{e'}, e \in q} NC_{e'} y_{q,e'} \leq v_e + f_e, \forall e \in E, \forall e' \neq e \in E$$

$$v_e \leq M^W w_e + g_e, \forall e \in E$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N$$

$$u_n \leq M^O y_n + h_n, \forall n \in N$$

El objetivo es minimizar el coste total, que es la suma del coste de fibra (primer término), el coste de los equipos WDM (segundo término), y el coste de las unidades OXC (tercer término).

Cuando falla el segmento e, el tráfico de trabajo sobre ese segmento debe ser reconducido a través de una de las posibles rutas de protección en J_e. Por lo tanto, se debe satisfacer el primer conjunto de restricciones. El segundo conjunto de restricciones asegura que, si falla el segmento e, la capacidad sobrante sobre otros segmentos debe ser suficiente para conducir el flujo que circula sobre las rutas de protección.

Esquemas de protección de camino compartido

En este esquema, es necesario restaurar cada camino de trabajo interrumpido. Una de las rutas de protección alternativas debe conducir el tráfico del camino de trabajo correspondiente a una demanda. Proponemos dos esquemas de protección de camino compartido: un esquema de protección en el que cada camino de trabajo tiene diversos conjuntos de rutas de protección alternativas dependiendo del enlace que haya fallado, y un esquema de protección en el que las rutas de protección y los caminos de trabajo correspondientes no pueden tener ningún enlace común. En este caso, sólo se genera un conjunto de rutas alternativas para cada camino de trabajo.

El modelo para el primero de estos esquemas usa las siguientes definiciones:

- $J_{e,od}$ = conjunto de rutas de protección alternativas entre el origen y el destino de la demanda $(o,d) \in D$ que no contienen al segmento e.
- $y_{q,e}^{od} = 1$ si la demanda (o,d) es reconducida sobre el camino $q \in J_{e,od}$ si el segmento e falla, y 0 en otro caso.
- WP_{od} = camino de trabajo de la demanda $(o,d) \in D$.

El modelo asume que la demanda conducida sobre cada camino de trabajo no puede ser dividida cuando se reconduce. El modelo se establece como sigue:

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right)$$

Sujeto a:

$$\sum_{q \in J_{e,od}} y_{q,e}^{od} = 1, \forall (o,d) \in D, \forall e \in WP_{od}$$

$$\sum_{(o,d) \in D, e' \in WP_{od}} \sum_{q \in J_{e',od}, e \in q} R^{od} y_{q,e'}^{od} \leq v_e + f_e, \forall e, e' \in E, e' \neq e$$

$$v_e \leq M^W w_e + g_e, \forall e \in E$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N$$

$$u_n \leq M^O y_n + h_n, \forall n \in N$$

El modelo para el segundo esquema de camino compartido usa las siguientes definiciones:

- J_{od} = conjunto de rutas de protección alternativas entre los nodos origen y destino de la demanda $(o, d) \in D$.
- $y_q^{od} = 1$ si la demanda (o, d) es reconducida sobre el camino $q \in J_{od}$ si falla el segmento e, y 0 en otro caso.

El modelo, cuyo objetivo es minimizar el coste total, se establece como sigue:

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right)$$

Sujeto to:

$$\sum_{q \in J_{od}} y_q^{od} = 1, \forall (o, d) \in D$$

$$\sum_{(o,d) \in D, e' \in WP_{od}} \sum_{q \in J_{od}, e \in q} R^{od} y_q^{od} \leq v_e + f_e, \forall e, e' \in E, e' \neq e$$

$$v_e \leq M^W w_e + g_e, \forall e \in E$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N$$

$$u_n \leq M^O y_n + h_n, \forall n \in N$$

En ambos modelos, las variables de decisión son enteras no negativas.

El problema de incremento de la capacidad de una red WDM para proporcionar supervivencia en la misma ha sido abordado en la literatura después de obtener la mejor red de trabajo que permite satisfacer los requerimientos de demanda. En este capítulo hemos propuesto una variante de la metaheurística híbrida desarrollada anteriormente. Este metaheurística, basada en la búsqueda dispersa, está dividida en dos fases. La primera fase resuelve el problema de expansión de capacidad para satisfacer el conjunto de demandas dadas, proporcionando un conjunto de referencia formado por las mejores redes de trabajo obtenidas durante la búsqueda. En la segunda fase, se considera cada una de las soluciones del conjunto de referencia y se resuelve el problema de protección comenzando con ellas. Dado que el problema de protección es también un problema de expansión de capacidad, pero que permite que la demanda sea compartida para restaurar fallos de distintos enlaces, se usa la metaheurística híbrida sin más que cambiar el método de evaluación de las soluciones.

Resultados Computacionales

Finalmente, se realiza la experiencia computacional usando los tres esquemas de protección. El primer experimento llevado a cabo en este capítulo consiste en resolver la formulación matemática presentada para el esquema de protección de enlace compartido. Para ello, se consideran como soluciones de partida las mejores redes de trabajo obtenidas para los ejemplos reales y aleatorios presentados en los capítulos anteriores. Además, se resuelven los mismos ejemplos usando el metaheurística híbrida, que proporciona, para cada ejemplo, un conjunto de referencia de redes con supervivencia.

Los resultados obtenidos demuestran que el mejor diseño de red obtenido, que permite satisfacer los requerimientos de demanda y que proporciona supervivencia, no es obtenido, en la mayoría de los casos, comenzando con la mejor red de trabajo, sino una red de trabajo con mayor coste.

El siguiente experimento realizado en este capítulo consiste en resolver el problema de protección para los esquemas de camino compartido. Para cada uno de estos esquemas se utiliza una versión modificada de la metaheurística híbrida que permite evaluar correctamente las soluciones. Se comparan los resultados obtenidos y se concluye que el esquema de protección que proporciona mejores diseños de red es el esquema de camino compartido en rutas de enlaces disjuntos.

CAPÍTULO 6.- Problema de Provisión y Conducción en redes WDM bajo Incertidumbres en las Demandas.

Una consideración importante al planificar redes ópticas WDM es la incertidumbre. Claramente, las demandas futuras de ancho de banda no se pueden considerar conocidas en un entorno como la industria de las telecomunicaciones.

Los algoritmos y modelos usualmente asumen que los datos son conocidos exactamente. Sin embargo, en la mayoría de las aplicaciones del mundo real esto no sucede. Este hecho es particularmente importante cuando los datos representan demandas futuras de tráfico y costes de productos que no se conocen con certeza. Las técnicas usuales que permiten tratar problemas de optimización con incertidumbre en los datos son el análisis de sensibilidad, la optimización difusa, la programación estocástica, y la optimización robusta.

Modelo de Programación Estocástica

En este apartado usamos el marco de modelización de la programación estocástica con el propósito de desarrollar un modelo usado para resolver el problema de expansión de capacidad con incertidumbre en las demandas. Denotamos por $S = 1, \cdots, s^*$ el conjunto de posibles escenarios de demandas de tráfico. El modelo usa las siguientes definiciones:

- z_{ods} = *bajo aprovisionamiento* para cada escenario $s \in S$ y cada demanda $(o,d) \in D_s$. El *bajo aprovisionamiento* es la cantidad de demanda *ods* que no puede ser conducida usando la capacidad instalada en la red.
- P_s = probabilidad de ocurrencia del escenario s.
- Cu = coste de *bajo aprovisionamiento*.
- $x_p^{ods} = 1$ si la demanda $(o,d) \in D_s$ es conducida sobre el camino p, y 0 en otro caso.

Para cada coste de *bajo aprovisionamiento* Cu, el objetivo a ser minimizado es la suma del coste de diseño más un coste de penalización.

$$\min \sum_{e\in E} 2C_e^F f_e + \sum_{e\in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n\in N} \left(C^O y_n + C^p u_n \right) +$$

$$Cu \left(\sum_{s\in S} P_s \sum_{od\in D_s} z_{ods} \right)$$

Sujeto a:

$$\sum_{p\in J_{od}^s} x_p^{ods} R_{od}^s + z_{ods} = R_{od}^s, \forall (o,d) \in D_s, \forall s \in S$$

$$\sum_{(o,d)\in D_s} \sum_{p\in J_{od}^s, e\in p} x_p^{ods} R_{od}^s \leq v_e + f_e, \forall e \in E, \forall s \in S$$

$$v_e \leq M^W w_e + g_e, \forall e \in E$$

$$\sum_{e\in A_n} (v_e + f_e) \leq u_n, \forall n \in N$$

$$u_n \leq M^O y_n + h_n, \forall n \in N$$

$$0 \leq z_{ods} \leq R^s_{od}, \forall (o,d) \in D_s, \forall s \in S$$

Desarrollamos una metaheurística de búsqueda dispersa para resolver el problema de expansión de capacidad con incertidumbres en las demandas. Comparamos los resultados obtenidos al resolver esta formulación mediante un optimizador comercial con los resultados proporcionados por un enfoque de resolución basado en búsqueda dispersa. La metaheurística permite considerar funciones de penalización, y por tanto, funciones objetivo tanto lineales como no lineales.

El enfoque de programación estocástica presentado obtiene soluciones que pueden no ser factibles para alguno o ninguno de los escenarios. Por este motivo, proponemos además un modelo robusto basado en el enfoque robusto propuesto por Kouvelis y Yu [76], que genera soluciones factibles incluso para el peor de los escenarios. El modelo se establece como sigue:

$$\min y$$

Sujeto a:

$$\sum_{e \in E} 2C^F_e f_e + \sum_{e \in E} \left(\left(C^F_e + C^W_e \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right) \leq y + z^s,$$
$$\forall s \in S$$

$$\sum_{p \in J^s_{od}} x^{ods}_p = 1, \forall (o,d,s) \in D_s, s \in S$$

$$\sum_{p \in J^s_{od}, e \in p} x^{ods}_p R^s_{od} \leq v_e + f_e, \forall e \in E, \forall s \in S$$

$$v_e \leq M^W w_e + g_e, \forall e \in E$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N$$

$$u_n \leq M^O y_n + h_n, \forall n \in N$$

Dado que para resolver este modelo se precisa tener la solución óptima para cada uno de los escenarios, se resuelve este modelo mediante un optimizador usando aquellos ejemplos para los que fue posible obtener la solución óptima en un tiempo de ejecución de dos horas.

Conclusiones (español)

La planificación en redes de telecomunicaciones ópticas constituye un campo de estudio muy reciente, mientras se trata de sustituir las redes existentes basadas en anillos físicos por mallas ópticas y mientras evoluciona la tecnología óptica. En esta memoria se aborda la resolución metaheurística del problema real de utilización efectiva de nueva tecnología diseñada para incrementar la capacidad en redes de telecomunicaciones ópticas. El problema surge cuando la demanda de ancho de banda en una red óptica excede la capacidad existente en la red. La tecnología de multiplexación por división de longitudes de onda (WDM) proporciona una alternativa eficiente para expandir la capacidad de redes ópticas existentes. El problema de decisión consiste en determinar la mejor combinación de nueva fibra y sistemas de multiplexación por división de longitudes de onda que permita incrementar la capacidad a mínimo coste satisfaciendo ciertos requerimientos de demanda.

En la planificación de redes de telecomunicaciones, como en otros campos de aplicación industrial, surgen problemas de optimización a los que hay que dar respuesta adecuada en un tiempo moderado. La modelización matemática y el uso de optimizadores comerciales proporciona soluciones óptimas para algunos problemas específicos y de pequeño tamaño, pero no es rentable para los problemas complejos de mayor tamaño que surgen en circunstancias reales. Las metaheurísticas tratan de dar respuesta a estas cuestiones evitando métodos heurísticos excesivamente dependientes de la estructura del problema considerado y procedimientos diseñados para cada problema particular. Una metaheurística es una estrategia que guía y modifica otras heurísticas con el propósito de generar soluciones más allá de las generadas normalmente en una búsqueda de optimalidad local. Los procedimientos derivados de las metaheurísticas, diseñados para obtener óptimos globales, no garantizan que la solución obtenida al finalizar la búsqueda sea la solución óptima del problema. Sin embargo, permiten obtener soluciones de alta calidad en problemas reales difíciles con un coste computacional moderado.

Las conclusiones más importantes obtenidas durante la elaboración de esta memoria son expuestas a continuación.

1. Tras la realización de un estudio en el capítulo 1 sobre los problemas reales que surgen en general en el área de las telecomunicaciones, y en particular los que aparecen al planificar redes con tecnología de multiplexación por división de longitudes de onda, se identifica el problema a considerar como uno de los más relevantes.

2. En la revisión bibliográfica, recogida en el capítulo 3, se constata que el problema de expansión de capacidad en redes ópticas mediante el uso de tecnología WDM presentado en esta memoria no ha sido abordado suficientemente. El trabajo más importante propone un modelo de programación lineal entera precisamente para este problema y lo aborda mediante el desarrollo de un algoritmo genético basado en permutaciones de las demandas. Incluso una búsqueda exhaustiva del espacio de permutaciones de las demandas puede no proporcionar la solución óptima del problema, por lo que se precisa un enfoque de solución alternativo.

3. En otros trabajos se resuelven problemas similares al tratado en esta memoria simplemente proponiendo un modelo matemático y usando un optimizador comercial para ejemplos de pequeñas dimensiones. Los optimizadores comerciales no proporcionan soluciones óptimas a los modelos de programación matemática propuestos para llevar a cabo la resolución del problema sobre redes reales en un tiempo computacional razonable.

4. Un estudio de las características y propiedades deseables de las metaheurísticas, realizado en el capítulo 2, permite seleccionar las más apropiadas para abordar el problema que surge en la planificación de redes ópticas WDM y desarrollar sistemas híbridos que aprovechen las ventajas de cada una de ellas.

5. En el capítulo 2 también se constata que la paralelización de metaheurísticas es una de las vías en las que se puede mejorar el rendimiento de las mismas en su aplicación práctica a problemas reales.

6. Para evitar las deficiencias establecidas en los puntos anteriores sobre el tratamiento que han recibido los problemas de planificación de redes de telecomunicaciones, se desarrolla en el capítulo 3 un procedimiento metaheurístico híbrido que combina las buenas propiedades y características de la búsqueda dispersa (scatter search), de la búsqueda multiarranque (multistart), y de la búsqueda tabú (tabu search).

7. Se comprueba la efectividad del metaheurístico híbrido al ser comparado con dos variantes de un procedimiento metaheurístico basado en permutaciones y con las soluciones óptimas proporcionadas por un optimizador comercial al resolver el modelo propuesto por otros autores.

8. Se propone en el capítulo 4 un modelo heurístico de programación lineal entera alternativo con el fin de determinar la capacidad requerida para conducir un conjunto de requerimientos de demanda a través de una red. Este modelo tiene la ventaja de incluir significativamente menos variables y restricciones que la formulación anterior.

9. Del análisis comparativo entre los resultados obtenidos usando el metaheurístico híbrido y los resultados obtenidos resolviendo la formulación matemática de forma óptima, se concluye que el modelo alternativo propuesto permite resolver ejemplos reales de dimensiones mayores que el anterior.

10. El problema de supervivencia en redes de telecomunicaciones ópticas de múltiples longitudes de onda se aborda en el capítulo 5. Se proponen tres modelos matemáticos alternativos para resolver el problema usando tres esquemas de protección diferentes.

11. De la revisión bibliográfica de este problema, presentada en el capítulo 1 de esta memoria, se observa que el problema de expansión de capacidad para proteger la red ante cualquier fallo de un elemento de la misma se aborda una vez obtenido el diseño de red que permite satisfacer el conjunto de demandas. El estudio realizado para concluir si la mejor red de servicio o red de trabajo (la que satisface el conjunto de requerimientos de demanda) proporciona la mejor red con supervivencia lleva a la conclusión de que esto no sucede en la mayoría de los casos.

12. Se propone una versión modificada de la metaheurística híbrida desarrollada en el capítulo 3, que incluye características de la búsqueda dispersa y proporciona un conjunto de referencia de buenas redes de servicio alternativas. Entonces se resuelve el problema de protección considerando cada diseño de red del conjunto de referencia como solución de partida. Se corrobora que la mejor red que permite satisfacer las demandas y que además protege el tráfico en caso de fallo, no se obtiene, en la mayoría de los ejemplos, al considerar la mejor red de servicio. Se concluye además que la mejor alternativa para garantizar la supervivencia es uno de los esquemas de protección propuestos.

13. Los modelos desarrollados en los capítulos anteriores y las herramientas metaheurísticas correspondientes se extienden para contemplar una importante consideración en esta área que es la de abordar la incertidumbre presente en las aplicaciones reales.

14. De la revisión bibliográfica se deduce que es la primera vez que se considera precisamente este problema con incertidumbre en las demandas.

15. Se proponen varios modelos matemáticos con el fin de resolver el problema de expansión de capacidad con demandas inciertas.

16. Se propone un modelo de programación estocástica, un procedimiento metaheurístico

que permite considerar funciones objetivo lineales y no lineales, y un enfoque robusto que obtiene soluciones factibles incluso para el peor caso.

Líneas Futuras (español)

Las posibles líneas de investigación futura se resumen en los siguientes puntos.

- La aplicación de paralelismo a los procedimientos metaheurísticos híbridos desarrollados para resolver el problema de expansión de capacidad en redes ópticas WDM.

- El diseño de metaheurísticas híbridas para abordar otros problemas relevantes que surgen al planificar redes ópticas WDM, tales como el problema de conducción y asignación de longitudes de onda y el problema de localización de convertidores.

- La aplicación de metodología *Fuzzy* para tratar el problema de provisión y conducción bajo incertidumbres en las demandas.

Aportaciones (español)

Esta memoria incluye los resultados de diversos trabajos de investigación que han sido presentados en Congresos Internacionales o en Revistas Internacionales de los campos de las heurísticas y de las telecomunicaciones.

El capítulo 2 introduce un estudio detallado de algunos procedimientos metaheurísticos básicos. Algunas metaheurísticas, tales como la búsqueda multiarranque, permiten incorporar las ventajas de la estructura global de la función objetivo de determinados problemas. Esto permite identificar un conjunto de regiones del espacio de soluciones en las cuales un procedimiento de búsqueda local desde cualquiera de sus soluciones converge con alta probabilidad a un óptimo local determinado. Estos resultados han sido desarrollados en el artículo "A Multistart Clustering Technique for Combinatorial Optimization" presentado en el congreso internacional MS'2000 y publicado en la serie *The best of MS2000 International Conference on Modelling and Simulation* (2000).

El capítulo 2 también describe la posibilidad de considerar las ventajas proporcionadas por los métodos de paralelización, que permiten incrementar la exploración del espacio de soluciones o bien reducir el tiempo de computación total. Estas ventajas han sido probadas mediante el uso de las metaheurísticas de búsqueda dispersa y de entornos variables. Los artículos The Parallel Variable Neighborhood Search for the p-median Problem" (2002), que ha sido publicado en la revista *Journal of Heuristics*, y "Parallelization of the Scatter Search for the p-median problem" (2003), que ha sido publicado en la revista *Parallel Computing* son los resultados de esta investigación. Además, un artículo sobre búsqueda tabú, resultado de la colaboración con el profesor Fred Glover, aparecerá en la revista internacional *Inteligencia Artificial*.

El capítulo 3 introduce el problema de provisión y conducción que surge en la planificación de redes ópticas WDM. Con el propósito de resolver este problema eficientemente se desarrolló un procedimiento metaheurístico híbrido, que combina características de la búsqueda dispersa, del multiarranque y de la búsqueda tabú. El híbrido ha sido comparado con un procedimiento basado en permutaciones teniendo en cuenta las consideraciones propuestas por otros autores y con los resultados obtenidos al resolver de forma óptima el modelo propuesto por otros autores. Los resultados obtenidos, que corroboran la efectividad del método, se recogen en el artículo "Minimizing the Cost of Placing and Sizing Wavelength Division Multiplexing and Optical Cross-Connect Equipment in a Telecommunications Network", que está en segunda revisión en la revista *Networks*. Una versión preliminar de este trabajo fue presentada en el congreso *9th International Conference on Telecommunication Systems, Modelling and Analysis*, en Marzo de 2001.

En el capítulo 4 se propone un modelo matemático que incluye un menor número de variables y restricciones que el propuesto por otros autores. El metaheurístico híbrido se comparó con los resultados obtenidos al resolver la formulación propuesta con un optimizador existente, y los resultados se recogen en el artículo "Capacity Expansion of Fiber Optic Networks with WDM Systems: Problem Formulation and Comparative Analysis". Una versión preliminar de este trabajo fue presentada en el congreso internacional de telecomunicaciones *Symposium on Informatics and Telecommunications September* (SIT'02) y una versión mejorada fue presentada como conferencia invitada en el congreso INFORMS Annual Meeting 2002. El artículo ha sido aceptado para su publicación en la revista *Computers and Operations Research*.

Finalmente, el documento de trabajo "Provisioning of Survivable WDM Mesh Networks Under Demand Uncertainty and Single Link Failure Protection", que incluye los resultados de los capítulos 5 y 6, será presentado en el congreso *International Network Optimization Conference* (INOC2003), que tendrá lugar el próximo mes de octubre.

Preface

In the last four decades, problems related to planning and routing in telecommunication networks have become a fertile ground for developing and applying optimization techniques. Two main events have driven these efforts: (1) the large investment in telecommunication, which offers significant opportunities for reducing costs and improving network designs, and (2) the rapid changes in technology, which result in new operational environments.

Several developments that occurred during the last decade as well as the increase in the volume of demand have triggered a new era in telecommunications where analog technology is being substituted with digital technology. The replacement of traditional copper wires with optical transmission equipment not only increased bandwidth but also reduced operational cost.

It is evident that the world of networking in telecommunications is rapidly changing as new technologies are introduced. Where once demand was measured in kilobits, it is now measured in megabits. Where customers demanded megabits, they now demand gigabits. And where gigabits of capacity are now available, customers will soon be demanding terabits. As economies evolve from an industrial past to an information-based future, the economics of businesses and institutions are also shifting on their axes and the demand for more bandwidth is becoming a key driver in most societies.

Optical fiber technology has become the preferred choice to build telecommunication networks, due to its cost effectiveness, reliability and its almost unlimited capacity. Such combination enables telecommunication companies to offer a multitude of new services that require large amounts of bandwidth. Technology advances have motivated the development of appropriate planning tools that respond to the characteristics of a new structural and operational environment.

An *optical network* is a telecommunications network with optical fibers as transmission links, and with an architecture that is designed to exploit the features of fibers.

Despite the fact that the fiber has extremely high bandwidth, since the rate at which an end user can access the network is limited by electronic speed, only speeds of a few gigabits per second can be achieved. Therefore, the optical-electronic conversion does not allow to exploit the bandwidth of a single fiber. *Wavelength Division Multiplexing* (WDM) and *the erbium-doped fiber amplifier* (EDFA) are two recent developments that overcome these limitations.

Wavelength Division Multiplexing is the transmission of multiple laser signals at different wavelengths (colors) in the same direction, at the same time and over the same strand of fiber. While traditional technologies such as Synchronous Digital Hierarchy (SDH) or its American equivalent Synchronous Optical Network (SONET) are often based on networks consisting of interconnected rings, WDM technology usually does not restrict to any special network connection. The DWDM system is equipped with amplifiers that allow transmission in one channel. The most important recent development has been the erbium-doped fiber amplifiers (EDFAs), which amplify signals at many different wavelengths simultaneously.

The current optical communications networks do not have enough capacity to cover the bandwidth demands. WDM and its related equipment are a new technology for increasing the capacity of an existing fiber network while reducing costs. The problem that we address in this dissertation is a real world problem that results from the need to expand capacity of telecommunication networks built with fiber optics technology. This problem is referred to as the *Provisioning and Routing Problem in WDM networks.*

The high capacity in optical fiber links provided by WDM systems results in network designs that are sparser than those associated with copper technology. Consequently, the traffic routed through a single link is significantly larger in optical fiber networks and so is the disruption of services if any single link or node were to fail. Therefore, along with cost reduction, survivability against failures is one of the most important aspects that are considered when designing fiber networks. Survivability is defined as the ability to restore network services in the event of a catastrophic failure such as a faulty link or a defective node.

An important consideration for additional research in this area deals with tackling uncertainty. Clearly, the demands cannot be considered known in an environment such as the telecommunications industry. The availability of a MIP formulation that can be used to find near-optimal solutions to the capacity expansion problem represents a stepping-stone toward the solution of a stochastic version of the problem that treats the demands as uncertain.

Chapter 1 provides an overview on optical telecommunication networks and the problems that arise when planning optical networks as well as a literature review. It also provides a description of the optical technology that is used in this dissertation.

Chapter 2 presents an introduction of Metaheuristics, that are master strategies that guide and modify other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality. Metaheuristics provide a means for approximately solving complex optimization problems, as those that arise in Communication Network Planning. These methods are designed to search for global optima. However, they cannot guarantee that the best solution found after termination criteria are satisfied is indeed a global optimal solution to the problem. Experimental testing of metaheuristic implementations show that the search strategies embedded in such procedures are capable of finding solutions of high quality to hard problems in industry, business and science.

Chapter 3 introduces the general provisioning and routing problem, giving a wide description without considering the technology that has to be installed, including its applications in the real world, and a detailed literature review. This chapter also provides a literature review and a description of the provisioning and routing problem in WDM optical networks. It provides a mathematical model and a hybrid metaheuristic solution approach, that is then compared with others metaheuristics and with the model solved to optimality. Chapter 4 provides an alternative mathematical model for solving the problem. The proposed metaheuristic in then compared with the model solved to optimality. The Survivability problem in WDM optical networks is approached in Chapter 5, which provides several integer linear programming models for some protection schemes and proposes solution procedures. Finally, the provisioning and routing problem under demand uncertainties is approached in Chapter 6, in which several alternative mathematical models are developed.

Acknowledgements

I would like to thank Professors José A. Moreno Pérez and Manuel Laguna for their guidance in my optical networking and combinatorial optimization research and for helping me to become a more mature researcher. I deeply appreciate their time and effort.

I would like to acknowledge Manuel Laguna and people in the Leeds School of Business at the University of Colorado.

I wish to thank professors Marcos Moreno Vega and Félix García López for their valuable collaboration on several research topics and their contributions to Chapter 2.

I thank my colleagues from the *Departamento de Estadística, Investigación Operativa y Computación* for the nice working environment.

I would also like to thank to *Universidad de La Laguna*.

Finally, I thank my family and friends for their constant support and encouragement throughout my life.

To all these wonderful people I owe a deep sense of gratitude especially now that this project has been completed.

Chapter 1

Optical Telecommunications Networks

This chapter provides an introduction of Optical Telecommunication Networks. The main aim is to present different problems that arise when planning WDM optical networks.

1.1 Introduction

Telecommunications networks and systems have evolved and grown at an explosive rate in the last decades creating new opportunities for modeling and the application of optimization in planning problems. Our society requires that we have access to information *when we need it* and *where we need it*. The information is provided to us through the global mesh of communication networks, whose current implementations do not have the capacity to support the bandwidth demands.

Several developments that occurred during the last decade as well as the increase in the volume of demand have triggered a new era in telecommunications where analog technology is being substituted with digital technology. The replacement of traditional copper wires with optical transmission equipment has not only increased bandwidth but also reduced operational cost. During this time, the industry also has experienced both increased competition between providers of telecommunication services and the evolution of a wide range of user services that combine voice, data, graphics and video. In addition, corporative intranets and extranets have felt the need to increase the connectivity and bandwidth among several locations of a company, its clients and its providers.

The world of networking in telecommunications is rapidly changing as new technologies are introduced. As economies evolve from an industrial past to an information-based future, the economics of businesses and institutions are also shifting on their axes and the demand for more bandwidth is becoming a key driver in most societies.

Most of the telecommunication networks are divided into three main levels [3], [41], [93]:

1. The long distance or backbone network that typically connects pairs of cities through the gateway nodes;
2. The metropolitan interoffice network that interconnects switching centers in different clusters (groups of clients) and allows the access to the gateway nodes; and
3. The local access network that connects individual clients, that belong to a cluster, to the corresponding switching center.

These three different levels in a telecommunication system differ in several aspects, including their design criteria. Ideally, the design of a telecommunication network should simultaneously consider these three levels. However, due to its complexity, the overall planning problem is typically divided into subproblems, where each level is considered separately.

Two special network flow problems have been the focus in the telecommunications literature: the feasibility problem and the analysis problem. Given a graph and capacitated links, the first problem refers to the feasibility of a set of flows. If a set of demand requirements is also considered, the second problem refers to the determination of feasible flows such that the demand requirements are satisfied. In this thesis, however, we deal with a third type of network flow problem: the network synthesis or provisioning problem, which consists of minimizing the total cost of installing capacity on links of a given network so that demand requirements are satisfied. In this problem, both the physical network topology and the demand requirements are given and the decision variables relate only to adding capacity to links at minimum cost. When the problem includes also the design of the network topology, that is, determining which links to install, then a complete graph and installation costs are considered.

A given network topology, a cost structure and a set of demand requirements characterize a typical instance of a network synthesis problem. The cost structure depends on the problem at hand and on the available technology. It is customary to assume that the system does not add routing costs once the equipment has been installed. Also, these problems typically deal with commodities involving a single source and a single sink. A requirement between two nodes is a single commodity flow requirement. Multicommodity flow requirements are also considered as long as the commodities involve different origins and destinations while sharing the capacity of the network.

1.2 Optical Networks

Networks allow a geographically distributed community of users to communicate and share information. Although most of the telecommunication networks in use today have optical fiber in them, this does not make them optical networks. In almost all cases in which fiber is used, it is deployed in transmission links as a substitute for copper wires. An *optical network* is a telecommunications network with optical fibers as transmission links, and with an architecture that is designed to exploit the features of fibers. These architectures involve combinations of both electronic and optical devices. Therefore, the term optical or lightwave network does not necessarily imply purely optical network, but it does imply more than a set of fibers terminated by electronic switches.

Whenever a new technological development comes on the scene, there is a technology push toward deploying new systems that make use of it. This has been the case with fiber optic technology. Since the fabrication of the first low-loss optical fiber in a Corning lab in 1970, which made fiber practical for communications, the optical global information industry has grown at an explosive rate. Ever since this event, a vision of an all-optical network has intrigued researchers, service providers, and the general public. Despite the deployment of optical fiber throughout the world, its capacity has not been exploited efficiently.

Optical fiber technology is nowadays the premier component in telecommunication networks, due to its cost effectiveness, reliability and its, for all practical purposes, unlimited capacity. Such combination enables telecommunication companies to offer a multitude of new services that require large amounts of bandwidth. Technology advances have motivated the development of appropriate planning tools that respond to the characteristics of a new structural and operational environment.

A large percentage of the total fiber capacity remains unused. One reason is that the distribution network, that is the bridge between the high-speed fiber backbone and the end users, is still in a relatively primitive state at this time. Another reason is that the fibers in the ground are not organized into an architecture that makes their huge capacity available for new broadband services. In addition, economic, legal, administrative and political impediments make it difficult to succeed in using fiber capacity. However, the main barrier is that the demand for services that could be supported by a new optical infrastructure awaits the realization of that infrastructure, but the massive investment required to realize it will not materialize until the potential investors see some sign of a market for these new services.

1.3 Wavelength Division Multiplexing - WDM

In theory, fiber has extremely high bandwidth (about 25 million MHz) in the 1.55 low-attenuation band. However, since the rate at which an end user can access the network is limited by electronic speed, only speeds of a few gigabits per second can be achieved. Therefore, it is not possible to exploit the bandwidth of a single fiber due to the optical-electronic conversion. *Wavelength Division Multiplexing* (WDM) and *the erbium-doped fiber amplifier* (EDFA) are two major recent developments that overcome these limitations.

Wavelength Division Multiplexing is the transmission of multiple laser signals at different wavelengths (colors) in the same direction, at the same time and over the same strand of fiber. WDM with more than eight frequencies, called Dense Wavelength Division Multiplexing (DWDM), which creates multiple bi-directional 'virtual fibers' per physical fiber, currently enables a low cost per bit [116]. DWDM solves the bandwidth bottleneck resulting from growth in data traffic, because it is an emerging technology that increases transportation capacity while preserving optical fiber equipments previously installed. Hence, DWDM provides carriers the flexibility and scalability they need to deploy capacity when and where it is needed. While traditional technologies such as Synchronous Digital Hierarchy (SDH) or its American equivalent Synchronous Optical Network (SONET) are often based on networks consisting of interconnected rings, WDM technology usually does not restrict to any special network connection, i.e., it assumes a so-called mesh topology.

The DWDM system is equipped with amplifiers that allow transmission in one channel. The most important recent development has been the commercialization of erbium-doped fiber amplifiers (EDFAs), which amplify signals at many different wavelengths simultaneously, regardless of their modulation scheme or speed.

Most of the capacity cost when using DWDMs is related to the channel cards, which are added as needed and their cost is charged to the design accordingly. This means that a system capable of handling up to 96 channels can be installed where only eight channels are active, and the design would only consider the cost of equipping the eight active channels. To use WDM technology, an equipment “unit” must be placed at both endpoints of each fiber link. For each wavelength, or channel in use, channel equipment must be also placed at both endpoints of the channel. Each WDM channel is bidirectional and has the same capacity as a pair of fibers. Amplification is the process restoring the optical signal to its original optical power and without distortion after the signal has lost power when passing through a strand of fiber. This process is particularly important in DWDM environments. The typical amplifiers do not have electronic elements, that is, they are completely optical. Consequently, they do not require the classical electrical-optical and optical-electrical conversions, thereby avoiding the associated need for additional bandwidth.

1.3.1 Layers of the multiwavelength network architecture

The operations research literature approaches the network planning problem as a multilayered model due to the size and complexity of the networks. Figure 1.1 shows the different layers and sublayers of a multiwavelength network. The higher layers represent the logical network. The architecture presented in this section is representative of a Wide Area Network (WAN). In smaller networks, some of the functions may not exist.

The logical-physical layer interface is located at the external ports of a network access station (NAS), which connects user terminals and other nonoptical end systems to the network. Then, a NAS is the boundary between the electronic and optical domain. The interface between the physical and logical layers is in the ports of the NASs. Each NAS connects an optical network node (ONN) through one or more pairs of fibers. Each node is connected to other nodes using pairs of fibers, which are the network links. The graph that consists of the network links and the network nodes is called the physical topology of the network. If the network link is too long, in order to avoid the attenuation in

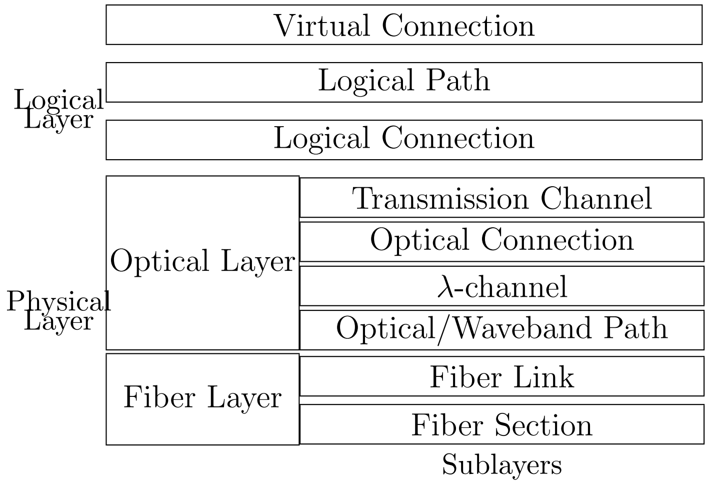


Figure 1.1: Layered view of optical network connections

the fiber, optical amplifiers are installed, so that a *fiber link* consists of a series of *fiber sections* between amplifiers.

The sublayers in the optical layer include the λ-*channels*, each of which has an assigned wavelength and is routed separately by the optical network node. Each point-to-point *optical connection* (OC) is routed on a λ-channel. The transceivers at origin and destination nodes must be tuned to a selected wavelength. In addition, the optical connection is created by establishing an *optical path* (OP) through a sequence of network nodes to carry that wavelength from the origin to the destination.

A *logical connection* (LC) is a connection between ports on a pair of source and destination network access stations. Each logical connection is carried on an optical connection through the intermediary of a *transmission channel*. The optical connections are established through the actions of the nodes and stations. The nodes also create an optical path on the assigned wavelength.

