



HAL
open science

Communication in Wavelength Multiplexed Optical Networks

Bruno Beauquier

► **To cite this version:**

Bruno Beauquier. Communication in Wavelength Multiplexed Optical Networks. Networking and Internet Architecture [cs.NI]. Université Nice Sophia Antipolis, 2000. English. NNT : . tel-00504392

HAL Id: tel-00504392

<https://theses.hal.science/tel-00504392>

Submitted on 20 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

présentée à

L'UNIVERSITÉ DE NICE – SOPHIA ANTIPOLIS
U.F.R. SCIENCES

pour obtenir le titre de

DOCTEUR

Spécialité : **INFORMATIQUE**

COMMUNICATIONS DANS LES RÉSEAUX OPTIQUES PAR MULTIPLEXAGE EN LONGUEUR D'ONDE

par **Bruno BEAUQUIER**

Soutenue le 17 janvier 2000 devant le jury composé de :

M.	Ioan	BOND	Président
M.	Jean-Claude	BERMOND	Directeur
Mme.	Dominique	SOTTEAU	Rapporteur
M.	François	TILLEROT	Rapporteur
M.	Ugo	VACCARO	Rapporteur
M.	Pierre	FRAIGNIAUD	Examineur
M.	Jean-Claude	KÖNIG	Examineur

À mes parents.

Remerciements

Je tiens particulièrement à remercier toutes les personnes qui m'ont aidé, tant sur le plan professionnel que sur le plan personnel, à mener à terme cette épreuve :

- Johny Bond, responsable de l'équipe PaCom, pour l'honneur qu'il me fait de présider mon jury, et pour l'extrême gentillesse avec laquelle il m'a accueilli lors de mon arrivée à Sophia Antipolis ;
- Jean-Claude Bermond, pour m'avoir accueilli dans son équipe et avoir accepté de diriger cette thèse. Je lui suis profondément reconnaissant de l'aide et du temps précieux qu'il m'a consacrés, de la confiance qu'il m'a toujours accordée, et surtout de l'ambiance particulièrement amicale qu'il sait instaurer et qui m'a permis d'effectuer mon doctorat dans les meilleures conditions possibles ;
- Dominique Sotteau, pour l'honneur qu'elle me fait de participer à mon jury, malgré ses contraintes professionnelles, pour la gentillesse et le courage d'avoir rapporté cette thèse, et pour ses nombreux contacts cordiaux et utiles tout au long de mes travaux ;
- François Tillerot, pour l'honneur qu'il me fait de participer à mon jury, pour la gentillesse et le courage d'avoir rapporté cette thèse dans des délais aussi courts, et pour ses nombreuses remarques très profitables sur la partie technologique de ce manuscrit ;
- Ugo Vaccaro, pour l'honneur qu'il me fait de participer à mon jury, pour la gentillesse et le courage d'avoir rapporté cette thèse, et pour l'hospitalité chaleureuse dont il a fait preuve lors de mon séjour à Salerno ;
- Pierre Fraigniaud et Jean-Claude König, pour l'honneur qu'ils me font de faire partie de mon jury ;
- Stéphane Pérennes, pour sa précieuse collaboration et son contact très enrichissant, et pour nos nombreuses discussions amicales ;
- Michel Syska, pour avoir accepté d'être mon tuteur de monitorat, pour tous les précieux conseils qu'il m'a fournis, pour son amitié sincère, voire (même) pour son hospitalité généreuse lors de nombreuses soirées footballistiques ;
- Olivier Delmas, pour le plaisir d'avoir partagé son bureau et d'avoir travaillé avec lui, et pour toutes nos discussions enjouées ;
- Eric Darrot, pour le plaisir d'avoir collaboré avec lui et pour avoir apprécié son obstination acharnée et son sens de la répartie ;

- Nausica Marlin, pour le plaisir de l'avoir côtoyée dans la bonne humeur, pour tous les moments passés ensemble et la complicité que nous avons partagée.
- Afonso Ferreira, pour ses conseils avisés, son contact amical et pour nos nombreuses discussions extra-professionnelles ;
- les autres membres du projet SLOOP, David Coudert, Tania Jiménez, Philippe Mussi, Jean-Noël Petit, ses ex-membres, Françoise Baude, Denis Caromel, Nathalie Furmento, Bruno Gaujal, Peter Mammay, Yves Roudier, David Sagnol, Günther Siegel, Laurent Villefranche, et les nouveaux, Sébastien Choplin et Xavier Hardy, pour l'ambiance sympathique qu'ils font ou ont fait régner dans les bureaux et les couloirs du projet ;
- Ephie Deriche, Zohra Khalafi et Patricia Lachaume, tant pour leur diligence et leur efficacité à résoudre mes embarras administratifs, que pour leur gentillesse et leur disponibilité ;
- François Brémond, Olivier Dalle et Cyril Godart, dits les Marmules, pour leur accueil sympathique, leurs bouffes soit diététiques, soit pantagruéliques, et pour l'amitié forgée autour de leur table ou encore à la sueur des bourres tirées dans les côtes sphiapolitaines ;
- mes autres compagnons de course à pied, Jérôme Galtier, Igor Litovski et Philippe Nain, pour leur bonne humeur permanente et leur saine émulation lors de nos sorties à l'heure du déjeuner ;
- Sandrine Julia et Bruno Martin, pour leur accueil chaleureux et leur amitié, entretenue malgré l'éloignement, à coups de bouffes et de sorties au ski ;
- Joseph Yu, pour le plaisir d'avoir fait sa connaissance dès mon arrivée, pour ses conseils avisés et sa gentillesse perpétuelle ;
- Luisa Gargano, Pavol Hell et David Peleg, pour leur contact à la fois enrichissant sur le plan professionnel et très amical sur le plan personnel ;
- toutes les personnes rencontrées sur Sophia et qui m'ont permis de passer de très bons moments : Soraya Arias, Thomas Bonald, Sandrine Chevrin, Jérôme Durand-Lose, Patrick Itay, Nathanael Rota, et j'en oublie ;
- mes amis de plus longue date, Alexandre, Anne-Emmanuelle, Frédéric, Julia et Yves, pour le plaisir que j'ai eu à les connaître, pour toutes les aventures vécues ensemble et pour leur amitié sincère ;
- et bien sûr mes parents, pour l'aide et l'affection qu'ils n'ont jamais cessé de m'apporter.

Table des matières

Introduction	1
1 La technologie optique pour les télécommunications	11
1.1 Aperçu historique	11
1.2 Systèmes de transmission optiques	16
1.2.1 Fibre optique	16
1.2.2 Émetteurs optiques	18
1.2.3 Récepteurs optiques	20
1.2.4 Amplificateurs optiques	21
1.3 Multiplexage optique	23
1.3.1 Multiplexage temporel (TDM)	23
1.3.2 Multiplexage en longueur d'onde (WDM)	24
1.3.3 Comparaison entre TDM et WDM	25
1.4 Commutation optique	26
1.4.1 Générations de réseaux	26
1.4.2 Commutateurs optiques	27
1.4.3 Convertisseurs optiques	30
2 Problématique et modélisation	35
2.1 Problématique	35
2.2 Graphes et réseaux	37
2.2.1 Notions élémentaires de théorie des graphes	38
2.2.2 Réseaux usuels	40
2.3 Optimisation et approximation	42
2.4 Formulation des problèmes étudiés	44
2.4.1 Problème du routage tout-optique	45
2.4.2 Problème de la charge	47
2.4.3 Relations entre les problèmes	49
2.5 Communications structurées	51
3 Routage WDM tout-optique	59
3.1 Réseaux généraux	59
3.1.1 Instances quelconques	59

3.1.2	Permutations et k -relations	68
3.1.3	Diffusion et multicast	72
3.1.4	Échange total	73
3.2	Réseaux particuliers	75
3.2.1	Arbres	75
3.2.2	Cycles	81
3.2.3	Grilles	83
Conclusion et perspectives		91
A Optimal wavelength–routed multicasting		
B. Beauquier, P. Hell et S. Pérennes		
<i>Discrete Applied Mathematics</i> , volume 84, pages 15–20, Elsevier Science, 1998. Editors’ Choice 1998.		
Rapports de recherche I3S 97-32 et INRIA 3276.		95
B All-to-all communication in some wavelength–routed all–optical networks		
B. Beauquier		
<i>Networks: An International Journal</i> , volume 33, numéro 3, pages 179–187, Wiley-Interscience, 1999.		
Rapports de recherche I3S 98-06 et INRIA 3452.		103
C All-to-all routing and coloring in weighted trees of rings		
B. Beauquier, S. Pérennes et D. Tóth		
In <i>Proceedings of the 11th ACM Symposium on Parallel Algorithms and Architectures (SPAA’99)</i> , pages 185–190, Saint-Malo, France, 27-30 juin 1999.		
Rapports de recherche I3S 99-12 et INRIA 3700.		121
D Tight bounds for broadcasting in the linear cost model		
B. Beauquier, O. Delmas et S. Pérennes		
Rapports de recherche I3S 99-20 et INRIA 3827.		137
E Broadcasting in WDM optical rings and tori		
B. Beauquier		
In P.-J. Wan, D.-Z. Du et P. M. Pardalos, editors, <i>DIMACS Workshop on Multichannel Optical Networks: Theory and Practice</i> , Rutgers University, NJ, 16-19 mars 1998, volume 46 of <i>DIMACS Series in Discrete Mathematics and Theoretical Computer Science</i> , pages 63–73. American Mathematical Society, 1998.		
Rapports de recherche I3S 98-05 et INRIA 3410.		159

F On arbitrary Waksman networks and their vulnerability

B. Beauquier et E. Darrot

Soumis à *Parallel Processing Letters* (World Scientific). Version étendue de l'article paru dans *Actes des 1ères Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel'99)*, pages 95–100, Roscoff, France, 5-7 mai 1999.

Rapports de recherche I3S 99-17 et INRIA 3788.

175

Introduction

Les résultats obtenus dans cette thèse, qui s'est déroulée au sein du projet SLOOP¹, portent principalement sur l'étude des **communications dans les réseaux optiques par multiplexage en longueur d'onde**. Ils s'inscrivent dans une thématique d'allocation des ressources en vue de réaliser des communications dans un réseau. La problématique générale que nous avons considérée peut se résumer de la manière suivante. Il s'agit de satisfaire dans un réseau optique une famille de requêtes de connexion, appelée *instance* de communication et formée de couples de nœuds (source, destination). La satisfaction d'une requête passe par l'attribution d'un chemin dans le réseau et d'une longueur d'onde sur les liens utilisés, avec la contrainte que deux requêtes ne peuvent pas utiliser le même lien avec la même longueur d'onde. L'objectif dans ce cadre est de minimiser l'utilisation des ressources optiques, c'est-à-dire le nombre total de longueurs d'onde permettant de satisfaire l'instance donnée.

La nécessité pressante de disposer de réseaux de communication à très haut débit, dont les performances dépassent largement celles que peuvent fournir les réseaux ATM actuels, se fait jour face au nombre croissant d'utilisateurs et à l'émergence d'applications en réseau intensives, telles que la récupération de données sur Internet, les applications Java, les conférences à distance, l'imagerie en temps réel ou encore le calcul intensif distribué. La technologie de la fibre optique peut apporter une solution pour subvenir à ce besoin, grâce à ses caractéristiques de transmission exceptionnelles, sans comparaison avec les capacités des réseaux classiques. Son intérêt majeur réside dans sa bande passante énorme qui dépasse celle des transmissions électroniques de plusieurs ordres de grandeur. Alors que la fibre optique était une curiosité technologique pendant les dernières décennies, le défi est à présent de rendre ses promesses effectives afin de pourvoir à la demande des réseaux de communication du futur.

En tenant compte de la contrainte pour un site connecté – qui peut être une station de travail ou une porte d'accès vers des sous-réseaux locaux – de ne pouvoir transmettre et recevoir des données qu'au débit électronique, la clé du problème dans la conception de réseaux de communication exploitant la bande

¹SLOOP est un projet commun CNRS / INRIA / UNIVERSITÉ DE NICE - SOPHIA ANTIPOLIS.

passante optique est de parvenir à introduire dans les architectures et les protocoles des moyens de transmission des données concurrents. Dans un réseau de communication optique, cette concurrence peut être établie par la technique du *multiplexage en longueur d'onde* (*Wavelength-Division Multiplexing*, WDM).

La technologie WDM est une approche qui permet de gérer l'importante discordance opto-électronique en termes de bande passante, en permettant aux équipements des sites connectés d'accéder au réseau au débit de l'électronique seulement, ceci en multiplexant sur la même fibre optique plusieurs *canaux* WDM connectant différents sites. Le spectre optique d'une fibre est découpé en un certain nombre d'intervalles de longueurs d'onde (ou fréquences) distincts, de telle sorte que chaque intervalle supporte un canal de communication transmettant au débit désiré, en particulier celui de l'électronique.

Il est communément admis que la technologie WDM constitue actuellement le procédé de multiplexage favori pour les réseaux de communication optiques. Le problème devient alors de concevoir et de développer des architectures de réseau, des protocoles et des algorithmes appropriés, destinés à assurer une utilisation efficace des nouvelles potentialités offertes, notamment en ce qui concerne l'allocation des ressources optiques. Plus spécifiquement, le problème du routage optique des communications, qui consiste à attribuer une route et une longueur d'onde aux différentes requêtes entre les sites connectés, requiert la résolution de nombreux problèmes algorithmiques de nature combinatoire.

Pour résoudre ce type de problèmes, nous nous sommes placés dans un contexte d'optimisation algorithmique *déterministe* et *statique* (*off-line*). En d'autres termes, nous supposons les données du réseau et de l'instance fixées une fois pour toutes, et nous cherchons un routage optique satisfaisant des contraintes et minimisant une fonction de coût. Cette situation se rencontre aussi bien dans le cadre des réseaux de transport, des réseaux locaux ou des interconnexions de processeurs, lorsque les communications sont connues à l'avance. En revanche, nous n'avons pas considéré de méthodes *probabilistes* de résolution, ni l'approche *dynamique* (*on-line*) du problème, pour laquelle les requêtes de connexion se présentent au cours du temps.

Dans les réseaux optiques WDM, le nombre de longueurs d'onde disponibles est fortement limité, généralement de l'ordre de la trentaine (on trouve cependant aujourd'hui une offre commerciale de Lucent à 80 canaux), et les perspectives futures ne laissent pas envisager de dépasser l'ordre de la centaine. De plus, le coût et la complexité des nœuds de commutation augmentent significativement avec le nombre de longueurs d'onde à traiter. C'est pourquoi il est crucial de chercher à minimiser l'utilisation de ces ressources optiques. Ce type d'optimisation revêt une importance particulière dans le cadre de la conception et du dimensionnement

d'un réseau optique : d'une part, en vue de satisfaire une certaine demande de trafic prévue à l'avance, et afin de minimiser le coût global des équipements nécessaires ; d'autre part, pour gérer l'utilisation des ressources optiques en assurant la possibilité d'établir de nouvelles connexions. Cette problématique est celle du projet PORTO ² qui a récemment démarré.

Les points d'accès au réseau optique communiquent entre eux via des canaux WDM, dits *tout-optiques* lorsqu'ils ne subissent pas de conversion électronique intermédiaire, et qui sont également appelés *chemins optiques*. Une telle liaison peut recouvrir une série de liens en fibre optique et fournir ainsi une connexion de type *commutation de circuits* (*circuit-switching*, en anglais) entre deux sites distants dans la topologie physique. Chaque nœud de routage intermédiaire doit pour cela posséder une fonction de commutation optique adaptée. Les commutateurs optiques les plus simples actuellement commercialisés (par exemple par Alcatel ou Lucent Technologies) sont capables de diriger une fibre d'entrée sur n'importe quelle fibre de sortie (*répartiteur* ou *fiber optical cross-connect* (F-OXC)). D'autres, plus évolués, peuvent extraire d'une fibre certaines longueurs d'onde et en réinsérer (*multiplexeur à insertion/extraction* (MIE) ou *optical add/drop multiplexer* (OADM)). D'autres enfin, plus sophistiqués encore, appelés *brasseurs*, permettent d'envoyer un canal WDM indépendamment des autres sur n'importe quelle fibre de sortie, sans modifier sa longueur d'onde (*wavelength routing optical cross-connect* (WR-OXC)), ou en la convertissant en une autre (*wavelength translating optical cross-connect* (WT-OXC)).

En fonction de ces types de nœuds de commutation, différentes architectures pour les réseaux tout-optiques ont été proposées dans la dernière décennie. Toutes présupposent cependant l'utilisation de liens entre les nœuds de routage qui sont constitués de paires de fibres optiques unidirectionnelles. Ainsi, nous modélisons un réseau tout-optique par un graphe ou multi-graphe orienté symétrique, dont les sommets représentent indifféremment les points d'accès et les nœuds de routage. Dans ce cadre, chaque arc correspond à une fibre optique unidirectionnelle qui supporte un nombre fixé de longueurs d'onde, que nous supposons uniforme dans tout le réseau. Notons toutefois qu'un arc ne représente pas nécessairement une liaison physique point-à-point, mais peut correspondre à une liaison virtuelle pré-établie. Ainsi la topologie fournie au départ peut être celle librement choisie par le concepteur du réseau. De plus, nous considérons l'une des architectures les plus communes et les plus souvent étudiées, qui est constituée de nœuds de commutation formés de brasseurs et de multiplexeurs à insertion/extraction, sans toutefois autoriser la conversion des longueurs d'onde. La possibilité de conversion totale

²PORTO (Planification et Optimisation des Réseaux de Transport Optiques) est un projet précompétitif du RNRT (Réseau National de Recherche en Télécommunications) entre le projet SLOOP, ALCATEL (Corporate Research Center) et FRANCE TELECOM (CNET).

dans tout le réseau a été cependant abordée, mais nous avons laissé de côté l’hypothèse de conversion partielle et limitée à certains nœuds.

Dans le **chapitre 1**, nous présentons la technologie optique pour les télécommunications, afin de préciser le cadre technique de notre recherche et d’aider le lecteur informaticien à la compréhension des contraintes physiques sous-jacentes à la modélisation théorique. Cette présentation vise à caractériser les problèmes essentiels rencontrés en pratique, en gardant toutefois à l’esprit que les progrès technologiques dans le domaine sont rapides.

Nous commençons dans la section 1.1 par donner un aperçu historique de l’évolution des télécommunications et de l’avènement de la fibre optique comme support de transmission privilégié. Nous poursuivons dans la section 1.2 en décrivant les différents éléments d’un système de transmission optique – que sont la fibre, les émetteurs, les récepteurs et les amplificateurs – et en donnant quelques caractéristiques techniques.

Dans la section 1.3, nous présentons deux techniques de multiplexage qui permettent d’exploiter la bande passante spectrale de la fibre optique : le multiplexage temporel (*Time-Division Multiplexing*, TDM) et le multiplexage en longueur d’onde (WDM). Nous expliquons pourquoi cette dernière approche semble préférable. Enfin, nous donnons dans la section 1.4 une classification des différents types de nœuds de commutation envisagés dans la conception de réseaux dits tout-optiques, dans lesquels les signaux sont acheminés à très haut débit sans conversion opto-électronique intermédiaire entre la source et la destination.

Dans le **chapitre 2**, nous posons la problématique étudiée au cours de la thèse et nous donnons la modélisation qui a servi de base à nos recherches. Nous commençons dans la section 2.1 par rappeler le contexte qui est celui des réseaux de télécommunications tout-optiques, dans lesquels il est possible de réaliser des communications sous le mode de la “commutation de circuit” (*circuit switching*), déjà abondamment étudié pour les réseaux d’interconnexion entre processeurs. Ainsi une requête de connexion entre deux sites distants peut être établie par l’allocation d’un chemin optique, qui consiste en l’attribution d’une route dans le réseau et d’une longueur d’onde sur les liens utilisés. Un tel routage optique présuppose une certaine fonctionnalité des nœuds de commutation. Nous précisons le type de commutateurs optiques retenu dans notre étude, en insistant sur sa sélectivité en longueur d’onde. Cette hypothèse permet de commuter les chemins optiques indépendamment les uns des autres, avec la contrainte que deux chemins ne peuvent utiliser la même longueur d’onde sur le même lien. De plus, sans conversion dans les nœuds de routage, la même longueur d’onde doit être utilisée tout le long d’un chemin optique.

Dans la section 2.2, nous donnons les notions élémentaires de la théorie des graphes nécessaires à la modélisation des communications dans les réseaux optiques, ainsi que la définition de topologies particulières étudiées par la suite. Nous rappelons dans la section 2.3 quelques notions sur les problèmes d'optimisation et les algorithmes d'approximation.

Nous formulons dans la section 2.4 les problèmes d'optimisation étudiés dans le chapitre suivant et les annexes. Nous commençons par modéliser un réseau tout-optique par un graphe orienté, la plupart du temps symétrique, puis nous définissons formellement d'une part le problème du routage optique, et d'autre part celui de la charge. Ce dernier problème correspond au cas où la conversion en longueur d'onde est possible. Dans les deux cas, l'objectif est de trouver un ensemble R de chemins optiques pour un ensemble I de requêtes de connexion donné dans un graphe G donné, et de minimiser le nombre total de longueurs d'onde utilisées (l'optimal est noté $\vec{w}(G, I)$ dans le premier cas, noté $\vec{\pi}(G, I)$ et appelé *charge* dans le second). Ces problèmes de minimisation sont reliés à des problèmes de maximisation associés, pour lesquels il s'agit de trouver un ensemble de requêtes réalisables de cardinalité maximale, étant donné un nombre de longueurs d'onde fixé. Plus généralement, tous ces problèmes sont intrinsèquement connectés à des problèmes de routage disjoint et de multiflot entier.

Nous terminons ce chapitre par la section 2.5 qui expose certains schémas de communication étudiés par la suite. Ce type de communications, dites structurées, possède une régularité qui permet une formalisation et une résolution plus aisées, et peut se révéler intéressant à étudier dans la mesure où il se rapproche des situations réelles plus irrégulières.

Le **chapitre 3** est une synthèse des résultats obtenus dans la littérature concernant principalement le problème du routage optique. Une version préliminaire de cet état de l'art avait été fournie dans [1]. Nous avons tenté de présenter les apports les plus significatifs selon notre point de vue, en espérant n'avoir oublié aucune contribution importante. L'accent a été mis sur les rapports entre les paramètres \vec{w} et $\vec{\pi}$, c'est-à-dire sur les relations entre le problème du routage optique et celui de la charge, pour lequel il existait déjà nombre de méthodes de résolution algorithmiques, de par son antériorité dans les domaines de la théorie des graphes et de l'optimisation combinatoire.

Nous commençons dans la section 3.1 par donner des résultats de complexité montrant la difficulté du problème étudié dans sa généralité, à cause de ses liens étroits avec les problèmes de multiflot entier et de coloration, connus comme étant difficiles. Nous mentionnons ensuite des bornes inférieures et supérieures générales relatives à certains paramètres des graphes. Nous poursuivons cette approche en restreignant les instances de communication considérées à celles des

permutations, des k -relations, de la diffusion, du multicast, puis de l'échange total. Dans la section 3.2, nous en venons à considérer des topologies de réseaux particulières, que sont les arbres, les cycles et les grilles, dont il est généralement plus facile d'exploiter les propriétés structurelles.

La conclusion présente des perspectives de nouveaux axes de recherche sur les communications optiques, qu'il semble d'ores et déjà intéressant d'approfondir au vu de l'évolution de la technologie et de son application effective. Le reste de la thèse est constituée des annexes qui rassemblent les articles publiés, dans le format des rapports de recherche.

L'**annexe A** correspond à l'article [2] et présente un algorithme polynômial pour résoudre de façon optimale le problème du routage optique pour les instances de multicast dans les graphes orientés, avec un nombre de longueurs d'onde égal à la charge ($\vec{w} = \vec{\pi}$). Une conséquence de ce résultat, comme d'autres qui suivent, est que la conversion des longueurs d'onde n'est pas utile pour router optiquement ce type de communications. La particularité du multicast, qui possède une source unique pour toutes les requêtes de connexion, a permis de ramener la résolution du problème à celle d'un flot entier simple dans un réseau associé, pour lequel une méthode optimale est connue. Un argument de coupe permet de relier cette optimalité à celle de la charge.

L'**annexe B** correspond à l'article [3] et donne des méthodes constructives optimales (ou quasi-optimales) pour router optiquement l'instance de l'échange total dans diverses topologies régulières : les grilles toriques carrées multi-dimensionnelles, les grilles simples associées, et les sommes cartésiennes de graphes complets. La structure régulière de ces réseaux a permis une résolution de type algébrique. Ici encore nous avons montré l'égalité entre les paramètres \vec{w} et $\vec{\pi}$.

L'**annexe C** correspond à l'article [4] et montre, encore une fois, que le problème du routage optique se résout aussi bien que le problème de la charge pour l'échange total, cette fois-ci pour une classe de réseaux très utilisée dans les télécommunications, celle des arbres de cycles. Ce résultat possède en plus l'avantage de voir la définition de l'échange total généralisée à un réseau dont les nœuds sont pondérés, ce qui permet de considérer un trafic non uniforme dépendant du produit des poids des couples de nœuds communiquant.

L'**annexe D** correspond à l'article [5] qui étudie une modélisation des communications différente de celle abordée jusqu'à présent. Cette étude concernant la diffusion est valable aussi pour tous les modèles de réseaux d'interconnexion utilisant le mode de la commutation de circuit. L'hypothèse de travail est la réalisation d'une instance de requêtes en plusieurs étapes successives. Dans le cadre des réseaux optiques, une telle contrainte survient lorsqu'il n'est pas pos-

sible de réaliser l'instance donnée en une seule étape, par exemple à cause d'un nombre de longueurs d'onde insuffisant. On parle alors de réseaux *multihop*. Le coût d'un protocole réalisant un ensemble de requêtes peut alors être exprimé comme la somme des coûts des différentes étapes effectuées pour les satisfaire. Le coût d'une étape de communication est supposé dépendre de manière affine de la quantité d'information transmise, c'est-à-dire composé d'un coût fixe et d'un coût de transmission. Dans ce contexte, nous fournissons des protocoles de diffusion pour le graphe complet et nous obtenons des bornes inférieures correspondantes très fines, qui restent valables dans toute topologie. Il est ainsi observé une décroissance exponentielle du coût global de transmission lorsqu'un petit nombre d'étapes est ajouté par rapport au minimum, et une décroissance linéaire pour un très grand nombre d'étapes supplémentaires.

L'**annexe E** correspond à l'article [6] et met en application pour les réseaux optiques WDM les résultats obtenus dans l'annexe précédente. Sa lecture se suffit cependant à elle-même. Nous obtenons, sous un modèle de coût affine, des familles de protocoles quasi-optimaux pour la diffusion dans des réseaux optiques multihop, qui sont des cycles et des grilles toriques bidimensionnelles. Ces algorithmes exploitent le gain de connectivité fourni par le routage optique, de manière à transposer l'étude précédente relative au graphe complet, et apportent des performances significatives par rapport à la diffusion classique sur un arbre couvrant.

Pour terminer, nous présentons dans l'**annexe F**, qui correspond à l'article [7], des résultats issus d'une collaboration avec la société Alcatel Space Industries et qui sortent du cadre de l'étude des réseaux optiques. Ces travaux ont été motivés par la conception de réseaux de permutation tolérants aux pannes et destinés à être embarqués dans des satellites de télécommunications. Le problème général consiste à minimiser le nombre de commutateurs utilisés, à cause de leur coût prohibitif, pour réaliser un réseau de permutation réarrangeable. Nous avons ainsi généralisé la famille des réseaux de Waksman qui était définie et démontrée valide seulement pour les puissances de deux. Nous avons également considéré le cas d'un blocage quelconque de commutateur et fourni une construction pour le tolérer.

Bibliographie

- [1] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes et U. Vaccaro. Graph problems arising from wavelength-routing in all-optical networks. In *Proc. of IPPS'97, 2nd Workshop on Optics and Computer Science (WOCS'97)*, Geneva, Switzerland, avril 1997. IEEE Press.
- [2] B. Beauquier, P. Hell et S. Pérennes. Optimal wavelength-routed multicasting. *Discrete Applied Mathematics* (Elsevier Science), 84:15–20, 1998. Editors' Choice 1998.
- [3] B. Beauquier. All-to-all communication in some wavelength-routed all-optical networks. *Networks* (Wiley-Interscience), 33(3):179–187, mai 1999.
- [4] B. Beauquier, S. Pérennes et D. Tóth. All-to-all routing and coloring in weighted trees of rings. In *Proc. of the 11th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pages 185–190, Saint-Malo, France, juin 1999. ACM Press.
- [5] B. Beauquier, O. Delmas et S. Pérennes. Tight bounds for broadcasting in the linear cost model. Rapport de recherche Sloop, novembre 1999. En cours de dépôt.
- [6] B. Beauquier. Broadcasting in WDM optical rings and tori. In D-Z. Du P-J. Wan and P.M. Pardalos, editors, *Proc. the DIMACS Workshop : "Multichannel Optical Networks : Theory and Practice"*, volume 46 of *DIMACS Series on Discrete Mathematics and Theoretical Computer Science*, pages 63–73, Rutgers University, NJ, mars 1998. American Mathematical Society.
- [7] B. Beauquier and E. Darrot. Arbitrary size Waksman networks and their vulnerability. Soumis à *Parallel Processing Letters* (World Scientific). Version étendue de l'article paru dans *Actes des 1ères Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel'99)*, pages 95–100, Roscoff, France, mai 1999.

CHAPITRE 1

La technologie optique pour les télécommunications

1.1 Aperçu historique

Les premières liaisons servant à communiquer entre les continents furent les câbles télégraphiques, installés sous la mer depuis l'époque de la guerre de Sécession. Les câbles coaxiaux leur ont succédé pour acheminer les conversations téléphoniques. Le premier câble coaxial traversant l'Atlantique fut posé en 1955 et correspondait à 48 lignes téléphoniques. Depuis, quel chemin parcouru grâce aux fibres optiques ! Une seule paire de ces minces et longs fils de verre, qui propagent l'information sous forme d'impulsions de lumière, transmet simultanément près d'un demi-million de communications téléphoniques d'un continent à l'autre.

L'idée de fabriquer des fibres en verre de silice suffisamment pur pour transporter la lumière sur de grandes distances a fait son chemin dès la fin des années 1960. Le principe simplifié est le suivant : la lumière envoyée dans le cœur de la fibre se réfléchit sur ses parois, ce qui permet de guider le faisceau lumineux le long de la fibre, malgré sa courbure.

La concrétisation de cette idée a pris du temps. Il a fallu des innovations technologiques successives touchant à la fois le support matériel et la manière d'y faire circuler l'information. Il a fallu en particulier développer des sources laser miniatures (diodes laser) et des dispositifs de réception (photodiodes), ainsi que l'électronique numérique des circuits intégrés. Aussi, pour les communications à longue distance, les liaisons radio par satellite sont longtemps restées les plus utilisées et elles n'ont cédé le pas devant les câbles optiques qu'à la fin des années 1980. Mais aujourd'hui, l'essentiel des communications intercontinentales passe par des câbles optiques sous-marins déposés sur le fond des océans, qui tissent une véritable toile sur la planète. Les fibres optiques ont ainsi complète-

ment remplacé les câbles coaxiaux. Pour l'utilisateur, un signe tangible de cette mutation a été la disparition, en 1988, du temps mort de 0,4 seconde, dû à la liaison vers le satellite relais.

En quoi réside l'intérêt des liaisons optiques et quelle est la technologie nécessaire à leur utilisation ? Rappelons d'abord les principaux éléments d'une liaison optique. Elle comprend une source de lumière laser, qui est une diode laser semblable à celle des lecteurs de disques compacts, fonctionnant près de l'infrarouge (à des longueurs d'onde allant de 1,2 à 1,6 μm). La lumière émise est modulée par un transmetteur, système commandé par le signal électrique qui apporte l'information. Ces impulsions lumineuses sont envoyées dans la fibre et récupérées à l'autre extrémité par un récepteur (une photodiode) qui reconvertit le signal optique en signal électrique, lequel est enfin transformé en son, image ou texte, dans le téléphone, la télévision ou l'écran d'ordinateur.

Comme dans tout système de communication numérique, l'information est codée sous la forme d'une succession de *bits* (de *binary digit*, pour désigner 0 ou 1). Dans une fibre optique, les bits sont acheminés physiquement par une onde lumineuse dont on module l'intensité : le temps est divisé en créneaux de même durée, pendant lesquels le bit 1 est codé par une impulsion lumineuse et le bit 0 par une absence de lumière, comme pour la lecture des disques laser compacts.

Le téléphone standard fonctionne à 64 kilobits (64 mille bits) par seconde, la future télévision numérique à quelque 100 mégabits (cent millions de bits) par seconde. Dans les communications numériques à haut débit, on compte maintenant en gigabits par seconde (Gbit/s), soit un milliard de bits transmis par seconde sous forme d'impulsions lumineuses. Pour donner une idée, 1 Gbit/s représente environ quinze mille conversations téléphoniques simultanées.

Pourquoi le débit est-il limité ? N'oublions pas que les impulsions constituant les signaux sont découpées dans une onde lumineuse. Ce découpage n'a de sens que si chaque créneau contient au moins une période de l'onde. Plus formellement, la théorie des communications nous dit que le nombre d'informations transmises par seconde ne peut excéder la fréquence de l'onde porteuse (soit au plus un bit par période de l'onde). Cette propriété montre l'intérêt d'utiliser des signaux optiques, dont les fréquences vont de 10^{14} à 10^{15} Hertz (Hz), plutôt que les ondes radio de fréquences plus faibles (de l'ordre de 10^5 à 10^{11} Hz).

Un débit supérieur n'est pas le seul avantage de la lumière. Les fibres optiques présentent des pertes bien moindres que les câbles en cuivre utilisés pour les transmissions électriques, jusqu'à certaines distances. Prenons un exemple : pour des débits de 1 Gbit/s, la fréquence de l'onde porteuse doit être au moins de

l'ordre du gigahertz (10^9 Hz). Or dans cette gamme, un câble coaxial perd 99,9% de l'énergie sur 1 kilomètre, et plus encore aux fréquences plus élevées. Ce support ne sied que pour des signaux de plus basse fréquence (1-100 MHz), donc de faibles débits. À l'inverse, les fibres de verre présentent à la fréquence de 2×10^{14} Hz une transmission de 95% par kilomètre, ce qui autorise une très longue portée. L'ampleur et la rapidité des progrès apparaissent clairement quand on sait que la première fibre développée par la société Corning en 1970 perdait également 99% de la lumière au bout d'un kilomètre seulement, à comparer avec la même perte pour 100 kilomètres actuellement.

En dépit de ces performances remarquables, le signal finit par s'affaiblir après une certaine distance et se rapproche trop du bruit de fond des détecteurs. Au-dessous d'un certain seuil d'énergie, le récepteur devient incapable de distinguer les bits sans ambiguïté. Les standards habituels des télécommunications imposent un maximum d'une erreur par milliard de bits reçus, et cela correspond grossièrement à 500 photons par bit. Dans les communications à fort débit, on atteint vite cette limite. La puissance optique d'une diode laser est de quelques milliwatts, ce qui représente quelque 10^{16} photons par seconde. À un taux de transmission de 1 Gbit/s, il y a donc 10^7 photons par bit. On calcule facilement qu'avec une transparence de 95% par kilomètre, le seuil de 500 photons par bit est franchi au bout d'environ 200 kilomètres.

Pour étendre la portée des liaisons sur plusieurs centaines de kilomètres sans dégrader le taux d'erreur, il suffit de maintenir le nombre de photons par bit au-dessus de la valeur fatidique au moyen de dispositifs qui régénèrent le signal affaibli. Des "répéteurs" optoélectroniques sont arrivés sur le marché dès les années 1980. Ils comportent une photodiode (puce en matériau semi-conducteur mise sous tension) qui convertit le signal optique en signal électrique, un amplificateur, une bascule électronique qui identifie les bits, et finalement une diode laser commandée par le signal électrique résultant pour la conversion inverse.

C'est en 1988 que le tout premier système de ce type a été installé entre la France, l'Angleterre et les États-Unis. Il est constitué d'un câble sous-marin d'une longueur de 7 500 km, doté d'environ 110 répéteurs placés tous les 70 km, et il opère à 0,28 Gbits/s, soit l'équivalent de quarante mille lignes téléphoniques. Mais pour les distances transocéaniques, le débit maximal de ce type de liaisons avec répéteurs optoélectroniques était limité non par la fibre elle-même, mais par la rapidité de l'électronique. En effet, à des fréquences de fonctionnement supérieures au gigahertz, les circuits électroniques deviennent coûteux et leur fiabilité diminue.

Cet obstacle a été contourné à la fin des années 1980 grâce à l'apparition révolutionnaire des amplificateurs optiques à fibre de verre dopée à l'erbium [Gla93,

Des94]. Un développement qui fut rapidement suivi en 1995 par leur mise en service dans le domaine des câbles sous-marins. Un amplificateur optique fonctionne sur le principe de l'émission stimulée comme un laser. Les atomes d'erbium contenus dans la fibre sont excités par une diode laser auxiliaire qui les porte vers un état d'énergie supérieur, énergie qu'ils peuvent céder en retour pour amplifier le signal affaibli passant dans la fibre. Le signal optique voit ainsi sa puissance multipliée par un facteur allant de 100 à 10 000.

Dans les télécommunications optiques, l'intérêt d'utiliser de tels amplificateurs optiques plutôt qu'électroniques est énorme. Tout d'abord, ils se raccordent par simple soudure aux fibres de transmission. Mais surtout, ils évitent les conversions optoélectroniques effectuées par les répéteurs et donc la limitation du débit associée à ces derniers. La gamme de fréquences qu'accepte l'amplificateur optique s'étend sur plusieurs téraHertz (1 THz = mille milliards de Hertz), ce qui englobe très largement le signal à amplifier. Autre propriété, dont nous verrons l'intérêt par la suite : plusieurs canaux optiques de longueurs d'onde différentes peuvent être amplifiés simultanément.

En sus du pompage efficace par diode laser, la fibre dopée à l'erbium possède deux qualités supplémentaires. D'abord, le gain - c'est-à-dire le rapport d'amplification - est très peu sensible à la polarisation du signal incident (la polarisation caractérise la direction dans laquelle vibre le champ électrique associé à l'onde lumineuse). C'est un atout essentiel, car l'état de polarisation des signaux se modifie de façon aléatoire au fur et à mesure qu'ils se propagent dans la fibre.

Ensuite, l'amplificateur ne déforme pas les signaux, il les amplifie à l'identique (y compris le bruit engendré par l'émission spontanée qui peut être limité par le placement d'un filtre optique en sortie). Cette propriété subsiste dans des conditions de fonctionnement extrêmes. Par exemple, avec une puissance d'entrée trop élevée, le gain diminue mais le signal ne subit pas de distorsion, contrairement au cas des amplificateurs électroniques.

Aux trois gros atouts de l'amplificateur à fibre dopée à l'erbium (pompage efficace, insensibilité à la polarisation, absence de distorsion) s'ajoutent la compatibilité avec les fibres standards, la faiblesse des pertes d'énergie dans les connexions, le bruit minimal, l'insensibilité à la température. Selon les applications, la plage de gain exploitable autour de la longueur d'onde $1,5 \mu\text{m}$ s'étale sur 100 à 3 000 GHz.

La course vers les hauts débits a également bénéficié du développement des techniques optiques de **multiplexage** et de **commutation**. Le multiplexage consiste à transporter sur un même support physique plusieurs signaux. La commutation est une opération de routage au niveau du réseau global qui permet d'acheminer les signaux de chaque émetteur vers chaque destinataire.

Le multiplexage dit temporel demeure très utilisé. Il consiste à imbriquer temporellement les différents flots d'information en trames successives. On peut réaliser électroniquement les fonctions de multiplexage/démultiplexage temporels avec des circuits intégrés ultra-rapides (40 Gbits/s en laboratoire). Toutefois, le coût prohibitif de ces circuits pour les très hauts débits suggère d'effectuer le multiplexage temporel par des moyens purement optiques, une voie actuellement explorée.

Dans les générations de câbles optiques en développement, l'accroissement du débit est assuré par une technique qui se superpose à la première : le multiplexage en longueur d'onde, connu sous l'appellation anglaise WDM (Wavelength-Division Multiplexing). Celui-ci consiste à envoyer plusieurs signaux de longueurs d'onde différentes simultanément dans la même fibre optique. Multiplexage et démultiplexage en longueur d'onde sont effectués par des composants optiques passifs, de façon similaire à la décomposition et recombinaison des couleurs de l'arc-en-ciel par un prisme. La technique WDM ouvre également des perspectives de routage optique dans les réseaux. Les communications peuvent être ainsi aiguillées dans telle ou telle direction suivant leur longueur d'onde.

Deux questions majeures se posent : quelle est la portée maximale des liaisons amplifiées ? Jusqu'à quel débit d'information peut-on aller ? Car lorsqu'on atteint des débits de plus de 10 Gbits/s sur des milliers de kilomètres, d'autres difficultés majeures d'ordre physique se font jour, en particulier la dispersion chromatique et les effets non-linéaires (variation de la vitesse de propagation liée à la longueur d'onde d'une part, liée à l'intensité d'autre part). Le **soliton** (impulsion très brève, cf. [LT76]) permet d'échapper à ce dilemme en conjuguant idéalement les deux types de dispersion de manière à les neutraliser, ceci en calculant judicieusement son intensité et sa longueur d'onde. Le soliton excite l'imagination des ingénieurs du monde des télécommunications. Mais son utilisation suppose certains développements techniques pointus, relatifs aux dispositifs d'émission et au filtrage en réception.

Les techniques progressant continuellement, il est hasardeux de faire un pronostic sur les capacités ultimes des systèmes de communication optiques. À chaque record de transmission, de nouveaux effets limitants sont découverts, et de nouvelles parades mises en œuvre pour chacun. Il n'en reste pas moins que les débits faramineux de l'ordre du Tbits/s déjà démontrés en laboratoire, équivalent à plusieurs dizaines millions de connexions téléphoniques simultanées dans une seule fibre. De tels débits ouvrent des perspectives pratiquement illimitées aux "autoroutes de l'information".

1.2 Systèmes de transmission optiques

Un système de transmission optique élémentaire est constitué de trois parties, comme il est montré sur la figure 1.1 : un émetteur, une fibre optique et un récepteur. Comme plusieurs livres excellents traitent en détails des systèmes de transmission optiques, par exemple [JJ96, Gre93, Agr92, vEvdP91], nous ne décrivons dans cette section que les concepts de base.

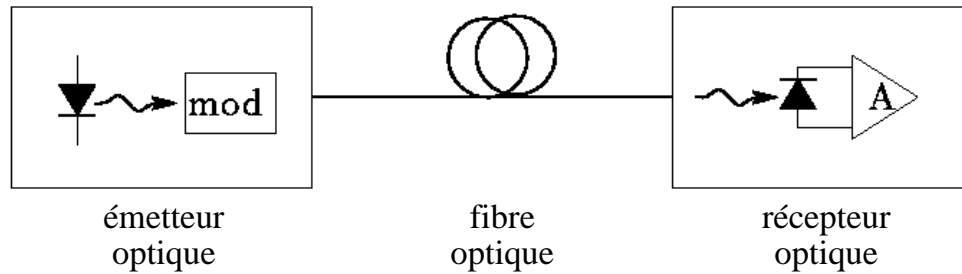


Figure 1.1: Système de transmission optique élémentaire.

1.2.1 Fibre optique

La fibre optique possède un grand nombre de propriétés remarquables qui en font un support physique excellent pour les télécommunications. La qualité d'un milieu physique pour la transmission de signaux se fonde sur deux principaux facteurs qui sont l'**atténuation** et la **dispersion**.

Atténuation

L'atténuation représente la réduction de la puissance du signal au cours de sa propagation. Ce facteur est déterminant pour connaître la distance maximale que peut parcourir un signal, étant données la puissance d'émission et la sensibilité en réception.

Notons $P(D)$ la puissance de l'impulsion lumineuse dans une fibre optique à la distance de D km de l'émetteur, P_r la sensibilité du récepteur (puissance minimale requise pour détecter le signal) et A le facteur d'atténuation de la fibre. D'après [Hen85], l'atténuation est caractérisée par

$$P(D) = 10^{-AD/10} P(0)$$

Pour une longueur de fibre de D km, la puissance $P(D)$ doit être au moins égale à la sensibilité P_r , d'où nous déduisons

$$D_{\max} = \frac{10}{A} \log_{10} \frac{P(0)}{P_r}$$

La distance maximale entre l'émetteur et le récepteur (ou entre les amplificateurs) dépend donc bien davantage du facteur d'atténuation que de la puissance d'émission ou de la sensibilité en réception.

La figure 1.2(a) montre une courbe typique de l'atténuation (exprimée en décibels par kilomètre (dB/km)) de la puissance du signal lumineux en fonction de sa longueur d'onde (exprimée en μm) dans une fibre optique utilisée en régime monomode. La bande passante spectrale est d'environ 25 terahertz (THz) autour de la longueur d'onde $1.3 \mu\text{m}$ pour une atténuation inférieure à 0.5 dB/km , et d'autant autour de la longueur d'onde $1.5 \mu\text{m}$ pour une atténuation inférieure à 0.2 dB/km . Ces valeurs sont deux ordres de grandeur plus faibles en atténuation et quatre ordres de grandeur plus élevés en bande passante que celles des câbles coaxiaux.

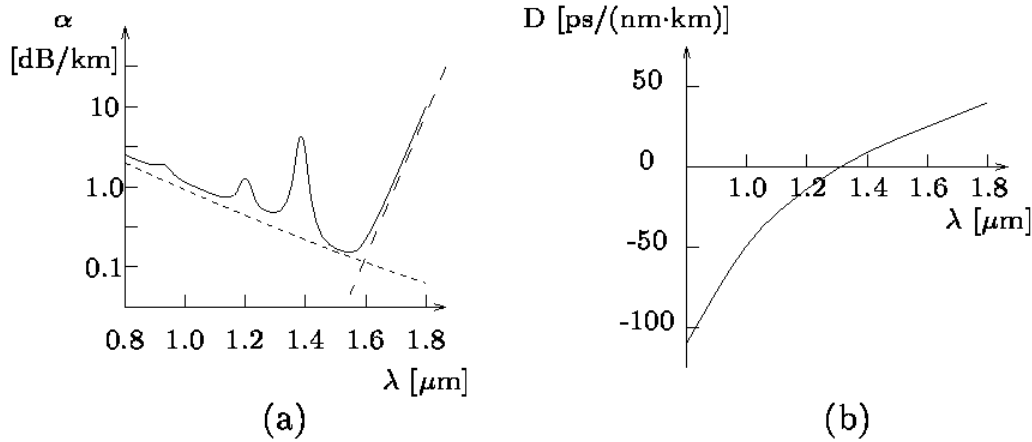


Figure 1.2: Courbes typiques de l'atténuation (a) et de la dispersion (b) en fonction de la longueur d'onde pour une fibre optique en mode simple.

Dispersion

La dispersion représente la dépendance de la constante de propagation par rapport à la fréquence. Elle permet de calculer l'élargissement de la durée des impulsions au cours de leur propagation. Pour des signaux lumineux, on parle de dispersion **chromatique**, ce phénomène étant responsable de la décomposition de la lumière blanche par un prisme de verre.

En effet, la vitesse de propagation de la lumière dans la matière transparente, définie par l'indice de réfraction optique, est fonction de la longueur d'onde. Or une impulsion lumineuse dans une fibre optique n'est pas parfaitement monochromatique, puisqu'un laser ne transmet pas sur une fréquence unique, et puisqu'un

signal transportant de l'information a une largeur spectrale non nulle. Par conséquent, les différentes longueurs d'onde constituant le signal lumineux vont se propager à des vitesses différentes, ce qui entraîne l'élargissement temporel des impulsions qui peuvent alors se chevaucher, provoquant des erreurs à la détection. En outre, plus une impulsion est brève, plus sa gamme de fréquences est étendue. Aussi la dispersion chromatique est un facteur d'autant plus limitatif que les débits sont élevés, car les impulsions sont alors très brèves et proches les unes des autres dans le temps.

La figure 1.2(b) montre une courbe typique de la dispersion chromatique (exprimée en picosecondes par nanomètre par kilomètre (ps/(nm.km))) du signal lumineux par rapport à sa longueur d'onde (exprimée en μm) dans une fibre optique standard (G.652) utilisée en mode simple. La dispersion est quasiment nulle autour de la longueur d'onde $1.3 \mu\text{m}$ et conserve des valeurs faibles autour de la longueur d'onde $1.5 \mu\text{m}$. Par chance, la fibre optique combine donc dans les mêmes zones de fréquences ses très bonnes performances en dispersion et en atténuation.

1.2.2 Émetteurs optiques

Dans un système de transmission optique, un émetteur a deux fonctions primordiales : la génération d'un signal optique et la modulation de ce signal par l'information à émettre (voir Figure 1.1). Une qualité supplémentaire très utile d'un émetteur optique est sa capacité d'être **accordable en fréquence**.

Principe du laser

Le mot **laser** est l'acronyme de *Light Amplification by Stimulated Emission of Radiation*, signifiant littéralement "amplification de lumière par émission stimulée de rayonnement". Le but est de produire un faisceau très intense de lumière cohérente monochromatique.

Pour comprendre le principe du laser et de l'émission stimulée, nous devons rappeler quelques notions sur les niveaux d'énergie des particules élémentaires. Un atome est dit *stable* lorsque ses électrons sont dans les niveaux d'énergie les plus faibles. Pour chaque type d'atome, les électrons ont un nombre fini de niveaux d'énergie possibles, appelés états. Lorsqu'un atome absorbe de l'énergie, ces électrons sont portés dans des états d'énergie plus élevés. L'atome devient alors instable et généralement il retourne dans un état inférieur rapidement en libérant un photon. Cependant, certaines molécules sont dites *métastables* car elles peuvent demeurer dans un état excité plus longtemps. Cette propriété permet l'émission stimulée comme nous allons le voir immédiatement.

La figure 1.3 montre une représentation schématique de la structure d'un laser. Il est constitué de deux miroirs qui forment une cavité contenant une substance quasi-stable. L'excitation des électrons de cette substance par apport d'énergie (par exemple par un courant électrique) provoque dans un premier temps l'émission de photons, qui vont ensuite se réfléchir sur les miroirs aux extrémités de la cavité.

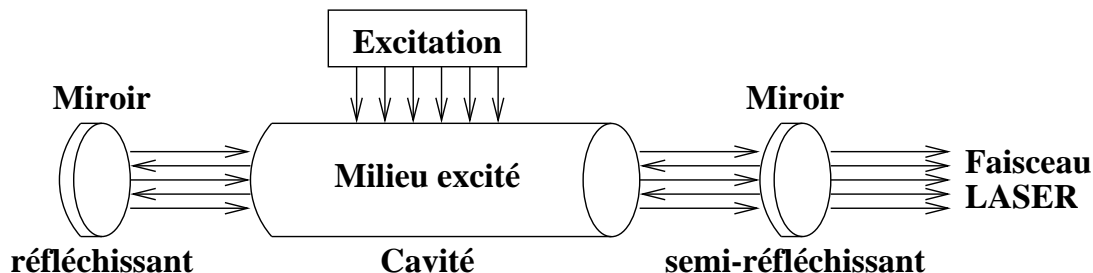


Figure 1.3: La structure générale d'un laser.

L'émission stimulée se produit lorsqu'un photon rencontre un électron excité. Cette probabilité est d'autant plus forte dans un milieu quasi-stable, puisque les électrons excités peuvent y devenir majoritaires si suffisamment d'énergie est apportée. L'électron peut alors retourner dans un état inférieur et libérer son énergie sous la forme d'un autre photon qui possède la même direction et la même cohérence que le photon stimulant. Si la longueur de la cavité est multiple entier de la demi-longueur d'onde des photons, ceux-ci vont pouvoir se combiner de façon cohérente et se dupliquer, de manière à produire un faisceau de plus en plus intense à une fréquence précise. Ce faisceau laser est récupéré en utilisant un miroir semi-réfléchissant à l'une des extrémités de la cavité, et sa fréquence peut être ajustée en modifiant la taille de la cavité. On peut également accorder un laser plus rapidement en modifiant l'indice optique du milieu excité.

Modulation

La **modulation** est le procédé qui permet au laser de transmettre de l'information dans une fibre optique, en faisant varier certaines caractéristiques du signal lumineux émis, en particulier son amplitude, sa fréquence ou sa phase. La **fréquence de modulation** détermine le débit de données qui peut être transmis. On parle de modulation **directe** ou **externe**, selon que le laser lui-même ou un composant extérieur (modulateur) fait varier le signal.

Une étude détaillée de la modulation directe peut être trouvée dans [Pet88]. Les **lasers à semi-conducteurs** permettent des fréquences de modulation directe jusqu'à quelques dizaines de gigahertz, mais cette modulation est limitée en pratique par des effets de dispersion. Pour des débits supérieurs à quelques gigabits

par seconde, la modulation externe est préférable, notamment à l'aide de modulateurs par **électro-absorption** (voir [Woo88] pour une présentation exhaustive).

Accordabilité

Il existe plusieurs techniques pour accorder dynamiquement la longueur d'onde d'émission des signaux laser. Les deux paramètres importants sont le temps de réglage sur une fréquence donnée et la largeur de la bande de fréquences accessible. Sans entrer dans les détails techniques, nous donnons dans le tableau 1.1 les performances actuelles, susceptibles d'évoluer, des différents émetteurs laser accordables communément utilisés [Bra90, Muk92]. Notons l'existence d'un compromis entre les deux paramètres d'accordabilité.

<i>Laser accordable</i>	<i>Bande de fréquences (nm)</i>	<i>Temps de réglage</i>
Mécanique	500	1–10 ms
Acousto-optique	83	$\sim 10 \mu s$
Electro-optique	7	1–10 ns (estimé)
Injection DBR ¹	10	1–10 ns

Tableau 1.1: Caractéristiques des émetteurs laser.

1.2.3 Récepteurs optiques

La fonction d'un récepteur dans un système de transmission optique est de détecter et de démoduler un signal lumineux transmis sur une fibre.

La **détection** consiste en la conversion du signal optique en signal électrique. La **démodulation** est généralement accomplie ensuite par les techniques habituelles des systèmes de transmission électriques.

Détection

Deux techniques de détection sont utilisées. La détection **directe** est réalisée par une diode photo-sensible qui convertit un flot de photons en un flot d'électrons. Le courant électrique résultant est ensuite amplifié puis soumis à un test de seuil pour déterminer si l'information logique correspondante est un bit 0 ou 1.

Une alternative est la détection **cohérente** qui utilise un laser auxiliaire comme oscillateur local. Une photodiode reçoit alors un signal issu de la combinaison des deux signaux laser, qui est plus facile à détecter et qui est compatible avec les trois types de modulation (en amplitude, fréquence et phase). Ce système

¹réflecteur de Bragg distribué (*Distributed Bragg Reflector*).

est néanmoins plus complexe et beaucoup plus coûteux. L'avantage de la détection cohérente est la réception de signaux faibles, rendue difficile par le bruit de fond, l'inconvénient étant sa complexité de mise en place. Nous renvoyons à [Agr92, LG88] pour davantage d'information sur la détection cohérente.

