Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 EECS Building
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 97-048

Designs for Wavelength Division Multiplexing Lightwave Networks with Tunable Transceivers

by: Feng Cao
(Ph.D. Thesis)

DESIGNS FOR WAVELENGTH DIVISION MULTIPLEXING LIGHTWAVE NETWORKS WITH TUNABLE TRANSCEIVERS

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

FENG CAO

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Ding-Zhu Du, Adviser

September 1997

Abstract

The need of high speed networks, for applications incorporating high performance distributed computing, multimedia communication and real time network services, has provided the impetus for the study of optical networks. Wavelength Division Multiplexing (WDM) has been used widely for studying the throughput performance of optical networks. We studied WDM lightwave networks with tunable transceivers including designs for lightwave networks with limited tuning ranges for transceivers.

Transmission schedules and virtual topology embeddings are needed to support high performance distributed computing. How to design the transmission schedule depends on how the virtual topology is embedded in the physical lightwave network. We developed general graph theoretic results and algorithms and using these built optimal embeddings and optimal transmission schedules for de Bruijn graphs and undirected de Bruijn graphs, assuming certain conditions on the network parameters. We proved our transmission schedules are optimal over all possible embeddings.

Partitioned Optical Passive Stars(POPS) topology is a physical architecture to scale up local optical passive star networks. POPS data channel can be efficiently utilized for random permutation-based communication patterns. Reliability is important for such a scaled-up network. We analyzed the fault tolerant routing properties of POPs networks. We demonstrated some worst cases due to link errors and the lower bound for connectivity is obtained. Some sufficient approaches were proposed to detect and keep connectivity of the whole system.

The current technology only allows the transceivers to be tunable in a small range, a fact ignored in previous studies. We focused on the design of WDM optical passive star networks with tunable transmitters of limited tuning range and

fixed wavelength receivers. The limited tuning range has big effects on the maximum delay, the total number of wavelengths which can be used, and the topological embedding. We proposed efficient communication protocols for systems with limited tuning ranges. Different network topologies were analyzed in our study. The relationship between the total number of wavelengths which can be utilized and the embedded topology is established. The optimal embedding algorithms are given for the systems embedded with different virtual topologies.

Acknowledgements

First of all, I would like to thank my advisor Professor Ding-Zhu Du for his invaluable guidance and tutelage at every stage of this thesis. He introduced me to exciting research areas and his continuous support and encouragement make me focus on the research work.

Special thanks go to Professor David H. Du, Dr. A. Pavan and Dr. Al Borchers. Their suggestions and advice have been invaluable in finishing this research work. It has been my great pleasure to work with them.

I am grateful to Department of Computer Science of University of Minnesota for its environment and resources which made this research possible.

At last, I would like to thank my wife Jie for her love and support, and my parents for their love and inspiration.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Fiber Optical Networks

The need of high speed networks, for applications incorporating high performance parallel computing, multimedia communication and real time network services, has provided the impetus for the study of optical networks. Optical fibers are attractive for their high bandwidth, immunity to electromagnetic radiation, and security. The available bandwidth for optical fibers is up to 30 THZ. Optical fibers are not only the medium of choice for wide area networks, but are also being deployed extensively in local area environments. However, while optical networks present new possibilities for high speed networks, they also present new network design problems that must be surmounted before they become practical.

Wavelength Division Multiplexing (WDM) has been used widely for studying the throughput performance of optical networks [4, 10, 21, 22, 42, 47, 48, 53, 56, 57], especially those employing optical passive star couplers [7, 11, 30, 39, 40, 58] . The reason is that electronically operated interface devices can not match the speed of the optical transmission media. WDM is used to divide the total available bandwidth into different wavelengths, each running at a speed compatible with electronic devices. Different wavelengths are used for transmission in the optical fiber simultaneously with enough spacing between adjacent wavelengths. This makes WDM attractive for optical networks since they alleviate the optical-to-electronic bottleneck. In this study, we use the terms wavelength and channel interchangeably.

Time Division Multiplexing (TDM) is another approach which uses the time slots to assign the bandwidth to different applications in optical networks. Time Wavelength Division Multiplexing (TWDM) is the mix of TDM and WDM, which receives more and more attention [4, 42].

An optical passive star coupler is one of the most common ways to interconnect an optical network via optical fibers [16, 18, 24, 31, 20, 25, 30, 34, 35, 36, 54, 55]. Lots of optical networks are based on star couplers [26]. Figure 1.1 shows an optical passive star network with N stations. Each station sends its signal to the passive star coupler through its transmitter on a specific channel. In the passive star coupler, all the signals are combined and broadcast to all the stations. The signals in a certain channel can be received by the stations whose receivers occupy the corresponding channel. There are a number of advantages for using an optical passive star network. First, it is very simple and is completely passive. Second, its scalability is very good up to a certain number of stations. Lastly optical passive stars offer flexibility of topological embeddings. Different topological embeddings are required for high performance distributed computing and application topology dependent communications.