The optical layer is subdivided in several sublayers to provide multiplexing, multiple access at several layers, and switching. Multiplexing allows to combine various logical channels on a λ-channel originating from one station. The multiple access allows λ-channel originating from distinct stations carry multiple logical connections to the same destination. Finally, through switching, distinct optical paths may be created on different fibers in the network, using λ-channels on the same wavelength.

An optical path is routed on a succession of fiber links connecting network nodes. Each of the internodal connections along an optical path is called an *optical hop*.

Figure 1.2 shows the concepts described above. A virtual connection (VC) between a pair of end systems is carried on a logical path (LP), which consists of two logical connections terminated by logical switching nodes. The first and second LCs are carried on two- and three-hop optical connections, respectively. The access links are not counted as hops.

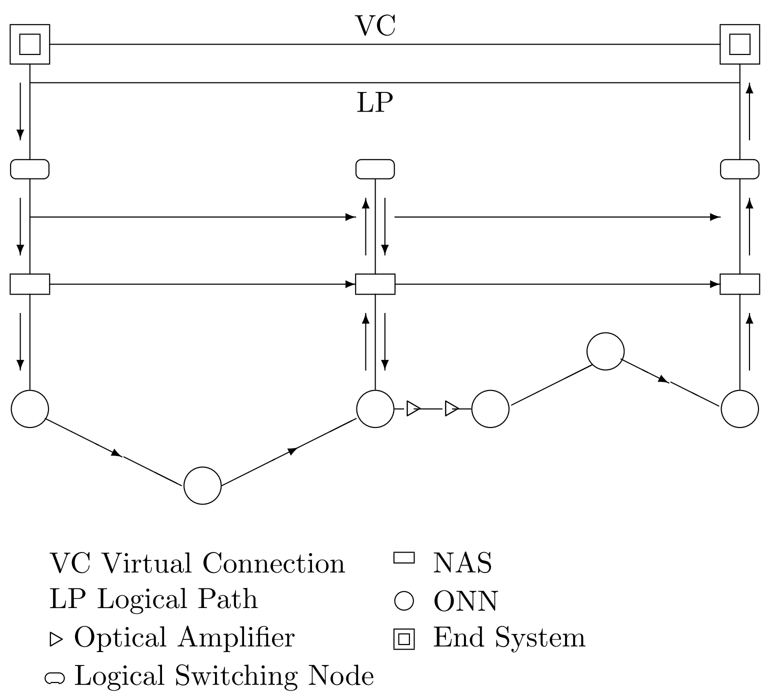


Figure 1.2: Typical connection

1.3.2 Network Links

Each network link can support a number of concurrent connections through successive levels of multiplexing. Typically, a link consists of a cable containing several bidirectional pairs of fiber. This way if installing the fibers is an instance of *space-division multiplexing* in the fiber layer. We consider wavelength division as the multiplexing technique used to carry on each fiber several connections on many distinct wavelengths (channels). The wavelengths assigned to each channel must be spaced sufficiently to avoid signal overlapping, that is, interference among the signals at the optical receiver.

1.3.3 Optical Network Nodes

The functions of the optical path layer are implemented in the optical network nodes. Therefore, the more functionality these nodes have, the more flexible the network is in reacting to fluctuating user demand, changing loads, and equipment problems. The different types of node functionality, classified in increasing order of complexity, are the following:

- Static nodes: Directional Couplers and Static Routers.

  A 2×2 directional coupler is an optical four-port, represented as in Figure 1.3, with ports 1 and 2 designated as input ports and $1'$ and $2'$ designated as output ports. An optical signal enters the coupler through fibers connected to the input ports and leaves through fibers connected to the output ports. Static nodes can also be constructed by interconnecting 2×2 couplers.

  A static router has n input and n output fibers, each carrying up to n distinct channels with different wavelengths. Each input fiber is connected to an $1 \times n$ wavelength demultiplexer (WDMUX), which separates the wavelengths on the fiber. Each output fiber is connected to an $n \times 1$ wavelength multiplexer (WMUX), which combines the different wavelengths on that fiber. The WDMUXs, WMUXs and their interconnecting fibers act as a single node.

- Dynamic (switching) nodes.

  The simplest dynamic node is a space-division switch, that is commonly called *optical cross-connect* (OXC). WDM networks are capable of switching data optically by using optical cross-connects (OXCs). An optical cross-connect can be one of the following types:

  1. OXCs with opto-electro-opto (OEO) conversion, which are equipped with transmitters and receivers and convert data from optical domain to electronic domain, switch data with an electrical switch and convert data back to optical domain. This type of OXCs are called "OEO switches".
  2. All-optical OXCs, which have photonic switch fabrics and switch data entirely in the optical domain. All-optical OXCs are also called "all-optical switches", "OOO OXCs", and "OOO switches".

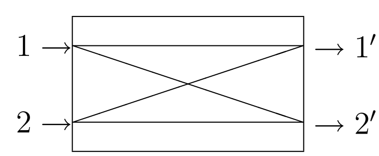


Figure 1.3: Directional Coupler

When an optical signal has to remain on a fixed wavelength from end to end, the condition called *wavelength continuity* is satisfied. One possible way to overcome this constraint is to use wavelength converters inside an all-optical switch. A wavelength converter is an optical device capable of shifting an input wavelength to a possibly different output wavelength among the W wavelengths in the system (see Figure 1.4). Wavelength converters relax the continuity constraint at a node. Therefore, they reduce the number of wavelengths needed to route a set of demands, resulting in better bandwidth utilization. Wavelength convertes have been referred to in the literature as wavelength shifters, wavelength translators, wavelength changers and frequency converters.

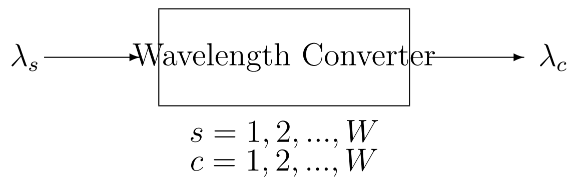


Figure 1.4: Functionality of a Wavelength Converter

For the purpose of illustrating the wavelength continuity constraint, consider a wavelength routed network with five nodes and two wavelengths per fiber as shown in Figure 1.5.

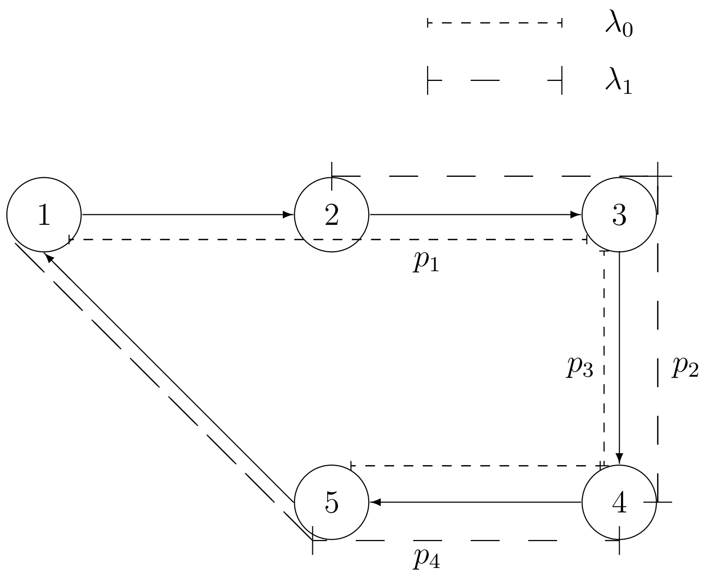


Figure 1.5: Wavelength Routed Network

The mechanism of communication in a wavelength network is a *lightpath*, which is an all-optical communication channel between two nodes in the network and may span on more than one fiber link. Lightpaths have to be established for the node pairs $(1, 3)$, $(2, 4)$, $(3, 5)$, $(4, 1)$, and $(5, 2)$. There is only one physical path for each node pair. Figure 1.5 shows the lighpaths for node pairs $(1, 3)$, $(2, 4)$, $(3, 5)$, $(4, 1)$, which are denoted by p_1, p_2, p_3, and p_4, respectively. However, due to the wavelength continuity constraint it is not possible to establish lightpaths for all five node pairs. Despite the fact that bandwidth is available on links $5 \rightarrow 1$ and $1 \rightarrow 2$, the wavelengths on these two links are different and therefore no lightpath can be obtained for the node pair $(5, 2)$. This bandwidth loss caused by the mentioned constraint can be overcome by using a wavelength converter.

1.4 WDM Optical Network Architectures

This section summarizes how connectivity is created, taking into account the relations among physical constraints, device functionality, and connectivity. Based on the level of controllability in network nodes, three classes of WDM optical network architectures are considered: broadcast-and-select networks, wavelength routed networks, and linear lightwave networks.

1.4.1 Broadcast-and-Select Network

In the broadcast-and-select method, a receiver can make a logical connection with any transmitter by selecting the information which is interesting for it and discarding the rest. Different nodes transmit messages at the same time on different wavelength and the coupler combines all messages and broadcasts the combination to all nodes. A node selects a wavelength to receive its message.

1.4.2 Wavelength Routed Network (WRN)

A Wavelength Routed Network includes either static or dynamic wavelength selectivity in the network nodes, wherein wavelength reuse becomes possible through the use of appropriate connection control algorithms. These networks are not restricted to a particular physical topology. A wavelength can be selected in the optical network node and routed individually. The nodes contain transceivers tunable over the range of wavelengths. Then the optical paths are point-to-point and to get multipoint connectivity it is necessary the use of multiple point-to-point optical connections (WDM) and an optical transceiver for each connection. There also could be wavelength conversion in the network.

In a WRN, the traffic is sent from one node to another using a lightpath without requiring any optical-electronic-optical conversion. The lightpath is routed through the intermediate nodes using their wavelength cross-connects. The origin and destination nodes for a lightpath use a transceiver and a receiver that are tuned to the wavelength on which the lightpath operates. As mentioned before, if the wavelength continuity constraint is imposed, the same wavelength must be used on all the links along the selected route.

1.4.3 Linear Lightwave Networks (LLN)

The properties of the LLN are more general than the properties of WRN, since they are waveband rather than wavelength routed, and they can support multipoint optical connections in addition to point-to-point connection. The routing nodes in a WRN switch wavelength, whereas in a LLN switch wavebands, but not wavelengths within a waveband. Then the number of optical switches at nodes reduces from the number of wavelengths to the number of wavebands. The individual wavelengths in a waveband are separated at the end nodes using optical receivers.

1.5 Issues in Wavelength-Routed Networks

This section briefly describes some relevant issues in wavelength-routed networks including routing and wavelength assignment, minimizing the bandwidth loss due to wavelength continuity constraint, design, reconfiguration, and survivability of virtual topology, and control and management.

1.5.1 Routing and Wavelength Assignment (RWA)

The RWA problem is one of the fundamental subproblems when designing wavelength routed networks (WRNs). The wavelength assignment problem consists in allocating an available wavelength to a connection and tuning the transmitter and receiver to the assigned wavelength. Solving the routing problem consists of determining a path for the selected wavelength and setting the switches in the intermediate nodes. Since an optical path is associated to a wavelength, such path cannot be established until a wavelength is allocated. A survey for the RWA problem may be found in [142], [143].

A connection is supported by a lightpath in a wavelength-routed WDM network and it may span multiple fiber links. The traffic demands between pairs of nodes can be either static or dynamic. In static routing, traffic requirements is a set of known point-to-point demands. In this case, the goal is to assign routes and wavelength to each requirement minimizing the number of wavelengths used. This problem, that is also known as *Static Lightpath Establishment* (SLE) problem, is NP-complete [24] and, therefore, algorithms that reach a solution close to the optimal in reasonable computational time are used. The SLE problem under the *wavelength continuity constraint* can be formulated as an Mixed Integer Linear Program (MIP) [114] in which the objective is to minimize the flow on each link. The wavelength continuity constraint increases the blocking probability, that is the percentage of connections rejected.

In dynamic routing, traffic requirements have to be routed or not in a random manner.

The established lightpaths remain in the network for a finite time. The dynamic problem arises under changes in traffic patterns or network components failures. The goal of this problem is to set up lightpaths and assign wavelengths in order to maximize the number of connections established in the network. This problem is also called *Dynamic Lightpath Establishment* (DLE). Dynamic RWA algorithms reach worst solutions than static RWA algorithms since a dynamic algorithm has no knowledge about future demand requirements.

The SLE problem can be divided into two subproblems: (1) routing and (2) wavelength assignment; and each problem can be solved independently. Banerjee, et al. [4] proposed algorithms to solve the SLE problem for large networks, and used graph-coloring algorithms to assign wavelengths to lightpaths once the they have been routed. The DLE problem is more difficult to solve and heuristics need to be employed. Algorithms for solving the routing subproblem can be broadly classified in three types: *fixed routing*, *alternate routing*, and *exhaust routing* [90], [117]. The fixed routing provides for each node pair only a candidate route, which is computed offline. The route for a node pair does not change with changing traffic conditions. The alternate routing generates, for each node pair, a set of k candidate routes, which are also computed offline. For the wavelength assignment subproblem, several heuristics have been proposed [15], [23], [67], [69]. These heuristics are *Random Wavelength Assignment*, *First-Fit*, *Least-Used/SPREAD*, *Most-Used/PACK*, *Min-Product*, *Least Loaded*, *MAX-SUM*, *Relative Capacity Loss*, *Wavelength Reservation*, and *Protecting Threshold*.

1.5.2 Wavelength-Convertible Networks

To overcome the bandwidth loss due to the wavelength continuity constraint, wavelength converters can be installed at network nodes. Wavelength routed networks with this capability are referred to as *wavelength-convertible networks* [115]. A node with wavelength conversion capability is called a *wavelength converting node*. If a wavelength converter can convert any wavelength to any other (it is said that the converter has full degree of conversion) and if there is a wavelength converter for each fiber link in every node of the network, then the network is said to have *full wavelength-conversion capabilities*. If a wavelength-convertible network with full wavelength-conversion capabilities is considered, then only the routing problem needs to be addressed, and the wavelength assignment is not an issue.

Wavelength-convertible networks reduce the bandwidth loss due to the wavelength continuity constraint. However, such a network is too costly since the converters are expensive. This leads to questioning how many nodes should be wavelength converting nodes, how do we choose the converting nodes, and *where* (optimally) to place a few converters in an arbitrary network.

1.5.3 Virtual Topology Design (VTD)

The virtual or lightpath topology (optical layer), which is superposed on the fiber layer, consists of a set of lightpaths established between a subset of node pairs in the network. The RWA problem and the VTD problem are similar, except that the design has to be done within the optical layer.

Two nodes can communicate on one-(light)hop if they are connected by a lightpath. On the other hand, if they are connected by a sequence of lightpaths, they communicate on muti-(light)hop. In this case, the connection between two consecutive lightpaths is performed via electronic processing.

The virtual topology is designed to route the traffic in such a way as to optimize a certain performance metric. This metric can be the average message delay, which can be a measure of the number of lightpaths traversed by it, or network congestion, which can be the maximum traffic carried by any lightpath.

1.5.4 Virtual Topology Reconfiguration

Since the virtual topology is designed to satisfy the estimated traffic between the node pairs, this topology may not be optimal for different patterns of flow. Therefore, it is necessary to perform the reconfiguration of the virtual topology when the traffic pattern changes by removing existing lightpaths and adding new ones. Wavelength cross-connects, which allow changing the switching patterns of wavelengths, facilitate the process of reconfiguration.

Two different approaches have been proposed to migrate one topology to another avoiding expensive service disruption. The first approach consists of designing an optimal topology for the new traffic pattern and obtaining the number of steps required to migrate the current topology to the new one. The second approach selects the feasible topology that requires the minimum changes to the current topology.

1.5.5 Survivable Networks

The high capacity in WDM networks results in network designs that are more sparse than those associated with copper technology. As a result, the traffic routed through a single link is significantly larger and so is the disruption of services if any single link or node were to fail. Hence, survivability, together with fault protection and restoration, is a growing area of concern.

Along with cost reduction, survivability against failures is one of the most important aspects that are considered when designing fiber networks. Survivability is defined as the ability to restore network services in the event of a catastrophic failure such as a faulty link or a defective node. Network providers typically use survivability as a competitive differentiator or offer it as a premium service.

Failures in an optical network can be links or switching devices failures. Links faults are often caused by external causes, whereas equipment failures in the network nodes are due to internal causes.

The fault-recovery techniques can be roughly classified in two types: *restoration* and *protection*. *Restoration* techniques should reroute the affected traffic dynamically and rapidly using the capacity provided in the network [66]. *Protection* techniques pre-compute in advance backup routes for each possible failure [38], [116]. If only single failures are considered, protection techniques can be used, whereas restoration is used if two or more failures occur at the same time. However, since services can be restored in hours or days, most network planners do not consider more than a single node or link failure when designing for survivability because the probability of another failure during the repair period is small. We consider only protection mechanisms for single-link failures.

Some criteria can be assumed in order to classify the protection techniques. A first classification criterium may be associated with the network topology. Two alternatives can be distinguished: ring network protection and mesh network protection. The ring network architecture has been widely used to implement survivability in the event of single link or node failures, because inherently it provides two different paths between any two points in the network. Although rings have desirable features in terms of network survivability, users who are transporting sensitive information across the network might be reluctant to be on the same ring as another user in a different building.

The approaches to survive single-link failures in an optical network are based on two basic survivability mechanisms: *path protection/restoration* and *link protection/restoration*.

Path protection statically allocates a backup path between the origin and destination nodes of a connection (see Figure 1.6). In path restoration, the origin and destination nodes dynamically find a backup route after the link failure. If sharing among backups is not allowed and a link-disjoint backup path is dedicated to the fault, then we are under a *dedicated-path protection scheme* (also called 1+1 protection). The switches on the backup paths can be configured at the beginning. This type of recovery is very fast, but is not capacity efficient. If sharing among backups is allowed, then we are under a *shared-path protection scheme*. A link-disjoint backup path is also reserved to recover the fault. The switches on the backup paths cannot be configured until the fail occurs. The recovery time is longer, but the system is more capacity efficient.

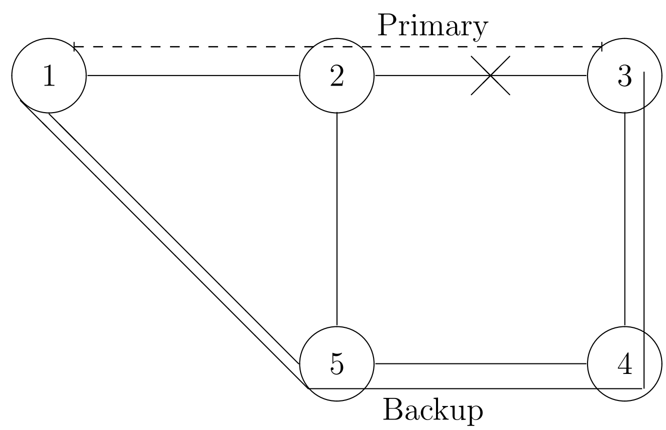


Figure 1.6: Path Protection

In link protection, all the traffic that is routed on the failed link is rerouted around that link (see figure 1.7). During demand setup, backup paths are reserved around each link in the primary path. In link restoration, the origin and destination nodes of a link dynamically reroute the traffic around the link. If at the time of demand setup, a backup path is reserved around each link of the primary path, this is known as a *dedicated-link protection scheme*. In *shared-link protection*, at the time of demand setup, for each link of the primary path, a backup path is reserved around the link, allowing other backup paths to share the reserved capacity.

There have been several authors who have applied ring-protection schemes in WDM mesh networks. One such approach include the development of p-cycles (preconfigured

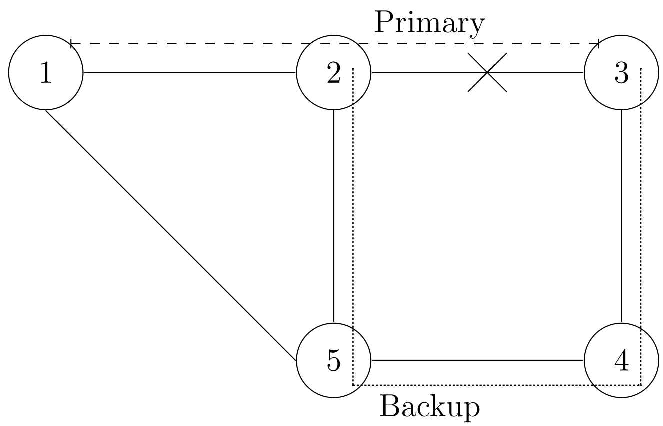


Figure 1.7: Link Protection

protection cycles). The p-cycle concept was introduced by Grover and Stamatelakis in [58], [59], [129]. Techniques based on ring protection offer very fast recovery times, but the ratio between the spare capacity and the primary resources is at leat 100% [59]. However, for mesh protection techniques, the ratio can typically be in the range of only 50-70% [59]. Using p-cycles both the efficiency of mesh protection and the recovery speed of ring networks are achieved [58]. The applicability of the p-cycle concept has been focused on SDH/SONET Networks. Schupke, et al. in [125] study the efficiency of p-cycles applied to WDM networks with and without wavelength conversion. The results obtained by these authors in particular for wavelength converting networks show that p-cycles achieve high efficiency. Their goal is to minimize the cost of using the spare channels to generate the p-cycles used to protect the network. However, they do not consider the possibility of increasing the capacity of the network by adding more fiber, WDM and OXC systems. If wavelength conversion is not allowed, then the problem becomes more complex since it is necessary to take care of the individual wavelengths of the system.

1.5.6 Optical Multicasting Routing

So far it has been assumed that the traffic demands are point-to-point. However, in many real applications the traffic has to be routed from an origin to several destinations. This one-to-many communication is referred to as *multicasting*, which is used for video conferencing, distance learning, real time work groups, etc.

1.5.7 Network Control and Management

The control mechanism should be able to find a route and assign a wavelength to the connection to respond dynamically to customer demand and configure the switches along that route. The objectives are to maximize the number of connections, minimize the connection setup times, and minimize the bandwidth used for control signals.

1.6 Economic Advantages in Multiwavelength Optical Networks

WDM technology can reduce the cost of adding fiber capacity in long-haul carrier networks. There is no question about the usefulness of this technology, as it would cost millions to lay new fiber to meet constantly increasing demand. In fact, the first systems in the market were optimized for long-haul inter-exchange applications. Competitive inter-exchange carriers have already created new markets for optical capacity by leasing specific wavelengths to other carriers. The WDM/DWDM equipment deployed in North America has already helped to relieve long-haul network congestion experienced in 1997 and 1998 as Internet and data traffic put unprecedented demands on existing capacity. Perhaps more importantly, WDM is enabling a new "optical layer" of telecommunication networks, in which traffic flows can be aggregated and routed more efficiently (by suitable wavelength grooming) and restored more quickly and reliably after network failures.

The local network presents several challenges to carriers. Rapid growth in demand and changes in traffic patterns imply that the access network must support many different service interfaces and at the same time, it must be scalable, it must be able to allocate bandwidth on demand, and it must be reliable. It is important to note that the metropolitan interoffice network is not less dynamic than the access network, because the changes in the access network directly impact the interoffice network. Therefore, rapid response to changes in the access network is crucial for the telecommunications industry.

While long-haul carriers have generally understood the economic advantages of DWDM, local access companies did not immediately perceive an advantage of deploying this technology when compared to adding more fiber. Recently, however, some local access companies are starting to see metro optical networks based on DWDM technology as an attractive alternative to new fiber pulls. Meanwhile, vendors have been able to reduce costs to a point where WDM/DWDM links and associated equipment that route and switch wavelengths, such as optical cross-connects (OXCs), have become clearly cost-advantageous for short-haul (metropolitan area and local access) networks.

We study the economic advantages for the architectures *WDM point-to-point systems* and *WDM cross-connected mesh networks*.

Currently, the long-haul networks and some local exchange networks do not have enough capacity to support the traffic that has to be routed on them. In a conduit there can be several fiber cables (e.g., 4), each of which containing several single fibers (e.g., 96). The term *fiber exhaust* means that a single cable is full of fibers, but there might be space in the conduit to install more cables. The term *conduit exhaust* means that there is not room to install any more cables in the physical structure. It is costly to dig up the ground to build new conduit facilities. Under this scenario there are two ways to increase the capacity of an existing network: dig up the ground to install new fiber conduits and deploy traditional systems or increase the capacity of the remaining by installing WDM systems.

1.6.1 Point-to-Point WDM Networks

WDM point-to-point systems are being deployed by several carriers to increase the capacity of their existing fiber networks without installation of new fiber due to the increasing demands on communication bandwidth.

Figure 1.8 shows a links between two nodes that is running at exhaust capacity. Laying new fiber in the ground currently costs about 60,000 dollars per mile. Then the cost of laying new fiber to increase the capacity of the link $A \longrightarrow B$ would cost 6 million dollars.

Figure 1.8: Network Link

The capacity of the fiber link $A \longrightarrow B$ in figure 1.9 is now increased by a factor of 3, the number of wavelength channels used. The first systems were deployed with 8 wavelengths per fiber. However, the number of wavelengths carried by recently installed systems has increased to 128. The cost of a WDM terminal with 16 wavelengths can be less than 1 million dollar. The total cost (two WDM terminal are required in a point.to-point configuration) is less than 2 million dollars.

Even when there is still fiber on links, WDM can be an attractive solution, especially in long-haul carrier networks. There is no question about the usefulness of this technology if there is fiber exhaust, as it would cost millions to lay new fiber to meet constantly increasing demand. In addition, the length of a long-haul transmission link might be as

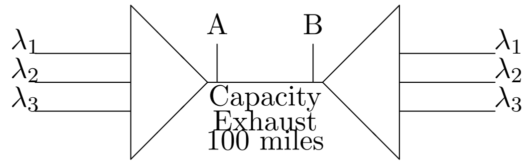


Figure 1.9: Point-to-Point WDM System

long as 600 km, with several electronic regenerators, which are required every 40 km, whereas the optical amplifiers (EDFAs) can be spaced as far as 120 km. Finally, only one optical amplifier is required to amplify multiple wavelength, whereas in the electronic case is required one regenerator for each wavelength.

1.6.2 WDM Cross-Connect Networks

Figure 1.10 shows the current mode of operation in a central office using WDM. A central office is a facility that contains switching, multiplexing, transmission, and end system equipment. The switching and multiplexing equipment showed in the figure is electronic based on DCSs (Digital Cross-Connects).

Figure 1.10 shows an electronic switching system installed in the network node. This situation is compared with the situation in figure 1.11 with an optical cross-connect approach.

In Figure 1.10 each transport system carries 16 wavelengths, each of which terminates on a SONET terminal (MUX in the figure). Each SONET terminal demultiplexes the 2.5-Gbps (OC-48) signals into their constituent DS3 tributaries (48 DS3s in an OC-48). In this mode of operation, the individual DS3s (running at 45 Mbps) are terminated onto a large digital cross-connect, which is responsible for grooming, provisioning, and protecting the individual DS3 channels. This makes sense if there is a need to access these individual DS3s; that is, if all of them terminate at the switching node at DS3 speeds, and none are terminated at higher speeds or passed through. However, in the current market, switches with direct OC-48 or higher speed interfaces will be available. It is

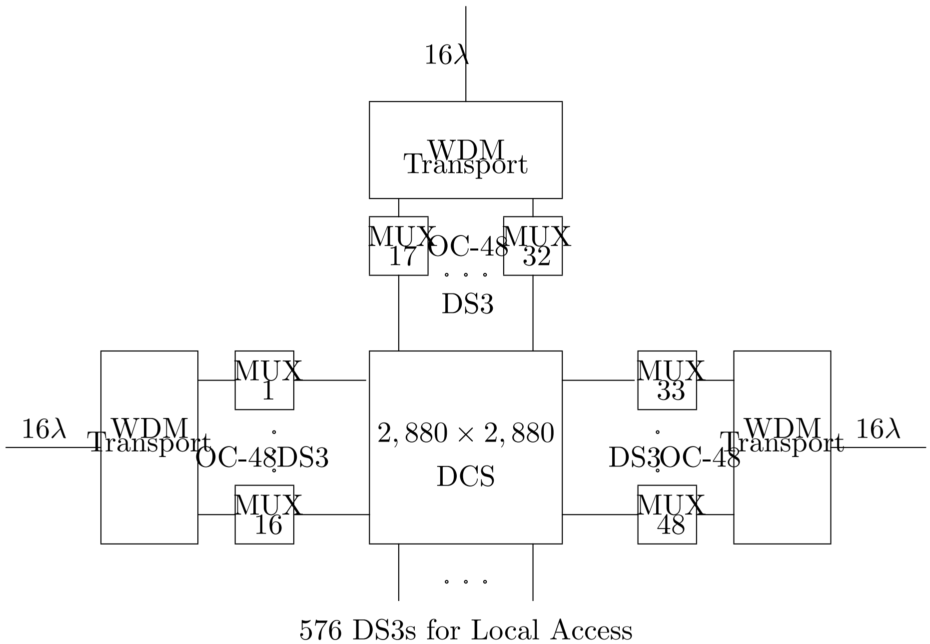


Figure 1.10: Network node with DCS

therefore unnecessary and costly to demultiplex these optical signals into their low-speed components.

To compare the optical and electronic switching approaches, let us assume that the network is dominated by data traffic that terminates on equipments with OC-48 interfaces, so that demultiplexing to DS3 speeds at the cross-connect is unnecessary. Let the cost of each DS3 termination on the digital cross-connect be $1,000$, and let the cost of the SONET terminal (unprotected; protection provided at the cross-connect) be 40,000 dollars. In the electronic switching approach of Figure 1.10, a $2,880 \times 2,880$ DCS is required to switch signals at the DS3 level, assuming that most of the traffic is pass-through, with 576 DS3 local access ports. Now compare this with the case in which

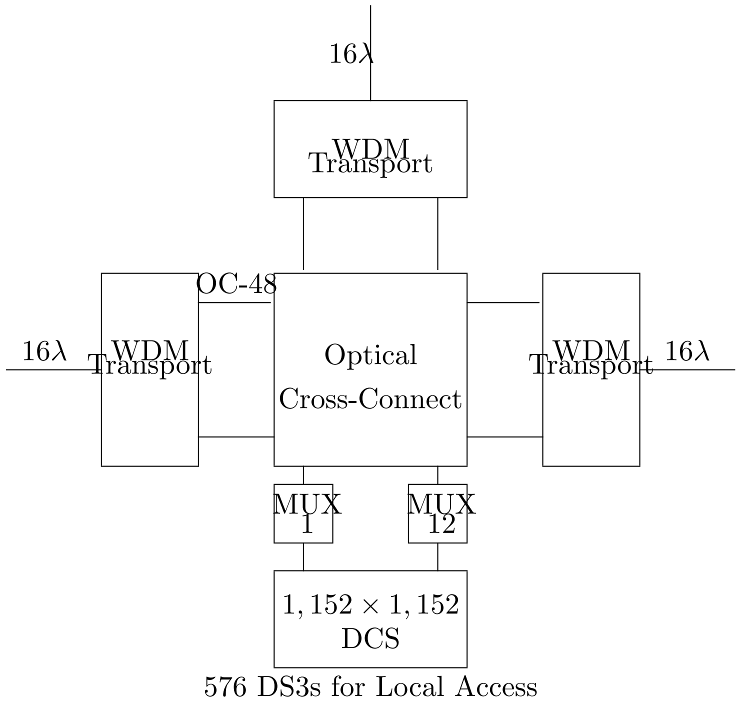


Figure 1.11: Network node with OXC

an optical cross-connect is installed at the node to switch the high-speed optical signals without demultiplexing them to DS3 level. The connections between the WDM terminal and the OXC are short lengths of fiber. As shown in figure 1.11, the local traffic is now dropped at the node in the form of 12 OC-48s, which are then demultiplexed into 576 DS3s. Replacing the DCS by the optical cross-connect reduces the number of SONET terminals from 48 to 12 and the size of the low-speed digital cross-connect from $2,880 \times 2,880$ to $1,152 \times 1,152$. (The smaller DCS is still needed for switching local traffic.) This can result in savings of more than \$3 million, depending on the cost of the cross-connect.

Chapter 2

Metaheuristics

2.1 Introduction

A *Metaheuristic* is a master strategy that guides and modifies other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality. Metaheuristics provide a means for approximately solving complex optimization problems, as those that arise in Communication Network Planning. These methods are designed to search for global optima. However, they cannot guarantee that the best solution found after termination criteria are satisfied is indeed a global optimal solution to the problem. Experimental testing of metaheuristic implementations show that the search strategies embedded in such procedures are capable of finding solutions of high quality to hard problems in industry, business and science.

2.1.1 Optimization

The *theory of optimization* refers to the quantitative study of optima and the methods for finding them. The technical verb optimize means to achieve the optimum and optimization is the act of optimizing. To achieve the optimum entails in some cases to obtain the most of some measure of success (e.g., revenue) or in some other cases to obtain the least of another measure (e.g., cost). Choosing a quantitative measure of effectiveness and then optimizing it has become the typical way in which many important decisions are made. Decisions involving how to design, build or operate a physical or economics system are reached in three steps:

1. Identify the decision variables in the system and determine, accurately and qualitatively, how they interact.

2. Identify a measure of system effectiveness that can be expressed in terms of the system variables. This measure is often referred to as the objective function.

3. Choose those values of the system variables that yield optimum effectiveness.

In classical optimization methods, such as linear programming, these three steps result in a model formulation of the type:

$$\text{Minimize } f(s), \text{ subject to } g(s) \geq b$$

In this formulation, $f(s)$ is the quantitative measure of quality (or objective function) and s are the decision variables. The set of constraints is also formulated in terms of the decision variables and represented as bounds on the function $g(s)$. In the case of linear programming both $f(s)$ and $g(s)$ are linear functions.

2.1.2 Linear Programming

Linear programming is considered a general-purpose tool because the only requirement is to represent the optimization model as a linear objective function subject to a set of linear constraints. The state-of-the-art linear programming solvers are quite powerful and can successfully solve models with thousands and even millions of variables employing reasonable amounts of computer effort. Evidently, however, not all business, industrial and scientific problems can be expressed by means of a linear objective and linear equalities or inequalities. Many complex systems may not even have a convenient mathematical representation, linear or nonlinear. Techniques such as linear programming and its cousins (nonlinear programming and integer programming) generally require a number of simplifying assumptions about the real system to be able to properly frame the problem.

Linear programming solvers are designed to exploit the structure of a well-defined and carefully studied problem. The disadvantage to the user is that in order to formulate the problem as linear program, simplifying assumptions and abstractions may be necessary. This leads to the well-known dilemma of choosing between finding the optimal solution to a model that does not represent the real system accurately and developing a model that is a good abstraction of the real system but for which only inferior sub-optimal solutions can be obtained. When dealing with the optimization of complex systems, a course of action taking for many years has been to develop specialized heuristic procedures that, in general, do not require a mathematical formulation of the problem. These procedures were appealing from the standpoint of simplicity, but generally lacked the power to provide high quality solutions to complex problems.

Example

Let us consider the routing and wavelength assignment (RWA) problem described in chapter 1. Many problems in WDM networks have RWA as a subproblem. The RWA problem consists of selecting routes and wavelengths to establish lightpaths. The static RWA problem consists of assigning routes and wavelengths (lightpaths) to source-destination pairs of a set of static connection requests (traffic demands).

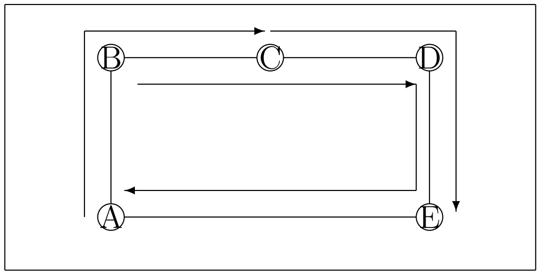
In a wavelength routed WDM optical network, a connection is realized by a lightpath. The requirement that the same wavelength must be used on all the links along the selected route is referred to as the *wavelength continuity constraint*. The *distinct wavelength assignment constraint* is that two lightpaths cannot be assigned the same wavelength on any fiber. Therefore, a lightpath between two nodes is a pair consisting of a route between these nodes and a wavelength.

Consider the problem of assigning routes and wavelengths (lightpaths) to source-destination pairs of a set of static connection requests (traffic demands). The objective is to assign lightpaths to all the demands in order to minimize the required number of wavelengths. This is the static lightpath assignment problem that have been shown to be NP-complete (see [24]).