Accordabilité

Comme les émetteurs laser, les récepteurs optiques sont susceptibles d'être accordables en fréquence de réception. Une documentation sur les différentes technologies peut être trouvée dans [Gre93] et [Bra90]. Le tableau 1.2 donne les caractéristiques des récepteurs accordables communément utilisés, en termes de bande de fréquences et de temps de réglage. Ici encore, les deux paramètres de réglage sont inversement proportionnels.

<i>Récepteur accordable</i>	<i>Bande de fréquences (nm)</i>	<i>Temps de réglage</i>
Fabry-Perot	500	1–10 ms
Acousto-optique	250	~10 μ s
Electro-optique	16	1–10 ns

Tableau 1.2: Caractéristiques des récepteurs optiques.

1.2.4 Amplificateurs optiques

Bien que la figure 1.1 ne mentionne pas la présence d'amplificateurs dans un système de transmission optique élémentaire, ceux-ci sont nécessaires lorsque les signaux optiques parcourent de longues distances dans la fibre optique, à cause du phénomène d'atténuation. Ils sont également utiles dans les réseaux optiques locaux pour restaurer la qualité des signaux affectée par la dispersion et le bruit.

Les deux principaux paramètres qui déterminent la qualité d'un amplificateur optique sont : le **gain**, exprimé en décibels (dB), qui mesure le rapport entre la puissance du signal entrant et celle du signal sortant, et la **largeur de bande** des fréquences amplifiables. Notons que les amplificateurs peuvent être également caractérisés par leur **gain efficace**, exprimé en décibels par milliwatt (dB/mW), qui mesure le rapport entre le gain et la puissance utilisée pour l'amplification.

L'amplification optique repose sur le principe de l'émission stimulée, comme pour les émetteurs laser. Les deux grands types de composants sont les **amplificateurs à semi-conducteur** et les **amplificateurs à fibre dopée**. Une présentation générale des amplificateurs optiques se trouve dans [O'M93].

Les amplificateurs à semi-conducteur sont constitués d'un laser à semi-conducteur, modifié en supprimant la cavité résonante. Le signal lumineux d'entrée traverse la région active du semi-conducteur qui, par émission stimulée, l'amplifie.

Les amplificateurs à fibre dopée sont composés d'une pompe laser, d'un coupleur, d'une portion de fibre dopée avec des ions de terre rare, et d'un filtre, comme il est représenté sur la figure 1.4. La pompe laser émet un signal puissant qui excite les ions dopants, afin que le signal transmettant les données puisse agir sur les ions excités par le principe de l'émission stimulée. Le bruit engendré par l'émission spontanée peut être limité par le placement d'un filtre optique en sortie.

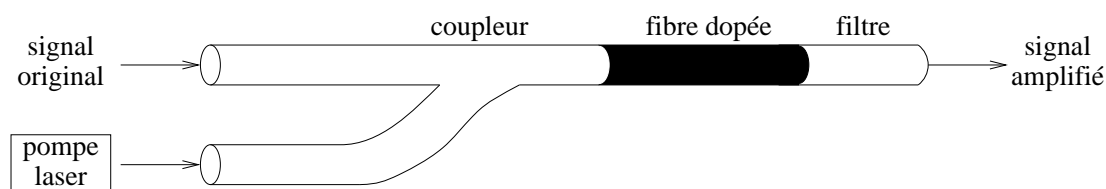


Figure 1.4: Amplificateur à fibre dopée.

L'élément dopant de terre rare le plus utilisé est l'**erbium**, qui permet d'amplifier les longueurs d'onde entre 1525 nm et 1560 nm. Les amplificateurs à fibre dopée à l'erbium (AFDE) ont des gains de l'ordre de 25 dB, et même 51 dB expérimentalement. Les gains efficaces varient entre 5 et 10 dB/mW. La puissance requise de ces pompes laser reste donc très faible, quoique supérieure à celle des autres lasers d'émission (250 mW contre 1 à 2 mW).

Pour amplifier les signaux autour de la longueur d'onde 1300 nm, les amplificateurs à fibre fluorée dopée au praseodymium (AFFDP) ont récemment retenu l'attention des chercheurs. Les derniers développements sont présentés dans [Whi95].

<i>Amplificateur</i>	<i>Fréquences</i>	<i>Largeur de bande</i>	<i>Gain</i>
Semi-conducteur	Toutes	40 nm	25 dB
AFDE	1525–1560 nm	35 nm	25–51 dB
AFFDP	1280–1330 nm	50 nm	20–40 dB

Tableau 1.3: Caractéristiques des amplificateurs optiques.

1.3 Multiplexage optique

Cette partie répond à la question de savoir comment exploiter les dizaines de térahertz de bande passante spectrale disponibles dans la fibre optique. La technique générale utilisée est le multiplexage dans le domaine optique, ce qui signifie que la capacité de la fibre est divisée par des moyens optiques en plusieurs canaux accessibles individuellement et indépendamment.

La division de la bande passante en canaux peut être réalisée, comme en électronique, dans la dimension temporelle ou dans la dimension des fréquences (ou longueurs d'onde). Dans le premier cas, on parle de multiplexage temporel (*Time-Division Multiplexing, TDM*) et dans le second de multiplexage en longueur d'onde (*Wavelength-Division Multiplexing, WDM*).

1.3.1 Multiplexage temporel (TDM)

Le multiplexage TDM consiste à imbriquer temporellement différents canaux de communication en trames successives. Si l'on se représente un flot d'information par les dents d'un peigne, le multiplexage temporel revient à superposer les peignes des différents canaux en les décalant les uns par rapport aux autres (voir figure 1.5). Cela nécessite une synchronisation précise. À la réception, chaque canal temporel est démultiplexé puis acheminé vers sa destination. On peut réaliser électriquement les fonctions de multiplexage/démultiplexage temporels avec des circuits intégrés ultra-rapides (40 Gbits/s en laboratoire). Toutefois, le coût prohibitif de ces circuits pour les très hauts débits suggère d'effectuer le multiplexage temporel par des moyens purement optiques, une voie actuellement explorée.

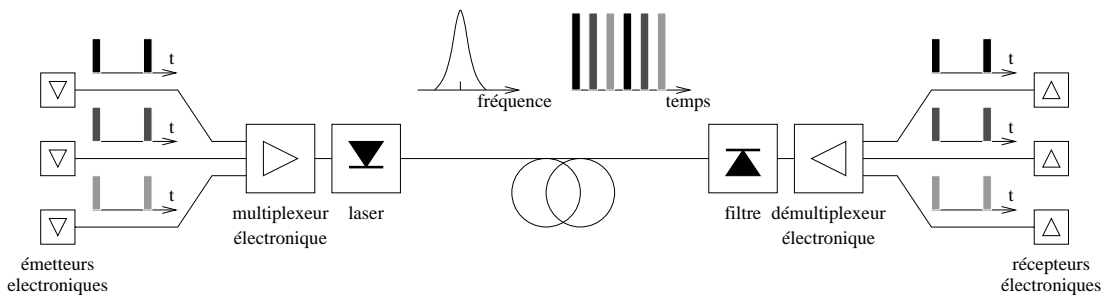


Figure 1.5: Multiplexage temporel (TDM).

1.3.2 Multiplexage en longueur d'onde (WDM)

Le multiplexage en longueur d'onde (*Wavelength-Division Multiplexing, WDM*) est une technique de multiplexage qui peut se superposer à la précédente. Le principe est le même que celui du multiplexage en fréquences (*Frequency-Division Multiplexing, FDM*) dans les transmissions électriques. Plusieurs signaux sont générés simultanément sur des fréquences optiques (ou longueurs d'onde) différentes et peuvent être modulés individuellement.

Le multiplexage WDM est une approche qui permet de gérer l'importante discordance opto-électronique en termes de bande passante. Le spectre optique d'une fibre est découpé en un certain nombre d'intervalles de longueurs d'onde (ou fréquences) distincts, dans les régions de faible atténuation (voir figure 1.2), de telle sorte que chaque intervalle supporte un canal de communication transmettant au débit désiré, par exemple celui de l'électronique. Ainsi, en autorisant la coexistence de plusieurs canaux WDM sur une même fibre, l'énorme bande passante optique peut être exploitée, en permettant aux équipements d'accéder au réseau au débit de l'électronique seulement.

Les canaux WDM sont accessibles par des émetteurs laser réglés sur des longueurs d'onde spécifiques. Le multiplexage et le démultiplexage en longueur d'onde sont effectués par des composants optiques passifs, de façon similaire à la décomposition et recombinaison des couleurs de l'arc-en-ciel par un prisme.

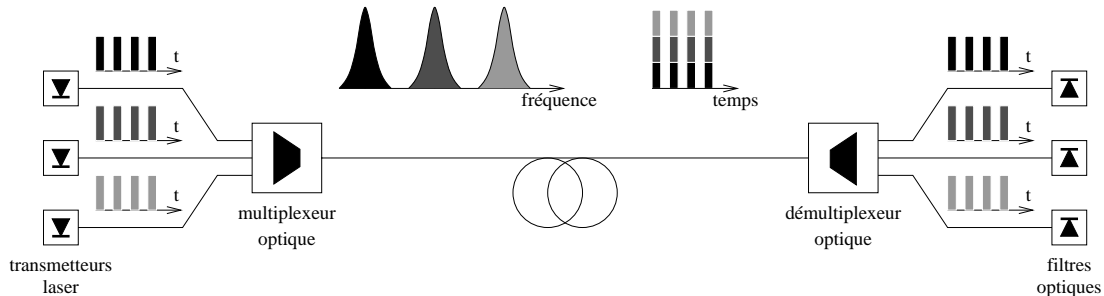


Figure 1.6: Multiplexage en longueur d'onde (WDM).

Canaux WDM

Un facteur important dans la conception des réseaux optiques WDM est le nombre de longueurs d'onde utilisables, qui est limité par la technologie des composants.

La bande passante de la fibre optique, comme il est mentionné dans la section 1.2.1, est limitée par les régions de faible atténuation autour des longueurs

d'onde $1.3 \mu\text{m}$ et $1.5 \mu\text{m}$. Ces régions ont une bande passante d'environ 25 THz chacune. Cependant, les réseaux optiques ne vont pas pouvoir bénéficier de toute cette largeur de bande en raison de la limitation des composants optiques et de l'espacement des canaux WDM.

Les amplificateurs opèrent sur des bandes de fréquences de 35 nm à 40 nm, les émetteurs laser réglables sur 10 nm, et les filtres réglables ne couvrent pas toujours toute la gamme des fréquences disponibles (16 nm pour les filtres électro-optiques).

Par ailleurs, le nombre de canaux utilisables dépend fortement de leur espacement. D'après [Bra90], l'espacement des canaux doit valoir au moins 6 fois leur bande passante pour éviter les interférences. En supposant une efficacité en modulation de 0.5 bps/Hz (ce qui signifie qu'un bit est codé sur deux intervalles de temps), il en résulte un facteur 8 entre la bande des fréquences utilisées et le débit d'information qui va pouvoir circuler sur les canaux.

Un nombre élevé de canaux fournit au réseau davantage de capacité, ainsi qu'un coût plus élevé des composants et une plus grande complexité des mécanismes de gestion.

1.3.3 Comparaison entre TDM et WDM

Au premier abord, on pourrait penser que les deux approches de multiplexage optique TDM et WDM sont semblables. Elles le sont au niveau formel puisqu'elles permettent la superposition sur le même support physique de transmission de plusieurs canaux de communication, identifiables selon leur décalage temporel pour TDM et selon leur longueur d'onde pour WDM. Cependant, au niveau technologique l'approche TDM présente des inconvénients significatifs par rapport à l'approche WDM.

En premier lieu, les canaux TDM ne sont pas transparents pour le débit de modulation ni pour le type de modulation. Ils sont uniquement modulables numériquement en amplitude et leur débit est imposé par le multiplexeur temporel. À l'inverse, chaque canal WDM peut être modulé individuellement, numériquement ou analogiquement, en amplitude ou en phase. Le débit de chaque canal WDM peut en outre être choisi arbitrairement, du moment que les signaux ne se recouvrent pas spectralement. Il en résulte une plus grande flexibilité.

Une autre caractéristique désavantageuse du multiplexage optique TDM provient du très haut débit du signal multiplexé, qui résulte de l'agrégation des canaux entrelacés temporellement. Cette conséquence inhérente au multiplexage temporel constitue un inconvénient majeur pour les systèmes de transmission optiques, lorsque le signal multiplexé se met à couvrir des dizaines de gigahertz. Le traitement électronique constitue alors un frein aux opérations de multiplexage

et démultiplexage temporels, alors que l'approche WDM effectue celles-ci optiquement et passivement. De plus, le phénomène de dispersion limite d'autant plus la propagation du signal TDM multiplexé que son débit est important. Les débits moindres de chaque canal WDM permettent d'éviter ce problème.

Enfin, l'approche TDM souffre d'un manque d'extensibilité. L'addition d'un nouveau canal TDM nécessite une modification des décalages temporels et une resynchronisation de tous les canaux déjà existants. Inversement, l'indépendance des canaux WDM autorise la création d'un nouveau canal simplement par l'ajout d'un émetteur laser et d'un filtre optique appropriés, sans affecter les autres canaux.

En conclusion, l'approche WDM s'avère la technique de multiplexage préférentielle pour les systèmes de transmission optiques, en raison de la transparence, de la flexibilité et de l'extensibilité des canaux WDM.

1.4 Commutation optique

La plupart des réseaux actuels procèdent électroniquement au traitement des données et utilisent la fibre optique seulement comme un support de transmission. Les opérations de commutation et de routage sont réalisées en convertissant les signaux optiques sous leur forme électronique originelle. Il en résulte une grande flexibilité, cependant le débit de l'électronique n'est pas compatible avec l'énorme bande passante de la fibre optique.

De plus, la conversion opto-électronique introduit un délai supplémentaire dans le routage des flots d'information. Pour s'affranchir de ces limitations des composants de commutation optiques ont été développés, capables de traiter des signaux optiques à très haut débit sans conversion opto-électronique.

Pour cette raison, les réseaux utilisant cette technologie de commutation sont appelés **réseaux tout-optiques**. Dans l'état actuel, le contrôle des fonctions de commutation demeure cependant réalisé de façon électronique.

1.4.1 Générations de réseaux

Au vu de leur évolution, on peut classer les réseaux de communication câblés en trois générations, selon le rôle des fibres optiques dans leur topologie. Cette classification a été établie dans [Gre91].

La première génération de réseaux ne fait aucune utilisation de la fibre optique. Les réseaux locaux traditionnels LAN (*Local Area Network*), comme Ethernet, tombent dans cette catégorie, de même que certains grands réseaux WAN (*Wide Area Network*) comme ARPAnet. Ils ont été conçus sur la base d'un milieu de

transmission en cuivre plus ou moins fiable. Une caractéristique principale de ces réseaux est le contrôle en chaque nœud des flots d'information afin de réaliser leur commutation.

Dans la deuxième génération de réseaux, les transmissions électriques des données sont remplacées par des systèmes de transmission optiques. L'architecture traditionnelle des réseaux est cependant conservée. Les avantages immédiats sont un débit beaucoup plus élevé, des distances de transmission plus grandes et une fiabilité accrue. Il en a résulté l'émergence de réseaux métropolitains MAN (*Metropolitan Area Network*), ainsi que les premiers réseaux SONET (*Synchronous Optical NETWORK*) et SDH (*Synchronous Digital Hierarchy*).

La troisième et dernière génération de réseaux permet de profiter pleinement des propriétés uniques de la transmission optique par fibre. L'exceptionnelle bande passante spectrale peut en effet être exploitée par la technique du multiplexage en longueur d'onde WDM que nous avons évoquée dans la section 1.3.2. De plus, non seulement les fonctions d'émission et de réception sont réalisées par des composants optiques (diodes laser et photo-diodes), mais également les fonctions de commutation en chaque nœud du réseau.

Pour cette raison, les réseaux de la troisième génération sont appelés **réseaux tout-optiques**. Nous allons décrire à présent les différents composants optiques de commutation qui sont à l'origine de leur avènement.

1.4.2 Commutateurs optiques

Répartiteur

Un répartiteur (*Fiber Optical Cross-Connect, F-OXC*) permet d'effectuer une fonction de commutation entre les fibres d'entrée et de sortie du routeur (voir figure 1.7). C'est le plus élémentaire des commutateurs optiques et le moins coûteux en fabrication. Ce type d'équipements ne permet pas d'effectuer les opérations de démultiplexage sur les signaux entrants ni de multiplexage sur les signaux sortants. C'est pourquoi le répartiteur est dit *insensible aux longueurs d'onde*. Les fonctions d'extraction et d'insertion de signaux avec le terminal relié au routeur sont possibles, mais c'est alors l'ensemble des canaux WDM multiplexés qui est extrait d'une fibre d'entrée ou inséré dans une fibre de sortie.

Multiplexeur à insertion/extraction (MIE)

Un multiplexeur à insertion/extraction (*Optical Add/Drop Multiplexer, OADM*) permet d'extraire certains canaux WDM en transit sur une fibre optique et d'en insérer d'autres. Il est généralement constitué d'un démultiplexeur optique passif, de commutateurs 2×2 dédiés aux différentes longueurs d'onde et d'un mul-

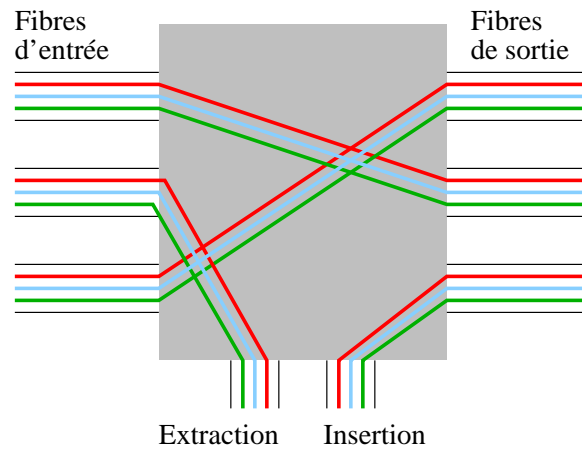


Figure 1.7: Répartiteur (F-OXC).

tiplexeur optique passif, comme il est montré schématiquement sur la figure 1.8. Les états des commutateurs intermédiaires contrôlés électroniquement déterminent quels canaux WDM poursuivent leur chemin, lesquels sont extraits en vue de leur réception locale et lesquels peuvent être insérés après émission locale.

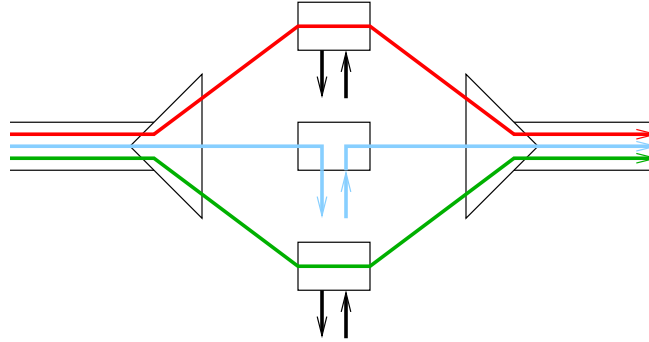


Figure 1.8: Multiplexeur à insertion/extraction (MIE).

Le composant MIE peut être greffé directement sur une fibre optique, ou faire partie d'un nœud de routage plus complexe. En particulier, associés avec un répartiteur, plusieurs MIE peuvent former un routeur à MIE, comme il est représenté sur la figure 1.9.

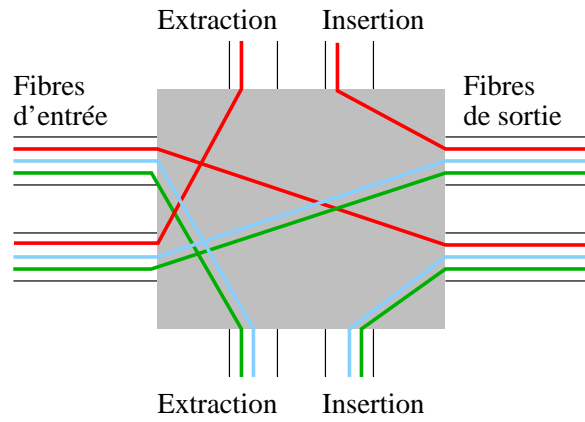


Figure 1.9: Répartiteur avec MIE.

Brasseur

Un brasseur (*Wavelength Routing Optical Cross-Connect*, *WR-OXC*) est un composant de commutation **sélectif en longueur d'onde**. Cela signifie que chaque canal WDM peut être dirigé vers une fibre de sortie indépendamment des autres canaux multiplexés sur la même fibre d'entrée.

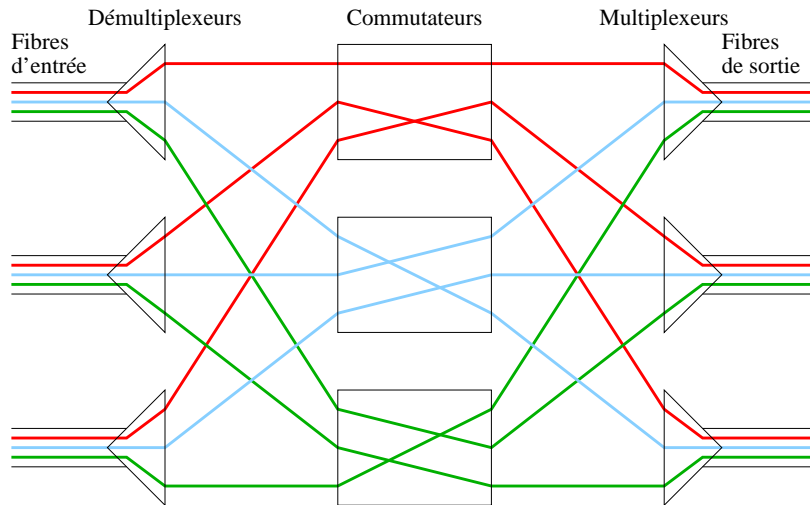


Figure 1.10: Brasseur reconfigurable (WR-OXC).

La composition d'un brasseur est représentée sur la figure 1.10. Une série de démultiplexeurs situés aux extrémités des fibres d'entrée permet dans un premier temps de démultiplexer les signaux entrants et de diriger spatialement chaque groupe de canaux WDM à la même longueur d'onde vers un commuta-

teur photonique particulier. Ces commutateurs sont contrôlés électroniquement et appliquent sur les groupes de canaux des fonctions de commutation indépendantes les unes des autres. Les canaux WDM sont enfin remultiplexés sur chaque fibre de sortie.

Associé à des multiplexeurs à insertion/extraction, un brasseur possède une capacité de routage plus étendue, puisqu'ainsi les canaux WDM peuvent être commutés, extraits ou insérés, comme il est représenté sur la figure 1.11.

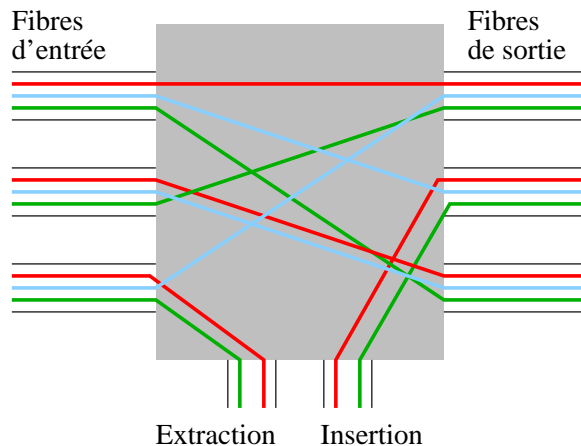


Figure 1.11: Brasseur avec MIE.

1.4.3 Convertisseurs optiques

Les différents composants optiques de commutation que nous venons de décrire imposent aux canaux WDM de respecter la **contrainte de continuité en longueur d'onde**. En d'autres termes, les canaux commutés conservent en sortie la longueur d'onde qu'ils possèdent en entrée. On peut s'affranchir de cette contrainte et augmenter ainsi les capacités de routage optique par l'utilisation de convertisseurs de longueur d'onde.

Les différentes techniques permettant la conversion de longueur d'onde ont été classifiées et comparées dans différentes études [DMJ⁺96, SI96, Wie96, Yoo96]. On peut distinguer deux grands types de technologies : la **conversion opto-électronique**, pour laquelle le signal optique doit être préalablement converti en un signal électrique, et la **conversion tout-optique**, pour laquelle le signal demeure dans le domaine optique. Les techniques de conversion tout-optique peuvent à leur tour être divisées en celles basées sur les *effets cohérents* et celles qui utilisent la *modulation croisée*.

Conversion opto-électronique

Dans la conversion de longueur d'onde opto-électronique, le signal optique à convertir doit d'abord être traduit dans le domaine électrique à l'aide d'une photodiode (voir figure 1.12). Le flot électronique résultant est ensuite réinjecté, après stockage éventuel dans une mémoire tampon, sur la commande de modulation d'un émetteur laser réglé sur la longueur d'onde désirée.

Cette méthode a été expérimentée pour des débits allant jusqu'à 10 Gbps [Yoo96]. Cependant, elle est plus complexe et consomme davantage de puissance que les autres méthodes décrites ci-dessous [DMJ⁺96]. De plus, le procédé de conversion opto-électronique affecte la transparence du signal, en lui imposant un format de modulation et un débit spécifiques. Toute information modulée en phase, en fréquence ou analogiquement, est perdue durant ce processus de conversion.

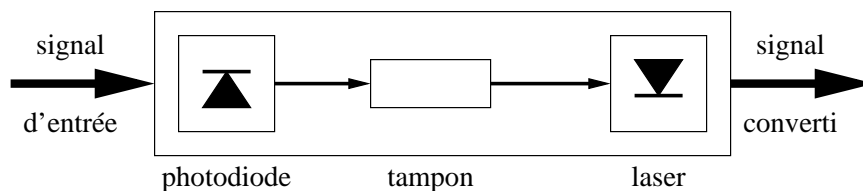


Figure 1.12: Convertisseur de longueur d'onde opto-électronique.

Conversion tout-optique

Sans entrer dans les détails technologiques, nous donnons ici brièvement les propriétés des deux principales techniques de conversion tout-optique.

Les méthodes de conversion basées sur les effets cohérents tolèrent tous les formats de modulation, offrant ainsi une transparence totale du signal. C'est de plus le seul type d'approche qui permet la conversion simultanée d'un ensemble de longueurs d'onde vers un autre et qui peut tolérer des débits dépassant les 100 Gbps [Yoo96].

Les techniques de conversion basées sur la modulation croisée utilisent des composants optiques actifs à semi-conducteurs, tels que des amplificateurs ou des émetteurs laser. Les débits autorisés sont de l'ordre de 10 Gbps. L'avantage de cette approche réside dans sa facilité d'utilisation.

Brasseur convertisseur

Un brasseur convertisseur (*Wavelength Translating Optical Cross-Connect*, *WT-OXC*) est un composant de commutation sélectif en longueur d'onde et qui permet de plus leur conversion. Le brassage des canaux WDM ne s'effectue donc pas nécessairement à longueur d'onde constante, comme dans un brasseur simple. Ces nœuds de commutation présentent ainsi davantage de flexibilité, mais une structure beaucoup plus complexe.

Les architectures des brasseurs convertisseurs étant aussi variées que sophistiquées, nous ne donnons pas ici de représentation détaillée. Il faut mentionner de plus que le domaine technologique de la conversion de longueur d'onde est en pleine évolution, et les techniques actuelles peuvent changer rapidement.

Cependant, nous représentons sur la figure 1.13 la capacité de routage optique d'un brasseur convertisseur total, c'est-à-dire capable d'effectuer n'importe quelle conversion d'un canal WDM, associé à des multiplexeurs à insertion/extraction. Cette fonctionnalité puissante sera considérée dans les chapitres suivants et nous étudierons son intérêt par rapport à celle des brasseurs simples.

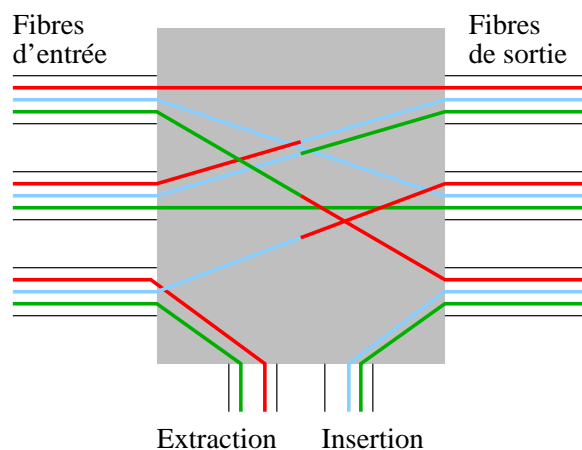


Figure 1.13: Brasseur convertisseur total avec MIE.

REMARQUE. Les systèmes de transmission WDM point-à-point aujourd'hui disponibles utilisent à l'entrée des liaisons des convertisseurs opto-électroniques appelés *transpondeurs*. Utilisés à l'origine comme interface avec les autres réseaux (SDH en l'occurrence), ils prennent de l'importance dans le brassage optique puisqu'ils peuvent être utilisés autour de brasseurs reconfigurables, parant ainsi à l'incompatibilité entre les longueurs d'onde propriétaires des différents constructeurs.

Bibliographie

- [Agr92] G. P. Agrawal. *Fiber-Optic Communication Systems*. Wiley, 1992.
- [Bra90] C. A. Brackett. Dense wavelength division multiplexing networks: principles and applications. *IEEE Journal on Selected Areas in Communications*, 8(6):948–964, 1990.
- [Des94] E. Desurvire. *Erbium-Doped Fiber Amplifiers: Principles and Applications*. Wiley, New-York, 1994.
- [DMJ+96] T. Durhuus, B. Mikkelsen, C. Joergensen, S. L. Danielsen, and K. E. Stubkjaer. All-optical wavelength conversion by semiconductor optical amplifiers. *IEEE/OSA Journal of Lightwave Technology*, 14:942–954, 1996.
- [Gla93] A. M. Glass. Fiber optics. *Physics Today*, January 1993.
- [Gre91] P. E. Green. The future of fiber-optic computer networks. *IEEE Computer*, 24(9):78–87, September 1991.
- [Gre93] P. E. Green. *Fiber Optic Networks*. Prentice-Hall, 1993.
- [Hen85] P. S. Henry. Lightwave primer. *IEEE Journal of Quantum Electronics*, 21:1862–1879, December 1985.
- [JJ96] I. Joindot and M. Joindot. *Les Télécommunications par Fibre Optique*. Collection Technique et Scientifique des Télécommunications. Dunod, 1996.
- [LG88] R. A. Linke and A. H. Gnauk. High-capacity coherent lightwave systems. *IEEE Journal of Lightwave Technology*, 6(11):1750–1769, November 1988.
- [LT76] A. Libchaber and G. Toulouse. Le retour des solitons. *La Recherche*, December 1976.
- [Muk92] B. Mukherjee. WDM-based local lightwave networks, Part 1: Single-hop systems. *IEEE Network Magazine*, 6(3):12–27, May 1992.

- [O'M93] M. J. O'Mahony. Optical amplifiers. In J.E. Midwinter, editor, *Photonics in Switching*, volume 1, pages 147–167. Academic Press, San Diego, CA, 1993.
- [Pet88] K. Petermann. *Laser Diode Modulation and Noise*. Kluwer Academic Publishers, 1988.
- [SI96] R. Sabella and E. Iannone. Wavelength conversion in optical transport networks. *Journal of Fiber and Integrated Optics*, 15(3):167–191, 1996.
- [vEvdP91] W. van Etten and J. van der Plaats. *Fundamentals of Optical Fiber Communications*. Prentice-Hall, 1991.
- [Whi95] T. J. Whitley. A review of recent system demonstrations incorporating 1.3- μm praseodymium-doped fluoride fiber amplifiers. *IEEE/OSA Journal of Lightwave Technology*, 13:744–760, May 1995.
- [Wie96] J. M. Wiesenfeld. Wavelength conversion techniques. In *Proc. of Optical Fiber Communication (OFC'96)*, pages 71–72, San Jose, CA, 1996.
- [Woo88] T. H. Wood. Multiple Quantum Well (MQW) waveguide modulators. *IEEE Journal of Lightwave Technology*, 6(6):143–157, June 1988.
- [Yoo96] S. J. B. Yoo. Wavelength conversion technologies for WDM network applications. *IEEE/OSA Journal of Lightwave Technology*, 14:955–966, June 1996.

CHAPITRE 2

Problématique et modélisation

2.1 Problématique

Un réseau de communication en fibre optique est déployé sur une région très étendue par un opérateur en télécommunications en tant qu'épine dorsale de son réseau global. Un point d'accès est rattaché au réseau fibré par l'intermédiaire d'un nœud de routage optique. Dans ce contexte, un point d'accès ne fournit pas nécessairement une porte d'entrée à un terminal particulier, mais plutôt à une collection de terminaux issus de sous-réseaux locaux, dont l'activité se trouve agrégée, de manière à ce que chaque émetteur ou récepteur du point d'accès fonctionne à un débit de transmission électronique.

Les points d'accès au réseau optique communiquent entre eux via des canaux WDM (cf. section 1.3.2), dits *tout-optiques* car ils ne subissent pas de conversion électronique intermédiaire, et qui sont également appelés *chemins optiques*. Une telle liaison peut recouvrir une série de liens en fibre optique et fournir ainsi une connexion de type *commutation de circuits* (*circuit-switching*, en anglais) entre deux sites distants dans la topologie physique. Chaque nœud de routage intermédiaire doit pour cela posséder une fonction de commutation optique adaptée. Les différents types de commutateurs optiques existants ont été présentés dans la section 1.4. Nous ne nous intéressons pas ici aux mécanismes de commutation permettant l'établissement des connexions, mais nous supposons que le contrôle du réseau est effectué de manière centralisée en vue d'assurer des liaisons durables.

Plusieurs architectures pour les réseaux tout-optiques ont été proposées dans la dernière décennie. Toutes présupposent cependant l'utilisation de liens entre les nœuds de routage qui sont constitués de paires de fibres optiques unidirectionnelles, puisque la pose de ces fibres en pratique s'effectue facilement de manière symétrique. La distinction entre les différentes architectures intervient au niveau

du type des nœuds de commutation utilisés. Pour notre part, nous considérons l'une des architectures les plus communes et les plus souvent étudiées, par exemple dans [CNW90, Pan92, VD93, RS95], qui est constituée de nœuds de commutation formés à partir de brasseurs avec multiplexeurs à insertion/extraction (nœuds WR-OXC avec MIE), tels qu'ils sont présentés dans la section 1.4.2 (voir figure 1.11, page 30). Cette hypothèse permet ainsi une grande liberté de commutation optique, sans toutefois autoriser la conversion des longueurs d'onde, qui est en revanche possible avec les brasseurs convertisseurs (nœuds WT-OXC, voir figure 1.13, page 32).

L'ensemble de nos travaux est basé sur ce type d'architecture. À l'époque de leur rédaction, les possibilités de conversion tout-optique étaient alors encore considérées comme relativement immatures et très coûteuses [BDO⁺98]. Il faut cependant noter que la recherche technologique actuelle se focalise sur ce nouvel enjeu afin de rendre les composants de conversion utilisables à grande échelle et dans un futur proche. Cette nouvelle perspective amène de nouveaux problèmes algorithmiques – notamment si les possibilités de conversion sont limitées – que nous n'avons pas abordés dans notre étude, si ce n'est le cas de la conversion totale. Nous invitons le lecteur intéressé par les problèmes résultant de la conversion limitée à se reporter sur les travaux [Gar98, ACKP98a, ACKP98b, RS98, KK99].

Dans ce qui suit, nous allons présenter la modélisation théorique qui a servi de support à notre recherche. Dans une certaine mesure, elle vise à simplifier la structure réelle des réseaux optiques qui peuvent se révéler hétérogènes, tout en conservant cependant les caractéristiques essentielles des problèmes divers rencontrés en pratique.

Ainsi, nous modéliserons un réseau tout-optique par un graphe ou multigraphe orienté, symétrique la plupart du temps, dont les sommets représenteront indifféremment les points d'accès et les nœuds de routage. Dans ce cadre, chaque arc correspond à une fibre optique unidirectionnelle qui supporte un nombre fixé de longueurs d'onde, que nous supposons uniforme dans tout le réseau.

Le type de commutateurs retenu dans nos études (brasseurs WR-OXC) possède la propriété d'être *sélectif en longueur d'onde*. Cela signifie que les différents signaux optiques arrivant sur une même fibre d'entrée peuvent être distingués selon leurs longueurs d'onde et commutés chacun sur une fibre de sortie particulière. Sans conversion optique ou opto-électronique, un signal doit conserver la même longueur d'onde tout le long d'un chemin reliant la source et la destination. De plus, à cause de l'interférence électro-magnétique, deux signaux ayant la même longueur d'onde ne peuvent pas être transmis sur la même fibre optique. Une longueur d'onde unique doit donc être réservée sur tous les liens d'un chemin

pour satisfaire une requête de connexion. Cette dernière contrainte disparaît en cas de conversion possible en des nœuds intermédiaires.

Dans les réseaux optiques WDM, le nombre de longueurs d'onde disponibles est fortement limité, généralement de l'ordre de la trentaine, et les perspectives futures ne laissent pas envisager de dépasser l'ordre de la centaine. De plus, le coût et la complexité des commutateurs et des multiplexeurs augmentent significativement avec le nombre de longueurs d'onde à traiter. C'est pourquoi il est crucial de chercher à minimiser l'utilisation de ces ressources optiques.

En général, le problème se posant dans un réseau WDM tout-optique consiste à satisfaire simultanément un ensemble de requêtes de connexion, appelé *instance de communication*, qui est formé de couples source-destination. Pour chacune des requêtes, il s'agit d'attribuer un chemin dans le réseau ainsi qu'une longueur d'onde, de manière à respecter les contraintes du routage optique formulées ci-dessus.

L'objectif de cette allocation de ressources est de minimiser le nombre de longueurs d'onde utilisées. Cette optimisation est importante de deux points de vue : d'une part, dans le cadre de la conception et du dimensionnement d'un réseau optique, en vue de satisfaire une certaine demande de trafic prévue à l'avance et afin de minimiser le coût global des équipements nécessaires (fibres, brasseurs, multiplexeurs) ; d'autre part, dans un réseau existant, pour gérer l'utilisation des ressources optiques en assurant la possibilité d'établir de nouvelles connexions ultérieurement.

2.2 Graphes et réseaux

Un réseau optique de télécommunications peut se modéliser à l'aide d'un graphe ou d'un multigraphe orienté, tels que nous allons les définir formellement. Les sommets représentent les nœuds du réseau, qui peuvent être des points d'accès ou des commutateurs. Les arcs représentent les liens physiques unidirectionnels en fibre optique. Un certain nombre de paramètres usuels en théorie des graphes sont associés à des propriétés physiques ou à des performances des réseaux modélisés.

En règle générale, le graphe physique du réseau optique ne comporte pas de structure particulière, hormis une certaine connexité nécessaire à l'acheminement du trafic et à la sécurisation du réseau. Notons cependant que l'étude de topologies régulières peut s'avérer intéressante à partir du moment où on les retrouve comme sous-structures du réseau global. De même, l'étude de schémas de communication réguliers, plus aisément formalisables et solubles en général, présente un intérêt dans la mesure où ils se rapprochent de situations réelles.

2.2.1 Notions élémentaires de théorie des graphes

Nous donnons seulement les définitions nécessaires par la suite. Le lecteur pourra trouver les notions non rappelées ici dans le livre de l'action RUMEUR [dR94], dans les livres de théorie des graphes de Berge [Ber83], de Bondy et Murty [BM76] ou dans l'ouvrage de Leighton consacré au parallélisme [Lei92].

Définitions

- Un **graphe orienté** (*digraph* en anglais) $G = (V(G), A(G))$ est constitué d'un ensemble fini $V(G) = \{x_1, x_2, \dots, x_N\}$ d'éléments, appelés **sommets**, et d'une famille finie $A(G) = \{a_1, a_2, \dots, a_m\}$ de représentants de couples de sommets, appelés **arcs**.
- Notons qu'un couple de sommets $(x, y) \in V(G) \times V(G)$ peut être représenté plusieurs fois dans la famille $A(G)$; on parle alors d'**arc multiple** (x, y) et de **multigraphe** G .
- Un **graphe non orienté** $G = (V(G), E(G))$ est constitué d'un ensemble fini $V(G)$ de sommets et d'une famille finie $E(G) = \{e_1, e_2, \dots, e_m\}$ de représentants de paires de sommets, appelées **arêtes**.
- Le nombre de sommets d'un (multi)graphe (orienté ou non) est appelé l'**ordre** du graphe, et est noté N .
- Si $a = (x, y)$ est un arc, alors x est son **extrémité initiale** et y son **extrémité finale**.
- Si $a = (x, y)$ est un arc, le sommet y est un **successeur** du sommet x et x est un **prédécesseur** de y . On s'autorise à dire également que y est **adjacent** à x .
- Si $e = \{x, y\}$ est une arête, les sommets x et y sont dits **adjacents** l'un à l'autre et l'arête e est dite **incidente** à x et y .
- Nous appellerons par la suite **liens** les arcs ou les arêtes d'un graphe, qu'il soit orienté ou non.
- Un graphe orienté est dit **symétrique** si l'existence d'un arc $a = (x, y)$ implique l'existence de l'arc $a' = (y, x)$.
- On note G^* le graphe orienté symétrique obtenu à partir d'un graphe non orienté G en prenant dans $A(G^*)$ les arcs $a = (x, y)$ et $a' = (y, x)$ pour chaque arête $e = \{x, y\}$ de $E(G)$.
- Un arc de la forme (x, x) ou une arête de la forme $\{x, x\}$ est appelé(e) **boucle**.

- Un (multi)graphe ne comportant ni boucle, ni arc ou arête multiple, est appelé **graphe simple**.
- Nous appellerons par la suite **graphes** (sans autre adjectif) les graphes simples non orientés.
- On appelle **degré sortant** (resp. **entrant**) d'un sommet x dans un graphe orienté, noté $d^+(x)$ (resp. $d^-(x)$) le nombre d'arcs d'extrémité initiale (resp. finale) x .
- On appelle **degré** d'un sommet x dans un graphe, noté $d(x)$, le nombre d'arêtes incidentes à x .
- On appelle **degré maximum** (resp. **minimum**) d'un graphe, noté Δ (resp. δ), le maximum (resp. minimum) des degrés des sommets.
- On appelle **chemin** (*dipath* en anglais) dans un graphe orienté une suite $P = (a_1, a_2, \dots, a_q)$ d'arcs, telle que l'extrémité finale de a_i est l'extrémité initiale de a_{i+1} , pour $1 \leq i < q$. La **longueur** du chemin P est alors le nombre d'arcs qui le composent.
- On peut également définir un chemin $P = (a_1, a_2, \dots, a_q)$ dans un graphe orienté simple par la suite des sommets (x_0, x_1, \dots, x_q) , telle que $a_i = (x_{i-1}, x_i)$. On dit alors que P est un chemin de x_0 vers x_q .
- On appelle **chaîne** (*path* en anglais) dans un graphe une suite $P = (e_1, e_2, \dots, e_q)$ d'arêtes, telle que deux arêtes consécutives sont incidentes à un même sommet. La **longueur** de la chaîne P est alors le nombre d'arêtes qui la composent.
- On s'autorise par la suite à appeler une chaîne un chemin, la distinction dépendant de la nature du graphe considéré.
- Un chemin qui ne comporte pas deux fois le même lien est dit **simple**. Un chemin qui ne comporte pas deux fois le même sommet est dit **élémentaire**.
- On appelle **distance** entre deux sommets x et y dans un graphe (orienté ou non), notée $d(x, y)$, la longueur minimale d'un chemin de x vers y .
- On appelle **diamètre** d'un graphe (orienté ou non), noté D , le maximum des distances entre les sommets.
- On appelle **circuit** dans un graphe orienté un chemin d'un sommet vers lui-même.
- On appelle **cycle** dans un graphe une chaîne d'un sommet vers lui-même.

- Un graphe orienté est dit **fortement connexe** s'il existe un chemin de tout sommet vers tout autre sommet.
- Un graphe est dit **connexe** s'il existe une chaîne entre toute paire de sommets.

Constructions classiques

- La **somme cartésienne** (*cartesian product* en anglais) de deux graphes simples orientés $G = (V, A)$ et $G' = (V', A')$, notée $G \square G'$, est le graphe ayant pour ensemble de sommets le produit cartésien $V \times V'$ et pour ensemble d'arcs les couples $((x, x'), (y, y'))$ tels que $x = y$ et $(x', y') \in A'$ ou tels que $x' = y'$ et $(x, y) \in A$. La somme cartésienne de deux graphes non-orientés se définit de manière analogue.
- Soient Γ un groupe dont la loi de composition est notée multiplicativement et S un ensemble (ou système) de générateurs de Γ ne contenant pas l'élément neutre e de Γ . Alors le **graphe de Cayley orienté** sur Γ muni du système de générateurs S , noté $G(\Gamma, S)$, est défini par :
 - l'ensemble des sommets est Γ ,
 - l'ensemble des arcs est constitué des couples (x, sx) , pour $x \in \Gamma$ et $s \in S$.

2.2.2 Réseaux usuels

Nous définissons ici certains graphes classiques qui sont le plus souvent étudiés. Dans le cadre des réseaux optiques, il s'agira de considérer les graphes orientés symétriques associés.

- On note C_N le **cycle** (ou **anneau**) (élémentaire) d'ordre N , de longueur N (voir figure 2.1).
- On note P_N la **chaîne** (élémentaire) d'ordre N , de longueur $N - 1$.
- On note K_N le **graphe complet** d'ordre N , ayant N sommets deux à deux adjacents (voir figure 2.1).
- On note H_n l'**hypercube** de dimension n , d'ordre $N = 2^n$, qui est la somme cartésienne de n copies du graphe K_2 . Il se définit récursivement à partir de K_2 :

$$H_n = K_2 \square H_{n-1} = \underbrace{K_2 \square K_2 \square \dots \square K_2}_{n \text{ fois}}$$

Il peut également se définir comme le graphe dont les sommets sont les mots de longueur n sur l'alphabet $\{0, 1\}$, tel que deux sommets sont adjacents si et seulement si leurs mots diffèrent en une seule lettre (voir figure 2.2).

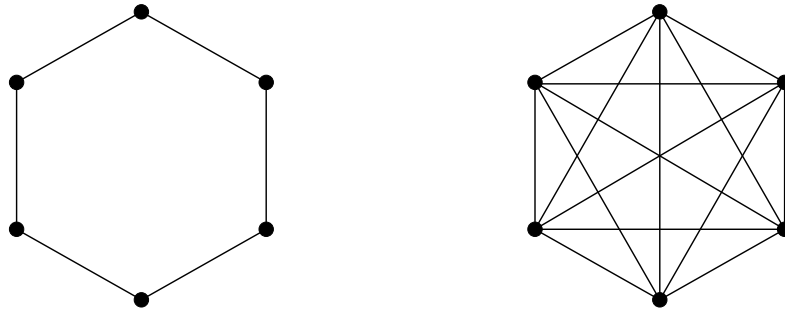


Figure 2.1: Cycle C_6 et graphe complet K_6 .

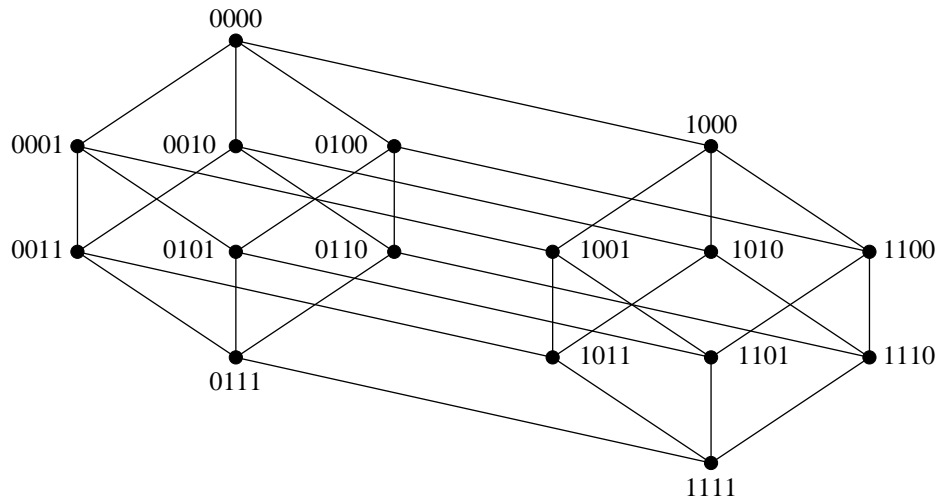


Figure 2.2: Hypercube H_4 .

- On note $M(p_1, p_2, \dots, p_n)$ la grille (*mesh* en anglais) de dimension n , d'ordre $N = \prod_{i=1}^n p_i$, qui est la somme cartésienne des n chaînes P_{p_i} ($i = 1, 2, \dots, n$), soit $P_{p_1} \square P_{p_2} \square \dots \square P_{p_n}$ (voir figure 2.3).
- On note $TM(l_1, l_2, \dots, l_n)$ la grille torique (*toroidal mesh* en anglais) ou tore de dimension n , d'ordre $N = \prod_{i=1}^n l_i$, qui est la somme cartésienne des n cycles C_{l_i} ($i = 1, 2, \dots, n$), soit $C_{l_1} \square C_{l_2} \square \dots \square C_{l_n}$ (voir figure 2.3).
- Un arbre de cycles est un graphe formé par une union de cycles qui s'intersectent deux-à-deux en au plus un sommet, et tel que deux sommets quelconques peuvent être reliés par exactement deux chemins arête-disjoints (voir figure 2.4).

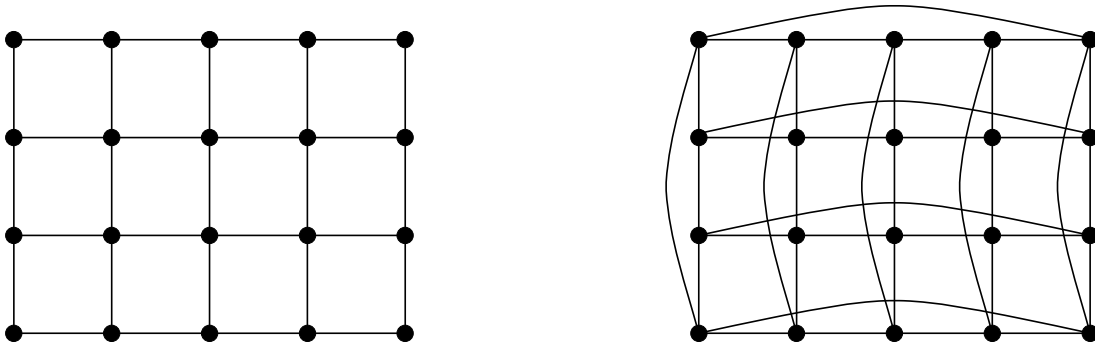


Figure 2.3: Grille $M(5,4)$ et tore $TM(5,4)$.

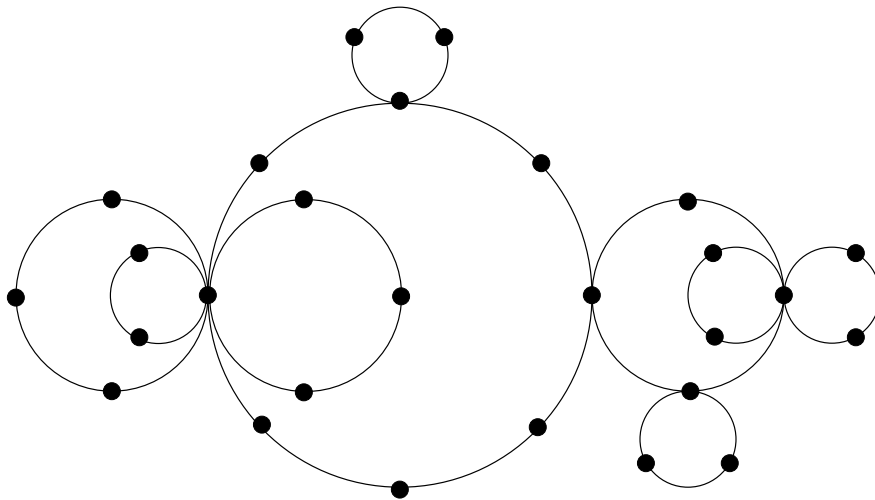


Figure 2.4: Arbre de cycles.

2.3 Optimisation et approximation

Un *problème d'optimisation* P est donné par un ensemble \mathcal{I} d'instances (entrées), un ensemble \mathcal{S} de solutions (sorties), une fonction $s : \mathcal{I} \rightarrow \mathcal{P}(\mathcal{S})$ des instances vers les parties des solutions acceptables, une fonction $val : \mathcal{I} \times \mathcal{S} \rightarrow \mathbb{N}$ mesurant la qualité $val(I, S)$ de la solution $S \in s(I)$ pour l'instance I , et un objectif *min* ou *max*. Si l'objectif est *min* (resp. *max*), on souhaite trouver une solution $S \in s(I)$ qui minimise (resp. maximise) $val(I, S)$. Nous ne considérons que des problèmes d'optimisation pour lesquels il est facile de vérifier l'appartenance à \mathcal{I} , l'appartenance à une partie de solutions acceptables, et de calculer la fonction val . Davantage de détails sur les définitions formelles de ces problèmes peuvent être trouvés dans [MPS98]. Lorsque nous aborderons par la suite des problèmes particuliers d'optimisation, nous les présenterons parfois

d'une manière plus informelle, en gardant toutefois à l'esprit ces définitions.

Nous rappelons qu'un *problème de décision* P est \mathcal{NP} -complet s'il est dans la classe de complexité \mathcal{NP} et s'il existe une réduction polynomiale de tout autre problème de \mathcal{NP} vers P . La version décisionnelle d'un problème d'optimisation P est le problème de décision obtenu en ajoutant un entier K à l'instance I et en se demandant "Existe-t-il une solution $S \in s(I)$ telle que $val(I, S) \geq K$?" dans le cas d'un problème de maximisation, ou bien "Existe-t-il une solution $S \in s(I)$ telle que $val(I, S) \leq K$?" dans le cas d'un problème de minimisation. Un problème d'optimisation est dit \mathcal{NP} -dur ou \mathcal{NP} -difficile si sa version décisionnelle est \mathcal{NP} -complète. Nous invitons le lecteur intéressé à se référer au livre de Garey et Johnson [GJ79] pour une introduction à la théorie de la \mathcal{NP} -complétude.

Pour un problème d'optimisation donné, un algorithme qui s'exécute en temps polynômial et qui produit toujours une solution optimale est dit *exact*, ou simplement *optimal*. L'existence d'un algorithme exact pour un problème \mathcal{NP} -dur impliquerait l'égalité $\mathcal{P} = \mathcal{NP}$. Par conséquent, les algorithmes polynômiaux d'approximation sont intéressants pour les problèmes d'optimisation \mathcal{NP} -durs.

Un *algorithme d'approximation* A pour un problème d'optimisation P est un algorithme déterministe dont le temps d'exécution est polynômial en la taille de l'entrée et qui rend toujours une solution acceptable. Notons $OPT(I)$ la valeur d'une solution optimale, i.e., $OPT(I) = \max_{S \in s(I)} val(I, S)$ pour les problèmes de maximisation et $OPT(I) = \min_{S \in s(I)} val(I, S)$ pour les problèmes de minimisation. Dans le premier cas, on dit que l'algorithme A possède un *facteur d'approximation (absolu)* ρ si $OPT(I)/val(I, A(I)) \leq \rho$ pour toute instance I , $A(I)$ représentant la sortie de A sur l'entrée I . Dans le second cas, la condition est $val(I, A(I))/OPT(I) \geq \rho$.