2.1.3 Constructive Heuristics

Constructive heuristics are procedures that heuristically select the components of initially empty structure to get a solution of the problem. A simple constructive heuristic for this problem may be to select the highest demand and finding the shortest path from the origin to the destination and choose a wavelength of the set of used wavelengths that can be used (distinct wavelength assignment constraint). If there is not such a wavelength choose a new one, that is introduced in the set of used wavelengths. Then choose from a candidate list the highest demand and repeat the process. Once a demand is chosen, it is deleted from the candidate list of unsatisfied demands. While this heuristic may occasionally give acceptable results in some problems in general, its performance is predicted to be poor. This procedure falls within the class of heuristics called myopic, because they make decisions based on limited (also called "local") information without considering the consequences of implementing those decisions.

Consider a network consisting of 5 nodes A, B, C, D, and E in a cycle. Let the set of demands be: A-C, D-E, C-E, B-D. Using the above rule to assign lightpaths in this order we will need 3 wavelengths. The assignments would be: A-C: λ_1, D-E: λ_1, C-E: λ_2, B-D: λ_3. However, using the same rule with the permutation of demands A-C, C-E, D-E, B-D we will need only 2 wavelengths. The assignments now are: A-C: λ_1, C-E: λ_1, D-E: λ_2, B-D: λ_2.



2.1.4 Local Improving

In addition to heuristics designed to construct solutions, there are also procedures for improving solutions. The most common way of improving a solution is via the application of a local search (LS). To continue with our RWA illustration, let us consider the network with 5 nodes A, B, C, D, and E in a cycle. Suppose that a manager is asked to try to satisfy the demands in the following order:

1:A-C, 2:D-E, 3:C-E, 4:B-D.

Also suppose that the permutation of demands was constructed using the "highest demand" rule discussed above. A local search procedure would attempt to modify the current permutation by performing a move (or change). One possible move is to exchange the position of two demands in the order in which they are satisfied and measure the impact on the objective function.

If we limit our local search to moves that exchange demands that immediately follow each other in the current solution, then we only have to test 4 moves, i.e., (1,2), (2,3), (3,4), and (4,5). We would select the "best" of those moves. Note that move (2,3) will result in the solution with two wavelength.

However, if we want to test all possible exchanges of two demands as part of the local search effort, there are 10 moves to be examined. The amount of exploration, which is directly related to the amount of computational effort, is an important design issue in local search procedures. The effort to explore the neighborhood of a solution (that is the set of solutions reachable from the current solution by applying a move mechanism) can vary considerably. In a problem with n demands, there are $n-1$ neighbors if the move is defined as exchanging the positions of two demands that immediately follow each other in the current solution. However, if the move is defined as the exchange of positions of any two demands, the size of the neighborhood (i.e., the cardinality of the set of solutions reachable with such a move) is $(n^2 - n)/2$.

Regardless of the move mechanism, local search typically explores only a small fraction of the solution space. In the case of this problem, for example, the solution space consists of $n!$ solutions. Then, local search procedures that explore in the order of n^2 or even n^3 solutions are only dealing with a fairly small fraction of the entire solution space as the dimension of the problem increases.

Heuristics designed for constructing solutions are typically combined with improving local search procedures to create what is called a hill climbing method. These methods start from a solution and apply a local search in an attempt to find an improved solution. A local search is a procedure that analyze the neighborhood of the current solution to continue the search. If an improved solution is found, the search moves to it and the local search is applied again. The method stops when the local search is not capable of finding a solution that improves upon the current solution, i.e., when the "best" possible move cannot improve upon the objective function value of the current solution. The hill climbing terminology refers to the trajectory of the objective function values in a maximization problem.

The basic LS pseudocode can be stated as shown in figure 2.1.

The usual stopping criteria are designed to guess when the current solution is a local minimum; a solution that has not a better neighbor. Typical conditions are related to the number of iterations without improvements.

The usual greedy strategies try to improve as much as possible at every iteration. The steps of the greedy local search are shown in figure 2.2.

Initialization.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Find an initial solution s and compute $f(s)$.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) *Exploration of neighborhood*.

Find a neighbor s' of s ($s' \in \mathcal{N}(s)$).

(2) *Move or not*.

If $f(s') < f(s)$then $s \leftarrow s'$.

Figure 2.1: Basic Local Search.

The stoping criterion is that there is not a better solution in the neighborhood of the current solution; i.e., the current solution is known a local minimum.

The main shortcoming of an improving local search method is its inability to escape local optimality. The search strategies proposed by metaheuristic methodologies result in iterative procedures with the ability to escape local optimal points. Three are the main ways to do it: to generate a new starting solution to apply the improving local search, to consider a new set of moves or neighborhood definition and to allow non improving moves. These provide the three basic metaheuristics: the Multistart Search, the Variable Neighborhood search and the Global Search. In the global search are included those that allow back moves that are controlled by memory (like Tabu Search) or probabilistic (like Simulated Annealing). In addition, several metaheursitics use sets of solutions that interact and evolve in the solution space (like Genetic Algorithms, Scatter Search or Estimation Distribution Algorithms)

Initialization.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Find an initial solution s and compute $f(s')$.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) *Exploration of neighborhood*.

Find the best neighbor s' of s ($s' \in \mathcal{N}(s)$) and compute $f(s')$.

(2) *Move or not*.

If $f(s') < f(s)$ then set $s \leftarrow s'$. Otherwise stop.

Figure 2.2: Greedy Local Search.

.

2.2 Basic Metaheuristics

In this section we briefly describe the following well-known basic metaheuristics: multistart search, tabu search, simulated annealing, variable neighborhood search, scatter search and genetic algorithms. The descriptions focus on the main features of these methodologies.

2.2.1 Multistart Metaheuristic

A very simple metaheuristic is the Multistart Search (MS) [17], [51], [94]. A Multistart Search consists of applying a search to a series of initial solutions. They are designed from any search procedure by including it in a greater loop. Usually, the search procedure is an improving local search. The stopping condition of the local search is then taken as restarting criterium.

Initialization.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Take a set of starting solutions.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) *Start the LS*.

Take a new starting solution s and compute $f(s)$.

(2) *Performing the LS*.

Apply a the local search from s to get s' If s' is better than the current best solution s^* set $s^* \leftarrow s'$.

Figure 2.3: Basic Multistart

As indicated above, the main disadvantage of the improving local search is the possibility of being trapped in a non-optimal local minimum. The Multistart procedure is an easy way to scape a local minimum.

The basic MS pseudocode is shown in figure 2.3.

Recent analysis of objective function surfaces in some problems show that as its sizes grow large, random local minima are almost surely of *average* quality, implying that current random multistart heuristics, which rely on random starting solutions, are doomed to a *central limit catastrophe*. A key question for the performance of the Multistart Metaheuristics is to use the information about the topology of the neighborhood corresponding to the distance between solutions defined using the move used in the local search. A study in this matter can be found in [95]. The Greedy Multistart heuristic can be described as shown in figure 2.4.

Initialization.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Take a set of starting solutions.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) *Start the greedy LS*.

Take a new starting solution s and compute $f(s)$.

(2) *Performing the greedy LS*.

Repeat the following steps until no improvement is met

(a) Find the best neighbor s' of s and compute $f(s')$.

(b) If $f(s') < f(s)$ then set $s \leftarrow s'$ and go to (a).

(3) *Improve or not*.

If s' is better than the current best solution s^* set $s^* \leftarrow s'$.

Figure 2.4: Greedy Multistart Search.

2.2.2 Tabu Search

Basic Tabu Search (TS) [42], [43], [44], [45], [47], [48], [49], [50], [52] maintains a selective history H of the states encountered during the search, and replaces the neighborhood of the current solution $N(s)$ by a modified neighborhood, which may be denoted $N(H, s)$. History therefore determines which solutions may be reached by a move from the current solution, selecting s' from $N(H, s)$.

In the TS strategies based on short term considerations, $N(H, s)$ characteristically is a subset of $N(s)$, and the tabu classification serves to identify elements of $N(s)$ excluded from $N(H, s)$. In the intermediate and longer term strategies, $N(H, s)$ may contain solutions not in $N(s)$, generally consisting of selected elite solutions (high quality local optima) encountered at various points in the solution process. Such elite solutions typically are identified as elements of a regional cluster in intermediate term intensification strategies, and as elements of different clusters in longer term diversification strategies.

TS also uses history to create a modified evaluation of currently accessible solutions. This may be expressed formally by saying that TS replaces the objective function $f(s)$ by a function $f(H, s)$, which has the purpose of evaluating the relative quality of currently accessible solutions. It is provided by the use of frequency based memory. The relevance of this modified function occurs because TS uses aggressive choice criteria that seek a best s'; i.e., one that yields a best value of $f(H, s)$, over a candidate set drawn from $N(H, s)$. Moreover, modified evaluations often are accompanied by systematic alteration of $N(H, s)$, to include neighboring solutions that do not satisfy customary feasibility conditions.

For large problems, where $N(H, s)$ may have many elements, or for problems where these elements may be costly to examine, the aggressive choice orientation of TS makes it highly important to isolate a candidate subset of the neighborhood, and to examine this subset instead of the entire neighborhood. Because of the significance of the candidate subset, we refer to it explicitly by the notation $N'(H, s)$.

Initialization.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Find an initial solution s.

Start with the history record H empty.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) *Exploration of neighborhood*.

Select a non tabu neighbor s' of s ($s' \in \mathcal{N}'(H, s)$) to minimize $f(H, x)$.

(2) *Update*.

Update the history record H.

Figure 2.5: Basic Tabu Search.

Therefore, the Tabu Search procedure, instead simply selecting the best neighbor s' of s with respect to f as the greedy local search do, selects the solution in $N'(H, s)$ that minimizes $f(H, s)$. The selected solution s' is called a highest evaluation candidate.

The TS pseudocode is shown in figure 2.5.

Formally the tabu search method is quite straightforward to state. The essence of the method depends on how the history record H is defined and utilized, and on how the candidate neighborhood $N'(H, s)$ and the evaluation function $f(H, s)$ are determined.

2.2.3 Scatter Search

Scatter Search (SS) [35], [77], [79], [80], [81], [82], [83], [84], [85], [86], [87] is a population-based metaheuristic that uses a reference set to combine its solutions and construct others. The method generates a reference set from a population of solutions. Then a subset is selected from this reference set. The selected solutions are combined to get starting solutions to run an improvement procedure. The result of the improvement can motivate the updating of the reference set and even the updating of the population of solutions.

The initial population must be a wide set of disperse solutions. However, it must also include good solutions. Several strategies can be applied to get a population with these properties. The solutions to be included in the population can be created, for instance, by using a random procedure to achieve a certain level of diversity. Then a simple improvement heuristic procedure must be applied to these solutions in order to get better solutions. The initial population can also be obtained by a procedure that provides at the same time disperse and good solutions like GRASP procedures [36].

A set of good representative solutions of the population is chosen to generate the reference set. The good solutions are not limited to those with the best objective values. By good representative solutions we mean solutions with the best objective values as well as disperse solutions. Disperse solutions should reach different local minima by the local search. Indeed, a solution may be added to the reference set if the diversity of the set improves. The criteria for updating the reference set, when necessary, must be based on comparisons and measures of diversity between the new solutions and the existing solutions.

A subset of solutions from the reference set is selected as input data for performing a combination method to get good starting solutions for an improvement procedure. In general, the method consists of selecting all the subsets of a fixed size. The combination procedure tries to combine good characteristics of the selected solutions to get new current solutions. The aim is to get better solutions, which are not similar to those already in the reference set.

The possible improvement solution methods applied to the solutions range from simplest local searches to a very specialized search. A very simple procedure is a local search based on basic moves consisting of selecting the best improving move or a first found

Repeat the following sequence until the stopping condition is met:

Generate a population of solutions P.

Repeat the following sequence until a new population needs to be obtained:

- Generate a reference set R from the population.
- Repeat the following sequence until a new reference set needs to be obtained:

  1 Select a subset of solutions S from the reference set.

  2 Apply the combination procedure to the subset S to get s.

  3 Apply the improvement method to s to get s'.

  4 Update the reference set according to the results of the improvements.

Figure 2.6: The Basic Scatter Search

improving move. The procedure must allow to use tools like recent or intermediate memory, variable neighborhoods, or hashing scanning methods of the neighborhood. Then the method applied could be a Tabu Search [43], [44], a Variable Neighborhood Search [61], [62] or any sophisticated hybrid heuristic search.

The metaheuristic strategy includes the decision on how to update the reference set taking into account the state of the search. The algorithm must also realize when the reference set does not change and seek to diversify the search by generating a new set of solutions for the population.

Usual stopping conditions are based on allowing a total maximum computational time or a maximum computational effort since the last improvement. The computational effort is measured by the number of iterations, number of local searches or real time.

A pseudocode of the Scatter Search is shown in figure 2.6.

The Scatter Search strategy involves six procedures and three stopping criteria to solve an optimization problem. The procedures are the following:

1. **The Initial Population Creation Method.** This procedure creates a random initial population P of good and disperse solutions.

2. **The Reference Set Generation Method.** This procedure selects some of the best representative solutions in the population to be included in the reference set R.

3. **The Subset Generation Method.** This procedure generates subsets, which consist of good solutions in the reference set, to apply the combination procedure.

4. **The Solution Combination Method.** This procedure, which includes parameters used to modulate the intensification and/or diversification, combines the solutions in the previously selected subset to get the new current solution s.

5. **The Improvement Solution Method.** This procedure, which includes parameters to modulate the specialization of the method, improves the current solution s to get an improved solution s'.

6. **The Reference Set Updating Method.** This procedure updates the reference set by deciding when and how the obtained improved solutions are included in the reference set replacing some solutions already in it.

In addition to these six procedures, the metaheuristic involves three stopping procedures that implement the criteria to decide when generating a new reference set, a new population or when stopping the search.

1. **New Reference Set Criterion.** The first criterion decides when to generate a new reference set from the population.

2. **New Population Criterion.** The second criterion decides when to generate a new population.

3. **Termination Criterion.** Finally, the third criterion decides when to stop the whole search.

2.2.4 Simulated Annealing

Simulated Annealing (SA) [75], [133], [134] is an important metaheuristic technique for solving optimization problems. The fundamentals of SA were introduced by Kirkpatrick et al. (1983) and by Cerny (1985) following an analogy with the physical annealing process used to find low-energy states of solids.

Simulated Annealing applies a global search where probabilistic criteria for solution acceptance are used. The probabilistic criteria consist of accepting any movement that improves the current solution and only with probability p $(0 < p < 1)$ any other movement. Normally p depends on the modification of the objective function and is modified dynamically. The search starts with an initial solution and an initial value of a control parameter, called *temperature* and denoted by c. For each value of c a loop is performed l times, that can be fixed or dynamically updated. For every fixed temperature c, the search is a Markov chain and the parameter l is referred to as the *length* of the chain. The loop consists of the following steps. Generate at random a solution s' from the neighborhood of the current solution s, and evaluate the new solution. Let $\Delta = f(s') - f(s)$. If $\Delta \leq 0$ then do $s \leftarrow s'$, otherwise do it with probability $p = e^{-\Delta/T}$. This is done by generating a random number $x \in (0, 1)$ and accepting s' (i.e., $s' \leftarrow s$) if $x < p$ and rejecting it otherwise. After each execution of the loop the temperature is updated by decreasing it in a multiplicative factor α (i.e., $\alpha \leftarrow \alpha T$).

The basic SA pseudocode is shown in figure 2.7.

The SA belongs to a wider class of threshold accepting local search algorithms [33] that continually select a neighbor of a current solution and compare the difference between the objective value of the neighbor and the objective value of the current solution to a threshold. If the difference is within the threshold, the neighbor replaces the current

---

Initialization.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Find an initial solution s and an initial temperature T.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) Perform the following loop L times:

(a) *Exploration of neighborhood*. Find and evaluate a neighbor s' of s ($s' \in \mathcal{N}(s)$) and compute $\Delta = f(s') - f(s)$.

(b) *Move or not*.

If the solution thus obtained s' is better than s ($\Delta < 0$), set $s \leftarrow s'$. Otherwise, generate at random $x \in (0, 1)$ and if $x < e^{-\Delta/t}$ set $s \leftarrow s'$.

(2) Set $T \leftarrow \alpha T$.

Figure 2.7: Basic Simulated Annealing.

<u>*Initialization*</u>.

Select a neighborhood structure $\mathcal{N}$ that will be used in the search.

Set $k \to 0$.

<u>*Iterations*</u>.

Repeat the following sequence until the stopping condition is met:

(1) Set $k \leftarrow k + 1$.

(2) Perform the following loop l times:

(a) <u>*Exploration of neighborhood*</u>. Find a neighbor s' of s ($s' \in \mathcal{N}(s)$) and compute $\Delta = f(s') - f(s)$.

(b) <u>*Move or not*</u>. Get the new threshold value t_k. If $\Delta < t_k$ set $s \leftarrow s'$.

Figure 2.8: Basic Threshold accepting.

solution. Otherwise, the search continues with the current solution. The pseudocode for these threshold-accepting algorithm is presented in figure 2.8.

The number l of loops that use the same threshold can be only one. The sequence of thresholds $t_k, k = 1, 2, ...$, used at iteration k of the algorithm are modified along the execution. The basic SA takes t_k from a random variable with exponential distribution and expected value c_k. This is a control parameter called temperature that is decreased according to a cooling rule. Other threshold algorithms use deterministic rules to modify the threshold t_k.

The *cooling schedule* specifies the parameters that govern the convergence of the Simulated Annealing. Namely, the cooling schedule gives:

- The initial value of the control parameter.

- The decrement function for lowering the control parameter.

- The final value of the control parameter.
- The length of the loop for every value of the control parameter.

Typical cooling schedules start at a sufficient large value of c_k such that all movements are virtually accepted. There must be a trade-off between the decrement of the control parameter and the length of the loops. One can use small decrements of c_k with large lengths l for the loops or large decrements with small lengths. Usually the decrement of the control parameter is made by a geometric rule like $c_{k+1} = \alpha c_k$ where α is a positive constant smaller but near to 1 (typically a value between 0.8 and 0.99). Typical values for c_0 are related to the maximal difference between the objective function for two solutions. The final values for c_k are related to the smallest possible difference between the objective function for two solutions. The values for the lengths are related to the size of the neighborhoods [1].

2.2.5 Variable Neighborhood Search

Variable Neighborhood Search (VNS) [62], [63], [64] is a recent metaheuristic for solving combinatorial and global optimization problems based upon a simple principle: systematic change of neighborhood within the search. Its development has been rapid, with a lot of papers already published and its applications have been numerous and successful. Many extensions have been made, mainly to be able to solve large problem instances. In most of them, an effort has been made to keep the simplicity of the basic scheme.

Let $\mathcal{N}_k$, $(k = 1, \ldots, k_{max})$ be a finite set of neighborhood structures, and $\mathcal{N}_k(s)$ be the set of solutions in the k^{th} neighborhood of a solution s. Neighborhoods $\mathcal{N}_k$ may be induced from metric functions introduced into a solution space S. If $d(.,.)$ is this distance then take increasing values d_k, $k = 1, ..., k_{max}$ and set $N_k(s) = \{s' \in S : d(s, s') \leq d_k\}$. Most local search heuristics use only one neighborhood structure $\mathcal{N}$. Therefore a series of nested neighborhoods are obtained from a single neighborhood by taking $\mathcal{N}_1(s) = \mathcal{N}(s)$

and $\mathcal{N}_{k+1}(s) = \mathcal{N}(\mathcal{N}_k(s))$, for every solution s. This means that a move to the k-th neighborhood is performed by repeating k times a move to the original neighborhood. A solution $s' \in S$ is a *local minimum* with respect to $\mathcal{N}_k$ if there is no solution $s \in \mathcal{N}_k(s') \subseteq S$ better than s' (i.e., such that $f(s) < f(s')$ where f is the objective function of the problem).

The Variable Neighborhood Search metaheuristic is based on three simple facts:

1 A local minimum with respect to one neighborhood structure is not necessary a local minimum with another.

2 A global minimum is a local minimum with respect to all possible neighborhood structures.

3 Local minima with respect to one or several neighborhood structures are usually similar to each other.

The last observation is empirical. By similar local minima we mean, for instance, that they have several variables with the same value in common. This implies that a local optimum often provides some information about the global one. However, it is usually not known which ones are such. However, a search in these "neighborhoods" of a local minimum will meet other local minima, and the global minimum among them.

By applying the VNS principle to an improving local search we get the **Variable Neighborhood Descent** (VND). The method consists of changing the neighborhoods systematically within a local search. The basic VND is presented in figure 2.9.

The final solution should be a local minimum with respect to all k_{max} neighborhoods, and thus the probability of reaching the global minimum is higher than by using a single structure. Beside this *sequential* order of neighborhood structures in VND above, one can develop a *random* strategy by choosing the successive values for k at random.

Initialization.

Select the set of neighborhood structures $\mathcal{N}_k$, for $k = 1, \ldots, k_{max}$, that will be used in the descent.

Find an initial solution s.

Iterations.

Repeat the following sequence until no improvement is obtained:

(1) Set $k \leftarrow 1$;

(2) Repeat the following steps until $k = k_{max}$:

(a) *Exploration of neighborhood*.

Find the best neighbor s' of s ($s' \in \mathcal{N}_k(s)$).

(b) *Move or not*.

If the solution thus obtained s' is better than s, set $s \leftarrow s'$ and $k \leftarrow 1$; otherwise, set $k \leftarrow k + 1$.

Figure 2.9: Variable Neighborhood Descent.

Basic Variable Neighborhood Search

Most local search heuristics use in their descents a single or sometimes two neighborhoods ($k_{max} \leq 2$). A usual strategy with two neighborhoods consists of performing local searches for the first neighborhood from points s' that belong to the second neighborhood of the current solution (i.e. $s' \in \mathcal{N}_2(s)$). The perturbation strategy consists of applying a local search using the first neighborhood and then perturbing the current solution by choosing a random solution in the second neighborhood to perform a new local search. The **Basic Variable Neighborhood Search** (BVNS) method uses deterministic changes in the neighborhood structure for perturbation or shaking. Its steps are given in Figure 2.10.

The stopping condition may be the maximum CPU time allowed, the maximum number of iterations, or the maximum number of iterations between two improvements. The Reduced Variable Neighborhood Search (RVNS) method is obtained if random points are selected from $\mathcal{N}_k(s)$, without being followed by descent. It is useful for very large instances for which local search is costly.

The local search step (2b) in the basic VNS may be replaced by VND. This gives the **General Variable Neighborhood Search** (GVNS) that is the version with the most recent success. Its steps are shown in figure 2.11.

Several extensions of the VNS have also been proposed. The basic VNS is a first improvement descent method with randomization. It is transformed into a descent-ascent method if, Step 2c sets also $s \leftarrow s$" with some probability even if the solution is worse than the incumbent (or the best solution found so far). It is changed into a best improvement method by making a move to the best neighborhood k^* among all k_{max} of them. Other variants of the basic VNS are to find a solution s' in Step 2a as the best among b (a parameter) randomly generated solutions from the k^{th} neighborhood, or to introduce k_{min} and k_{step}, two parameters that control the change of the neighborhood process. In the previous algorithm, instead of setting $k \leftarrow 1$ set $k \leftarrow k_{min}$ and instead of

<u>*Initialization*</u>.

Select the set of neighborhood structures $\mathcal{N}_k$, for $k = 1, \ldots, k_{max}$.

Find an initial solution s.

Choose a stopping condition.

<u>*Iterations*</u>.

Repeat the following sequence until the stopping condition is met:

(1) Set $k \leftarrow 1$.

(2) Repeat the following steps until $k = k_{max}$:

(a) <u>*Shaking*</u>.

Generate a point s' at random from the k^{th} neighborhood of s ($s' \in \mathcal{N}_k(s)$).

(b) <u>*Local search*</u>.

Apply some local search method with s' as initial solution; denote with s'' the so obtained local optimum.

(c) <u>*Move or not*</u>.

If this local optimum is better than the incumbent, move there ($s \leftarrow s''$), and continue the search with $\mathcal{N}_1$ ($k \leftarrow 1$); otherwise, set $k \leftarrow k + 1$.

Figure 2.10: Basic Variable Neighborhood Search.

Initialization.

Select the set of neighborhood structures $\mathcal{N}_k$, for $k = 1, \ldots, k_{max}$ for the shaking.

Select the set of neighborhood structures $\mathcal{N}'_j$, for $j = 1, \ldots, j_{max}$ for the descent.

Find an initial solution s.

Choose a stopping condition.

Iterations.

Repeat the following sequence until the stopping condition is met:

(1) Set $k \leftarrow 1$.

(2) Repeat the following steps until $k = k_{max}$:

   (a) *Shaking*.

   Generate a point s' at random from the k^{th} shaking neighborhood of s ($s' \in \mathcal{N}_k(s)$).

   (b) *Descent*.

   Apply to s' the VND with $\mathcal{N}'_j, j = 1, \ldots, j_{max}$ as neighborhoods to get a new solution s''.

   (c) *Move or not*.

   If $f(s'') < f(s)$ set $s \leftarrow s''$ and $k \leftarrow 1$; otherwise, set $k \leftarrow k + 1$.

Figure 2.11: General Variable Neighborhood Search.

setting $k \leftarrow k+1$ set $k \leftarrow k+k_{step}$.

The **Variable Neighborhood Decomposition Search** (VNDS) method extends the basic VNS into a two-level VNS scheme based upon decomposition of the problem. The only difference between the general or basic VNS and VNDS is in step 2b. Instead of applying the VND or other local search descent method in the whole solution space $\mathcal{S}$ (starting from $s' \in \mathcal{N}_k(s)$), the VNDS solves at each iteration a subproblem in some subspace $V_k \subseteq \mathcal{N}_k(s)$ with $s' \in V_k$. When the procedure used in this step is also VNS, the two-level VNS-scheme arises. VNDS can be viewed as embedding the classical successive approximation scheme

2.2.6 Genetic Algorithms

A genetic algorithm (GA) GA [21], [31], [55], [65], [99], [103] is an iterative procedure for solving an optimization problem that uses an evolving constant-size population of individuals. Each individual of the population is represented by a finite string of symbols, known as the genome, and encodes a possible solution of search space. Solutions to a problem were originally encoded as binary strings due to certain computational advantages associated with such encoding. Also the theory about the behavior of algorithms was based on binary strings. However, the solution representation has been extended in recent years to include character-based encoding, real-valued encoding, and tree representations.

The standard genetic algorithm proceeds as follows. An initial population of individuals is generated at random or heuristically. Every evolutionary step, known as a generation, the individuals in the current population are decoded and evaluated according to some predefined quality criterion, referred to as the fitness, or fitness function. To form a new population (the next generation), individuals are selected according to their fitness. Many selection procedures are currently in use, one of the simplest being fitness-proportionate selection, where individuals are selected with a probability proportional to their relative fitness. This ensures that the expected number of times an individual is chosen is approximately proportional to its relative performance in the population. Thus, high-fitness ("good") individuals stand a better chance of "reproducing", while low-fitness ones are more likely to disappear.

Genetically inspired operators are used to introduce new individuals into the population, i.e., to generate new points in the search space. The best known of such operators are crossover and mutation. Crossover is performed, with a given probability p_c (the "crossover probability" or "crossover rate"), between two selected individuals, called parents, by exchanging parts of their genomes (i.e., encoding) to form two new individuals, called offspring; in its simplest form, sub-strings are exchanged after a randomly selected crossover point. This operator tends to enable the evolutionary process to move toward "promising" regions of the search space. The mutation operator is introduced to prevent premature convergence to local optima by randomly sampling new points in the search space. Mutation entails flipping bits at random, with some (small) probability p_m. Genetic algorithms are stochastic iterative processes that are not guaranteed to converge; the termination condition may be specified as some fixed, maximal number of generations or as the attainment of an acceptable fitness level for the best individual.

Let us consider the following example to illustrate the genetic algorithm. The population consists of 4 individuals, which are binary-encoded strings (genomes) of length 8. The fitness value equals the number of ones in the bit string, with $p_c = 0.7$, and $p_m = 0.001$. More typical values of the population size and the genome length are in the range 50-1000. Also note that fitness computation in this case is extremely simple since no complex decoding nor evaluation is necessary. The initial (randomly generated) population might look like this:

Label	Genome	Fitness
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3

Using fitness-proportionate selection we must choose 4 individuals (two sets of parents), with probabilities proportional to their relative fitness values. In our example, suppose that the two parent pairs are B, D and B, C (note that A did not get selected as our procedure is probabilistic). Once a pair of parents is selected, the crossover operation is performed with probability p_c, resulting in two offspring. If the crossover operation is not performed (with probability $1 - p_c$), then the offspring are exact copies of each parent. Suppose, in our example, that crossover takes place between parents B and D at the (randomly chosen) first bit position, forming offspring $E = 10110100$ and $F = 01101110$, while the crossover operation is not performed between parents B and C, forming offspring that are exact copies of B and C. Next, each offspring is subject to mutation with probability pm per bit. For example, suppose offspring E is mutated at the sixth position to form $E' = 10110000$, offspring B is mutated at the first bit position to form $B' = 01101110$, and offspring F and C are not mutated at all. The next generation population, created by the above operators of selection, crossover, and mutation is therefore:

Label	Genome	Fitness
E'	10110000	3
F	01101110	5
C	00100000	1
B'	01101110	5

Note that in the new population, although the best individual with fitness 6 has been lost, the average fitness has increased. Iterating this procedure, the genetic algorithm will eventually find a perfect string, i.e., with maximal fitness value of 8. More sophisticated implementations of Genetic Algorithms include the use of local search and several crossover operators that are chosen probabilistically to be applied to each pair of selected parents.

2.2.7 Other Metaheuristics

In the literature, there is a high number of other Metaheuristics. Some of the most recent books on Metaheuristics are the following: [11], [25], [52], [78], [92], [100], [26], [118], [120], [135]. Among these other metaheuristics we can mention the following. Memetic Algorithms, that combine the features of genetic algorithms with local searches [106]. The *Estimation of Distribution Algorithm* (EDA) is a evolutionary algorithm in which populations of individuals are created by estimation and simulation of the joint probability distribution of the selected individuals [88]. *Greedy randomized adaptive search heuristic* (GRASP) is an iterative process in which each iteration consists of two phases, a construction phase in which a feasible solution is produced, and a local search phase in which a local optimum in the neighborhood of the constructed solution is sought [36], [119]. The *Reactive Search* that also uses special types of memory [7], [8], [9]. The *Path Relinking*, that is an intensification strategy for exploring trajectories connecting high quality solutions mainly used in TS [50] and SS [53]. The neural network (NN) is another metaheuristic that uses an associative form of memory [127], [139]. *Ant System* is a metaheuristic that takes inspiration from the behavior of real ant colonies [32]. *Iterated Local Search* (ILS) is based on building a sequence of locally optimal solutions by perturbing the current solution applying a local search to that modified solution [91]. *Extreme Optimization* (EO) is another evolutionary local-search algorithm [18], [19], [20]. *Concentration Heuristic* that uses the information from a set of local minima to concentrate the search in a region of the search space [122], [123], [124].

The *Guided Local Search* (GLS) metaheuristic explores efficiently and effectively the search space by exploiting prior information known about the problem in conjunction with historical information gathered during the search process [101], [102], [136], [140]. The *Fuzzy Adaptive Nieghborhood Search* (FANS) metaheuristics use a new mechanism to scape a local minimum where solutions are evaluated in terms of fuzzy valuation [16]. The *Particle Swarm Optimization* (PSO) is an evolutionary computation technique inspired by social behavior of bird flocking or fish schooling [34], [70]. The *Very Large-Scale Neighborhood Search* (VSLN) metaheuristic uses special tools for searching very vast space solutions [111], [112].

2.2.8 Parallelization of Metaheuristics

A key question in the applications of the Metaheuristics for real problems is to exploit the possible parallelization of the procedures. The Metaheuristics based of local searches like Tabu Search, VNS, Simulated Annealing, and Scatter Search can be parallelized following three main strategies. A first parallelization is the low-level parallelization consisting of replacing the local search by the parallel version of the local search. This is done by dividing the neighborhood of the current solutions in subsets that are assigned to the processors and each returns an improving neighbor in its part of neighborhood. A second parallelization consists of replicating the metaheuristic in each processors. It corresponds to a natural parallelization of the hybrid between the metaheuristic and a Multistart search. The third strategy for the parallelization of the metaheuristcs is a specific parallelization for the metaheuristic that consists of parallelizing a characteristic tool of the search process. We have analyzed these three parallelization strategies in [39] and [40].

In [40] the three strategies for parallelizing the Scatter Search are compared. The third considered strategy of parallelization for the Scatter Search consists of selecting several subsets from the reference set that are then combined and improved by the available processors. These procedures are replicated as many times as the number of available processors. The local optima found by the processors are used to update the reference set. This procedure, called *Replicated Combination Scatter Search*, shows better results than parallelizing the local search and than replicating the scatter search, where each processor uses its own population and reference sets.

In [39], where several parallelizations for the VNS are developed, two of them got better performance than the other strategies. The first one consists of increasing the number of solutions drawn from the current neighborhood in the shake stage and doing local search in parallel from each of them. The second one do the same as the first, but updating the information about the best solution found. The second of these parallelizations gave the best results and its steps are presented in figure 2.12.

2.3 The Role of Metaheuristics

Metaheuristics provided a way of considerably improving the performance of simple heuristic procedures. The search strategies proposed by metaheuristic methodologies result in iterative procedures with the ability to escape local optimal points. Metaheuristics have been developed to solve complex optimization problems in many areas, with combinatorial optimization being one of the most fruitful. Generally, the best procedures achieve their efficiencies by relying on context information. The solution methods can be viewed as the result of adapting several metaheuristic strategies to specific optimization problems.

The term metaheuristic was coined by Fred Glover in 1986 and has come to be widely applied in the literature, both in the titles of comparative studies and in the titles of volumes of collected research papers. A metaheuristic refers to a master strategy that guides and modifies other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality. The heuristics guided by such a meta-strategy may be high level procedures or may embody nothing more than a description of available

---

<u>*Initialization*</u>.

Select the set of neighborhood structures $\mathcal{N}_k$, for $k = 1, \ldots, k_{max}$.

Find an initial solution s.

Choose a stopping condition.

<u>*Iterations*</u>.

Repeat the following sequence until the stopping condition is met:

(1) Set $k \leftarrow 1$.

(2) Repeat, for each processor $i = 1, ..., p$ in parallel, the following steps until $k = k_{max}$:

   (a) <u>*Shaking*</u>.

   Generate a point s_i at random from the k^{th} neighborhood of s ($s_i \subseteq \mathcal{N}_k(s)$).

   (b) <u>*Local search*</u>.

   Apply the local search method with s_i as initial solutions to get the local minimum s'_i.

   (c) <u>*Move or not*</u>.

   Denote with s' the best among the local optima s'_i obtained by the processors. If s' is better than s set $s \leftarrow s'$ and set $k \leftarrow 1$; otherwise, set $k \leftarrow k + 1$.

Figure 2.12: Parallel Variable Neighborhood Search.

moves for transforming one solution into another, together with an associated evaluation rule.

The contrast between the metaheuristic orientation and the "local optimality" orientation is significant. For many years, the primary conception of a heuristic procedure (a conception still prevalent today) was to envision either a clever rule of thumb or an iterative rule that terminates as soon as no solutions immediately accessible could improve the last one found. Such iterative heuristics are often referred to as descent methods, ascent methods, or local search methods. (A sign of the times is that "local search" now sometimes refers to search that is not limited to being local in character.) Consequently, the emergence of methods that departed from this classical design - and that did so by means of an organized master design - constituted an important advance. Widespread awareness of this advance only began to dawn during the last decade, though its seeds go back much farther.

2.3.1 Metaheuristic Features

The evolution of metaheuristics during the past ten years has taken an explosive upturn. Metaheuristics in their modern forms are based on a variety of interpretations of what constitutes "intelligent" search. These interpretations lead to design choices that in turn can be used for classification purposes. However, a rigorous classification of different metaheuristics is a difficult and risky enterprise, because the leading advocates of alternative methods often differ among themselves about the essential nature of the methods they espouse. This is later illustrated by considering the classification of metaheuristics in terms of their features with respect to three basic design choices:

(1) the use of adaptive memory,

(2) the kind of neighborhood exploration used, and

(3) the number of current solutions carried from one iteration to the next.