Parfois il peut être intéressant de considérer le comportement asymptotique d'un algorithme d'approximation. Pour un problème de maximisation, un algorithme A possède un *facteur d'approximation asymptotique* ρ si $\limsup_{OPT(I) \rightarrow \infty} A(I)/OPT(I) \leq \rho$. Pour un problème de minimisation, la condition requise est $\limsup_{OPT(I) \rightarrow \infty} OPT(I)/A(I) \leq \rho$. Une présentation générale des algorithmes d'approximation pour les problèmes \mathcal{NP} -durs peut être trouvée dans le livre [Hoc97] publié par Hochbaum.

Les algorithmes d'approximation requièrent habituellement que l'entrée soit donnée entièrement à l'avance, et dans ce cas ils sont appelés *statiques (off-line)*, en anglais). Pour les problèmes d'optimisation considérés dans cette thèse, l'entrée consiste généralement en un ensemble ou une collection de requêtes de communication. En revanche, pour certaines applications il est réaliste de supposer que les requêtes se présentent au fur et à mesure et l'algorithme doit les traiter sans

connaître les requêtes à venir. De tels algorithmes sont appelés *dynamiques* (*on-line*, en anglais). Pour d'autres applications encore, l'ensemble des requêtes est connu à l'avance, mais la durée d'une communication est inconnue jusqu'à sa fin. Une introduction aux algorithmes dynamiques et un état-de-l'art peuvent être trouvés dans le livre de Borodin et El-Yaniv [BEY98] et dans celui publié par Fiat et Woeginger [FW98].

2.4 Formulation des problèmes étudiés

Nous sommes à présent en mesure de définir les problèmes d'optimisation étudiés dans le chapitre suivant et les annexes. Ces problèmes sont reliés à l'allocation des ressources pour des requêtes de connexion dans un réseau de communication optique. Un réseau donné est modélisé par un graphe ou multigraphe orienté $G = (V, A)$. Les sommets correspondent aux nœuds du réseau et les arcs aux liens physiques en fibre optique. La plupart du temps, les graphes étudiés seront orientés symétriques. Leur graphe non-orienté associé pourra alors être considéré également, par souci de simplicité.

Dans leur généralité, les problèmes se posant dans les réseaux WDM tout-optiques consistent à satisfaire simultanément un ensemble de requêtes de connexion, appelé *instance*, formé de couples de nœuds. Pour chaque requête, il s'agit d'attribuer un chemin dans le réseau et d'allouer une longueur d'onde, soit tout le long du chemin, soit sur chaque lien formant le chemin en cas de conversion possible, de manière à ce que deux requêtes distinctes n'utilisent pas la même longueur d'onde sur le même lien. L'objectif dans ce cadre est de minimiser l'utilisation des ressources optiques, c'est-à-dire le nombre total de longueurs d'onde permettant de satisfaire l'instance donnée.

Les premières études dans la littérature ont représenté les réseaux optiques par des graphes non-orientés. Cette modélisation ne correspond pas précisément à la réalité physique, puisque les communications optiques sont unidirectionnelles, mais elle peut être prise en considération sous les hypothèses restrictives suivantes :

REMARQUE. Le modèle non-orienté des réseaux de communication tout-optiques correspond au cas où les requêtes de connexion sont symétriques et sous la contrainte d'être routées deux-à-deux par des chemins symétriques et sur la même longueur d'onde.

Les problèmes d'optimisation définis dans toute cette section peuvent être étudiés pour les différentes variantes de graphes (orientés, orientés symétriques,

orientés non-symétriques, non-orientés).

2.4.1 Problème du routage tout-optique

Dans un réseau tout-optique, une requête de connexion entre un nœud u et un nœud v est représentée par le couple (u, v) . Elle est établie en réservant une longueur d'onde sur tous les liens d'un chemin de u vers v . Les chemins correspondant aux différentes connexions ne doivent pas utiliser la même longueur d'onde sur le même lien. Si les commutateurs ne permettent pas de conversion (donc de type WR-OXC), une connexion doit utiliser la même longueur d'onde tout le long du chemin de la source vers la destination.

Nous appelons *problème du routage (tout-)optique* le problème de l'allocation des chemins et des longueurs d'onde (ou couleurs) pour une collection de requêtes (instance) dans les réseaux tout-optiques sans conversion.

PROBLÈME du Routage Optique

- Entrée :** un (multi)graphe G et une instance I de requêtes dans G
Sortie : une allocation de chemins et de couleurs aux requêtes, telle que deux chemins utilisant le même lien ont deux couleurs différentes
Objectif : minimiser le nombre de couleurs utilisées

On note $\tilde{w}(\mathbf{G}, \mathbf{I})$ le nombre de couleurs d'une solution optimale si le graphe G est orienté, et $w(\mathbf{G}, \mathbf{I})$ si le graphe G est non-orienté.

L'allocation de chemins aux requêtes de l'instance donnée est appelée *routage*. Par abus de langage, le terme routage désigne également l'ensemble des chemins ainsi formés. Notons que le routage doit être calculé et fourni en sortie. Si le routage est donné comme une partie de l'entrée (au lieu de permettre à l'algorithme de le choisir), on obtient un problème de *coloration de chemins*. Ceci revient à résoudre le problème de la coloration des sommets du *graphe de conflit* associé, de telle manière que deux sommets adjacents sont colorés différemment. Le nombre minimum de couleurs nécessaires pour colorer les sommets d'un graphe est appelé **nombre chromatique**.

DÉFINITION. Le **graphe de conflit** associé à un routage R dans un graphe G (orienté ou non) est le graphe non-orienté dont les sommets représentent les chemins de R et tel que deux sommets sont adjacents si et seulement si les chemins correspondants partagent un lien de G .

On note $\tilde{w}(\mathbf{G}, \mathbf{I}, \mathbf{R})$ (resp. $w(\mathbf{G}, \mathbf{I}, \mathbf{R})$) le nombre chromatique du graphe de conflit d'un routage R réalisant une instance I dans un graphe orienté G (resp.

non-orienté). Ainsi $\vec{w}(G, I) = \min_R \vec{w}(G, I, R)$, et de même pour $w(G, I)$.

Pour les graphes qui sont des arbres, le routage optique et la coloration de chemins sont deux problèmes équivalents, puisque un chemin élémentaire associé à une requête quelconque est unique. La figure 2.5 montre un exemple simple d'instance de requêtes dans un arbre orienté symétrique à six sommets, chaque lien représentant une paire d'arcs symétriques. Cinq chemins sont dessinés correspondant aux requêtes $(0, 2)$, $(1, 3)$, $(4, 3)$, $(4, 5)$ et $(0, 5)$. Une coloration correcte possible attribuée à ces chemins les couleurs vert, rouge, vert, rouge et bleu, respectivement. En outre, trois couleurs sont effectivement nécessaires puisque le graphe de conflit est un cycle de longueur impaire.

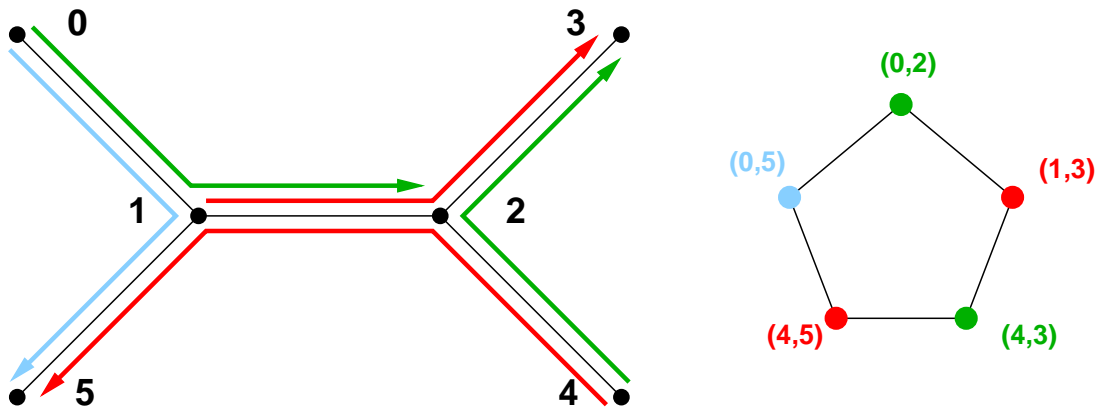


Figure 2.5: Un routage dans un arbre et son graphe de conflit associé.

Le problème du routage optique vaut la peine d'être étudié dans le cadre de la conception et du dimensionnement d'un réseau optique, en vue de satisfaire une certaine demande de trafic prévue à l'avance. De même, il est digne d'intérêt dans un réseau possédant une bande passante optique suffisante pour satisfaire un ensemble de requêtes de connexion, en vue de minimiser l'utilisation des ressources optiques, qui se révèlent limitées par la technologie, et d'assurer ainsi la possibilité d'établir de nouvelles connexions ultérieurement.

Cependant, la situation est différente dans un réseau optique qui n'a pas la capacité suffisante pour satisfaire tout un ensemble de requêtes simultanément. Plus exactement, ceci se produit lorsque le nombre de longueurs d'onde disponibles est plus petit que le nombre nécessaire. Il peut alors être souhaitable de satisfaire un sous-ensemble de l'instance initiale, le but étant de maximiser son cardinal.

PROBLÈME du Routage Optique Maximum

- Entrée :** un (multi)graphe G , une instance I de requêtes dans G et un entier $W > 0$
- Sortie :** une partie $I' \subseteq I$ et une allocation aux requêtes de I' de chemins et de couleurs parmi W possibles, telle que deux chemins utilisant le même lien ont deux couleurs différentes
- Objectif :** maximiser le cardinal de I'

REMARQUE. Une autre possibilité de routage, si le nombre de longueurs d'onde est insuffisant, consiste à réaliser l'instance de communication en plusieurs étapes successives. Si l'on conserve un routage tout-optique entre les sources et les destinations, alors le nombre d'étapes nécessaires devient $\lceil \bar{w}/W \rceil$ si W longueurs d'onde sont disponibles. Si l'on s'autorise des reconversions opto-électroniques en des nœuds intermédiaires, on se place alors dans le cadre des réseaux dits *multi-hop* qui est abordé dans les annexes D et E.

2.4.2 Problème de la charge

Dans les réseaux optiques permettant toute conversion d'une longueur d'onde en une autre, dans un nœud intermédiaire d'un chemin optique, la contrainte de la coloration des chemins alloués aux requêtes disparaît. Il en résulte en fait une multicoloration de chaque chemin, qui consiste à attribuer une couleur spécifique sur chaque lien utilisé, de telle sorte que deux chemins qui partagent un lien n'ont pas la même couleur sur ce lien. Ceci ne pose aucun problème algorithmique, puisque toute stratégie d'allocation gloutonne fonctionne. Le nombre total de longueurs d'onde nécessaires pour établir les requêtes de connexion est alors égal au nombre maximum de chemins utilisant le même lien.

DÉFINITIONS. La **charge d'un lien** α pour un routage R réalisant une instance I dans un graphe G (orienté ou non) désigne le nombre de chemins de R qui utilisent le lien α . Ce nombre est noté $\bar{\pi}(G, I, R, \alpha)$ si G est orienté et $\pi(G, I, R, \alpha)$ si G est non-orienté.

La **charge du routage** R est alors définie comme la charge maximale des liens pour R . C'est le nombre maximum de chemins de R qui utilisent le même lien de G . Il est noté $\bar{\pi}(G, I, R)$ en orienté et $\pi(G, I, R)$ en non-orienté.

Nous appelons *problème de la charge* le problème d'optimisation résultant du problème du routage optique avec l'hypothèse de conversion totale des longueurs d'onde. Il s'agit de trouver un routage R réalisant l'instance I donnée et minimisant la charge $\bar{\pi}(G, I, R)$ (ou $\pi(G, I, R)$).

PROBLÈME de la Charge

- Entrée :** un (multi)graphe G et une instance I de requêtes dans G
Sortie : une allocation de chemins aux requêtes (un routage)
Objectif : minimiser la charge du routage

On note $\tilde{\pi}(\mathbf{G}, \mathbf{I})$ la charge d'une solution optimale si le graphe G est orienté, et $\pi(\mathbf{G}, \mathbf{I})$ si le graphe G est non-orienté. Ainsi $\tilde{\pi}(G, I) = \min_R \max_\alpha \tilde{\pi}(G, I, R, \alpha)$, et de même pour $\pi(G, I)$.

Pour l'exemple de la figure 2.5, nous avons $\tilde{\pi}(G, I) = 2$.

De manière évidente, la charge optimale d'un routage réalisant une instance donnée minore le nombre optimal de longueurs d'onde pour le problème du routage optique appliqué à la même instance. Ceci résulte immédiatement du fait que le problème de la charge est un problème d'allocation de chemins et de longueurs d'onde, avec la liberté supplémentaire de la conversion.

PROPRIÉTÉ. Pour tout problème (G, I) , $\vec{w}(G, I) \geq \tilde{\pi}(G, I)$.

La propriété analogue est aussi valable en non-orienté. Une autre démonstration consiste à dire que le nombre de longueurs d'onde nécessaires pour le problème du routage optique est au moins égal au nombre maximum de chemins devant nécessairement utiliser le même lien, ce qui correspond à la définition de la charge minimale d'un routage.

REMARQUE. Notons que répondre à la question " $\tilde{\pi}(G, I) \leq p$?", étant donné un entier p , revient à déterminer si le multiflot entier correspondant aux requêtes de I est réalisable dans le réseau de flot F_p obtenu en affectant à chaque arc de G une capacité égale à p . Il s'agit donc d'un problème très classique en optimisation (cf. le livre [AMO93] ou le chapitre [GTT90]).

Comme le problème du routage optique, le problème de la charge possède son problème de maximisation associé :

PROBLÈME du Routage Maximum

- Entrée :** un (multi)graphe G et une instance I de requêtes dans G et un entier $W > 0$
Sortie : une partie $I' \subseteq I$ et une allocation de chemins aux requêtes de I' , de charge au plus W
Objectif : maximiser le cardinal de I'

Dans le cas extrême où le nombre W de longueurs d'onde disponibles est réduit à 1, les deux problèmes de maximisation précédents deviennent équivalents et peuvent être énoncés par une formulation qui a été très étudiée dans la littérature (cf. par exemple le chapitre de Frank [Fra95b] dans [GGL95]) :

PROBLÈME du Routage Disjoint Maximum

- Entrée :** un (multi)graphe G et une instance I de requêtes dans G
- Sortie :** une partie $I' \subseteq I$ et une allocation de chemins aux requêtes de I' , telle que deux chemins n'utilisent pas le même lien
- Objectif :** maximiser le cardinal de I'

La version décisionnelle du problème devient :

PROBLÈME du Routage Disjoint

- Données :** un (multi)graphe G et une instance I de requêtes dans G
- Question :** existe-t-il une allocation de chemins aux requêtes de I' , telle que deux chemins n'utilisent pas le même lien ?

2.4.3 Relations entre les problèmes

Nous commençons par montrer brièvement en quoi le problème du routage optique peut être vu comme un problème de multiflot entier dans un graphe associé, et donc également comme un problème de routage disjoint.

Par souci de simplicité, nous avons représenté sur la figure 2.6 la réduction dans le cas d'une instance de multicast, pour laquelle toutes les requêtes possèdent le même nœud source. Une construction semblable est utilisée en annexe A. Il s'agit en l'occurrence de résoudre dans un graphe orienté symétrique G le problème du routage optique pour l'instance $I = \{x\} \times \{v, w, y, z\}$.

Nous avons déjà remarqué que le problème de la charge pour (G, I) revient à déterminer le plus petit entier p rendant réalisable un certain problème associé F_p de multiflot entier. Nous dirons qu'un sommet $x \in V(G)$ est une source (resp. une destination) s'il existe $y \in V(G)$ tel que $(x, y) \in I$ (resp. $(y, x) \in I$). Le réseau de multiflot F'_p associé au problème du routage optique (G, I) est construit à partir de G de la manière suivante.

- on commence par prendre p copies indépendantes du graphe G dont les arcs sont munis d'une capacité unitaire. Les p copies d'un sommet $u \in V(G)$ sont notées u_1, u_2, \dots, u_p .
- pour chaque source $x \in V(G)$, on ajoute un sommet x_{out} qui est relié à x_i , pour $1 \leq i \leq p$, par un arc (x_{out}, x_i) de capacité infinie.

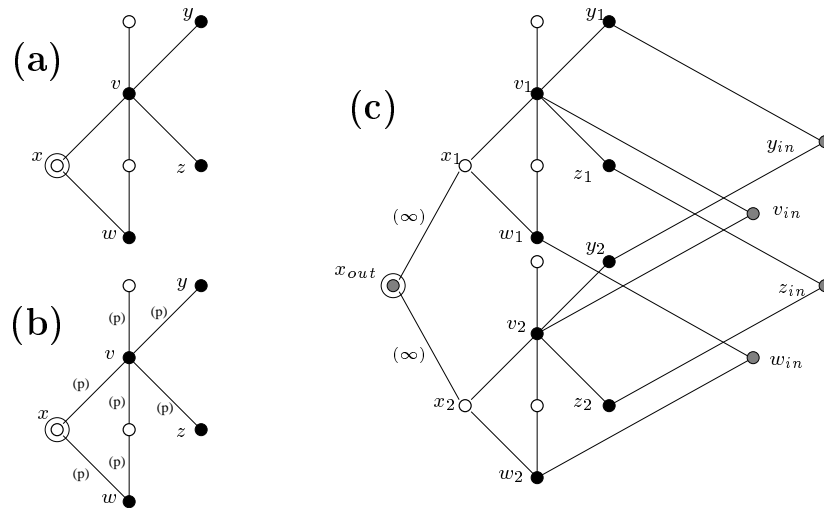


Figure 2.6: (a) Un graphe orienté G avec une source x et des destinations $\{v, w, y, z\}$. (b) Le réseau de flot F_p . (c) Le réseau de flot F'_2 . Toutes les capacités non marquées valent 1 et toutes les orientations des arcs sont omises.

- pour chaque destination $y \in V(G)$, on ajoute un sommet y_{in} , et le sommet y_i , pour $1 \leq i \leq p$, est relié à y_{in} par un arc (y_i, y_{in}) de capacité unitaire.
- à l'instance de requêtes I dans G , nous associons l'instance $I' = \{(x_{out}, y_{in}) \mid (x, y) \in I\}$ dans F'_p .

Il est alors facile de voir que $\vec{w}(G, I) = \min \{p \mid \text{le multiflot } F'_p \text{ est réalisable}\}$.

Nous avons donc établi que problème du routage optique (resp. maximum) se réduit au problème du routage disjoint (resp. maximum).

REMARQUE. Dans le cas particulier d'une instance multicast, le problème de routage optique se réduit à un problème de simple flot. Nous renvoyons à l'annexe A pour une présentation plus détaillée.

PROPRIÉTÉ. S'il existe, pour une classe de graphes, un algorithme pour le problème du routage disjoint maximum ayant un facteur d'approximation ρ (non nécessairement constant), alors il existe, pour cette même classe de graphes, un algorithme pour le problème du routage optique ayant un facteur d'approximation en $O(\log |I| \cdot \rho)$.

REMARQUE. La technique utilisée est semblable à celle colorant les sommets d'un graphe à l'aide d'un algorithme sachant approximer le problème de trouver

un ensemble de sommets indépendants de cardinalité maximale (*maximum independent set*, MIS). Elle a été utilisée par Aumann et Rabani dans [AR95] afin de traiter le problème du routage optique dans les grilles bidimensionnelles.

IDÉE DE PREUVE. Soit I une instance dans un graphe G et posons $\vec{w} = \vec{w}(G, I)$. Pour toute partie $I' \subseteq I$, $\vec{w}(G, I') \leq \vec{w}$. Donc dans toute solution optimale du problème du routage optique (G, I) , il existe au moins $|I'|/\vec{w}$ requêtes utilisant la même couleur, et donc routées par des chemins arc-disjoints. Par conséquent, une solution optimale du problème du routage disjoint maximum (G, I') est de taille au moins $|I'|/\vec{w}$. Comme nous supposons l'existence d'un algorithme polynômial d'approximation pour ce problème, nous en déduisons que pour toute partie I' , on peut trouver en temps polynômial une partie $I'' \subseteq I'$ de taille $\Omega(|I'|/(\rho\vec{w}))$ qui peut être routée par des chemins arc-disjoints.

On peut alors définir un algorithme d'approximation glouton pour le routage optique de la manière suivante. Initialement, $I' = I$ et on sait router avec la même couleur une partie I_0 de taille $\Omega(|I|/\rho\vec{w})$. En posant ensuite $I' = I \setminus I_0$, on sait router avec une autre couleur une partie $I_1 \subseteq I'$ de taille $\Omega(|I'|/(\rho\vec{w}))$, en sachant que $|I'| = O(|I|(1 - \frac{1}{\rho\vec{w}}))$. L'ensemble courant I' décroît donc d'un facteur $\Omega(\rho\vec{w}/(\rho\vec{w} - 1))$ à chaque étape. Le calcul montre que le nombre d'étapes de l'algorithme, jusqu'à épuisement de toutes les requêtes de I , est $O(\rho\vec{w} \log |I|)$, ce qui correspond au nombre de couleurs utilisées. \square

2.5 Communications structurées

Lorsqu'un ensemble de connexions doit s'établir dans un réseau de télécommunications, il peut être plus ou moins aléatoire ou posséder une structure prédéterminée. Dans le cas de communications prévisibles à l'avance, il est souhaitable de connaître quel routage utiliser avant même que le problème ne se pose, et d'optimiser la solution fournie afin d'exploiter au mieux les ressources de transmission du réseau.

C'est pourquoi les problèmes de routage dans les réseaux, très étudiés dans le cadre des réseaux d'interconnexion entre processeurs [FL94, HHL86, HKMP95], peuvent être regroupés en deux grandes classes [Fra95a, Lei92] : les routages *pré-calculés* ou *statiques* (en anglais, *off-line*) et les routages *temps-réel* ou *dynamiques* (en anglais, *on-line*).

Le terme temps-réel est utilisé lorsque les requêtes de communication ne sont pas connues à l'avance et qu'il faut résoudre le problème de routage au moment où elles se présentent. Dans ce cas de figure, les communications peuvent intervenir à tout moment, a priori sans aucune synchronisation ni cohérence entre elles,

et on pourra parler de communications anarchiques. L'étude de cette classe de problèmes, qui n'est pas notre objet, passe par l'étude (difficile) de fonctions de routage qui essaient de garantir au mieux l'établissement des communications en évitant les blocages.

À l'opposé, la classe des routages précalculés concerne des problèmes réguliers qui apparaissent souvent, notamment dans les réseaux de transport, et dont il est bénéfique de déterminer et d'optimiser la solution au préalable. Ces instances de communication particulières sont appelées *communications structurées*. On parle également de communications *globales* ou *collectives*, lorsque celles-ci font intervenir la totalité des nœuds du réseau.

Nous définissons à présent les communications structurées les plus fréquemment mises en œuvre :

- **Diffusion** (*One-to-All* ou *broadcasting*) : opération qui consiste à envoyer d'un nœud du réseau une même information vers tous les autres nœuds.
- **Distribution** (diffusion personnalisée, *personalized One-to-All*, *distributing* ou *scattering*) : opération qui consiste à envoyer d'un nœud du réseau une information différente vers chacun des autres nœuds.
- **Rassemblement** (*gathering*) : opération qui consiste à envoyer de chaque nœud du réseau une information vers un nœud particulier. C'est l'opération inverse de la distribution.
- **Échange total** (*All-to-All*, *total exchange* ou *gossiping*) : opération qui consiste à envoyer de chaque nœud du réseau une même information vers tous les autres nœuds. Cela revient à effectuer une diffusion à partir de tous les nœuds du réseau simultanément.
- **Multidistribution** (échange total personnalisé, *personalized All-to-All*, *complete exchange* ou *multiscattering*) : opération qui consiste à envoyer de chaque nœud du réseau une information différente vers chacun des autres nœuds. Cela revient à effectuer une distribution à partir de tous les nœuds du réseau simultanément.
- **Multicast** (diffusion partielle, *One-to-Many* ou *multicasting*) : opération qui consiste à envoyer d'un nœud du réseau une même information vers d'autres nœuds.
- **Permutation** : opération pour laquelle chaque nœud du réseau envoie et reçoit une information exactement.
- **k -Relation** : opération pour laquelle chaque nœud du réseau envoie et reçoit au plus k informations.

REMARQUE. Dans le cadre de notre étude sur le routage tout-optique, nous attirons l'attention sur le fait que les schémas de communication structurée que sont, d'une part la diffusion et l'échange-total, d'autre part la distribution et la multidistribution, correspondent respectivement aux mêmes instances de communication, puisque la donnée des requêtes ne tient pas compte de la nature de l'information véhiculée. Ainsi, un ensemble de chemins optiques réalisant une diffusion (un échange total) peut tout aussi bien réaliser une distribution (une multidistribution). Cependant, nous continuerons d'utiliser les termes les plus restrictifs par (mauvaise) habitude. Notons toutefois que la distinction devient importante dans le cadre des réseaux optiques multi-hop, qui sont étudiés pour la diffusion en annexe E.

DÉFINITIONS. Étant donné l'ensemble V des sommets d'un graphe G ,

- Une instance de diffusion ou de distribution est notée I_0 et correspond à un ensemble $\{(x_0, y) \mid y \in V\}$ pour un sommet x_0 fixé.
- L'instance de l'échange total ou de la multidistribution est notée I_A et correspond à l'ensemble $\{(x, y) \mid x \in V, y \in V\}$.
- Une instance de multicast ou de multicast personnalisé est notée I_M et correspond à un ensemble $\{(x_0, y) \mid y \in V'\}$ pour un sommet x_0 fixé et une partie $V' \subseteq V$.
- Une instance de permutation est notée I_1 et correspond à un ensemble $\{(x, \phi(x)) \mid x \in V\}$ pour une bijection ϕ de V .
- Une 1-relation est une instance notée également I_1 et qui correspond à une partie d'une instance de permutation.
- Une k -relation est une instance notée I_k et qui correspond à une union de k instances qui sont des 1-relations.

Bibliographie

- [ACKP98a] V. Auletta, I. Caragiannis, C. Kaklamanis, and G. Persiano. Efficient wavelength routing with low-degree converters. In *Proc. the DIMACS Workshop : "Multichannel Optical Networks : Theory and Practice"*, volume 46 of *DIMACS Series on Discrete Mathematics and Theoretical Computer Science*, pages 1–14, Rutgers University, NJ, March 1998. American Mathematical Society.
- [ACKP98b] V. Auletta, I. Caragiannis, C. Kaklamanis, and P. Persiano. On the complexity of wavelength converters. In *Proc. of the 23rd Symposium on Mathematical Foundations of Computer Science (MFCS'98)*, volume 1450 of *Lecture Notes in Computer Science*, pages 771–779. Springer-Verlag, 1998.
- [AMO93] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, 1993.
- [AR95] Y. Aumann and Y. Rabani. Improved bounds for all optical routing. In *Proc. of the 6th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '95)*, pages 567–576. ACM Press, January 1995.
- [BDO⁺98] G-L. Bona, W. E. Denzel, B. J. Offrein, R. Germann, H. W. M. Salemink, and F. Horst. Wavelength division multiplexed add/drop ring technology in corporate backbone networks. Technical Report RZ3046, IBM Research Division, August 1998.
- [Ber83] C. Berge. *Graphes*. Gauthiers-Villars, 1983.
- [BEY98] A. Borodin and R. El-Yaniv. *Online Computation and Competitive Analysis*. Cambridge University Press, 1998.
- [BM76] J. A. Bondy and U. S. R. Murty. *Graph Theory with Applications*. MacMillan Press, 1976.
- [CNW90] N. K. Cheung, K. Nosu, and G. Winzer. Special issue on dense WDM networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, 8, 1990.

- [dR94] Jean de Rumeur. *Communications dans les Réseaux de Processeurs*. Masson, Paris, Collection Etudes et Recherche en Informatique edition, 1994.
- [FL94] P. Fraigniaud and E. Lazard. Methods and problems of communication in usual networks. *Discrete Applied Mathematics*, 53:79–133, 1994.
- [Fra95a] P. Fraigniaud. *Vers un principe de localité pour les communications dans les réseaux d'interconnexion*. Thèse d'habilitation, Université de Lyon, École Normale Supérieure de Lyon, 1995.
- [Fra95b] A. Frank. *Handbook of Combinatorics*, volume 1, chapter Connectivity and Network Flows, pages 111–177. North-Holland, 1995.
- [FW98] A. Fiat and G. J. Woeginger, editors. *Online Algorithms: The State of the Art*, volume 1442 of *Lecture Notes in Computer Science*. Springer-Verlag, 1998.
- [Gar98] L. Gargano. Limited wavelength conversion in all-optical tree networks. In *Proc. of 25th International Colloquium on Automata, Languages and Programming (ICALP'98)*, volume 1443 of *Lecture Notes in Computer Science*, pages 544–555. Springer-Verlag, 1998.
- [GGL95] R. L. Graham, M Grötschel, and L Lovász, editors. *Handbook of Combinatorics*, volume 1. North-Holland, 1995.
- [GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability - A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.
- [GTT90] A. V. Goldberg, R. E. Tarjan, and É. Tardos. Network flow algorithms. In B. Korte, L. Lovász, H. J. Prömel, and A. Schrijver, editors, *Paths, Flows and VLSI-Layout*, pages 101–164. Springer-Verlag, 1990.
- [HHL86] S. M. Hedetniemi, S. T. Hedetniemi, and A. L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18:319–349, 1986.
- [HKMP95] J. Hromkovic, R. Klasing, B. Monien, and R. Peine. *Combinatorial Network Theory*, chapter Dissemination of information in interconnection networks (broadcasting and gossiping), pages 125–212. Kluwer Academic, 1995.
- [Hoc97] D. S. Hochbaum, editor. *Approximation Algorithms for NP-Hard Problems*. PWS Publishing Company, 1997.

- [KK99] J. Kleinberg and A. Kumar. Wavelength conversion in optical networks. In *Proc of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'99)*, 1999.
- [Lei92] F. T. Leighton. *Introduction to Parallel Architectures : Arrays, Trees, Hypercubes*. Morgan Kaufmann Publishers, 1992.
- [MPS98] E. W. Mayr, H. J. Prömel, and A. Steger, editors. *Lectures on Proof Verification and Approximation Algorithms*, volume 1367 of *Lecture Notes in Computer Science*. Springer-Verlag, 1998.
- [Pan92] R. K. Pankaj. *Architectures for Linear Lightwave Networks*. PhD thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1992.
- [RS95] R. Ramaswami and K. N. Sivarajan. Routing and wavelength assignment in all-optical networks. *IEEE/ACM Transactions on Networking*, 3(5):489–500, 1995.
- [RS98] R. Ramaswami and G. Sasaki. Multiwavelength optical networks with limited wavelength conversion. *IEEE/ACM Transactions on Networking*, 6(6):744–754, December 1998.
- [VD93] R. J. Vetter and D. H. C. Du. Distributed computing with high-speed optical networks. *IEEE Computer*, 26(2):8–18, February 1993.

CHAPITRE 3

Routage WDM tout-optique

3.1 Réseaux généraux

3.1.1 Instances quelconques

Trouver un routage optique utilisant un minimum de longueurs d'onde (ou couleurs) n'est pas un problème facile. Dans le cas général, c'est-à-dire pour un graphe orienté G et une instance I quelconques, le problème du routage optique et de la détermination de $\vec{w}(G, I)$ a été montré \mathcal{NP} -complet par Erlebach et Jansen dans [EJ96]. En fait, ils ont montré cette propriété pour les arbres et les cycles orientés symétriques. Ils l'ont étendue aux grilles dans [EJ97] et au cas des arbres binaires avec un nombre de couleurs limité à 3, par réduction du problème de la coloration des arêtes d'un multigraphe, \mathcal{NP} -complet d'après Holyer [Hol81].

Théorème 3.1.1 (Erlebach et Jansen [EJ97]) — *Le problème suivant est \mathcal{NP} -complet :*

Données : *Un arbre binaire orienté symétrique T et une instance I .*

Question : *Existe-t-il un routage optique pour (T, I) utilisant au plus 3 couleurs ?*

Corollaire 3.1.2 — *Le problème suivant est \mathcal{NP} -complet :*

Données : *Un graphe orienté G , une instance I et un entier k .*

Question : *Existe-t-il un routage optique pour (G, I) utilisant au plus k couleurs ?*

D'un certain point de vue, la difficulté du problème de routage optique provient de la difficulté intrinsèque de deux problèmes sous-jacents : le problème du routage et le problème de la coloration des chemins.

Étant donné un graphe orienté G et une instance I , la recherche d'un routage R , réalisant I et minimisant la charge maximale $\vec{\pi}(G, I, R)$ des arcs de G , est intéressante pour le problème de routage optique puisque la charge d'un routage minore le nombre de couleurs nécessaires pour le colorer : $\vec{\pi}(G, I, R) \leq \vec{w}(G, I, R)$.

La détermination de $\vec{\pi}$ est étroitement reliée au problème du multiflot à capacités entières. Étant donné un graphe G muni de capacités entières sur ses liens et un ensemble de requêtes munies de demandes entières, il s'agit de trouver un flot de valeur correspondante de chaque source vers sa destination, de manière à ne pas dépasser les capacités des liens. D'après Even, Itai et Shamir [EIS76], même le bi-flot à capacités unitaires est un problème \mathcal{NP} -complet, dans les graphes orientés ou non. En transformant chaque requête munie de sa demande d en d copies de la même requête unitaire, le problème du multiflot consiste alors à router une instance I dans le graphe G sans dépasser les charges des liens autorisées. Il s'en suit immédiatement la \mathcal{NP} -complétude du problème de routage et de la détermination de $\vec{\pi}(G, I)$ ou de $\pi(G, I)$, puisqu'il s'agit de faire passer un multiflot unitaire dans un graphe à capacités constantes.

Théorème 3.1.3 (d'après Even, Itai et Shamir [EIS76]) — *Le problème suivant est \mathcal{NP} -complet :*

Données : *Un graphe G (orienté ou non), une instance I et un entier k .*

Question : *Existe-t-il un routage pour (G, I) de charge maximale k sur les liens de G ?*

Notons que dans le cas orienté, ce dernier résultat peut aussi s'obtenir à partir d'un théorème plus puissant de Fortune, Hopcroft et Willie [FHW80], établissant la \mathcal{NP} -difficulté de trouver deux chemins arc-disjoints pour deux requêtes dans un graphe orienté. En effet, router une instance dans un graphe sans dépasser une certaine charge c est équivalent à router les requêtes par des chemins disjoints dans le multi-graphe associé de multiplicité uniforme c .

À l'aide de cette propriété, nous pouvons donner un résultat encore plus fort démontrant la non-approximabilité du problème du routage optique dans le modèle orienté, comme cela a été remarqué par Jarry dans [Jar99].

Théorème 3.1.4 — *Le problème suivant est \mathcal{NP} -complet :*

Données : *Un graphe orienté G et une instance I .*

Question : *Se trouve-t-on dans l'une des deux situations suivantes ?*

- *Il existe deux chemins arc-disjoints dans G pour deux requêtes de I .*
- *Il n'existe pas de routage pour (G, I) par des chemins arc-disjoints.*

L'idée de la preuve de ce théorème consiste à construire une semi-grille G orientée (non-symétrique) en plaçant aux intersections des lignes et des colonnes des copies d'un sous-graphe \mathcal{G} pour lequel il est \mathcal{NP} -difficile d'après [FHW80] de décider si deux requêtes peuvent le traverser par deux chemins arc-disjoints. Une représentation schématique se trouve sur la figure 3.1. On définit alors un ensemble I de requêtes $\{(x_i, y_i) \mid 1 \leq i \leq l\}$ telles que leur réalisation par des chemins arc-disjoints est possible si et seulement si deux chemins arc-disjoints

existent pour traverser le sous-graphe \mathcal{G} . Dans le cas contraire, il n'existe pas deux chemins arc-disjoints dans G pour router deux requêtes de I .

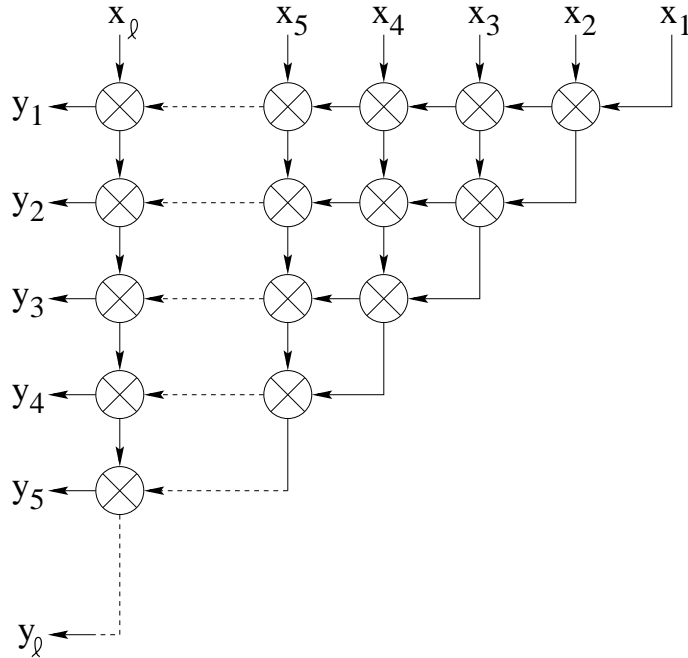


Figure 3.1: Semi-grille et routage disjoint non-approximable.

Ainsi, il est \mathcal{NP} -difficile de décider si une seule ou toutes les requêtes peuvent être routées avec une charge maximale unitaire sur les arcs du graphe, donc de décider si $\vec{w}(G, I) = 1$ ou si $\vec{w}(G, I) = |I|$, et nous en déduisons le corollaire suivant :

Corollaire 3.1.5 — *Il n'existe pas d'algorithme d'approximation pour le problème du routage optique dans un graphe orienté, à moins que $\mathcal{P} = \mathcal{NP}$.*

La question analogue dans le cas orienté symétrique reste ouverte.

Cependant, lorsque le nombre de requêtes est borné, le problème de routage devient polynômial, dans le cas non-orienté d'après un résultat de Robertson et Seymour [RS95], et dans le cas orienté symétrique d'après des travaux récents de Jarry [Jar99], avec toutefois des constantes cachées énormes qui rendent les algorithmes impraticables.

Nous invitons le lecteur intéressé par la complexité des problèmes de routage (et la solvabilité en temps polynômial de certains cas particuliers) à se référer au livre édité par Korte, Lovász, Prömel et Schrijver [KLPS90], et en particulier

au chapitre écrit par Frank [Fra90]. Une technique générale pour obtenir des algorithmes d'approximation (randomisés) pour le problème de routage (sous certaines conditions) est la technique randomisée d'arrondi due à Raghavan et Thompson. Les détails de leur méthode et les applications au routage VLSI et aux réseaux de multiflot peuvent être trouvés dans [RT87].

Une fois un routage connu pour réaliser une instance de communication, savoir colorer efficacement ses chemins est nécessaire dans le cadre du problème de routage optique. Malheureusement, ce problème de coloration restreint à des chemins est aussi difficile que le problème bien connu de la coloration des sommets d'un graphe, comme il a été remarqué par Chlamtac, Ganz et Karmi dans [CGK92]. Étant donné un graphe non-orienté H quelconque, il suffit de définir un ensemble R de chemins ad-hoc dans un graphe G également construit à dessein, de telle sorte que H soit le graphe de conflit de R . Par ailleurs, le résultat découle immédiatement du premier théorème de ce chapitre, puisque le routage est fixé dans un arbre.

Proposition 3.1.6 — *Le problème suivant est \mathcal{NP} -complet :*

Données : *Un graphe G (orienté ou non) et un routage R .*

Question : *Existe-t-il une coloration des chemins de R utilisant 3 couleurs ?*

Dans le modèle non-orienté, la \mathcal{NP} -complétude du problème de routage optique provient de résultats bien antérieurs à l'avènement de la technologie WDM. Concernant le problème de la coloration de chemins dans un arbre, Golubic et Jamison [GJ85b] ont établi que la détermination de $\mathfrak{w}(G, I)$ est un problème \mathcal{NP} -complet même pour les étoiles (arbres dont tous les sommets sauf un ont degré 1), en réduisant à ce problème celui de la coloration des arêtes d'un multigraphe qui est \mathcal{NP} -complet d'après Holyer [Hol81] (voir section 3.2.1, page 78). Ce résultat a été étendu dans [EJ96] au cas des cycles non-orientés, alors que le problème devient polynômial pour les arbres de degré borné d'après [EJ97].

Au vu de ce dernier résultat et de la \mathcal{NP} -difficulté à déterminer $\vec{\mathfrak{w}}(G, I)$ pour les arbres binaires orientés, on pourrait croire que le problème de routage optique est plus difficile dans le modèle orienté. Ce n'est pas vrai en général. Par exemple, pour les étoiles, le problème non-orienté est \mathcal{NP} -complet [GJ85b] alors que le problème orienté est polynômial, par réduction au problème de couplage dans les graphes bipartis (voir section 3.2.1, page 76), soluble d'après le théorème de König-Hall.

Prenons un graphe orienté G et un routage R réalisant une instance I . Soient L la longueur maximale des chemins dans R et Δ le degré maximum du graphe de conflit associé à R . Puisque tout chemin est en conflit avec au plus $\vec{\pi}(G, I, R) - 1$ autres chemins sur chacun de ses arcs, il est clair que $\Delta \leq L(\vec{\pi}(G, I, R) - 1)$. Une

coloration gloutonne du graphe de conflit permet d'utiliser au plus $\Delta + 1$ couleurs. Nous en déduisons une borne asymptotique pour le problème de routage optique (G, I) : $\vec{w} = O(L\vec{\pi})$, où L représente ici la plus petite longueur maximale des chemins d'un routage R tel que $\vec{\pi}(G, I, R) = \vec{\pi}(G, I)$. Le même raisonnement dans le modèle non-orienté conduit également à $\mathbf{w} = O(L\pi)$.

Une stratégie gloutonne plus élaborée pour colorer les chemins d'un routage est proposée dans [ABNC⁺94, ABNC⁺96]. Elle s'avère plus efficace que la précédente si le paramètre L est suffisamment grand. Reprenons les mêmes notations et notons m le nombre d'arcs du graphe G . La somme des charges des arcs de G vaut au plus $m\vec{\pi}(G, I, R)$, or elle est égale d'autre part à la somme des longueurs des chemins de R . Par conséquent, le nombre de chemins de longueur au moins \sqrt{m} ne peut excéder $\vec{\pi}(G, I, R)\sqrt{m}$. Une couleur spécifique est attribuée à chacun de ces chemins. En considérant le graphe de conflit des chemins restants et en appliquant la même coloration gloutonne que précédemment, on peut compléter la coloration en utilisant au plus $\vec{\pi}(G, I, R)\sqrt{m}$ couleurs supplémentaires. Il en découle le résultat suivant, formulé initialement dans le modèle non-orienté :

Théorème 3.1.7 (Aggarwal et al. [ABNC⁺94]) — *Pour tout problème (G, I) , où G a m arcs,*

$$\vec{w}(G, I) \leq 2\vec{\pi}(G, I)\sqrt{m}.$$

Dans le même temps, il est montré dans [ABNC⁺94, ABNC⁺96] que les deux bornes supérieures issues des colorations gloutonnes mentionnées ci-dessus sont optimales à un facteur constant près. Plus précisément, il est fourni une famille de problèmes non-orientés, dont une représentation est donnée par la figure 3.2, telle que $\mathbf{w} = \Omega(\pi \cdot \min\{L, \sqrt{m}\})$, où L représente la distance maximale entre deux sommets formant une requête et m le nombre d'arêtes du graphe.

Cette famille de cas pathologiques repose sur une topologie en "pseudo-grille", telle qu'elle est représentée sur la figure 3.2. Pour un entier $n \geq 2$ donné, la pseudo-grille G_n est constituée de n nœuds sources $\{s_1, s_2, \dots, s_n\}$, de $2n - 2$ colonnes de $2n - 1$ nœuds chacune et de n nœuds destinations $\{t_1, t_2, \dots, t_n\}$, reliés entre eux comme il est montré sur la figure 3.2 pour $n = 4$. Le nombre d'arêtes de G_n est ainsi $m = O(n^2)$. Notons I_n l'instance $\{(s_i, t_i)\}_{1 \leq i \leq n}$ et I_n^k l'instance formée par k copies de l'instance I_n . Dans tous les cas, la distance maximale pour une requête vaut $L = 4n - 3 = O(n)$. Remarquons d'abord que deux requêtes quelconques dans I_n entrent nécessairement en conflit, puisqu'il n'existe pas deux chemins arête-disjoints reliant deux paires source-destination. On en déduit que $\mathbf{w}(G_n, I_n) = n = \Omega(\min\{L, \sqrt{m}\})$. Par ailleurs, notons que l'instance I_n peut être routée avec une charge maximale de $\pi(G_n, I_n) = 2 = O(1)$ (voir figure 3.2). En considérant à présent l'instance I_n^k , on obtient pour la

même raison que $\vec{w}(G_n, I_n^k) = kn$, or $\pi(G_n, I_n^k) = O(k)$. Nous avons donc bien $\vec{w}(G_n, I_n^k) = kn = \Omega(\pi \cdot \min \{L, \sqrt{m}\})$.

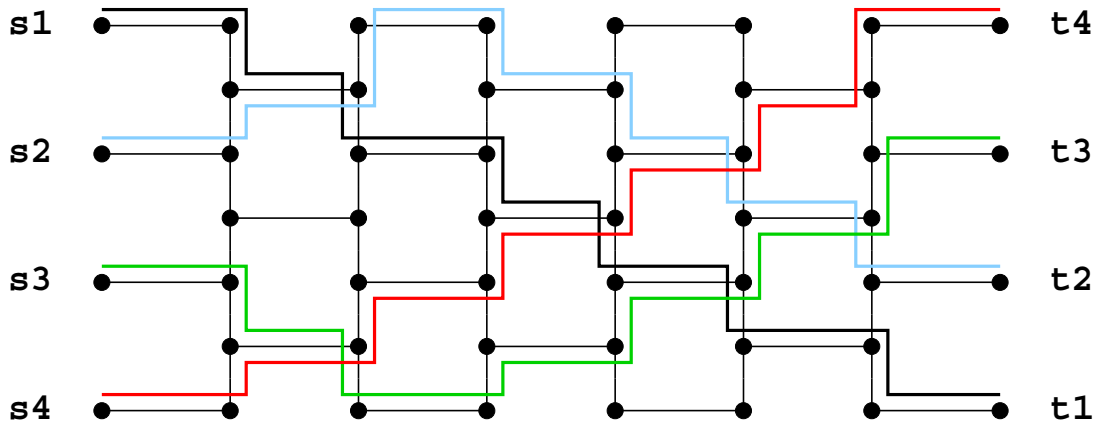


Figure 3.2: Instance pathologique dans une pseudo-grille non-orientée.

La question se pose alors de savoir si ce dernier résultat est transposable dans le modèle orienté. Nous pouvons répondre par l'affirmative, mais le problème reste ouvert dans le cas orienté symétrique.

Proposition 3.1.8 — *Pour tout π et tout L , il existe un graphe orienté G et une instance I tels que*

$$\vec{\pi}(G, I) \geq \pi, \max_{(x,y) \in I} d(x, y) \geq L \text{ et } \vec{w}(G, I) = \Omega(\pi L) .$$

La preuve de cette proposition est en tous points semblable à la précédente démonstration. Elle repose sur la construction de la figure 3.3, identique à celle de la figure 3.2 si ce n'est l'orientation des arcs. Cette orientation permet de conserver la propriété que deux requêtes entrent nécessairement en conflit, du fait qu'il n'existe pas deux chemins arc-disjoints reliant deux couples source-destination. Notons que la même topologie orientée symétriquement ne possède plus cette propriété, puisqu'alors deux chemins ont la possibilité de se croiser verticalement en sens inverse sans générer de conflit.

Nous avons vu précédemment que la résolution du problème de routage optique peut être décomposée en deux étapes (trouver un routage de charge minimale et colorer ses chemins), chacune d'elles demeurant cependant intrinsèquement difficile à optimiser. Il est à noter que cette approche souvent utilisée ne conduit pas nécessairement à une solution optimale, même si chacun des deux sous-problèmes est résolu de façon optimale. En effet, nous pouvons

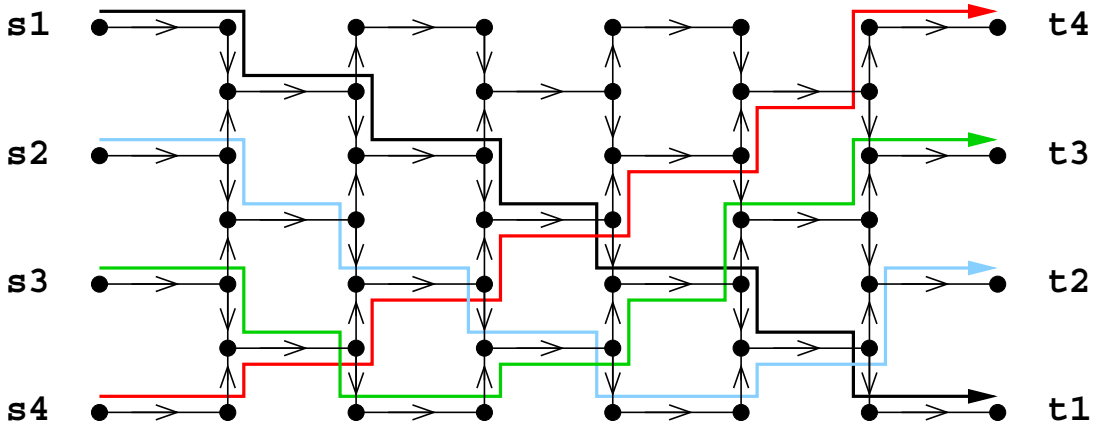


Figure 3.3: Instance pathologique dans une pseudo-grille orientée non-symétrique.

construire un problème (G, I) tel que tout routage R réalisant I , de charge minimale $\vec{\pi}(G, I, R) = \vec{\pi}(G, I)$, n'est pas colorable avec $\vec{w}(G, I)$ couleurs. Autrement dit, un routage optique optimal pour ce problème particulier n'est pas un routage optimal en termes de charge (ce qui répond à une question posée dans [BBG⁺97]).

Proposition 3.1.9 — *Il n'existe pas toujours de routage R réalisant une instance I dans un graphe orienté G , tel que $\vec{\pi}(G, I, R) = \vec{\pi}(G, I)$ et $\vec{w}(G, I, R) = \vec{w}(G, I)$.*

PREUVE. Nous donnons une construction dans un graphe orienté non-symétrique, qui reprend la structure en pseudo-grille venant d'être présentée. Pour tout entier $n \geq 2$, le graphe H_n est construit à partir du graphe G_n représenté sur la figure 3.3, auquel on ajoute deux sommets s et t ainsi que l'arc (s, t) et tous les arcs (s_i, s) , (t, t_i) , pour $1 \leq i \leq n$, comme le montre la figure 3.4 pour $n = 4$.

Nous reprenons la notation I_n pour l'instance $\{(s_i, t_i)\}_{1 \leq i \leq n}$. Comme précédemment, nous avons $\vec{\pi}(H_n, I_n) = 2$. Par conséquent, dans tout routage R_n réalisant I_n , de charge maximale $\vec{\pi}(H_n, I_n, R_n) = \vec{\pi}(H_n, I_n)$, au moins $n - 2$ chemins n'utilisent pas l'arc (s, t) , d'où $\vec{w}(H_n, I_n, R_n) \geq n - 2$. Or il n'est pas difficile de voir que $\vec{w}(H_n, I_n) = \lceil n/2 \rceil$, puisqu'en routant x requêtes par l'arc (s, t) et $n - x$ requêtes par le sous-graphe G_n , on obtient un routage colorable avec $\max\{x, n - x\}$ couleurs.

Notons enfin qu'en définissant le multigraphe H'_n à partir du graphe H_n en multipliant l'arc (s, t) par un facteur $c = n/4 - 1$ (ou en reliant s à t par c chemins arc-disjoints, afin que H'_n soit un graphe simple), on obtient alors que $\vec{\pi}(H'_n, I_n) = 2$ et $\vec{w}(H'_n, I_n) = 4$, alors que tout routage R_n de charge 2 nécessite

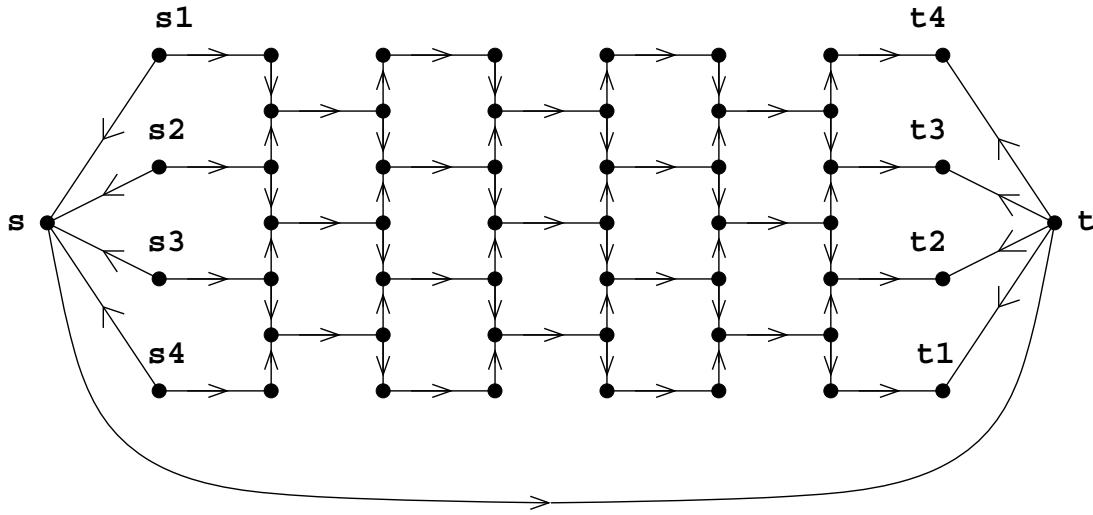


Figure 3.4: Pseudo-grille H_4 avec raccourci.

$n/2 - 2$ couleurs. Le rapport $\vec{w}(H'_n, I_n, R_n)/\vec{w}(H'_n, I_n)$ est ainsi de l'ordre de $n/8$. \square

Question 3.1.10 — Qu'en est-il des propositions 3.1.8 et 3.1.9 dans le cas orienté symétrique ?

La première partie de cette question revient à se trouver des constructions orientées symétriques telles que les paramètres \vec{w} et $\vec{\pi}$ diffèrent significativement. De telles constructions pourraient alors éventuellement être exploitées pour répondre à la seconde partie de la question, comme il a été fait précédemment pour la proposition 3.1.9.

Dans cette perspective, nous pouvons fournir un élément de réponse en considérant le cycle C_N . Nous allons définir une instance I_N telle que $\vec{w}(C_N, I_N) = N/2 = 2\vec{\pi}(C_N, I_N) - 1$. Ainsi, dans le cas orienté symétrique, \vec{w} et $\vec{\pi}$ peuvent différer d'un facteur 2.

Posons $N = 2(2k + 1)$ et représentons les sommets de C_N par les entiers modulo N . Soient $J_N = \{(2i \bmod N, 2i + N/2 \bmod N)\}_{1 \leq i \leq N/2}$ et I_N l'instance formée par deux copies de J_N . Nous avons donc dans I_N un total de N requêtes et clairement toutes les requêtes qui sont routées dans le même sens entrent en conflit. Nous en déduisons que $\vec{w}(C_N, I_N) = N/2$. Par ailleurs, nous pouvons montrer qu'un routage optimal pour la charge consiste à router dans les deux sens opposés chaque paire de requêtes identiques, comme il est montré sur la figure 3.5 pour $N = 6$, ce qui donne une charge maximale de $k + 1 = \lceil N/4 \rceil$. Comme tout routage pour I_N comporte N chemins de longueur $N/2$ et le cycle C_N contient $2N$

arcs, $N^2/4N = N/4$ est la charge moyenne des arcs et donc $\vec{\pi}(C_N, I_N) = \lceil N/4 \rceil$.

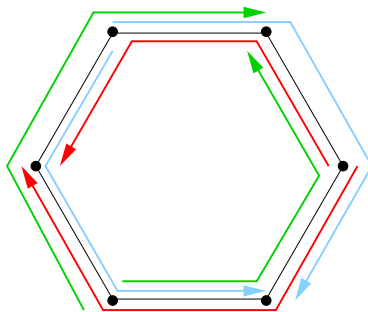


Figure 3.5: Routage optique pour (C_6, I_6) .

Nous allons à présent nous intéresser aux liens entre le problème de routage optique et l'*arc-connexité*.

DÉFINITIONS. Soit $G = (V, A)$ un graphe orienté fortement connexe.

- Pour tout $S \subseteq V(G)$, on note $m(S)$ le nombre d'arcs de $A(G)$ ayant seulement leur extrémité initiale dans S .
- L'*arc-connexité* $\lambda(G)$ de G est le nombre minimum d'arcs dont la suppression rend le graphe non fortement connexe : $\lambda(G) = \min_{\emptyset \subset S \subset V(G)} m(S)$.
- On dit que G est *k-arc-connexe* si $\lambda(G) \geq k$.
- L'*arc-connexité* $c_G(r)$ d'un sommet $r \in V(G)$ est le nombre minimum d'arcs dont la suppression rend un sommet de $V(G)$ non accessible à partir de r : $c_G(r) = \min_{r \in S \subset V(G)} m(S)$.

Les versions orientées des théorèmes de Menger, que l'on peut trouver par exemple dans le livre de Berge [Ber83], permettent de relier les notions globale et locale de l'*arc-connexité* avec le problème du routage arc-disjoint. Nous donnons ici la version locale.

Théorème 3.1.11 (Menger) — *Pour tout couple (r, t) de sommets dans un graphe orienté G , le nombre maximum de chemins arc-disjoints de r vers t est $c_G(r)$.*

Dans un graphe orienté, un **arbre de recouvrement** enraciné est un arbre qui permet d'atteindre tout sommet à partir de la racine. Le théorème suivant montre que l'*arc-connexité* donne une propriété encore plus forte que celle décrite par le théorème de Menger.

Théorème 3.1.12 (Edmonds [Edm72]) — *Pour tout sommet r dans un graphe orienté G , le nombre maximum d'arbres de recouvrement arc-disjoints enracinés en r est $c_G(r)$.*

Une conséquence très intéressante de ce théorème fondamental a été établie dans [Shi79].

Corollaire 3.1.13 (Shiloach (et Tarjan) [Shi79]) — *Pour k couples quelconques de sommets $(s_1, t_1), \dots, (s_k, t_k)$ dans un graphe orienté k -arc-connexe, il existe k chemins arc-disjoints P_1, \dots, P_k tels que P_i est un chemin de s_i vers t_i , pour $1 \leq i \leq k$.*

PREUVE. Ajoutons au graphe G un sommet r ainsi que les arcs (r, s_i) , pour $1 \leq i \leq k$, pour former le graphe G' . Le graphe G étant k -arc-connexe, $c_{G'}(r) = k$. En appliquant le théorème 3.1.12, on obtient k arbres de recouvrement arc-disjoints enracinés en r et un chemin dans chacun respectivement de r vers chaque t_i en passant par s_i , ce qui induit le résultat. \square

Tous les résultats mentionnés sur l'arc-connexité sont accompagnés dans la littérature d'algorithmes polynômiaux permettant d'exhiber des solutions effectives. Le dernier résultat nous permet d'en déduire une méthode efficace pour résoudre un problème de routage optique (G, I) dans un graphe d'arc-connexité λ . En partitionnant l'ensemble des requêtes en $\lceil |I|/\lambda \rceil$ sous-ensembles de taille au plus λ , on peut alors router chacun par des chemins arcs-disjoints et lui attribuer ainsi une longueur d'onde spécifique.

Corollaire 3.1.14 — *Pour tout graphe orienté G d'arc-connexité λ et pour toute instance I , le problème de routage optique (G, I) peut être résolu en temps polynômial en utilisant $\lceil |I|/\lambda \rceil$ longueurs d'onde.*

3.1.2 Permutations et k -relations

Dans sa thèse [Pan92] et dans [PG95], Pankaj a déterminé une borne inférieure sur la charge nécessaire dans le pire des cas pour router une instance de permutation dans un graphe quelconque de degré maximum donné. La technique utilisée repose sur la construction d'une instance de permutation dont les requêtes sont distantes, de manière à maximiser la somme des distances de l'instance. En divisant cette somme par le nombre de liens dans le graphe, on obtient une minoration de la charge moyenne d'un routage, et donc de la charge maximale.

Plus précisément, il est montré que dans tout graphe orienté G de degré entrant ou sortant maximum $\Delta \geq 2$, il existe une instance de permutation I_1

telle que $\vec{\pi}(G, I_1) \geq \frac{\lfloor \log_{\Delta} N/2 \rfloor}{2\Delta}$. La preuve donnée dans [PG95] peut être modifiée afin d'améliorer le résultat. Aussi nous donnons directement le théorème optimisé et sa preuve :

Proposition 3.1.15 — *Pour tout graphe orienté G de degrés maximum sortant et entrant $\Delta^+, \Delta^- \geq 2$, avec $\Delta = \max\{\Delta^+, \Delta^-\}$ et $\delta = \min\{\Delta^+, \Delta^-\}$, il existe une instance de permutation I_1 telle que*

$$\vec{w}(G, I_1) \geq \vec{\pi}(G, I_1) \geq \frac{\lfloor \log_{\Delta} (N-1) \rfloor}{\delta}.$$

PREUVE. Supposons que $\Delta = \Delta^+$ (et $\delta = \Delta^-$). Le cas inverse est identique par symétrie. Pour tout sommet x dans G , il y a au plus Δ sommets à distance 1 de x , et plus généralement au plus Δ^d à distance d . Donc le nombre de sommets à distance $< d$ de x est au plus

$$\sum_{i=0}^{d-1} \Delta^i = \frac{\Delta^d - 1}{\Delta - 1} \leq \Delta^d$$

Ainsi, en posant $d_0 = \lfloor \log_{\Delta} (N-1) \rfloor$, on obtient que le nombre de sommets à distance $< d_0$ de x est au plus $N-1$. Par conséquent, il existe un sommet à distance au moins d_0 de tout sommet fixé.

Construisons à présent un graphe biparti $H = (V_1, V_2, E)$ où V_1 et V_2 sont deux copies de $V(G)$ et tel que l'arête $\{x_1, y_2\}$ est dans E si et seulement si $d_G(x, y) \geq d_0$, où x et y sont les deux sommets correspondants dans G . D'après ce qui précède, tous les sommets de V_1 ont un degré au moins 1, et les sommets de V_2 aussi par le raisonnement symétrique, puisque $d_0 \leq \lfloor \log_{\Delta^-} (N-1) \rfloor$. Le théorème de König-Hall (dont un corollaire affirme qu'il existe dans tout (multi)graphe biparti un couplage saturant tous les sommets de degré maximum, voir par exemple [Ber83]) implique alors l'existence d'un couplage parfait dans H , qui induit de façon naturelle une permutation I_1 sur les sommets de G (toute arête $\{x_1, y_2\}$ du couplage induit la requête (x, y)).

Nous avons donc montré l'existence d'une instance de permutation dans le graphe G qui nécessite N chemins, chacun de longueur au moins d_0 . La charge d'un routage pour cette instance est minorée par le rapport entre la somme des longueurs de ses chemins et le nombre d'arcs valant au plus δN , ce qui donne $\vec{w}(G, I_1) \geq \vec{\pi}(G, I_1) \geq \frac{d_0 N}{\delta N} = \frac{\lfloor \log_{\Delta} (N-1) \rfloor}{\delta}$. \square

La même technique de démonstration permet d'obtenir un résultat encore plus satisfaisant dans le cas des graphes orientés symétriques et *sommet-transitifs*.

DÉFINITION. Un graphe orienté $G = (V, A)$ est **sommet-transitif** si pour tout couple de sommets (a, b) , il existe une bijection ϕ sur l'ensemble $V(G)$, telle que $\phi(a) = b$ et $(x, y) \in A(G)$ si et seulement si $(\phi(x), \phi(y)) \in A(G)$.

Proposition 3.1.16 (Pankaj [Pan92]) — *Pour tout graphe orienté symétrique et sommet-transitif G de diamètre D et de degré maximum Δ , il existe une instance de permutation I_1 telle que*

$$\vec{w}(G, I_1) \geq \vec{\pi}(G, I_1) \geq \frac{D}{\Delta}.$$

REMARQUE. Notons que cette dernière borne ne peut être améliorée, puisqu'elle s'applique exactement au cas de l'hypercube, où $D/\Delta = 1$. En effet, comme nous le verrons ultérieurement, l'instance formée des couples de sommets antipodaux (de charge uniforme 1) peut être routée avec une seule longueur d'onde.

Intéressons-nous à présent aux k -relations. Pour rappel, une k -relation est une instance de communication où chaque sommet est origine et destination d'au plus k requêtes. Par le théorème de König-Hall, on peut partitionner une k -relation en k instances de permutation. Réciproquement, l'union de k instances de permutation donne une k -relation (qui peut être une multi-instance, où certaines requêtes peuvent apparaître plusieurs fois). Ainsi, les deux théorèmes précédents peuvent être facilement adaptés au cas des k -relations, en multipliant les bornes inférieures par k .

En s'inspirant de la construction de la figure 3.2 donnée par Aggarwal et al. [ABNC⁺94, ABNC⁺96], Raghavan et Upfal ont obtenu dans [RU94] une borne inférieure pour le routage optique non-orienté des k -relations, qui tient compte de l'*arête-expansion* du graphe. En adaptant leur preuve à l'aide de la construction de la figure 3.3, on obtient un résultat équivalent dans le modèle orienté, qui tient compte de l'*arc-expansion* du graphe.

DÉFINITION. Soit $G = (V, A)$ un graphe orienté. Pour tout $S \subseteq V(G)$, soit $m(S)$ le nombre d'arcs ayant seulement leur extrémité initiale dans S . L'*arc-expansion* de G est alors définie par

$$\beta(G) = \min_{1 \leq |S| \leq N/2} \frac{m(S)}{|S|}.$$

Ce paramètre a été étudié dans le cas des graphes non-orientés par Mohar [Moh89] sous le nom de nombre isopérimétrique. Une définition légèrement différente a été considérée par Solé dans [Sol95]. Son inverse peut se voir comme un minorant de la charge nécessaire pour router une permutation dans le pire cas. On peut de même montrer facilement que pour tout graphe G orienté d'ordre N , $\beta(G) \cdot \vec{\pi}(G, I_A) \geq N/2$, où I_A est l'instance de l'échange total. Ce paramètre vaut par exemple $\beta = N/2$ pour le graphe complet d'ordre N et $\beta = 2/N$ pour le

chemin de longueur N , et pour ces deux graphes on a l'égalité dans la relation précédente si N est pair. Dans le cas du chemin, il est facile de construire une k -relation comprenant $kN/2$ requêtes devant toutes utiliser le même arc central. Le nombre de couleurs pour cette instance sera alors égal à $\vec{w} = kN/2 = k/\beta$. Le résultat suivant montre qu'on peut se trouver dans une situation encore pire, en termes d'arc-expansion.

Proposition 3.1.17 ((d'après Raghavan et Upfal [RU94])) — *Pour tout $\beta \leq 1$ et tout $1 \leq k \leq N$, il existe un graphe orienté G planaire, de degré borné et d'arc-expansion β , et une k -relation I_k , tels que*

$$\vec{w}(G, I_k) = \Omega(k/\beta^2) .$$

PREUVE. Considérons la pseudo-grille orientée G_n , représentée pour $n = 4$ sur la figure 3.3, sur laquelle on greffe un chemin orienté (symétrique) de longueur n sur chaque s_i et t_i , pour $1 \leq i \leq n$. Le graphe orienté H_n ainsi obtenu possède une arc-expansion β approximativement égale à $1/n$, pour n grand. Constituons à présent l'instance I_k à l'aide de k requêtes (s_i^j, t_i^j) pour $1 \leq i, j \leq n$, où s_i^j et t_i^j sont les j -èmes sommets des chemins reliés respectivement à s_i et t_i . Il s'en suit que I_k est une k -relation et que toutes ses requêtes entrent en conflit, pour les mêmes raisons que précédemment. Nous avons donc $\vec{w}(H_n, I_k) = kn^2 = \Omega(k/\beta^2)$. \square

La dernière proposition repose sur l'adaptation au cas orienté qui a permis d'obtenir la proposition 3.1.8 et qui ne connaît pas d'équivalent dans le cas orienté symétrique. La question suivante se pose donc sur le même plan que la question 3.1.10.

Question 3.1.18 — La proposition 3.1.17 est-elle vérifiée pour les graphes orientés symétriques ?

Étant donné un problème de routage optique (G, I) , nous avons établi au début de ce chapitre que $\vec{w}(G, I) \leq L(R)\vec{\pi}(G, I, R)$ pour tout routage R réalisant l'instance I , où $L(R)$ est la longueur maximale des chemins de R . Ainsi, il est intéressant de savoir router une instance en minimisant à la fois la charge et la longueur des chemins. Ce problème a été traité par Leighton et Rao [LR88] dans le modèle non-orienté et dans le cadre des réseaux de multiflot. Le résultat qui nous intéresse concerne les instances de permutation et peut être transposé au cas orienté symétrique, c'est pourquoi nous le formulons sous cette forme.

Théorème 3.1.19 (Leighton et Rao [LR88]) — *Pour tout graphe orienté symétrique G de degré borné et d'arc-expansion β , et pour toute instance de permutation I_1 , on peut obtenir en temps polynômial un routage R réalisant I_1 tel que $\vec{\pi}(G, I_1, R) = O(\log N/\beta)$ et $L(R) = O(\log N/\beta)$, où $L(R)$ est la longueur maximale des chemins de R .*

Malheureusement, aucun résultat équivalent pour les graphes orientés non-symétriques n'est connu, les problèmes de multiflot étant en général plus difficiles à résoudre dans ce cas. En utilisant le théorème précédent, Aumann et Rabani [AR95] ont obtenu une borne supérieure qui découle de la remarque ci-dessus et du fait qu'une k -relation peut être décomposée en k permutations. Au vu de la proposition 3.1.17, le résultat suivant constitue une bonne approximation asymptotique du problème de routage optique pour les k -relations dans les cas les pires.

Corollaire 3.1.20 (Aumann et Rabani [AR95]) — *Pour tout graphe orienté symétrique G de degré borné et d'arc-expansion β , et pour toute k -relation I_k , le problème de routage optique (G, I_k) peut être résolu en temps polynomial en utilisant $O(k \log^2 N / \beta^2)$ longueurs d'onde.*

3.1.3 Diffusion et multicast

Une instance de *diffusion* dans un graphe $G = (V, A)$ est constituée de tous les couples (u_0, v) pour un sommet u_0 fixé : $I_0(u_0) = \{(u_0, v) \mid v \in V(G), v \neq u_0\}$. Il s'agit donc de router et colorer $N - 1$ requêtes de communication. Si le degré sortant de u_0 est noté $d^+(u_0)$, au moins $|I_0(u_0)|/d^+(u_0)$ chemins doivent nécessairement partager un arc issu de u_0 , d'où $\vec{w}(G, I_0(u_0)) \geq \vec{\pi}(G, I_0(u_0)) \geq \lceil \frac{N-1}{d^+(u_0)} \rceil$. Par ailleurs, si G est k -arc-connexe, en utilisant le corollaire 3.1.14 on a $\vec{w}(G, I_0(u_0)) \leq \lceil \frac{N-1}{k} \rceil$. Donc, comme il a été montré dans [BGP⁺96], cela nous donne la valeur exacte de $\vec{w}(G, I_0(u_0))$ pour la diffusion dans un graphe k -arc-connexe à partir d'un sommet de degré sortant k .

Proposition 3.1.21 (Bermond et al. [BGP⁺96]) — *Pour tout graphe orienté G k -arc-connexe et pour une instance de diffusion $I_0(u_0)$ à partir d'un sommet u_0 de degré sortant k ,*

$$\vec{w}(G, I_0(u_0)) = \vec{\pi}(G, I_0(u_0)) = \left\lceil \frac{N-1}{k} \right\rceil .$$

Ce dernier résultat est intéressant car il s'applique pour diffuser à partir d'un sommet de degré sortant minimum dans les graphes dits *d'arc-connexité maximale*, pour lesquels l'arc-connexité est justement égale au degré minimum sortant. D'après un théorème de Mader [Mad71], les graphes sommet-transitifs font partie de cette famille, et donc également tous les graphes de Cayley.

Une généralisation de la proposition précédente a été établie dans [BHP98] et fait l'objet de l'article en annexe A. Une instance de *multicast* est un multi-ensemble de requêtes tel que la projection sur la première coordonnée est réduite

à un sommet unique. Ainsi une instance de diffusion est un cas particulier de multicast.

Théorème 3.1.22 (Beauquier, Hell et Pérennes [BHP98]) — *Pour tout graphe orienté G et toute instance de multicast I_M , le problème de routage optique (G, I_M) peut être résolu de façon optimale en temps polynômial, et de plus,*

$$\vec{w}(G, I_M) = \vec{\pi}(G, I_M) .$$

3.1.4 Échange total

Nous rappelons que l'instance de l'échange total dans un graphe G est définie par $I_A = V(G) \times V(G)$. Tous les résultats de cette section sont accompagnés d'algorithmes optimaux résolvant les problèmes en temps polynômial. Notons que la complexité de la détermination des paramètres $\vec{\pi}(G, I_A)$ et $\vec{w}(G, I_A)$ en général demeure un problème ouvert.

Bermond et al. dans [BGP⁺96] et indépendamment Wilfong dans [Wil96] ont traité le cas des cycles.

Théorème 3.1.23 — *Dans tout cycle orienté symétrique C_N ,*

$$\vec{w}(C_N, I_A) = \vec{\pi}(C_N, I_A) = \lceil \lfloor N^2/4 \rfloor / 2 \rceil .$$

Le cas de l'hypercube a été résolu par Pankaj dans [Pan92] et indépendamment par Bermond et al. dans [BGP⁺96].

Théorème 3.1.24 — *Dans tout hypercube orienté symétrique H_d de dimension d et d'ordre $N = 2^d$,*

$$\vec{w}(H_d, I_A) = \vec{\pi}(H_d, I_A) = 2^{d-1} .$$

Les trois résultats suivants font l'objet de l'article en annexe B. Le premier généralise le théorème précédant et a été montré également en partie dans [Tog98b, Tog98a].

Théorème 3.1.25 (Beauquier [Bea99]) — *Soient n_1, n_2, \dots, n_d des entiers tels que $2 \leq n_1 \leq n_2 \leq \dots \leq n_d$. En notant par $K(n_1, n_2, \dots, n_d)$ la somme cartésienne des d graphes complets K_{n_i} ($1 \leq i \leq d$) orientés symétriques,*

$$\vec{w}(K(n_1, n_2, \dots, n_d), I_A) = \vec{\pi}(K(n_1, n_2, \dots, n_d), I_A) = \prod_{i=1}^d n_i .$$

Théorème 3.1.26 (Beauquier [Bea99]) — Dans toute grille torique carrée C_n^d de dimension d , de côté n et d'ordre $N = n^d$,

$$\vec{w}(C_n^d, I_A) = \vec{\pi}(C_n^d, I_A) = n^{d+1}/8, \text{ si } n \text{ est pair, et sinon,}$$

$$\vec{\pi}(C_n^d, I_A) = (n^2 - 1)n^{d-1}/8 \leq \vec{w}(C_n^d, I_A) \leq (n + 1)^{d+1}/8 = \vec{w}(C_{n+1}^d, I_A)$$

Théorème 3.1.27 (Beauquier [Bea99]) — Dans toute grille carrée P_n^d de dimension d , de côté n et d'ordre $N = n^d$,

$$\vec{w}(P_n^d, I_A) = \vec{\pi}(P_n^d, I_A) = n^{d+1}/4, \text{ si } n \text{ est pair, et sinon,}$$

$$\vec{\pi}(P_n^d, I_A) = (n^2 - 1)n^{d-1}/4 \leq \vec{w}(P_n^d, I_A) \leq (n + 1)^{d+1}/4 = \vec{w}(P_{n+1}^d, I_A)$$

Enfin, le cas des arbres de cycles a été considéré dans [BPT99] et fait l'objet de l'annexe C. Notons que cette étude traite de graphes munis d'une pondération sur les sommets et d'un échange total pondéré en conséquence.

DÉFINITION. Un **arbre de cycles** est un graphe formé par une union de cycles qui s'intersectent deux-à-deux en au plus un sommet, et tel que deux sommets quelconques peuvent être reliés par exactement deux chemins arête-disjoints.

Théorème 3.1.28 (Beauquier, Pérennes et Tóth [BPT99]) — Dans tout arbre de cycles pondéré \mathcal{T} , $\vec{w}(\mathcal{T}, I_A) = \vec{\pi}(\mathcal{T}, I_A)$.

Au vu de ces résultats variés, une perspective de recherche intéressante serait de montrer l'égalité entre les paramètres \vec{w} et $\vec{\pi}$ pour l'échange total dans tout graphe orienté, à moins de trouver un contre-exemple. La question reste ouverte.

Question 3.1.29 — L'égalité $\vec{w}(G, I_A) = \vec{\pi}(G, I_A)$ est-elle vérifiée pour l'instance d'échange total I_A dans tout graphe orienté G ?

Nous pensons que c'est au moins le cas pour les graphes orientés symétriques.

Conjecture 3.1.30 — L'égalité $\vec{w}(G, I_A) = \vec{\pi}(G, I_A)$ est vérifiée pour l'instance de l'échange total I_A dans tout graphe orienté symétrique G .

L'étude du paramètre $\pi(G, I_A)$ en non-orienté, qui a été introduit initialement (pour la charge des sommets) par Chung et al. dans [CCRS87] sous le nom d'indice de transmission (*forwarding index*) et qui est souvent relié au cas orienté, a donné lieu à des résultats généraux par Heydemann, Meyer et Sotteau dans [HMS89] et par Heydemann et al. dans [HMOS94].

3.2 Réseaux particuliers

Lorsque la topologie du réseau est prédéterminée, il est souvent plus facile d'exploiter ses propriétés structurelles connues à l'avance afin de résoudre efficacement le problème du routage optique. Quand ce problème demeure \mathcal{NP} -complet malgré la restriction à une certaine classe de réseaux et/ou à une certaine famille d'instances de communication, on parvient cependant à obtenir de meilleures approximations que dans le cas général. Nous considérons en premier lieu le cas des arbres, très fréquemment utilisés dans les réseaux de télécommunications longue distance, puis nous traiterons certaines familles de graphes plus adaptés aux réseaux locaux, de par leur forte connexité.

3.2.1 Arbres

Dans notre contexte, un **arbre** est un graphe orienté symétrique fortement connexe et sans circuit. Cela revient à dire qu'il existe un seul chemin élémentaire reliant un couple de sommets donnés. Par conséquent, le problème du routage ne se pose pas dans les arbres : toutes les requêtes sont routées par les plus courts chemins.

L'exemple le plus simple d'arbre est le chemin. Dans ce cas, la coloration des requêtes peut se faire indépendamment pour celles qui vont dans un sens et dans l'autre, et chacun de ces deux sous-problèmes devient un problème polynômial bien connu : la coloration des sommets du **graphe d'intervalles** correspondant. C'est le graphe de conflit constitué d'un sommet par requête et d'une arête entre requêtes qui se chevauchent. Une coloration gloutonne dans l'ordre des sources des requêtes donne une solution optimale. Ainsi $\vec{w}(P_N, I) = \vec{\pi}(P_N, I)$ pour toute instance I dans tout chemin orienté symétrique P_N de longueur $N - 1$.

Une autre famille d'arbres relativement simple est celle des étoiles. Une **étoile** est un arbre ayant un seul sommet de degré supérieur à 1. L'étoile orientée symétrique d'ordre $N \geq 3$ est notée S_N .

L'exemple de la figure 2.5 (page 46) montre que l'égalité $\vec{w}(\mathcal{T}, I) = \vec{\pi}(\mathcal{T}, I)$ n'est pas vérifiée pour tout arbre \mathcal{T} et toute instance I . Nous pouvons cependant caractériser les arbres pour lesquels cette égalité est satisfaite pour toute instance.

DÉFINITION. Une **subdivision d'étoile** (*spider* en anglais) est un arbre ayant au plus un sommet de degré supérieur à 2.

Théorème 3.2.1 — *Soit \mathcal{T} un arbre orienté symétrique. Les deux assertions suivantes sont équivalentes :*

(a) \mathcal{T} est une subdivision d'étoile.

(b) Pour toute instance I , $\vec{w}(\mathcal{T}, I) = \vec{\pi}(\mathcal{T}, I)$.

PREUVE. On peut déduire de sa définition qu'une subdivision d'étoile est formée par l'union de chemins qui s'intersectent tous en un même sommet. Les chemins de longueurs ≥ 2 et les étoiles sont des subdivisions d'étoiles particulières. On a vu que la condition (b) est satisfaite pour les chemins. Montrons qu'elle est aussi vérifiée pour les étoiles.

Notons $V(S_N) = \{u_1, u_2, \dots, u_N\}$ l'ensemble des sommets de l'étoile S_N , de telle manière que tous les sommets sauf u_1 soient de degré 1. Soit I une instance de communication dans S_N . On peut supposer sans perte de généralité que le sommet u_1 ne fait partie d'aucune requête de I . Dans le cas contraire, les requêtes concernant u_1 se colorent en dernier facilement et en conservant la propriété désirée. Soit $G(I) = (V_1, V_2, E)$ le (multi)graphe biparti tel que $V_1 = \{a_i\}_{1 < i \leq N}$, $V_2 = \{b_j\}_{1 < j \leq N}$, et l'arête $[a_i, b_j] \in E(G)$ pour toute requête $(u_i, u_j) \in I$. De cette définition on en déduit que deux arêtes de $G(I)$ sont adjacentes si et seulement si les deux requêtes correspondantes dans I sont en conflit (voir figure 3.6, avec une notation différente). Par conséquent, trouver une coloration des requêtes de I revient à trouver une coloration des arêtes de $G(I)$, telle que deux arêtes adjacentes sont colorées différemment.

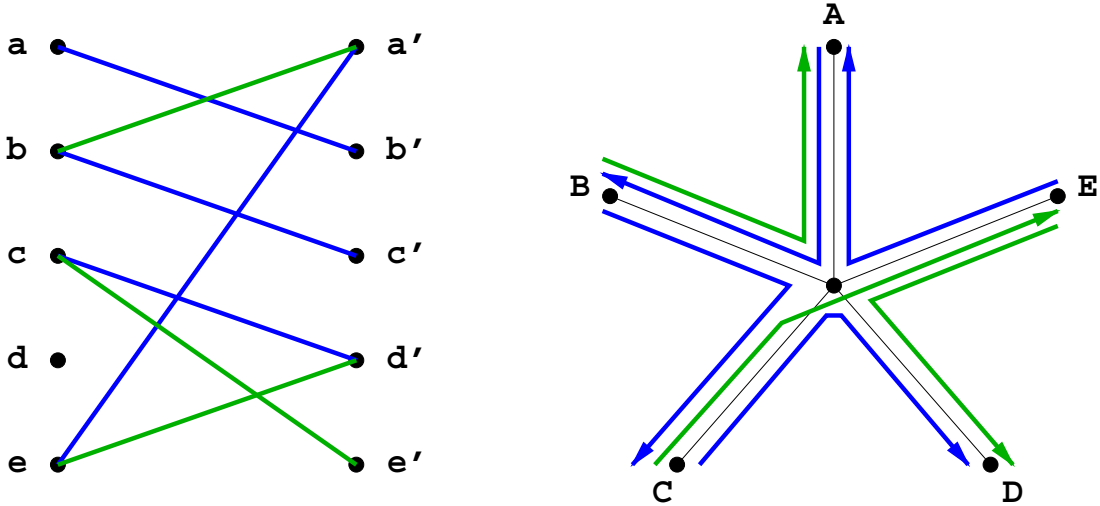


Figure 3.6: Couplage biparti et routage optique dans une étoile orientée.

Pour ce problème, connu également sous le nom de problème de *couplage dans un graphe biparti*, il existe le théorème de König-Hall (voir [Ber83] par exemple) qui nous dit que dans tout (multi)graphe biparti de degré maximum Δ , on peut

trouver une coloration des arêtes qui utilise Δ couleurs. En remarquant que le degré maximum de $G(I)$ correspond exactement à la charge maximale des arcs de l'étoile S_N pour l'instance I , on obtient que $\vec{w}(S_N, I) = \vec{\pi}(S_N, I)$.

Pour généraliser cette propriété aux subdivisions d'étoiles, la technique employée repose sur une combinaison des deux méthodes appliquées aux étoiles et aux chemins respectivement. Étant donnée une instance dans une subdivision d'étoile, on peut colorer les requêtes entre les différentes branches par la technique utilisée pour les étoiles, et les requêtes restantes à l'intérieur des branches par la technique utilisée pour les chemins, de manière à conserver l'égalité entre le nombre de couleurs utilisées et la charge maximale des arcs. L'implication (a) \Rightarrow (b) peut être ainsi démontrée.

Pour montrer la contraposée, il suffit de montrer que si la condition (a) n'est pas satisfaite, alors la condition (b) non plus. Soit \mathcal{T} un arbre possédant au moins deux sommets de degrés ≥ 3 , notés u et v . Soient u_1 et u_2 (resp. v_1 et v_2) deux sommets voisins de u (resp. de v) non situés sur le plus court chemin reliant u et v . Alors l'instance $I = \{(u_1, v), (u, v_1), (v_2, v_1), (v_2, u_2), (u_1, u_2)\}$ est telle que $\vec{\pi}(\mathcal{T}, I) = 2$ et $\vec{w}(\mathcal{T}, I) = 3$, de la même façon que sur la figure 2.5, page 46. D'où (b) \Rightarrow (a). \square

Dans sa généralité, le problème de routage optique dans les arbres orientés symétriques a été montré \mathcal{NP} -dur par Erlebach et Jansen [EJ96] et par Kumar et al. [KPRS97], même si la charge de l'instance vaut 3. Il le demeure pour les arbres binaires [EJ97, KPRS97] et pour les arbres de profondeur 2 [KPRS97], sans restriction sur la charge. Il devient cependant polynômial pour les arbres de taille bornée [KPRS97].

Différents travaux ont tenté d'approximer ce problème difficile. Nous commençons par présenter les résultats obtenus dans le modèle non-orienté.

Dans ce cas, la coloration des chemins correspond à la coloration des graphes EPT (pour *Edge intersection graphs of Paths in a Tree*, en anglais) étudiés par Golumbic et Jamison dans [GJ85a, GJ85b], qui ont montré notamment que leur reconnaissance et leur coloration sont \mathcal{NP} -difficiles. En fait, ils ont montré que le problème de coloration est déjà \mathcal{NP} -dur dans les étoiles, par réduction du problème de la coloration des arêtes d'un multigraphe, de la manière suivante.

Notons $V(S_N) = \{u_1, u_2, \dots, u_N\}$ l'ensemble des sommets de l'étoile S_N centrée en u_1 . Soit I une instance dans S_N et soit $G(I) = (V(S_N), E(I))$ le multigraphe tel que l'arête $[u_i, u_j] \in E(I)$ pour toute requête $(u_i, u_j) \in I$. De cette définition on en déduit que deux arêtes de $G(I)$ sont adjacentes si et seulement si les deux requêtes correspondantes dans I sont en conflit (voir figure 3.7, avec une notation différente). Par conséquent, trouver une coloration des requêtes

de I revient à trouver une coloration des arêtes de $G(I)$, telle que deux arêtes adjacentes sont colorées différemment.

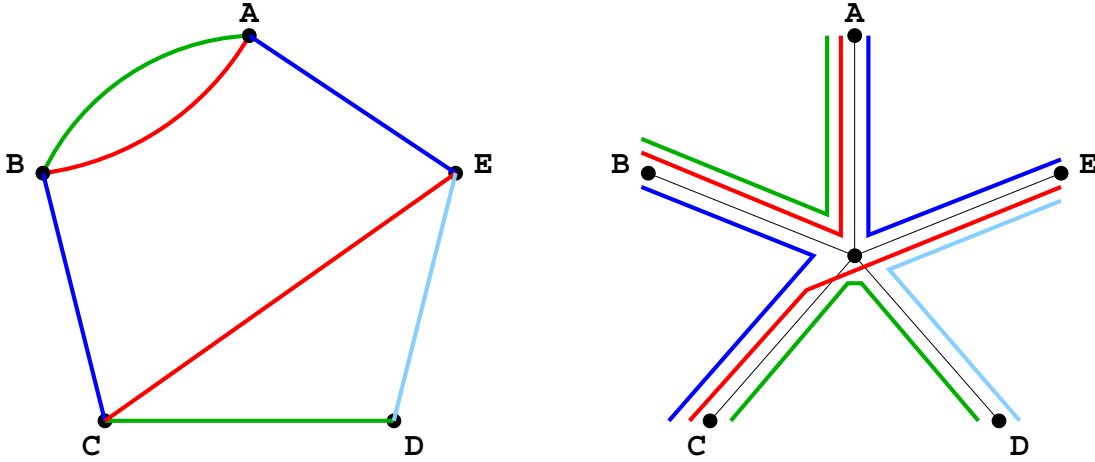


Figure 3.7: Arête-coloration d'un multigraphe et routage optique dans une étoile non-orientée.

Comme Tarjan précédemment dans [Tar85], Raghavan et Upfal [RU94] ont montré que $w(\mathcal{T}, I) \leq 3\pi(\mathcal{T}, I)/2$ pour toute instance dans tout arbre. Ce résultat s'appuie sur le théorème de Shannon [Sha49] qui permet de colorer les arêtes d'un multigraphe de degré maximum Δ à l'aide de $3\Delta/2$ couleurs (voir par exemple [FW77] pour une présentation de l'algorithme).

Il a été observé dans [MKR95] que ce facteur d'approximation $3/2$ peut être amélioré asymptotiquement en $9/8$ grâce à l'algorithme de coloration des arêtes de Goldberg [Gol84a, Gol84b]. De même, en utilisant un théorème d'approximation de Nishizeki et Kashiwagi [NK90], Erlebach et Jansen ont donné dans [EJ96] un algorithme ayant pour facteur d'approximation 1.1 asymptotiquement et $4/3$ dans l'absolu. D'après [Hol81], il est \mathcal{NP} -dur de décider si les arêtes d'un graphe de degré 3 sont colorables avec 3 ou 4 couleurs. Le facteur $4/3$ d'approximation absolue est donc le meilleur possible.

Théorème 3.2.2 (Erlebach et Jansen [EJ96]) — *Pour tout arbre non-orienté \mathcal{T} et pour toute instance I , le problème de routage optique (\mathcal{T}, I) peut être résolu en temps polynômial en utilisant $w(\mathcal{T}, I)$ couleurs si $w(\mathcal{T}, I) \leq 2$ et au plus $\lceil 1.1w(\mathcal{T}, I) + 0.8 \rceil$ couleurs sinon.*

REMARQUE. Dans le modèle non-orienté, il est difficile de colorer un ensemble de chemins qui partagent un sommet, puisque cela revient à colorer les arêtes d'un

multigraphe. En revanche, il est aisé de combiner de telles colorations locales pour obtenir une coloration globale de tous les chemins considérés, sans augmenter le nombre de couleurs utilisées. Inversement, dans le modèle orienté, il est facile de colorer des chemins qui partagent un sommet, puisque cela revient à colorer les arêtes d'un graphe biparti, mais il est difficile de combiner plusieurs colorations locales optimales pour former une coloration globale optimale.

Mihail, Kaklamanis et Rao [MKR95] ont été les premiers à considérer le cas orienté pour les réseaux optiques. Contrairement aux graphes EPT, les graphes de conflit de chemins dans les arbres orientés symétriques n'avaient pas retenu l'attention auparavant. Leur résultat sur les arbres est une borne de $15\vec{\pi}(\mathcal{T}, I)/8$ pour $\vec{w}(\mathcal{T}, I)$, obtenue en réduisant le problème d'allocation des longueurs d'onde à un cas particulier de coloration des arêtes d'un graphe biparti, comme il a été montré précédemment pour les étoiles. En utilisant la même technique générale, le facteur $15/8$ a pu être amélioré en $7/4$ par Kaklamanis et Persiano dans [KP96] (et indépendamment par Kumar et Schwabe dans [KS97]), puis en $5/3$ par Erlebach et al. dans [EJKP97, EJK⁺99].

Théorème 3.2.3 (Erlebach et al. [EJKP97, EJK⁺99]) — *Pour tout arbre orienté symétrique \mathcal{T} et pour toute instance I , le problème de routage optique (\mathcal{T}, I) peut être résolu en temps polynômial en utilisant $\lceil 5\vec{\pi}(\mathcal{T}, I)/3 \rceil$ longueurs d'onde.*

L'algorithme permettant d'atteindre ce résultat présente en outre l'avantage d'être glouton. Il procède par phases en parcourant les sommets de l'arbre en profondeur et en colorant pour chacun d'eux tous les chemins qui le contiennent et qui n'ont pas été préalablement colorés. Cette procédure est réduite à un problème contraint de coloration des arêtes d'un graphe biparti, d'où est issu le résultat. Un tel algorithme ne nécessite pas de contrôle global, dans le sens où chaque sommet décide localement, à son tour, quelles couleurs il attribue aux requêtes qui le concernent et qui ne sont pas déjà colorées.

Notons que des algorithmes plus simples ont été fournis pour les arbres binaires dans [CKP97] et [Jan97], permettant de colorer toute instance I à l'aide d'au plus $\lceil 5\vec{\pi}(\mathcal{T}, I)/3 \rceil$ couleurs, ce qui présente une très légère amélioration par rapport au cas général. Il est donné de plus dans [Jan97] une construction non triviale dans un arbre binaire, pour laquelle $\vec{\pi} = 3$ et $\vec{w} = 5$.

L'importance du théorème 3.2.3 est soulignée par la propriété, montrée également dans [CKP97, Jan97, EJKP97, EJK⁺99], qu'aucun algorithme glouton de ce type ne peut procurer de meilleure approximation par rapport à la charge :

Proposition 3.2.4 (Erlebach et al. [EJKP97, EJK⁺99]) — *Pour tout $\pi > 1$, tout $\epsilon > 0$ et tout algorithme glouton \mathcal{A} , il existe un arbre orienté symétrique \mathcal{T} (binaire) et une instance I , tels que $\vec{\pi}(\mathcal{T}, I) = \pi$ et \mathcal{A} résoud le problème (\mathcal{T}, I) en utilisant au moins $(5/3 - \epsilon)\pi$ longueurs d'onde.*

Par ailleurs, comme dans le modèle non-orienté, en se basant sur la \mathcal{NP} -difficulté de la coloration des arêtes d'un multigraphe de degré 3 [Hol81] et par une réduction de ce problème, Erlebach a montré dans sa thèse [Erl98] qu'on ne peut espérer obtenir un algorithme polynômial d'approximation inférieure à $4/3$:

Théorème 3.2.5 (Erlebach [Erl98]) — *Décider si $\vec{w}(\mathcal{T}, I) = 3$ pour une instance I dans un arbre \mathcal{T} orienté symétrique est \mathcal{NP} -complet. Par conséquent, il n'existe pas d'algorithme polynômial pour le problème de routage optique ayant un facteur d'approximation inférieur à $4/3$, à moins que $\mathcal{P} = \mathcal{NP}$.*

Toujours pour apprécier le facteur d'approximation $5/3$ du théorème 3.2.3, on peut exhiber des problèmes de routage optique dans des arbres ayant une charge arbitrairement grande et tels que le rapport $\vec{w}/\vec{\pi}$ vaut $5/4$. La construction de tels exemples est basée sur l'exemple (\mathcal{T}, I) de la figure 2.5, page 46. En notant I^n l'instance formée par n copies de I , pour tout $n \geq 2$, le problème (\mathcal{T}, I^n) revient à multi-colorer les sommets d'un pentagone. D'après un résultat montré dans [HR82], nous en déduisons la proposition suivante, donnée dans [BBG+97] et également par Kumar et Schwabe dans [KS97] :

Proposition 3.2.6 — *Pour tout entier π il existe un problème (\mathcal{T}, I) dans un arbre orienté symétrique \mathcal{T} , tel que*

$$\vec{\pi}(\mathcal{T}, I) \geq \pi \text{ et } \vec{w}(\mathcal{T}, I) \geq 5\vec{\pi}(\mathcal{T}, I)/4.$$

Conjecture 3.2.7 — La constante $5/4$ de la proposition 3.2.6 est fine.

Le problème de l'échange total dans les arbres orientés symétriques a été résolu par Gargano, Hell et Pérennes [GHP97], par un algorithme polynômial permettant de colorer toutes les requêtes de l'instance I_A à l'aide de $\vec{\pi}(\mathcal{T}, I_A)$ couleurs. En fait, le résultat a été démontré pour les arbres *pondérés*, pour lesquels les sommets ont des poids entiers et l'instance I_A comporte pour chaque couple de sommets autant de copies de la requête correspondante que le produit de leurs poids respectifs.

La démonstration est inductive et repose sur le fait que tout arbre peut être obtenu à partir d'une étoile pondérée, de même poids total et de même charge maximale pour l'échange total, à l'aide de deux opérations sur les feuilles qui sont : 1) l'ajout d'une nouvelle feuille sur une ancienne, et 2) la division d'une feuille en deux feuilles. Ces deux opérations sur les feuilles d'un arbre pondéré maintiennent constant le poids total des sommets en jeu ainsi que la charge maximale des arcs pour l'échange total. Pour chacune d'elles, une modification adéquate de la coloration des requêtes est fournie, notamment à l'aide du théorème de König-Hall pour l'opération de division.

Théorème 3.2.8 (Gargano, Hell et Pérennes [GHP97]) — *Pour tout arbre orienté symétrique \mathcal{T} et pour l'instance d'échange total I_A , le problème de routage optique (\mathcal{T}, I_A) peut être résolu de façon optimale en temps polynômial et*

$$\vec{w}(\mathcal{T}, I_A) = \vec{\pi}(\mathcal{T}, I_A) .$$

Dans le modèle non-orienté, ce dernier résultat n'est pas vérifié en ce qui concerne l'égalité entre \vec{w} et $\vec{\pi}$. En effet, pour la subdivision d'étoile S ayant trois branches de longueur 2 chacune, et donc 7 sommets au total, nous avons une charge $\pi(S, I_A) = 2.(2.5) = 20$, atteinte sur les arêtes intérieures, or au moins 24 couleurs sont nécessaires pour colorer les 3 groupes de 8 requêtes entre les branches distinctes, qui entrent en conflit deux à deux.

Cependant, il n'a pas été démontré à notre connaissance que ce problème ne pouvait pas être résolu efficacement, aussi nous formulons la question.

Question 3.2.9 — Pour tout arbre non-orienté \mathcal{T} et pour l'instance d'échange total I_A , le problème de routage optique (\mathcal{T}, I_A) peut-il être résolu de façon optimale en temps polynômial ?

3.2.2 Cycles

Dans notre contexte, sauf précisé explicitement, un `cycle` est un graphe orienté symétrique, qui peut se définir comme étant fortement connexe et 2-régulier. Contrairement au cas des arbres où le problème de routage ne se pose pas, il intervient dans les cycles sous la forme de choix binaires : pour router chaque requête par un chemin, il faut décider quelle orientation du cycle utiliser.

Étant donné un routage pour réaliser une instance de communication, le problème de la coloration de ses chemins est alors équivalent au problème bien connu de la coloration des sommets des graphes *arc-circulaires*, montré \mathcal{NP} -complet dans [GJMP80]. Au vu de la liberté supplémentaire accordée par le choix du routage, il n'est pas immédiat que le problème de routage optique dans les cycles soit également \mathcal{NP} -difficile. Erlebach et Jansen l'ont cependant démontré dans [EJ96], aussi bien dans le modèle orienté que dans le modèle non-orienté, par une réduction du problème de la coloration des graphes arc-circulaires. Une preuve plus simple de \mathcal{NP} -complexité pour le cas orienté symétrique pourra toutefois être trouvée dans la version journal de [WW98].

Un algorithme d'approximation de facteur 2 pour le routage optique peut être obtenu facilement dans les cycles, comme il a été remarqué précédemment dans [RU94] pour le cas non-orienté et dans [MKR95] pour le cas orienté. Il suffit de router toutes les requêtes par des chemins qui n'utilisent pas un lien

particulier du cycle (une paire d’arcs symétriques ou une arête). De cette façon, la charge obtenue sur les liens du cycle est au plus double de la charge optimale pour l’instance considérée. Le graphe de conflit des chemins résultants est alors un graphe d’intervalles, qui se colore en temps polynômial avec autant de couleurs que la charge du routage. La charge optimale minorant le nombre de couleurs optimal, le facteur d’approximation est bien de 2.

À ce jour, il n’est malheureusement pas connu d’algorithme déterministe qui possède un meilleur facteur d’approximation, ce qui peut paraître surprenant. Cependant, Kumar a récemment présenté dans [Kum98] un algorithme randomisé d’approximation asymptotique $1 + 1/e \simeq 1.37$, avec forte probabilité, sous l’hypothèse $\log |I| = o(\bar{\pi}(G, I))$. Un résultat complémentaire est fourni dans [Che99].

En ce qui concerne le problème de routage, Wilfong et Winkler dans [WW98] ont récemment donné un algorithme polynômial pour router toute instance I dans un cycle orienté symétrique C avec une charge maximale $\bar{\pi}(C, I)$. Leur méthode consiste à utiliser une relaxation du programme linéaire entier correspondant et une technique d’arrondi qui préserve l’optimalité entière de la solution.

Théorème 3.2.10 (Wilfong et Winkler [WW98]) — *Pour tout cycle C orienté symétrique et toute instance I , on peut trouver en temps polynômial un routage R réalisant I tel que $\bar{\pi}(C, I, R) = \bar{\pi}(C, I)$.*

Le résultat équivalent dans le modèle non-orienté avait déjà été démontré par Frank et al. dans [Fra85, FNS⁺92], à partir d’un résultat de Okamura et Seymour [OS81].

Théorème 3.2.11 (Frank et al. [FNS⁺92]) — *Pour tout cycle non-orienté C et pour toute instance I , on peut trouver en temps linéaire un routage R pour I tel que $\pi(C, I, R) = \pi(C, I)$.*

Le facteur d’approximation déterministe de 2 mentionné ci-dessus résulte d’une coloration à l’aide d’un nombre de couleurs au plus double de la charge (on peut même économiser une couleur facilement, voir [Tuc75]). Il s’avère qu’on ne peut mieux faire par rapport à la charge, comme le montre l’exemple donné à la page 66, pour lequel $\bar{w}(C, I) = 2\bar{\pi}(C, I) - 1$. Aussi les deux théorèmes précédents ne permettent pas d’envisager une meilleure approximation dans le cas général.

En revanche, étant donné un routage dans un cycle, on peut approximer la coloration de ses chemins par un facteur $3/2$. En effet, un résultat très peu connu de Karapetyan [Kar80] permet de colorer les sommets d’un graphe arc-circulaire avec $\lceil 3\varphi/2 \rceil$ couleurs, où φ représente la taille maximale d’une clique.

Le meilleur algorithme précédemment connu (bien qu'antérieur!) était celui de Shih et Hsu [SH90] offrant un facteur d'approximation égal à $5/3$.

Théorème 3.2.12 (Karapetyan [Kar80]) — *On peut colorer en temps polynômial toute famille F d'arcs sur un cercle en utilisant $\lfloor \frac{3}{2}\varphi(F) \rfloor$ couleurs.*

Il serait donc souhaitable d'obtenir un algorithme de routage qui minimise la taille de la clique maximale engendrée dans le graphe de conflit des chemins, afin de tirer profit de ce résultat.

Enfin, notons que dans le cas où un routage donné est *propre*, c'est-à-dire si aucun chemin n'est contenu dans un autre, Orlin, Bonucelli et Bovet ont donné un algorithme optimal pour sa coloration [OBB81].

3.2.3 Grilles

Kramer et van Leeuwen ont donné dans [KvL84] une preuve de \mathcal{NP} -complétude pour un problème de routage intervenant dans le contexte de la théorie VLSI. Elle permet d'en déduire à l'aide de légères modifications que le même résultat pour le problème du routage disjoint maximum dans les grilles bidimensionnelles orientées symétriques ou non-orientées. Par conséquent, sont également \mathcal{NP} -difficiles les problèmes du routage optique et du routage dans ces topologies, et il n'existe aucun algorithme d'approximation ayant un facteur inférieur à 2, à moins que $\mathcal{P} = \mathcal{NP}$.

Le modèle non-orienté a été étudié essentiellement. Pour toute grille \mathcal{G} de dimension bornée et d'ordre N , et pour toute instance I , Aumann et Rabani ont donné dans [AR95] un algorithme d'approximation pour résoudre le problème du routage optique (\mathcal{G}, I) à l'aide de $O(\log N \log |I| \mathfrak{w}(\mathcal{G}, I))$ couleurs. La technique utilisée repose sur l'utilisation d'une routine pour le problème du routage disjoint maximum, ayant un facteur d'approximation en $O(\log N)$. Kleinberg et Tardos ont amélioré cette routine dans [KT95] pour une classe de graphes incluant les grilles bidimensionnelles, en obtenant une approximation constante, ce qui mène à une approximation en $O(\log N)$ pour le routage optique. La dernière amélioration, toujours pour une dimension de 2, a été apportée par Rabani :

Théorème 3.2.13 (Rabani [Rab96]) — *Pour toute grille carrée non-orientée P_n^d d'ordre $N = n^d$ et pour toute instance I , le problème de routage optique (P_n^d, I) peut être résolu en temps polynômial et en utilisant $\text{poly}(\log \log N) \cdot \mathfrak{w}(P_n^d, I)$ longueurs d'onde, où poly désigne un polynôme. De plus, $\mathfrak{w}(P_n^d, I)$ peut être calculé à un facteur constant près.*

Bibliographie

- [ABNC⁺94] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing and scheduling algorithms for optical networks. In *Proc. of the 5th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'94)*, pages 412–423. ACM Press, January 1994.
- [ABNC⁺96] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing in optical networks. *Journal of the ACM*, 46(6):973–1001, November 1996.
- [AR95] Y. Aumann and Y. Rabani. Improved bounds for all optical routing. In *Proc. of the 6th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'95)*, pages 567–576. ACM Press, January 1995.
- [BBG⁺97] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes, and U. Vaccaro. Graph problems arising from wavelength-routing in all-optical networks. In *Proc. of IPPS'97, 2nd Workshop on Optics and Computer Science (WOCS'97)*, Genève, April 1997. IEEE Press.
- [Bea99] B. Beauquier. All-to-all communication in some wavelength-routed all-optical networks. *Networks*, 33(3):179–187, May 1999.
- [Ber83] C. Berge. *Graphes*. Gauthier-Villars, Paris, 1983.
- [BGP⁺96] J-C. Bermond, L. Gargano, S. Pérennes, A. A. Rescigno, and U. Vaccaro. Efficient collective communication in optical networks. In *Proc. of 23rd International Colloquium on Automata, Languages and Programming (ICALP'96)*, volume 1099 of *Lecture Notes in Computer Science*, pages 574–585. Springer-Verlag, July 1996.
- [BHP98] B. Beauquier, P. Hell, and S. Pérennes. Optimal wavelength-routed multicasting. *Discrete Applied Mathematics*, 84:15–20, 1998.
- [BPT99] B. Beauquier, S. Pérennes, and D. Tóth. All-to-all routing and coloring in weighted trees of rings. In *Proc. of 11th ACM Symposium on Parallel Algorithms and Architectures (SPAA'99)*, pages 185–190, Saint-Malo, France, June 1999. ACM Press.

- [CCRS87] F. R. K. Chung, E. G. Coffman, M. I. Reiman, and B. E. Simon. The forwarding index of communication networks. *IEEE Transactions on Information Theory*, 33:224–232, 1987.
- [CGK92] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: An approach to high bandwidth optical WAN's. *IEEE Transactions on Communications*, 40(7):1171–1182, July 1992.
- [Che99] C. T. Cheng. A new approximation algorithm for the demand routing and slotting problem with unit demands on rings. In *Proc. of the International Workshop on Approximation Algorithms for Combinatorial Optimization (APPROX'99)*, volume 1671 of *Lecture Notes in Computer Science*, pages 209–220. Springer-Verlag, 1999.
- [CKP97] I. Caragiannis, C. Kaklamanis, and P. Persiano. Bounds on optical bandwidth allocation on directed fiber tree topologies. In *Proc. of IPPS'97, 2nd Workshop on Optics and Computer Science (WOCS'97)*, Geneva, Switzerland, April 1997. IEEE Press.
- [Edm72] J. Edmonds. Edge-disjoint branchings. In R. Ruskin, editor, *Combinatorial Algorithms*, pages 91–96. Algorithmic Press, NY, 1972.
- [EIS76] S. Even, A. Itai, and A. Shamir. On the complexity of timetable and multicommodity flow problems. *SIAM Journal of Computing*, 5(4):691–703, December 1976.
- [EJ96] T. Erlebach and K Jansen. Scheduling of virtual connections in fast networks. In *Proc. of 4th Workshop on Parallel Systems and Algorithms (PASA'96)*, pages 13–32. World Scientific, 1996.
- [EJ97] T. Erlebach and K Jansen. Call scheduling in trees, rings and meshes. In *Proc. of 30th Hawaii International Conference on System Sciences (HICSS'97)*, volume 1, pages 221–222. IEEE Computer Society Press, 1997.
- [EJK⁺99] T. Erlebach, K. Jansen, C. Kaklamanis, M. Mihail, and P. Persiano. Optimal wavelength routing on directed fiber trees. *Theoretical Computer Science*, 221(1–2):119–137, 1999.
- [EJKP97] T. Erlebach, K. Jansen, C. Kaklamanis, and P. Persiano. Constrained bipartite edge coloring with applications to wavelength routing. In *Proc. of 24th Internat. Colloq. on Automata, Languages and Programming (ICALP'97)*, volume 1256 of *Lecture Notes in Computer Science*, pages 493–504. Springer-Verlag, 1997.