In addition to the three basic design elements used in the classification, metaheuristics incorporate other strategies with the goal of guiding the search. A metaheuristic may strategically modify the evaluation provided by a component heuristic (which normally consists of identifying the change in an objective function value produced by a move). For example, simulated annealing relies on a problem objective function to provide each evaluation, but then amends this evaluation based on the current solution. In the amended form, all improving moves are considered equally attractive, and any such move encountered is accepted. Moves that deteriorate the value of the objective function are accepted or rejected by a probabilistic criterion that initially assigns a high probability (when the temperature is high) to accepting any move generated, regardless of its quality. However, a bias is incorporated that favors smaller deteriorating moves over larger ones, and over time this bias is increased, ultimately reducing the probability of accepting a non-improving move to zero. The set of available moves can be taken from another heuristic, but classical Simulated Annealing preempts all other move generation processes to generate moves randomly from the proposed domain.

A metaheuristic may also modify the neighborhood of moves considered to be available, by excluding some members and introducing others. This amended neighborhood definition may itself necessitate a change in the nature of evaluation. The strategic oscillation approach of tabu search illustrates this intimate relationship between changes in neighborhood and changes in evaluation.

A standard neighborhood that allows moves only among feasible solutions is enlarged by this approach to encompass infeasible solutions. The search is then strategically driven to cross the feasibility boundary to proceed into the infeasible region. After a selected depth is reached, the search changes direction to drive back toward feasibility, and upon crossing the feasibility boundary similarly continues in the direction of increased feasibility. (One-sided oscillations are employed in some variants to remain predominantly on a particular side of the boundary.) To guide these trajectories, the approach modifies customary evaluations to take account of the induced direction of movement and the region in which the movement occurs. The result generates a controlled behavior that exploits the theme of non-monotonic exploration.

The emphasis on guidance differentiates a metaheuristic from a simple random restart procedure or a random perturbation procedure. However, sometimes these naive restarting and perturbation procedures are also classed as low-level metaheuristics, since they allow an opportunity to find solutions that are better than a first local optimum encountered. "Noising" procedures, which introduce controlled randomized changes in parameters such as cost or resource availability coefficients, provide one of the popular mechanisms for implementing such approaches. Another popular mechanism is simply to randomly modify evaluations, or to choose randomly from evaluations that fall within a chosen window. Such randomized processes are also applied to selecting different types of moves (neighborhood definitions) at different junctures.

In contrast to an orientation that still often appears in the literature, the original conception of a metaheuristic does not exclude consideration of constructive moves for generating initial solutions, but likewise allows these moves to be subjected to metaheuristic guidance. (A popular orientation in the literature is to suppose that metaheuristics are only used in connection with "transition" moves, which operate on fully constructed solutions.) ¿From a broader perspective, a partial solution created by a constructive process is simply viewed as a solution of a particular type, and procedures for generating such solutions are natural candidates to be submitted to higher-level guidance. This view has significant consequences for the range of strategies available to a metaheuristic approach.

Strategic oscillation again provides an illustration. By the logical restructuring theme of tabu search, constructive moves are complemented by creating associated destructive moves, allowing the oscillation to proceed constructively to (and beyond) a stipulated boundary, and then to reverse direction to proceed destructively to various depths, in alternating waves. Transition moves permit refinements at varying levels of construction and destruction.

The perspective that restricts attention only to transition moves is gradually eroding, as researchers are coming to recognize that such a restriction can inhibit the development of effective methods. However, there remain pockets where this recognition is slow to dawn. (For example, methods that alternate between construction and transition moves - affording a simple subset of options provided by strategic oscillation - have recently been characterized in a segment of the literature as a "new development.")

The use of population-based strategies and adaptive memory strategies, are often taken to be fundamental distinctions in the literature. Population-based strategies manipulate a collection of solutions rather than a single solution at each stage. Such procedures are now often referred to as composing the class of evolutionary methods. A prominent subclass of these methods is based on strategies for "combining" solutions, as illustrated by genetic algorithms, scatter search and path relinking methods. Another prominent subclass consists of methods that are primarily driven by utilizing multiple heuristics to generate new population members. This incorporation of multiple heuristics for generating trial solutions, as opposed to relying on a single rule or decision criterion, is a very old strategy whose origins are probably not traceable.

Some of the recent evolutionary literature occasionally cites work of the mid 1960s as embodiments of such ideas, but such work was clearly preceded by earlier developments. The key to differentiating the contributions of such methods obviously rests on the novelty of the component heuristics and the ingenuity of the strategies for coordinating them. Such concerns are more generally the focus of parallel processing solution methods, and many "evolutionary" contributions turn out chiefly to be a subset of the strategies that are being developed to a higher level of sophistication under the parallel processing rubric.

The adaptive memory classification provides a more precise means of differentiation, although it is not without pitfalls. From a naive standpoint, virtually all heuristics other than complete randomization induce a pattern whose present state depends on the sequence of past states, and therefore incorporate an implicit form of "memory." Given that the present is inherited from the past, the accumulation of previous choices is in a loose sense "remembered" by current choices. This sense is slightly more pronounced in the case of solution combination methods such as genetic algorithms and scatter search, where the mode of combination more clearly lends itself to transmitting features of selected past solutions to current solutions. Such an implicit memory, however, does not take a form normally viewed to be a hallmark of an intelligent memory construction. In particular, it uses no conscious design for recording the past and no purposeful manner of comparing previous states or transactions to those currently contemplated. By contrast, at an opposite end of the spectrum, procedures such as branch and bound and A* search use highly (and rigidly) structured forms of memory - forms that are organized to generate all non-dominated solution alternatives with little or no duplication.

Adaptive memory procedures, properly conceived, embody a use of memory that falls between these extremes, based on the goal of combining flexibility and ingenuity. Such methods typically seek to exploit history in a manner inspired by (but not limited to) human problem solving approaches. They are primarily represented by tabu search and its variations that sometimes receive the "adaptive memory programming" label. In recent years, as previously intimated, other approaches have undertaken to incorporate various aspects of such memory structures and strategies, typically in rudimentary form. Developments that produce hybrids of tabu search with other approaches at a more advanced level have become an important avenue for injecting adaptive memory into other methods, and constitute an active area of research.

Another distinction based on memory is introduced by neural network (NN) approaches. Such methods emphasize an associative form of memory, which has its primary application in prediction and pattern matching problems. Neural network procedures also implicitly involve a form of optimization, and in recent years such approaches have been adapted to several optimization settings. Performance is somewhat mixed, but researchers in optimization often regard neural networks as appropriate to be included within the metaheuristic classification. Such an inclusion is reinforced by the fact that NN-based optimization approaches sometimes draw on standard heuristics, and produce solutions by transformations that are not limited to ordinary notions of local optimality. A number of initiatives have successfully combined neural networks with simulated annealing, genetic algorithms and, most recently, tabu search.

Metaheuristics are often viewed as composed of processes that are intelligent, but in some instances the intelligence belongs more to the underlying design than to the particular character (or behavior) of the method itself. The distinction between intelligent design and intelligent behavior can be illustrated by considering present day interior point methods of linear programming. Interior point methods (and more general barrier function methods) exploit a number of ingenious insights, and are often remarkably effective for achieving the purposes for which they were devised. Yet it seems doubtful whether such methods should be labelled intelligent, in the sense of being highly responsive to varying conditions, or of changing the basis for their decisions over time as a function of multiple considerations. Similar distinctions arise in many other settings. It must be conceded that the line that remarks intelligent methods from other methods is not entirely precise. For this reason it is not necessary for a master procedure to qualify as intelligent in a highly rigorous sense in order to be granted membership in the category of metaheuristics.

2.3.2 Tabu Search features

Tabu search is based on principles of intelligent search. The TS premise is that problem solving is qualified as intelligent because it incorporates adaptive memory and responsive exploration. Consider these two features:

- The *adaptive memory* feature of tabu search allows the implementation of procedures that are capable of searching the solution space economically and effectively.

Adaptive memory is shown in the local choices in tabu search that are guided by information collected during the search. The use of adaptive memory contrasts with memoryless methods like simulated annealing and with the use of rigid memory typical of branch and bound strategies. Some types of evolutionary procedures that operate by combining solutions embody a form of implicit memory.

- The emphasis on *responsive exploration* in tabu search, whether in a deterministic or probabilistic implementation, derives from the supposition that a bad strategic choice can yield more information than a good random choice. In a system that uses memory, a bad choice based on strategy can provide useful clues about how the strategy may profitably be changed. Even in a space with significant randomness a purposeful design can be more adept at uncovering the imprint of structure. Responsive exploration integrates the basic principles of intelligent search, i.e., exploiting good solution features while exploring new promising regions.

Tabu search is concerned with finding new and more effective ways of taking advantage of the mechanisms associated with both adaptive memory and responsive exploration. The development of new designs and strategic mixes makes TS a fertile area for research and empirical study.

Dimensions of memory

The memory structures in tabu search operate by reference to four principal dimensions, consisting of: recency, frequency, quality, and influence.

1. Recency-based memory refers to the length of the interval of time since the last time a feature appeared in the search.

2. Frequency-based memory refers to the number of times a feature appeared in the search. Recency-based and frequency-based memory complement each other.

3. The quality dimension refers to the ability to differentiate the merit of solutions visited during the search. In this context, memory can be used to identify elements

that are common to good solutions or to paths that lead to such solutions. Operationally, quality becomes a foundation for incentive-based learning, where inducements are provided to reinforce actions that lead to good solutions and penalties are provided to discourage actions that lead to poor solutions. The flexibility of these memory structures allows the search to be guided in a multi-objective environment, where the goodness of a particular search direction may be determined by more than one function. The tabu search concept of quality is broader than the one implicitly used by standard optimization methods.

4. The fourth dimension, influence, considers the impact of the choices made during the search, not only on quality but also on structure. Recording information about the influence of choices on particular solution elements incorporates an additional level of learning. By contrast, in branch and bound, for example, the separation rules are pre-specified and the branching directions remain fixed, once selected, at a given node of a decision tree. It is clear however that certain decisions have more influence than others as a function of the neighborhood of moves employed and the way that this neighborhood is negotiated (e.g., choices near the root of a branch and bound tree are quite influential when using a depth-first strategy). The assessment and exploitation of influence by a memory more flexible than embodied in such tree searches is an important feature of the TS framework.

Types of memory

The types of memory used in tabu search are both: explicit and attributive.

- *Explicit memory* records complete solutions, typically consisting of elite solutions visited during the search. An extension of this memory records highly attractive but unexplored neighbors of elite solutions. The memorized elite solutions (or their attractive neighbors) are used to expand the local search.

- *Attributive memory* records information about solution attributes that change in moving from one solution to another. TS uses this type of memory for guiding purposes. For example, in a graph or network setting, attributes can consist of nodes or arcs that are added, dropped or repositioned by the moving mechanism. In production scheduling, the index of jobs may be used as attributes to inhibit or encourage the method to follow certain search directions.

Intensification and Diversification

Two highly important components of tabu search are intensification and diversification strategies.

- *Intensification* strategies are based on modifying choice rules to encourage move combinations and solution features historically found good. They may also initiate a return to attractive regions to search them more thoroughly. Since elite solutions must be recorded in order to examine their immediate neighborhoods, explicit memory is closely related to the implementation of intensification strategies. The main difference between intensification and diversification is that during an intensification stage the search focuses on examining neighbors of elite solutions.

  Here the term "neighbors" has a broader meaning than in the usual context of "neighborhood search." That is, in addition to considering solutions that are adjacent or close to elite solutions by means of standard move mechanisms, intensification strategies generate "neighbors" by either grafting together components of good solution or by using modified evaluation strategies that favor the introduction of such components into a current (evolving) solution.

- The *diversification* stage on the other hand encourages the search process to examine unvisited regions and to generate solutions that differ in various significant ways from those seen before. Again, such an approach can be based on generating subassemblies of solution components that are then "fleshed out" to produce full solutions, or can rely on modified evaluations as embodied, for example, in the use of penalty / incentive functions.

Intensification strategies require a means for identifying a set of elite solutions as basis for incorporating good attributes into newly created solutions. Membership in the elite set is often determined by setting a threshold that is connected to the objective function value of the best solution found during the search. However, considerations of clustering and "anti- clustering" are also relevant for generating such a set, and more particularly for generating subsets of solutions that may be used for specific phases of intensification and diversification. The TS notions of intensification and diversification are beginning to find their way into other metaheuristics. It is important to keep in mind that these ideas are somewhat different than the old control theory concepts of "exploitation" and "exploration," especially in their implications for developing effective problem solving strategies.

2.3.3 Scatter Search Features

Scatter search, from the standpoint of metaheuristic classification, may be viewed as an evolutionary (or also called population-based) algorithm that constructs solutions by combining others. It derives its foundations from strategies originally proposed for combining decision rules and constraints (in the context of integer programming). The goal of this methodology is to enable the implementation of solution procedures that can derive new solutions from combined elements. The way scatter search combines solutions and updates the set of reference solutions used for combination sets this methodology apart from other population-based approaches.

Combining solutions

The approach of combining existing solutions or rules to create new solutions originated in the 1960s. In the area of scheduling, researchers introduced the notion of combining rules to obtain improved local decisions. Numerically weighted combinations of existing rules, suitably restructured so that their evaluations embodied a common metric, generated new rules. The conjecture that information about the relative desirability of alternative choices is captured in different forms by different rules motivated this approach. The combination strategy was devised with the belief that this information could be exploited more effectively when integrated than when treated in isolation (i.e., when existing selection rules are selected one at a time). In general, the decision rules created from such combination strategies produced better empirical outcomes than standard applications of local decision rules. They also proved superior to a "probabilistic learning approach" that used stochastic selection of rules at different junctures, but without the integration effect provided by generating combined rules.

In integer and nonlinear programming, associated procedures for combining constraints were developed, which likewise employed a mechanism for creating weighted combinations. In this case, nonnegative weights were introduced to create new constraint inequalities, called surrogate constraints. The approach isolated subsets of constraints that were gauged to be most critical, relative to trial solutions based on the surrogate constraints. This critical subset was used to produce new weights that reflected the degree to which the component constraints were satisfied or violated.

The main function of surrogate constraints was to provide ways to evaluate choices that could be used to create and modify trial solutions. A variety of heuristic processes that employed surrogate constraints and their evaluations evolved from this foundation. As a natural extension, these processes led to the related strategy of combining solutions. Combining solutions, as manifested in scatter search, can be interpreted as the primal counterpart to the dual strategy of combining constraints.

<u>Reference set</u>

Scatter search operates on a set of solutions, the reference set, by combining these solutions to create new ones.

Unlike a "population" in genetic algorithms, the reference set of solutions in scatter search tends to be small. In genetic algorithms, two solutions are randomly chosen from the population and a "crossover" or combination mechanism is applied to generate one or more offspring. A typical population size in a genetic algorithm consists of 100 elements, which are randomly sampled to create combinations. In contrast, scatter search chooses two or more elements of the reference set in a systematic way with the purpose of creating new solutions.

Since the combination process considers at least all pairs of solutions in the reference set, there is a practical need for keeping the cardinality of the set small. Typically, the reference set in scatter search has 20 solutions or less.

In general, if the reference set consists of b solutions, the procedure examines approximately $(3b-7)b/2$ combinations of four different types. The basic type consists of combining two solutions; the next type combines three solutions, and so on and so forth. Limiting the scope of the search to a selective group of combination types can be used as a mechanism for controlling the number of possible combinations in a given reference set.

Scatter Search Template

The scatter search process, building on the principles that underlie the surrogate constraint design, is organized to

(1) capture information not contained separately in the original vectors,

(2) take advantage of auxiliary heuristic solution methods to evaluate the combinations produced and to generate new vectors.

Specifically, the main scatter search features may be sketched as follows:

- *Initial Population.* Generate a starting set of solution vectors to guarantee a critical level of diversity and apply heuristic processes designed for the problem considered as an attempt for improving these solutions. Designate a subset of the best vectors to be reference solutions. (Subsequent iterations of this operation, transferring from below, incorporate advanced starting solutions and best solutions from previous history as candidates for the reference solutions.) The notion of "best" in this step is not limited to a measure given exclusively by the evaluation of the objective function. In particular, a solution may be added to the reference set if the diversity of the set improves even when the objective value of such solution is inferior to other solutions competing for admission in the reference set.

- *Structured combinations.* Create new solutions consisting of structured combinations of subsets of the current reference solutions. The structured combinations are:

  a) chosen to produce points both inside and outside the convex regions spanned by the reference solutions.

  b) modified to yield acceptable solutions. (For example, if a solution is obtained by a linear combination of two or more solutions, a generalized rounding process that yields integer values for integer-constrained vector components may be applied. Note that an acceptable solution may or may not be feasible with respect to other constraints in the problem.)

- *Further improvements.* Apply the heuristic processes used for the initial population Step 1 to improve the solutions created by combinations. (Note that these heuristic processes must be able to operate on infeasible solutions and may or may not yield feasible solutions.)

- *"Best solutions".* Extract a collection of the "best" improved solutions and add them to the reference set. The notion of "best" is once again broad; making the objective value one among several criteria for evaluating the merit of newly created points.

- *Diversifying the reference set*

  Repeat generation, combination and improvements until the reference set does not change. Diversify the reference set, by re-starting from initial population generation. Stop when reaching a specified iteration limit.

The first notable feature in scatter search is that its structured combinations are designed with the goal of creating weighted centers of selected subregions.

This adds non-convex combinations that project new centers into regions that are external to the original reference solutions. The dispersion patterns created by such centers and their external projections have been found useful in several application areas.

Another important feature relates to the strategies for selecting particular subsets of solutions to combine in Step 2. These strategies are typically designed to make use of a type of clustering to allow new solutions to be constructed "within clusters" and "across clusters". Finally, the method is organized to use ancillary improving mechanisms that are able to operate on infeasible solutions, removing the restriction that solutions must be feasible in order to be included in the reference set.

The following principles summarize the foundations of the scatter search methodology:

- Useful information about the form (or location) of optimal solutions is typically contained in a suitably diverse collection of elite solutions.

- When solutions are combined as a strategy for exploiting such information, it is important to provide mechanisms capable of constructing combinations that extrapolate beyond the regions spanned by the solutions considered. Similarly, it is also important to incorporate heuristic processes to map combined solutions into new solutions. The purpose of these combination mechanisms is to incorporate both diversity and quality.

- Taking account of multiple solutions simultaneously, as a foundation for creating combinations, enhances the opportunity to exploit information contained in the union of elite solutions.

The fact that the mechanisms within scatter search are not restricted to a single uniform design allows the exploration of strategic possibilities that may prove effective in a particular implementation.

The success of scatter search and related strategies is evident in a variety of application areas such as vehicle routing, arc routing, quadratic assignment, financial product design, neural network training, job shop scheduling, flow shop scheduling, crew scheduling, graph drawing, linear ordering, unconstrained optimization, bit representation, multi-objective assignment, optimizing simulation, tree problems, mixed integer programming.

2.3.4 MultiStart Features

Recent studies confirmed that intelligent uses of adaptive memory would create improved forms of multistart methods [51]. ¿From a perspective in global metaheuristics, a multistart approach is an extreme version of the strategic oscillation principle. Strategic oscillation operates by alternating constructive and destructive phases. Solutions are generated by a constructive phase and then they are dismantled to a detail degree by the destructive phase, after a new constructive builds the solutions anew. This strategies can be applied to oscillation patterns that destroy large parts of solutions during destructive phases.

The principle of Persistent Attractiveness says that good choices derive from making decisions that often appeared attractive, but that have not previously been made within a particular phase of search. A way to take advantage of this principle is by creating measures of attractiveness for the purpose of modifying customary evaluations of constructive moves. It is made by using persistent attractiveness measures derived from an operation of creating a component evaluator.

The principle of Marginal Conditional Validity derives from the facts that constructive methods make decisions sequentially, and the evaluation of potential decisions depends on those decisions made earlier. Therefore the effect of conditionally is one of the primary determinant of the effectiveness of sequential constructive procedures.

2.3.5 A Classification

The evolution of metaheuristics during the past ten years has taken an explosive upturn. Metaheuristics in their modern forms are based on a variety of interpretations of what constitutes “intelligent” search. These interpretations lead to design choices that in turn can be used for classification purposes. However, a rigorous classification of different metaheuristics is a difficult and risky enterprise, because the leading advocates of alternative methods often differ among themselves about the essential nature of the methods they espouse. This may be illustrated by considering the classification of metaheuristics in terms of their features with respect to three basic design choices:

(1) the use of adaptive memory,

(2) the kind of neighborhood exploration used, and

(3) the number of current solutions carried from one iteration to the next.

These options can be embedded in a classification scheme of the form $x/y/z$, where the choices for x are A (if the metaheuristic employs adaptive memory) and M (if the method is “memoryless”).

The choices for y are N (for a method that employs some systematic neighborhood search either to select the next move or to improve a given solution) and S (for those methods relying on random sampling). Finally, z may be 1 (if the method moves from one current solution to the next after every iteration) or P (for a population-based approach with a population of size P). This simple 3-dimensional scheme gives us a basis of classification, which discloses that agreement on the proper way to label various metaheuristics is far from uniform.

Metaheuristic	Classification 1	Classification 2
Genetic algorithms	M/S/P	M/N/P
Scatter search	M/N/P	A/N/P
Simulated annealing	M/S/1	M/N/1
Tabu search	A/N/1	A/N/P
MultiStart	M/N/1	A/S/1
Variable Neighborhhod	M/N/1	A/N/1

Table 2.1: Metaheuristic classification

We show this by providing classifications for some well-known metaheuristics in Table 2.1.

Two different ways are given for classifying each of these procedures. The first classification most closely matches the "popular conception" and the second is favored by a significant (if minority) group of researchers. The differences in these classifications occur for different reasons, depending on the method. Some differences have been present from the time the methods were first proposed, while others represent recent changes that are being introduced by a subgroup of ardent proponents. For example, the original form of simulated annealing has come to be modified by a group that believes stronger elements of neighborhood search should be incorporated. A similar change came about in genetic algorithms, a few years before it was introduced in simulated annealing, in the mid 1980s. Still, it should be pointed out that not all the advocates of simulated annealing and genetic algorithms view these changes as appropriate.

On the other hand, among those examples where different classifications were present from the start, the foundation papers for tabu search included population-based elements in the form of strategies for exploiting collections of elite solutions saved during the search. Yet a notable part of the literature has not embraced such population-based features of tabu search until recently.

Similarly, scatter search was accompanied by adaptive memory elements as a result of being associated with early tabu search ideas, but this connection is likewise only beginning to be pursued.

A few proponents of simulated annealing and genetic algorithms have recently gone farther in modifying the original conceptions than indicated in Table 2.1, to propose the inclusion of elements of adaptive memory as embodied in tabu search.

Such proposals are often described by their originators as hybrid methods, due to their marriage of aspects from different frameworks.

2.3.6 Desirable properties

Desirable properties of metaheuristics are those which would guarantee both their practical and theoretical interest. A possible list of such properties is the following:

1. *Simplicity:* The metaheuristic should be based on a simple and clear principle.

2. *Precision:* The metaheuristic should be formulated in precise mathematical terms.

3. *Coherence:* The steps of the procedure should follow naturally from the metaheuristic's principle.

4. *Efficiency:* The algorithm should take moderate computing time to provide the solutions;

5. *Effectiveness:* The heuristics should provide optimal or near-optimal solutions for all or at least most realistic instances.

6. *Efficacy:* The heuristics should provide optimal solutions for a great majority of problems.

7. *Generality:* The metaheuristic should be largely applicable to a wide set of problems.

8. *Adaptive:* The metaheuristic should have tools to change and fit different kind of problems or instances
9. *Robustness:* The performance of heuristics should be consistent over a variety of instances.
10. *Interactive:* Heuristics should be user-friendly, easy to understand and easy to use.
11. *Multiple:* The Metaheuristic should provide several high quality and different solutions.
12. *Autonomous:* The heuristics should be parameter-free or have as few parameters as possible.
13. *Innovative:* The metaheuristic characteristics should lead to new types of applications that lead to innovation.

Chapter 3

Provisioning and Routing Problem

An introduction on the Provisioning and Routing Problem (PRP) is provided in this chapter. The aim is to give a wide description of the general problem. Section 3.2 provides a literature review of the general problem and of the problem when using WDM technology. Next section describes the main features of the problem when considering WDM technology. Section 3.4 describes an Integer Linear Programming model proposed for the PRP. Section 3.5 reports the development of a metaheuristic solution approach for solving the problem. Last section summarizes the comparative analysis between the proposed metaheuristic, two variants of a permutation-based approach, and the lower bounds generated by solving the formulation with Cplex using real and randomly generated data.

3.1 Introduction

Two special network flow problems have been the focus in the telecommunications literature: the feasibility problem and the analysis problem. Given a graph and capacitated links, the first problem refers to the feasibility of a set of flows. If a set of demand requirements is also considered, the second problem refers to the determination of feasible flows such that the demand requirements are satisfied. We deal with a third type of network flow problem: the network provisioning problem (also referred to as synthesis or dimensioning), which consists of minimizing the total cost of installing capacity on links of a given network so that demand requirements are satisfied. In these problems, both the physical network topology and the demand requirements are given and the decision variables relate only to adding capacity to links and nodes at minimum cost. When the problem includes also the design of the network topology, that is, determining which links to install, then a complete graph and installation costs are considered.

A given network topology, a cost structure and a set of demand requirements characterize a typical instance of a network provisioning problem. The cost structure depends on each situation as well as on the available technology. It is customary to assume that the system does not add routing costs once the equipment has been installed. Also, these problems typically deal with commodities involving a single source and a single destination. A requirement between two nodes is a single commodity flow requirement. Multi-commodity flow requirements are also considered as long as the commodities involve different origins and destinations while sharing the capacity of the network.

Another typical assumption is that the optical traffic is expressed in OC-48 units, i.e., Optical Carrier level 48 SONET channels. Each such channel carries 2.488×10^9 bits per second, equivalent to 48,672 voice-grade digital channels digitized at 64,000 bits per second each, after subtracting out overhead bits used for routing and control.

When the set of requirements consists of a single demand, the provisioning and routing problem reduces to solving a shortest path problem on a graph with incremental costs as arc weights. The incremental costs are associated with the equipment required to route the smallest allowed demand increment (i.e., one OC-48 if demand splitting is allowed, and the entire demand if no splitting is allowed). More likely, however, the set of demand requirements consists of several origin-destination pairs. In this case, the demand requirements are routed taking into account the spare capacity in the current network. The spare capacity problem is carefully studied because the equipment installed on links and nodes to route a given demand under consideration can also be used to route another demand to be considered later, making the design more cost effective. Provisioning problems also consider non-simultaneous demand requirements, where the capacity installed to route one demand at a given time can be used to route another demand at a different time without additional cost.

The provisioning problem has the following additional applications in the world of telecommunications networking: Multi-hour problems and Survivability problems.

Multi-hour problems. The set of nodes at which demands originate and terminate depends on the time of the day. The capacity installed on the network to satisfy the demand for bandwidth at a given time can be used to route the requirements at a different time without additional cost. Therefore, the total cost of adding capacity is minimized by solving a non-simultaneous multiple requirements problem where each multiple requirement corresponds to a matrix of demands in a period.

The study of this problem was initiated in 1961 by Gomory and Hu [56], who work on a special case of the multi-hour problem with single commodity flow requirements, fractional capacities and the graph is complete with capacity installation costs on all edges equal to 1. In [57] they consider multi-commodity flow requirements and capacity costs given by a linear function of the edge length. The larger instance studied in this work has 10 nodes and 20 edges.

The problem introduced by Gomory and Hu in 1961 has been also studied by Talluri [130] . Some applications have a certain setup cost associated with edges, and it is of interest to design the network with as few edges as possible. This variant of the network synthesis problem was studied by Gusfield [60], who developed two algorithms to build networks with the number of edges less than or equal to those in the networks constructed by Gomory and Hu. Talluri proposes for this variant a new algorithm, which guarantees to use, at most, the number of edges as in the networks of Gusfield's algorithms. In addition, Sridhar and Chandrasekaran [128] gave a polynomial-time algorithm that solves the integer network synthesis problem.

The **Survivability problems** have been described above.

3.2 Provisioning and Routing Problem in WDM Networks: Previous Contributions

WDM/DWDM technology and its related equipment have several advantages when telecommunication companies consider increasing the capacity of an existing fiber network. The first step is increasing the capacity of point-to-point links by using the multiple channels provided by the WDM system. The next step is the switching of the channels in the optical layer by using all-optical cross-connects. In this work we consider WDM networks with wavelength conversion in the cross-connects. In this case, the path on which a point-to-point static demand is carried can have different wavelengths on subsequent links.

When planning WDM optical networks it is important to define the number of wavelengths per fiber to be used (8, 16, 32) and the channel spacing. We assume that only one type of WDM system with 32 wavelengths is used when solving the problem, i.e., we consider only homogeneous networks. Then the planning consists of selecting the paths to route the set of estimated point-to-point demands and provisioning the network. In order to increase the capacity of the network at a minimum cost, it is necessary to decide:

- Where to place WDM and OXC systems;
- How to route the traffic within the resulting network; and
- How to restore the network in the event of any single link failure.

In the literature there are several works related to provisioning of WDM optical networks. This works can be roughly classified in two categories: the case of limited deployed fiber, where provisioning seeks to minimize the number of wavelengths [4], [5], [108], [137]; and the case of limited number of wavelengths per fiber, where provisioning seeks to minimize the amount of required fiber [109] or to maximize accommodated traffic [114].

Caenegem, et al. [131] propose a simulated annealing metaheuristic for designing a fiber topology and optical path layer for WDM Networks minimizing the total cost for a given static traffic demand. They consider two different types of WDM networks; networks that do not use wavelength conversion and networks that use wavelength conversion in the cross-connects. They also consider the protection problem using three rerouting strategies for single link failures: link protection, path protection and path protection with link-disjoint route. They do not take into account uncertainty in traffic demands, which is under the scope of this thesis and will be studied in chapter 6.

Alanyali and Ayanoglu [2] focus on heuristic methods for provisioning a static set of connections on a given WDM optical network topology. They consider that there is a fixed set of wavelengths available on each fiber, which has a cost reflecting the fiber material, optical amplifiers, and the optical termination equipment at both endpoints of the link. However, they consider a fixed positive weight as cost for each link without taking into account the costs of the channels cards required to route traffic on a WDM channel. The design should only consider the cost of equipping the active channels.

Baroni, et al. [6] present several ILP models that attempt to minimize the total number of fibers needed to meet demand for different variations of the routing and wavelength assignment problem. They develop heuristics for large instances and solve to optimality small instances.

Kennington, et al. [74] present an empirical study comparing solution that forbids wavelength conversion with those that permit conversion for the routing and wavelength assignment problem. They address a planning problem with connection requests known with certainty at the time the network is planned and they seek to minimize or maximize some objective function measuring the quality of the assignments. For the extreme cases of no conversion and conversion at every node the instances are solved using [28]. For the more difficult instances, they used a tabu search heuristic. In [72] they developed an optimization model and heuristic for the version without wavelength conversion.

Kennington, et al. [73] study issues concerning uncertain demand forecasts by using robust optimization. They model the wavelength division multiplexing routing and provisioning problem with uncertain demands and a fixed budget as a multicriteria optimization problem. The primary objective is to minimize a quadratic regret function that models the total amount of over and/or under provisioning in the network resulting from uncertainty in the demand forecast. The secondary objective is to minimize the equipment cost that achieves the optimal value for regret. They propose a two-phase robust optimization strategy based on mixed integer linear programs. In the basic provisioning model for each scenario, the objective is to minimize the total cost for provisioning the network with terminal equipment located at each node and optical amplifiers and regenerators associated with the needed links.

DWDM telecommunication network planning is often divided into four main phases: design of the network, routing of the demands, multiplexing and survivability. We assume that there is a current network design such that our problem consists of dealing with the remaining steps of the planning process. In this chapter we do not tackle the problem of protecting the network for link failures. However, since the protection problem can also be treated as a provisioning problem, the proposed formulation for the service network is essentially the same for the protection network, which is typically obtained after the service network has been configured.

Cox, et al. [27] proposed the planning problem that simultaneously addresses the provisioning, routing and survivability problems. The problem was approached using a genetic algorithm (GA), which is based on incrementally adding equipment to minimize the cost of routing each demand. The GA uses permutations to represent solutions. A permutation represents the ordering in which the demands are considered, one by one, for routing purposes. Therefore, a permutation is mapped into an actual solution by a procedure that uses the given order to route the demands in the most cost-effective way. Since the equipment is added to satisfy the current demand without considering the demands that are yet to be routed, each permutation typically results in a different network design. (It is possible, but unlikely, for two different permutations to be mapped to the same network design.) The approach cannot guarantee the existence of an ordering of the demands that would result in an optimal design. In other words, even an exhaustive search of all permutations may result in a sub-optimal network design.

3.3 Problem Description

This dissertation is concerned with the provisioning and routing problem in WDM networks. This optimization problem deals with a set of demands to be routed through the existing optical network. Associated with each demand is an origin node, a destination node, and a size, expressed in OC-48 units. Optical fiber joining pairs of nodes is used to route demands through the network. Each demand can be routed either entirely on one or more bare fibers, over one or more channels of a WDM system or it can be switched from a WDM to another through OXCs. The goal of the network planner is to minimize the total cost, which consists of the cost of additional fiber, WDM systems and OXC equipment.

The existing physical network design (i.e., the set of exiting links) constrains where new optical fibers and WDM systems can be placed. A segment is defined as a sequence of individual links that do not pass through any OXC system. In this case, any intermediate node will be called *glass-through node*, meaning that fiber or a WDM system passes through the node without adding or dropping traffic and without requiring additional equipment. Each OC-48 unit uses two bare fibers or a channel of a WDM system.

For convenience, we will refer to the capacity required for an OC-48 unit as a channel, regardless of whether a pair of fibers or a channel of a WDM system is actually used. All links within a segment must carry the same amount of traffic running from its origin to its destination. In an optimal network design, each segment should follow a least-cost path (with respect to fiber cost) from its origin to its destination. Since the shortest path from any node to any other is treated as a potential segment, the network of segments results in a complete graph, which is intractable in most cases. Therefore, it is useful to generate a subset of promising segments as one of the search strategies.

Once an OXC is reached, wavelengths and fibers can be rearranged. Therefore, the capacity constraints on each segment are simply that enough fiber and WDM equipment must be available on the segment to handle the number of OC-48 units assigned to it. Each individual link must have enough channel capacity to cover all demands routed over segments that uses it.

We will use the network in Figure 3.1 to illustrate the segment-based formulations in the next section. The following three demands are to be routed through the network using some combination of fiber, WDM systems and OXC equipment:

- Route 1 OC-48 from node A to node C.
- Route 1 OC-48 from node A to node D.
- Route 1 OC-48 from node A to node E.

The first step is to determine the set of promising segments that are to be used during the optimization. The set of segments consists of (A,B), (B,C), (B,D), (B,E), which correspond to the individual links in the network, and (A,C), (A,D), (A,E), (C,D), (C,E), (D,E), which correspond to the segments generated by the connection of two individual links. Let us consider as promising segments: (A,B), (B,C), (B,D), (B,E), (A,C), (A,D), and (A,E) (see Figure 3.2). The three segments generated by the connection of two individual use B as a glass-through node. Suppose that WDM equipment with only 3

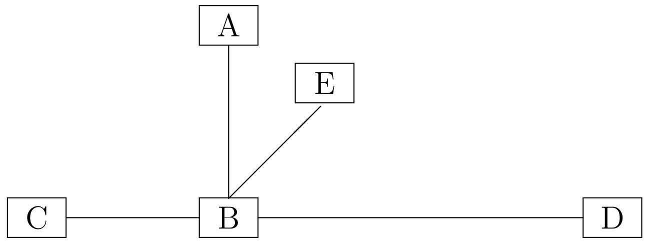


Figure 3.1: Individual-links Network

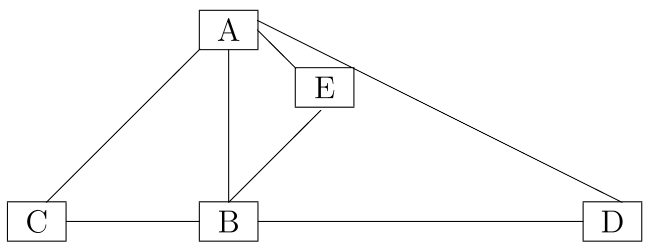


Figure 3.2: Segments Network

channels is available. Depending on the cost of WDM systems, the cost of fiber on (A,B), and the cost of an OXC, the optimal solution would either be:

Option 1

- Route demand A-C on segments (A,B) and (B,C).
- Route demand A-D on segments (A,B) and (B,D).
- Route demand A-E on segments (A,B) and (B,E).
- Put a WDM system on (A,B), (B,C), (B,D), and (B,E).
- Place an OXC with six OXC ports at node B.

- There will be a total of six ports used at nodes A, C, D, and E to add and drop the traffic.

Cost = [fiber costs for links (A,B), (B,C), (B,D) and (B,E)] + [WDM cost for links (A,B), (B,C), (B,D) y (B,E)] + [cost of twelve OXC ports]

Option 2

- Route demand A-C on segment (A,C).
- Route demand A-D on segment (A,D).
- Route demand A-E on segment (A,E).
- Put a WDM system on (A,C) glassed-through at B. The fiber and WDM is groomed to carry only this traffic.
- Put a WDM system on (A,D) glassed-through at B. The fiber and WDM is groomed to carry only this traffic.
- Put a WDM system on (A,E) glassed-through at B. The fiber and WDM is groomed to carry only this traffic.
- There will be a total of six ports used at nodes A, C, D, and E to add and drop the traffic.

Cost = [cost of three fibers on (A,B), one fiber on (B, C), one fiber on (B,D) and one fiber on (B,E)] + [cost of three WDM systems, one for the (A,C) fiber, one for the (A,D) fiber and the other for the (A,E) fiber] + [cost of six OXC ports to add and drop traffic at demand origins and destinations].

Option 1 requires four WDM systems to be installed (one from A to B and one each from B to C, from B to D, and from B to E). It also requires an OXC to split off the wavelengths coming off the AB WDM system and put them on the three other (BC, BD

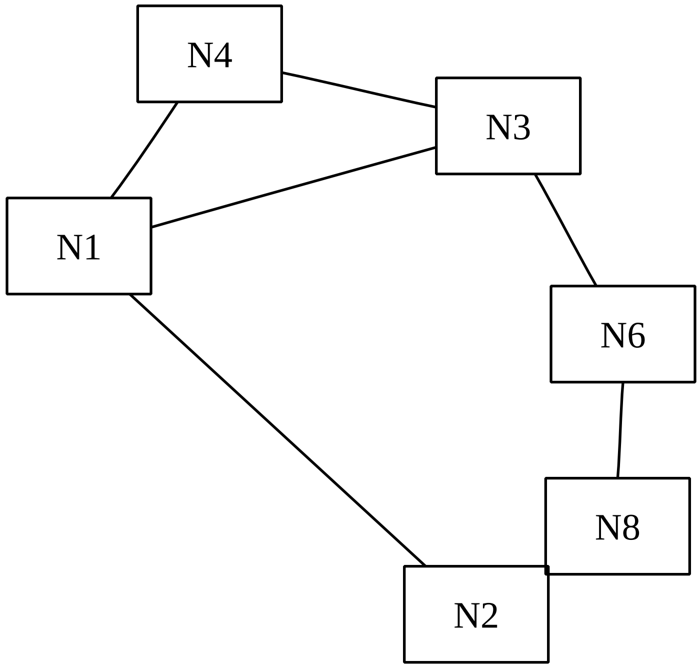


Figure 3.3: Test Glassthrough

and BE) WDM systems. Option 2 saves the cost of an OXC at B and one WDM system, but involves longer fiber routes.

The rationale behind generating a set of potential segments given a network topology is that it provides a cost reduction. Let us consider the network topology in Figure 3.3 and let us suppose that 1 OC-48 unit has to be routed from node $N3$ to node $N2$. If the segment joining the nodes $N3$ and $N2$ is not added to the physical network, then a WDM must be put on segment $(N3, N1)$, another must be put on segment $(N1, N2)$, and an OXC system has to be placed at node $N1$. There will be two ports used at node $N1$, one port at node $N3$ for adding traffic, and one port at node $N2$ for dropping traffic. The total cost of this design is 405,600 dollars. If segment $(N3, N2)$ is added to the physical topology, then only one WDM system glassed-through at $N1$ has to be installed on it and there will be a total of two ports for adding and dropping traffic at nodes $N2$ and $N3$. The total cost of this design is 262,800 dollars.

3.4 Node-Segment Model

We present a mixed integer linear programming (MILP) cost model for the provisioning and routing problem. We provide a node-segment formulation for the problem with multicommodity flow requirements without uncertainty in key data and without protection.

Cox, et al. in [27] propose a node-segment formulation for the planning problem that tackles simultaneously the multiplexing, routing and survivability problems. Our node-segment formulation is based on Cox, et al.'s formulation with the difference that we do not tackle simultaneously the provisioning, routing and survivability problems but instead deal with survivability after provisioning and routing of the service network. We will show that our formulation of the problem using a path-assignment approach to represent solutions results in improved outcomes when compared to tackling the whole problem with the permutation based approach proposed by Cox, et al.

The network topology is represented as an undirected graph $G = (N, E)$, where N denotes the set of nodes and $E \subseteq N \times N$ denotes the set of segments. In our formulation links and segments are equivalent in that they represent a directed connection between two points. The cost of using an individual link or segment is correctly computed in the objective function. A non-simultaneous multicommodity flow requirement, consists of a set of demands $D = (o_1, d_1, R_1), (o_2, d_2, R_2), \cdots, (o_q, d_q, R_q)$ to be routed through the graph. Each single demand consists of an origin node, o_i, a destination node, d_i, and a size, R_i.

Most of the cost sources are mapped to the segment and node costs. The segment cost is subdivided into two components: the cost related to the fiber and the cost related to the channels if WDM systems have to be installed on the fiber. Furthermore, the fiber cost is also subdivided in three different quantities: the fiber terminating equipment (WDM systems) cost, which is a fixed amount; the fiber cost, which depends on the fiber length; and the amplifiers cost, which depends on the number of amplifiers installed on the fiber. These three different costs have to be added to the total cost for each fiber used on each segment. For each WDM channel used in a fiber, the cost for the channel cards and for the wavelength converter has to be added to the total cost. In addition, the node costs include the cost of installing the all-optical cross-connects and the ports required as origin and destination of each channel, which can be either a WDM channel or a pair of bare fibers.

The formulation presented in this section use the following definitions.

3.4.1 Data

Cost Input Data

- C_e^F = cost of a fiber on segment e (sum of costs per link along that segment).
- $C_e^{W_j}$ = cost of a type $j \in J$ WDM unit on segment e.
- C^{O_l} = cost of a type $l \in L$ OXC unit.
- C^{c_j} = channel cost of a type j WDM unit.
- C^{p_l} = port cost of a type l OXC unit.

Capacity Data

- M^{w_j} = capacity of a type j WDM unit.
- M^{o_l} = capacity of a type l OXC unit.

Existing Infrastructure

- g_e^j = spare WDM channels on WDM systems of type j on segment e.
- h_n^l = spare OXC ports on OXC systems of type l at node n.

3.4.2 Decision Variables

- x_{ie} = amount of demand i routed on segment e.
- x_{ie}^F = amount of demand i routed on segment e in the forward direction.
- x_{ie}^R = amount of demand i routed on segment e in the reverse direction.
- f_e = number of stand-alone (no WDM) fiber pairs on segment e.
- w_e^j = number of type j WDM units on segment e.
- v_e^j = number of channels on type j WDM units on segment e.
- y_n^l = number of type l OXC units installed at node n.
- u_n^l = number of ports on type l OXC units installed at node n.

3.4.3 Objective Function

The objective function to be minimized is the sum of fiber costs (first term), WDM costs (second term) and the OXC costs (third term).

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \sum_{j \in J} \left(\left(C_e^F + C_e^{W_j} \right) w_e^j + C^{c_j} v_e^j \right) +$$

$$\sum_{n \in N} \sum_{l \in L} \left(C^{O_l} y_n^l + C^{p_l} u_n^l \right) \tag{3.1}$$

3.4.4 Constraints

The following constraints require that all demand must be carried, that no link should be assigned more demand than its capacity allows it to carry and that no switching element should be assigned more traffic than its capacity allows.

Conservation of Service Flow

$$\sum_{\substack{e \in E \\ o_i = start(e)}} x_{ie}^F + \sum_{\substack{e \in E \\ o_i = end(e)}} x_{ie}^R - \sum_{\substack{e \in E \\ o_i = end(e)}} x_{ie}^F - \sum_{\substack{e \in E \\ o_i = start(e)}} x_{ie}^R = R_i, \quad \forall (o_i, d_i, R_i) \in D \tag{3.2}$$

$$\sum_{\substack{e \in E \\ d_i = end(e)}} x_{ie}^F + \sum_{\substack{e \in E \\ d_i = start(e)}} x_{ie}^R - \sum_{\substack{e \in E \\ d_i = start(e)}} x_{ie}^F - \sum_{\substack{e \in E \\ o_i = end(e)}} x_{ie}^R = R_i, \quad \forall (o_i, d_i, R_i) \in D \tag{3.3}$$

$$\sum_{\substack{e \in E \\ j = end(e)}} x_{ie}^F + \sum_{\substack{e \in E \\ j = start(e)}} x_{ie}^R - \sum_{\substack{e \in E \\ j = start(e)}} x_{ie}^F - \sum_{\substack{e \in E \\ j = end(e)}} x_{ie}^R = R_i, \quad \forall (o_i, d_i, R_i) \in D, \forall j \in N, j \neq o_i, d_i \tag{3.4}$$

$$x_{ie}^F + x_{ie}^R = x_{ie}, \forall (o_i, d_i, R_i) \in D, \forall e \in E \tag{3.5}$$

Segment Capacity

$$\sum_{(o_i,d_i,R_i) \in D} x_{ie} \leq f_e + \sum_{j \in J} v_e^j, \forall e \in E \tag{3.6}$$

$$v_e^j \leq w_e^j M^{w_j} + g_e^j, \forall e \in E, \forall j \in J \tag{3.7}$$

Switch Requirements

$$\sum_{n=end(e)} \left(f_e + \sum_{j \in J} v_e^j \right) \leq \sum_{l \in L} u_n^l, \forall n \in N \tag{3.8}$$

$$u_n^l \leq y_n^l M^{O_l} + h_n^l, \forall n \in N, \forall l \in L \tag{3.9}$$

Integrality Constraints. All variables are nonnegative integer.

This formulation assumes an undirected graph. For directed graphs, the variables x_{ie}^F and x_{ie}^R can be eliminated. That is, the R-variables are entirely eliminated and x is used in place of the F-variables. With this definition, each segment needs to be listed only once, but can be used in either direction. A WDM system on that segment can also be used in either direction.

The studied formulation thus requires the following number of variables:

$$|E| + 2|E||J| + 2|N||L| + 3|D||E|,$$

and the following number of constraints:

$$2|D| + |D|(|N| - 2) + |D||E| + |E| + |E||J| + |N| + |N||L|.$$

The number of variables and constraints scale with the number of links and nodes in the network. They also scale with the number of demands. The integer constraints are necessary, otherwise it would be most likely to find solutions that are not feasible because of noninteger values for capacities.

The preceding objective function, decision variables, and constraints specify a formal version of the provisioning and routing problem. In practice, however, demands and costs are typically uncertain, while available technology options, such as OXC and WDM system capacities, change frequently as new products are introduced. Therefore, this formal version of the problem only approximates a more complicated provisioning and routing problem with uncertainty in key data and changing constraints. This problem will be discussed in chapter 6.

Rather than pursuing a true multi-period optimization approach with formulations that explicitly model uncertainty, many designers prefer to work with a simpler formulation and re-run the associated optimization procedure frequently as conditions change. Previous values of decision variables may then become initial conditions for a new optimization run, which would possibly involve changes in costs and an expanded set of technically feasible options. Repeatedly running a static optimization procedure with changing inputs is, in principle, a sub-optimal approach to adaptive planning. However, in practice, such "rolling optimization" is often preferred to theoretically more realistic dynamic formulations for which required input data cannot be estimated with an appropriate accuracy level.

Additional, and perhaps more realistic, formulations are obtained by constraining the optimization as follows:

Maximum length allowed for WDM systems The signal reach stimulated by WDM equipment spans a maximum of about 400 miles without electronic regenerators. If the length of the segment, with an installed WDM system, exceeds that distance, additional WDM systems must be placed back to back along the segment. Our proposed metaheuristic procedure handles this constraint, and the additional WDM systems and channel cards are subsumed in the formulation by adding the appropriate cost.

Allowed technologies The environment in which the network design problem arises may limit the fiber, WDM systems, and OXC equipment that can be assumed that are available for use.

3.5 Metaheuristic Solution Approach

This section summarizes the development of a metaheuristic procedure that searches for optimal solutions to the provisioning and routing problems. As indicated above, the protection problem is tackled as a provisioning problem, which is solved after the service problem, so that a similar procedure can be used to design the working and protection networks. The MIP model presented in the previous section has a very large number of variables and constraints, making it impractical for the exact solution of real instances of moderate or large size. For small planning problems, the MIP formulation can be solved in reasonable amount of computer time, as shown in our computational experiments. However, the exact solution of the model is only a lower bound on the optimal solution to the real problem.

Our solution procedure employs the notion of a *base network*, which initially consists of the current network design. A base is an incomplete network design that does not satisfy the set of demand requirements that a complete design should be capable of handling. As the process iterates, the base network evolves and the estimated cost of routing a demand becomes more accurate. An evolved base network includes additional equipment, which has been tentatively added to the original base. When a demand is considered for routing on an evolved base network, this demand can share the additional capacity with other demand requirements, making the cost estimates more accurate, due to a decreasing fraction of the capacity that is not shared for costing purposes. The evolution of the base network is linked to an adaptive memory mechanism that keeps track of where new equipment is added in the best solutions recorded during the search. The solution approach that we propose builds a list of paths for each demand by making use of an efficient implementation of the k-shortest path algorithm. This procedure identifies a controlled set of feasible paths for each demand [46] and is a variant of the k-shortest path algorithm reported in [89]. The paths for a given demand are found calculating the incremental cost of routing the entire demand in the base network. For example, one of the possible paths would be to add the necessary fiber and WDM systems to create a segment from the origin to the destination of a given demand. Other paths are created using alternative ways of carrying the demand from origin to destination, which would most likely imply adding WDMs and OXCs.

Four basic elements are common to heuristic searches, regardless of the specific methodology or strategic design choices: (1) a solution representation, (2) an objective, (3) an evaluation function, and (4) a move mechanism. The specifications for our proposed search procedure are:

Solution representation. The construction of a solution starts with the selection of a path for each demand requirement. Once each demand is assigned to a path, the cost of the resulting design is calculated. The cost is associated with the equipment that is required to satisfy the demands using the chosen paths. A solution is fully determined by a data structure that stores the path assignments and the equipment required in each element of the original network.

Objective. The goal of the DWDM planning problem is to minimize the sum of additional fiber cost, WDM equipment cost and its terminal equipment (OXC units) cost, subject to the appropriate technology constraints.

Evaluation. Once each demand has been assigned to a path in its list of potential paths, the evaluation of the solution consists of calculating the increase of capacity required in the elements of the network that route the demands through the assigned paths. The increased capacity is then translated into cost of installing additional fiber and adding WDMs and OXCs.

Move mechanism. Every solution has a neighborhood, which consists of all the feasible solutions that are reached by changing a demand from one path to another.

Our overall solution strategy consists of an adaptive metaheuristic method that combines ideas from scatter search [82], multistart [59], and tabu search [50]. The hybrid metaheuristic takes advantage of strategies that can explore a large solution space effectively. Specifically, tabu search contributes with a short term memory component that is designed to avoid cycling. Scatter search adds a mechanism to generate new solutions from the combination of solutions in an updated reference set of solutions. Finally, the multistart component uses a long term memory that forces construction of new solutions in a wider range of the solution space.

Figure 3.4 shows the main steps of our proposed procedure. The procedure starts with the generation of a set of promising segments using the shortest path algorithm (with distances as weights). Segments corresponding to any existing WDM systems are also included in the promising set. The procedure uses these segments to execute the k-shortest path algorithm for each demand (with incremental costs from a base network B as weights). After the execution of this step (line 5) each demand has a set of paths that are used as the basis for building solutions. Given the network of segments, the spare capacity on the segments and nodes is determined. Obtaining spare capacities allows the procedure to assess incremental costs of routing demands in each segment.

The initial reference set is constructed in lines 6 and 7. The set is populated using a constructive procedure (line 7) that attempts to assign demands to paths in order to efficiently utilize the spare capacity in the original base network. The rationale behind this initialization is that spare capacity for channels in the final network design should be zero except for channels on WDM systems covering a segment without slack. The strategy acknowledges that spare capacity in the original network simply accounts for existing network infrastructure. The solutions in the reference set are ordered according to their total cost, where the first solution, labelled RefSet1, is the one with lowest cost. The reference set is updated as the process iterates (lines 16 and 20). The notion of a reference set is the same as the one used in the scatter search methodology.

We use the current solution S, which at the beginning is the first solution in the reference set (line 8), to obtain an ordering of the demands according to their unit cost (where a unit is an OC- 48). In this ordering, the demand that contributes most (per OC-48) to the total cost of the design is first and the one that contributes least (per OC-48) is last. The demand ordering is important, because the local search, which is based on changing one demand from its current path to another, starts with the demand that has the largest unit cost. To calculate the unit cost, the demands are examined one by one. The examination consists of deleting the demand from the current solution and calculating the cost reduction. The cost reduction is then divided by the bandwidth requirement of the demand under consideration. Once all demands have been examined, the unit cost associated with each demand is known.

The neighborhood search (line 12) within the local search in lines 11 to 18 examines moves employing the ordering of the demands determined in line 10. That is, the first candidate move is to reassign the demand that is at the top of the unit cost list. If reassigning this demand leads to an improving move, the move is executed to change the current solution (lines 13 and 14). If the new solution is better than the worst in the current reference set, then the reference set is updated (line 16). If an improving move that involves reassigning the first demand in the list cannot be found, then the second demand is considered. The process continues until a demand is found for which a reassignment of paths leads to an improving move. If all the demands are examined and no improving move is found, the local search is abandoned. Once the local search is abandoned, the procedure compares the current reference set with the reference set before the last time the local search was executed (line 19). If the reference set did not change after the last execution of the local search, the set is rebuilt (line 20). The process of evolving the base network from a reference set is mainly deterministic and therefore if the reference set does not change, then the base does not evolve properly. Rebuilding of the reference set entails keeping the top $|RefSet|/2$ solutions and generating new solutions to substitute the worst $|RefSet|/2$ in the set, as typically done in implementations of scatter search.

Numerous studies show that effective metaheuristic procedures keep a balance between search intensification and diversification, that is, between reinforcing attributes associated with good solutions and driving the search into regions not yet visited. To achieve this balance, the original base is evolved (line 21) employing the information embedded in the reference set. One of the main criteria used to evolve the base network relates to the number of times a segment appears in the paths assigned to the demands in the $RefSet$ solutions. The procedure also uses global (referred to the whole search process) and local (referred to the current reference set) information in the form of counters that keep track of the number of channels used in each segment in order to decide where to add equipment to the current base. The difference between the maximum global and the maximum local number of channels used in each segment shows its importance. The smaller the difference the more important the segment is in the final network design.

For each demand, the k-shortest paths are again calculated (line 22) by using the incremental costs of routing the demand through the new base network. This step updates the list of best candidate paths according to the current base network. The new paths take advantage of the additional channels included in the evolved base network that can be used without increasing the total cost of the design. The local search now starts from an initial solution constructed to best utilize the spare resources in the new base (line 23). The procedure includes intensification and diversification strategies in the evolution of the base network and in the utilization of the spare capacity during the construction of a starting solution for the local search. The procedure terminates after a pre-specified number of iterations.

3.6 Computational Results

In this section, we present and discuss our computational experiments. We first describe the problem instances that were used to carry out the experimentation. Then we report the results of three experiments. All programs were implemented in C and compiled with Microsoft Visual C++ 6.0. All experiments were performed on a PC with one Pentium 4 processor at 2.53 GHz.

The problem instances used for testing are both real (shared by Dr. Leonard Lu of AT&T Labs) and randomly generated. The random instances are based on the networks corresponding to the real instances, with the demands and existing equipment randomly generated. The motivation for generating random instances is to study the performance of our methods on instances with various characteristics. We consider four different network sizes (with number of nodes varying from 11 to 113) and we generate links to create several densities. Finally, several sets of uniform and clustered sets of demands are randomly created. Uniform demands are generated by randomly selecting an origin and a destination, where each pair has the same probability of being selected. Clustered demands are generated selecting a subset of nodes as high traffic locations and then generating a demand pattern that has a higher density around those nodes. The problem instances used for testing are summarized in Table 3.1. For each set, Table 3.1 shows the name, the number of nodes N , links L , and the different numbers of demands D . The last set in the table has one problem only, corresponding to a real set of demands. No artificial demands were generated for this network.

Table 3.2 summarizes the data regarding the equipment cost used in the solution of the problem instances listed in Table 3.1.

Table 3.3 shows the characteristics for the instance MetroD of the mathematical model described above.

3.6.1 Metaheuristic - Permutation Based Procedure

For comparison purposes, we have implemented a permutation-based algorithm that follows the same structure as the one proposed in [27]. In this approach, a permutation represents the ordering in which the demands are considered for routing. A permuta-

Set Name	$\|N\|$	$\|E\|$	$\|D\|$
MetroD	11	16	10, 20, 30, 54
		27	10, 20, 30, 54
		42	10, 20, 30, 48, 54
Extant0D	12	17	15, 19, 21, 44, 66
		33	15, 21, 44, 66
		46	15, 19, 21, 44, 66
Example2D	17	26	27, 36, 79, 81, 135
		68	27, 36, 81, 135
NationalD	50	63	45, 65, 91, 112
108_annealed-3D	113	137	130

Table 3.1: Test problem characteristics

Constant	Cost	Description
C_e^F	$\$1,400 * length(e)$	Cost of a fiber on a segment e
C_e^W	$95,000	cost of a WDM unit
C^O	$120,000	cost of an OXC unit
C^c	$18,000	channel cost of a WDM unit
C^p	$10,000	port cost of an OXC unit

Table 3.2: Description of costs

tion is mapped into a solution by a procedure that uses the given order to route the demands in the most cost-effective way. When the first demand is considered for routing the current design consists of the original network. The demands are considered one by one as specified by the order in the current permutation. Additional equipment is added as required and the design is updated. The permutation is fully mapped when all the demands have been considered. The approach has the goal of locally minimize the addition of equipment as each demand is routed through the network. More details of the approach can be found in [27] where the permutation search is conducted using a genetic algorithm. Since the genetic algorithm used in [27] is a proprietary code of Cox

Set Name	$\lvert N\rvert$	$\lvert E\rvert$	$\lvert D\rvert$	Total Constraints	Integer Variables
MetroD	11	16	10	324	550
			20	594	1030
			30	864	1510
			54	1512	2662
		27	10	456	913
			20	836	1723
			30	1216	2533
			54	2128	4477
		42	10	636	1408
			20	1166	2668
			30	1696	3928
			48	2650	6196
			54	2968	6952

Table 3.3: MIP characteristics for the instance MetroD

and Associates, Inc., we employ OptQuest [110], a commercial scatter search solver that is capable of searching a permutation space.

Our first experiment consists of comparing the solutions obtained by the permutation-based metaheuristic and our hybrid metaheuristic approach applying the Wilcoxon Signed Ranks Test [29]. The objective is to determine if we may conclude from sample evidence that there is a significant difference between these two procedures. We apply both procedures to the problems in Table 3.1 and record the objective function values obtained by each procedure. Then, compute the absolute objective function value differences (without regard of the sign) for each problem and all differences of zero are omitted. Let the number of pairs remaining by denoted by n. Ranks from 1 to n are assigned to these n pairs according to the relative size of the absolute difference, as follows. Rank 1 is given to the pair with the smallest absolute difference; rank 2 is given to the pair with the second smallest difference; and so on, until rank n is assigned to the pair with the largest absolute difference. If several pairs have absolute differences that are equal to each other, we assign to each of these several pairs the average of the ranks that they would have been assigned if ties were broken arbitrarily.

Wilcoxon suggested a T statistic, which has the approximate quantiles given by the normal distribution, under the null hypothesis that there are no significant differences between the two compared procedures. The critical region of approximate size $\alpha = 0.001$ corresponds to all values of T less than -3.0902. Since in our case $T = -3.217$, the null hypothesis is rejected and we may conclude that there are significant differences between the two metaheuristic procedures.

We have now established that our procedure performs significantly better than the permutation-based approach, as indicated by our statistical test. In our second experiment, we assess the quality of the solutions obtained by the application of our hybrid metaheuristic. For this experiment, we give the MIP formulation presented in section 4 to the Cplex 8.0 MIP solver. The solution of this model provides a lower bound because the number of intermediate nodes for paths between origin and destination pairs is not bounded. The costs shown in Tables 3.4 and 3.5 do not include the cost of additional equipment needed for the restoration of traffic after a single link failure. The first column in Tables 3.4 and 3.5 identifies the problem set. Columns 2, 3, and 4, contain the number of nodes, number of segments, and number of demands. Under the headings "PERM" and "METAH" we report the total costs in million dollars and CPU times corresponding to the permutation based procedure and the proposed metaheuristic, respectively. Under the heading Cplex we report the total cost obtained by solving our MIP formulation. Also under the Cplex heading we report either the total solution time or the optimality gap, if Cplex is unable to find an optimal solution after 2 hours of execution. For the instance 108-Annealed, Cplex was able to find an integer solution only after 8 hours of CPU time. The solution to this problem reported in Table 3 was found after 10 hours of execution. The last two columns in this table show the deviation between the PERM solution and the METAH solution, and between the METAH solution and Cplex solution, respectively.

Several observations can be made regarding the results shown in Tables 3.4 and 3.5. First, the instance-by-instance comparison between PERM and METAH shows that only in 2 instances (Extant0D-12-17-66 and Extant0D-12-33-44) out of 40 the PERM solutions are better than the METAH solutions. The largest difference in favor of the METAH approach occurs in the largest problem (10.04% for the Annealed instance). The deviations of the METAH solutions from the lower bounds found with Cplex are small, ranging from 0 to 5.7%. For the problem for which Cplex could not terminate within 2 CPU hours, our hybrid metaheuristic was able to always improve upon the best upper bound. The results shown in the previous tables corroborate the merit of the solutions found by the proposed hybrid metaheuristic.

One of the main components of our hybrid metaheuristic is the local search. In our last experiment, we use the permutation-based procedure to independently test the effectiveness of the local search. That is, we use the procedure to isolate the local search from other components of our hybrid metaheuristic in order to assess its effectiveness. As described in the previous section, the local optimizer performs a first-improving local search in the neighborhood of the current solution. When a network design is obtained using a permutation of the demands, it is possible to execute our local search to try to improve upon the given solution. Our experiments show that the designs obtained after executing the local search are typically better than the initial designs obtained using the permutation procedure alone. We applied the permutation based approach augmented with the local search to all the instances in Table 1 and use the results to test for significant differences between this approach (PERM+LS) and the one that does not use the local search (PERM).

We once again use Wilcoxons test with $a = 0.001$. Since for the procedures PERM and PERM+LS, $T = -3.516 < -3.0902$, we conclude that the differences between the permutation based procedure and the permutation based procedure with the local search are statistically significant. Therefore, we may say that the application of the local search to the designs obtained by the permutation based procedure generally improves upon the final network designs by reducing their total costs.

Our previous experiment and associated statistical test have determined that the performance of the permutation-based procedure is enhanced with the application of our local search. In our final comparison, we test if there is a significance difference between PERM+LS and our hybrid metaheuristic (METAH). The application of Wilcoxons test results in $T = -0.355$. The critical region of size $a = 0.05$ corresponds to values of T less than -1.6449. Since $T > -1.6449$, the test leads to the conclusion that the performance of PERM+LS is not significantly different than the performance of METAH. This conclusion indicates that the local search that we have designed is quite effective and can be used to improve upon solutions yielded by construction procedures.

3.6.2 MSTS - Multistart

This section shows the effectiveness of developing a hybrid metaheuristic for solving the provisioning and routing problem in WDM mesh networks. The hybrid metaheuristic MSTS developed above takes advantage of the strategies scatter search and tabu search.

Figure 3.5 summarizes the main steps of a multistart procedure, which has been obtained by omitting the components of the scatter search and tabu search of Figure 3.4. The multistart procedure starts generating the set of promising segments and paths for the given initial base in lines 3 and 5, respectively. The constructive method that attempts to efficiently utilize the spare capacity in the original base constructs the initial current solution in line 6.

The current solution S is then used for obtaining an ordering of the demands according to their unit cost in line 9. The local search in lines 10 to 17 examines moves employing the ordering of the demands. If reassigning demands to different paths leads to a solution better than S, then the move is executed to change the current solution S. If the new solution is better than the best solution, $BestSol$, then the best solution is changed.

Once the local search is completed, the constructive method constructs a new starting solution for the local search in line 18. Finally, the procedure terminates after a pre-specified number of iterations.

Tables 3.6, 3.7, 3.8, and 3.9 report the design costs in million dollars and CPU times in seconds for both the MSTS and the MS metaheuristics. Both metaheuristics have been executed using as stopping condition the same pre-specified number of iterations, where an iteration is a local search.

For small instances both metaheuristics reach the same solutions. However, our experiments corroborate the effectiveness of the MSTS metaheuristic as the size of the problems increases.

3.7 Conclusions

We have addressed an important and current problem in the telecommunications industry. We have provided the motivation for studying this optimization problem and have discussed the technology behind it. Our segment-based formulation is used as the framework for developing heuristic procedures and as a means for finding lower bounds.

Our experiments with real and randomly generated data show the merit of our proposed solution procedure when compared to a permutation-based approach and to the lower bounds generated by solving an MIP formulation with Cplex. We used a nonparametric statistical test to compare our procedure and two variants of a permutation-based approach. The test revealed the effectiveness of our local search, which is capable of improving solutions constructed with the permutation-based approach to a point that the resulting method is statistically comparable to the proposed hybrid metaheuristic.

Although our general approach contemplates solving the protection problem as part of the design process, in the scope of this chapter we have not included the implementation and experimentation associated with network survivability. An extension of our work will include solving the protection problem using the last reference set obtained when the termination criterion (line 24 in Figure 3.4) is satisfied. Specifically, given a network design (i.e., a solution in the reference set), the protection problem would consist of finding the most cost-efficient way of routing demands after a link failure. We believe that the lessons learned while tuning the procedure for finding good solutions to the working or service problem will be valuable in the development of a comprehensive procedure that includes protection. This is addressed in chapter 5.

```
Procedure Multistart Scatter Tabu Search
{
    generate (segments);
    B = initial_base;
    Paths = find_k-shortest_ paths ( B );
    for (i = 1, · · ·, |RefSet|)
        RefSet_i = constructive ( B, Paths );
    S = RefSet_1;
    do {
        find_demand_order (S, B);
        while (improving move) {
            move = find_next_improving_move (S,Paths);
            if (move) {
                S = execute (move);
                if (S better than RefSet_last)
                    RefSet = update (S);
            }
        }
        if (equal(RefSet))
            rebuild (RefSet);
        B = evolve (RefSet);
        Paths = find_k-shortest_paths (B);
        S = constructive (B,Paths);
    } until (stoppingcondition)
}
```

Figure 3.4: MSTS

				Perm		MSTS		Cplex				
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Cost	Time	Gap(%)	Dev(P,M)	Dev(M,C)
MetroD	11	16	10	4.09	5.12	4.09	0.13	4.09	0.20	–	0	0
			20	4.38	9.77	4.38	4.90	4.38	0.75	–	0	0
			30	8.42	15.02	8.42	7.37	8.42	0.39	–	0	0
			54	14.13	27.02	14.12	16.60	14.03	6.53	–	0.07	0.64
		27	10	2.75	5.21	2.75	1.96	2.75	0.29	–	0	0
			20	4.31	10.28	4.26	4.30	4.03	1.09	–	1.17	5.70
			30	6.46	15.35	6.46	6.35	6.40	1.76	–	0	0.93
			54	11.39	27.52	11.24	14.69	10.84	17.34	–	1.33	3.69
		42	10	1.96	5.49	1.96	2.03	1.94	0.38	–	0	1.03
			20	3.08	11.16	3.08	3.48	3.06	1.21	–	0	0.65
			30	5.40	16.24	5.38	4.86	5.38	7.21	–	0.37	0
			48	7.31	26.72	7.11	12.24	6.99	10.89	–	2.81	1.71
			54	8.80	29.86	8.50	13.18	8.35	39.44	–	3.52	1.79
Extant0D	12	17	15	3.69	8.16	3.69	3.25	3.69	0.75	–	0	0
			19	6.26	10.41	6.26	9.99	6.26	8.04	–	0	0
			21	6.21	11.43	6.21	9.46	6.21	3.45	–	0	0
			44	14.48	23.55	14.36	27.19	14.36	96.11	–	0.83	0
			66	11.99	35.71	12.14	41.00	11.83	81.81	–	-1.23	2.62
		33	15	3.69	7.82	3.69	3.57	3.69	20.09	–	0	0
			21	7.32	11.60	7.32	7.21	6.03	269.68	–	0	2.13
			44	13.94	23.69	14.23	19.99	14.23	–	7.46	-2.03	–
			66	11.83	35.81	11.83	29.93	13.62	–	19.83	0	–
		46	15	3.69	7.94	3.69	2.30	3.69	38.61	–	0	0
			21	7.32	11.61	7.32	3.83	6.03	770.15	–	0	2.13
			44	13.97	24.43	13.95	18.54	23.43	–	51.16	0.14	–
			66	11.83	36.19	11.77	28.82	25.83	–	63.12	0.50	–

Table 3.4: Comparative Results 11 and 12 nodes

				Perm		MSTS		Cplex				
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Cost	Time	Gap(%)	Dev(P,M)	Dev(M,C)
Ex2D	17	26	27	23.22	26.11	23.22	21.59	22.47	38.42	–	0	3.33
			36	81.84	33.34	81.84	20.69	81.84	492.57	–	0	0
			79	180.70	78.27	178.05	93.94	182.94	–	5.30	2.11	–
			81	98.80	75.44	97.37	89.57	96.65	5438.10	–	1.46	0.74
			135	177.42	127.73	173.03	191.46	182.04	–	7.46	2.59	–
		68	27	24.43	27.12	24.43	18.11	19.27	4406	–	0	2.67
			36	68.15	35.19	68.10	20.20	69.51	–	12.23	0.07	–
			81	84.09	80.05	82.65	80.24	102.12	–	31.92	1.74	–
			135	149.71	135.11	144.14	169.60	–	–	–	3.86	–
National	50	63	45	37.63	125.12	37.36	61.77	44.08	–	39.42	0.72	–
			65	51.16	192.28	50.87	141.70	56.25	–	35.43	0.57	–
			91	59.18	267.46	59.13	162.75	62.77	–	31.59	0.08	–
			112	44.44	296.16	42.88	230.70	51.51	–	51.96	3.6	–
Annealed	113	137	130	118.30	1124.07	107.50	682.47	127.16	10H	57.28	10.04	–

Table 3.5: Comparative Results 17, 50, and 113 nodes

```
Procedure Multistart
{
    generate (segments);
    B = initial_base;
    Paths = find_k-shortest_ paths ( B );
    S = constructive ( B, Paths );
    BestSol = S;
    do {
        find_demand_order (S, B);
        while (improving move) {
            move = find_next_improving_move (S,Paths);
            if (move) {
                S = execute (move);
                if (S better than BestSol)
                    BestSol = S;
            }
        }
        S = constructive (B,Paths);
    } until (stoppingcondition)
}
```

Figure 3.5: Multistart

				MSTS		Multistart		
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Dev(MSTS,MS)
MetroD	11	16	10	4.09	0.13	4.09	2.32	0
			20	4.38	4.90	4.38	4.86	0
			30	8.42	7.37	8.42	6.76	0
			54	14.12	16.60	14.12	14.97	0
		27	10	2.75	1.96	2.75	2.06	0
			20	4.26	4.39	4.26	4.55	0
			30	6.46	6.35	6.46	5.59	0
			54	11.22	14.69	11.25	13.81	0.26
		42	10	1.96	2.03	1.96	2.31	0
			20	3.08	3.48	3.08	3.15	0
			30	5.38	4.86	5.38	6.20	0
			48	7.11	12.24	7.11	12.27	0
			54	8.49	13.18	8.50	14.13	0.11

Table 3.6: MSTS - MS with the 11-node networks

				MSTS		Multistart		
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Dev(MSTS,MS)
Extant0D	12	17	15	3.69	3.25	3.69	5.81	0
			19	6.26	9.99	6.30	5.27	0.63
			21	6.21	9.46	7.53	5.74	21.25
			44	14.36	27.19	14.68	26.46	2.22
			66	12.14	41.00	12.58	21.47	3.62
		33	15	3.69	3.57	3.69	6.72	0
			21	7.32	7.21	8.30	8.70	13.38
			44	14.23	19.99	14.45	12.95	1.54
			66	11.83	29.93	14.51	33.17	22.65
		46	15	3.69	2.30	3.69	4.16	0
			21	7.33	3.83	7.96	8.93	8.59
			44	13.95	18.54	14.79	10.69	6.02
			66	11.77	28.82	13.03	31.43	10.70

Table 3.7: MSTS - MS with the 12-node networks

				MSTS		Multistart		
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Dev(MSTS,MS)
Ex2D	17	26	27	23.22	21.59	24.28	24.38	4.56
			36	81.84	20.69	83.77	33.20	2.35
			79	176.96	93.94	180.44	105.72	1.96
			81	97.37	89.57	100.00	105.14	2.70
			135	172.94	191.46	175.62	263.36	1.54
		68	27	24.43	18.11	25.56	24.33	4.62
			36	68.10	20.20	70.18	46.38	3.05
			81	82.65	80.24	86.65	130.87	4.83

Table 3.8: MSTS - MS with the 17-node networks

				MSTS		Multistart		
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Dev(MSTS,MS)
National	50	63	45	37.36	61.77	40.17	78.17	7.52
			65	51.09	141.70	52.30	157.46	2.36
			91	59.13	162.75	59.82	199.07	1.16
			112	42.88	230.70	45.41	206.69	5.90

Table 3.9: MSTS - MS with the 50-node networks

Chapter 4

Provisioning and Routing Problem: An Alternative Model

In this chapter, we present the segment-path model that determines the equipment required to route a set of point-to-point demands through the network. The goal of this chapter is to compare the solutions obtained by solving the capacity expansion problem on WDM networks with the metaheuristic approach developed above and the alternative proposed formulation.