- [Erl98] Thomas Erlebach. *Scheduling Connections in Fast Networks*. PhD thesis, Technischen Universität München, Germany, October 1998.
- [FHW80] S. Fortune, J. Hopcroft, and J. Wyllie. The directed subgraph homeomorphism problem. *Theoret. Computer Science*, 10:111–121, 1980.
- [FNS⁺92] A. Frank, T. Nishizeki, N. Saito, H. Suzuki, and E. Tardos. Algorithms for routing around a rectangle. *Discrete Applied Mathematics*, 40:363–378, 1992.
- [Fra85] A. Frank. Edge-disjoint paths in planar graphs. *Journal of Combinatorial Theory, Series B*, 38:164–178, 1985.
- [Fra90] A. Frank. Packing paths, circuits and cuts – A survey. In B. Korte, L. Lovász, H. J. Prömel, and A. Schrijver, editors, *Paths, Flows and VLSI-Layout*, pages 47–100. Springer-Verlag, 1990.
- [FW77] S. Fiorini and R. J. Wilson. *Edge-Coloring of Graphs*, volume 16 of *Research Notes in Mathematics*. Pitman, London, 1977.
- [GHP97] L. Gargano, P. Hell, and S. Pérennes. Colouring paths in directed symmetric trees with applications to WDM routing. In *Proc. of 24th Int. Colloq. on Automata, Languages and Programming (ICALP'97)*, volume 1256 of *Lecture Notes in Computer Science*, pages 505–515. Springer-Verlag, July 1997.
- [GJ85a] M. C. Golumbic and R. E. Jamison. Edge and vertex intersection of paths in a tree. *Discrete Mathematics*, 55:151–159, 1985.
- [GJ85b] M. C. Golumbic and R. E. Jamison. The edge intersection graphs of paths in a tree. *Journal of Comb. Theory, Series B*, 38:8–22, 1985.
- [GJMP80] M. R. Garey, D. S. Johnson, G. L. Miller, and C. H. Papadimitriou. The complexity of coloring circular arcs and chords. *SIAM Journal on Algebraic and Discrete Methods*, 1(2):216–227, June 1980.
- [Gol84a] M. Goldberg. An approximate algorithm for the edge-coloring problem. *Congressus Numerantium*, 43:317–319, 1984.
- [Gol84b] M. Goldberg. Edge-coloring of multigraphs: Recoloring technique. *Journal of Graph Theory*, 8:123–137, 1984.
- [HMOS94] M-C. Heydemann, J-C. Meyer, J. Opatrny, and D. Sotteau. Forwarding indices of consistent routings and their complexity. *Networks*, 24:75–82, 1994.

- [HMS89] M-C. Heydemann, J-C. Meyer, and D. Sotteau. On forwarding indices of networks. *Discrete Applied Mathematics*, 23:103–123, 1989.
- [Hol81] I. Holyer. The NP-completeness of edge coloring. *SIAM Journal of Computing*, 10(4):718–720, 1981.
- [HR82] P. Hell and F. S. Roberts. Analogues of Shannon capacity. *Annals of Discrete Mathematics*, pages 155–168, 1982.
- [Jan97] K. Jansen. Approximation results for wavelength routing in directed trees. In *Proc. of IPPS'97, 2nd Workshop on Optics and Computer Science (WOCS'97)*, Geneva, Switzerland, April 1997. IEEE Press.
- [Jar99] A. Jarry. Théorie des graphes: chemins disjoints. Rapport de stage de 1ère année, École Normale Supérieure de Lyon, July 1999.
- [Kar80] I. A. Karapetyan. On coloring of arc graphs. *Doklady Akad. Nauk Armianskoi CCP*, 70(5):306–311, 1980. In Russian.
- [KLPS90] B. Korte, L. Lovász, H. J. Prömel, and A. Schrijver, editors. *Paths, Flows and VLSI-Layout*. Springer-Verlag, Berlin, 1990.
- [KP96] C. Kaklamanis and P. Persiano. Efficient wavelength routing on directed fiber trees. In *Proc. of European Symposium on Algorithms (ESA '96)*, volume 1136 of *Lecture Notes in Computer Science*, pages 460–470, 1996.
- [KPRS97] S. R. Kumar, R. Panigrahy, A. Russel, and R. Sundaram. A note on optical routing on trees. *Information Proc. Letters*, 62:295–300, 1997.
- [KS97] V. Kumar and E. J. Schwabe. Improved access to optical bandwidth in trees. In *Proc. of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '97)*, pages 437–444. ACM Press, January 5–7 1997.
- [KT95] Jon Kleinberg and Éva Tardos. Disjoint paths in densely embedded graphs. In *Proc. of 36th Annual Symposium on Foundations of Computer Science (FOCS'95)*, pages 52–61. IEEE, October 1995.
- [Kum98] V. Kumar. Approximating circular arc coloring and bandwidth allocation in all-optical ring networks. In *Proc. of the International Workshop on Approximation Algorithms for Combinatorial Optimization (APPROX'98)*, volume 1444 of *Lecture Notes in Computer Science*, pages 147–158. Springer-Verlag, 1998.

- [KvL84] M. E. Kramer and J. van Leeuwen. The complexity of wire routing and finding the minimum area layouts for arbitrary VLSI circuits. In F. P. Preparata, editor, *Advances in Computing Research; VLSI Theory*, volume 2, pages 129–146. JAI Press Inc., 1984.
- [LR88] F. T. Leighton and S. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms'. In *Proc. of 29th Annual Symposium on Foundations of Computer Science (FOCS'88)*, pages 422–431. IEEE, October 1988.
- [Mad71] W. Mader. Minimale n -fach kantenzusammenhängende graphen. *Math. Ann.*, 191:21–28, 1971.
- [MKR95] M. Mihail, C. Kaklamanis, and S. Rao. Efficient access to optical bandwidth — wavelength routing on directed fiber trees, rings, and trees of rings. In *Proc. of 36th IEEE Annual Symposium on Foundations of Computer Science (FOCS'95)*, pages 548–557. IEEE Press, October 1995.
- [Moh89] B. Mohar. Isoperimetric number of graphs. *Journal of Combinatorial Theory, Series B*, 47:274–291, 1989.
- [NK90] T. Nishizeki and K. Kashiwagi. On the 1.1 edge-coloring of multigraphs. *SIAM Journal on Discrete Mathematics*, 3(3):391–410, August 1990.
- [OBB81] J. B. Orlin, M. A. Bonuccelli, and D. P. Bovet. An $O(n^2)$ algorithm for coloring proper circular arc graphs. *SIAM Journal of Algorithmic Discrete Methods*, 2(2):88–93, June 1981.
- [OS81] H. Okamura and P. D. Seymour. Multicommodity flows in planar graphs. *Journal of Combinatorial Theory, Series B*, 31:75–81, 1981.
- [Pan92] R. K. Pankaj. *Architectures for Linear Lightwave Networks*. PhD thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1992.
- [PG95] R. K. Pankaj and R. G. Gallager. Wavelength requirements of all-optical networks. *IEEE/ACM Transactions on Networking*, 3:269–280, 1995.
- [Rab96] Y. Rabani. Path coloring on the mesh. In *Proc. of 37th IEEE Annual Symposium on Foundations of Computer Science (FOCS'96)*, October 1996.

- [RS95] N. Robertson and P. D. Seymour. Graph minors XIII. The disjoint paths problem. *Journal of Combinatorial Theory, Series B*, 63, 1995.
- [RT87] P. Raghavan and C. D. Thompson. Randomized routing: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.
- [RU94] P. Raghavan and E. Upfal. Efficient routing in all-optical networks. In *Proc. of the 26th Annual ACM Symposium on Theory of Computing (STOC'94)*, pages 134–143. ACM Press, May 1994.
- [SH90] W. K. Shih and W. L. Hsu. An approximation algorithm for coloring circular-arc graphs. In *Proc. of SIAM Conference on Discrete Mathematics*, San Francisco, CA, 1990.
- [Sha49] C. E. Shannon. A theorem on coloring the lines of a network. *J. Math. Phys.*, 28:148–151, 1949.
- [Shi79] Y. Shiloach. Edge-disjoint branching in directed multigraphs. *Information Processing Letters*, 8(1):24–27, 1979.
- [Sol95] P. Solé. Expanding and forwarding. *Discrete Applied Mathematics*, 58:67–78, 1995.
- [Tar85] R. E. Tarjan. Decomposition by clique separators. *Discrete Mathematics*, 55(2):221–232, 1985.
- [Tog98a] O. Togni. *Force des graphes – Indice optique des réseaux*. PhD thesis, Université de Bordeaux 1, 1998.
- [Tog98b] O. Togni. Optical all-to-all communication in inflated networks. In *Proc. of the 5th Annual International Workshop on Graph-Theoretic Concepts in Computer Science (WG'98)*, volume 1517 of *Lecture Notes in Computer Science*, pages 78–87. Springer-Verlag, 1998.
- [Tuc75] A. Tucker. Coloring a family of circular arcs. *SIAM Journal of Applied Mathematics*, 29(3):493–502, 1975.
- [Wil96] G. Wilfong. Minimizing wavelengths in an all-optical ring network. In *Proc. of the 7th International Symposium on Algorithms and Computation (ISAAC'96)*, volume 1178 of *Lecture Notes in Computer Science*, pages 346–355. Springer-Verlag, 1996.
- [WW98] G. Wilfong and P. Winkler. Ring routing and wavelength translation. In *Proc. of the 9th Ann. ACM-SIAM Symposium on Discrete Algorithms (SODA'98)*, pages 333–341. ACM Press, January 1998.

Conclusion et perspectives

Nous avons étudié dans cette thèse des problèmes d'optimisation modélisant l'allocation des ressources pour la satisfaction de requêtes de connexion dans les réseaux de communication optiques. Dans le premier chapitre, nous avons présenté la technologie optique pour les télécommunications, afin de préciser le cadre technique de notre recherche et d'aider le lecteur informaticien à la compréhension des contraintes physiques sous-jacentes à la modélisation théorique. Dans le deuxième chapitre, nous avons posé la problématique étudiée au cours de la thèse et décrit formellement le modèle et les problèmes qui ont servi de base à nos recherches. Nous avons ensuite fourni dans le troisième chapitre une synthèse des résultats obtenus dans la littérature concernant le problème du routage optique et ses relations avec le problème de la charge. Les résultats que nous avons obtenus sont détaillés dans les annexes qui suivent, sous la forme d'articles en anglais et dans le format des rapports de recherche.

Un grand nombre de perspectives de recherches semblent prometteuses. Nous avons montré au début du troisième chapitre pourquoi le problème de routage optique n'est pas approximable dans sa généralité dans le cas des graphes orientés. Cependant, aucun résultat de ce type n'est connu dans le cas orienté symétrique. De plus, la question reste ouverte de savoir dans quelle mesure le rapport $\bar{w}/\bar{\pi}$ (entre le nombre de longueurs d'onde et la charge nécessaires) peut être élevé, ne serait-ce que supérieur à 2.

De manière générale, il serait intéressant de chercher à améliorer les facteurs d'approximation des algorithmes fournis dans la littérature. En particulier, le facteur de $5/3$, valable pour toute instance du problème de routage optique dans les arbres orientés symétriques, est issu d'un algorithme local et glouton. La question de savoir si un algorithme centralisé peut atteindre une meilleure approximation mérite d'être abordée.

Plusieurs de nos résultats présentés dans les annexes confortent notre sentiment que l'égalité $\bar{w}(G, I_A) = \bar{\pi}(G, I_A)$ est vérifiée dans tout graphe orienté symétrique G pour l'instance I_A de l'échange total. À moins de trouver un contre-exemple qui contredirait notre conjecture, il serait souhaitable de parvenir à montrer cette propriété dans le cas général.

Notre modélisation théorique tend à simplifier la structure réelle des réseaux de communication optiques et les objectifs d'optimisation, tout en conservant les caractéristiques essentielles des problèmes rencontrés en pratique. Des généralisations vers des modèles plus larges seraient souhaitables. En particulier, nous pourrions considérer des réseaux tout-optiques possédant des capacités limitées de conversion et des ensembles de longueurs d'onde disponibles différents selon les liens. De plus, nous pourrions chercher à optimiser des fonctions de coût sur les routages optiques qui prennent en compte la complexité des nœuds de commutation nécessaires pour les réaliser.

Ainsi, dans le cadre du projet PORTO du RNRT, nouvellement contracté avec Alcatel (Corporate Research Center) et France Telecom (CNET), le domaine de la conception et du dimensionnement des réseaux optiques de télécommunications, qui vise à minimiser le coût global des équipements mis en place pour assurer une certaine matrice de trafic, nous apparaît comme le plus propice à valider pratiquement les compétences que nous avons acquises par l'étude des communications tout-optiques multiplexées en longueur d'onde. De plus, cette étude comporte de nouveaux aspects pratiques intéressants à formaliser, comme ceux issus des différentes politiques de sécurisation (protections, restaurations) des réseaux de transport, destinées à garantir l'acheminement du trafic malgré d'éventuelles pannes de nœuds ou de liens.

Nous avons considéré dans cette thèse des problèmes relativement réguliers, sous une approche statique et centralisée. Il serait intéressant de prolonger ces travaux dans une perspective dynamique et/ou distribuée. Nous souhaiterions de même faire évoluer nos recherches vers l'analyse de topologies et d'instances de communication irrégulières. Ceci implique de se familiariser avec de nouvelles méthodes de résolution, telles que la programmation linéaire en nombres entiers ou l'algorithmique randomisée, afin de les associer avec les techniques plus classiques que nous avons utilisées jusqu'ici.

ANNEXES

ANNEXE A

Optimal wavelength-routed multicasting

Optimal Wavelength-Routed Multicasting

Bruno Beauquier^{*}, Pavol Hell^{**}, Stéphane Pérennes^{*}

Thème 1 — Réseaux et systèmes
Projet SLOOP

Octobre 1997

Abstract: Motivated by wavelength division multiplexing in all-optical networks, we consider the problem of finding a set of paths from a fixed source to a multiset of destinations, which can be coloured by the fewest number of colours so that paths of the same colour do not share an arc. We prove that this minimum number of colours is equal to the maximum number of paths that share one arc, minimized over all sets of paths from the source to the destinations. We do this by modeling the problems as network flows in two different networks and relating the structure of their minimum cuts. The problem can be efficiently solved (paths found and coloured) using network flow techniques.

Key-words: All-optical networks, WDM routing, multicast, graphs, flows.

(Résumé : *tsvp*)

Submitted to *Discrete Applied Mathematics*.

^{*} SLOOP is a joint project with the CNRS and the University of Nice-Sophia Antipolis (I3S laboratory). This work has been partially supported by the French GDR/PRC Project PRS, by the Galileo Project. Emails : {beauquier, speren}@sophia.inria.fr

^{**} Simon Fraser University, School of Computing Science, Burnaby, B.C., V5A 1S6, Canada.
Email : pavol@cs.sfu.ca

Algorithme optimal de multicast tout-optique

Résumé : Ce rapport est motivé par l'étude du routage par multiplexage en longueur d'onde (en anglais *Wavelength Division Multiplexing : WDM*) dans les réseaux tout-optiques. Nous considérons le problème du *multicast* dans lequel un processeur fixé désire communiquer simultanément avec un certain nombre d'autres processeurs. Il s'agit de trouver dans le graphe associé des chemins reliant la source du multicast aux destinations et de colorer ces chemins, de manière à ce qu'aucun lien du réseau ne soit traversé par deux chemins de la même couleur. Nous résolvons ce problème en utilisant des techniques de flots. Il en découle un algorithme polynomial pour calculer une solution optimale.

Mots-clé : Réseaux tout-optiques, multicast, graphes, routage, flots.

1 Motivation and Definitions

Optics is emerging as a key technology in communication networks, promising very high speed local or wide area networks of the future. A single optical wavelength supports rates of gigabits per second, which in turn support multiple channels of voice, data and video [4, 5]. Multiple laser beams that are propagated over the same fiber on distinct optical wavelengths can increase this capacity even further. This is achieved through *Wavelength Division Multiplexing* (or *WDM*) [2], by partitioning the optical bandwidth into several channels and allowing the transmission of multiple data streams concurrently along the same optical fiber.

All-optical (or *single-hop* [6]) communication networks provide all source-destination pairs with end-to-end transparent channels that are identified through a wavelength and a physical path. Wavelengths being a limited resource, solutions to the problem of efficient routing and wavelengths allocation are of importance for the future development of optical technology.

The problem we consider here is motivated by *switched* networks with reconfigurable wavelength selective optical switches, without wavelength converters, where different signals may travel on the same communication link (but on different wavelengths) into a node, and then exit from it on different links, keeping their original wavelengths. See the recent survey [1] for an account of the theoretical problems and results obtained for this all-optical model.

We use the standard terminology of digraphs and flow networks [3]. A *dipath* ('directed path') in digraph $G = (V, A)$ is a sequence of nodes $P = (v_1, v_2, \dots, v_k)$, $k \geq 1$, such that $(v_i, v_{i+1}) \in A$ for $1 \leq i \leq k - 1$. For any sets S and S' in a digraph, we denote by $m(S, S')$ the number of arcs beginning in S and ending in S' . In a flow network [3], we denote by $c(u, v)$ the capacity of the arc (u, v) , and by $c(S, \bar{S})$ the capacity of the cut (S, \bar{S}) .

We are now ready to formulate our problems. Let G be a digraph. A *request* in G is an ordered pair of nodes (x, y) (corresponding to a message to be sent from x to y). An *instance* I in G is a collection (multiset) of requests (a request (x, y) may appear more than once). A *routing* R of an instance I is a collection of dipaths $R = \{P(x, y) \mid (x, y) \in I\}$, where $P(x, y)$ is a dipath from x to y .

Let G be a digraph and I an instance in G . For each routing R of I , we denote by $\vec{w}(G, I, R)$ the minimum number of wavelengths ('colours') that can be assigned to the dipaths in R , so that no two dipaths of the same wavelength share an arc. The *WDM problem* (G, I) asks for a routing R of I which minimizes $\vec{w}(G, I, R)$. We denote by $\vec{w}(G, I)$ this minimum of $\vec{w}(G, I, R)$ over all routings R for I .

Let again G be a digraph and I an instance. For each routing R of I , the *load* of an arc $\alpha \in A$ in the routing R , denoted by $\vec{\pi}(G, I, R, \alpha)$, is the number of dipaths of R containing α . The load of a routing R , denoted by $\vec{\pi}(G, I, R)$, is the maximum of $\vec{\pi}(G, I, R, \alpha)$ over all arcs $\alpha \in A$. The *congestion problem* (G, I) asks for a routing R of I which minimizes its load. We denote by $\vec{\pi}(G, I)$ this minimum of $\vec{\pi}(G, I, R)$ over all routings R for I .

The relevance of the parameter $\bar{\pi}$ to our problem is shown by the following lemma:

Lemma 1 $\bar{w}(G, I) \geq \bar{\pi}(G, I)$ for any instance I in any digraph G .

Proof. Indeed, to solve a given WDM problem (G, I) one has to use a number of wavelengths at least equal to the maximum number of dipaths having to share an arc. \square

In this paper, we are interested in a special case of instances where the collection of requests has the form $\{(x, y) \mid y \in Y\}$ for a fixed node $x \in V$, called the *originator*, and a multiset Y of nodes in V . Such an instance is called a *multicast* (or a *one-to-many*) *instance*. The particular instance where Y is the set V , is called a *broadcast*.

2 Multicasting and Network Flows

In the case of multicasting, we can model both of the above problems (the WDM problem as well as the congestion problem) by network flows. Consider a digraph $G = (V, A)$ with a multicast instance $I = \{(x, y) \mid y \in Y\}$, with originator x . In what follows we assume that Y is a set, i.e., that each node y appears at most once in Y . Indeed, if Y is a general multiset, we can transform the problem by adding to each destination y which appears in Y $\mu(y)$ times, a set of $\mu(y) - 1$ vertices of degree one, adjacent to y , each (as well as y) having multiplicity one.

We begin by modeling the congestion problem (G, I) (cf. Figure 1 (b)). Let s, t be two new vertices that will play the roles of source and sink. For every positive integer p , we define the network F_p to have the vertex set $V \cup \{s, t\}$, the arc set $\{(s, x)\} \cup A \cup (\bigcup_{y \in Y} \{(y, t)\})$, and the capacities $c(s, x) = \infty$, $c(u, v) = p$ for all $(u, v) \in A$, and $c(y, t) = 1$ for all $y \in Y$.

Proposition 2 $\bar{\pi}(G, I) \leq p$ if and only if F_p has a flow of value $|Y|$.

Proof. From the definitions. \square

For future reference we also consider the capacities of cuts in the network F_p . Suppose S is a subset of V and $\bar{S} = V \setminus S$. This gives rise to the cut $(S \cup \{s\}, \bar{S} \cup \{t\})$ in F_p . The capacity of this cut is infinite when $x \in \bar{S}$; otherwise it is $|S \cap Y| + p \cdot m(S, \bar{S})$.

According to the above proposition, $\bar{\pi}(G, I)$ is the smallest integer p such that F_p admits a flow of value $|Y|$. Combining this with the max-flow min-cut theorem of [3] we obtain

$$|S \cap Y| + \bar{\pi}(G, I) \cdot m(S, \bar{S}) \geq |Y|, \text{ for any } S \subseteq V. \quad (1)$$

Next we discuss the network flow model for the WDM problem (G, I) (cf. Figure 1 (c)). Let λ be a positive integer. For $1 \leq i \leq \lambda$, let $G_i = (V_i, A_i)$ be a copy of $G = (V, A)$, and let Y_0 be a copy of Y . For each $v \in V$, let v_i be the copy of v in V_i , and for each $y \in Y$, let y_0 be the copy of y in Y_0 . Let s, t be two new vertices which will be the source and sink. We define the network F'_λ to have the vertex set $\mathcal{V} = \{s, t\} \cup (\bigcup_{i=1}^\lambda V_i) \cup Y_0$, the arc set $\mathcal{A} = \bigcup_{i=1}^\lambda \{(s, x_i)\} \cup (\bigcup_{i=1}^\lambda A_i) \cup (\bigcup_{y \in Y} \bigcup_{i=1}^\lambda \{(y_i, y_0)\}) \cup (\bigcup_{y \in Y} \{(y_0, t)\})$, and except for the arcs (s, x_i) ($1 \leq i \leq \lambda$) of infinite capacity, all other arcs having capacity one.

Proposition 3 $\vec{w}(G, I) \leq \lambda$ if and only if F'_λ has a flow of value $|Y|$.

Proof. Suppose first that G admits dipaths from x to each $y \in Y$, which can be coloured by integers from $\{1, \dots, \lambda\}$ in such a way that no two dipaths sharing the same arc have the same color. Define the values $f'(u_i, v_i) = 1$, for each arc $(u, v) \in A$ which belongs to a dipath of color i ($1 \leq i \leq \lambda$). These values can be extended in a natural way to a flow f' of value $|Y|$ in F'_λ .

Conversely, assume that there exists a flow f' of value $|Y|$ in F'_λ . By construction of F'_λ , this implies that there is for each node $y \in Y_0$ an incoming arc with flow one. Hence for each $y \in Y_0$ there is exactly one i such that the node y_i has an incoming arc with flow one in G_i . In that G_i there must be a dipath from x_i to y_i consisting of arcs *all* having flow one, and we can assign the corresponding dipath in G the color i . Thereby are defined dipaths in G from x to each $y \in Y$, which are coloured by λ colours as required. \square

Theorem 4 Let G be a digraph and I a multicast instance in G . Then $\vec{w}(G, I) = \vec{\pi}(G, I)$, and an optimal solution to the WDM problem (G, I) can be found in polynomial time.

Proof. In view of lemma 1 and the above proposition, it will suffice to show that the network F'_λ , with $\lambda = \vec{\pi}(G, I)$, has a flow of value $|Y|$. We shall do this by showing that every cut has capacity at least $|Y|$, cf. [3].

Consider a cut $(S \cup \{s\}, \bar{S} \cup \{t\})$ in F'_λ , and let $S_i = S \cap V_i$ and $\bar{S}_i = \bar{S} \cap V_i$, for $1 \leq i \leq \lambda$, and let $S_0 = S \cap Y_0$ and $\bar{S}_0 = \bar{S} \cap Y_0$. Clearly, we may assume that each $x_i \in S_i$, $1 \leq i \leq \lambda$, otherwise the capacity of the cut will be infinite. Then the capacity is

$$c(S \cup \{s\}, \bar{S} \cup \{t\}) = \sum_{i \geq 1} m(S_i, \bar{S}_i) + \sum_{i \geq 1} m(S_i, \bar{S}_0) + |S_0|$$

For each $y \in Y$, let $S_y = \{y_i \mid i \geq 1, y_i \in S\}$. If Y_i denotes the copy of Y in V_i , then $\sum_{i \geq 1} |S_i \cap Y_i| = \sum_{y \in Y_0} |S_y|$. Note also that $\sum_{i \geq 1} m(S_i, \bar{S}_0) = \sum_{y \in Y_0} m(S_y, \bar{S}_0)$, and that $m(S_y, \bar{S}_0)$ is 0 if $y \in S_0$, and is $|S_y|$ if $y \in \bar{S}_0$. For every $y \in Y$, we have $|S_y| \leq \lambda$, and hence $\sum_{y \in S_0} |S_y| \leq \lambda |S_0|$, i.e., $|S_0| \geq \frac{1}{\lambda} \sum_{y \in S_0} |S_y|$. Summarizing, we obtain:

$$\begin{aligned} c(S \cup \{s\}, \bar{S} \cup \{t\}) &\geq \sum_{i \geq 1} m(S_i, \bar{S}_i) + \sum_{y \in \bar{S}_0} |S_y| + \frac{1}{\lambda} \sum_{y \in S_0} |S_y| \\ &\geq \sum_{i \geq 1} m(S_i, \bar{S}_i) + \frac{1}{\lambda} \sum_{y \in Y_0} |S_y| \\ &\geq \sum_{i \geq 1} \left(m(S_i, \bar{S}_i) + \frac{1}{\lambda} |S_i \cap Y_i| \right) \end{aligned}$$

Recall that we set $\lambda = \vec{\pi}(G, I)$. By the inequality (1), we have $|S_i \cap Y_i| + \lambda \cdot m(S_i, \bar{S}_i) \geq |Y|$ for each $i \geq 1$, and thus $c(S \cup \{s\}, \bar{S} \cup \{t\}) \geq |Y|$.

Our proof implies, in an obvious way, a flow-based algorithm to compute $\vec{w}(G, I)$ (and solve the WDM problem (G, I)) in polynomial time, for any multicast instance I in any digraph G . (Of course, the same applies to $\vec{\pi}(G, I)$ and the congestion problem.) \square

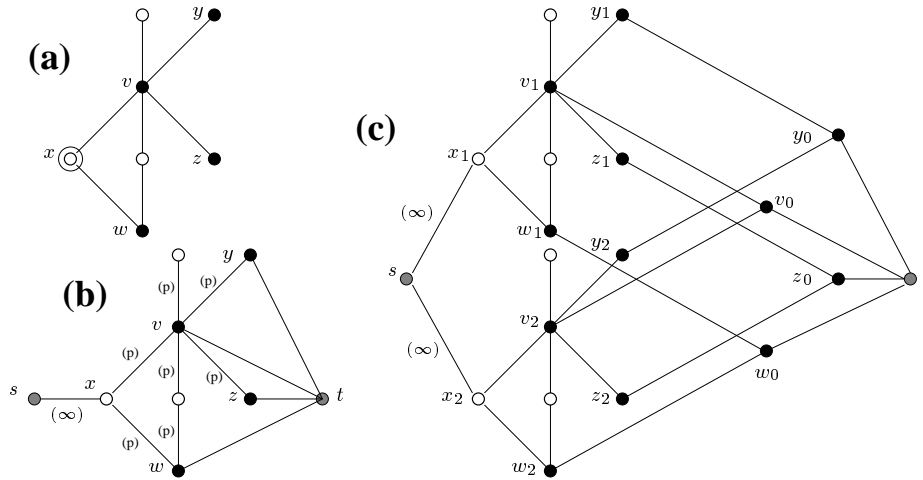


Figure 1: (a) A digraph G with originator x and where $Y = \{v, w, y, z\}$. (b) The network F_p . (c) The network F'_2 . All capacities not marked are equal to one. All arc orientations are omitted.

References

- [1] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes, and U. Vaccaro. Graph problems arising from wavelength–routing in all–optical networks. In *Proc. of 2nd Workshop on Optics and Computer Science*, Geneva, April 1997.
- [2] N. K. Cheung, K. Nosu, and G. Winzer. Special issue on dense WDM networks. *Journal on Selected Areas in Communications*, 8, 1990.
- [3] L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.
- [4] P. E. Green. *Fiber–Optic Communication Networks*. Prentice–Hall, 1993.
- [5] D. Minoli. *Telecommunications Technology Handbook*. Artech House, 1991.
- [6] B. Mukherjee. WDM-based local lightwave networks, Part 1: Single-hop systems. *IEEE Network Magazine*, 6(3):12–27, May 1992.

ANNEXE B

All-to-all communication in some wavelength-routed all-optical networks

All-to-All Communication for some Wavelength-Routed All-Optical Networks

Bruno Beauquier*

Thème 1 — Réseaux et systèmes
Projet Sloop

Juillet 1998

Abstract: This paper studies the problem of All-to-All Communication for optical networks. In such networks the vast bandwidth available is utilized through *wavelength division multiplexing* (WDM) : a single physical optical link can carry several logical signals, provided that they are transmitted on different wavelengths. In this paper we consider *all-optical* (or *single-hop*) networks, where the information, once transmitted as light, reaches its destination without being converted to electronic form in between, thus reaching high data transmission rates. In this model, we give optimal all-to-all protocols, using minimum numbers of wavelengths, for particular networks of practical interest, namely the d -dimensional square tori with even side, the corresponding meshes and the Cartesian sums of complete graphs.

Key-words: Optical networks, WDM, routing, all-to-all.

(Résumé : *tsvp*)

To appear in NETWORKS: An International Journal (Wiley-Interscience).

* SLOOP, joint project I3S-CNRS/UNSA/INRIA. Work partially supported by the French GDR/PRC Project PRS and by Galileo Project. Email : beauquier@sophia.inria.fr

Routage tout-optique pour l'échange complet dans des réseaux WDM

Résumé : Ce rapport étudie le problème réaliser simultanément toutes les communications possibles dans certaines classes de réseaux optiques. Dans ces réseaux, la forte bande passante disponible est utilisée par la technique du *multiplexage en longueur d'onde* (en anglais, *Wavelength Division Multiplexing* : WDM) : un seul lien physique en fibre optique peut transporter plusieurs signaux logiques, du moment qu'ils sont transmis à des longueurs d'onde différentes. Sont considérés ici des réseaux *tout-optiques*, où l'information, une fois convertie en lumière, atteint sa destination sans reconversion électronique intermédiaire. Cela permet des taux de transmission de données plus élevés. Pour ce modèle de réseaux optiques, nous donnons des protocoles de communication pour réaliser toutes les connexions possibles en même temps. Les topologies étudiées sont celles des tores et des grilles multidimensionnelles carrées, et des produits cartésiens de graphes complets. La plupart des résultats fournis sont optimaux quant au nombre de longueurs d'onde utilisées.

Mots-clé : Réseaux optiques, multiplexage en longueur d'onde, WDM, routage.

1 Introduction

Motivation. Optics is emerging as a key technology in communication networks, promising very high speed local or wide area networks in the future. A single optical wavelength supports rates of gigabits-per-second (which in turn support multiple channels of voice, data and video [11, 14]). Multiple laser beams that are propagated over the same fiber on distinct optical wavelengths can increase this capacity much further. This is achieved through WDM (Wavelength Division Multiplexing) [6], by partitioning the optical bandwidth into a large numbers of channels whose rates match those of the electronic transmission. This allows multiple data streams to be transferred concurrently along the same optical fiber.

All-optical (or *single-hop* [15]) communication networks provide all source-destination pairs with end-to-end transparent channels that are identified through a wavelength and a physical path. Maintaining the signal in optical form allows the elimination of the “electronic bottleneck” of networks with electronic switching.

It is worth pointing out the severe limitations that current optical technologies impose on the amount of available wavelengths per fiber. Therefore, solutions to the problem of efficient routing and wavelengths allocation have importance for the future development of the optical technology.

The Optical Model. In general, a WDM optical network consists of routing nodes interconnected by point-to-point fiber-optic links, which can support a certain number of wavelengths. Two optical signals on the same wavelength incoming on two input ports *must* be routed to different output ports, otherwise it is not possible to distinguish them later. In this paper we consider *switched* networks with reconfigurable wavelength selective optical switches, without wavelength converters, which can be based on acousto-optic filters, as done in [1, 2, 18]. In this kind of networks, signals for different requests may travel on the same communication link into a node (on different wavelengths) and then exit along different links, keeping their original wavelength. The only constraint on the solution is that no two paths in the network sharing the same optical link have the same wavelength assignment. See the recent survey [3] for an account of the theoretical results obtained for this all-optical model.

Some authors considered topologies with single undirected fiber links carrying undirected paths [1, 2, 16, 18]. However, it has since become apparent that optical amplifiers placed on the fiber will be directed devices. In this paper, each link is bidirectional and actually consists of a pair of unidirectional links. Hence we model the underlying fiber network as a symmetric directed graph $G = (V(G), A(G))$, where each arc represents a point-to-point unidirectional fiber-optic link.

A *solution* consists of settings for the switches in the network, and an assignment of wavelengths to the requests, so that there is a directed path (dipath) between the nodes of each request, and that no arc will carry two different signals on the same wavelength. If it is not possible to establish all requested connections at the same time because of resource limitations, some connections are to be deferred and established later. It is thus important to minimize the number of wavelengths used to service a requested communication pattern.

Contributions of this work. In this paper we study the problem of designing efficient routing and wavelength allocation for all-to-all communication in some all-optical networks. In particular we consider the multi-dimensional square tori and meshes, and the Cartesian sums of complete graphs. This work has to be considered as a step in understanding the complexity of the all-to-all problem in optical topologies relevant to local, metropolitan and wide area networks.

We consider the design of efficient algorithms for two widely used global (or structured) communication operations : the *total exchange* (also called *all-to-all* or *gossiping*) and the *complete exchange* (also called *personalized all-to-all* or *multi-scattering*). Formally these processes can be described as follows :

⊙ *Total exchange* : Each node in the network has a message and every node has to get all the messages.

⊙ *Complete exchange* : Each node u in the network has some messages $B(u, v)$, to be sent respectively to all the other nodes v .

In fact, these two communication schemes are equivalent in the all-optical model, if every node is able to send simultaneously different messages on different links and wavelengths. We briefly call *All-to-All protocol* any solution suitable to both problems.

In this paper, we obtain optimal All-to-All protocols for the d -dimensional hyper-square tori with even side, for the corresponding meshes, and for the Cartesian sums of complete graphs. Note that the same problem was solved for the binary hypercubes in [5, 16]. We refer to the recent survey [3] for other topologies and other communication patterns.

2 Preliminaries

We model an all-optical network by a *symmetric digraph*, that is a directed graph, with vertex set $V(G)$ and arc set $A(G)$, such that if $\alpha = (u, v) \in A(G)$ then $\alpha' = (v, u) \in A(G)$. The number of vertices in G is always denoted by $N = |V(G)|$. The following notation is also used (for more details, see [7] or any textbook in graph theory like [4]) :

- $P(x, y)$ denotes a *dipath* in G from node x to node y , that is, a directed path which consists of a sequence of consecutive arcs beginning in x and ending in y .
- $\delta(x, y)$ denotes the *distance* from x to y in G , that is, the minimum length of a dipath $P(x, y)$.
- The *Cartesian sum* (also called *Cartesian product*) of two digraphs G and G' is the digraph whose vertices are the ordered pairs (x, x') where x is a vertex of G and x' is a vertex of G' . Thus its vertex set is the usual Cartesian sum of the vertex sets $V(G)$ and $V(G')$. Furthermore there is an arc from (x, x') to (y, y') if and only if $x = y$ and (x', y') is an arc of G' , or $x' = y'$ and (x, y) is an arc of G .
- $Z_n = \{0, 1, \dots, n - 1\}$ denotes the set of integers modulo n .
- Z_n^d denotes the Cartesian sum of d copies of Z_n .

- C_N and M_N denote respectively the cycle and the chain with N vertices.
- G^d denotes the Cartesian sum of d copies of a digraph G . Hence C_n^d is the d -dimensional hyper-square torus with side n , and M_n^d is the d -dimensional hyper-square mesh with side n .
- K_N denotes the complete graph with N vertices.

Wavelength-routing problem

- A *request* is an ordered pair of nodes (x, y) in G (corresponding to a message to be sent by node x to node y).
- An *instance* I is a collection of requests.
- A *routing* R for an instance I in G is a set of dipaths $R = \{P(x, y) \mid (x, y) \in I\}$, realizing the requests of I .
- The *conflict graph* associated to a routing R is the undirected graph (R, E) with vertex set R and such that two dipaths of R are adjacent if and only if they share an arc of G .
- Let G be a digraph and I an instance. The *problem* (G, I) consists of finding a routing R for the instance I and assigning each request $(x, y) \in I$ a wavelength, so that no two dipaths of R sharing an arc have the same wavelength. If we think of wavelengths as colours, the problem (G, I) seeks a routing R and a vertex colouring of the conflict graph (R, E) , such that two adjacent vertices are coloured differently. We denote by $\vec{w}(G, I, R)$ the chromatic number of (R, E) , and by $\vec{w}(G, I)$ the smallest $\vec{w}(G, I, R)$ over all routings R . Thus $\vec{w}(G, I, R)$ is the minimum number of wavelengths for a routing R and $\vec{w}(G, I)$ the minimum number of wavelengths over all routings for (G, I) .

For a general network G and an arbitrary instance I , the problem of determining $\vec{w}(G, I)$ has been proved to be NP-hard in [8]. In particular, it has been proved that determining $\vec{w}(G, I)$ is NP-hard for trees and cycles. In [9] these results have been extended to binary trees and meshes.

A related parameter

- Given a digraph G and a routing R for an instance I , the *load* of an arc $\alpha \in A(G)$ for the routing R , denoted by $\vec{\pi}(G, I, R, \alpha)$, is the number of dipaths of R containing α . The *load* of G for the routing R (also called *congestion*), denoted by $\vec{\pi}(G, I, R)$, is the maximum load of an arc of G for the routing R , that is, $\vec{\pi}(G, I, R) = \max_{\alpha \in A(G)} \vec{\pi}(G, I, R, \alpha)$.
- The *load* of G for an instance I , denoted by $\vec{\pi}(G, I)$, is the minimum load of G for a routing R for I , that is, $\vec{\pi}(G, I) = \min_R \vec{\pi}(G, I, R)$.

The relevance of this parameter to our problem is shown by the following lemma :

Lemma 1 $\vec{w}(G, I) \geq \vec{\pi}(G, I)$ for any instance I in any digraph G .

Proof. To solve the problem (G, I) , first a routing R for the instance I is needed. Then the number $\vec{w}(G, I, R)$ of wavelengths necessary to be properly assigned to the dipaths of R is at least the maximum number of dipaths sharing an arc of G , that is $\vec{\pi}(G, I, R)$. As $\vec{w}(G, I)$ is some $\vec{w}(G, I, R)$ for some routing R and by definition of $\vec{\pi}(G, I)$, the lemma holds. \square

The inequality in Lemma 1 can be strict, as shown by Figure 1.

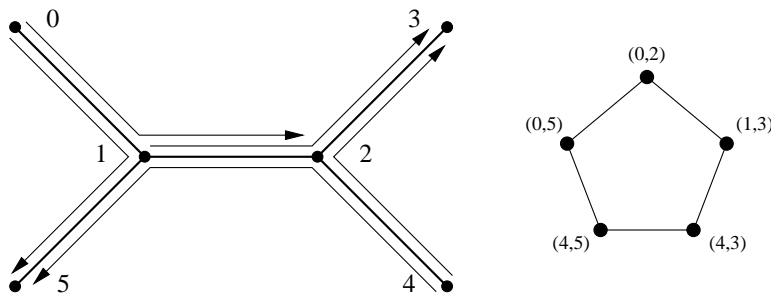


Figure 1: A routing for five requests in a tree G and its associated conflict graph.

Indeed, for this instance I in this tree G , the load is $\vec{\pi}(G, I) = 2$ but $\vec{w}(G, I) = 3$, since the conflict graph is a cycle of length 5 which has chromatic number 3.

In general, minimizing the number of wavelengths is not the same problem as that of realizing a routing that minimizes the number of dipaths sharing an arc. Indeed, our problem is made much harder due to the further requirement of wavelengths assignment on the dipaths. This is the case for trees. In order to get equality in Lemma 1, a routing R such that $\vec{\pi}(G, I, R) = \vec{\pi}(G, I)$ should be found, for which the associated conflict graph is $\vec{\pi}(G, I)$ -vertex colourable.

In this paper, we consider a specific instance I_A in G , called the *All-to-All* instance, which consists of all the different ordered pairs of vertices in $V(G)$. Hence $I_A = \{(x, y) \mid x \in V(G), y \in V(G), x \neq y\}$. In this case, $\vec{\pi}(G, I_A)$ is called the *arc-forwarding index* [13] of G . The same parameter defined similarly for undirected graphs is called the *edge-forwarding index* and has been studied in [12, 20].

3 Hypersquare torus C_n^d

In this section we consider the problem (C_n^d, I_A) for the All-to-All instance I_A in the d -dimensional hypersquare torus with side n . We first solve optimally the case where n is even. Then we will give a nearly optimal solution for the case where n is odd. Note that for n odd and $d = 2$ the problem has been optimally solved in [19].

Theorem 2 *In the d -dimensional hypersquare torus C_n^d , if n is even then we have :*

$\vec{w}(C_n^d, I_A) = \vec{\pi}(C_n^d, I_A) = n^{d+1}/8$, and if n is odd then :

$\vec{\pi}(C_n^d, I_A) = n^{d-1} \lfloor n^2/4 \rfloor / 2 = (n^2 - 1)n^{d-1}/8 \leq \vec{w}(C_n^d, I_A) \leq (n+1)^{d+1}/8 = \vec{w}(C_{n+1}^d, I_A)$.

Before presenting the proof, we introduce two more lemmas. The first one is adapted from an analogous property given in [12] in the undirected case.

Lemma 3 *Let G be a digraph and R a routing in G for the All-to-All instance I_A . Then $\vec{\pi}(G, I_A, R) = (\sum_{x \neq y} \delta(x, y)) / |A(G)|$ if and only if R is a routing of shortest dipaths loading equally each arc of G .*

Proof. The sum of the loads of all the arcs for $R = \{P(x, y) \mid x \neq y\}$ is equal to the sum of the lengths of all the dipaths in R . If R is a routing of shortest dipaths loading equally each arc, then the load is $\vec{\pi}(G, R) = (\sum_{x \neq y} \delta(x, y)) / |A(G)|$.

Conversely, if the maximum load of an arc for R is $\vec{\pi}(G, R) = (\sum_{x \neq y} \delta(x, y)) / |A(G)|$ then by summing the loads over all the arcs we get $\sum_{x \neq y} |P(x, y)| \leq \sum_{x \neq y} \delta(x, y)$. Therefore the equality holds and each arc is maximally loaded. Thus R is a routing of shortest dipaths loading equally each arc. \square

A set S of dipaths in a digraph G is said to be *covering* if each arc in $A(G)$ is contained in a dipath of S .

Lemma 4 *Let G be a digraph such that $\vec{\pi}(G, I_A) = (\sum_{x \neq y} \delta(x, y)) / |A(G)|$.*

We have $\vec{w}(G, I_A) = \vec{\pi}(G, I_A)$ if and only if there exist a routing R of shortest dipaths for I_A and a partition of R in some covering sets of arc-disjoint dipaths.

Proof. A partition of a routing of shortest dipaths $R = \{P(x, y) \mid x \neq y\}$ in some covering sets of arc-disjoint dipaths provides a solution to the problem (G, I_A) by assigning a different colour to each set of the partition. Thereby the number of colours used is the load of any arc for R , equal to $\vec{\pi}(G, I_A)$ according to Lemma 3. From Lemma 1 it follows that $\vec{w}(G, I_A) = \vec{\pi}(G, I_A)$.

Conversely if $\vec{w}(G, I_A) = \vec{\pi}(G, I_A)$, then there exist a routing R for I_A and a colouring of its dipaths with $\vec{\pi}(G, I_A)$ colours, so that no two dipaths with the same colour share an arc. This colouring gives a partition of R in $\vec{\pi}(G, I_A)$ sets of arc-disjoint dipaths. Each of them is a covering set because the sum of the lengths of the coloured dipaths is at least $\sum_{x \neq y} \delta(x, y) = \vec{\pi}(G, I_A) \cdot |A(G)|$. As the load of G for R is $\vec{\pi}(G, I_A)$, R is a routing of shortest dipaths, according to Lemma 3. \square

A straight way to prove the first part of Theorem 2 would be to assign properly a coloured dipath to each request in the instance I_A , using a total number of colours (or wavelengths) equal to $\bar{\pi}(C_n^d, I_A)$. In view of Lemma 4, we have found more convenient to provide an optimal solution by assigning to each colour a covering set of arc-disjoint shortest dipaths, so that each request in I_A gets exactly one coloured dipath. Thereby the condition on a suitable partition stated in Lemma 4 is satisfied and we ensure optimality without counting the total number of colours used. Before going further, some more definitions and notation are necessary.

Definitions 5 The vertex set $V(C_n^d)$ is represented by Z_n^d . The elements in Z_n^d may be expressed in the canonical base $\{e_i\}_{1 \leq i \leq d}$, that is, we may denote $x = (x_1, x_2, \dots, x_d) \in Z_n^d$ by $(\sum_{i=1}^d x_i \cdot e_i)$. Let $J = \sum_{i=1}^d e_i$. Given two subsets X and Y of Z_n^d , we define $X + Y$ as the subset $\{x + y \mid x \in X, y \in Y\}$.

Remark 6 C_n^d can be defined as a Cayley digraph on the Abelian group Z_n^d . It is an arc-transitive digraph where each vertex $x = (x_1, \dots, x_d)$ is joined by $2d$ arcs to the vertices $x \pm e_i = (x_1, \dots, x_i \pm 1, \dots, x_d)$ for $1 \leq i \leq d$.

We distinguish two kinds of dipaths to be assigned to the requests in the instance I_A :

Definitions 7 Given an arc α from x to y , we say that α is in the dimension i and in the *progressive (regressive)* direction if $y = x + e_i$ (if $y = x - e_i$). These definitions are extended to the dipaths containing only arcs of the same kind.

Given a request (x, y) such that $y = x + \sum_{i=1}^d a_i \cdot e_i$, we define *the ascending dipath* $P_a(x, y)$ as the concatenation $(P_1 P_2 \dots P_d)$ of d dipaths, such that :

- P_1 is the shortest dipath from x to $(x + a_1 \cdot e_1)$, in the progressive direction if $a_1 = n/2$.
- P_i ($2 \leq i \leq d$) is the shortest dipath from $(x + \sum_{j=1}^{i-1} a_j \cdot e_j)$ to $(x + \sum_{j=1}^i a_j \cdot e_j)$, in the progressive direction if $a_i = n/2$.

Given a request (x, y) , we define *the descending dipath* $P_d(x, y)$ as the reverse dipath of the ascending dipath $P_a(y, x)$, that is the dipath made of all the symmetric arcs of those of $P_a(y, x)$.

For example in C_4^2 , the ascending dipath from node $(0, 0)$ to $(2, 1)$ goes through nodes $(1, 0)$ and $(2, 0)$, while the descending dipath connecting the same nodes goes through $(0, 1)$ and $(3, 1)$.

In order to define the different subsets of requests to be assigned covering sets of arc-disjoint dipaths, we need the following definitions, when n is even :

Definitions 8 Let $n = 2k$. For every node $x = (x_1, x_2, \dots, x_d) \in Z_n^d$, *the level of x* is defined as $L(x) = \sum_{i=1}^d x_i \in Z_n$. For every request (x, y) , *the move of (x, y)* is defined as $m(x, y) = y - x \in Z_n^d$. Given $m \in Z_n^d$, let $\bar{m} = (k, \dots, k) - m$. An equivalence relation E is defined in Z_n^d as follows :

$$m \sim m' \text{ if and only if } (m' = m \text{ or } m' = -m \text{ or } m' = \bar{m} \text{ or } m' = -\bar{m})$$

Let us denote by $E(m)$ the E -class of m . In each E -class a special move m_0 is chosen.

Now we distinguish three different kinds of moves. First let $M = \{m \in Z_n^d \mid m \neq -m \text{ and } m \neq \overline{m}\}$. If $m \in M$ then $E(m)$ has 4 elements all in M . For example in Z_8^2 , $m = (2, 1) \in M$ and $E(m) = \{(2, 1), (-2, -1), (2, 3), (-2, -3)\}$. Now let $K = \{m \in Z_n^d \mid m = -m\}$. If $m \in K$ then $E(m)$ has 2 elements all in K . For example in Z_8^2 , $m = (4, 0) \in K$ and $E(m) = \{(4, 0), (0, 4)\}$. At last let $H = \{m \in Z_n^d \mid m = \overline{m}\}$ (H is empty if k is odd). If $m \in H$ then $E(m)$ has 2 elements all in H . For example in Z_8^2 , $m = (2, -2) \in H$ and $E(m) = \{(2, -2), (-2, 2)\}$. Note that every move belongs to exactly one of these sets M , K or H .

Requests with move in M

For $l \in Z_n$, \mathcal{HP}_l denotes the hyperplane orthogonal to J containing all the nodes of level l , that is $\mathcal{HP}_l = \{x \in Z_n^d \mid L(x) = l\}$.

A node y is said to be translated from a node x by a vector v if $y = x + v$. In what follows, we use a generalization of this definition for dipaths : the dipath translated by a vector v from a dipath P is induced by all the nodes translated from those of P .

Given a node x_0 and a special move $m_0 \in M$, we first consider the two symmetric dipaths $P_1 = P_a(x_0, x_0 + m_0)$ and $P'_1 = P_d(x_0 + m_0, x_0)$ and the two symmetric dipaths $P_2 = P_d(x_0 + m_0, x_0 + k.J)$ and $P'_2 = P_a(x_0 + k.J, x_0 + m_0)$. In what follows we show that all the dipaths translated from P_1, P'_1, P_2, P'_2 by all the vectors in $(\mathcal{HP}_0 \cup \mathcal{HP}_k)$ form a covering set of arc-disjoint dipaths. We give an example by drawing, as shown in Figure 2-b, for $d = 2$, $n = 8$ and $m_0 = (2, 1)$. So $-m_0 = (-2, -1)$, $\overline{m_0} = (2, 3)$ and $-\overline{m_0} = (-2, -3)$. Figure 2-a shows only the dipaths translated by the vectors in \mathcal{HP}_0 . Wrap-around connections of the torus are omitted for clarity and the dipaths P_1, P'_1, P_2 and P'_2 are drawn in bold.

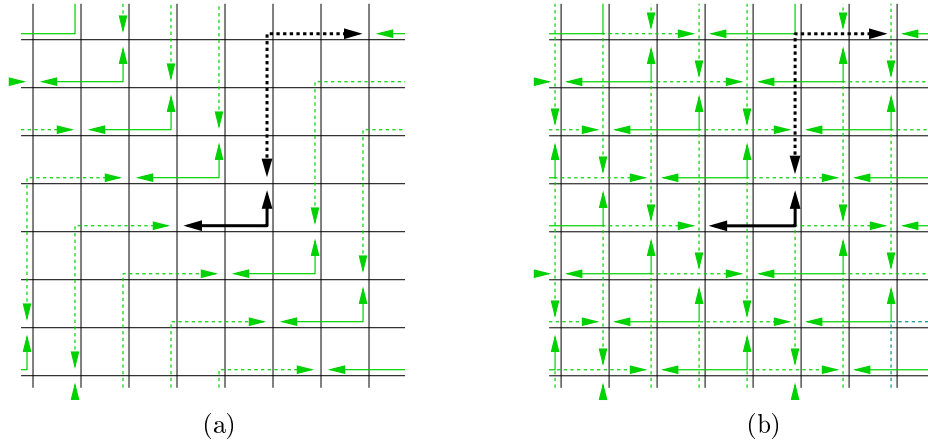


Figure 2: Construction of a covering set of arc-disjoint dipaths in C_8^2 .

Such a set of dipaths is assigned the same colour and contains one dipath for each request (x, y) such that :

- $L(x) = L(x_0) \bmod k$ and $m(x, y) = m_0$, or
- $L(x) = L(x_0 + m_0) \bmod k$ and $m(x, y) = -m_0$, or
- $L(x) = L(x_0 + m_0) \bmod k$ and $m(x, y) = \overline{m_0}$, or
- $L(x) = L(x_0) \bmod k$ and $m(x, y) = -\overline{m_0}$.

Such a set depends thus only on the value of $(L(x_0) \bmod k)$ and on m_0 . The corresponding colour is represented by an element of $C_M = (Z_k \times E(M))$, where $E(M)$ is the set of the classes of M . Notice that in this way every request with move in M is assigned exactly one coloured dipath. Given a colour $c \in C_M$, every dipath coloured by c is called a c -dipath.

Claim 9 *For every colour c in C_M , the set of c -dipaths is a covering set of arc-disjoint dipaths.*

Proof. For $1 \leq i \leq d$, A_i^+ denotes the set of the arcs in the dimension i and in the progressive direction. Let us show that each arc in A_i^+ is contained in exactly one c -dipath. As the c -dipaths are defined as symmetric pairs, the property will also hold for the arcs in the opposite direction.

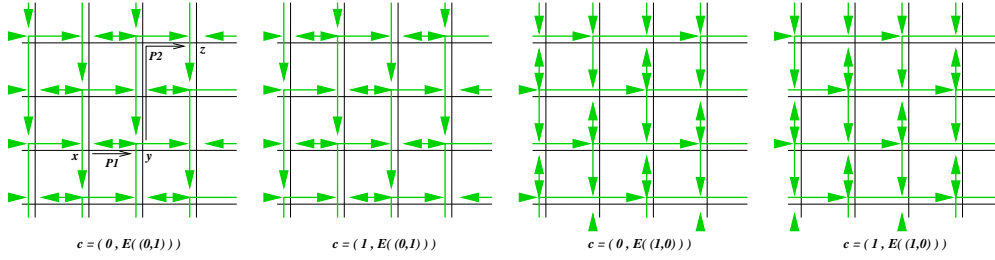
There are as many arcs in A_i^+ as nodes in C_n^d , that is, $n^d = N$. On one hand, we show that the sum over all the c -dipaths of the number of arcs in A_i^+ is equal to N . On the other hand, we show that each arc in A_i^+ is contained in a c -dipath.

Let $c = (\lambda, E(m)) \in C_M$, with $\lambda \in Z_k$, and $m_0 = \sum_{j=1}^d a_j \cdot e_j \in E(m)$ the special move chosen in $E(m)$. We assume that $a_i \in \{0, 1, \dots, k\}$, otherwise the proof is similar. Then the only c -dipaths using arcs in A_i^+ have their move equal to m_0 or $\overline{m_0}$.

There are exactly $(2n^{d-1})$ c -dipaths having move m_0 , since there are $2n^{d-1}$ vectors in $(\mathcal{HP}_0 \cup \mathcal{HP}_k)$. As many c -dipaths have move $\overline{m_0}$. Each of those having move m_0 (respectively $\overline{m_0}$) uses \tilde{a}_i (resp. $k - \tilde{a}_i$) arcs in A_i^+ , where \tilde{a}_i is the integer representative of a_i in $\{0, \dots, k\} \subset Z$. Therefore the c -dipaths use $2kn^{d-1} = N$ times an arc in A_i^+ .

As $x + (\mathcal{HP}_0 \cup \mathcal{HP}_k) = \mathcal{HP}_{L(x)} \cup \mathcal{HP}_{L(x)+k}$, if X is a set of k nodes whose levels form the entire set Z_k , then $X + (\mathcal{HP}_0 \cup \mathcal{HP}_k) = \bigcup_{l=0}^{n-1} \mathcal{HP}_l = Z_n^d$. Thus, if there exist k arcs in A_i^+ contained in c -dipaths and whose origins have all the different levels modulo k , then each arc in A_i^+ is contained in a c -dipath.

Let $x \in Z_n^d$ such that $L(x) = \lambda \bmod k$, $y = x + m_0$ and $z = y + \overline{m_0}$. Let $m_1 = \sum_{j=1}^{i-1} a_j \cdot e_j$ ($m_1 = 0$ if $i = 1$) and $m_2 = \sum_{j=i+1}^d a_j \cdot e_j$ ($m_2 = 0$ if $i = d$), so that $m_0 = m_1 + a_i \cdot e_i + m_2$. First, $P_1 = P_a(x, y)$ is an ascending c -dipath containing the \tilde{a}_i consecutive arcs in A_i^+ between $x + m_1$ and $x + m_1 + a_i \cdot e_i$. Secondly, $P_2 = P_d(y, z)$ is a descending c -dipath containing the $(k - \tilde{a}_i)$ consecutive arcs in A_i^+ between $y + \overline{m_2}$ and $y + \overline{m_2} + (k - a_i) \cdot e_i$. On the whole, these k arcs have origins with all the different levels modulo k , because $L(x + m_1 + a_i \cdot e_i) = L(y - m_2) = L(y + \overline{m_2}) \bmod k$. \square

Figure 3: Sets of c -dipaths in C_4^2 for requests with move in M .

Requests with move in K

Let us now consider the case of requests with move $m = \sum_{j=1}^d a_j \cdot e_j \in Z_n^d$ such that $2m = 0$. In this case we have $a_j \in \{0, k\}$ for each $j \in \{1, \dots, d\}$ and the cardinality of $K = \{m \in Z_n^d \mid 2m = 0\}$ is 2^d . As noticed before $E(m)$ has 2 elements for every $m \in K$ and the set of classes $E(K)$ has 2^{d-1} elements. The construction of sets of arc-disjoint dipaths covering all the arcs is similar to the previous one, except that requests with moves in two E -classes are involved in the same set.

Given $\lambda \in Z_k$, m and m' two moves taken in two different E -classes, we consider all the dipaths $P_a(x, y)$ for $L(x) = \lambda \bmod k$ and $m(x, y) \in E(m)$, and all the dipaths $P_d(x, z)$ for $L(x) = \lambda \bmod k$ and $m(x, z) \in E(m')$. Recall that these ascending (descending) dipaths contain only arcs in the progressive (regressive) direction, in accordance with the definitions 7. Obviously, in any dimension and in any direction there exist k consecutive arcs used by one of these dipaths. From the same arguments as in the previous proof, it follows that these dipaths form a covering set of arc-disjoint dipaths.

Note that this dipaths assignment does not lead to a symmetric routing (where the dipaths for any two requests (x, y) and (y, x) are symmetric).

Requests with move in H

At last we consider the case of requests with move $m = \sum_{j=1}^d a_j \cdot e_j \in Z_n^d$ such that $m = \bar{m} = k \cdot J - m$. As noticed before, such requests exist only if k is even, say $k = 2h$, and we have $a_j \in \{h, -h\}$ for $1 \leq j \leq d$.

Given $\mu \in Z_h$ and a special move $m_0 \in H$, we define the set of all the symmetric dipaths $P_a(x, x + m_0)$ and $P_d(x + m_0, x)$ such that $L(x) = \mu \bmod h$. By the same arguments again, it can be shown that these dipaths form a covering set of arc-disjoint dipaths.

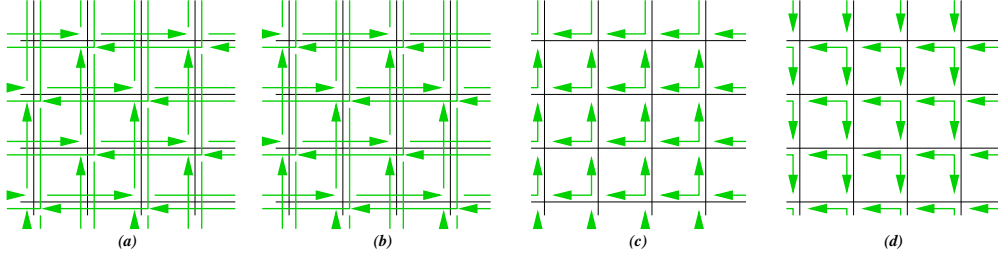


Figure 4: Sets of c -dipaths in C_4^2 for requests with move in K (a)(b) and in H (c)(d).

Proof of Theorem 2. Following a result given in [12] for the undirected torus, we have $\vec{\pi}(C_n^d, I_A) = n^{d-1} \lfloor n^2/4 \rfloor / 2 = (\sum_{x \neq y} \delta(x, y)) / |A(G)|$. The assumption of Lemma 4 is thus satisfied. In the case where n is even, we have shown how to obtain some sets of arc-disjoint shortest dipaths, each set covering all the arcs of the torus C_n^d , so that every request in the All-to-All instance I_A is assigned exactly one dipath. As a consequence of Lemma 4, we have $\vec{w}(C_n^d, I_A) = \vec{\pi}(C_n^d, I_A)$. \square

Now we consider the case where n is odd and prove the second part of Theorem 2.

Proposition 10 In the d -dimensional hypersquare torus C_n^d , if n is odd we have :
 $\vec{\pi}(C_n^d, I_A) = n^{d-1} \lfloor n^2/4 \rfloor / 2 = (n^2 - 1)n^{d-1}/8 \leq \vec{w}(C_n^d, I_A) \leq (n + 1)^{d+1}/8 = \vec{w}(C_{n+1}^d, I_A)$.

Proof. From Lemma 1, we have $\vec{\pi}(C_n^d, I_A) \leq \vec{w}(C_n^d, I_A)$. To prove the upper bound we make use of the routing and the colouring given by the solution for $n + 1$. We consider C_n^d as obtained from C_{n+1}^d by removing a symmetric cycle in each dimension and joining up each pair of disconnected links. Formally, we remove from $V(C_{n+1}^d)$ all the nodes having a component equal to 0, which form the set denoted by $V_0(C_{n+1}^d)$. In the induced subgraph we connect together with two symmetric new arcs all the pairs of remaining nodes of type $\{(\sum_{j \neq i} x_j \cdot e_j) - e_i, (\sum_{j \neq i} x_j \cdot e_j) + e_i\}$. Note that if a dipath in C_{n+1}^d uses two consecutive arcs in different dimensions around a node in $V_0(C_{n+1}^d)$, then either its source or its destination is in $V_0(C_{n+1}^d)$, according to the definitions 7. Thus, all the dipaths with source and destination out of $V_0(C_{n+1}^d)$ can be modified from C_{n+1}^d to C_n^d by following the arc transformation.

Therefore there is for every odd value of n a solution in C_n^d using $\vec{w}(C_{n+1}^d, I_A)$ colours. \square

4 Hypersquare mesh M_n^d

In this section we consider the problem (M_n^d, I_A) for the All-to-All instance I_A in the d -dimensional hypersquare mesh with side n . The construction of coloured dipaths is based on the solution for the torus given in the previous section. Note that the proof given in [19] for the mesh is not sound.

Theorem 11 *In the d -dimensional hypersquare mesh M_n^d , if n is even then we have : $\vec{w}(M_n^d, I_A) = \vec{\pi}(M_n^d, I_A) = n^{d+1}/4$, and if n is odd then : $\vec{\pi}(M_n^d, I_A) = n^{d-1} \lfloor n^2/4 \rfloor = (n^2 - 1)n^{d-1}/4 \leq \vec{w}(M_n^d, I_A) \leq (n + 1)^{d+1}/4 = \vec{w}(M_{n+1}^d, I_A)$.*

Proof. Consider the mesh M_n^d as obtained from the torus C_n^d by removing all the pairs of wrap-around symmetric arcs. For every wrap-around arc contained in a dipath P in C_n^d , let p be the longest sub-dipath of P using this arc and using arcs in the same dimension only. Let p' be the dipath connecting the same nodes as p but using arcs in the other direction. By replacing p by p' for every wrap-around arc contained in P , we obtain a new dipath P' not using any wrap-around arc. Any such dipath in the torus C_n^d induces a dipath in the mesh M_n^d . This gives a routing R' for the instance I_A in M_n^d from the routing R in C_n^d given in the previous section.

Let us now define the new colour assignment. The following property holds in the solution given previously to the problem (C_n^d, I_A) : every set of c -dipaths can be partitioned into two subsets, so that no two arcs with opposite directions in the same dimension are used respectively by two c -dipaths in the same subset. Thus, by construction the dipaths in each of the two corresponding subsets obtained in the mesh M_n^d are pairwise arc-disjoint and can be assigned two new colours c_1 and c_2 , respectively. So twice as many colours as in the torus C_n^d are used in the mesh M_n^d . As from [12] $\vec{\pi}(M_n^d, I_A) = 2\vec{\pi}(C_n^d, I_A)$, the theorem holds, according to Theorem 2. \square

5 Cartesian sum of complete graphs

The following theorem generalizes the result obtained in [16, 5] for hypercubes.

Theorem 12 *Let n_1, n_2, \dots, n_d be integers such that $2 \leq n_1 \leq n_2 \leq \dots \leq n_d$. Let us denote by $K(n_1, n_2, \dots, n_d)$ the Cartesian sum of the d complete graphs K_{n_i} ($1 \leq i \leq d$). We have : $\vec{w}(K(n_1, n_2, \dots, n_d), I_A) = \vec{\pi}(K(n_1, n_2, \dots, n_d), I_A) = \prod_{i=2}^d n_i$.*

Proof. In the sequel $K(n_1, n_2, \dots, n_d)$ is denoted simply by G . The vertex set is represented by $(Z_{n_1} \times Z_{n_2} \times \dots \times Z_{n_d})$. There is a pair of symmetric arcs between two nodes if they differ in exactly one component. The elements in $V(G)$ may be expressed in the canonical base $\{e_i\}_{1 \leq i \leq d}$.

The graph G can be seen as n_1 copies of $K(n_2, n_3, \dots, n_d)$ connected together with $n_1(n_1 - 1) \prod_{i=2}^d n_i$ arcs. Because there are $n_1(n_1 - 1)(\prod_{i=2}^d n_i)^2$ requests between the pairwise distinct copies, it follows that one of these arcs must be contained in at least $(\prod_{i=2}^d n_i)$ dipaths. Therefore we have $\vec{\pi}(G) \geq \prod_{i=2}^d n_i$. According to Lemma 1, it remains to show that $\vec{w}(G, I_A) \leq \prod_{i=2}^d n_i$.

A dipath (u_0, u_1, \dots, u_k) from u_0 to u_k is called *ascending* if for $1 \leq i \leq k$ the node u_i is obtained from u_{i-1} by changing the component in position p_i , so that $p_1 < p_2 < \dots < p_k$. We assign to each request (x, y) the ascending dipath $P(x, y)$.

Given a_1 in Z_{n_1} , let α_1 be the integer representative of a_1 in $\{0, 1, \dots, n_1 - 1\} \subset Z$. For $2 \leq i \leq d$, $(a_1)_{n_i}$ denotes the element of Z_{n_i} having α_1 as integer representative.

The set of colours is represented by $(Z_{n_2} \times Z_{n_3} \times \dots \times Z_{n_d})$. To each dipath $P(x, y)$, with $x = (x_1, x_2, \dots, x_d)$ and $y = (y_1, y_2, \dots, y_d)$, we assign the colour $c(x, y) = ((y_j - x_j) + (x_1)_{n_j})_{2 \leq j \leq d}$.

We prove now that each arc $\alpha = (z, z + \lambda_i \cdot e_i)$, with $\lambda_i \in Z_{n_i}$, is not contained in two dipaths with the same colour. As ascending dipaths are considered, the arc α is contained only in dipaths $P(x, y)$ with $x = (x_1, \dots, x_{i-1}, z_i, \dots, z_d)$ and $y = (z_1, \dots, z_{i-1}, (z_i + \lambda_i), y_{i+1}, \dots, y_d)$. So we have :

$$c(x, y) = ((z_2 - x_2) + (x_1)_{n_2}, \dots, (z_{i-1} - x_{i-1}) + (x_1)_{n_{i-1}},$$

$$\lambda_i + (x_1)_{n_i}, (y_{i+1} - z_{i+1}) + (x_1)_{n_{i+1}}, \dots, (y_d - z_d) + (x_1)_{n_d}).$$

Consider now any other dipath $P(x', y')$ containing the arc α . Again, we have :

$$c(x', y') = ((z_2 - x'_2) + (x'_1)_{n_2}, \dots, (z_{i-1} - x'_{i-1}) + (x'_1)_{n_{i-1}},$$

$$\lambda_i + (x'_1)_{n_i}, (y'_{i+1} - z_{i+1}) + (x'_1)_{n_{i+1}}, \dots, (y'_d - z_d) + (x'_1)_{n_d}).$$

Assume that $c(x, y) = c(x', y')$. Thus, $\lambda_i + (x_1)_{n_i} = \lambda_i + (x'_1)_{n_i} \pmod{n_i}$ and so $x_1 = x'_1 \pmod{n_1}$ because $n_1 \leq n_i$. From the identities of the other components, it follows that $x = x'$ and $y = y'$. \square

6 Final remarks

In this paper we have obtained optimal and nearly optimal All-to-All protocols in some switched all-optical networks. It remains to prove that the equality $\vec{w} = \vec{\pi}$ also holds for square tori and meshes with odd side (some bidimensional cases of tori have been solved in [19]), and more generally for any torus and any mesh.

We considered topologies that are all Cartesian sums of simple graphs, namely cycles, chains and complete graphs. An interesting issue deserves to be investigated : to obtain results for general Cartesian sums and to find a way to design an efficient solution for the Cartesian sum of two graphs, for which efficient solutions are known.

A recent work is worth to be pointed out. In [10] it is proved that the equality $\vec{w}(T, I_A) = \vec{\pi}(T, I_A)$ holds for any tree T . In view of all the results obtained by now, it is likely that the equality can be achieved for any symmetric digraph, indeed for any digraph.

Finally we give two other future lines of research. The computation complexity of the quantity $\vec{w}(G, I_A)$ remains to be determined. It is likely that this problem is NP-hard. Therefore, it will be of interest to design approximation algorithms. Fault tolerant issues have to be considered too. See the survey [17] for an account of the vast literature on fault-tolerance in traditional networks.

References

- [1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing and scheduling algorithms for optical networks. In *Proc. of the 5th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 412–423. ACM Press, January 1994.
- [2] Y. Aumann and Y. Rabani. Improved bounds for all optical routing. In *Proc. of the 6th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 567–576. ACM Press, January 1995.
- [3] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes, and U. Vaccaro. Graph problems arising from wavelength-routing in all-optical networks. In *Proc. of 2nd Workshop on Optics and Computer Science (WOCS), part of IPPS*, Genève, April 1997.
- [4] C. Berge. *Graphs and Hypergraphs*. North Holland, Amsterdam, 1973.
- [5] J-C. Bermond, L. Gargano, S. Pérennes, A. A. Rescigno, and U. Vaccaro. Efficient collective communication in optical networks. In *Automata, Languages and Programming, 23rd International Colloquium (ICALP)*, volume 1099 of *Lecture Notes in Computer Science*, pages 574–585, Paderborn, Germany, July 1996. Springer-Verlag.
- [6] N. K. Cheung, K. Nosu, and G. Winzer. Special issue on dense WDM networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, 8, 1990.
- [7] Jean de Rumeur. *Communications dans les réseaux de processeurs*. Collection Études et Recherches en Informatique. Masson, Paris, 1994. English version to appear.
- [8] T. Erlebach and K. Jansen. Scheduling of virtual connections in fast networks. In *Proc. of Parallel Systems and Algorithms (PASA)*, pages 13–32, 1996.
- [9] T. Erlebach and K. Jansen. Call scheduling in trees, rings and meshes. In *HICSS: Hawaii International Conference on System Sciences*. IEEE Computer Society Press, 1997.
- [10] L. Gargano, P. Hell, and S. Pérennes. Colouring paths in directed symmetric trees with applications to WDM routing. In *Automata, Languages and Programming, 24th International Colloquium (ICALP)*, volume 1256 of *Lecture Notes in Computer Science*, pages 505–515, Bologna, Italy, July 1997. Springer-Verlag.
- [11] P. E. Green. *Fiber Optic Networks*. Prentice Hall, 1993.
- [12] M-C. Heydemann, J-C. Meyer, and D. Sotteau. On forwarding indices of networks. *Discrete Applied Mathematics*, 23:103–123, 1989.

-
- [13] Y. Manoussakis and Z. Tuza. The forwarding index of directed networks. *Discrete Applied Mathematics*, 68:279–291, 1996.
 - [14] D. Minoli. *Telecommunications Technology Handbook*. Artech House, 1991.
 - [15] B. Mukherjee. WDM-based local lightwave networks, Part I: Single-hop systems. *IEEE Network Magazine*, 6(3):12–27, May 1992.
 - [16] R. K. Pankaj. *Architectures for Linear Lightwave Networks*. PhD thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1992.
 - [17] A. Pelc. Fault-tolerant broadcasting and gossiping in communication networks. *Networks*, 28, 1996.
 - [18] P. Raghavan and E. Upfal. Efficient routing in all-optical networks. In *Proc. of the 26th Annual ACM Symposium on Theory of Computing (STOC)*, pages 134–143, May 1994.
 - [19] H. Schröder, O. Sýkora, and I. Vrt'o. Optical all-to-all communication for some product graphs. In *Proc. of 24th Seminar on Current Trends in Theory and Practice of Informatics (SOFSEM)*, 1997.
 - [20] P. Solé. Expanding and forwarding. *Discrete Applied Mathematics*, 58:67–78, 1995.

ANNEXE C

All-to-all routing and coloring in weighted trees of rings

All-to-All Routing and Coloring in Weighted Trees of Rings

Bruno Beauquier^{*}, Stéphane Pérennes^{*}, David Tóth^{**}

Thème 1 — Réseaux et systèmes
Projet Sloop

Juin 1999

Abstract: A *tree of rings* is an undirected graph obtained from the union of rings, which intersect two by two in at most one node, such that any two nodes are connected by exactly two edge-disjoint paths. In this paper, we consider symmetric directed trees of rings with weighted nodes. A *routing* for a weighted digraph is a collection of directed paths (dipaths), such that for each ordered pair of nodes (x_1, x_2) with respective weights w_1 and w_2 , there are $w_1 w_2$ dipaths (possibly not distinct) from x_1 to x_2 . Motivated by the *Wavelength Division Multiplexing* (WDM) technology in all-optical networks, we study the problem of finding a routing which can be colored by the fewest number of colors so that dipaths of the same color are arc-disjoint. We prove that this minimum number of colors (wavelengths) is equal to the maximum number of dipaths that share one arc (load), minimized over all routings. The problem can be efficiently solved (dipaths found and colored) using cut properties.

Key-words: optical networks, WDM, routing, coloring, cut.

(Résumé : *tsvp*)

Accepted in the 11th ACM Symposium on Parallel Algorithms and Architectures (SPAA'99).

This work has been partially supported by the AFIRST in the framework of the French-Israeli project *Communication Algorithms in Optical Networks* and by the French RNRT project *PORTO*.

^{*} SLOOP is a joint project with the CNRS and the University of Nice-Sophia Antipolis (I3S laboratory).
E-mail: {Bruno.Beauquier,Stephane.Perennes}@inria.fr.

^{**} Technical University of Budapest, research done while visiting INRIA at Sophia Antipolis. E-mail: tocsa@math.bme.hu.

Routage Optique WDM dans les arbres d'anneaux pondérés

Résumé : Un *arbre d'anneaux* est un graphe non-orienté obtenu par l'union d'anneaux, qui s'intersectent deux-à-deux en au plus un sommet, de telle sorte que deux sommets sont reliés par exactement deux chemins arête-disjoints. Dans ce rapport, nous considérons des arbres d'anneaux orientés symétriques avec des sommets pondérés. Un *routage* pour un graphe orienté pondéré est une collection de chemins orientés, telle que pour tout couple de sommets (x_1, x_2) de poids respectifs w_1 et w_2 , elle contient $w_1 w_2$ chemins (éventuellement distincts) allant de x_1 vers x_2 . Motivés par la technologie de multiplexage en longueurs d'onde WDM (*Wavelength-Division Multiplexing*) dans les réseaux tout-optiques, nous étudions le problème de trouver un routage et une coloration minimale des chemins, avec la contrainte que deux chemins de la même couleur sont arc-disjoints. Nous montrons que le nombre minimum de couleurs (longueurs d'onde) pour une solution est égal au nombre maximum de chemins qui partagent un arc (charge), parmi tous les routages possibles. Le problème est résolu par une méthode constructive et efficace, en utilisant des critères de coupe.

Mots-clé : réseaux optiques, WDM, routage, coloration, coupe.

1 Introduction and Motivation

Optics is emerging as a key technology in communication networks, promising very high speed local or wide area networks of the future. A single optical wavelength supports rates of gigabits per second, which in turn support multiple channels of voice, data and video [10, 13]. Multiple laser beams that are propagated over the same fiber on distinct optical wavelengths can increase this capacity even further. This is achieved through *Wavelength Division Multiplexing* (or *WDM*) [6], by partitioning the optical bandwidth into a number of channels and allowing the transmission of multiple data streams concurrently along the same optical fiber on different channels, i.e., different wavelengths. (Two streams of the same wavelength cannot be routed on the same fiber due to electromagnetic interference.)

All-optical (or *single-hop* [14]) communication networks provide all source-destination pairs with end-to-end transparent channels that are identified through a wavelength and a physical path. They exploit photonic technology for the implementation of both the switching and the transmission functions [9], and maintain the signal in optical form through the transmission, thus allowing much higher data transmission rates (since there is no prohibitive overhead due to conversions to and from the electronic form). Wavelengths being a limited resource, solutions to the problem of efficient routing and wavelengths allocation are of importance for the future development of optical technology.

The problem we consider here is motivated by *switched* networks with reconfigurable wavelength selective optical switches, without wavelength converters, where different signals may travel on the same communication link (but on different wavelengths) into a node, and then exit from it on different links, keeping their original wavelengths. Since optical links are inherently bidirectional, the problem has initially been modeled by undirected graphs, as in [1, 2, 15]. However, it has since become clear that the use of amplifiers will actually make each bidirectional optical link into a pair of unidirectional links [16], and hence the new models of the situation tend to represent optical networks by symmetric directed graphs, as in [5, 12, 11].

In an all-optical network one needs to set up a number of communications (dipaths) between given pairs of nodes, with each communication being transmitted on one particular wavelength, and so that all communications that share a link have different wavelengths. Specifically, the general problem is the following : Given a collection of *requests* $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$, connect each x_i to the corresponding y_i by a dipath P_i , and assign a wavelength to each P_i , so that dipaths with the same wavelength do not share a link. The objective is to use the minimum possible number of wavelengths. This parameter is considered of importance in evaluating the competitiveness of the wavelength division multiplexing technology [12]. A survey of graph theoretic problems associated with routing in optical networks can be found in [4].

The problem of minimizing the number of wavelengths has been proved to be NP-complete for a general collection of requests, even in rings [7]. Here we consider the *All-To-All* communication process in the weighted case : for each ordered pair of nodes (x, y) with respective weights $w(x)$ and $w(y)$, $w(x)w(y)$ dipaths (possibly not distinct) from x to y have to be found and assigned wavelengths. Optimal results for this problem have been proved in

the unweighted case for rings [5, 18], for some cartesian products [3] including hypercubes and some toroidal meshes, for some compound graphs [17] and for trees [8].

In this paper we solve the case of the weighted symmetric directed trees of rings. These networks are interesting since tree-like structures are often used in the telecommunications industry [12] and emerging SDH (Synchronous Digital Hierarchy) local area networks generally consist of SONET (Synchronous Optical NETWORK) rings connected together.

2 Preliminaries

In this paper, we model an all-optical network as a *weighted symmetric digraph* (G, w) , that is a directed graph G , with node set $V(G)$ and arc set $A(G)$ (such that if $(u, v) \in A(G)$ then $(v, u) \in A(G)$), together with a non negative integral weight function w on the set of nodes. We always denote by N the total weight $\sum_{v \in V(G)} w(v)$ of (G, w) , and for any subset of nodes $S \subset V(G)$ we define the weight of S as $w(S) = \sum_{s \in S} w(s)$.

Definition 1 *A tree of rings is an undirected graph obtained from the union of rings, which intersect two by two in at most one node, such that any two nodes are connected by exactly two edge-disjoint paths.*

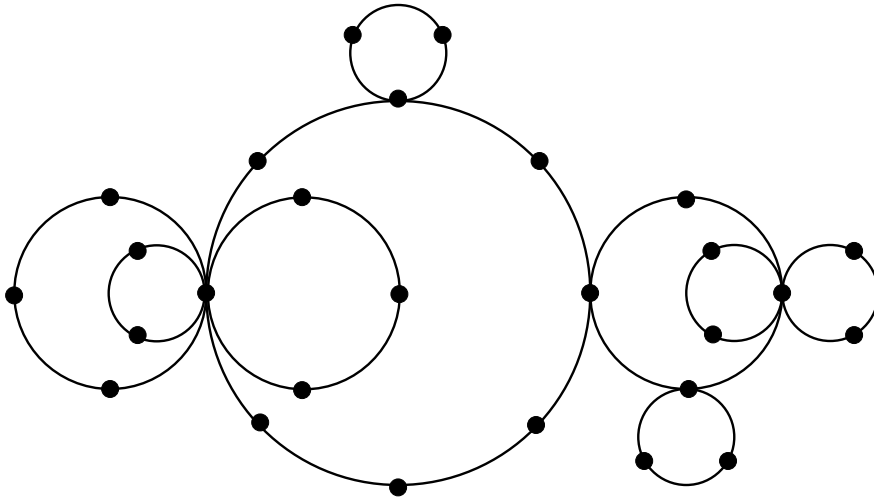


Figure 1: A tree of rings.

A symmetric digraph is naturally induced by an undirected graph by replacing each edge by two opposite arcs. In this paper, we consider weighted symmetric directed trees of rings.

Definition 2 A routing for a weighted digraph (G, w) is a collection of directed paths (dipaths) in G , such that for each ordered pair of nodes (x, y) there are $w(x)w(y)$ dipaths (possibly not distinct) from x to y .

The *wavelength-routing problem* for a weighted digraph (G, w) consists of finding a routing and assigning each dipath a wavelength (or color), so that no two dipaths sharing an arc have the same wavelength. We denote by $\vec{w}(G, w)$ the minimum number of wavelengths used to solve the wavelength-routing problem in (G, w) .

Given a weighted digraph (G, w) and a routing R , the *load* of an arc $\alpha \in A(G)$ is the number of dipaths in R containing α . The *load* of (G, w) , denoted by $\vec{\pi}(G, w)$, is the minimum over all routings R of the maximum load of the arcs of G . Clearly, $\vec{\pi}(G, w)$ is a lower bound on $\vec{w}(G, w)$.

Property 3 $\vec{w}(G, w) \geq \vec{\pi}(G, w)$ for any weighted digraph (G, w) .

Given a weighted digraph (G, w) and any partition (S, \bar{S}) of the set of nodes $V(G)$, called a *cut*, we denote by $c(S, \bar{S})$ the number of arcs beginning in S and ending in \bar{S} .

At least $w(S)w(\bar{S})$ dipaths in any routing for (G, w) must use these arcs. So we have $\vec{\pi}(G, w) \geq \frac{w(S)w(\bar{S})}{c(S, \bar{S})}$ for any cut (S, \bar{S}) .

As far as a weighted tree of rings (\mathcal{T}, w) is concerned, we consider only cuts (S, \bar{S}) such that $c(S, \bar{S}) = 2$. Therefore S and \bar{S} are two connected components of \mathcal{T} . Setting $M = w(S)$ and $m = w(\bar{S})$, we have $M + m = N$ and we assume that $M \geq m$ in the sequel. The product Mm is called the *load of the cut* and the difference $\delta = M - m$ is called the *gap of the cut*. A *tight cut* is a cut with maximum load (or equivalently with minimum gap).

Given two nodes a and b of (\mathcal{T}, w) with positive weights, we denote by $w_{a,b}$ the weight function obtained from w by decreasing the weights of nodes a and b by one unit (hence $(\mathcal{T}, w_{a,b})$ has total weight $N - 2$). The minimum gap of $(\mathcal{T}, w_{a,b})$ is denoted by $\delta_{a,b}$. The following property will be useful :

Property 4 Let (\mathcal{T}, w) be a weighted tree of rings with minimum gap $\delta \geq 2$. Let a and b be two nodes in $V(\mathcal{T})$ with positive weights. If there exists a tight cut (S, \bar{S}) of (\mathcal{T}, w) such that $a, b \in S$ then $\delta_{a,b} = \delta - 2$, otherwise $\delta_{a,b} \geq \delta$.

Proof. Consider any cut (S, \bar{S}) of \mathcal{T} . If (S, \bar{S}) is a tight cut of (\mathcal{T}, w) , then (S, \bar{S}) has gap $\delta - 2 \geq 0$ in $(\mathcal{T}, w_{a,b})$ if $a, b \in S$, and gap at least δ in $(\mathcal{T}, w_{a,b})$ if not. If (S, \bar{S}) is not a tight cut of (\mathcal{T}, w) , then it has gap at least $\delta + 2$ in (\mathcal{T}, w) and gap at least δ in $(\mathcal{T}, w_{a,b})$. \square

Half the load of a tight cut is clearly a lower bound on both $\vec{\pi}(\mathcal{T}, w)$ and $\vec{w}(\mathcal{T}, w)$. In this paper, we show that this bound is indeed exact :

Main Theorem For any weighted tree of rings (\mathcal{T}, w) we have $\vec{w}(\mathcal{T}, w) = \vec{\pi}(\mathcal{T}, w) = \lceil \frac{Mm}{2} \rceil$, where Mm is the load of a tight cut.

The proof is done by induction on the total weight. We first prove the result in the case of weighted rings and then proceed with the general case.

3 Weighted rings

In this section we prove the theorem in the case of weighted rings. The proof for arbitrary weighted trees of rings will follow the same lines.

Due to the upper integral part appearing in Main Theorem, we need to introduce the notion of *half-colors*. Any color assigned to a set of arc-disjoint dipaths can be seen as the union of two opposite half-colors : a positive one for clockwise dipaths and a negative one for anti-clockwise dipaths. Conversely, any two opposite half-colors can give rise to one color. Therefore, half-coloring a routing with h^+ positive half-colors and h^- negative half-colors leads to a coloring with $\max\{h^+, h^-\}$ colors. Half-colors for a routing are said to be *balanced* if $|h^+ - h^-| \leq 1$. We will manage to have $h^+ = \lceil \frac{Mm}{2} \rceil$ and $h^- = \lfloor \frac{Mm}{2} \rfloor$, so that we can get a solution to the wavelength-routing problem using a total of $\max\{h^+, h^-\} = \lceil \frac{Mm}{2} \rceil$ colors. The following proposition implies Main Theorem for weighted rings.

Proposition 5 *Let (C, w) be a weighted ring and Mm be the load of a tight cut. There exists a routing that can be colored with $h^+ = \lceil \frac{Mm}{2} \rceil$ and $h^- = \lfloor \frac{Mm}{2} \rfloor$ half-colors.*

Our proof is by induction on the total weight N : a coloring for (C, w) is obtained from a coloring of the weighted ring $(C, w_{a,b})$ for some two nodes a and b .

Definition 6 *Let (C, w) be a weighted ring with total weight N . Given two nodes a and b with positive weights, let A and B be the two connected sets of nodes of $C - \{a, b\}$. Then nodes a and b are said to be antipodal if $w(A) \leq \lfloor N/2 \rfloor$ and $w(B) \leq \lfloor N/2 \rfloor$.*

Intuitively, antipodal nodes can be seen in the following way. Assume that the nodes of the weighted ring are labelled with disjoint intervals of integers modulo N , so that each node a has an interval $I(a)$ of length equal to $w(a)$ and adjacent nodes have adjacent intervals. Then, any two nodes a and b are antipodal if there are $x \in I(a)$ and $y \in I(b)$, such that $y = x + \lfloor \frac{N}{2} \rfloor \pmod N$ or $y = x + \lceil \frac{N}{2} \rceil \pmod N$.

Case A

Keeping the same notation as in Property 4, we assume that :

Condition A *There exist two antipodal nodes a and b such that $\delta_{a,b} \geq \delta$.*

Note that Condition A is fulfilled if $\delta \leq 1$, or if there exist two antipodal nodes a and b , such that for any tight cut (S, \bar{S}) , $a \in S$ if and only if $b \in \bar{S}$.

Lemma 7 *If Proposition 5 holds for all weighted rings with weight less than N , then it holds for any weighted ring with weight N satisfying Condition A.*

Proof. Let (C, w) be a weighted ring with weight N satisfying Condition A. (The result is obviously true if $N \leq 1$.) According to Condition A, there are two nodes a and b such that the load of a tight cut of $(C, w_{a,b})$ is at most $(M-1)(m-1) = Mm - (m+M) + 1$. Using the inductive hypothesis, we color a routing of $(C, w_{a,b})$ with $Mm - (N-1)$ balanced

half-colors. Let A and B be the two connected sets of nodes of $C - \{a, b\}$. To obtain a colored routing for (C, w) satisfying the lemma, it remains to assign in a balanced way $N - 1$ half-colors to a set D_1 of $4(w(A) + w(B))$ dipaths connecting a and b with the nodes in A and B and to a set D_2 of $2(w(a) + w(b) - 1)$ dipaths connecting a with b .

Given $a' \in A \cup \{a, b\}$ and $b' \in B \cup \{a, b\}$, two half-colors (one in each direction) are sufficient to color eight dipaths (possibly less if $a' \in \{a, b\}$ or $b' \in \{a, b\}$) connecting a and b with a' and b' (see Figure 2).

If N is even, as a and b are antipodal, we can choose $\frac{N-2}{2}$ times $a' \in A \cup \{a, b\}$ and $b' \in B \cup \{a, b\}$ and use $N - 2$ half-colors, so that sets D_1 and D_2 , but two dipaths connecting a with b , are found and colored. At last, one half-color (whose direction can be chosen to ensure the balance condition) is used in this case for the two last dipaths connecting a with b .

In the other case, N is odd and we can assume w.l.o.g. that $w(A) \leq \lfloor N/2 \rfloor$ and $w(B) \leq \frac{N-3}{2}$. By choosing $\frac{N-3}{2}$ times $a' \in A \cup \{a, b\}$ and $b' \in B \cup \{a, b\}$ and using $N - 3$ half-colors, we can define and color set D_1 , but four dipaths connecting a and b with one node in A , and set D_2 , but two dipaths connecting a with b . These last six dipaths can be defined and colored with two more half-colors. From the inductive hypothesis, we can also ensure that the same number of half-colors is used in the two directions of the ring (N is odd and Mm even).

In both cases, we get a colored routing of (C, w) with $h^+ + h^- = Mm$ half-colors, so that $h^+ = \lceil \frac{Mm}{2} \rceil$ and $h^- = \lfloor \frac{Mm}{2} \rfloor$. \square

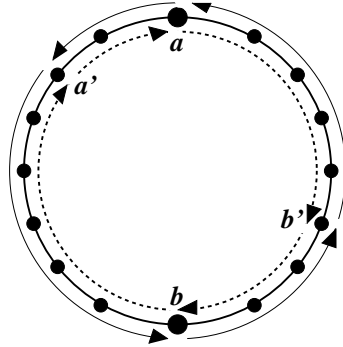


Figure 2: Coloring eight dipaths between $\{a, b\}$ and $\{a', b'\}$ with two half-colors.

As proved in [5, 18], we deduce the following result :

Corollary 8 *Proposition 5 holds for any weighted ring (C, w_1) with uniform weight equal to 1.*

Proof. The minimum gap of (C, w_1) is at most one, so we are in Case A and the result follows inductively from the proof of Lemma 7. \square

Corollary 9 *Proposition 5 holds for any weighted ring (C, w_k) with uniform weight equal to k .*

Proof. Let p be the number of nodes of C . The load of a tight cut of (C, w_k) is k^2 times the load of a tight cut of (C, w_1) equal to $\lceil p/2 \rceil \lfloor p/2 \rfloor$. If p is odd or if p and $p/2$ are even, we can get a solution for (C, w_k) by taking k^2 times the half-colored routing for (C, w_1) , where $h^+ = h^-$. If p is even and $p/2$ odd, we have $h^+ = h^- + 1$ in the half-colored routing for (C, w_1) . To ensure the balance condition for (C, w_k) , we take $\lceil \frac{k^2}{2} \rceil$ times the half-colored routing for (C, w_1) and $\lfloor \frac{k^2}{2} \rfloor$ times the routing for (C, w_1) where clockwise and anti-clockwise half-colors are swapped. \square

Case B

Let (C, w) be a weighted ring with minimum gap δ . We assume that we are not in Case A, so $\delta > 1$ and the following condition is satisfied :

Condition B *For any two antipodal nodes a and b , there exists a tight cut (S, \bar{S}) such that $a \in S$ and $b \in \bar{S}$.*

Lemma 10 *Any weighted ring satisfying Condition B has uniform weight and an odd number of nodes.*

Proof. First notice that for any connected set X of nodes, $w(X) < M$ implies $w(X) \leq m$ and $w(X) > m$ implies $w(X) \geq M$. Note also that any node is antipodal to a least one node.

Let a, b and d be three nodes, such that b and d are adjacent. Let B and D be the two connected sets of nodes of $C - \{a, b, d\}$, between a and b , and between a and d , respectively.

If a and b are antipodal and if (S, \bar{S}) is a tight cut such that $a \in S$ and $b \in \bar{S}$, then $S = \{a\} \cup B \cup \{b\}$. Indeed, otherwise we have $w(a) + w(B) + w(b) < M$, so $w(a) + w(B) + w(b) \leq m = N - M$ and $w(a) + w(B) + w(b) < N/2$, but as a and b are antipodal, we have also $w(d) + w(D) \leq \lfloor N/2 \rfloor$ (contradiction).

Condition B implies that every node has weight at least δ . Indeed, assuming a and b antipodal, we have $w(a) + w(B) + w(b) = M = m + \delta$ and $w(B) + w(b) < M$, so $w(B) + w(b) \leq m$ and $w(a) \geq \delta$. Note also that if node a is antipodal to itself, then $w(a) > N/2$, there is only one tight cut and we are clearly in Case A.

Assume that both nodes b and d are antipodal with node a . Then $w(a) + w(B) + w(b) = M$ and $w(a) + w(D) + w(d) = M$. So $N + w(a) = 2M = M + m + \delta = N + \delta$, hence $w(a) = \delta$.

Assume that b is the only antipodal node with a . Then $w(a) + w(B) + w(b) = M$ and $w(B) + w(b) \leq m \leq N/2$. As a and d are not antipodal and since $w(D) < m \leq N/2$, we have $w(B) + w(b) > \lfloor N/2 \rfloor$ (contradiction).

Therefore, Condition B implies that every node has at least two antipodal nodes and weight δ . The number of nodes cannot be even, otherwise $\delta = 0$ (Case A). So, in Case B the weight is uniform and the number of nodes is odd. \square

Proof of Proposition 5

The proof is made by induction. Let (C, w) be a weighted ring with weight N . If $N \leq 1$ then the result is trivial. Assume that Proposition 5 holds for all weighted rings with weight less than N . If (C, w) satisfies Condition A, then we use Lemma 7, else (C, w) has uniform weight and a number of nodes odd, according to Lemma 10, we use then Corollary 9 to conclude.

4 General case

Now we extend Proposition 5 (and thus Main Theorem) to the case of weighted trees of rings.

Proposition 11 *Let (T, w) be a weighted tree of rings and Mm be the load of a tight cut. Then there exists a routing which can be colored with $h^+ + h^- = Mm$ half-colors, with $h^+ = \lceil \frac{Mm}{2} \rceil$ and $h^- = \lfloor \frac{Mm}{2} \rfloor$.*

Definition 12 *Let (T, w) be a weighted tree of rings and P a closed Eulerian dipath in T visiting each node. The interval labelling I for P (see Figure 2) is the assignment of an interval $I(a)$ of integers in $\{1, \dots, N\}$ to each node a of T , such that :*

- *distinct nodes are assigned disjoint intervals and each node a is assigned an interval of length $w(a)$;*
- *for any two integers i and j in $\{1, \dots, N\}$, $i < j$, if $i \in I(a)$ and $j \in I(b)$ for some nodes a and b , then $a = b$ or a is visited for the first time before b along P .*

Definition 13 *Let (T, w) be a weighted tree of rings with total weight N , let P be a closed Eulerian dipath in T and let I be the interval labelling for P . Two nodes a and b are said to be antipodal if there exist $i \in I(a)$ and $j \in I(b)$ such that $j = i + \lfloor \frac{N}{2} \rfloor \bmod N$ or $j = i + \lceil \frac{N}{2} \rceil \bmod N$.*

This definition of antipodal nodes of weighted trees of rings is equivalent to that given previously for weighted rings. Also, half-colors can be defined for trees of rings. Note that a closed Eulerian (and Hamilton) dipath P for a weighted tree of rings (T, w) induces an orientation of each ring of T . Indeed, every arc used by P can be defined *clockwise* and every other arc can be defined *anti-clockwise*. We can then call *half-color* a color used for routing a set of arc-disjoint dipaths using exclusively clockwise or anti-clockwise arcs. Nevertheless, we will need a slightly different definition of half-colors in Case B. Indeed, to achieve Main Theorem it is only required that half-colors can be combined two by two to obtain colors. This will be the case in the proof of Lemmas 17 and 18 where each half-color will be used for routing a set of arc-disjoint dipaths using *in each ring* exclusively clockwise or anti-clockwise arcs.

As for weighted rings, we first assume that Condition A is satisfied (Case A).

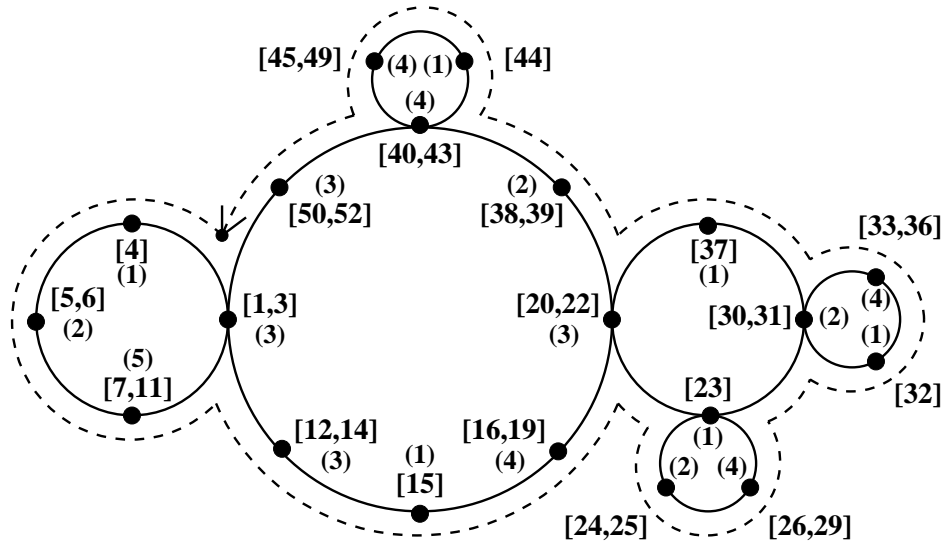


Figure 3: Interval labelling of a weighted tree of rings (weights are in brackets) induced by a closed Eulerian dipath.

Lemma 14 *If Proposition 11 holds for all weighted trees of rings with weight less than N , then it holds for any weighted tree of rings with weight N satisfying Condition A.*

Proof. The same arguments as in the proof of Lemma 7 are used and we give here a short proof. Given a weighted tree of rings (\mathcal{T}, w) , let P be a closed Eulerian dipath in \mathcal{T} . We consider two antipodal nodes a and b and we can assume w.l.o.g. that a is visited for the first time before b along P . Let A be the set of nodes visited after a and before b and let B be the set of nodes visited before a or after b . By defining A and B in this manner and by routing all requests clockwise or anti-clockwise along P , the proof of Lemma 7 can be adapted and leads to the same result. \square

Let (\mathcal{T}, w) be a weighted tree of rings with minimum gap δ . If Condition A is not satisfied, then $\delta > 1$ and Condition B holds (Case B).

Definition 15 *A main ring of (\mathcal{T}, w) is any ring where some tight cut occurs. A block of a main ring \mathcal{C} is any set of nodes remaining connected after removing all the arcs of \mathcal{C} .*

Whereas for weighted rings Condition B implies an odd number of nodes of the same weight, for weighted trees of rings we need to add the condition that every block of a main ring has weight at most $N/2$ (Case B1).

Lemma 16 Consider a weighted tree of rings satisfying Condition B and having a main ring with blocks all of weight at most $N/2$. Then there is an odd number of blocks all of the same weight.

Sketch of proof. See the proof of Lemma 10. (Blocks of a tree of rings correspond to nodes of a ring and antipodal blocks are defined as containing antipodal nodes.) \square

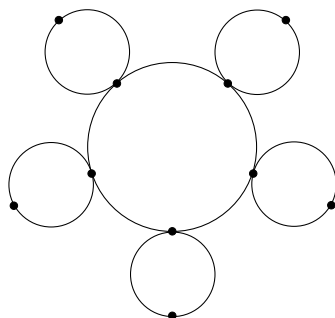


Figure 4: A weighted tree of rings in Case B1. All weights are equal to 1.

The remaining case to study is that of a weighted tree of rings satisfying Condition B and having a main ring with one block of weight greater than $N/2$ (Case B2). So there is exactly one tight cut in the main ring separating this block from the other blocks. Moreover Condition B implies the existence of another tight cut out of the main ring, as two antipodal nodes can be chosen respectively in and out of the main ring. So we can see the weighted tree of rings as the union of weighted trees of rings (at least two of them having the same maximum weight) which intersect two by two in the same weighted node.

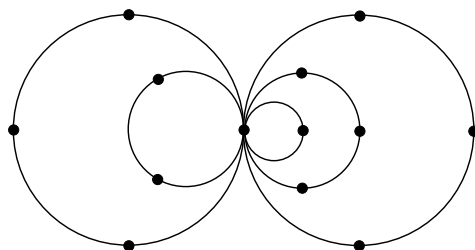


Figure 5: A weighted tree of rings in Case B2. All weights are equal to 1.

In order to complete the proof of Main Theorem we need to solve Case B1 and Case B2.

Lemma 17 Proposition 11 holds for any weighted tree of rings (\mathcal{T}, w) with $2p + 1$ blocks ($p > 0$) of weight k .

Proof. The load of a tight cut is $p(p+1)k^2$. The half-coloring of (\mathcal{T}, w) follows from the solution obtained for the ring (C_{2p+1}, w_1) with $2p+1$ nodes and uniform unit weight, that uses $p(p+1)$ half-colors.

We number blocks of \mathcal{T} with integers from 1 to $2p+1$ like the nodes of C_{2p+1} and for each integer i we take a closed Eulerian dipath P_i of block number i and its interval labelling I_i . Moreover, for $1 \leq i \leq 2p+1$ and $1 \leq x \leq k$, we denote by $l_i(x)$ the node of block number i such that $x \in I_i(l_i(x))$. We establish a correspondence between each half-color used in a solution for (C_{2p+1}, w_1) and a set of k^2 half-colors for (\mathcal{T}, w) . Every dipath in \mathcal{T} will be routed in every block along its Eulerian dipath.

Assume that c is a half-color used in (C_{2p+1}, w_1) for clockwise dipaths. For each pair of integers x and y such that $1 \leq x, y \leq k$, we use in (\mathcal{T}, w) a half-color to color dipaths 1) from node $l_i(x)$ to node $l_i(y)$ for each i , and 2) from node $l_i(y)$ to node $l_j(x)$ if there is in (C_{2p+1}, w_1) a dipath from node i to node j colored by c . Note that these dipaths are defined clockwise in the main ring, clockwise in the blocks if $x \leq y$ and anti-clockwise in the blocks if $y < x$.

Similarly, if c is used in (C_{2p+1}, w_1) for anti-clockwise dipaths then for $1 \leq x, y \leq k$ we use in (\mathcal{T}, w) a half-color to color dipaths 1) from node $l_i(y)$ to node $l_i(x)$ for each i , and 2) from node $l_i(x)$ to node $l_j(y)$ if there is in (C_{2p+1}, w_1) a dipath from node i to node j colored by c . Note that these dipaths are defined anti-clockwise in the main ring, anti-clockwise in the blocks if $x \leq y$ and clockwise in the blocks if $y < x$.

This symmetric assignment ensures that half-colors can be combined two by two to obtain colors, so Proposition 11 holds in Case B1. \square

Lemma 18 *Proposition 11 holds for any weighted tree of rings (\mathcal{T}, w) that is the union of weighted trees of rings (at least two of them having the same maximum weight) which intersect two by two in the same weighted node.*

Sketch of proof. Let v be the common weighted node of the weighted subtrees of rings whose union forms (\mathcal{T}, w) . Note that v is at the border of every tight cut. In each subtree of rings we take a closed Eulerian dipath and every dipath in \mathcal{T} will be routed in every subtree of rings along its Eulerian dipath. A good routing strategy is to combine the different routings for each subtree of rings obtained by adding to the weight of node v the weights of the other subtrees of rings (Condition A is then satisfied in each so modified subtree of rings). The main problem consists in coloring the set of all dipaths going to or from or through v . This is obviously equivalent to color a set of dipaths in a subdivided star, i.e. in a graph that is the union of paths (twice more than subtrees of rings) which intersect two by two in the same node. Theorem 21 of [4] states the existence of a coloring in this case with the right number of colors. The remaining dipaths go neither to neither from nor through v and can be colored more easily. \square

References

- [1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing and scheduling algorithms for optical networks. In *Proc. of the 5th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'94)*, pages 412–423. ACM Press, January 1994.
- [2] Y. Aumann and Y. Rabani. Improved bounds for all optical routing. In *Proc. of the 6th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'95)*, pages 567–576. ACM Press, January 1995.
- [3] B. Beauquier. All-to-all communication in some wavelength-routed all-optical networks. *Networks (Wiley-Interscience)*, 33(3):179–187, May 1999.
- [4] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes, and U. Vaccaro. Graph problems arising from wavelength-routing in all-optical networks. In *Proc. of 2nd IEEE Workshop on Optics and Computer Science (WOCS, part of IPPS'97)*. IEEE Press, April 1997.
- [5] J.-C. Bermond, L. Gargano, S. Pérennes, A.A. Rescigno, and U. Vaccaro. Efficient collective communication in optical networks. In *Proc. of 23rd Int. Colloq. on Automata, Languages and Programming (ICALP'96)*, volume 1099 of *Lecture Notes in Computer Science*, pages 574–585. Springer-Verlag, July 1996.
- [6] N. K. Cheung, Nosu K., and G. Winzer. Special issue on dense WDM networks. *Journal on Selected Areas in Communications*, 8, 1990.
- [7] T. Erlebach and K Jansen. Scheduling of virtual connections in fast networks. In *Proc. of Parallel Systems and Algorithms (PASA)*, pages 13–32, 1996.
- [8] L. Gargano, P. Hell, and S. Pérennes. Colouring paths in directed symmetric trees with applications to WDM routing. In *Proc. of 24th Int. Colloq. on Automata, Languages and Programming (ICALP'97)*, volume 1256 of *Lecture Notes in Computer Science*, pages 505–515. Springer-Verlag, July 1997.
- [9] P. E. Green. The future of fiber-optic computer networks. *IEEE Computer*, 24(9):78–87, September 1991.
- [10] P. E. Green. *Fiber-Optic Communication Networks*. Prentice-Hall, 1993.
- [11] C. Kaklamanis and P. Persiano. Efficient wavelength routing on directed fiber trees. In *Proc. of European Symp. on Algorithms (ESA'96)*, volume 1136 of *Lecture Notes in Computer Science*, pages 460–470, 1996.
- [12] M. Mihail, C. Kaklamanis, and S. Rao. Efficient access to optical bandwidth—wavelength routing on directed fiber trees, rings, and trees of rings. In *Proc. of 36th Ann. Symp. on Foundations of Computer Science (FOCS'95)*, pages 548–557. IEEE, October 1995.

- [13] D. Minoli. *Telecommunications Technology Handbook*. Artech House, 1991.
- [14] B. Mukherjee. WDM-based local lightwave networks, Part 1: Single-hop systems. *IEEE Network Magazine*, 6(3):12–27, May 1992.
- [15] P. Raghavan and E. Upfal. Efficient routing in all-optical networks. In *Proc. of the 26th Ann. ACM Symp. on Theory of Computing (STOC'94)*, pages 134–143, May 1994.
- [16] R. Ramaswami. Multiwavelength lightwave networks for computer communication. *IEEE Communications Magazine*, 31(2):78–88, February 1993.
- [17] O. Togni. Optical all-to-all communication in inflated networks. In *Proc. of the 5th Ann. Int. Workshop on Graph-Theoretic Concepts in Computer Science (WG'98)*, volume 1517 of *Lecture Notes in Computer Science*, pages 78–87. Springer-Verlag, 1998.
- [18] G. Wilfong. Minimizing wavelengths in an all-optical ring network. In *Proc. of the 7th Int. Symp. on Algorithms and Computation (ISAAC'96)*, volume 1178 of *Lecture Notes in Computer Science*, pages 346–355. Springer-Verlag, 1996.

ANNEXE D

Tight bounds for broadcasting in the linear cost model

Tight Bounds for Broadcasting in the Linear Cost Model

Bruno Beauquier , Olivier Delmas , Stéphane Pérennes

Thème 1 — Réseaux et systèmes

Projet Sloop

Novembre 1999

Abstract: This work considers broadcast protocols made of successive communication rounds in the *linear cost* model: the time needed to send a message of length L is defined as $\alpha + L\tau$. In this model, the *communication time* of any algorithm \mathcal{A} is expressed as the sum $R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, where $R_{\mathcal{A}}$ is the *number of rounds* and $T_{\mathcal{A}}$ the *transmission cost* of the algorithm. In order to design an efficient algorithm realizing a given communication pattern, it appears that minimizing $R_{\mathcal{A}}$ and $T_{\mathcal{A}}$ are antinomic goals. We study this trade-off issue for broadcast protocols. Surprisingly, such a general theoretical study has almost never been done. In the literature, only the two opposite issues are actually considered: minimizing the number of rounds in the case of short messages, or minimizing the transmission cost in the case of large messages. Our results concern the fully-connected N -nodes network K_N , with $N = (k + 1)^T$, in the bidirectional k -ports mode. We derive tight bounds on the communication time for broadcasting in $T + r$ rounds, our lower bounds holding for any network.