4.1 Introduction

The proposed formulation has significantly less integer variables than the segment-based formulation presented above. That segment-based formulation is mainly used as a mathematical definition of the problem instead of a mechanism for solving it. In the node-segment formulation there are three traffic variables for each pair demand-segment, which specify the amount of a demand routed on a segment, the amount routed in the forward direction and the amount routed in the reverse direction, respectively. However, in the segment-path formulation presented in this section there is only a traffic variable for each pair demand-path, where the maximum number of paths is a limited value significantly smaller than the number of segments on the network. The formulation presented here is intended for finding optimal or near-optimal solutions to the capacity expansion problem on hand. Our formulation uses the following definitions.

4.2 The segment-path model

4.2.1 Data

The network topology is represented as a graph $G = (N, E)$, where N denotes the set of nodes and $E \subseteq N \times N$ denotes the set of segments. In this formulation as well as in the node-segment formulation, links and segments are equivalent. For each $n \in N$, A_n denotes the set of segments adjacent to node n. The origin/destination node pairs corresponding to the point-to-point demands are given by $D \subseteq N \times N$.

- J_{od} = the set of possible paths from the origin o to the destination d that can be used to route this demand.

Since the set of paths used for each demand may not consist of all the possible paths from o to d, the formulation described in this section may be used as a heuristic model for the provisioning and routing problem.

Cost Input Data

- C_e^F = cost of a fiber on segment e (sum of costs per link along that segment).
- C_e^W = cost of a WDM unit on segment e.
- C^O = cost of an OXC unit.
- C^c = channel cost of a WDM unit.

- C^p = port cost of an OXC unit.

Capacity Data

- M^w = capacity of a WDM unit.
- M^o = capacity of an OXC unit.

Existing Infrastructure

- g_e = spare WDM channels on WDM systems on segment e.
- h_n = spare OXC ports on OXC systems at node n.

4.2.2 Decision Variables

- x_p^{od} = 1 if demand (o, d) is routed on path p and 0 otherwise.
- f_e = number of stand-alone (no WDM) fiber pairs on segment e.
- w_e = number of WDM units on segment e.
- v_e = number of channels on WDM units on segment e.
- y_n = number of OXC units installed at node n.
- u_n = number of ports on OXC units installed at node n.

4.2.3 Objective Function

The objective function to be minimized is the sum of fiber costs (first term), WDM costs (second term) and the OXC costs (third term).

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right) \tag{4.1}$$

4.2.4 Constraints

There are five sets of constraints in this model. The first set of constraints, labelled as (4.2), ensures demand satisfaction and does not allow splitting demands. Constraint set (4.3) converts path capacity to segment capacity and segment capacity into fibers and channels. Constraint set (4.4) converts segment capacity to WDM units. The fourth set of constraints, labelled (4.5), accumulates channels on links to add the required number of ports to each node. The last set of constraints (4.6) converts node capacity to OXC units.

$$\sum_{p \in J_{od}} x_p^{od} = 1, \forall (o,d) \in D \tag{4.2}$$

$$\sum_{(o,d) \in D} R_{od} \sum_{p \in J_{od}, e \in p} x_p^{od} \leq v_e + f_e, \forall e \in E \tag{4.3}$$

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{4.4}$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{4.5}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{4.6}$$

All decision variables are nonnegative integer.

This formulation thus requires the following number of constraints:

$$|D| + 2|E| + 2|N|,$$

and the following number of variables

$$3|E| + 2|N| + \sum_{i=1}^{|D|} k_i.$$

In this formula k_i represents the number of possible paths generated to route demand i. The number of variables and constraints scale with the number of links and nodes in the network. They also scale with the number of demands and paths used for each demand.

Set Name	$\lvert N\rvert$	$\lvert E\rvert$	$\lvert D\rvert$	Total Constraints	Integer Variables	Binary Variables
MetroD	11	16	10	64	70	100
			20	74	70	200
			30	84	70	300
			54	108	70	540
		27	10	86	81	100
			20	96	81	200
			30	106	81	300
			54	130	81	540
		42	10	116	148	100
			20	126	148	200
			30	136	148	300
			48	154	148	480
			54	160	148	540

Table 4.1: MIP characteristics for MetroD

Table 4.1 summarizes the number of constraints and the number of integer and binary variables required for the instance MetroD if a set of 10 paths is generated for each demand and the segment-path model is used.

4.3 Computational Results

The goal of our experimental testing is to compare the solutions obtained by solving the capacity expansion problem on WDM networks with the metaheuristic approach and the formulation presented above. Even if the previous MIP model is solved to optimality, the solution is not guaranteed to be optimal for the original problem because the model includes only a subset of all possible segments and a subset of all possible paths that can be found with the given segments.

The set of instances is the set used for solving the problem with the node-segment formulation. The artificial instances use the same number of nodes as in the networks of the real instances but the demands and the exiting equipment are randomly generated. The intent in generating random instances is to analyze the performance of both methods on instances with various characteristics. We consider four different numbers of nodes and for each we generate networks with different densities according to the number of segments. Then, for each network, several sets of uniform and clustered sets of demands are randomly created. Uniform demands are generated by randomly selecting an origin and a destination, where each pair has the same probability of being selected. Clustered demands are generated selecting a subset of nodes as "high traffic" locations and then generating a demand pattern that clusters around those nodes. For each set of instances, we show the number of nodes —N—, segments —E—, and demands —D—. For both the metaheuristic method and the mathematical model we have considered the same number of paths for each demand pair to make the results comparable. In addition, demands are not split in both cases.

The MIP formulation was solved with Cplex 7.5. All experiments were performed on a PC (with one processor at 1.0 GHz and 256 Mbytes of RAM). Table 4.2 summarizes the computational results for the three networks with 11 nodes. The first network is 30% dense, that is, it consists of 30% of the links in the completely dense network. The second and third networks are 50% and 76% dense, respectively. For each network we have created four random sets of demands to simulate possible situations on a telecommunications network. The first set of demands, which consists of 54 demands, consists of uniformly distributed requirements. The other three sets of demands are generated in clusters, where either only a few nodes generate demand requirements or there are a few "high traffic" nodes with demands to other nodes in the network. The instance consisting of 11 nodes, 42 segments and 18 demands corresponds to a real instance.

Table 4.3 shows the results for three networks with 12 nodes. These networks are 30%, 50%, and 70% dense, respectively. For each network we have also created four random

				MetaH		Cplex		
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Dev(M,C)
MetroD	11	16	10	4.09	0.13	4.09	0.22	0
			20	4.38	0.3	4.38	0.21	0
			30	8.42	0.19	8.42	0.18	0
			54	14.12	2.01	14.03	1.25	0.64
		27	10	2.75	0.13	2.75	0.13	0
			20	4.26	0.25	4.26	0.29	0
			30	6.46	0.44	6.46	0.27	0
			54	11.22	10.76	11.16	2.87	0.53
		42	10	1.96	0.04	1.96	0.11	0
			20	3.08	16.16	3.08	0.13	0
			30	5.38	0.05	5.38	0.19	0
			48	7.11	4.65	7.08	0.51	0.42
			54	8.49	1.05	8.41	0.41	0.95

Table 4.2: Experiments with the 11-node networks

sets of demands. The set with 66 demands consists of uniformly distributed requirements. The other three sets have clustered demands. The instance consisting of 12 nodes, 17 segments and 19 demands corresponds to a real instance. Columns 2, 3, and 4, contain the number of nodes, number of segments, and number of demands in each instance. Columns 5 and 6 in Tables 4.2and 4.3 show the total costs and the CPU time (in seconds) corresponding to the metaheuristic procedure and Cplex, respectively. All instances in these two tables were solved employing a set of 10 paths for each demand. The results obtained with Cplex have an optimality gap of 0.0001, which is the default value in this optimizer. Last column in both tables shows the deviation between the Cplex solution and the metaheuristic solution. In the worst case, the deviation is no greater than 0.95% in Table 4.2 and 2.62% in Table 4.3. For almost every instance in Tables 4.2 and 4.3, the proposed metaheuristic procedure is able to reach the same solution obtained with

				MetaH		Cplex		
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Dev(M,C)
Extant0D	12	17	15	3.69	0.14	3.69	0.15	0
			19	6.26	0.36	6.26	0.71	0
			21	6.21	0.39	6.21	0.63	0
			44	14.36	7.32	14.36	10.70	0
			66	12.14	26.51	11.83	7.38	2.62
		33	15	3.69	0.43	3.69	1.41	0
			21	7.32	3.45	7.32	10.94	0
			44	14.23	38.17	13.94	81.16	2.08
			66	11.83	109.82	11.83	225.32	0
		46	15	3.69	0.22	3.69	0.83	0
			21	7.33	5.79	7.33	25.80	0
			44	13.95	58.16	13.95	467.72	0
			66	11.77	152.12	11.77	224.90	0

Table 4.3: Experiments with the 12-node networks

Cplex. Computational times generally favor Cplex in these relatively small networks.

Table 4.4 shows the results obtained for two networks with 17 nodes. These networks are 19% and 50% dense, respectively. As in the case of Tables 4.2 and 4.3, four sets of demands are randomly generated for each network to provide both uniformly distributed and clustered sets of demand requirements. These instances were solved using a maximum of 6 paths for both the metaheuristic procedure and the MIP formulation. The instance consisting of 17 nodes, 26 links, and 79 demands corresponds to a real instance. Table 4.5 displays the computational results obtained for a network with 50 nodes and 63 links for which four sets of demands have also been created. The instance consisting of 112 demands corresponds to a real instance.

In Tables 4.4 and 4.5, columns 5 and 6 summarize costs and running times for the metaheuristic. Under the heading Cplex we have reported the total cost obtained by

				MetaH		Cplex			
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Gap(%)	Dev(M,C)
Ex2D	17	26	27	23.22	2.14	23.22	5.99	–	0
			36	81.84	0.95	81.84	13.42	–	0
			79	176.96	185.10	180.08	–	3.14	–
			81	97.37	97.63	96.81	296.16	–	0.57
			135	172.94	65.18	173.04	–	1.97	–
		68	27	24.43	13.47	24.43	12.26	–	0
			36	68.10	11.69	67.79	292.93	–	0.45
			81	82.65	233.56	82.74	–	1.27	–

Table 4.4: Experiments with the 17-node networks

				MetaH		Cplex			
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Gap(%)	Dev(M,C)
National	50	63	45	37.36	50.80	36.85	257.54	–	1.38
			65	51.09	112.76	49.73	1467.1	–	2.73
			91	59.13	138.08	59.43	–	1.65	–
			112	42.88	230.70	42.88	262.41	–	0
			268	110.34	558.17	111.36	–	4.94	–

Table 4.5: Experiments with the 50-node networks

solving our MIP formulation and either the total solution time or the optimality gap, if Cplex cannot find an optimal solution after 2 h of execution. Last column in both tables shows the deviation between Cplex solution and the metaheuristic solution. In the worst case, the deviation is no greater than 0.57% in Table 4.4 and 2.73% in Table 4.5 when Cplex is able to find and confirm the optimal solution. When Cplex fails to complete the branch and bound optimization, the upper bound solutions found are in all cases inferior to the solutions found with the metaheuristic procedure.

Specifically, in Table 4.4, Cplex finds the optimal solution of the heuristic MIP model in five instances. For the other four instances, the execution of Cplex was stopped after

2 hours, obtaining inferior solutions than those obtained with the metaheuristic in a shorter time period. A similar pattern is observed in Table 4.5, where Cplex fails twice to find the optimal solution within the 2 h limit.

4.4 Conclusions

This chapter presents a heuristic optimization model for the capacity expansion problem in WDM networks. We have carried out a comparative analysis between the results obtained using a metaheuristic procedure for the problem on hand and the results obtained solving a mathematical model with Cplex. The mathematical model is solved as a relaxation of the original problem because we do not consider all possible segments or paths between each pair of demand requirements. Our experiments corroborate the effectiveness of the metaheuristic developed in chapter 3 as the size of the problems increases. For relatively small problems (i.e., with number of nodes equal to 12 or less), solving the MIP formulation seems to be a better alternative than running the metaheuristic procedure.

Chapter 5

Survivability Problem

This chapter focuses on the provisioning and routing problem with single link failure protection. We develop several models with the aim of solving the problem using both a shared-link protection scheme and shared-path protection schemes. Section 5.2 provides the model for the first scheme and section 5.3 provides two variants of the shared-path protection scheme. The models presented in this chapter start with a solution to the provisioning and routing problem without protection in order to reach the protection network. We propose the use of a modified version of the hybrid metaheuristic developed in chapter 3 for solving the problem. Section 5.4 summarizes the computational results for these schemes.

5.1 Introduction

We use both a *shared-link protection scheme* and a *shared-path protection scheme*. When *link-protection* is used, at the time of demand setup, for each link of the primary path, a backup path is reserved around that link, allowing other backup paths to share the reserved capacity. When *path-protection* is used, for each demand being routed through the failed link, a backup route is reserved between the end nodes of the affected demand. In this way, protection capacity all over the network is used and a lower capacity requirement is expected than in link-protection. In this situation, more network nodes are involved in the protection phase, since every working path needs to be protected separately. Kennington, et al. [71] use a *shared-path protection scheme* and to construct the model they begin with the solution to the model for the provisioning and routing problem without protection. Caenegem, et al [131] also begin with the solution obtained by the model without protection as the input for the model with shared protection.

5.2 Shared-Link Protection Scheme

This section presents the model for obtaining the backup paths for each failed segment beginning with a solution to the model without protection. The model uses the following definitions:

- NC_e = number of channels on segment e required to get the working paths. Then, if segment e fails, the traffic NC_e must be rerouted between the origin and destination nodes of e.
- J_e = set of possible paths between the end nodes of segment e that can be used to reroute the traffic.
- $z_q^e = 1$ if the traffic on segment e is rerouted on path $q \in J_e$, and 0 otherwise.

The model assumes that the traffic on a segment cannot be split when is rerouted. The model can be stated as follows:

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right) \tag{5.1}$$

Subject to:

$$\sum_{q \in J_e} z_q^e = 1, \forall e \in E \tag{5.2}$$

$$\sum_{q \in J_{e'}, e \in q} NC_{e'} z_q^{e'} \leq v_e + f_e, \forall e \in E, \forall e' \neq e \in E \tag{5.3}$$

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{5.4}$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{5.5}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{5.6}$$

All variables are integer and nonnegative.

The goal is to minimize the total cost, which is the sum of the fiber cost (first term), WDM cost (second term), and OXC cost (third term).

When segment e fails, the working flow of this segment must be rerouted through one of the possible protection routes in J_e. Hence constraint set (5.2) must be met. If segment e fails, the spare capacity on the other segments must be sufficient for the flow on the protection routes. Therefore, constraint set (5.3) must be met. As shown in previous chapters, constraint set (5.4) converts segment capacity to WDM units, (5.5) accumulates channels on links to add the required number of ports to each node, and (5.6) converts node capacity to OXC units.

5.3 Shared-Path Protection Schemes

In the path-protection scheme, every interrupted path needs to be restored. One of the alternative protection routes of the path must carry the traffic of the interrupted working path corresponding to a demand. We use two shared-path protection schemes: a protection scheme in which every working path has several sets of alternative protection paths depending on the failed link, and a protection scheme in which the protection routes and the working paths are link-disjoint. In this case, only a set of alternative routes is generated for each working path.

5.3.1 Non Link-Disjoint Alternative Routes

The model presented in this section also solves the protection problem starting with the best solution reached by the model without protection. For each failed segment e and each working path that passes through e, a set of protection routes is generated. These routes must allow to reroute the traffic if segment e fails. Hence, they cannot contain segment e.

This model uses the following definitions:

- $J_{e,od}$ = set of alternative protection routes between the end nodes of demand $(o,d) \in D$ that do not contain segment e. Therefore, it is the set of routes used to reroute demand (o,d) if segment e fails.
- $z_{q,e}^{od} = 1$ if demand (o,d) is rerouted on path $q \in J_{e,od}$ if segment e fails, and 0 otherwise.
- WP_{od} = working path of demand $(o,d) \in D$.

The model assumes that the demand on each working path cannot be split when it is rerouted. In this model the constraints are slightly different to the constraints in the shared-link protection model. The corresponding set of constraints is indicated in the same position as in the link-protection model presented above. The model can be stated as follows:

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right) \qquad (5.7)$$

Subject to:

$$\sum_{q \in J_{e,od}} z_{q,e}^{od} = 1, \forall (o,d) \in D, \forall e \in WP_{od} \tag{5.8}$$

$$\sum_{(o,d) \in D, e' \in WP_{od}} \sum_{q \in J_{e',od}, e \in q} R^{od} z_{q,e'}^{od} \leq v_e + f_e, \forall e, e' \in E, e' \neq e \tag{5.9}$$

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{5.10}$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{5.11}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{5.12}$$

The spare capacity and the flows on the protection routes or backup paths must be integer and nonnegative values.

Since the goal of the problem is to increase the capacity at a minimum cost for protecting the network, the objective is to minimize the total costs as in the previous model.

When a segment e fails, every interrupted path must be restored. Constraint set (5.8) ensures that the protection routes for the path carry the demand of the interrupted working path. If segment e fails, the spare capacity on the other segments must be sufficient for the flow on the protection routes of every interrupted path. Then constraint set 5.9 must be met. Constraint sets (5.4), (5.5), and (5.6) are unchanged applicable for path-protection.

5.3.2 Link-Disjoint Alternative Routes

In this case, we generate alternative protection routes for each demand that are link-disjoint with the working paths, i.e., routes that have no link in common with the working paths. Then, when two or more working paths have any common link, the protection paths assigned for rerouting each demand between its end nodes cannot share the spare capacity. On the contrary, if the working paths of any two or more demands do not have any common link, the protection paths can share the spare capacity. The model uses the following definitions:

- J_{od} = set of alternative protection routes between the end nodes of demand $(o, d) \in D$. These routes are link-disjoint with the corresponding working path. It is the set of routes used to reroute demand (o, d) if segment e fails.
- $z_q^{od} = 1$ if demand (o, d) is rerouted on path $q \in J_{od}$ if segment e fails, and 0 otherwise.

Then the model, whose objective is to minimize the total cost of adding spare capacity, can be stated as follows:

$$\min \sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) + \sum_{n \in N} \left(C^O y_n + C^p u_n \right) \tag{5.13}$$

Subject to:

$$\sum_{q \in J_{od}} z_q^{od} = 1, \forall (o, d) \in D \tag{5.14}$$

$$\sum_{(o,d) \in D, e' \in WP_{od}} \sum_{q \in J_{od}, e \in q} R^{od} z_q^{od} \leq v_e + f_e, \forall e, e' \in E, e' \neq e \tag{5.15}$$

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{5.16}$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{5.17}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{5.18}$$

The spare capacity and the flows are integer and nonnegative.

When a segment e fails, every demand that had been routed on the segment must be restored. Constraint set (5.14) ensures that one of the protection routes of the path carry the demand of the interrupted working path. Constraint set (5.15) ensures that if segment e fails, the spare capacity on the other segments must be sufficient for the flow on the protection routes of every interrupted path. Constraint sets (5.4), (5.5), and (5.6) are unchanged applicable for this path-protection scheme.

5.4 Computational Results

The metaheuristic was implemented in C and compiled with Microsoft Visual Studio.NET. All test runs in this chapter were performed on a Pentium 4 machine with one processor at 2.53 Ghz and 512 Mbytes of RAM.

5.4.1 Shared-Link Protection Scheme

The first experiment presented in this section consists of solving the MIP formulation developed for the link-protection scheme with Cplex starting with the solutions reported in chapter 4, where the set of instances are both real and randomly generated. Table 5.1 summarizes the computational results obtained by solving the MIP formulation with Cplex and the provisioning and routing problem with the hybrid metaheuristic approach described in chapter 3. This metaheuristic solves the routing and provisioning problem without protection and uncertainty in key data. The problem of provisioning the network to protect failed segments is also a provisioning and routing problem, where it is allowed sharing the capacity required to route the traffic from the origin to the destination of different failed segments. To solve the protection problem we run a modified version of this metaheuristic to allow sharing resources on every working network in the reference set.

Columns 2, 3, and 4, contain the number of nodes, number of segments, and number of demands in each instance. Next two columns report the total cost and CPU time (in seconds) for obtaining the working network. Columns 7 and 8 report the total cost and the CPU time (in seconds) for obtaining the spare capacity required to protect the working network in the event of any single link failure. Columns 9 and 10 show the total cost of the survivable network and the total CPU time required to obtain it using the model described above. Columns 11 and 12 summarizes the total cost of the working network and of the survivable network obtained using the proposed metaheuristic. Under the heading *Cplex Bound* we have reported the lower bound given by Cplex starting with the working solution provided by the metaheuristic. Last two columns show the total cost of the survivable network and the total CPU time obtained.

As indicated in chapters 3 and 4, when the size of the problems increases, the solver Cplex is not able to reach the optimal working network when the search is stopped after two hours of execution. Therefore, the experiments are performed on a set of problems solved to optimality.

The computational results reported in Table 5.1 show the effectiveness of using not only the best working network found, but also a set of good working networks for obtaining the backup routes.

Table 5.2 shows the total costs of designing survivable networks using the shared-link protection scheme for the instances in which Cplex was not able to reach the working network in less than two hours of execution. Columns 5 and 6 summarize the costs for obtaining the working network and the backup network, respectively. Under the heading *Cplex Bound* we report the total cost obtained by solving the link-protection MIP model starting with the working solution provided by the metaheuristic. Columns 8 and 9 show the total cost and the CPU time in seconds for solving both the working and protection problems with the metaheuristic procedure.

The solutions reported in Table 5.2 do not necessary correspond to the best working network reached after executing the metaheuristic developed in chapter 3. In most cases, the best survivable network was found using not only the best solution in the reference set, but a solution with a higher working cost.

5.4.2 Link-Disjoint and Non Link-Disjoint Alternative Routes

When a shared-path protection scheme with non link-disjoint routes is used, after obtaining the working network, alternative routes have to be generated for each demand requirement and each segment in its working path. The protection routes for each demand $(o, d) \in D$ and segment e in its working path cannot contain this segment. Hence, if segment e fails, demand (o, d) can be rerouted between its end nodes on any of the alternative routes. If segment e fails, the demands whose working paths contain segment e cannot share capacity in the final design.

Since the real and artificial networks that we are using to get the computational results are not necessary two-connected, there might be demands that cannot be protected using this scheme. If the path-protection scheme with link-disjoint routes is used, there can also be demands that cannot be rerouted. Since obtaining link-disjoint routes is a more restrictive condition than obtaining routes without a failed segment, using the path-protection scheme with link-disjoint routes implies that a higher number of demands may not be protected.

A modified version of the hybrid metaheuristic developed in chapter 3 is used for carrying out the experiments. As explained in chapter 3, once each demand has been assigned to a path in its list of potential paths, the evaluation of the solution consists of calculating the increase of capacity required in the elements of the network that route the demands through the assigned paths. The increased capacity is then translated into cost of installing additional fiber and adding WDMs and OXCs. In order to achieve a method for solving the protection problem using a shared-path protection scheme, the metaheuristic has to be modified by changing the evaluation of the solution. For the purpose of modifying the hybrid metaheuristic, we have to take into account that the demands whose working paths have any common link cannot share the spare capacity if any of those links fails.

Table 5.3 summarizes the total costs in million dollars of generating survivable networks using both a shared-path protection scheme with non link-disjoint routes and a shared-path protection scheme with link-disjoint routes. Under the heading *Non Link-Disjoint Path-Protection* we report the results corresponding to this protection scheme. Columns 5, 6, 7, and 8 summarize the working cost, protection cost, total cost, and CPU time in seconds, respectively. Column 9 shows the unprotected demand and the segment that causes this no-protection. Under the heading *Link-Disjoint Path-Protection* we report the results of the link-disjoint path-protection.

Comparing the results in Tables 5.1 and 5.3 we realize that when using path-protection schemes the spare capacity required to protect the network in the event of any single link failure is smaller than when using a shared-link protection scheme.

As explained above, the hybrid metaheuristic provides a reference set of working networks, which are then the initial solutions for solving the protection problem. Our experiments show that in most cases, the best survivable network is reached using not the best working network in the reference set, but a working network with higher working cost.

The results reported in Table 5.3 show that the link-disjoint shared-path protection scheme reaches network designs better than the other path-protection scheme. Furthermore, since the scheme with non link-disjoint routes has a higher number of demands, the CPU time required to solve the problem is higher.

5.5 Conclusions

This chapter focuses on the protection problem in WDM mesh networks. For the purpose of solving this problem we use three protection schemes: shared-link protection scheme and two variants of the shared-path protection scheme.

In the literature, the works related to planning WDM Networks deal with the protection problem after solving the provisioning and routing problem without protection when considering shared-protection schemes. We propose a hybrid metaheuristic that solves the protection problem using as starting solutions a reference set of "good" alternative working networks.

Our experiments corroborate that using the shared-path protection scheme with link-disjoint protection routes is better alternative than using any of the other two protection schemes presented in this chapter. However, the choice of one of the protection schemes presented depends on the decision maker.

				Cplex						MetaH				
				Work Paths		Backup Paths		Survivable Network						
Set	$\|N\|$	$\|E\|$	$\|D\|$	Cost	Time	Cost	Time	Cost	Time	Work Cost	Backup Cost	Cplex Bound	Total Cost	Time
MD	11	16	10	4.09	0.22	5.64	0.35	9.73	0.57	4.14	3.88	3.88	8.02	4.47
			20	4.38	0.21	6.23	0.22	10.61	0.43	4.48	4.93	4.93	9.41	6.03
			30	8.42	0.18	7.01	0.35	15.43	0.53	8.46	6.91	6.91	15.37	7.84
			54	14.03	1.25	14.05	0.35	28.08	1.6	14.25	12.63	12.49	26.88	14.07
		27	10	2.75	0.13	3.81	0.18	6.56	0.31	2.76	3.25	3.25	6.01	4.14
			20	4.26	0.29	4.32	0.56	8.58	0.85	4.30	4.23	4.23	8.53	8.25
			30	6.46	0.27	5.66	1.39	12.12	1.66	6.52	5.47	5.42	11.99	11.54
			54	11.16	2.87	10.23	1.29	21.39	4.16	11.26	10.12	10.12	21.38	17.80
		42	10	1.96	0.11	2.20	0.18	4.16	0.29	2.01	1.69	1.69	3.70	4.16
			20	3.08	0.13	2.65	0.23	5.73	0.36	3.54	2.81	2.52	6.35	12.62
			30	5.38	0.19	3.12	1.78	8.50	1.97	5.71	5.06	4.83	10.77	12.74
			48	7.08	0.51	4.84	5.21	11.92	5.72	7.91	5.80	5.48	13.71	21.41
			54	8.41	0.41	5.87	7.94	14.28	8.35	9.22	6.96	6.60	16.18	24.20
E0D	12	17	15	3.69	0.15	6.79	0.21	10.48	0.36	5.09	4.82	4.82	9.90	4.95
			19	6.26	0.71	4.85	0.27	11.11	0.98	7.03	3.71	3.71	10.74	5.00
			21	6.21	0.63	7.15	0.20	13.36	0.83	9.35	2.15	2.15	11.50	6.24
			44	14.36	10.70	16.55	0.63	30.91	11.33	15.99	11.51	11.51	27.50	15.11
			66	11.83	7.38	14.87	0.59	26.70	7.97	14.05	6.12	6.12	20.17	16.19
		33	15	3.69	1.41	3.82	0.24	7.51	1.65	4.90	2.06	2.06	6.96	2.61
			21	7.32	10.94	5.99	0.25	13.31	11.19	8.77	4.29	4.29	13.06	4.77
			44	13.94	81.16	16.48	9.43	30.42	90.59	17.05	9.90	9.41	26.95	12.96
			66	11.83	225.32	13.85	6.56	25.68	231.88	14.60	6.58	6.52	21.18	16.67
		46	15	3.69	0.83	3.82	0.24	7.51	1.07	4.17	2.92	2.92	7.09	4.01
			21	7.33	25.80	6.12	0.50	13.45	26.30	9.07	4.27	4.27	13.34	7.21
			44	13.95	467.72	16.56	4.25	30.51	471.97	16.91	7.95	7.95	24.86	14.52
			66	11.77	224.90	13.08	2.58	24.85	227.48	14.82	7.07	7.04	21.89	14.07
Ex2D	17	26	27	23.22	5.99	44.49	0.22	67.71	6.21	23.59	30.61	30.61	54.43	41.43
			36	81.84	13.42	153.99	1.44	235.83	14.86	81.84	153.35	152.93	235.19	52.80
			81	96.81	296.16	169.11	2.06	265.92	298.22	97.66	152.54	152.54	250.20	129.97
		68	27	24.43	12.26	33.37	8.10	57.80	20.36	27.65	27.20	26.33	54.67	44.46
			36	67.79	292.93	69.18	7.95	136.97	300.88	69.23	68.76	68.76	137.99	51.58

				Working Paths	Backup Paths		Survivable Network	
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	Working Cost	Backup Cost	Cplex Bound	Total Cost	Time
Ex2D	17	26	79	178.76	272.48	272.48	451.24	148.40
			135	176.22	236.17	235.20	412.39	248.95
		68	81	83.46	76.52	75.93	159.98	143.40
			135	143.07	129.27	126.71	272.345	275.00
National	50	63	45	40.19	28.92	28.02	69.11	148.07
			65	52.99	35.18	34.91	88.17	260.70
			91	60.05	47.22	46.90	107.27	352.40
			112	45.33	24.71	23.73	70.04	353.92

Table 5.2: Survivable Networks obtained with the MetaH for bigger instances

				Non Link-Disjoint Path-Protection					Link-Disjoint Path-Protection				
Set	$\lvert N\rvert$	$\lvert E\rvert$	$\lvert D\rvert$	Work Cost	Backup Cost	Total Cost	Time	Unmet Demand :S	Work Cost	Backup Cost	Total Cost	Time	Unmet Demand :S
MetroD	11	16	10	4.14	2.25	6.39	54.75	12:2	4.09	1.99	6.08	10.85	12
			20	4.43	2.28	6.71	169.40	0	4.38	2.17	6.55	116.26	0
			30	8.49	3.46	11.95	271.44	37:2	8.46	2.06	10.52	65.40	37
			54	14.12	3.90	18.02	2313.51	31:2	14.20	3.10	17.31	799.54	31
		27	10	2.86	1.86	4.72	57.32	0	2.80	1.85	4.65	29.66	0
			20	4.30	2.79	7.09	119.62	0	4.30	2.56	6.86	79.30	0
			30	6.57	3.41	9.98	220.93	0	6.52	3.38	9.91	198.89	0
			54	11.30	4.95	16.25	1192.56	0	11.26	5.06	16.33	747.00	0
		42	10	2.01	1.53	5.54	48.53	0	2.01	1.33	3.34	31.52	0
			20	3.62	2.06	5.68	139.52	0	3.62	2.01	5.63	96.92	0
			30	5.93	3.01	8.94	219.39	0	5.71	3.20	8.91	219.30	0
			48	7.90	5.55	13.45	572.75	0	7.91	5.24	13.16	523.38	0
			54	9.18	5.66	14.84	846.78	0	9.17	5.69	14.87	698.82	0
Extant0D	12	17	15	5.95	3.38	9.33	133.33	0	4.17	5.46	9.63	69.14	0
			19	7.51	3.01	10.52	117.53	0	6.26	4.17	10.43	75.35	0
			21	7.90	3.60	11.50	453.17	0	8.88	1.93	10.81	110.99	0
			44	16.06	4.74	20.80	4527.52	0	16.38	3.61	19.99	1057.49	0
			66	14.62	0.51	15.13	7409.31	0	14.62	0.54	15.16	2242.58	0
		33	15	4.17	5.57	9.74	134.43	0	4.17	4.71	8.88	55.44	0
			21	10.58	5.48	16.06	440.91	0	8.80	5.58	14.38	104.09	0
			44	16.91	12.28	29.19	2451.10	0	18.22	8.57	26.79	558.64	0
			66	17.10	6.35	23.45	4053.49	0	17.81	4.61	22.42	1331.77	0
		46	15	5.11	4.34	9.45	105.32	0	4.56	3.94	8.50	48.23	0
			21	11.72	5.40	17.12	275.93	0	10.31	4.39	14.79	93.98	0
			44	19.64	12.02	31.66	1283.07	0	18.81	9.73	28.54	495.95	0
			66	14.29	10.70	24.99	3962.38	0	16.41	10.24	26.65	1058.45	0
Ex2D	17	26	27	23.59	18.43	42.02	2403.31	0	27.05	13.66	40.71	355.14	0
			36	8.37	7.36	15.73	4981.29	0	8.37	5.69	14.06	737.25	0

Table 5.3: Non Link-Disjoint - Link-Disjoint Solutions

Chapter 6

Provisioning of WDM Mesh Networks Under Demand Uncertainty

An important consideration for additional research in this area deals with tackling uncertainty. Clearly, the demands cannot be considered known in an environment such as the telecommunications industry. The availability of a MIP formulation that can be used to find near-optimal solutions to the capacity expansion problem represents a stepping-stone toward the solution of a stochastic version of the problem that treats the demands as uncertain.

The aim of this chapter is to provide mathematical models for solving the provisioning and routing problem under demand uncertainty. Section 6.1 describes the usual techniques to deal with optimization problems with uncertainty. In sections 6.2 and 6.3 we develop a stochastic programming model and a scatter search solution approach for the problem, respectively. Section 6.4 provides a set of instances with the aim of comparing the scatter search solution and the solution reached by solving the model with Cplex. Section 6.5 develops an alternative stochastic model. The last section describes a robust approach for solving the problem.

6.1 Introduction

The problem that we address in this chapter is a real world problem that results from the need to expand capacity of telecommunication networks built with fiber optics technology in the presence of uncertain information. When deterministic information is considered, given a network physical topology and the estimate of the point-to-point demand traffic, the problem is to determine the routing for each demand and the least-cost WDM and OXC equipment configuration required to support the routes as shown in chapters 3 and 4. This problem is modelled as a mixed integer problem (MIP).

Algorithms and model formulations have usually assumed that the data for the given problems are known accurately. However, this is not true in most real applications due mainly to measurement and errors. This is particularly problematic for data representing future traffic demands and product costs that cannot be known with certainty. There are several ways to take into consideration the uncertainty when searching for optimal decisions. The usual techniques to deal with optimization problems with uncertainty in the data are sensitivity analysis, fuzzy optimization, stochastic programming and robust optimization.