Key-words: communication networks, broadcasting, linear cost.

(Résumé : *tsvp*)

An extended abstract of this paper written by the second author has been accepted in the *5th International Colloquium on Structural Information and Communication Complexity (SIROCCO'98)*.

This work has been partially supported by the AFIRST in the framework of the French-Israeli project *Communication Algorithms in Optical Networks* and by the French RNRT project *PORTO*.

Bornes Fines pour la Diffusion avec Coût Linéaire

Résumé : Ce travail considère des protocoles de diffusion se déroulant en *étapes* successives sous le modèle *linéaire* du coût de communication : le temps nécessaire pour envoyer un message de longueur L est défini par $\alpha + L\tau$. Dans ce modèle, le *temps de communication* d'un algorithme s'exprime par la somme $R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, où $R_{\mathcal{A}}$ est le *nombre d'étapes* et $T_{\mathcal{A}}$ le *coût de transmission* de l'algorithme. En vue de concevoir un algorithme efficace réalisant un schéma de communication donné, il apparaît que minimiser $R_{\mathcal{A}}$ et $T_{\mathcal{A}}$ sont des objectifs antinomiques. Nous étudions le compromis nécessaire entre ces deux coûts dans le cadre de la diffusion. Il s'avère qu'une telle étude théorique générale n'a pratiquement jamais été entreprise. Dans la littérature, seulement les deux cas extrêmes sont considérés : minimisation du nombre d'étapes dans le cas de messages courts, ou minimisation du coût de transmission dans le cas de messages longs. Nos résultats concernent le réseau complet K_N , avec $N = (k + 1)^T$, sous le mode k -ports bidirectionnel. Nous dérivons des bornes fines sur le temps de communication pour la diffusion en $T + r$ étapes, les bornes inférieures étant valables dans tout réseau.

Mots-clé : réseaux de communication, diffusion, coût linéaire.

1 Introduction

When studying information dissemination in communication networks, often considered are protocols made of successive communication *rounds* with a *linear* model for the communication *cost*: the time needed to send a message of size L from a node to another is defined as $\alpha + L\tau$, where α stands for a start-up time while $L\tau$ represents the cost of sending L bits of data in a channel with bandwidth $1/\tau$. The *cost of a round* is defined as the maximum of the cost of all the communications done during this round. In such a modelization, the **communication time** of any algorithm \mathcal{A} can be expressed as the sum $R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, where $R_{\mathcal{A}}$ is the **number of rounds** and $T_{\mathcal{A}}$ the **transmission cost** of the algorithm. This linear model of the communication cost is often used in the literature and references to algorithmic work can be found in [6, 7]. When trying to design an optimal algorithm performing a given communication pattern, it appears that minimizing $R_{\mathcal{A}}$ and $T_{\mathcal{A}}$ are often *antinomic* goals. In fact, in the literature many authors optimize only one of these two parameters by considering one of the two extremal cases:

- *Short messages (i.e. $\alpha \gg L\tau$)* – In this case, $R_{\mathcal{A}} \cdot \alpha$ is the leading term of the communication time. Hence, only the number of rounds is to be minimized. Many authors have proposed broadcast algorithms optimal with that respect (see the surveys [10, 11], the book [14], the chapter [13] or the theses [6, 7]).
- *Large messages (i.e. $L\tau \gg \alpha$)* – In this case, only the transmission cost is to be minimized. Techniques, like pipelined algorithms using “disjoint” spanning trees, ensure a nearly optimal cost for message transmission (see [8, 12, 15] for hypercube network or [4, 9] for grids, and [10, 14] for a general survey). Note that this kind of optimality always implies a large number of communication rounds.

Some authors considered also families of algorithms $\mathcal{A}(p)$ depending on some parameter p (for example, for pipelined algorithms p usually represents the size of packets, assuming that all packets have the same size – see the survey [10] and the associated references). In this case, the family $\mathcal{A}(p)$ uses $R_{\mathcal{A}}(p)$ rounds with a transmission cost $T_{\mathcal{A}}(p)$. Within a family, as the functions $R_{\mathcal{A}}(p)$ and $T_{\mathcal{A}}(p)$ have usually an opposite behavior, the optimal p minimizing the communication time can be chosen for given α , L and τ .

Here, in order to design optimal algorithms and to study this trade-off, we proceed in a different way. Given the number of rounds, we try to determine (at least to find a tight estimation of) the minimum transmission cost of any algorithm realizing the communication pattern within this given number of rounds. A general theoretical study about such a trade-off has almost never been done (except in [5]) and therefore no optimal protocol was known but in the case of short or large messages, by optimizing only the leading term.

The determination of the minimum transmission cost of an algorithm using a given number of rounds appears to be a difficult problem. So, we restrict ourselves to the communication scheme called broadcasting, where a single source has to send the same message of length L to all the other nodes in the network. We also assume that links are full-duplex (*i.e.* there is a bidirectional link between each pair of node) and that a node can simultaneously send and receive data on at most k -ports. We consider the case where the network is the complete graph K_N and we assume that $N = (k + 1)^T$, as this is the maximum number of nodes which can be informed in T rounds. With these hypotheses, any broadcast algorithm needs *at least* T rounds.

In this paper, we study the antinomy between the number of rounds and the transmission cost. When the number of rounds is small, more exactly of the form $T + r$ with $r \leq \log_{k+1}(T) + 1$, we show that an estimation of the transmission cost is given by an exponential decrease $\frac{T}{(k+1)^r} + \Theta(\frac{1}{k})$. Conversely, for a large number of extra rounds $r \geq T - 1$, we observe a linear decrease $\frac{T+r}{kr+1} - \mathcal{O}(\frac{1}{r^2})$, whereas for intermediate values no trade-off occurs, the transmission cost remaining close to $2/k$.

The article is organized as follows. In section 2 we give the definitions and hypotheses and show why our study is valid for both the store-and-forward and the circuit-switched routing modes. We also introduce a function allowing to study the trade-off issue formally. In section 3 we give an upper bound for broadcasting with a few extra rounds, while we derive tight lower bounds in section 4. Other tight bounds for many extra rounds are derived in section 5. In conclusion we summarize our results and discuss their applications.

2 Framework

Recall that we consider *synchronous* protocols made of successive communication rounds.

Definition 1 The **communication time** of a broadcast algorithm \mathcal{A} in a network G can be expressed as $b(\mathcal{A}) = R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, where $R_{\mathcal{A}}$ is the number of rounds and $T_{\mathcal{A}}$ the transmission cost. The minimum broadcast time over all algorithms is denoted by $b(G)$.

For any k -ports network with N nodes, the usual lower bound on the time for broadcasting a message of length L is expressed as the sum of the lower bounds valid for the number of rounds and for the transmission cost [6]:

$$b(G) \geq \lceil \log_{k+1} N \rceil \cdot \alpha + \frac{L}{k} \cdot \tau \quad (1)$$

As already noticed, existing upper bounds have quite a different form where a trade-off appears.

The function defined below will be useful to derive a relation between the *number of rounds* and the *transmission cost*.

Definition 2 Let G be a fixed network and T the minimum number of rounds for broadcasting in G . Then, $f(G, L, r)$ is defined as the **minimum transmission cost** of any algorithm broadcasting in G a message of length L and completing in r **extra rounds**, *i.e.* using $T + r$ rounds.

With the previous definition, we have $b(G) \geq (T + r) \cdot \alpha + f(G, L, r) \cdot \tau$, for any $r \geq 0$.

Hypothesis 3 Packet sizes are assumed to be **continuous**.

This does not hold formally in practice, as messages are made of integer numbers of bits. However, as these numbers are generally large, the continuity assumption makes sense.

Property 4 Assuming continuous packet sizes, the function f is linear in L , *i.e.* $f(G, L, r) = f(G, 1, r) \cdot L$. Moreover, $f(G, 1, r)$ is a non-increasing function of r .

Proof. Let M be a message of length L and let $\mathcal{A}_{L'}$ be an algorithm broadcasting a message M' of length L' with transmission cost $f(G, L', r)$.

Let us denote by h the function $x \mapsto x \cdot L/L'$ from $[0, L']$ to $[0, L]$. Note that h induces a function H from the set of submessages of M' to the set of submessages of M . With function H we are able to define submessages of M from submessages of M' . By following the same communication protocol as $\mathcal{A}_{L'}$ with the images by H , we obtain an algorithm \mathcal{A}_L broadcasting M with cost $f(G, L', r) \cdot L/L'$. Hence, $f(G, L, r) \leq f(G, L', r) \cdot L/L'$. By inverting roles of M' and M , we get the reverse inequality.

Note at last that any algorithm using $r + 1$ rounds can emulate an algorithm using r rounds by sending empty messages during its last round, so $f(G, 1, r + 1) \leq f(G, 1, r)$. \square

Remark 5 As $f(G, L, r)$ is linear in L , we will consider w.l.o.g. that $L = 1$.

Hypothesis 6 We consider the case of the **complete graph** K_N with N nodes.

Note that the behaviour of $f(G, 1, r)$ strongly depends on the topology of network G . Any link conflict is avoided in the complete graph. Therefore we will derive lower bounds valid for every network with N nodes. Moreover, as each pair of vertices can communicate

in the complete graph, the study is also valid for both the store-and-forward and the circuit-switched routing modes.

Finally, in order to simplify computations we will assume that:

Hypothesis 7 Under the k -ports communication mode, $N = (k + 1)^T$.

This is not a very restrictive hypothesis, as N is thereby the maximum number of nodes which can be informed by a broadcast protocol completing in T communication rounds.

With all these hypotheses, we give in the next sections results on $f(K_N, 1, r)$, with $N = (k + 1)^T$, that will be denoted simply by $f_T(r)$.

With this notation, we have immediately:

Proposition 8 $f_T(0) = T$.

Proof. As $N = (k + 1)^T$, every node other than the source can be informed only once during the T rounds and must receive the whole message whose length is 1. Hence, the size of any message sent during any round must be equal to 1, and so $f_T(0) \geq 0$. A greedy algorithm matching this bound can be defined as follows. During T rounds every informed node sends the whole message to k not yet informed nodes. The broadcast is thus completed and the resulting transmission cost is T . \square

3 Upper bound for broadcasting with a few extra rounds

Here, we construct inductively algorithms for broadcasting with r extra rounds. We initiate the induction by using the greedy broadcast algorithm given in Proposition 8.

Lemma 9 For $T > 0$ and $r > 0$, $f_T(r) \leq \frac{2}{k+1} + \frac{f_{T-1}(r-1)}{k+1}$.

Proof. For $t \geq 0$, let $N_t = (k + 1)^t$, hence $N = N_T$. The idea consists in defining an algorithm in K_N using $T + r$ rounds from an algorithm in $K_{N_{T-1}}$ using $T - 1 + r - 1$ rounds.

The vertex set $V(K_N)$ of the complete graph K_N can be seen as the cartesian product $\{1, 2, \dots, k + 1\} \times \{1, 2, \dots, N_{T-1}\}$ and can be represented by the following matrix with $k + 1$

rows and N_{T-1} columns:

$$V(K_N) = \begin{pmatrix} (1,1) & (1,2) & \cdots & (1, N_{T-1}) \\ (2,1) & (2,2) & \cdots & (2, N_{T-1}) \\ \vdots & \vdots & \ddots & \vdots \\ (k+1,1) & (k+1,2) & \cdots & (k+1, N_{T-1}) \end{pmatrix}$$

The source $(1,1)$ first partitions equally the message into $k+1$ packets each of size $1/(k+1)$, denoted by m_i for $1 \leq i \leq k+1$.

- During the first round, the source $(1,1)$ sends packet m_i to node $(i,1)$, for $2 \leq i \leq k+1$. The transmission cost of this round is $1/(k+1)$.
- During the next $(T-1+r-1)$ rounds and for $1 \leq i \leq k+1$, each packet m_i is broadcast from node $(i,1)$ to all the vertices in line i . This is realized by performing $k+1$ independent and parallel identical protocols broadcasting in $(T-1)+(r-1)$ rounds a message of length $1/(k+1)$ in $k+1$ disjoint complete graphs of cardinality N_{T-1} . By definition of f_{T-1} , the transmission cost of this phase is $f_{T-1}(r-1)/(k+1)$.
- Finally, as the original message is distributed on every column, each node (i,j) sends in the last round packet m_i to the k others nodes in column j (parallel total exchanges of the packets m_i are realized in the different columns). The transmission cost of this last round is $1/(k+1)$.

Summarizing, we obtain the inductive inequality $f_T(r) \leq \frac{2}{k+1} + \frac{f_{T-1}(r-1)}{k+1}$. □

Now the following proposition can be stated.

Proposition 10 For $r \leq T$, $f_T(r) \leq \frac{T-r}{(k+1)^r} + \frac{2}{k} \left(1 - \frac{1}{(k+1)^r}\right)$.

Proof. For $r = 0$ we use Proposition 8. For $1 \leq r \leq T$, applying inductively Lemma 9 leads to $f_T(r) \leq \sum_{i=1}^r \frac{2}{(k+1)^i} + \frac{f_{T-r}(0)}{(k+1)^r}$. Note that the condition $r \leq T$ follows from the fact that Lemma 9 can be applied only for $T > 0$. □

4 Lower bounds for broadcasting with a few extra rounds

4.1 Notation

For every $t \geq 0$, let $N_t = (k+1)^t$. Then N_t is an upper bound on the maximum number of nodes having received some information after the t first rounds. For $1 \leq t \leq T+r$, we need the following definitions:

$$\begin{aligned} l_t & : && \text{maximum size of the messages sent during round } t, \\ \mathcal{L}_t = \sum_{i=1}^t l_i & : && \text{transmission cost after the first } t \text{ rounds,} \\ i_t & : && \text{total information received during the round } t, \\ \mathcal{I}_t = \sum_{i=1}^t i_i & : && \text{total information received after the first } t \text{ rounds,} \end{aligned}$$

and by convention, $\mathcal{L}_0 = \mathcal{I}_0 = 0$.

4.2 Preliminaries

To study the transmission cost once some vertices have already been informed, we need a to consider particular protocols where a given number s of sources can simultaneously broadcast information, and which we call *broadcast protocols with s sources*.

Lemma 11 *For any broadcast protocol with s sources, for any $t \geq 1$ and for a fixed \mathcal{L}_t , we have $\mathcal{I}_t \leq s \cdot \frac{\mathcal{L}_t}{t} \cdot (N_t - 1)$. Moreover, equality holds only if all the costs l_i , with $1 \leq i \leq t$, are equal.*

Proof. The following holds for each round $t+1$ such that $t \geq 0$.

- At most sN_t nodes can send messages, hence $i_{t+1} \leq k(sN_t)l_{t+1}$.
- Let $i(x, t)$ be the total information received by node x after t rounds. The information sent by the sources is at most $s(kl_{t+1})$ and each of the other nodes can send no more than what it has already received, hence $i_{t+1} \leq s(kl_{t+1}) + k \sum_{x \in G} i(x, t) = skl_{t+1} + k\mathcal{I}_t$.

So for every t such that $0 \leq t \leq T$, we have:

$$\mathcal{I}_{t+1} \leq \mathcal{I}_t + ksN_t l_{t+1} \tag{2}$$

$$\mathcal{I}_{t+1} \leq (k+1)\mathcal{I}_t + ksl_{t+1} \tag{3}$$

Note that the lemma is true for $t=1$, as we have $N_0=1$, $\mathcal{I}_0=0$, leading to $\mathcal{I}_1 \leq ksl_1 = sl_1(N_1-1)$. We assume that it holds for a given t and we prove it for $t+1$.

We recall that \mathcal{L}_{t+1} is fixed. For a protocol with some transmission cost l_{t+1} at round $t + 1$, we have $\mathcal{L}_t = \mathcal{L}_{t+1} - l_{t+1}$ and by induction hypothesis

$$\mathcal{I}_t \leq s \cdot \frac{\mathcal{L}_{t+1} - l_{t+1}}{t} \cdot (N_t - 1)$$

From (2) and (3), we get:

$$\begin{aligned} \mathcal{I}_{t+1} &\leq \mathcal{I}_t + ksN_t l_{t+1} \\ &\leq s \left(\frac{\mathcal{L}_{t+1} - l_{t+1}}{t} (N_t - 1) + kN_t l_{t+1} \right) = \varphi_1(l_{t+1}) \\ \mathcal{I}_{t+1} &\leq (k+1)\mathcal{I}_t + ksl_{t+1} \\ &\leq s \left(\frac{\mathcal{L}_{t+1} - l_{t+1}}{t} (N_t - 1)(k+1) + kl_{t+1} \right) = \varphi_2(l_{t+1}) \end{aligned}$$

Functions φ_1 and φ_2 are two affine functions in l_{t+1} , φ_1 being increasing while φ_2 is non-increasing. Therefore, $\min(\varphi_1(l_{t+1}), \varphi_2(l_{t+1}))$ is smaller than the common value obtained by solving $\varphi_1(l_{t+1}) = \varphi_2(l_{t+1})$. The solution of this equation is $l_{t+1} = \frac{\mathcal{L}_{t+1}}{t+1}$ and we obtain:

$$\mathcal{I}_{t+1} \leq \varphi_1(l_{t+1}) = \varphi_2(l_{t+1}) = \frac{\mathcal{L}_{t+1}}{t+1} \cdot (N_{t+1} - 1)$$

To reach this bound, we must have by induction $l_{i+1} = \frac{\mathcal{L}_{i+1}}{i+1}$ for each $i \leq t$, which indeed implies that all the intermediate costs are equal. In this view, the transmission optimization may be called *greedy*. \square

Corollary 12 For $0 \leq p < q \leq T + 1$, the information transmission \mathcal{I}_q after q rounds of every protocol having fixed values of \mathcal{I}_p , \mathcal{L}_p and \mathcal{L}_q satisfies:

$$\mathcal{I}_q \leq \mathcal{I}_p + \frac{\mathcal{L}_q - \mathcal{L}_p}{q - p} \cdot (N_q - N_p) \quad (4)$$

Proof. First, note that after p rounds at most N_p nodes can be informed. So, after round p , an information transmission occurs from N_p sources during $q - p$ rounds and with a fixed cost $(\mathcal{L}_q - \mathcal{L}_p)$. According to Lemma 11 applied with $t = q - p$, $\mathcal{L}_t = \mathcal{L}_q - \mathcal{L}_p$ and $s = N_p$ we have $(\mathcal{I}_q - \mathcal{I}_p) \leq N_p \left(\frac{\mathcal{L}_q - \mathcal{L}_p}{q - p} (N_{q-p} - 1) \right)$ and the result holds. \square

Remark 13 For values of t greater than $T + 1$, Lemma 11 does not give an accurate estimation of \mathcal{I}_t . This is due to the fact that for $t \geq T + 1$, the maximum number of

informed nodes is N_T , which is much smaller than N_t . Nevertheless, lower bounds on $f_T(r)$ can still be derived from inequalities (2) and (3), but they are not tight enough and the refined analyses given in the next sections allow to obtain better results.

Consider a dissemination protocol with one source using $T+r$ rounds. After round $T+r$, each node must have received a complete message of length 1, hence, if $i(x, t)$ (as in the proof of Lemma 11) denotes the total information received by node x after t rounds, we have for any node x , $i(x, T+r) \geq 1$. By summing over all the nodes (except the source) we get the next inequality:

$$\mathcal{I}_{T+r} \geq N_T - 1 \tag{5}$$

Remark 14 Using the fact that any broadcast algorithm using T rounds satisfies $\mathcal{I}_T \geq N - 1$, we find again Proposition 8 from Lemma 11, *i.e.* the transmission cost \mathcal{L}_T is at least T .

4.3 One extra round

Proposition 15 $f_T(1) = \frac{T+1}{k+1}$

Proof. We prove that $f_T(1) \geq \frac{T+1}{k+1}$ by reduction to the absurd, assuming the existence of an algorithm using $T+1$ rounds with a transmission cost $\mathcal{L}_{T+1} < \frac{T+1}{k+1}$.

First we prove that $l_{T+1} < \frac{1}{k}$. With inequalities (5) and (3) and from Lemma 11, we have:

$$\begin{aligned} N_T - 1 &\leq \mathcal{I}_{T+1} \\ &\leq (k+1)\mathcal{I}_T + kl_{T+1} \end{aligned}$$

$$\mathcal{I}_T \leq \frac{\mathcal{L}_T}{T}(N_T - 1)$$

$$\begin{aligned} \text{by hypothesis, } \mathcal{L}_T &= \mathcal{L}_{T+1} - l_{T+1} \\ &< \frac{T+1}{k+1} - l_{T+1} \end{aligned}$$

Combining the three relations leads to $l_{T+1} < \frac{1}{k+1 - \frac{kT}{N-1}} \leq \frac{1}{k}$.

As $kl_{T+1} < 1$, a node without information cannot receive enough information during the last round. So, during the first T rounds each node must receive some information, exactly one message as $N = (k+1)^T$.

Let $l = \min \{l_t \mid 1 \leq t \leq T + 1\}$ be the minimum cost of all rounds. We have $(T + 1)l \leq \mathcal{L}_{T+1} < \frac{T+1}{k+1}$, so $l < \frac{1}{k+1}$. Now, we distinguish two cases.

- The minimum value l is attained for l_{T+1} . As a vertex can receive at most kl information in round $T + 1$, then it should be know at least $1 - kl$ information after round T and so, according to Proposition 8 we have $l_t \geq 1 - kl$ for $1 \leq t \leq T$.
- The minimum value l is attained for some t_0 with $1 \leq t_0 \leq T$. As the algorithm follows a spanning tree pattern during the first T rounds, exactly $(N_{t_0} - N_{t_0-1})$ nodes receive at most l information during round t_0 . Let S be the set of all the descendants of these nodes after round T ; and note that $|S| = (k + 1)^{T-t_0}(N_{t_0} - N_{t_0-1}) = N_T - N_{T-1}$. So, all the nodes in S have at most l information after round T . As $l < 1/(k + 1)$, no node in S can receive, in the last round, k messages from nodes in S . As $k(N - |S|) = |S|$, we deduce that during round $T + 1$ each node in S receives $(k - 1)$ messages of length at most l from nodes in S and one message of length at least $(1 - kl)$ from a node not in S . Consequently, any node not in S must have at least $(1 - kl)$ information after round T . So, $l_{T+1} \geq 1 - kl$ and $l_t \geq 1 - kl$ for $1 \leq t \leq T$, with $t \neq t_0$.

In both cases we have $\mathcal{L}_{T+1} \geq l + T(1 - kl) = \varphi(l)$. Finally, just note that $\varphi(l)$ is a strictly decreasing function in l . As $l < \frac{1}{k+1}$, we have $\mathcal{L}_{T+1} > \varphi(\frac{1}{k+1}) = \frac{T+1}{k+1}$, contradicting our main hypothesis. As the conditions used in this proof are necessary to perform any broadcast protocol, then, it is proved that $f_T(1) \geq \mathcal{L}_{T+1} \geq \frac{T+1}{k+1}$ and the result follows from Proposition 10. \square

4.4 Two extra rounds

Before to state a lower bound for two extra rounds, we need to introduce a new notion about the *transmission cost*.

Definition 16 Given a network G with one source vertex. A **transmission algorithm** using t rounds is a communication protocol such that:

- Initially the source has a message of infinite size, and all the other vertices in G have no information.
- During the t rounds of the protocol, each vertex with some information can send k different submessages such that each size is not greater than the known information.
- At the end of the protocol, *i.e.* after t rounds, all the vertices in G have received an amount of information at least equal to 1.

Remark 17 Clearly, these transmission constraints are weaker than those for broadcasting. Indeed, at the end of a transmission protocol, we make no assumption about what message reaches each vertex. We only restrict the final amount of information in each vertex.

Definition 18 Let $g_T(r)$ be the **minimal transmission cost** of any transmission algorithm for the complete graph K_N using $T + r$ rounds.

Property 19 *As any broadcast algorithm satisfies the transmission constraints, we have $f_T(r) \geq g_T(r)$.*

Our lower bound can now be stated.

Proposition 20 $f_T(2) = \frac{T}{(k+1)^2} + \frac{2}{k+1} - \Theta\left(\frac{1}{kT}\right)$.

Proof. From Proposition 10 we have $f_T(2) \leq \frac{T}{(k+1)^2} + \frac{2}{k+1}$. We consider an optimal broadcast algorithm using $T + 2$ rounds.

After the first round, at least $(1 - kl_1)$ information remains to be broadcast from the source in $T + 1$ rounds, so $l_1 + (1 - kl_1)f_T(1) \leq \mathcal{L}_{T+2} \leq \frac{T}{N_2} + \frac{2}{N_1}$ and from Proposition 15 we deduce:

$$l_1 \geq \frac{kT - 1 - k}{(kT - 1)N_1} = \frac{1}{N_1} - \Theta\left(\frac{1}{kT}\right) \quad (6)$$

Before the last round, every node must have at least $(1 - kl_{T+2})$ information. Therefore, by definition of the function g_T , we get that \mathcal{L}_{T+1} is at least $(1 - kl_{T+2})g_T(1)$. So l_{T+2} satisfies the same inequality as l_1 :

$$l_{T+2} \geq \frac{kT - 1 - k}{(kT - 1)N_1} = \frac{1}{N_1} - \Theta\left(\frac{1}{kT}\right) \quad (7)$$

Let λ be such that $\mathcal{L}_{T+2} = l_1 + T\lambda + l_{T+2}$. By applying Corollary 12 with $p = 1$ and $q = T + 1$, we obtain $\mathcal{I}_{T+1} \leq k + \lambda(N_{T+1} - N_1)$. Then from $N_T - 1 \leq \mathcal{I}_{T+2} \leq N_1\mathcal{I}_{T+1} + k$ we deduce:

$$\lambda \geq \frac{1}{N_2} - \Theta\left(\frac{1}{N}\right) \quad (8)$$

As T/N is small compared to $1/(kT)$, combining (6), (7) and (8) gives:

$$f_T(2) = \mathcal{L}_{T+2} = l_1 + T\lambda + l_{T+2} \geq \frac{T}{N_2} + \frac{2}{N_1} - \Theta\left(\frac{1}{kT}\right)$$

□

4.5 Some extra rounds

We have given in section 3 algorithms for broadcasting with r extra rounds, inducing upper bounds on $f_T(r)$. For large enough values of T , their transmission cost can be approximated by the sum of the two terms $\frac{T}{(k+1)^r}$ and $\frac{2}{k}$. The former is dominating for small values of r and decreases exponentially with r . We prove in the following that in this case the upper bounds of Proposition 10 are tight. Therefore we show the efficiency of our algorithms, which allow to decrease significantly the transmission cost with a few additive rounds.

Theorem 21 For $0 \leq r \leq T$, $f_T(r) = \frac{T}{(k+1)^r} + \Theta\left(\frac{1}{k}\right)$.

Proof. We derive here a lower bound on $f_T(r)$ matching the upper bound given by Proposition 10. Let consider a broadcast algorithm using $T + r$ rounds. After round $r - 1$ it can be seen as a broadcast protocol with (at most) N_{r-1} sources. So from inequality (3) with $t = T + r - 1$ we get:

$$\mathcal{I}_{T+r} \leq (k+1) \cdot \mathcal{I}_{T+r-1} + k \cdot N_{r-1} \cdot l_{T+r}$$

By iteration for $T \leq t < T + r - 1$ and as $l_{t+1} \geq 1$, we obtain:

$$\mathcal{I}_{T+r} \leq N_r \cdot \mathcal{I}_T + \sum_{i=0}^{r-1} (k \cdot N_i \cdot l_{T+r-i}) \leq N_r \cdot \mathcal{I}_T + (N_r - 1)$$

Then, we apply Lemma 11 with $t = T$,

$$N_T \leq \mathcal{I}_{T+r} + 1 \leq N_r \cdot (\mathcal{I}_T + 1) \leq N_r \cdot \left(\frac{\mathcal{L}_T}{T} (N - 1) + 1 \right)$$

and it follows that $\mathcal{L}_T \geq T \cdot \frac{N_{T+r-1}}{N-1} = \frac{T}{N_r} + \frac{T}{N-1} \cdot \frac{1-N_r}{N_r} = \frac{T}{N_r} + \Theta\left(\frac{T}{N}\right)$.

Now, we must compare this relation with the upper bound given by Proposition 9, that is we compare $\frac{T}{N}$ with $\frac{1}{k}$. As $\frac{T}{N}$ is small compared to $\frac{1}{k}$ the proposition holds. \square

The above proposition is meaningless when $r > \log_{k+1}(T)$ as it reduces to $f_T(r) = \Theta\left(\frac{1}{k}\right)$. In the following we investigate this case and we give a better bound.

Proposition 22 For $r < T - 1$, $f_T(r) \geq \frac{2}{k}$.

Proof. We consider a broadcast algorithm using $T + r$ rounds and having a transmission cost strictly less than $2/k$, and we show that $r \geq T - 1$. Let t_0 be the last round after which

it remains some information not yet broadcast from the source, hence $\mathcal{L}_{t_0+1} \geq 1/k$. After round t_0 , T rounds are necessary to broadcast this last piece of information in the entire network, so $t_0 + T \leq T + r$ and $t_0 \leq r$. Given the assumption on the total transmission cost, every node must have received some information after the round $t_0 + 1$, otherwise $\mathcal{L}_{T+r} \geq \mathcal{L}_{t_0+1} + 1/k \geq 2/k$. Therefore, we have $t_0 + 1 \geq T$, leading to $r \geq T - 1$. \square

Corollary 23 For $\log_{k+1}(T) + 1 < r < T - 1$, $f_T(r) = \frac{2}{k} + \mathcal{O}\left(\frac{1}{k^2}\right)$.

Proof. By Proposition 10 we have $f_T(r) \leq \frac{2}{k} + \frac{k(T-r)-2}{k(k+1)^r}$. As for $r \geq \log_{k+1}(T) + 2$, we have $(k+1)^r \geq T \cdot (k+1)^2$, then, $f_T(r) \leq \frac{2}{k} + \mathcal{O}\left(\frac{1}{k^2}\right)$ and the result holds with Proposition 22. \square

Remark 24 It has been proved that the minimum transmission cost $f_T(r)$ decreases exponentially when a small number of extra rounds is used. However, this behaviour stops as soon as r is greater than $\log_{k+1}(T)$. Indeed, the main cost becomes $2/k$. This can be explained as follows: to have an unit length message broadcast in a whole network under the k -ports mode, a transmission cost of $1/k$ is required to make the message go out from the source and also to make it enter into any other node. We could think that the pipeline techniques could decrease this cost as they should allow to have these two costs covered by each other, in order to go below a total cost of $2/k$. Unfortunately, Proposition 22 proves that no pipelined algorithm can be usefull if $r < T - 1$.

5 Broadcasting with many extra rounds

In this section, we investigate the case where many extra rounds are used, *i.e.* with $r \geq T - 1$. In this range pipelined protocols are possible and we study how such techniques allow to decrease the transmission cost.

5.1 Upper bounds: pipelined algorithms

Proposition 25 For $r \geq T - 1$, $f_T(r) \leq \frac{T+r}{kr+1}$.

Proof. This upper bound follows from an algorithm. In the following \mathbb{Z}_q will denote the set of integers modulo q . The elements of \mathbb{Z}_q will be taken in the set $\{0, 1, \dots, q - 1\}$. The vertex set of the complete graph K_N is represented by \mathbb{Z}_{k+1}^T . The elements in \mathbb{Z}_{k+1}^T may be expressed in the canonical base $\{e_i\}_{1 \leq i \leq T}$, that is, we may denote a vertex $x = (x_1, x_2, \dots, x_T) \in \mathbb{Z}_{k+1}^T$ by $\sum_{i=1}^T x_i \cdot e_i$. The source is the element with all components

equal to 0. In the following, we need to define a *labelling* of the arcs to be used in our algorithm. For each node x , for $1 \leq j \leq k$ and $1 \leq i \leq T$, the arc $(x, x + j \cdot e_i)$ will be labelled by i . In our algorithm, only arcs labelled by i are used during a round $t = i \bmod T$.

Now, let U be the set of vertices such that $U = \{u \in V(K_N) \mid u = j \cdot e_i, \text{ with } 1 \leq j \leq k \text{ and } 1 \leq i \leq T\}$. For each node $u \in U$, let $\mathcal{T}_0(u)$ be simply the vertex u . Now, for $1 \leq h \leq T$, let $\mathcal{T}_h(u)$ be the tree induced from $\mathcal{T}_{h-1}(u)$ by adding all the possible arcs labelled by $(i + h \bmod T)$. Thus, $\mathcal{T}_T(u)$ is a directed spanning tree rooted in u .

In this way, we have defined $k \cdot T$ arc-disjoint spanning trees and the algorithm mainly consists in pipelining concurrently in each of them. Moreover, to maintain the full use of the bandwidth until the end of the process, the last piece of information is broadcast during the last T rounds using a spanning tree rooted in the source. Now, we describe the algorithm.

- The source first cuts equally its unit length message into $kr + 1$ distinct submessages, denoted by m_l with $1 \leq l \leq kr + 1$.
- For $1 \leq t \leq r$, at round t the source sends submessage $m_{(t-1)k+j}$, for $1 \leq j \leq k$, to node $u = j \cdot e_{(t \bmod T)}$.
- When a submessage m_l , with $1 \leq l \leq kr$, has been sent from the source to a node $u \in U$, it is broadcast in $\mathcal{T}_T(u)$ during the next T rounds (using thereby all the arcs labelled by i at each round $t = i \bmod T$).
- The last submessage m_{kr+1} is broadcast from the source in the T last rounds, using all the possible arcs labelled by $t \bmod T$ at round t .

The different rounds use the same transmission cost equal to $1/(kr + 1)$, that is the length of each submessage. This leads to the expected total cost $(T + r)/(kr + 1)$. \square

5.2 Lower bounds

Lemma 26 For any $r \geq 0$, if $f_T(r) \geq \frac{T+r}{kr+1}$ then $f_T(r + T) \geq \frac{T+(r+T)}{k(r+T)+1}$.

Proof. Consider an optimal algorithm broadcasting an unit length message in $T+(T+r)$ rounds. After round T , at least $(1 - k\mathcal{L}_T)$ information remains to be broadcast from the source in $T + r$ extra rounds, hence:

$$f_T(r + T) \geq \mathcal{L}_T + f_T(r) \cdot (1 - k\mathcal{L}_T) \geq \mathcal{L}_T + \frac{T+r}{kr+1} \cdot (1 - k\mathcal{L}_T) = \varphi_1(\mathcal{L}_T)$$

Note also that after time T , an amount of $(N - 1 - \mathcal{I}_T)$ information has to be transmitted on at most $k(N - 1)$ links at each round (as those incoming the source are useless). From Lemma 11, we know that $\mathcal{I}_T \leq \frac{\mathcal{L}_T}{T}(N - 1)$, hence:

$$f_T(T + r) \geq \mathcal{L}_T + \frac{1}{k} \left(1 - \frac{\mathcal{L}_T}{T}\right) = \varphi_2(\mathcal{L}_T)$$

As in the proof of Lemma 11, φ_1 is a decreasing function while φ_2 is an increasing function. So $\min(\varphi_1(\mathcal{L}_T), \varphi_2(\mathcal{L}_T)) \geq \varphi_1(\xi)$, where ξ satisfies $\varphi_1(\xi) = \varphi_2(\xi)$. Solving the equation we get:

$$\begin{aligned} \xi + \frac{T+r}{kr+1}(1 - k\xi) &= \xi + \frac{1}{k}(1 - \frac{\xi}{T}) \\ \xi &= \frac{T}{k(r+T)+1} \\ \varphi_1(\xi) = \varphi_2(\xi) &= \frac{T+(r+T)}{k(r+T)+1} \end{aligned}$$

So we have $f_T(T + r) \geq \frac{T+(r+T)}{k(r+T)+1}$. □

Lemma 27 $f_T(T + 2) = \frac{T+(T+2)}{k(T+2)+1}$

Proof. From Proposition 25 we have $f_T(T + 2) \leq \frac{T+(T+2)}{k(T+2)+1}$ and by using the same arguments we get:

$$\begin{aligned} \frac{T+(T+2)}{k(T+2)+1} \geq f_T(T + 2) &\geq \mathcal{L}_{T+1} + f_T(1) \cdot (1 - k\mathcal{L}_{T+1}) \\ \frac{T+(T+2)}{k(T+2)+1} &\geq \mathcal{L}_{T+1} + \frac{T+1}{k+1}(1 - k\mathcal{L}_{T+1}) \\ \mathcal{L}_{T+1} &\geq \frac{T+1}{k(T+2)+1} \end{aligned}$$

Let us denote by \mathcal{L}'_{T+1} the sum of the transmission costs of the last $T + 1$ rounds, hence $f_T(T + 2) = \mathcal{L}_{T+1} + \mathcal{L}'_{T+1}$.

$$\begin{aligned} \frac{T+(T+2)}{k(T+2)+1} \geq f_T(T + 2) &\geq f_T(1) \cdot (1 - k\mathcal{L}'_{T+1}) + \mathcal{L}'_{T+1} \\ \frac{T+(T+2)}{k(T+2)+1} &\geq \frac{T+1}{k+1}(1 - k\mathcal{L}'_{T+1}) + \mathcal{L}'_{T+1} \\ \mathcal{L}'_{T+1} &\geq \frac{T+1}{k(T+2)+1} \end{aligned}$$

So the lemma holds. □

Theorem 28 For $r \geq T$, and $r = 0, 1, 2 \pmod T$, $f_T(r) = \frac{T+r}{kr+1}$.

Proof. We use Lemma 26 inductively. To initiate the induction for $r = 0 \pmod T$, we use Proposition 8, that is $f_T(0) = T$ verifying the condition of Lemma 26. For $r = 1 \pmod T$, we use Proposition 15, that is $f_T(1) = \frac{T+1}{k+1}$. Finally, for $r = 2 \pmod T$, we use Lemma 27 giving the exact value of $f_T(T+2)$, that verifies the needed condition too. For these three cases, we obtain $f_T(r) \geq \frac{T+r}{kr+1}$. The result follows from Proposition 25. \square

Corollary 29 For $T \in \{1, 2, 3\}$ and for any $r \geq T$, $f_T(r) = \frac{T+r}{kr+1}$.

Corollary 30 For $r \geq T$, $f_T(r) = \frac{T+r}{kr+1} - \mathcal{O}\left(\frac{1}{r^2}\right)$.

Proof. Assume that $r_1 = aT$, $r = aT + b$ and $r_2 = (a+1)T$ with $a \geq 1$ and $0 \leq b < T$. We know by Theorem 28 that $f_T(r_1) = \frac{(a+1)T}{kaT+1}$, and $f_T(r_2) = \frac{(a+2)T}{k(a+1)T+1}$. As by Property 4, f_T is a non-increasing function, we have:

$$\begin{aligned} f_T(r_1) &\geq f_T(r) \geq f_T(r_2) \\ \frac{(a+1)T}{kaT+1} &\geq f_T(r) \geq \frac{(a+2)T}{k(a+1)T+1} \end{aligned}$$

So we obtain for $r \geq T$,

$$\begin{aligned} f_T(r) &= \frac{1}{k} + \frac{1}{ka} - \frac{1}{k^2Ta} - \mathcal{O}\left(\frac{1}{a^2}\right) \\ f_T(r) &= \frac{1}{k} + \frac{T}{kr} - \frac{1}{k^2r} - \mathcal{O}\left(\frac{1}{r^2}\right) = \frac{T+r}{kr+1} - \mathcal{O}\left(\frac{1}{r^2}\right) \end{aligned}$$

\square

The above proposition and corollary make the following conjecture very likely to hold:

Conjecture 31 For any T and any $r \geq T$, $f_T(r) = \frac{T+r}{kr+1}$.

6 Conclusion

In this paper, we have shown the existence of a trade-off between the number of rounds and the transmission cost of broadcast protocols. We have derived tight bounds on the minimum communication time $(T+r) \cdot \alpha + f_T(r) \cdot L \cdot \tau$ of any algorithm using $T+r$ rounds for broadcasting a message of length L in the bidirectional k -ports complete network K_N , with $N = (k+1)^T$. The following table summarizes the different results obtained:

Number of extra rounds	Minimum transmission cost	Comment
$r = 0$	$f_T(0) = T$	greedy
$r = 1$	$f_T(1) = \frac{T+1}{k+1}$	exponential trade-off
$r = 2$	$f_T(2) = \frac{T}{(k+1)^2} + \frac{2}{k+1} - \Theta\left(\frac{1}{kT}\right)$	
$2 \leq r \leq \log_{k+1}(T) + 1$	$f_T(r) = \frac{T}{(k+1)^r} + \Theta\left(\frac{1}{k}\right)$	no trade-off
$\log_{k+1}(T) + 1 < r < T - 1$	$f_T(r) = \frac{2}{k} + \mathcal{O}\left(\frac{1}{k^2}\right)$	
$r \geq T, r = 0, 1, 2 \bmod T$	$f_T(r) = \frac{T+r}{k+1}$	linear trade-off
$r \geq T - 1$	$f_T(r) = \frac{T+r}{kr+1} - \mathcal{O}\left(\frac{1}{r^2}\right)$	

Our lower bounds have been obtained for complete networks, therefore they are valid for any topology. Conversely, our matching upper bounds derived from our algorithms require strong connectivity to be reached. For example, we can show that the exponential decrease of the transmission cost does not occur in a ring network. However, all routing mechanisms based on the circuit-switched mode can take advantage of our algorithms, due to the additional connectivity provided. In particular, the *Wavelength Division Multiplexed* (WDM) optical routing can offer full connectivity in various network topologies (see [3, 2]). Our study has thus been fruitfully taken into account in [1] for multi-hop optical ring and mesh networks. Further work should concentrate on other less connected topologies, for which greater lower bounds will arise.

Acknowledgements

We are grateful to J-C. Bermond for his helpful remarks and comments.

References

- [1] B. Beauquier. Broadcasting in WDM optical rings and tori. In *Proc. the DIMACS Workshop : "Multichannel Optical Networks : Theory and Practice"*, volume 46 of *DIMACS Series on Discrete Mathematics and Theoretical Computer Science*, pages 63–73, Rutgers University, NJ, March 1998. American Mathematical Society.
- [2] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes, and U. Vaccaro. Graph problems arising from wavelength-routing in all-optical networks. In *Proc. of 2nd IEEE Workshop on Optics and Computer Science (WOCS, part of IPPS'97)*, Geneva, Switzerland, April 1997. IEEE Press.
- [3] J-C. Bermond, L. Gargano, S. Pérennes, A. A. Rescigno, and U. Vaccaro. Efficient collective communication in optical networks. In *Proc. of 23rd International Colloquium on Automata, Languages and Programming (ICALP'96)*, volume 1099 of *Lecture Notes in Computer Science*, pages 574–585. Springer-Verlag, July 1996.
- [4] J-C. Bermond, P. Michallon, and D. Trystram. Broadcasting in wraparound meshes with parallel monodirectional links. *Parallel Computing*, 18:639–648, 1992.
- [5] J. Bruck, C-T. Ho, S. Kipnis, and D. Weathersby. Efficient algorithms for all-to-all communications in multi-port message-passing systems. In *Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '94)*, pages 298–309, Cape May, New Jersey, June 1994. ACM.
- [6] O. Delmas. *Communications par commutation de circuits dans les réseaux d'interconnexion*. Thèse de doctorat, Université de Nice - Sophia Antipolis, U.F.R. de Sciences, Ecole Doctorale - Sciences pour l'Ingénieur, France, January 1997.
- [7] E. Fleury. *Communications, routage et architectures des machines à mémoire distribuée - Autour du routage wormhole*. Thèse de doctorat, Université de Lyon, Ecole Normale Supérieure de Lyon, France, 1996.
- [8] P. Fraigniaud. Performance analysis of broadcasting hypercubes with restricted communication capabilities. *Journal of Parallel and Distributed Computing*, 16:15–26, 1992.
- [9] P. Fraigniaud and C. Laforest. Disjoint spanning trees of small depth. In G.R. Joubert, D. Trystram, F.J. Peters, and D.J. Evans, editors, *ParCo '93*, pages 105–112, 1993.
- [10] P. Fraigniaud and E. Lazard. Methods and problems of communication in usual networks. *Discrete Applied Mathematics*, 53:79–133, 1994. Special volume proceedings international workshop on broadcasting and gossiping 1990.
- [11] S.M. Hedetniemi, S.T. Hedetniemi, and A.L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18:319–349, 1986.

- [12] C.T. Ho and S.L. Johnsson. Optimum broadcasting and personalized communication in hypercubes. *IEEE Transactions on Computers*, 38(9):1249–1268, 1989.
- [13] J. Hromkovic, R. Klasing, B. Monien, and R. Peine. *Combinatorial Network Theory*, chapter Dissemination of information in interconnection networks (broadcasting and gossiping), pages 125–212. Kluwer Academic, 1995.
- [14] Jean de Rumeur. *Communication dans les réseaux de processeurs*. Collection Etudes et Recherches en Informatique. Masson, Paris, 1994.
- [15] Q. F. Stout and B. Wagar. Intensive hypercube communication, prearranged communication in link-bound machines. *Journal of Parallel and Distributed Computing*, 10:167–181, 1990.

ANNEXE E

Broadcasting in WDM optical rings and tori

Broadcasting in WDM Optical Rings and Tori

Bruno Beauquier*

Thème 1 — Réseaux et systèmes
Projet Sloop

Avril 1998

Abstract: The well-known spanning binomial tree broadcast algorithm is generalized to obtain two families of broadcast algorithms for optical rings and two-dimensional toroidal meshes (tori) using *Wavelength Division Multiplexing (WDM)*. These generalizations take advantage of the concurrent transmission through optical links offered by WDM. Their performances are measured under the *linear cost model*: the cost of sending a message of L bits is defined as $\alpha + L\tau$, where α is the latency and τ is the per-byte transmission cost. It is assumed that each node can concurrently transmit one message and receive one message. Our algorithms are based on the familiar spanning binomial tree and on the dimensional exchanges commonly used on hypercubes. We restrict the number of nodes in a ring and in each dimension of a torus to be a power of two. The algorithms described in this paper offer significant performance improvements over the basic spanning tree broadcast.

Key-words: Optical networks, WDM, broadcasting.

(Résumé : *tsvp*)

To appear in the DIMACS Series on Discrete Mathematics and Theoretical Computer Science published by the American Mathematical Society. Presented in the workshop on “Multichannel Optical Networks: Theory and Practice”, (DIMACS, March 16-19, 1998).

* SLOOP, joint project I3S-CNRS/UNSA/INRIA. Work partially supported by the AFIRST in the framework of the French-Israeli project *Communication Algorithms in Optical Networks*.
Email : beauquier@sophia.inria.fr

Diffusion par multiplexage en longueur d'onde dans les anneaux et les tores optiques

Résumé : L'algorithme bien connu de diffusion sur un arbre couvrant binomial est généralisé pour obtenir deux familles d'algorithmes de diffusion dans des anneaux et des tores optiques utilisant la technique du *multiplexage en longueur d'onde* (en anglais, *Wavelength Division Multiplexing* : WDM). Ces généralisations exploitent les avantages de la technologie WDM qui permet des transmissions de données concurrentes au sein d'une même fibre optique. Les performances des algorithmes sont mesurées par un modèle de coût affine : le temps pour envoyer un message de L bits est défini par $\alpha + L\tau$, où α est la latence et τ l'inverse de la bande passante. Il est supposé que chaque nœud peut en même temps envoyer et recevoir un seul message. Nos algorithmes sont basés sur des plongements des arbres couvrants binomiaux et des échanges dimensionnels, classiquement utilisés dans les hypercubes.

Mots-clé : Réseaux optiques, multiplexage en longueur d'onde, diffusion.

1 Introduction

Optics is emerging as a key technology in communication networks, promising very high speed local or wide area networks of the future. A single optical wavelength supports rates of gigabits per second, which in turn support multiple channels of voice, data and video [4, 6]. Multiple laser beams that are propagated over the same fiber on distinct optical wavelengths can increase this capacity even further. This is achieved through *Wavelength Division Multiplexing* (or *WDM*) [3], by partitioning the optical bandwidth into several channels and allowing the transmission of multiple data streams concurrently along the same optical fiber.

The problem we consider here is motivated by *multi-hop* communication networks [7] with reconfigurable wavelength selective optical switches, without wavelength converters. Different signals may travel on the same fiber-optic link (but on different wavelengths) into a node, and then exit from it on different links, keeping their original wavelengths.

In such communication networks, each source-destination pair can be connected by an end-to-end transparent channel that is identified through a wavelength and a physical path. This necessitates to tune the sender and the receiver to the same wavelength and to set the intermediate optical switches in the right configuration. Wavelengths being a limited resource, solutions to the problem of efficient routing and wavelengths allocation are of importance for the future development of optical technology.

A *broadcast* is an operation where one node of a distributed multicomputer network has a message that must be copied to each other node. The work presented here offers solutions to the broadcast problem for optical rings and toroidal meshes based on the familiar spanning binomial tree and on dimensional exchanges commonly used on hypercubes [5].

The paper is organized as follows. The next section describes the communication model on which the analysis of communication algorithms will be based. The broadcast algorithms and their costs are given in Section 3. In Section 4 all the results are summarized and comments are made.

2 Communication Model

This work considers solutions to the broadcast problem on rings of $n = 2^d$ processors and on two-dimensional toroidal meshes of $n = n_1 \times n_2 = 2^{d_1} \times 2^{d_2}$ processors. The nodes of a ring are numbered by integers i modulo n . The nodes of a toroidal mesh (*torus*) are numbered by couples (i, j) of an integer i modulo n_1 and an integer j modulo n_2 . Each node is connected by a pair of unidirectional fiber-optic links (one in each direction) to each of its immediate neighbours in the horizontal and vertical directions. Each node can concurrently transmit one message and receive one message. In this paper, a message in a torus is routed or only horizontally either only vertically to the receiver. Messages can be routed through a node without affecting its performance as a sender or receiver. It is assumed that all links have the same fixed bandwidth. *Wavelength conflict* occurs when the paths taken by two or more messages sent on the same wavelength have one or more links in common.

Finally, it is assumed that each receiving node allocates a buffer for each incoming message before it arrives. Under this assumption the sender can transmit a message without prior handshaking with the receiver.

A simple communication model describes the cost of sending a message of L bits as $\alpha + L\tau$, where α is the latency (including the *start-up*, the switching and the wavelength tuning delays) and τ is the per-byte transmission cost ($1/\tau$ corresponds to the bandwidth). A close examination of real message passing networks reveals a much larger collection of factors that can affect the cost of communication algorithms. Some of these factors are the length of the circuit over which the message travels, the effects of message packetization performed by the node operating system, the packet permutation costs within a node, the synchronization costs. The permutation costs generally affect the cost of global communication and remarks will be given in Section 4. The remaining factors, which have relatively small effects on cost, will not be considered here.

3 Broadcast Algorithms

Two families of broadcast algorithms for optical rings and toroidal meshes are considered. Each algorithm is based on communication patterns commonly used on hypercubes, such as spanning binomial trees and dimensional exchanges. It is assumed w.l.o.g. that node 0 of the ring (respectively node $(0, 0)$ of the torus) contains a message of length L to be broadcast to all the other nodes. The *cost* of a broadcast operation is measured from the time the originator begins the broadcast to the time the last node receives the message.

The familiar spanning tree broadcast algorithm is considered first. Improvements of that basic algorithm will take advantage of 1) the concurrent transmission through optical links offered by WDM, 2) the additional bandwidth offered by bidirectional links, and 3) the increased connectivity of toroidal meshes over rings.

3.1 Basic subroutines

3.1.1 Spanning tree broadcast

The first algorithm is based on the familiar spanning binomial tree that is used in the recursive broadcast algorithm for hypercubes [5]. This algorithm will be called the *spanning tree (ST) broadcast* and was described earlier in [1] for linear arrays and meshes. A ST broadcast on a ring or a torus of 2^d nodes takes d rounds. On the i^{th} round ($1 \leq i \leq d$), each node j that already has a copy of the message sends it to node $j \oplus 2^{d-i}$, where \oplus denotes bit-wise exclusive OR. This is illustrated for a ring in Figure 1(a). The corresponding spanning tree is shown in Figure 1(b). Each arc of the tree is labeled with the round at which it carries the message.

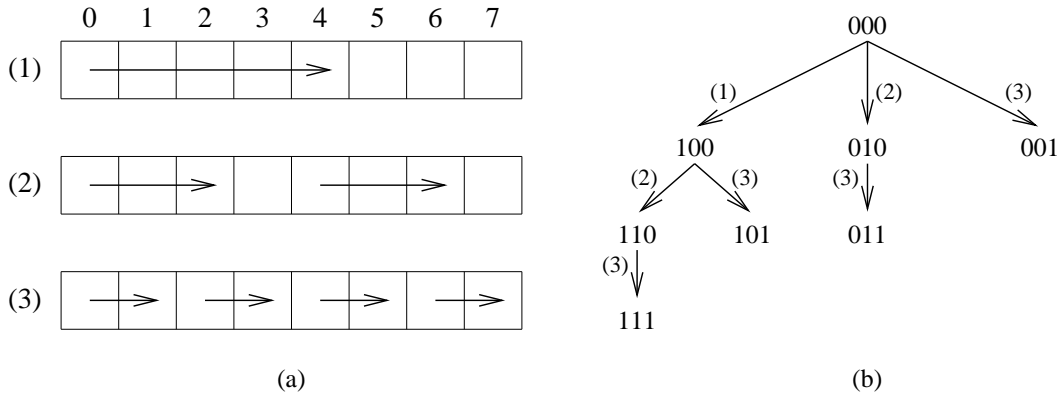


Figure 1: Spanning tree broadcast on a ring of 2^3 nodes.

The spanning binomial tree on which the ST broadcast algorithm is based is also used in all of the other broadcast algorithms considered here. The nodes of such trees are numbered in binary. (From a purely graph-theoretic point of view, these trees are not spanning trees, but they are useful for describing the scheduling and routing of messages and they will continue to be referred to as spanning trees here.) For purposes here, a *spanning tree with root 0* is a directed graph of $n = 2^d$ nodes in which each node j has children whose node numbers are obtained by complementing exactly one of the trailing zeros (if any) of j . To determine the node numbering of a spanning tree with root other than 0, exclusive OR the node number of each node in the tree with the node number of the root. For more details about the properties of these trees see [5].

In the spanning tree of a ring of 2^d nodes, some tree arcs represent directed paths of several links in the ring and some links are used in several tree arcs. Since some of the tree arcs carry messages simultaneously during this broadcast algorithm, there might be wavelength conflict. (This possibility does not arise in hypercube because there is a one-to-one correspondence between tree arcs and hypercube links.) Even though some arcs in the tree share the same links, at each round only disjoint sets of links are used, and so one wavelength is sufficient.

A ST broadcast on a torus of $2^{d_1} \times 2^{d_2} = 2^d$ nodes, based on the algorithm for rings, is shown in Figure 2. In this algorithm the message is first broadcast to the nodes in the row containing the originator, then each node in that row broadcasts the message to the nodes in its column. As in the case of the ring, wavelength conflict can be avoided with the use of one wavelength only.

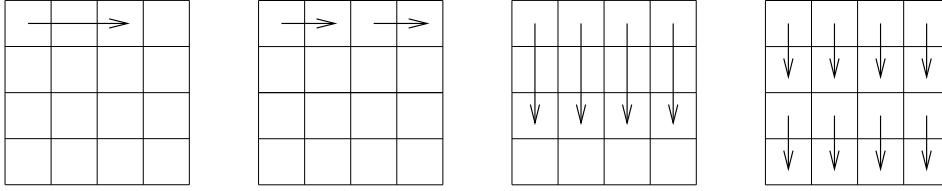


Figure 2: A simple spanning tree broadcast on a torus of $2^2 \times 2^2$ nodes.

Thus, for both a ring of 2^d nodes and a torus of $2^{d_1} \times 2^{d_2} = 2^d$ nodes, the ST broadcast algorithm uses one wavelength and has cost

$$d(\alpha + L\tau) \quad (1)$$

There are two other communication problems whose solutions are used frequently in the broadcast algorithms given here. Solutions to these two problems, based on spanning trees, are now described.

3.1.2 Distribute algorithm

Some of the algorithms that follow make use of the *distribute* operation in which one node sends a distinct message to each other node in the network. This operation is also called *scatter* or *one-to-all personalized communication* [5]. In all of the broadcast algorithms of Section 3.2, the messages to be distributed, called *packets*, will arise by partitioning the original message that is to be broadcast. In a ring or a torus of $n = 2^d$ nodes, each of the distributed packets, has length L/n , where L is the total length of the message to be broadcast. The distribute algorithm used here is based on a spanning tree: a message of length L is distributed to all the nodes in the tree by halving it at each round until each node has received its packet. For both a ring of 2^d nodes and a torus of $2^{d_1} \times 2^{d_2} = 2^d$ nodes, the cost of this distribute algorithm (using one wavelength) is

$$\sum_{i=1}^d \left(\alpha + \frac{L\tau}{2^i} \right) = d\alpha + \left(1 - \frac{1}{2^d} \right) L\tau \quad (2)$$

3.1.3 All-to-all broadcast

The other problem of interest is called the *all-to-all broadcast*. In this problem, *each* node has a message (typically, a packet of length L/n) that must be broadcast to all other nodes. This problem is easily solved by exchanging packets between nodes whose node numbers (written in binary) differ by one bit. On hypercubes this algorithm is known as a *dimensional exchange* and the same term will be used here. Note that the message length doubles at each round of this algorithm and so its cost is the same as the cost (2) of the distribute algorithm described above, provided that there is no wavelength conflict. However, on the

topologies considered in this paper a dimensional exchange can give rise to link contention, whose issue is addressed now.

In a ring of 2^d nodes, the all-to-all broadcast algorithm can be realized with 2^{d-2} wavelengths, that we number from 0 to $(2^{d-2} - 1)$, by making node numbered j communicate with wavelength numbered $[j] = j \bmod 2^{d-2}$ (see Figure 3). On the first round, each node has to send its packet to the corresponding opposite node in the ring. No wavelength conflict occurs if half of the nodes using the same wavelength communicate in the clockwise direction of the ring, while the other half of nodes use the other direction.

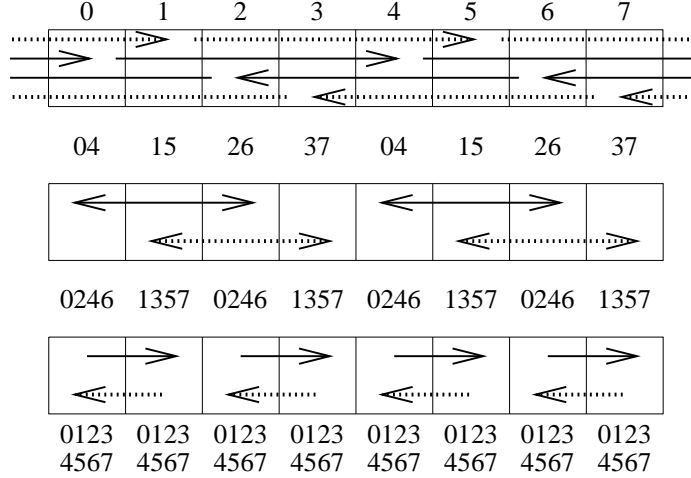


Figure 3: All-to-All broadcast algorithm using two wavelengths in a ring of 2^3 nodes.

In a torus of $2^{d_1} \times 2^{d_2}$ nodes (assuming that $d_1 \geq d_2$), the first $(d_1 - d_2)$ rounds of the all-to-all broadcast algorithm consist in exchanges of packets along the longest axis. As in the case of the ring, this can be done with 2^{d_1-2} wavelengths. Then it remains to realize a dimensional exchange in $2^{d_1-d_2}$ contiguous square blocks of $2^{d_2} \times 2^{d_2}$ nodes. Making half of the nodes communicate horizontally while the other half of nodes communicate vertically, and vice versa, as shown in Figure 4, allows to use only 2^{d_2-3} wavelengths in the last phase of the algorithm. An appropriate wavelength assignment can be easily obtained in a greedy fashion. So, in the special case of a square torus of $2^{d_1} \times 2^{d_1}$ nodes, only 2^{d_1-3} are thus sufficient for the dimensional exchange.

Thus, the all-to-all broadcast algorithm uses 2^{d-2} wavelengths for a ring of 2^d nodes, $2^{\max(d_1, d_2)-2}$ wavelengths for a torus of $2^{d_1} \times 2^{d_2} = 2^d$ nodes (only 2^{d_1-3} if $d_1 = d_2$), and has cost

$$\sum_{i=1}^d \left(\alpha + \frac{L\tau}{2^{d+1-i}} \right) = d\alpha + \left(1 - \frac{1}{2^d} \right) L\tau \quad (3)$$

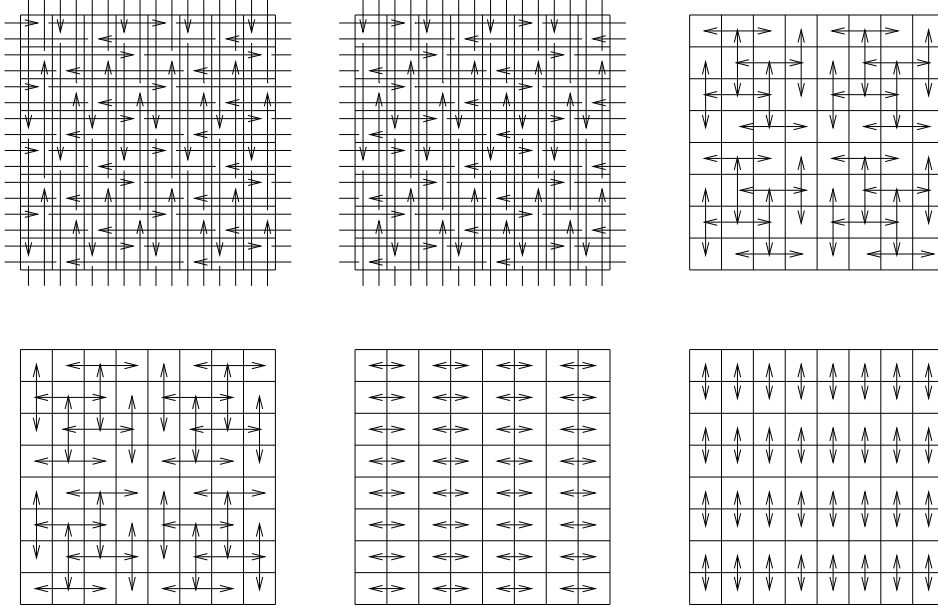


Figure 4: All-to-All broadcast algorithm using one wavelength in a torus of $2^3 \times 2^3$ nodes.

With the basic tools introduced above, we can now continue with the construction and analysis of broadcast algorithms.

3.2 Distribute-and-Exchange (DE) broadcast algorithms

The ST broadcast algorithm can be generalized to take better advantage of the available network optical bandwidth. In particular, broadcast algorithms that have bidirectional message flow are now described. They exploit the network property that messages moving in opposite directions do not contend with each other for optical communication links.

In this section we describe a family $\{\mathcal{A}(r)\}_{1 \leq r \leq d}$ of broadcast algorithms, so that $\mathcal{A}(r)$ completes in $d + r$ communication rounds in a ring or a torus of 2^d nodes. For rings, our algorithms are very similar to those given in [8] for linear arrays under a circuit-switched communication model that leads nearly to the same constraints as our WDM assumptions.

3.2.1 DE broadcast algorithms on a ring

For $1 \leq r \leq d$, the ring can be viewed as 2^r interleaved subrings each with 2^{d-r} nodes. The i^{th} of these subrings, for $0 \leq i < 2^r$, consists of nodes numbered $j \cdot 2^r + i$, for $0 \leq j < 2^{d-r}$.

To broadcast the message on these interleaved rings, the message is first distributed among nodes $0, 1, \dots, (2^r - 1)$ with one wavelength. From (2), the cost of this distribution

is

$$r\alpha + \left(1 - \frac{1}{2^r}\right) L\tau \tag{4}$$

Each of the informed nodes then acts as the source of a ST broadcast of a message of length $L/2^r$ on a subring of 2^{d-r} nodes (see Figure 5). By making the even numbered nodes transmitting to the right and the odd numbered nodes to the left, only 2^{r-1} wavelengths are sufficient. From (1), the cost of this phase is

$$(d - r) \left(\alpha + \frac{L\tau}{2^r}\right) \tag{5}$$

The final phase of the algorithm consists of collecting the packets to reconstruct the original message in each node. This is accomplished by a dimensional exchange on each contiguous subring of 2^r nodes using 2^{r-2} wavelengths. We use here the fact that at the end of the ST broadcasts the distribution of the packets is the same in each contiguous subring. Therefore, receiving the information of a node in the same subring is equivalent to receiving the information of the corresponding node in an adjacent subring. From (3), this phase has cost

$$r\alpha + \left(1 - \frac{1}{2^r}\right) L\tau \tag{6}$$

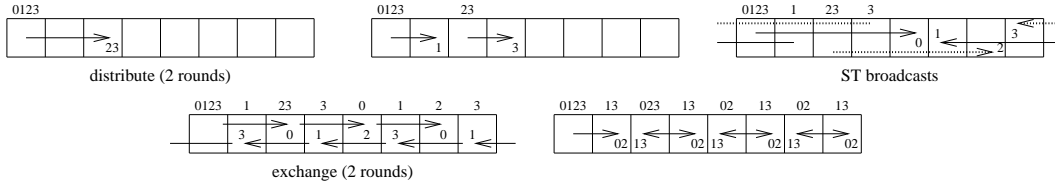


Figure 5: DE broadcast algorithm using two wavelengths in a ring of 2^3 nodes.

Thus, the DE broadcast algorithm $\mathcal{A}(r)$ in a ring of 2^d nodes uses 2^{r-1} wavelengths and its total cost is

$$(d + r)\alpha + \left(2 + \frac{d - r - 2}{2^r}\right) L\tau$$

3.2.2 DE broadcast algorithms on a torus

We describe here a family of broadcast algorithms similar to those of the previous section. They complete with the same total cost but they necessitate a smaller number of wavelengths, by taking advantage of the connectivity of the torus. As before, for $1 \leq r \leq d$ the algorithm $\mathcal{A}(r)$ uses $d + r$ rounds in a torus of $2^{d_1} \times 2^{d_2} = 2^d$ nodes and proceeds in three phases.

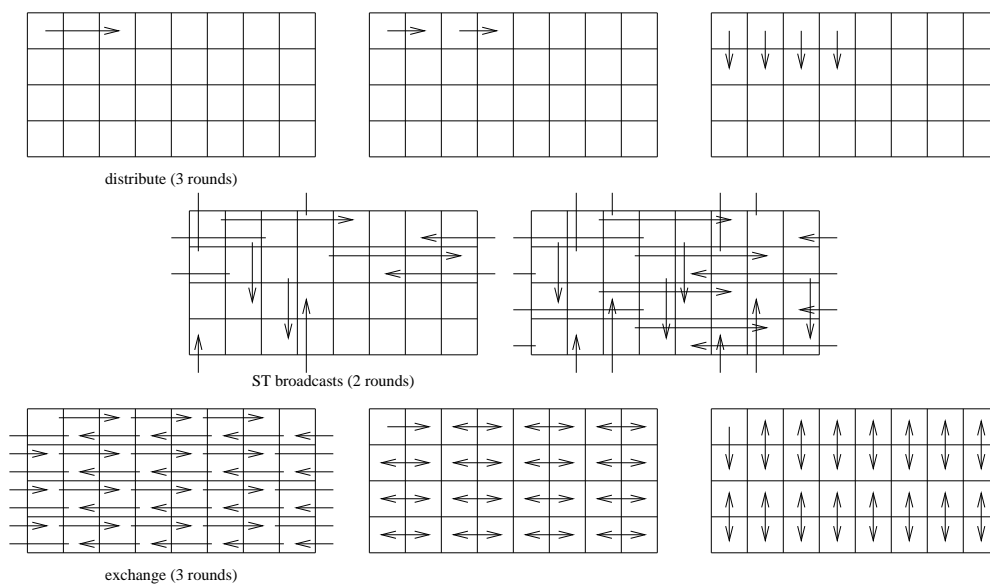


Figure 6: DE broadcast algorithm using one wavelength in a torus of $2^3 \times 2^2$ nodes.

Consider the $2^{d_1} \times 2^{d_2}$ torus as made up of 2^r interleaved subtori, each of size $2^{d_1-r_1} \times 2^{d_2-r_2}$, with $r_1 + r_2 = r$, so that each $2^{r_1} \times 2^{r_2}$ contiguous block of nodes has exactly one node from each subtorus. To have each subtorus as “square” as possible, we will choose to set $r_1 = \min(r, \lceil \frac{r-d_1+d_2}{2} \rceil)$ and $r_2 = \max(0, \lfloor \frac{r-d_1+d_2}{2} \rfloor)$ (hence $r_1 \geq r_2$), so that the difference $(d_1 - r_1) - (d_2 - r_2)$ is as small as possible.

In the first phase of the algorithm, the originator distributes the message as 2^r packets among the nodes in the $2^{r_1} \times 2^{r_2}$ block which it belongs to. See Figure 6 for an example. This necessitates only one wavelength and the same cost (4) as in a ring.

Each of the nodes in that block then acts as the root of a ST broadcast of a message of length $L/2^r$ in a subtorus of size $2^{d_1-r_1} \times 2^{d_2-r_2}$. If $(d_1 - r_1) > (d_2 - r_2)$, this phase can be realized with 2^{r_1-1} wavelengths. If $(d_1 - r_1) = (d_2 - r_2)$, alternating the orientation of the spanning trees of the 2^r subtori as in Section 3.1.3 (i.e. making one quarter of nodes communicate to the right, another downwards, another to the left and the last upwards) allows to use only 2^{r_1-2} wavelengths (See Figure 6). From Equation (1), this second phase has the same cost (5) as in a ring.

After the broadcast phase is completed, each node in each contiguous $2^{r_1} \times 2^{r_2}$ block contains one of the 2^r packets. These packets are then recombined using a dimensional exchange. As noticed before, each block can be seen as a torus of $2^{r_1} \times 2^{r_2} = 2^r$ nodes and 2^{r_1-2} wavelengths (only 2^{r_1-3} if $r_1 = r_2$) are sufficient. The last phase has the same cost (6) as in a ring.

Thus, the DE broadcast algorithm $\mathcal{A}(r)$ in a torus of $2^{d_1} \times 2^{d_2} = 2^d$ nodes uses 2^{r_1-1} wavelengths (only 2^{r_1-2} if $(r - |d_1 - d_2|)$ is non negative and even), with $r_1 = \min(r, \lceil \frac{r-d_1+d_2}{2} \rceil)$, and its total cost is

$$(d+r)\alpha + \left(2 + \frac{d-r-2}{2^r}\right)L\tau$$

3.3 Pipelined broadcast algorithms

We present in this section a generalization to rings and tori of the pipelined broadcast algorithms given in [5] for hypercubes. In addition, we provide a slight improvement on the total cost which may be also applied on hypercubes.

As before, we define a family $\{\mathcal{A}(r)\}_{r \geq d}$ of pipelined broadcast (PB) broadcast algorithms, so that $\mathcal{A}(r)$ completes in $d+r$ communication rounds in a ring or a torus of 2^d nodes. The communication pattern of all our PB algorithms are based on an embedding in rings or tori of the arc-disjoint spanning trees described in [5] for hypercubes.

However, we prefer for simplicity to describe our PB algorithms as induced by the communication pattern of the dimensional exchanges of Section 3.1.3. In consequence, the same number of wavelengths will be sufficient to ensure no wavelength conflict. It will be made use of the following property : the all-to-all broadcast is accomplished, however is chosen the order of the dimensional exchanges. This fact is easy to understand for hypercubes, where all dimensions are somewhat equivalent, and the same holds for rings and tori.

For $r \geq d$, the PB algorithm $\mathcal{A}(r)$ proceeds as follows (see Figure 7) for a ring and Figure 8 for a torus). The original message is first partitioned into $r+1$ packets, all of size $L/(r+1)$. During the first d rounds, the communication pattern of the ST broadcast is used, but the originator sends a different packet (to a different node) at each round. Each such packet is then broadcast from its first destination to all the network using a spanning tree induced by the all-to-all broadcast algorithm. During all the remaining rounds, the communication pattern of the dimensional exchanges is used cyclically. The originator continues sending a different packet at each round until round r and sends then the last packet during the last d rounds.

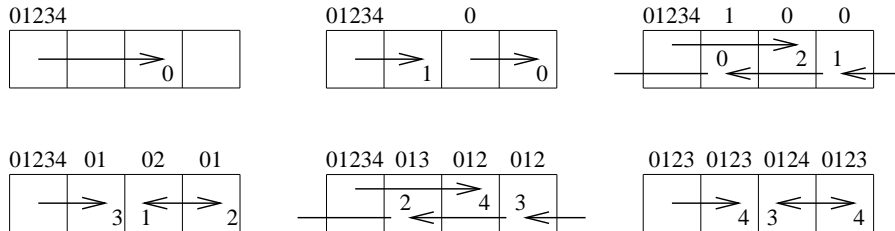
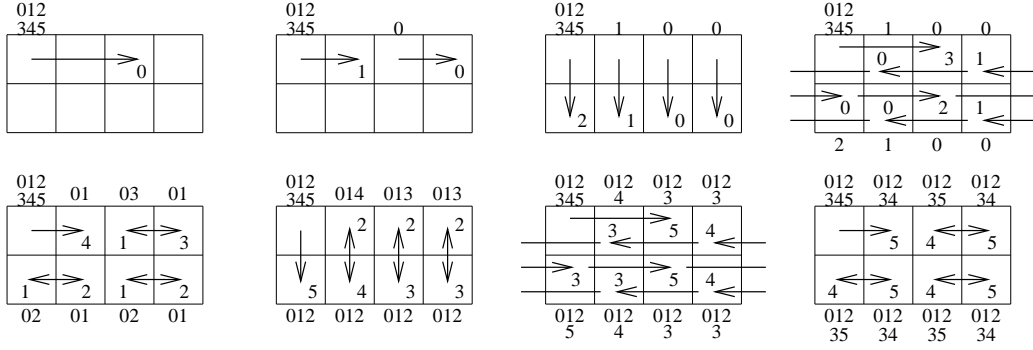


Figure 7: Pipelined broadcast algorithm in a ring of 2^2 nodes.

Figure 8: Pipelined broadcast algorithm in a torus of $2^2 \times 2^1$ nodes.

Thus, as packets of the same size $L/(r+1)$ are sent at any round, the PB algorithm $\mathcal{A}(r)$ has a total cost of

$$(d+r) \left(\alpha + \frac{L\tau}{r+1} \right)$$

For comparison, the algorithm of [5] completing in $(d+r)$ communication rounds has cost $(d+r)(\alpha + \frac{L\tau}{r})$. The slight improvement arises from the fact that in our algorithm the originator sends packets until the last round.

4 Result analysis

The following table summarizes the costs of the broadcast algorithms described in this paper and the maximum number of wavelengths that they use during a round. The costs of the algorithms designed for rings are given for $n = 2^d$ nodes while the torus algorithms are given for $n = 2^{d_1} \times 2^{d_2}$ nodes. It is assumed that the DE broadcast algorithms are defined for $1 \leq r \leq d$ and the PB algorithms for $d \leq r$.

Algorithm	Cost	Number of wavelengths
ring/torus ST	$d(\alpha + L\tau)$	1
ring DE	$(d+r)\alpha + (2 + \frac{d-r-2}{2^r})L\tau$	2^{r-1}
torus DE		2^{r_1-1} if $(r - d_1 - d_2)$ is negative or odd
		2^{r_1-2} else, with $r_1 = \min(r, \lceil \frac{r- d_1-d_2 }{2} \rceil)$
square torus DE		$2^{\lceil r/2 \rceil}$ if r is odd
	$2^{r/2-2}$ if r is even	
ring PB	$(d+r) \left(\alpha + \frac{L\tau}{r+1} \right)$	2^{d-2}
torus PB		$2^{\max(d_1, d_2)-2}$
square torus PB		$2^{d/2-2}$

Note that for $r \geq (\log_2 d + \Theta(1))$ the DE broadcast algorithms are not worthwhile. Indeed, their transmission cost has a constant coefficient close to two whenever $2^r = \Omega(d)$, and this cannot be improved unless $r \geq d - 1$.

For extremely short messages the ST broadcast algorithms have the lowest cost on both rings and tori because of their low latency. In fact, the ST broadcast is a DE broadcast for $r = 0$. However, the total cost function being convex, the best DE broadcast has lower cost than the ST broadcast if it is already the case for $r = 1$, that is, whenever

$$L > \frac{2\alpha}{(d-1)\tau}$$

For example, even in a 100Mb/sec network of $2^2 = 4$ nodes with latency $\alpha = 50\mu\text{sec}$, the ST broadcast has greater cost for broadcasting messages of length $L > 10\text{Kb}$.

In the case where sufficiently large messages are concerned, the PB algorithms presented in the previous section allow to decrease the coefficient of the transmission cost down to one, which is optimal. These algorithms are particularly interesting for applications such as broadcast TV or video conferencing.

In view of the results obtained in [2], all of our broadcast algorithms have nearly optimal transmission costs, and thus the best of them has the same optimality for the broadcast problem. More precisely, under the same communication model (actually not necessarily optical) it is shown in [2] that any broadcast algorithm completing in $d + r$ rounds has a transmission cost at least $(d/2^r + \Theta(1))L\tau$ if $1 \leq r \leq d$ and at least $(1 + o(1))(\frac{d+r}{r})L\tau$ if $r \geq d$.

In this paper we have not dealt with the issue of permutation costs that can affect our algorithms. As it can be noticed, internal data movements are required in the PB algorithms described in the previous section. Indeed, some of the nodes in the network receive, say, the packet numbered i up to d rounds after having received the packet numbered $(i + d - 1)$. A parameter representing the cost of moving data within a node should be subsequently included in the cost analysis of these algorithms. For our DE broadcast algorithms, this problem can be solved by modifying slightly the exchange phase. If the dimensional exchanges are realized in the reverse order of the one described in Section 3.1.3, then all the received packets can be concatenated within a node without permutation.

Nevertheless, we have chosen to present all the dimensional exchanges of our DE broadcast algorithms with that particular order for the following reason. If not enough wavelengths are available in the optical network, these algorithms can still be executed by dividing the not possible communication rounds into several rounds. For example, if one given round in the algorithm requires twice more wavelengths than it is available, then the same communication pattern can be realized in two rounds instead of one. Thus, it is important to have the communication cost of such rounds as small as possible, which is provided by our ordering of exchanges. Note that in our DE broadcast algorithms, only a few rounds require many wavelengths, therefore their total cost will not be significantly affected in case of wavelength limitation.

5 Conclusions

This study of the problem of broadcasting in WDM optical rings and two-dimensional toroidal meshes has yielded two families of algorithms. The well-known spanning tree (ST) broadcast algorithm was generalized by taking advantage of the increased bandwidth offered by the bidirectional WDM concurrent transmission. This gave rise first to a family of Distribute-and-Exchange (DE) algorithms having lower cost than the ST broadcast algorithm for all but the very short messages. Then was described a family of pipelined broadcast (PB) algorithms which perform better for long messages.

The additional connectivity of tori over rings has been fully used. Indeed, provided that the torus is quite square, the number of wavelengths used is approximatively the square root of that used in the ring with as many nodes.

Our algorithms require knowledge of machine-dependent constants for network latency and bandwidth to obtain good performance. However, implementations and performance measurements to be compared with the analytic predicted costs are desirable to design a more precise communication model. Further work should also generalize our algorithms to allow the number of nodes in a ring or in a dimension of a torus to be any number, not just a power of two.

References

- [1] M. Barnett, D. G. Payne, and R. van de Geijn. Optimal broadcasting in mesh-connected architectures. Technical Report TR-91-38, Dept. of Computer Sciences, Univ. of Texas, December 1991.
- [2] B. Beauquier, O. Delmas, and S. Pérennes. Tight bounds for broadcasting in the linear cost model. Technical report, Institut National de Recherche en Informatique et en Automatique, 1997. To appear.
- [3] N. K. Cheung, Nosu K., and G. Winzer. Special issue on dense WDM networks. *IEEE Journal on Selected Areas in Communications (JSAC)*, 8, 1990.
- [4] Paul E. Green. *Fiber Optic Networks*. Prentice Hall, 1993.
- [5] C.T. Ho and S.L. Johnsson. Optimum broadcasting and personalized communication in hypercubes. *IEEE Transactions on Computers*, 38(9):1249–1268, September 1989.
- [6] D. Minoli. *Telecommunications Technology Handbook*. Artech House, 1991.
- [7] B. Mukherjee. WDM-based local lightwave networks, Part II: Multihop systems. *IEEE Network Magazine*, 6(4):20–32, July 1992.
- [8] S. R. Seidel. Broadcasting on linear arrays and meshes. Technical Report ORNL/TM-12356, Oak Ridge National Laboratory, March 1993.

ANNEXE F

On arbitrary Waksman networks and their vulnerability

On Arbitrary Waksman Networks and their Vulnerability

Bruno Beauquier^{*}, Eric Darrot^{*}

Thème 1 — Réseaux et systèmes
Projet Sloop

Octobre 1999

Abstract: Motivated by problems in telecommunication satellites, we investigate rearrangeable permutation networks made of binary switches. A simple counting argument shows that the number of switches necessary to build a $n \times n$ rearrangeable network (i.e. capable of realizing all one-to-one mappings of its n inputs to its n outputs) is at least $\lceil \log_2(n!) \rceil = n \log_2 n - n \log_2 e + o(n)$ as $n \rightarrow \infty$. For $n = 2^r$, the r -dimensional Beneš network gives a solution using $n \log_2 n - \frac{n}{2}$ switches. Waksman, and independently Goldstein and Leibholz, improved these networks using $n \log_2 n - n + 1$ switches. We provide an extension of this result to arbitrary values of n , using $\sum_{i=1}^n \lceil \log_2(i) \rceil$ switches. The routing algorithm used in Beneš networks is also generalized for arbitrary values of n . Finally the fault-tolerance issue of these networks is discussed.

Key-words: switching networks, multistage networks, rearrangeable networks, permutation networks, fault tolerance, vulnerability.

(Résumé : tsvp)

Submitted to Parallel Processing Letters (World Scientific).

An extended abstract of this paper has been accepted in ALGOTEL'99, Roscoff, France, 5-7 mai 1999.

This work has been supported by Alcatel Space Industries (Toulouse, France).

^{*} SLOOP is a joint project with the CNRS and the University of Nice-Sophia Antipolis (I3S laboratory).
E-mail: {Bruno.Beauquier, Eric.Darrot}@inria.fr.

Réseaux de Waksman Généralisés et Vulnérabilité

Résumé : Motivée par des problèmes de conception de réseaux embarqués dans des satellites de télécommunications, notre étude porte sur les réseaux de permutation réarrangeables composés de commutateurs binaires. Un simple argument de comptage montre que le nombre de commutateurs nécessaires pour construire un réseau réarrangeable $n \times n$ (capable de réaliser toutes les permutations de ses n entrées vers ses n sorties) est au moins $\lceil \log_2(n!) \rceil = n \log_2 n - n \log_2 e + o(n)$, quand $n \rightarrow \infty$. Pour $n = 2^r$, le réseau de Beneš r -dimensionnel fournit une solution utilisant $n \log_2 n - \frac{n}{2}$ commutateurs. Waksman, et indépendamment Goldstein et Leibholz, ont amélioré cette construction en utilisant $n \log_2 n - n + 1$ commutateurs. Nous donnons une généralisation de ce résultat pour toute valeur de n , en utilisant $\sum_{i=1}^n \lceil \log_2(i) \rceil$ commutateurs. L'algorithme de routage utilisé dans les réseaux de Beneš est également généralisé pour notre construction. Enfin, nous traitons la tolérance aux pannes de ces réseaux.

Mots-clé : commutation, réseaux multi-étages, réseaux de permutation, réarrangeabilité, tolérance aux pannes.

1 Motivation

The following problem arises in network design for telecommunication satellites. Incoming audio or video signals are to be relayed to output amplifiers through an on-board network of waveguides and switching elements each having 4 ports. Each element can be seen as a 2×2 switch which can be set in a direct-connection state or a crossed-connection state. Furthermore, once on board, amplifiers can fail and the switches can lock in one state and remain blocked in it.

In [6, 12] it is assumed that amplifier failures can occur and that any input signal can be routed possibly to any output amplifier. In contrast, we consider in this paper the case where amplifiers do not fail and where each incoming signal needs a specific amplifier, according to the geographical areas targetted, the traffic variations and/or the compatibility between signals and amplifiers.

For many reasons (layout properties, reliability, energy saving, etc), but mainly to decrease launch costs, it is crucial to minimize the network physical weight, i.e. to minimize the number of switches, the number of links (waveguides) and their length. As launch costs are dramatically high, it is worth saving even one switch. In the case of switch blocking the problem becomes minimizing the number of additional switches needed to guarantee the satellite's purpose, despite a number of possible faulty switches determined by the expected satellite lifetime. This difficult problem is tackled at the end of this paper, in case of only one faulty switch and will be the subject of a forthcoming paper [1], in case of an arbitrary number of faults.

A practical way to realize such networks is to use the classical multistage permutation networks well studied for both telecommunication and parallel applications. In section 2 we recall some definitions on rearrangeable permutation networks and the known results. In section 3 we present a constructive way of building an arbitrary size permutation network and we give the associated routing algorithm. This improves on previous results in terms of the switch count, as shown by Table 1, and by computation in section 4. Finally in section 5 we extend our results in case of one faulty switch.

2 Rearrangeable Permutation Networks

A *switching network* is an arrangement of switches and transmission links allowing some input terminals (or signals) to be connected to some output terminals (or amplifiers) by edge-disjoint paths. Such a network can potentially perform all or only some of the possible connections of its inputs to its outputs, but is usually restricted to one-to-one connections. A switching network with n inputs and n outputs is said to be a *rearrangeable permutation network* if, for any one-to-one mapping π of the inputs on the outputs, there is a set of edge-disjoint paths connecting the input i to the output $\pi(i)$, for each $1 \leq i \leq n$.

As specified in the introduction we restrict ourselves in this paper to *binary permutation networks* constructed solely from binary switches. However, our results can be easily

extended to k -ary permutation networks where every switch can realize any mapping of its k inputs to its k outputs.

Since one binary switch has two possible settings, s switches have 2^s settings. Note that this does not imply that s switches will realize 2^s distinct network mappings, as different settings can sometimes produce the same network mapping. But at least $\lceil \log_2(n!) \rceil$, or at least $n \log_2 n - 1.443n$ from Stirling's formula, binary switches are needed to realize all $n!$ possible mappings of a network with n inputs and n outputs.

The first rearrangeable binary permutation network was designed by Beneš [2, 3, 4] and was based on the Clos 3-stage network [8]. The r -dimensional Beneš network can realize any permutation of its $n = 2^r$ inputs to its $n = 2^r$ outputs on edge-disjoint paths, through $2r - 1$ levels of 2^{r-1} switches, for a total of $n \log_2 n - \frac{n}{2}$ switches. Beneš networks are thus asymptotically optimal in terms of the switch count. The 3-dimensional 8×8 Beneš network is shown in Figure 1(a).

Further works by different authors [9, 10, 11, 13] have shown that one switch could be spared at each step of the recursive construction of Beneš networks, without affecting its rearrangeability. We call Waksman networks the corresponding optimized networks, having $n \log_2 n - n + 1$ switches, n being a power of two. The case $n = 8$ is shown in Figure 1(b).

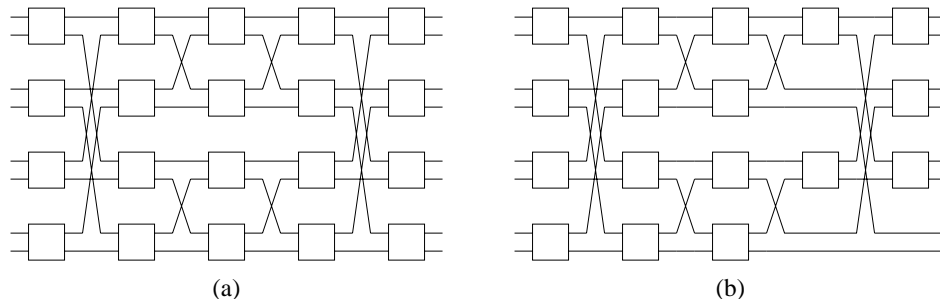


Figure 1: A 8×8 Beneš network (a) and a 8×8 Waksman network (b)

Recently Chang and Melhem [7] gave a generalization of Beneš networks for any number n of inputs/outputs. Although they wish to reduce the number of switches used, their construction does not take into account the possibility of sparing one switch in the even case of the recursion, that we present here.

We call arbitrary size Waksman networks (AS-Waksman) our permutation networks and $S(n)$ denoting the number of switches used for any number n of terminals, we obtain :

$$S(n) = S\left(\left\lceil \frac{n}{2} \right\rceil\right) + S\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n - 1 \quad (1)$$

$$= \sum_{i=1}^n \lceil \log_2(i) \rceil \quad (2)$$

Note that this result is mentioned in a footnote in [11] (page 447) and is attributed to M. W. Green, no publication being known though. Table 1 presents the values obtained by the different constructions.

n	Lower bound	Beneš [2]	AS-Beneš [7]	Waksman [13]	AS-Waksman
2	1	1	1	1	1
3	3		3		3
4	5	6	6	5	5
5	7		8		8
6	10		12		11
7	13		15		14
8	16	20	20	17	17
9	19		22		21
10	22		26		25
11	26		30		29
12	29		36		33
13	33		39		37
14	37		44		41
15	41		49		45
16	45	56	56	49	49
32	118	144	144	129	129

Table 1: Comparative switch count.

3 Network Construction

Three binary switches can be used to construct a 3×3 permutation network as shown in Figure 2(a). This network can be seen as being built from a 2×2 network (a switch) and a 1×1 network (a link). Figure 2(b) shows the 4×4 Waksman network using five binary switches. It can be seen as being built from two 2×2 permutation networks.

The procedures used to construct these networks can be generalized to recursively construct a network of any size. Specifically, a $n \times n$ AS-Waksman is constructed recursively from an AS-Waksman of size $\lceil \frac{n}{2} \rceil$ and an AS-Waksman of size $\lfloor \frac{n}{2} \rfloor$. When n is even, the construction is similar to that of Waksman. The n inputs are connected to $\frac{n}{2}$ switches and each switch is connected to two AS-Waksman networks of size $\frac{n}{2}$. Any $n - 2$ outputs are connected to

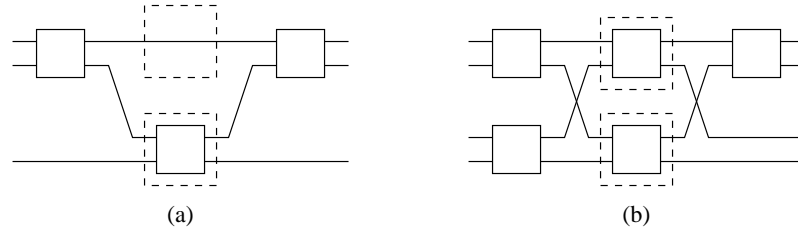


Figure 2: A 3×3 AS-Waksman network (a) and a 4×4 Waksman network (b)

$\frac{n}{2} - 1$ switches and each of them is connected to the two AS-Waksman subnetworks. The last two outputs are connected directly to the two subnetworks as shown in Figure 3(a).

To construct an AS-Waksman when n is odd, any $n - 1$ inputs are connected to $\lfloor \frac{n}{2} \rfloor$ switches and each switch is connected to the AS-Waksman of size $\lfloor \frac{n}{2} \rfloor$ and the AS-Waksman of size $\lceil \frac{n}{2} \rceil$. Similarly, any $n - 1$ outputs are connected to $\lfloor \frac{n}{2} \rfloor$ switches and each switch is connected to the two AS-Waksman subnetworks. The last input and the last output are connected directly to the AS-Waksman of size $\lceil \frac{n}{2} \rceil$ as shown in Figure 3(b).

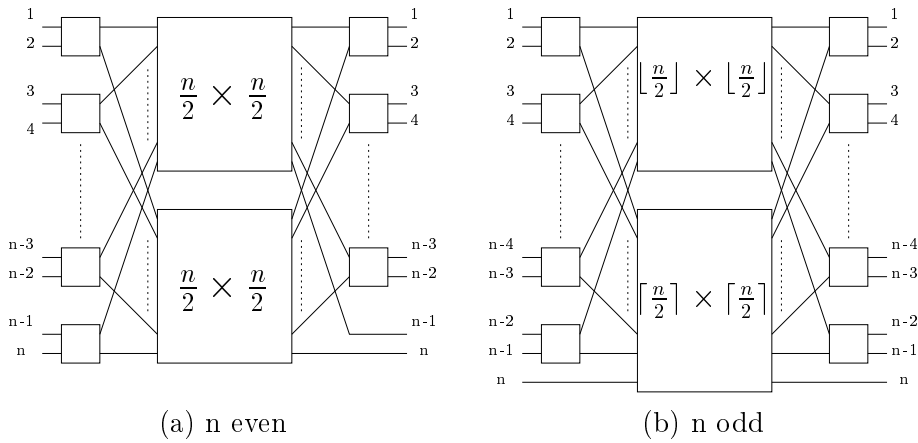


Figure 3: General construction of AS-Waksman networks

This recursive process is illustrated in Figure 4 where a 9×9 AS-Waksman is built from a 4×4 AS-Waksman and a 5×5 AS-Waksman. In general, a $n \times n$ AS-Waksman network may be constructed in this way, for any n .

Remark : Notice that our connection rules allow many possible constructions, according to which and how inputs/outputs are connected, and thus they define actually a family of AS-Waksman networks. Nevertheless, in the figures and for the main proof, we suppose for convenience that inputs $2k - 1$ and $2k$ are connected to the same switch for $1 \leq k \leq n/2$,

as well as outputs $2l - 1$ and $2l$ for $1 \leq l \leq (n - 1)/2$, while the remaining connections are direct as shown in Figure 3.

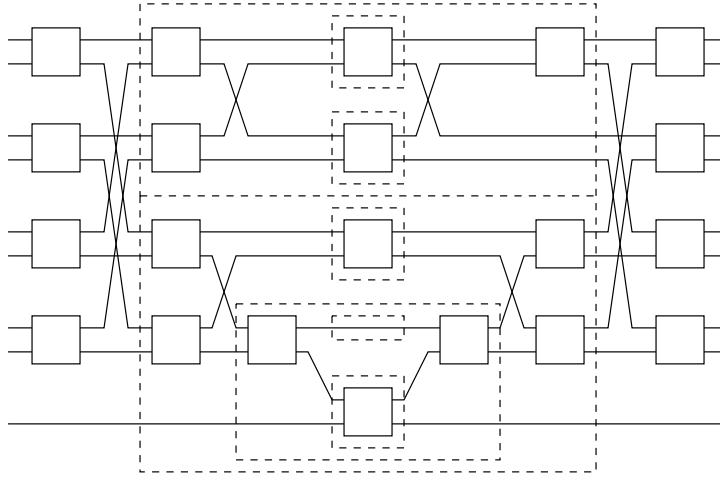


Figure 4: A 9×9 AS-Waksman network

Theorem 1 *Given any one-to-one mapping π of n inputs to n outputs in a $n \times n$ AS-Waksman network, there is a set of edge-disjoint paths from the inputs to the outputs connecting input i to output $\pi(i)$ for $1 \leq i \leq n$.*

Proof. The proof is by induction on n . If $n = 1$ or $n = 2$, the AS-Waksman network consists of a single link or a single binary switch respectively and the result is obvious. Hence, we assume that the result is true for the AS-Waksman networks of sizes $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$. The key of the induction is to observe that the middle part of an AS-Waksman network comprises two AS-Waksman subnetworks. Hence, it will be sufficient to decide whether each path is to be routed through the upper or the lower subnetwork.

The only constraints that we have on whether paths use the upper or lower subnetworks are that paths from inputs $2k - 1$ and $2k$ must use different subnetworks for $1 \leq k \leq n/2$, as well as paths to outputs $2l - 1$ and $2l$ for $1 \leq l \leq (n - 1)/2$. This is because each switch on the first and last levels of the AS-Waksman network has precisely one connection to each of the upper and lower subnetworks. There is no choice when either the input or the output is connected directly to a subnetwork.

The routing problem to realize any given permutation π can be reduced to a bipartite edge coloring problem as follows. If n is even then let $n = 2p$ and if n is odd then let $n = 2p - 1$. Consider the bipartite multigraph $G_\pi = (V, E)$ with vertex set $V = \{u_k, v_k\}_{1 \leq k \leq p}$ and with edge $[u_k, v_l] \in E$ if there are $i \in \{2k - 1, 2k\}$ and $j \in \{2l - 1, 2l\}$ such that $\pi(i) = j$. From this definition, it follows that G_π has maximum degree 2. It is a classical result from

graph theory (see e.g. [5]), known as König-Hall's theorem, that the edges of a bipartite multigraph of maximum degree Δ can be colored using exactly Δ colors, so that adjacent edges are assigned different colors.

Now we assume that the edge coloring problem is solved for G_π using two colors and we show how this leads to a routing of the permutation π in the $n \times n$ AS-Waksman network. Let color 1 be the color assigned to the edge $[u_{\lceil \pi^{-1}(n)/2 \rceil}, v_p]$. Note that the path connecting input $\pi^{-1}(n)$ to output n must use the lower subnetwork. Thus, for each i we decide to route the path from input i to output $\pi(i)$ through the lower subnetwork if the edge $[u_{\lceil i/2 \rceil}, v_{\lceil \pi(i)/2 \rceil}]$ has color 1 and through the upper subnetwork if it has color 2. In case of n odd, another constraint is that the path connecting input n to output $\pi(n)$ must use the lower subnetwork. However, the edge $[u_p, v_{\lceil \pi(n)/2 \rceil}]$ has necessarily color 1, due to a parity argument.

In this manner, all paths are assigned the upper or lower subnetworks without conflict, i.e. we can set the switches at the first and last levels of the AS-Waksman network so that both ends of every path are connected to the same subnetwork. The remainder of the path routing and switch setting is handled by induction in the subnetworks. Hence, we have established the inductive hypothesis, thereby proving the theorem. \square

As an example, we illustrate the routing algorithm in Figure 5 for the mapping

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 8 & 7 & 6 & 9 & 1 & 4 & 3 & 2 & 5 \end{pmatrix}$$

in a 9×9 AS-Waksman network. The associated bipartite multigraph G_π is drawn on the left and the edge coloring induces the subnetwork assignment for each connection between i and $\pi(i)$. The routing inside each subnetwork can be obtained recursively by the same procedure.

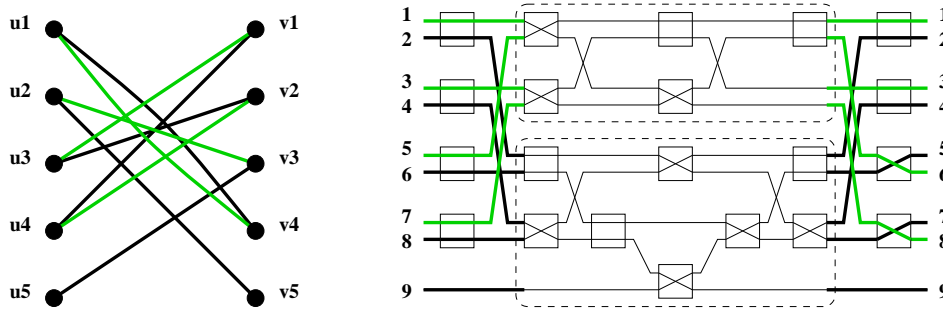


Figure 5: Routing a permutation in a 9×9 AS-Waksman by bipartite edge coloring.

4 Switch Count

In this section, we compute the number of binary switches used for a $n \times n$ AS-Waksman network, that we denote by $S(n)$.

Theorem 2 *For any $n \geq 1$, the number of binary switches used for the rearrangeable $n \times n$ AS-Waksman permutation network is $S(n) = \sum_{i=1}^n \lceil \log_2(i) \rceil$.*

Proof. The proof is by induction on n . As $S(1) = 0$, the result is true for $n = 1$. From the recursive construction given in the previous section, we have for any $n \geq 2$,

$$S(n) = S\left(\left\lceil \frac{n}{2} \right\rceil\right) + S\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n - 1$$

Note that for any two positive integers i and r , if $2^{r-1} < i \leq 2^r$ then $2^r < 2i - 1 < 2i \leq 2^{r+1}$, hence $\lceil \log_2(2i - 1) \rceil = \lceil \log_2(2i) \rceil = \lceil \log_2(i) \rceil + 1$. Thus, recalling that $\log_2(1) = 0$, we have :

$$\begin{aligned} S(n) &= \sum_{i=1}^{\lfloor n/2 \rfloor} \lceil \log_2(i) \rceil + \sum_{i=1}^{\lfloor n/2 \rfloor} \lceil \log_2(i) \rceil + n - 1 \\ &= \sum_{i=2}^{\lfloor n/2 \rfloor} (\lceil \log_2(i) \rceil + 1) + \sum_{i=1}^{\lfloor n/2 \rfloor} (\lceil \log_2(i) \rceil + 1) \\ &= \sum_{i=2}^{\lfloor n/2 \rfloor} \lceil \log_2(2i - 1) \rceil + \sum_{i=1}^{\lfloor n/2 \rfloor} \lceil \log_2(2i) \rceil = \sum_{j=1}^n \lceil \log_2(j) \rceil \end{aligned}$$

The inductive hypothesis being satisfied, the proposition holds. \square

Corollary 3 *For any $n \geq 1$, $S(n) \leq n \log_2 n - 0.91n + 1$.*

Indeed, a different computation approach can give a more expressive value : $S(n) = n \log_2 n - n(2^\lambda - \lambda) + 1$, where $\lambda = \lceil \log_2(n) \rceil - \log_2 n$. This allows to estimate the worst case obtained for $\lambda = -\log_2(\log_2 e) \sim 0.53$. Note that when n is a power of two, it can be computed by induction that $S(n) = n \log_2 n - n + 1$.

5 Fault-Tolerance Issue

Now we consider the possibility for switches to lock and remain definitively in the same state. A switching network with n inputs/outputs is said to be a k -tolerant permutation network if, for any one-to-one mapping π of the inputs on the outputs and for any k switches blocked each in any state, there is a set of edge-disjoint paths connecting the input i to the output $\pi(i)$, for each $1 \leq i \leq n$. We denote by $N(n, k)$ the minimum number of switches for such a network. In this paper we restrict ourselves to 1-tolerant permutation networks.

Theorem 4 For any $n \geq 1$, $N(2n, 1) \leq 2N(n, 1) + 2n$.

Proof. The proof is based on the construction illustrated in Figure 6. Both the $2n$ inputs and the $2n$ outputs are connected two by two to $2n$ switches and each switch is connected to two 1-tolerant permutation networks of size n . We claim that this $2n \times 2n$ network, made of two $n \times n$ 1-tolerant networks using each $N(n, 1)$ switches, is a 1-tolerant permutation network.

Actually we have a stronger result : any permutation can be realized despite one faulty switch in each subnetwork, plus one faulty switch at the first or last level. Indeed, in this case each subnetwork can still handle by assumption any mapping of its n inputs to its n outputs. Moreover, one switch locking at the first or last level corresponds to have the two associated inputs or outputs directly connected to the two subnetworks, similarly to the construction of even AS-Waksman networks. Thus, the arguments used in the proof of Theorem 1 can be easily adapted to decide through which subnetwork each path is to be routed, thereby proving our claim. \square

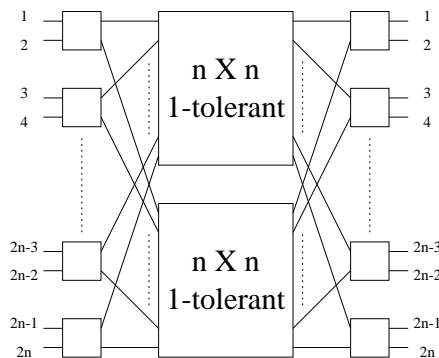


Figure 6: Construction of even 1-tolerant permutation networks.

Unfortunately, no similar good construction is known for the odd case. The last result is thus interesting only to build a twice larger 1-tolerant permutation network when a good one is already known. For this purpose we present now in Figure 7 some 1-tolerant networks obtained by hand for $n = 3, 4, 5, 6$. For space reasons proofs are omitted but can be done by restricted case analysis. These constructions together with the proof of Theorem 4 give rise to the summary of results presented in Table 2.

Size	2	3	4	5	6	8	10	12	16
Switches	2	5	7	11	14	22	32	40	50

Table 2: Number of switches for 1-tolerant permutation networks.

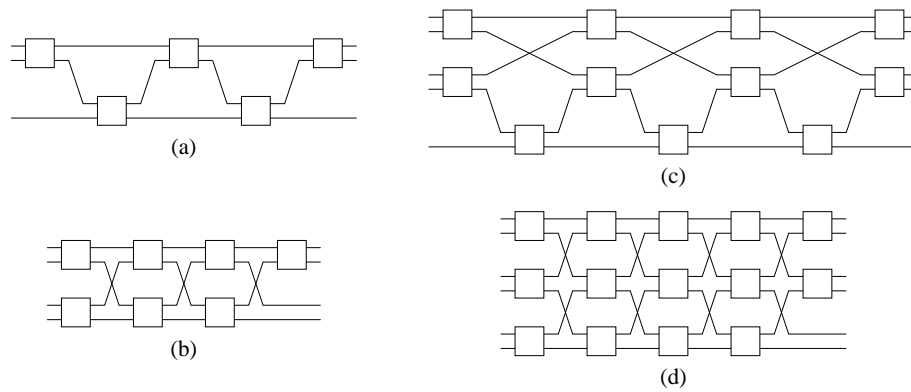


Figure 7: 1-tolerant permutation networks for $n = 3$ (a), $n = 4$ (b), $n = 5$ (c) and $n = 6$ (d).

6 Conclusion

This paper has provided a simple way for building rearrangeable permutation networks of arbitrary size n using $\sum_{i=1}^n \lceil \log_2(i) \rceil$ binary switches. An efficient algorithm for routing any permutation in such networks has been given. The fault-tolerance issue has also been tackled in the case of one locked switch. It is worth pointing out that all these results can easily be adapted for rearrangeable networks using p -ary switching elements.

In a forthcoming paper [1], the fault-tolerance results will be extended by presenting constructive ways for building k -tolerant permutation networks, for any number k of switch faults.

References

- [1] B. Beauquier and E. Darrot, Rearrangeable switching networks tolerant to switch locks, Manuscript (title may change).
- [2] V. Beneš, Permutation groups, complexes, and rearrangeable multistage connecting networks, *Bell Sys. Tech. Journal* **43** (1964) 1619–1640.
- [3] V. Beneš, *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York (1965).
- [4] V. Beneš, Applications of group theory to connecting networks, *Bell Sys. Tech. Journal* **54** (1975) 407–420.
- [5] C. Berge, *Graphs*, North-Holland (1985).
- [6] J-C. Bermond, E. Darrot and O. Delmas, Design of fault-tolerant on-board networks in satellites, Manuscript.

- [7] C. Chang and R. Melhem, Arbitrary size Beneš networks, *Parallel Processing Letters* **7** (No. 3, Sep. 1997) 279–284.
- [8] C. Clos, A study of non-blocking switching networks, *Bell Sys. Tech. Journal* **32** (1953) 406–424.
- [9] L. Goldstein and S. Leibholz, On the synthesis of signal switching networks with transient blocking, *IEEE Transactions on Electronic Computers* **16** (1967) 637–641.
- [10] A. Joel, On permutation switching networks, *Bell Sys. Tech. Journal* **47** (1968) 813–822.
- [11] W. Kautz, K. Levitt and A. Waksman, Cellular interconnection arrays, *IEEE Transactions on Computers* **17** (No.5, May 1968) 443–451.
- [12] S. Pérennes and D. Tóth, Design of fault-tolerant flow networks, Manuscript.
- [13] A. Waksman, A permutation network, *Journal of the ACM* **15** (1969) 159–163.

Résumé

Les résultats obtenus dans cette thèse portent principalement sur l'étude des **communications dans les réseaux optiques par multiplexage en longueur d'onde**. Ils s'inscrivent dans une thématique d'allocation des ressources en vue de réaliser des communications dans un réseau.

La problématique générale que nous avons considérée peut se résumer de la manière suivante. Il s'agit de satisfaire dans un réseau optique une famille de requêtes de connexion, appelée *instance* de communication et formée de couples de nœuds (source, destination). La satisfaction d'une requête passe par l'attribution d'un chemin dans le réseau et d'une longueur d'onde sur les liens utilisés, avec la contrainte que deux requêtes ne peuvent pas utiliser le même lien avec la même longueur d'onde. L'objectif dans ce cadre est de minimiser l'utilisation des ressources optiques, c'est-à-dire le nombre total de longueurs d'onde permettant de satisfaire l'instance donnée.

Dans le **chapitre 1**, nous présentons la technologie optique pour les télécommunications, afin de préciser le cadre technique de notre recherche et d'aider le lecteur informaticien à la compréhension des contraintes physiques sous-jacentes à la modélisation théorique. Dans le **chapitre 2**, nous posons la problématique étudiée au cours de la thèse et nous donnons la modélisation qui a servi de base à nos recherches. Le **chapitre 3** est une synthèse des résultats obtenus dans la littérature concernant principalement le problème du routage optique. Le reste de la thèse est constituée des annexes qui rassemblent les articles publiés, dans le format des rapports de recherche.

Mots clés : Communications, Réseaux optiques, Multiplexage en longueur d'onde (WDM), Routage, Coloration, Théorie des graphes, Mathématiques discrètes, Algorithmique, Flots, Multicast, Échange total.

Abstract

This thesis is mainly devoted to the study of **communication in wavelength multiplexed optical networks**, from which arise algorithmic and graph theoretic problems. The main issue in WDM (Wavelength Division Multiplexed) all-optical communication networks can be stated as follows. Connection requests between pairs of source-destination network nodes are to be established, by assigning each of them a colored path, so that no two paths use the same link with the same color. The goal is then to minimize the total number of colors (i.e. wavelength) used.

The first part made of three chapters is written in french, while the second part contains six articles given in annex. The first chapter contains an exposition of the basic physical and engineering principles on which modern optical networks are based. In the second chapter are presented the formulations of the basic problems in the area, their formalization as algorithmic and graph theoretic problems, and the relations between them. The third chapter contains a survey of the most relevant work in the area.

Keywords: Communication, Optical networks, Wavelength Division Multiplexing (WDM), Routing, Coloring, Graph theory, Discrete mathematics, Algorithmic, Flows, Multicast, Total exchange.