One of the oldest techniques to model uncertainty in an optimization problem is sensitivity analysis, which studies the way the optimal solution changes after a slight modification in the data. Since the beginning of linear programming, a half century ago, sensitivity analysis has been a part of the field of post-optimality in the theory, implementations and applications [10], [30]. As much as possible, these foundations have been extended to nonlinear, integer, stochastic, multicriteria, and other mathematical programming, though it is considered that those advances have so far not provided as rich a body of knowledge [138].

Fuzzy optimization uses the fuzzy set technology [22], [37], [141] and techniques to deal with uncertainty. A fuzzy optimization problem is an optimization problem where some of its components are uncertain and given by a membership function. The usual methods in fuzzy linear programming consider mainly fuzziness in the cost function, in the coefficients and in the inequalities of the constraints that can be verified in several degrees. When uncertain data are given by optimistic, pessimistic and intermediate values, triangular fuzzy numbers are a logical manner to deal with the problem.

Stochastic Programming is a framework for modelling optimization problems with uncertainty where probability distributions of data are known or can be estimated. The goal is usually to find a solution that is feasible for all (or almost all) the possible data instances and maximize the expectation of some function of the decision and random variables. A basic idea in most of methods in Stochastic Programming is the concept of recourse, that is the ability to take corrective actions after a random event has taken place. The simplest example is the two-stage problem, in which some decision variables are fixed before some events occur and other decisions are taken after the events. The decision maker takes some actions in the first stage, after which a random event occurs affecting the outcome of the first stage decisions. A recourse decision can then be made in the second stage that compensates for any bad effects that might have been experienced as a result of the first stage decision. The optimal decision for such a model consists of a single decision for the first stage and a collection of recourse decisions defining which second-stage decision should be made in response to each random outcome. These problems are extended to multi-stage problems [12], which are closely related to multi-stage decision analysis, Markov decision process, stochastic control theory and dynamic programming. Solution approaches to stochastic programming usually consist of obtaining a deterministic equivalent optimization problem that is solved by known techniques. These problems are typically very large scale problems and of a different type. Several survey articles [13] and [126], and books [68], [14], [113] are dedicated to stochastic programming. The bibliography by van der Vlerk includes more than 3500 references [132].

When using stochastic programming, it is necessary to ask the decision maker information on the probability value with which the instances or future scenarios might be realized. Assigning probabilities to different scenarios is not a trivial exercise for many decision makers.

Other techniques to deal with optimization problems that involve uncertainty are scenario-based, in which the scenarios correspond to the possible realizations of the uncertainty [121]. In practice, the uncertainty ranges from a few scenarios up to a precise joint probability distribution of all the random variables or data involved in the problem, where each possible scenario has its corresponding probability of occurrence. Uncertainty in key data is usually characterized by a probability distribution. The use of scenarios as a tool for modelling uncertainty has the advantage of not requiring knowledge of the underlying probability distributions associated with the random variables.

Robust optimization [107] belongs to the family of scenario-based optimization techniques. The main feature of robust optimization (RO) formulations is the flexibility to define the tradeoff between *solution robustness* and *model robustness*. A feasible solution to the problem is termed *robust solution* if it remains "close" to optimal for any realization of the scenario. The solution is also robust with respect to feasibility if it remains "almost" feasible for any realization of the scenario. Model robustness can be measured, for example, by the expected value of the infeasibility. Measuring the deviation of the proposed solution to the scenario-optimum gives an idea of its solution robustness.

Kouvelis and Yu [76] suggest a formal approach to decision making, which they refer to as the *robustness approach*. This approach produces decisions that will have a reasonable objective value under any likely input data scenario to the decision model. This approach identifies a set of possible scenarios without attempting to assign probabilities to scenarios, and then seeks to find the decision that performs well even in the worst case for the identified set of scenarios. The robustness concept proposed by these authors is a solution robustness concept in the Mulvey, et al. terminology. The robust optimization framework in [76] applies a *minmax regret criterion* to differentiate the performance of the various solutions over the given set of possible scenarios and it is mostly developed for models with discrete decision variables using state of the art combinatorial optimization techniques.

6.2 Stochastic Programming Model

We use the general modelling framework of stochastic programming for developing a model that will be used to solve the problem with uncertainty. The set of scenarios for a problem having s^* scenarios is denoted $S = 1, ..., s^*$. The model uses the following definitions:

- z_{ods} = under provisioning for each scenario $s \in S$ and origin/destination demand $(o, d) \in D_s$. The under provisioning value z_{ods} is the amount of demand *ods* that cannot be routed using the capacity currently installed in the network.
- P_s = probability of scenario s.
- Cu = under provisioning cost.
- $x_p^{ods} = 1$ if demand $(o, d) \in D_s$ is routed on path p, and 0 otherwise.

The amount of demand in each scenario that cannot be served with the installed capacity gives the under provisioning. The penalty cost of the unrouted demand measures the model robustness of the design.

For each *under provisioning cost* Cu, the objective to be minimized may be stated as the sum of the design cost (three first terms) and a penalty cost (fourth term) as follows:

$$\min \sum_{e\in E} 2C_e^F f_e + \sum_{e\in E} \left(\left(C_e^F + C_e^W\right) w_e + C^c v_e\right) + \sum_{n\in N} \left(C^O y_n + C^p u_n\right) +$$

$$Cu \left(\sum_{s\in S} P_s \sum_{od\in D_s} z_{ods} \right) \tag{6.1}$$

$$\sum_{p\in J_{od}^s} x_p^{ods} R_{od}^s + z_{ods} = R_{od}^s, \forall (o,d) \in D_s, \forall s \in S \tag{6.2}$$

$$\sum_{(o,d)\in D_s} \sum_{p\in J_{od}^s, e\in p} x_p^{ods} R_{od}^s \leq v_e + f_e, \forall e \in E, \forall s \in S \tag{6.3}$$

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{6.4}$$

$$\sum_{e\in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{6.5}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{6.6}$$

$$0 \leq z_{ods} \leq R_{od}^s, \forall (o,d) \in D_s, \forall s \in S \tag{6.7}$$

Note that since the cost of the design is being minimized in the objective function we do not use a variable to represent over provisioning.

The demand constraints can be modelled as shown in constraint set (6.2). Constraint set (6.3) converts path flows to segment flows and segment capacity into fibers and channels for each scenario. Constraint sets (6.4), (6.5) and (6.6) convert segment capacity to WDM units, accumulate channels on links to add the required number of ports to each node and convert node capacity to OXC units, respectively. These constraints are not affected by the scenarios. The last set of constraints (6.7) bounds the demands.

The stochastic programming model of minimizing (6.1) subject to (6.2) to (6.7) is only one of several models that can be used to design a network when the demand forecast is uncertain. Other possibilities include the *robust approach*, which will be presented in section 6.6.

6.3 Scatter Search Solution Approach for Stochastic Programming

This section summarizes a scatter search (SS) approach for solving the wavelength division multiplexing and optical cross-connect provisioning and routing problem under demand uncertainties.

The six procedures involved in the Scatter Search (see chapter 2) are the following:

1. *The initial population generation method,* that generates the initial population $InitPop$.

2. *The reference set generation method,* which selects the set $RefSet$ that consists of the "best" solutions in the population $InitPop$.

3. *The subset generation method,* which chooses a subset $SubSet$ that consists of r solutions in the reference set to apply the next combination procedure.

4. *The solution combination method,* which is a procedure that combine the solutions in $SubSet$ to get the current solution $CurSol$.

5. *The improvement solution method.* It is the procedure to improve the current solution $CurSol$ to get a better solution $ImpSol$.

6. *The reference set updating method.* It is the procedure to decide when and how to update the reference set taking into account the state of the search.

The initial population generation method generates solutions using the metaheuristic procedure developed in chapter 3 for the basic provisioning and routing problem.

For the provisioning and routing problem with uncertain demands every solution in the population is obtained by selecting a random number of demands for each scenario and routing them through the existing network design using the metaheuristic procedure described in chapter 3.

Once the design variables have been fixed, a multicommodity maximum flow problem must be executed for each scenario in order to achieve the amount of demand that cannot be carried using the capacity installed in the optical network. The resulting problem, that provides the under provisioning cost, is then stated as follows:

$$\min \sum_{od \in D_s} z_{ods} \tag{6.8}$$

Subject to:

$$\sum_{p \in J^s_{od}} R^s_{od} x^{ods}_p + z_{ods} = R^s_{od}, \forall (o,d) \in D_s \tag{6.9}$$

$$\sum_{p \in J^s_{od}, e \in p} R^s_{od} x^{ods}_p \leq v_e + f_e, \forall e \in E \tag{6.10}$$

The set of constraints labelled as (6.9) are demand constraints for a given scenario. Constraint set (6.10) converts path flows to segment flows and segment capacity into fibers and channels for a given scenario. The right hand coefficients in constraint set (6.10) are the sum of two design variables, which represent the number of WDM channels and fiber pairs on segment e. Therefore, they are constants since the design variables are fixed. We use Cplex 8.0 to solve this problem [28].

The notion of best solutions used in the *reference set generation method* is not limited to a measure given exclusively by the evaluation of the objective function. In particular, a solution may be added to the reference set if the diversity of the set improves even when

Combination Method

{

$maxnumchan(e) = \max_{s \in SubSet}\{numchan(e), e \in s\}$;

$minnumchan(e) = \min_{s \in SubSet}\{numchan(e), e \in s\}$;

if (InstallationCost($maxnumchan(e)$) ¡ PenaltyCost($maxnumchan(e)$))

install $maxnumchan(e)$ channels on segment e;

else if (InstallationCost($minnumchan(e)$) ¡ PenaltyCost($minnumchan(e)$))

install $minnumchan(e)$ channels on segment e;

else

$v_e = 0$; $f_e = 0$;

}

Figure 6.1: Combination method

the objective value of the solution is inferior to other solutions competing for admission into the reference set. The reference set that we generate consists of 10 solutions.

The combination parameters of the *solution combination method* are used to modulate the intensification and/or diversification of the search. For every segment $e \in E$, we obtain the maximum number of channels installed in the networks in $SubSet$. Then, if the installation cost of this number of channels is smaller than the penalty cost of not installing that capacity, the maximum number of channels is installed in the combined solution. Otherwise, we obtain the minimum number of channels installed in the networks in $SubSet$. Then, if the installation cost is smaller than the penalty cost, we add the minimum number of channels to the combined solution. If the decision is not to install any capacity on that segment, then the design variables v_e and f_e are equal to zero. This is shown in Figure 6.1.

The *improvement solution method* applied to the current solution is based on changing one demand from its current path to another. We execute a local search procedure based on these moves to achieve the improved solution. The set of possible paths for each demand ranges from 0 up to $kmax$, which indicates the maximum number of paths calculated for each demand. If the chosen path for a demand is equal to 0, then we do not install any capacity to route that demand and, therefore, the penalty term increases. The neighborhood search examines moves employing a given ordering of the demands. That is, the first candidate move is to reassign the demand that is at the top of the list. When reassigning this demand the design variables change, but the model that has to be solved in order to calculate the penalty term for a given scenario does not change except for the right hand coefficient in constraint set (6.10). Therefore, we solve the modified model. To do this, we do not have to start a new model from scratch, but instead we can take the existing model and change it to our needs. This is done by calling the Cplex modification methods.

If the new solution is better than the worst in the current reference set, then the reference set is updated. If an improving move that involves reassigning the first demand in the list cannot be found, then the second demand is considered. The process continues until a demand is found for which a reassignment of paths leads to an improving move. If all the demands are examined and no improving move is found, the local search is abandoned.

6.3.1 Illustrative Example

To illustrate the practical application of the stochastic programming methodology for the routing and provisioning problem under demand uncertainties, an illustrative example has been solved, for several values of the penalty parameter, using the scatter search procedure and Cplex. The metaheuristic was implemented in C and compiled with Microsoft Visual Studio.NET. All test runs in this chapter were performed on a Pentium 4 machine with one processor at 2.53 Ghz and 512 Mbytes of RAM. Table 6.1 summa-

Name	Extant0D
Total Nodes	12
Total Links	17
Total Demand Pairs	19
Number of Paths/Demand	5
Total Demand Scenarios	3

Table 6.1: Characteristics of Test Problem

rizes the characteristics of the test problem Extant0D and Figure 6.2 shows the network corresponding to this instance.

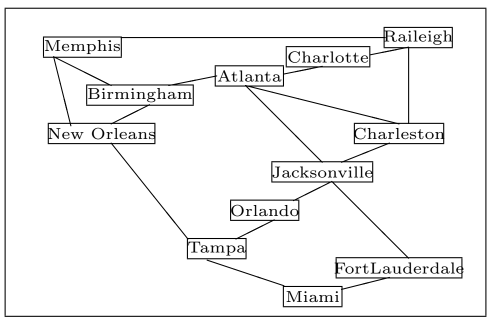
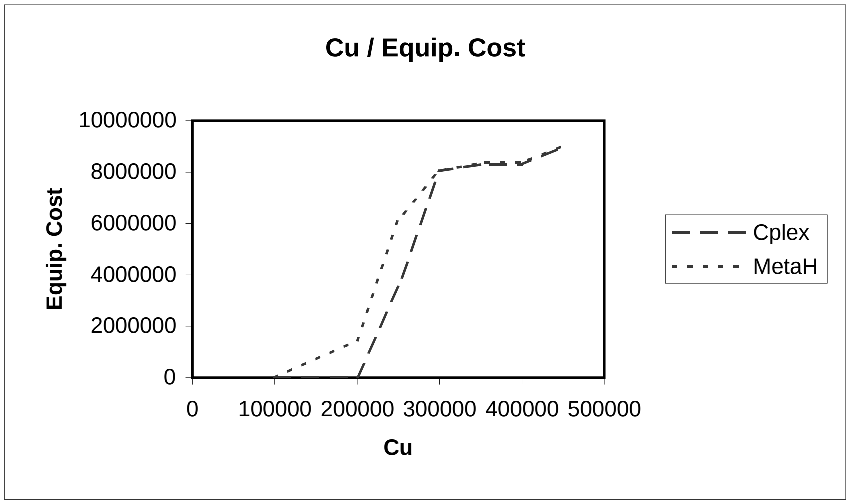


Figure 6.2: Network for Extant0D Problem

Note that if the penalty cost Cu is equal to zero, then the cost of not routing the set of demands for each scenario is equal to zero. In addition, the decision is not to increase the capacity of the network and the design cost is also equal to zero. When increasing the penalty cost Cu the design cost increases while the unmet demand decreases as is shown in Figure 6.3.

The first picture in Figure 6.3 compares the equipment costs provided by solving the problem with the scatter search and with Cplex for several values of the under provisioning cost Cu. For some values of Cu, the equipment costs provided by the metaheuristic are higher than the costs provided by Cplex. However, in those cases, the amount of unmet demand is smaller for the metaheuristic. When the under provisioning



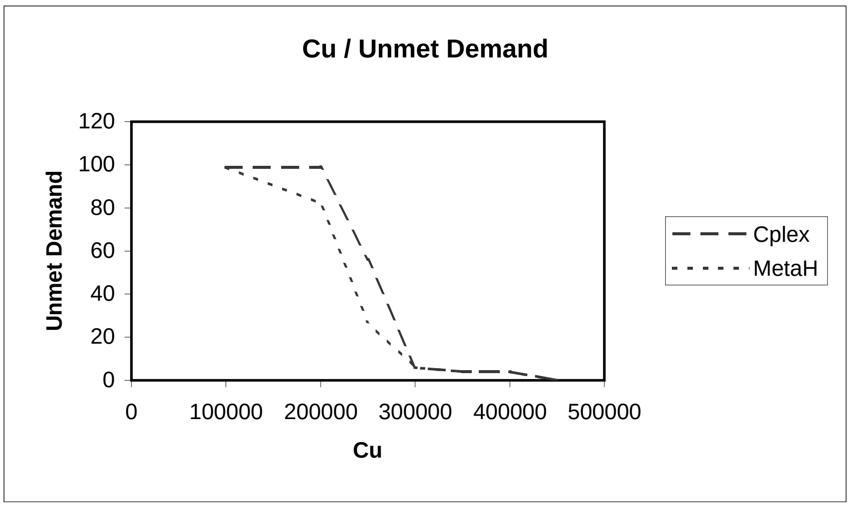


Figure 6.3: Results for the Extant0D Problem.

cost is equal to 450000, both methods reach a network design that is feasible for every scenario, and therefore the amount of unmet demand is equal to zero. As shown in Figure 6.3, for this value of Cu the equipment cost given by Cplex is slightly smaller than the one given by the metaheuristic.

The purpose of developing the scatter search metaheuristic is to allow the use of both linear and non linear penalty functions.

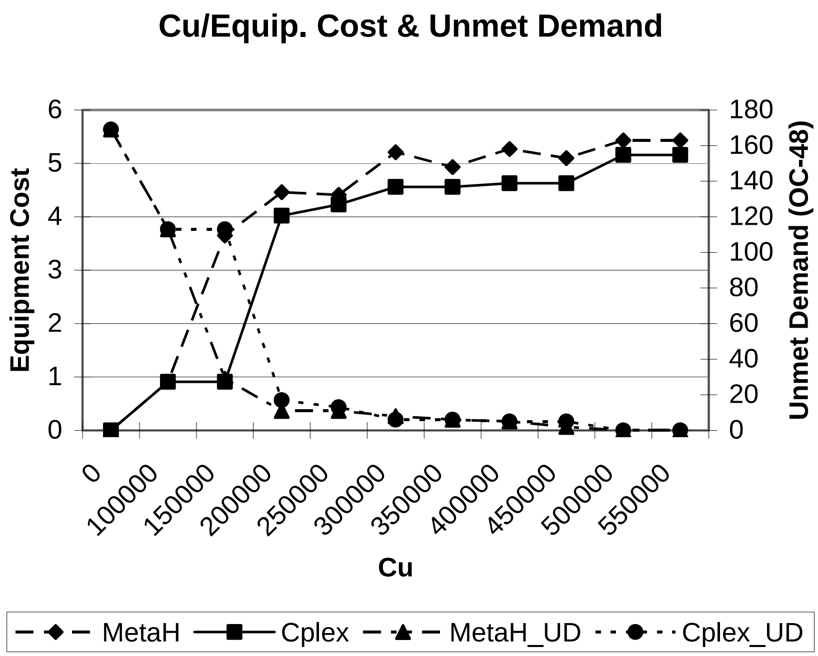


Figure 6.4: Instance consisting of 11 nodes, 16 segments and 10 demands

6.4 Computational Results

This section presents the computational results obtained solving some both artificial and real instances with the metaheuristic developed above and solving the stochastic programming model with Cplex. Figures 6.4, 6.5, and 6.6 show pictures in which the line labelled *MetaH* refers to the network equipment cost obtained for several under provisioning costs, and the line labelled *Cplex* refers to the equipment cost obtained by solving the stochastic program with Cplex. Then, the lines labelled *MetaH_UD* and *Cplex_UD* show the amount of unmet demand for the metaheuristic and Cplex, respectively. Figures 6.4 and 6.5 correspond to instances randomly generated, while Figure 6.6 corresponds to a real instance, which consists of 11 nodes, 42 segments and 48 demands.

Table 6.2 summarizes the equipment costs, the amount of unmet demand, and the CPU time in seconds for both the metaheuristic and Cplex, which was stopped after two hours of execution.

The computational results obtained in this section corroborate the effectiveness of the

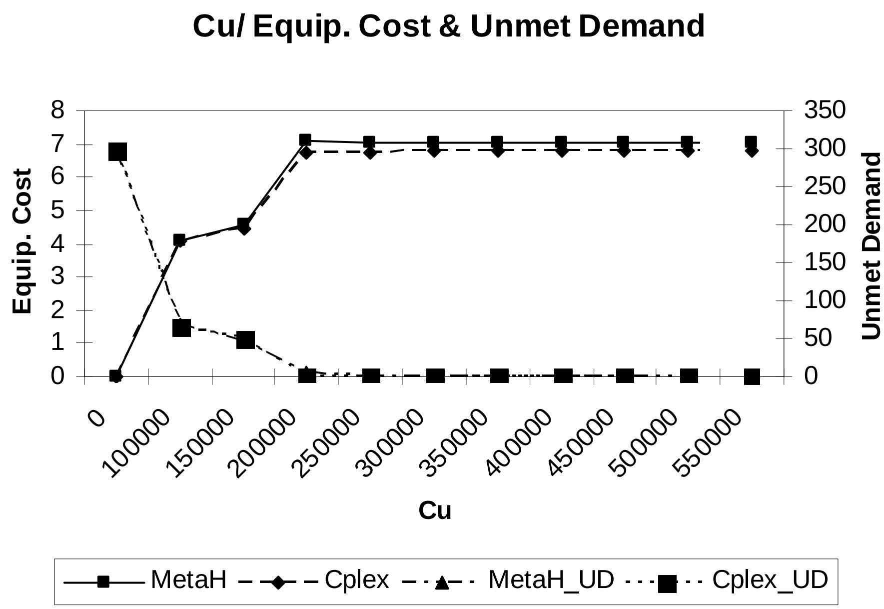


Figure 6.5: Instance consisting of 11 nodes, 16 segments and 20 demands

proposed scatter search metaheuristic to provide high quality solutions in a reasonable amount of time. Cplex requires a large amount of time to reach solutions comparable to those obtained by the metaheuristic procedure.

6.5 Alternative Stochastic Programming Model

The stochastic programming model previously developed in this chapter uses the variable z_{ods} to represent the under provisioning for each scenario $s \in S$ and origin/destination demand $(o, d, s) \in D_S$. The under provisioning value z_{ods} is the amount of demand *ods* that cannot be routed using the capacity currently installed in the network. Then, a penalty function of these variables is minimized. Since the total equipment cost of the design is being minimized in the objective function, the model does not use a variable to represent over provisioning. However, in order to penalize even more the over provisioning in the final network design, we add in the model the following variable:

- z^+_{es} = the over provisioning for each segment $e \in E$ under each scenario $s \in S$.

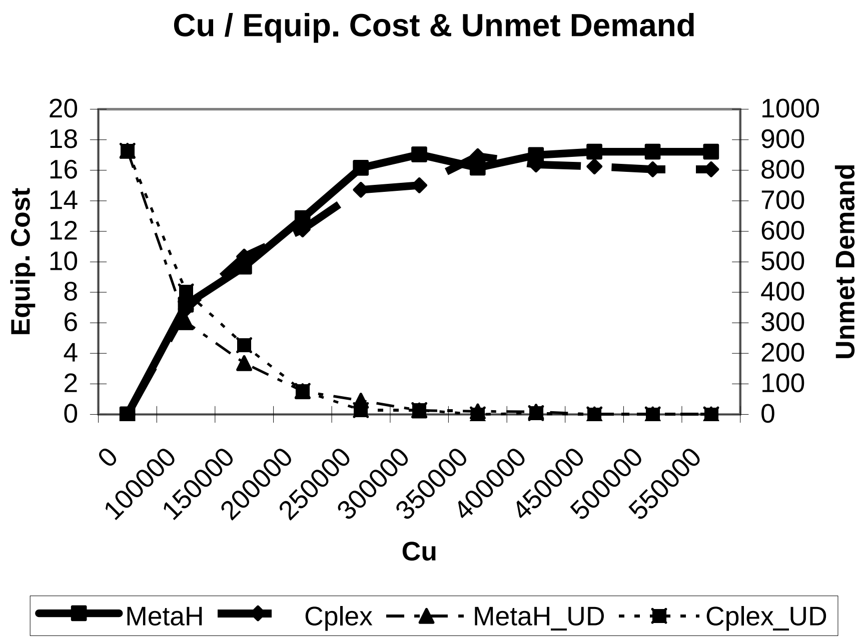


Figure 6.6: Real instance consisting of 11 nodes, 42 segments and 48 demands

For each *under provisioning cost Cu* and *over provisioning cost Co* the objective to be minimized may be stated as the sum of the design cost and a penalty cost (last two terms) as follows:

$$\min \sum_{e\in E} 2C_e^F f_e + \sum_{e\in E}\left(\left(C_e^F + C_e^W\right) w_e + C^c v_e\right) + \sum_{n\in N}\left(C^O y_n + C^p u_n\right) +$$

$$Cu\left(\sum_{s\in S} P_s \sum_{od\in D_s} z_{ods}\right) + Co\left(\sum_{s\in S} P_s \sum_{e\in E} z_{es}^+\right) \tag{6.11}$$

Subject to:

$$\sum_{p\in J_{od}^s} x_p^{ods} R_{od}^s + z_{ods} = R_{od}^s, \forall (o,d)\in D_s, s\in S \tag{6.12}$$

$$\sum_{p\in J_{od}^s, e\in p} x_p^{ods} R_{od}^s + z_{es}^+ = v_e + f_e, \forall e\in E, \forall s\in S \tag{6.13}$$

	MetaH			Cplex		
Cu	Equipment Cost	Unmet Demand	CPU Time	Equipment Cost	Unmet Demand	CPU Time
0	0	864	4.56	0	864	0.12
100000	7.19	302	1659.10	6.9	403	7209.47
150000	9.68	168	5564.06	10.34	227	7220.78
200000	12.86	76	4745.03	12.12	77	7234.49
250000	16,15	45	2577.12	14.72	14	7224.81
300000	17.04	13	1558.27	15.01	14	7257.85
350000	16.17	10	1435.64	16.92	0	7227.49
400000	16.99	9	2692.66	16.37	4	7269.93
450000	17.21	0	1642.70	16.25	0	7235.61
500000	17.21	0	4186.52	16.06	0	7248.57
550000	17.21	0	1274.86	16.06	0	7239.22

Table 6.2: Results Stochastic Programming Real Instance

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{6.14}$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{6.15}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{6.16}$$

$$0 \leq z_{ods} \leq R^s_{od}, \forall (o,d) \in D_s, s \in S \tag{6.17}$$

To test the effectiveness of this model we used the instance Extant0D presented in the previous section. First of all, we set a value for the under provisioning cost Cu and then Co ranges from $Co = 0$ up to $Co = 600000$. The cost Cu also ranges from 0 up to 600000.

Cu	Co						
	0	100000	150000	200000	250000	450000	600000
0	0	0	0	0	0	0	0
250000	0.22	0.22	0.22	0.22	0.22	0.22	0.22
300000	5.50	5.05	4.96	5.72	0.74	0.74	0.74
350000	7.56	7.10	6.85	5.72	5.72	5.76	5.76
400000	7.76	7.10	7.42	7.78	7.78	7.85	5.76
450000	7.97	8.24	7.42	8.66	8.66	8.73	7.79
500000	8.51	8.69	8.24	10.73	8.66	8.66	8.73
550000	8.51	8.88	10.73	10.73	10.73	10.73	8.73
600000	9.06	8.88	10.73	10.73	10.73	10.73	10.73

Table 6.3: Equipment Costs

The rationale behind adding the over provisioning variable, z^+_{es}, to the model is to minimize the maximum number of channels that are not used to route traffic when the possible scenarios are realized. If the value of the under provisioning cost Cu is fixed and the value of the over provisioning cost Co increases it is thought that the equipment cost of the final network design should decrease. However, we solved the model with Cplex using the illustrative example Extant0D and the results were not as expected. Intuitively, we expected that for a fixed value of Cu (under provisioning cost), the equipment cost was smaller when increasing the value of Co (over provisioning cost). Therefore, the lines in Figure 6.7 would not have to be intersected. The results are shown in Table 6.3 and in Figure 6.7. Therefore, the stochastic model proposed in this section is not a good alternative for solving the provisioning and routing problem under demand uncertainty.

6.6 Robustness Approach

The robustness approach proposed by Kouvelis and Yu [76] is a scenario-based technique, where an input data scenario represents a potential realization of the parameters of the

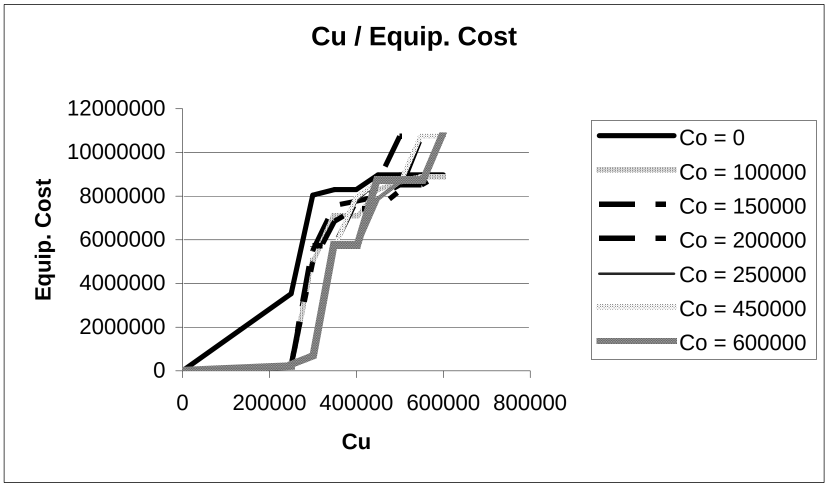


Figure 6.7: Cu / Equipment Cost

model.

Let S be the set of possible scenarios. Let X be the set of decision variables and D the set of input data. D^s denotes the instance of input data that corresponds to scenario s. Let F_s denote the set of all feasible decisions when scenario s is realized, and the quality of the decision $X \in F_s$ is evaluated with the function $f(X, D^S)$. Then, the optimal individual scenario decision X_s^* for the input data D^s is the solution to a deterministic optimization problem and it satisfies

$$z^s = f(X_s^*, D^s) = \min_{X \in F_s} f(X, D^s) \tag{6.18}$$

The *robust deviation decision* is defined as the one that provides the best worst case deviation from optimality, among all feasible decisions over all input data scenarios, i.e.,

$$z_D = \max_{s \in S}(f(X_D, D^s) - f(X_s^*, D^s)) = \min_{X \in \bigcap_{s \in S} F_s} \max_{s \in S}(f(X, D^s) - f(X_s^*, D^s)) \tag{6.19}$$

In this case, the robustness indicator of the decision is the worst observed deviation.

According to the above definition, finding the robust decisions implies to solve the following mathematical program:

$$z_D = \min\{y | f(X, D^s) \leq y + z^s, s \in S; X \in \bigcap_{s \in S} F_s\} \tag{6.20}$$

Then, this robust discrete optimization program consists of two sets of constraints:

Scenario Constraints, in which the objective function is restricted for each scenario to ensure good performance across scenarios.

Feasibility constraints across scenarios, which are the constraints of a single scenario problem.

6.6.1 Robust Deviation Decision

In this section we discuss the use of the above robustness definition to the provisioning and routing problem in WDM mesh networks.

$$\min y \tag{6.21}$$

Subject to:

$$\sum_{e \in E} 2C_e^F f_e + \sum_{e \in E} \left(\left(C_e^F + C_e^W \right) w_e + C^c v_e \right) +$$

$$\sum_{n \in N} \left(C^O y_n + C^p u_n \right) \leq y + z^s, \forall s \in S \tag{6.22}$$

$$\sum_{p \in J_{od}^s} x_p^{ods} = 1, \forall (o, d, s) \in D_s, s \in S \tag{6.23}$$

$$\sum_{p \in J_{od}^s, e \in p} x_p^{ods} R_{od}^s \leq v_e + f_e, \forall e \in E, \forall s \in S \tag{6.24}$$

$$v_e \leq M^W w_e + g_e, \forall e \in E \tag{6.25}$$

$$\sum_{e \in A_n} (v_e + f_e) \leq u_n, \forall n \in N \tag{6.26}$$

$$u_n \leq M^O y_n + h_n, \forall n \in N \tag{6.27}$$

The demand constraints can be modelled using the decision variables x_p^{ods} for the routing path of the demands in each scenario as shown in constraint set (6.23). Constraint set (6.24) converts path flows to segment flows and segment capacity into fibers and channels for each scenario. Constraint sets (6.25), (6.26) and (6.27) convert segment capacity to WDM units, accumulate channels on links to add the required number of ports to each node and convert node capacity to OXC units, respectively.

The robust discrete optimization program described in this chapter for the routing and provisioning problem under demand uncertainty differs from the stochastic program developed in the previous section in the following way. The main difference between the models presented in sections 6.2 and 6.6 is in the objective function. The stochastic programming model described in section 6.2 minimizes a balance, which is determined by the decision maker, between a design cost and a penalty factor. The objective function presented in section 6.6 attempts to find a solution that provides the best worst case deviation from optimality in each scenario. The best model depends on the goal of the decision maker and on the available information. Furthermore, if the decision maker is able to assign probabilities to the different input data scenarios, then it is possible to use the stochastic program to reach solutions that might be feasible only for a subset of all the possible scenarios. On the other hand, if the decision maker cannot obtain probabilities for each scenario, then both the stochastic programming model and the robust discrete optimization program can be developed.

6.6.2 Computational Results

In this section we provide the computational results obtained solving the robust discrete optimization program, for which five scenarios of demands are generated. To state the robust model, the optimal solutions have to be provided for each scenario. Hence, we report the results for those instances for which the scenario optimal solutions are reached in less than two hours of execution.

Tables 6.4 and 6.5 summarize the robust solutions obtained for real and artificial networks that consist of 11 and 12 nodes, respectively. The first four columns shows the set name, number of nodes, segments, and demands. Next five columns report the optimal solutions obtained for each scenario using the segment-path formulation proposed in Chapter 4. Column 10 provides the robust deviation value and column 11 the equipment cost of the network design reached by the model. The network designs obtained using this model are feasible for the five scenarios. Finally, the last two columns report either the CPU time in seconds or the optimality gap if the optimal solution was not reached after two hours of execution.

6.7 Conclusions

In this chapter the provisioning and routing problem under demand uncertainties is considered. We have provided several mathematical models for solving the problem. We develop a stochastic programming model and a scatter search solution approach for the problem. This metaheuristic is able to reach high quality solutions in a reasonable amount of time. Furthermore, we develop a robust deviation program that obtain feasible solutions even if the worst scenario is realized.

The results obtained in this chapter corroborate the effectiveness of using a scatter search metaheuristic to provide high quality solutions to problems that arise when planning WDM optical networks.

				Individual Scenarios					Robust Solution			
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	z^1	z^2	z^3	z^4	z^5	y	Equip. Cost	Time	gap (%)
MetroD	11	16	10	4.09	3.73	5.02	3.52	4.40	1.64	5.16	0.56	-
			20	4.38	4.97	5.31	3.91	6.52	2.87	6.78	1.40	-
			30	6.46	8.14	4.44	5.85	10.26	7.91	12.35	13.01	-
			54	14.03	16.36	9.80	15.55	16.70	9.10	18.90	883.16	-
		27	10	2.75	2.56	3.51	2.42	3.05	1.15	3.57	1.67	-
			20	4.26	4.74	5.22	3.81	6.10	2.83	6.65	164.57	-
			30	6.46	8.14	4.44	5.85	10.26	6.33	10.77	-	0.45
			54	11.15	13.37	7.37	12.84	13.32	10.28	17.65	-	27.94
		42	10	1.96	1.82	2.53	1.64	2.28	1.63	3.28	0.42	-
			20	3.08	3.87	3.93	3.00	4.79	3.85	6.85	221.14	-
			54	8.41	10.56	5.84	10.04	10.03	9.42	15.27	-	11.41

Table 6.4: Robust Approach with the 11-node networks

				Individual Scenarios					Robust Solution			
Set Name	$\|N\|$	$\|E\|$	$\|D\|$	z^1	z^2	z^3	z^4	z^5	y	Equip. Cost	Time	gap (%)
MetroD	12	17	15	3.69	4.20	7.03	6.68	6.20	3.62	7.32	1.86	–
			19	6.26	7.25	6.55	11.54	7.20	5.34	11.61	2.25	–
			21	6.20	7.24	7.11	8.17	11.75	5.54	11.75	3.19	–
			44	14.35	20.92	27.49	27.99	23.78	17.79	32.14	–	14.84
			66	11.83	15.83	29.32	20.97	30.09	19.70	31.54	709.26	–
		33	15	3.69	4.20	7.00	6.87	6.59	3.62	7.32	7.47	–
			21	7.32	8.30	8.45	9.82	12.06	4.74	12.06	14.96	–
		46	15	3.69	4.20	7.00	6.87	6.59	3.62	7.32	16.08	–
			21	7.32	8.30	8.45	9.78	12.16	4.83	12.16	572.45	–

Table 6.5: Robust Approach with the 12-node networks

Chapter 7

Contributions

This dissertation includes the results of several studies that have been either submitted to International Conferences or to International Journals both in the heuristics field and in the telecommunications field.

Chapter 2 is a comprehensive review of some basic metaheuristics. Some metaheuristics, such as multistart, take advantage of the global structure of the objective function, allowing to identify regions of the solution space in which a local from any of its solutions converges to a local optimum with high probability. These results have been reported in the article "A Multistart Clustering Technique for Combinatorial Optimization" submitted to the International Conference MS'2000 and then published in the series *The best of MS2000 International Conference on Modelling and Simulation*, (2000).

Chapter 2 also describes the possibility of taking advantage of parallelization methods, which allow either to increase the exploration of the solution space or to reduce the total computational time. These advantages have been proved when using both the Variable Neighborhood Search and the Scatter Search. The articles "The parallel variable Neighborhood search for the p-median problem" (2002), which has been published in the *Journal of Heuristics*, and "Parallelization of the Scatter Search for the p-median problem" (2003), that has been published in the journal *Parallel Computing* are the results of this research. Furthermore, a paper on tabu search with Fred Glover will appear in the International journal *Inteligencia Artificial*.

Chapter 3 introduces the provisioning and routing problem that arises when planning WDM optical networks. An hybrid metaheuristic procedure that combines ideas from scatter search, multistart, and tabu search has been proposed for efficiently solving this problem. This hybrid procedure has been compared with a permutation-based procedure based on the one proposed by other authors. The obtained results, that corroborate the effectiveness of developing a hybrid metaheuristic, are reported in the article "Minimizing the Cost of Placing and Sizing Wavelength Division Multiplexing and Optical Cross-Connect Equipment in a Telecommunications Network", which is on a second referee process in the journal *Networks*. A preliminary version of this work was presented at the *9th International Conference on Telecommunication Systems, Modelling and Analysis* in March 2001.

In chapter 4, a mathematical model with less variables and constraints than the one proposed by other authors is developed. The hybrid metaheuristic was compared with the solutions obtained solving the model with an existing optimizer and the results are reported in the article "Capacity Expansion of Fiber Optic Networks with WDM Systems: Problem Formulation and Comparative Analysis". A preliminary version of this article was presented at the International Telecommunications Conference *Symposium on Informatics and Telecommunications September* (SIT'02) and an improved version was presented as an invited conference at the INFORMS Annual Meeting 2002. The article has been accepted for its publication in the journal *Computers and Operations Research*.

Finally, the technical report "Provisioning of Survivable WDM Mesh Networks Under Demand Uncertainty and Single Link Failure Protection", which collects the results of chapters 5 and 6, will be presented in the *International Network Optimization Conference* (INOC2003), which will take place the coming October.

Chapter 8

Conclusions

Telecommunication Network Planning has become a fertile ground for developing and applying optimization techniques, while replacing the existing networks based on physical rings by optical meshes and evolving the optical technology. In this dissertation we have studied the metaheuristic optimization of a real world problem that arises in the telecommunications field: the Provisioning and Routing Problem in wavelength division multiplexing (WDM) mesh networks. This problem deals with the effective utilization of new technology designed to increase the capacity of an optical telecommunication network. The problem arises when the demand for bandwidth in a fiber optic network exceeds the current capacity. The new technology, wavelength division multiplexing, can expand the capacity of a fiber optic network without requiring additional fiber. The decision problem is to find the most cost-effective combination of WDM equipment and fiber that increases the capacity of the network to a point where all the expected demand can be handled.

When planning optical telecommunications networks, several optimization problems, which need to be solved in a reasonable amount of time, arise. Mathematical programming and existing solvers are able to provide optimal solutions for problems with small dimensions, but fail to optimize real instances. The metaheuristics overcome these difficulties. A Metaheuristic is a master strategy that guides and modifies other heuristics with the purpose of generating solutions beyond those that are normally generated in a quest for local optimality. Metaheuristics provide a means for approximately solving complex optimization problems, as those that arise in Telecommunication Network Planning. These methods are designed to search for global optima. However, they cannot guarantee that the best solution found after termination criteria are satisfied is indeed a global optimal solution to the problem. Experimental testing of metaheuristic implementations show that the search strategies embedded in such procedures are capable of finding solutions of high quality to hard problems in industry, business and science.

The most important conclusions obtained in this dissertation are the following.

1. After a comprehensive study in chapter 1 of several real problems that arise in telecommunications and, in particular, of those arising when planning wavelength division multiplexing optical networks, the provisioning and routing problem is considered.

2. The state of the art of the problem, reported in chapter 3, confirms that there are not many works attempting to solve the capacity expansion problem in WDM optical networks. An attempt to solve this problem appears in a previous contribution, in which a linear integer programming model has been proposed and a permutation-based genetic algorithm has been developed for solving the problem. Since even an exhaustive search of all permutations may result in a sub-optimal network design, an alternative solution approach is required.

3. Other previous contributions attempt to solve a similar capacity expansion problem by simply developing a mathematical model and solving it with existing optimization software for small dimension instances. The existing optimizers are not able to reach optimal solutions in a reasonable computational time for real problems when using the mathematical models proposed to solve the problem. Therefore,

metaheuristic procedures have to be proposed and developed.

4. In Chapter 2, it is also corroborated that the parallelization of metaheuristics is a good alternative for improving its performance when attempting to solve real problems.

5. After studying the features of the provisioning and routing problem, a hybrid metaheuristic is developed. This metaheuristic combines ideas from scatter search, multistart, and tabu search. The hybrid metaheuristic takes advantage of strategies that can explore a large solution space effectively. Specifically, tabu search contributes with a short term memory component that is designed to avoid cycling. Scatter search adds a mechanism to generate new solutions from the combination of solutions in an updated reference set of solutions. Finally, the multistart component uses a long term memory that forces construction of new solutions in a wider range of the solution space.

6. For the purpose of assessing its effectiveness, the metaheuristic is compared with two variants of a permutation-based procedure and with the optimal solutions provided by an existing optimizer using a mathematical model proposed by other authors.

7. The development of the hybrid procedure that allows to consider the features of the capacity expansion problem is the result of the exhaustive study of the features of the metaheuristics described in chapter 2

8. Since the mathematical model proposed in a previous work for solving the problem has a very large number of variables and constraints, it is proposed an alternative integer linear formulation in chapter 4. The proposed formulation has significantly less integer variables than previous formulation.

9. A comparative analysis between the results obtained using the metaheuristic procedure for the problem on hand and the results obtained solving the proposed mathematical model to optimality is carried out. The mathematical model is solved as a relaxation of the original problem because it does not consider all possible paths between each pair of demand requirements. This model allows to solve real instances to optimality with bigger dimensions than the model proposed by other authors.

10. The survivability problem in optical WDM networks is tackled in chapter 5, in which three mathematical models are proposed for three alternative protection schemes.

11. After making a deep study of the state of the art of the capacity expansion problem with protection, reporter in chapter 1, we have realized that the previous attempts solve the protection problem with shared-schemes beginning with the best network design obtained for the problem without protection.

12. We propose a modified version of the hybrid metaheuristic developed in chapter 3, which solves the protection problem beginning not only with the best working network, but also with a reference set of good working networks. We have concluded that the best network design that satisfies the set of demands and that protects the traffic in the event of any single link failure, is not obtained in most cases beginning with the best working network.

13. Finally, an important consideration for additional research in this area deals with tackling uncertainty. Clearly, the demands cannot be considered known in an environment such as the telecommunications industry. The availability of a MIP formulation that can be used to find near-optimal solutions to the capacity expansion problem represents a stepping-stone toward the solution of a stochastic version of the problem that treats the demands as uncertain.

14. Several alternative mathematical models are proposed for solving the provisioning and routing problem under demand uncertainties.

15. It has been proposed a stochastic programming model, a scatter search metaheuristic that can deal with linear and non linear objective functions, and a robust approach that provides feasible solutions even in the worst case.

Future Research

The topics for future research are the following.

- The application of parallelization to the hybrid metaheuristics developed for solving the capacity expansion problem in WDM optical networks.

- The design of efficient hybrid metaheuristics for solving other relevant problems that arise when planning WDM mesh networks, such as the routing and wavelength assignment problem and the converters placement problem.

- The application of Fuzzy methodology to deal with the provisioning and routing problem under demand uncertainties.

Chapter 9

Appendix. Data and Decision Variables.

Data

N = set of nodes.

E = set of segments, $E \subseteq N \times N$.

D = set of demands, $D = (o_1, d_1, R_1), (o_2, d_2, R_2), \cdots, (o_q, d_q, R_q)$.

NC_e = number of channels on segment e required to get the working paths.

J_e denotes the set of possible paths from the origin of e to the destination of e that can be used to reroute the traffic.

$J_{e,od}$ be the set of alternative protection routes between the end nodes of demand $(o, d) \in D$ which do not contain segment e.

<u>Cost Input Data</u>

C_e^F = cost of a fiber on segment e (sum of costs per link along that segment).

$C_e^{W_j}$ = cost of a type $j \in J$ WDM unit on segment e.

C^{O_l} = cost of a type $l \in L$ OXC unit.

C^{c_j} = channel cost of a type j WDM unit.

C^{p_l} = port cost of a type l OXC unit.

Capacity Data

M^{w_j} = capacity of a type j WDM unit.

M^{o_l} = capacity of a type l OXC unit.

Existing Infrastructure

g_e^j = spare WDM channels on WDM systems of type j on segment e.

h_n^l = spare OXC ports on OXC systems of type l at node n.

Decision Variables

x_{ie} = amount of demand i routed on segment e.

x_{ie}^F = amount of demand i routed on segment e in the forward direction.

x_{ie}^R = amount of demand i routed on segment e in the reverse direction.

f_e = number of stand-alone (no WDM) fiber pairs on segment e.

w_e^j = number of type j WDM units on segment e.

v_e^j = number of channels on type j WDM units on segment e.

y_n^l = number of type l OXC units installed at node n.

u_n^l = number of ports on type l OXC units installed at node n.

x_p^{od} = 1 if demand (o, d) is routed on path p and 0 otherwise.

z_q^e = 1 if the traffic on segment e is rerouted on path $q \in J_e$ and 0 otherwise.

$z_{q,e}^{od}$ = 1 if demand (o, d) is rerouted on path $q \in J_{e,od}$ if segment e fails and 0 otherwise.

Bibliography

[1] Aarts, E. and Korst, J., "Selected Topics in Simulated Annealing". In **Essays and surveys in Metaheuristics**, editors C. Ribeiro and P. Hansen, *Kluwer Academic Publishers*, 1:1-37 (2001).

[2] Alanyali, M. and Ayanoglu, E., "Provisioning algorithms for WDM Optical Networks". *IEEE/ACM Transactions on Networking*, 3: 767-778 (1999).

[3] Balakrishnan, A., Magnanti, T. L., Shulman, A. and Wong, R. T., "Models for Planning Capacity Expansion in Local Access Telecommunication Networks". *Annals of Operations Research*, 33:239-284 (1991).

[4] Banerjee, D. and Mukherjee, B., "A practical approach for routing and wavelength assignment in large wavelength routed optical networks". I*EEE Journal of Selected Areas in Communications*, 14(5): 903-908 (1996).

[5] Baroni, S. and Bayvel, P., "Wavelength requirements in arbitrarily connected wavelength-routed networks". *IEEE Journal of Lightwave Technology*, 15: 242-251 (1997).

[6] Baroni, S., Bayvel, P., Gibbens, R., and Korotky "Analysis and Design of Resilient Multifiber Wavelength-Routed Optical Transport Networks". *IEEE Journal of Lightwave Technology*, 17(5): 743-758 (1999).

[7] Battiti, R. and Tecchiolli, G., "The reactive tabu search". *ORSA J. Comput.*, 6:126-140 (1994).

[8] Battiti, R., "Reactive search: towards self-tuning heuristics". In **Modern heuristic search methods**, editors V.J. Rayward-Smith, I.H. Osman, C.R. Reeves, and G.D. Smith, 61-83, *Wiley* (1996).

[9] Battiti, R. and Protasi, M., "Reactive local search for the maximum clique problem". *Algorithmica*, 29(4):610-637 (2001).

[10] Beale, E. M. L., "On minimizing a convex function subject to linear inequalities". *J. of the Royal Statistical Society*, *Series B*, 17:173-184 (1955).

[11] Bhargava, H. K. and Ye, N., "Computational Modeling and Problem Solving in the Networked World". **Operations Research/Computer Science Interfaces**, 21, *Kluwer Academic Publishers* (2002).

[12] Birge, J. R., "Decomposition and partitioning methods for multi-stage stochastic linear programs". *Operations Research*, 33:989-1007 (1985).

[13] Birge, J.R., "Stochastic Programming Computation and Applications", *INFORMS Journal on Computing*, 9:111-133 (1997).

[14] Birge, J. R. and Louveaux, F. V., **Introduction to Stochastic Programming**. *Springer*, New York (1997).

[15] Birman, A. and Kershenbaum, A., "Routing and wavelength assignment methods in single-hop all-optical networks with blocking". In *Proceedings IEEE INFOCOM'95*, 2: 431-438 (1995).

[16] Blanco, A., Pelta, D., and Verdegay., J. L., "A Fuzzy Valuation-Based Local Search Framework for Combinatorial Optimization". *Fuzzy Optimization and Decision Making*, 1:177-193 (2002).

[17] Boender, C. G. E. and Rinnooy Kan, A. H. G., “Bayesian stopping rules for multi-start global optimization methods”. *Mathematical Programming*, 37:59-80 (1987).

[18] Boettcher, S. and Percus, A. G., “Extremal optimization: Methods derived from co-evolution”. In proceedings of the *Genetic and Evolutionary Computation Conference, GECCO-99*, editors Wolfgang Banzhaf, Jason Daida, Agoston E. Eiben, Max H. Garzon, Vasant Honavar, Mark Jakiela, and Robert E. Smith, 825-832, San Francisco, CA (1999).

[19] Boettcher, S. and Percus, A. G., “Nature’s way of optimizing”. *Artificial Intelligence*, 119:275-286 (2000).

[20] Boettcher, S. and Percus, A. G., “Extremal optimization: an evolutionary local-search algorithm”. In **Computational Modeling and Problem Solving in the Networked World: Interfaces in Computing and Optimization**, editors H.K. Bhargava and N. Ye, *Kluwer Academic Publishers*, Netherlands (2003).

[21] Buckles, B. P. and Petry, F. E., **Genetic Algorithms**. *IEEE Computer Society Press* (1992).

[22] Cadenas-Figueredo, J. M. and Verdegay-Galdeano, J. L., **Modelos de optimizacin con datos imprecisos**. Servicio de Publicaciones de la Universidad de Murcia, ISBN: 84-8371-808-3 (1999).

[23] Chlamtac, I., Ganz, A., and Karmi, G., “Purely optical networks for trabit communication”. In *Proceedings IEEE INFOCOM’89*, 3:887-896 (1989).

[24] Chlamtac, I., Ganz, A., and Karmi, G., “Lightpath communications: An approach to highbandwidth optical WAN’s”. *IEEE Transactions on Communications*, 40: 1171-1182 (1992).

[25] **New Ideas in Optimization**. Editors D. Corne, F. Glover, and M. Dorigo, *McGraw-Hill* (1999).

[26] **Telecommunications Optimization: Heuristic and Adaptive Techniques**. Editors D. W. Corne, M. J. Oates, and G. D. Smith, *Wiley* (2000).

[27] Cox, A. L., Jr., Sanchez, J., Lu, L. and Sun, X., "Cost Savings from Network Optimization of DWDM Facilities and Optical Switches". Technical Report, Cox Associates (2000).

[28] ILOG CPLEX, on-line documentation available at `http://www.ilog.com/products/cplex`.

[29] Daniel, W. W. **Applied Nonparametric statistics**. *PWS-Kent Publishing Company*, Boston (1990).

[30] Dantzig, G. B., "Linear programming under uncertainty". *Management Science*, 1:197-206 (1955).

[31] **Handbook of Genetic Algorithms**. Editor L. D. Davis, Van Nostrand Reinhold, New York (1991).

[32] Dorigo, M., Maniezzo, V., and Colorni, A.., "The ant system: An autocatalytic optimizing process". *Technical Report DEI-91-016*, Politecnico di Milano, Italy (1991).

[33] Dueck, G., "New optimization heuristics: The great deluge algorithm and the record-to-record travel". *Journal of Computational Physics*, 104:86-92 (1993).

[34] Eberhart, R. C. and Kennedy, J. A new optimizer using particle swarm theory. Proceedings of the Sixth International Symposium on Micromachine and Human Science, Nagoya, Japan. pp. 39-43, 1995

[35] El-Fallahi, A. and Mart, R., "Tabu and scatter search for artificial neural networks". To appera in **Interfaces in Computer Science and O.R.**, editors Bhargava and Ye, *Kluwer Academic Publishers.*

[36] Feo, T. A. and Resende, M. G. C., "Greedy randomized adaptive search heuristic". *Journal of Global Optimization*, 6:109-133 (1995).

[37] Fuller, R. and Carlsson, C., **Fuzzy Reasoning in Decision Making and Optimization**. *Physica Verlag* (2002).

[38] Fumagalli, A., Cerutti, I., Tacca, M., Masetti, F., Jagannathan, R., and Alagar, S., "Survivable Networks Based on Optimal Routing and WDM Self-Healing Rings". In *Proceedings of IEEE INFOCOM'99*, 2: 726-733 (1999).

[39] Garca Lpez, F., Melin Batista, B., Moreno Prez, J. A., and Moreno Vega, J. M., "The parallel variable neighborhood search for the p-median problem". *Journal of Heuristics*, 8: 377-390 (2002).

[40] Garca Lpez, F., Melin Batista, B., Moreno Prez, J. A., and Moreno Vega, J. M., "Parallelization of the Scatter Search for the p-median problem". *Parallel Computing*, 29: 575-589 (2003).

[41] Gavish, B., "Topological Design of Telecommunication Network. Local Access Design Methods". *Annals of Operations Research*, 33:17-71 (1991).

[42] Glover, F., "Future paths for integer programming and links to artificial intelligence". *Computers and Operations Research*, 5:533-549 (1986).

[43] Glover, F., "Tabu search-Part I". *ORSA Journal on Computing*, 1(3):190-206 (1989).

[44] Glover, F., "Tabu search-Part II". *ORSA Journal on Computing*, 2:4-32 (1990).

[45] Glover, F., Laguna, M., Taillard, E. D., and De Werra, D., "Tabu Search". *Annals of Operational Research*, 43, Baltzer (1993).

[46] Glover, F. and Laguna, M., "Bandwidth Packing: A Tabu Search Approach". *Management Science*, 39(4): 492-500 (1993).

[47] Glover, F. and Laguna, M., "Tabu search". In **Modern Heuristic Techniques for Combinatorial Problems**, editor C.R. Reeves, 70-150, *McGraw-Hill* (1993).

[48] Glover, F., Taillard, E., and De Werra, D., "A user's guide to tabu search". *Annals of Operations Research*, 41:3-28 (1993).

[49] Glover, F. and De Werra, D., "Tabu Search". *Annals of Operational Research*, 41, Baltzer (1993).

[50] Glover, F. and Laguna, M., **Tabu Search**. *Kluwer Academic Publishers* (1997).

[51] Glover, F., "Multi-Start Strategic Oscillation Methods Principles to Exploit Adaptive Memory". In **Computing tools for modeling, Optimization and Simulation: Interfaces in Computer Science and Operations Research**, eds. M. Laguna and J. L. Gonzlez-Velarde. *Kluwer Academic Publishers*, 1-23 (2000).

[52] **Handbook of Metaheuristics**. Editors F. Glover and G. Kochenberger, *Kluwer Academic Publishers* (2003).

[53] Glover, F., Laguna, M., and Mart, R., "Scatter Search and Path Relinking: Advances and Applications". **Handbook of MetaHeuristics**, editors F. Glover and G. Kochenberger, 1-36 (2003).

[54] Glover, F. and Melin Batista, B., "Bsqueda Tab". To appear in *Inteligencia Artificial* (2003).

[55] Goldberg, D. E., **Genetic Algorithms in Search, Optimization and Machine Learning**. *Addison Wesley* (1989).

[56] Gomory, R. E. and Hu, T. C., "Multi-Terminal Network Flows". *SIAM J. Appl. Math.*, 9: 551-570 (1961).

[57] Gomory, R. E. and Hu, T. C., "Synthesis of a Communication Network". *SIAM J. Appl. Math.*, 12: 348-369 (1964).

[58] Grover, W. D. and Stamatelakis, D., "Cycle-oriented distributed preconfiguration: ring-like speed with mesh-like capacity for self-planning network restoration". In *Proceedings of IEEE ICC* (1998).

[59] Grover, W. D. and Stamatelakis, D., "Bridging the ring-mesh dichotomy with p-cycles". In *Proceedings of Design of Reliable Communication Networks* (2000).

[60] Gusfield, D., "Simple Construction for Multi-Terminal Network Flow Synthesis". *SIAM Journal on Computing*, 12(1): 157-165 (1983).

[61] Hansen, P. and Mladenović, N., "An Introductionto Variable Neighborhood Search" In **Meta-heuristics: Advances and Trends in Local Search Paradigms for Optimization**, chapter 30, *Kluwer Academic Publishers* (1999).

[62] Hansen, P. and Mladenović, N., "Developments of variable neighborhood search". In **Essays and Surveys in Metaheuristics**, editors C. Ribeiro and P. Hansen, 415-440, *Kluwer Academic Publishers* (2001).

[63] Hansen, P. and Mladenović, N., "Variable neighborhood search". In **Handbook of Applied Optimization**, editors M. P. Pardalos and M. G. C. Resende, 221-234, *Oxford University Press* (2003).

[64] Hansen, P. and Mladenović, N., "Variable neighborhood search". In **Handbook of Metaheuristics**, editors F. Glover and G. A. Kochenberger, 6, *Kluwer Academic Publishers* (2003).

[65] Holland, J., "Adaptation in Natural and Artificial Systems". *University of Michigan Press, Ann Arbor* (1975).

[66] Iraschko, R. R. and Grover, W. D., "A Highly Efficient Path-Restoration Protocol for Management of Optical Network Transport Integrity". *IEEE Journal on Selected Areas in Communications*, 18(5): 779-794 (2000).

[67] Jeong, G. and Ayanoglu, E., "Comparison of wavelength-interchanging and wavelength-selective cross-connects in multiwavelength all-optical networks". In *Proceedings IEEE INFOCOM'89*, 1: 156-163 (1996).

[68] Kall, P. and Wallace, S. W., **Stochastic Programming**. *John Wiley & Sons*, Chichester, England (1994).

[69] Karasan, E. and Ayanoglu, E., "Effects of wavelength routing and selection algorithms on wavelength conversion gain in WDM optical networks". *IEEE/ACM Transactions on Networking*, 6: 186-196 (1998).

[70] Kennedy, J., Eberhart, R. C., and Shi, Y., **Swarm intelligence**. *Morgan Kaufmann Publishers* (2001).

[71] Kennington, J. L., Olinick, E. V., Ortynski, A. and Spiride, G., "Robust Solutions for the WDM Routing and Provisioning Problem with Uncertain Demands and Single Link Failure Protection". *Technical Report 01-EMIS-07*, EMIS Department, School of Engineering, SMU, Dallas, TX (2001).

[72] Kennington, J. L., Olinick, E. V., Ortynski, A., and Spiride, G., "Wavelength Routing and Assignment in a Survivable WDM Mesh Network". *Operations Research*, 51(1): 67-79 (2003).

[73] Kennington, J. L., Lewis, K. R., Olinick, E. V., Ortynski, A., and Spiride, G., "Robust Solutions for the WDM Routing and Provisioning Problem: Models and Algorithms". To Appear in *Optical Networks Magazine* (2003).

[74] Kennington, J. L. and Olinick, E. V., "Wavelengths Translation in WDM Networks: Optimization Models and Solution Procedures". To appear in *INFORMS Journal on Computing* (2003).

[75] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., "Optimization by simulated annealing". *Science*, 220:671-680 (1983).

[76] Kouvelis, P. and Yu, G., **Nonconvex Optimization and its Applications. Robust Discrete Optimization and its Applications**. *Kluwer Academic Publishers* (1997).

[77] Laguna, M., Campos, V., and Mart, R., "Scatter search for the linear ordering problem". In **New Ideas in Optimization**, editors M. Dorigo, D. Corne, and F. Glover, 331-341, *McGraw-Hill* (1999).

[78] **Computing Tools for Modeling, Optimization and Simulation**. Editors M. Laguna and J.L. González-Velarde, *Kluwer Academic Publishers* (2000).

[79] Laguna, M. and Mart, R., "Experimental testing of advanced scatter search designs for global optimization of multimodal functions". Technical Report, University of Colorado (2000).

[80] Laguna, M., Glover, F., and Mart, R., "Fundamentals of scatter search and path relinking". *Control and Cybernetics*, 39(3):653-684 (2000).

[81] Laguna, M., Campos, V., Glover, F., and Mart, R., "An experimental evaluation of a scatter search for the linear ordering problem". *Journal of Global Optimization*, 21:397-414 (2001).

[82] Laguna, M., "Scatter Search". In **Handbook of Applied Optimization**, editors P. M. Pardalos and M. G. C. Resende, *Oxford University Press* (2002).

[83] Laguna, M., Glover, F., and Mart, R., "Scatter search and path relinking: Advances and applications". In **Handbook of MetaHeuristics**, editors F. Glover and G. Kochenberger, *Kluwer Academic Publishers* (2003).

[84] Laguna, M., Glover, F., and Mart, R., "Scatter Search". To appear In **Theory and Applications of Evolutionary Computation: Recent Trends**, editors A. Ghosh and S. Tsutsui, *Springer-Verlag*.

[85] Laguna, M. and Mart, R., **Scatter Search: Methodology and Implementations in C**. *Kluwer Academic Publishers* (2003).

[86] Laguna, M., Glover, F., and Mart, R., New ideas and applications of scatter search and path relinking. To appear in **New Optimization Techniques in Engineering**, Godfrey Onwubolu, Springer-Verlag.

[87] Laguna, M., Glover, F., and Mart, R., "Scatter search and path relinking: Foundations and advanced designs". To appear in **New Optimization Techniques in Engineering**, editor G. Onwubolu, *Springer-Verlag*.

[88] Larrañaga, P. and Lozano, J. A., **Estimation of Distribution Algorithms**, *Genetic Algorithms and Evolutionary Computation*, volume 2, *Kluwer Academic Publishers* (2002).

[89] Lawler, E. S., **Combinatorial Optimization: Networks and Matroids**. Holt, Rinehart and Winston, New York (1976)

[90] Li, L. and Somani, A. K., "Dynamic wavelength routing using congestion and neighborhood information". *IEEE/ACM Transactions on Networking*, 7:779-786 (1999).

[91] Lourenco, H. R., Martin, O., and Stützle, T., "A beginner's introduction to iterated local search". In *Proceedings of MIC 2001*, 1-6, Porto, Portugal (2001).

[92] Lozano, J. A. and Mendiburu, A., **Estimation of Distribution Algorithms Applied to the Job Shop Scheduling Problem**, volume 2 of **Genetic Algorithms and Evolutionary Computation**, chapter 11, *Kluwer Academic Publishers* (2002).

[93] Magnanti, T. L., Wong, R. T., "Network Design and Transportation Planning: Models and Algorithm", *Transportation Science*, 18, 1-55 (1984).

[94] Mart, R., "Multi-Start Methods". In **Handbook of MetaHeuristics**, editors F. Glover and G. Kochenberger, *Kluwer Academic Publishers* (2003).

[95] Melián, B., Moreno-Pérez, J. A., and Moreno-Vega, J. M., "A Multistart Clustering Tecnique for Combinatorial Optimization". In *The best of MS2000 International Conference on Modelling and Simulation*, pp. 175-185 (2000).

[96] Melián, B., Laguna, M., and Moreno-Pérez, J. A., "Minimizing the Cost of Placing and Sizing Wavelength Division Multiplexing and Optical Cross-Connect Equipment in a Telecommunications Network". *Technical Report*, University of La Laguna, Spain (2002) (Submitted to *Networks*).

[97] Melián, B., Laguna, M., and Moreno-Pérez, J. A., "Capacity Expansion of Fiber Optic Networks with WDM Systems: Problem Formulation and Comparative Analysis". To appear in *Computers and Operations Research*, 2002.

[98] Melián, B., Laguna, M., and Moreno-Pérez, J. A., "Provisioning of Survivable WDM Mesh Networks Under Demand Uncertainty and Single Link Failure Protection". *Technical Report*, University of La Laguna, Spain (2003).

[99] Michalewicz, Z., **Genetic Algorithms + Data Structures = Evolution Programs**, *Springer Verlag* (1992).

[100] Michalewicz, Z. and Fogel, D. B., **How to Solve It: Modern Heuristics**, *Springer Verlag* (2000).

[101] Mills, P. and Tsang, E. P. K., "Guided local search for solving sat and weighted max-sat problems". *Journal of Automated Reasoning*, 24:205-223 (2000).

[102] Mills, P., Tsang, E. P. K., and Ford, J., "Applying an extended guided local search on the quadratic assignment problem". *Annals of Operations Research*, 118:121-135 (2003).

[103] Mitchel, M., **An introduction to Genetic Algorithms**. *MIT Press* (1998).

[104] Mladenović, N., "A variable neighborhood algorithm - a new metaheuristic for combinatorial optimization". In *Abstracts of Papers Presented at Optimization Days* (1995).

[105] Mladenović, N. and Hansen, P., "Variable neighborhood search". *Computers and Operations Research*, 24:1097-1100 (1997).

[106] Moscato, P., "Memetic algorithms: a short introduction". In **New Ideas in Optimization**, editors D. Corne, F. Glover, and M. Dorigo, 219-234, *McGraw-Hill* (1999).

[107] Mulvey, J. M., Vanderbei, R. J., and Zenios, S. A., "Robust Optimization of Large-Scale Systems". *Operations Research*, 43(2), March-April (1999).

[108] Nagatsu, N., Hamazumi, Y., and Sato, K., "Number of wavelengths required for constructing large-scale optical paths networks". *Electronics and Communications in Japan*, Part1, 78(9): 1-11 (1995).

[109] Nagatsu, N. and Sato, K., "Optical Path accommodation design enabling cross-connect system scale evaluation". *IEICE Transactions on Communications*, E78-B(9): 1339-1343 (1995).

[110] OptQuest Callable Library, OptTek Systems Inc., www.opttek.com (2001).

[111] Orlin, J. B., Ahuja, R. K., and Sharma, D., "Very large-scale neighborhood search". *International Transactions in Operations Research*, 7:301-317 (2000).

[112] Orlin, J. B., Ahuja, R. K., Ergun, O., and Punnen, A. P., "A survey of very large-scale neighborhood search techniques". To appear in *Discrete Applied Mathematics* (2002).

[113] Prkopa, A., **Stochastic Programming**, *Kluwer Academic Publishers*, Dordrecht (1995).

[114] Ramaswami, R. and Sivarajan, K. N., "Routing and Wavelength Assignment in All-Optical Networks". *IEEE/ACM Transactions on Networking*, 3(5): 489-500 (1995).

[115] Ramamurthy, S. and Mukherjee, B., "Wavelength conversion in WDM netowrking". *IEEE Journal on Selected Areas in Communications*, 16: 1061-1073 (1998).

[116] Ramamurthy, S. and Mukherjee, B., "Survivable WDM mesh networks, part I - protection". In *Proceedings of IEEE INFOCOM'99*, 2: 744-751 (1999).

[117] Ramamurthy, S. and Mukherjee, B., "Fixed-alternate routing and wavelength conversion in wavelength-routed optical networks". *IEEE/ACM Transactions on Networking*, 10: 351-367 (2002).

[118] **Genetic Algorithms - Principles and Perspectives A Guide to GA Theory**. Editors C. R. Reeves and J. E. Rowe, volume 20 of *Operations Research/Computer Science Interfaces*, *Kluwer Academic Publishers* (2002).

[119] Resende, M. G. C. and Ribeiro, C.C., "Greedy randomized adaptive search procedures". In **Handbook of Metaheuristics**, editors F. Glover and G. G. Kochenberger, *Kluwer Academic Publishers* (2003).

[120] **Essays and Surveys in Metaheuristics**. Editors C. C. Ribeiro and P. Hansen, *Kluwer Academic Publishers* (2001).

[121] Rockafellar, R. T. and Wets, R. J.-B., "Scenarios and policy aggregation in optimization under uncertainty", *Math. of Oper. Res.* 16: 119-147 (1991).

[122] Rosing, K. E. and ReVelle, C. S., "Heuristic concentration: two stage solution construction". *European Journal of Operations Research*, 97:75-86 (1997).

[123] Rosing, K. E., ReVelle, C. S., Rolland, E., Schilling, D. A., and Current, J. R., "Heuristic concentration and tabu search: A head to head comparison", *European Journal of Operations Research*, 104(1):93-99 (1998).

[124] Rosing, K. E. and Hodgson, M. J., "Heuristic concentration for the *p*-median: an example demonstrating how and why it works". *Computers and Operations Research*, 29(10):1317-1330 (2002).

[125] Schupke, D. A., Gruber, C. G., and Autenrieth, A., "Optimal Configuration of p-Cycles in WDM Networks". In *Proceeding of IEEE ICC* (2002).

[126] Sen, S. and Higle, J. L., "An Introductory Tutorial on Stochastic Linear Programming Models", *Interfaces*, 29:33-61 (1999).

[127] Sharda, R. and Rampal, R., "Neural networks and management science/operations research". http://catt.okstate.edu/itorms/guide/nnpaper.html (1995).

[128] Sridhar, S. and Chandrasekaran, R., "Integer Solution to Synthesis of Communications Networks". *Mathematics of Operations Research*, 17(3): 581-585 (1992).

[129] Stamatelakis, D. and Grover, W. D., "IP Layer Restoration and Network Planning Based on Virtual Protection Cycles". *IEEE Journal on Selected Areas in Communications*, 18(10): 1938-1949 (2000).

[130] Talluri, K. T., "Network synthesis with few edges". *Networks*, vol. 27(2): 109-115 (1996).

[131] Van Caenegem, B., Van Parys, W., De Turck, F., and Demeester, P. M., "Dimensioning of WDM Networks". *IEEE Journal on Selected Areas in Communications*, 16(7): 1146-1157 (1998).

[132] Van der Vlerk, M. H., “Stochastic Programming Bibliography”, World Wide Web, http://mally.eco.rug.nl/biblio/stoprog.html (1996-2002).

[133] Van Laarhoven, P. J. M. and Aarts, E. H. L., **Simulated Annealing: Theory and Applications**. *Kluwer Academic Publishers* (1987).

[134] Vidal, R. V. V., “Applied Simulated Annealing”. *Lecture Notes in Econ. and Math. Systems* 396, *Springer Verlag* (1993).

[135] **Optimization Software Class Libraries**. Editors S. Voss and D. L. Woodruff, volume 18 of *Operations Research/Computer Science Interfaces*, *Kluwer Academic Publishers* (2002).

[136] Voudouris, C. and Tsang, E. P. K., “Guided local search”. *European Journal of Operational Research*, 113(2):469-499 (1999).

[137] Wauters, N. and Demeester, P., “Design of the optical path layer in multiwavelength cross-connected networks”. *IEEE Journal on Selected Areas in Communications*, 14:881-892 (1996).

[138] Winston, W. L., “Introduction to Mathematical Programming: Applications and Algorithms”. *Munirpallam Venkataramanan*, *Brooks/Cole Publishing* (2003).

[139] Xu, Z. B., Jin, H. D., Leung, K. S., Leung, Y., and Wong, C. K., “An automata network for performing combinatorial optimization”. *Neurocomputing*, 47:59-83 (2002).

[140] Yagiura, M. and Ibaraki, T., “Local search”. In **Handbook of Applied Optimization**, editors P. M. Pardalos and M. G. C. Resende, 104-123, *Oxford University Press* (2002).

[141] Zadeh, L.A., “Fuzzy sets”, *Information and Control*, 8:338-353 (1965).

[142] Zang, H., Jue, J., and Mukherjee, B., "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks". *Optical Networks Magazine*, 1(1):4760 (2000).

[143] Zang, H., **WDM Mesh Networks: Management and Survivability**. *Kluwer Academic Publishers* (2003).