An issue in an optical passive star network is whether the network is single-hop or multi-hop. In a single-hop architecture [50], fast tunable transceivers are employed and the wavelength assignments are performed on a per-packet basis. When the sending station knows the wavelength occupied by the receiver of the receiving station, the sending station need only tune its transmitter to the receiver's wavelength and send the message. The transmitting station thus needs one hop to communicate with any receiving station. In a multi-hop architecture [37, 38, 47, 48], fixed or slowly-tunable transceivers are employed. Because the sending station may not be able to tune its transmitters to the wavelength occupied by the receiver of the receiving station directly, the sending station may need several hops to communicate with one receiving station. There are tradeoffs between single-hop and multi-hop architectures. For example, single-hop architectures introduce more overhead for each packet due to transmitter-receiver coordination and the tuning time, while multi-hop architectures introduce higher average end-to-end delay.

1.2 Transmission Schedule and Virtual Topology Embedding

To support high performance distributed computing, transmission schedules and virtual topology embeddings must be provided to efficiently utilize the high bandwidth of optical fibers. The embedding of a virtual topology and the transmission schedule in the lightwave network embedded with the virtual topology are closely related. How to design the transmission schedule depends on how the virtual topology is embedded in the physical lightwave network. Different embeddings of the virtual topology will induce different transmission schedules.

We want to embed a virtual network topology, or graph, in a physical network. This entails mapping the virtual vertices of the topology to the physical nodes of the network. Given this embedding, we then want to schedule transmissions in a repeating cycle, so that during each cycle a packet is transmitted along each edge of the virtual topology. This transmission cycle is called *all-to-neighbor Transmission.* The reason is that each node/station in the system has a dedicated slot to talk to each of its neighbors.

All-to-neighbor transmission is one of most common communication patterns in distributed computing, parallel computing, network protocols and network operation and maintain(OAM). In distributed computing, if one-to-all broadcasting is done from one node to all other nodes along a spanning tree, it is actually an *all-to-neighbor transmission* for a directed tree (people often call this tree top-down for the transmission starts from the root and passes down). Similarly, if all-to-one reduction is done along a spanning tree for distributed computing, it is an *all-to-neighbor transmission* for a reverse directed tree (people often call this tree bottom-up for the transmission starts from the leaf nodes and passes up). In a lot of communication protocols, flow specifications and information about availability of resources are needed for better integrated network services. An Efficient all-to-neighbor transmission can improve the performance of communication protocols. Some examples are teleconferencing and Available-Bit-Rate Asychronous Transmission Mode(ABR ATM). In network operation and maintain, each station or router wants to know the status of each of its neighbors and send out its status and requirements to its neighbors. *all-to-neighbor Transmission* can be used to exchange the information and reduce bandwidth for this kind of operations.

How to design an *all-to-neighbor transmission* is more important for high speed networks, especially for fiber optical networks. If we want to take advantage of the huge bandwidth of fiber channels, the transmission cycle for any kind of communication patterns should be finished as soon as possible in a TDWM lightwave network. That means the time slots for all-to-neighbor transmission cycle should be as small as possible. This makes bandwidth available for other applications or another round of all-to-neighbor transmission.

The first objective of our study is to find the optimal transmission schedule, i .e. the minimum number of time slots for each cycle. The second objective is to minimize the number of tuning times of tunable transmitters during each cycle. The minimum number of time slots for each cycle means the best utilization of lightwave networks bandwidth. The fewer number of tuning times may reduce the complexity of the transmission schedule. If tuning time is large, fewer tunings give a shorter overall schedule. We can ask for an optimal transmission schedule given an embedding, or even better, we would like an optimal transmission schedule over all possible embeddings. The transmission schedules we give are optimal over all possible embeddings.

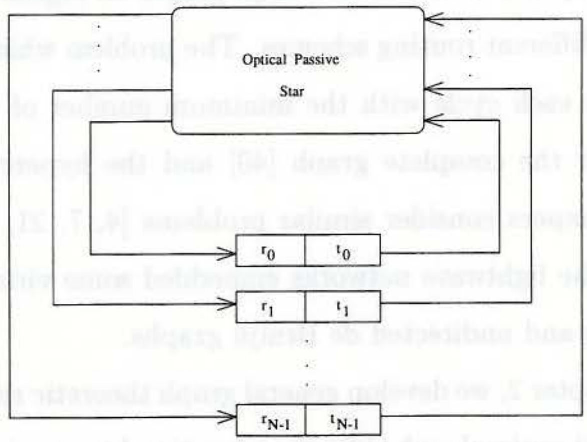
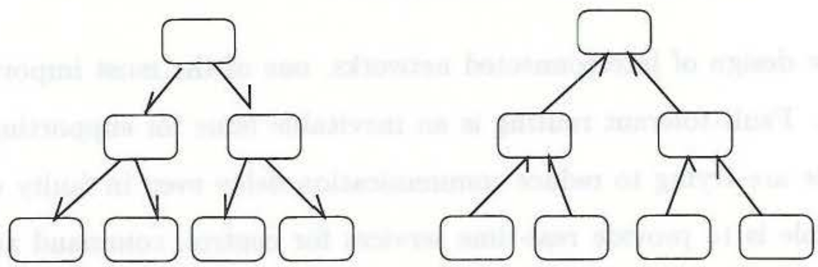


Figure 1.1: A N−station Optical Passive Star Network



one-to-all broadcasting all-to-one reduction

Figure 1.2: An example for all-to-neighbor transmission

There is previous work related with the lightwave networks embedded de Bruijn graphs or other topologies. Sivarajan and Ramaswami [53] studied the throughput and delay performance of de Bruijn graphs as logical topologies in lightwave networks under different routing schemes. The problem which is to find the transmission schedule with each cycle with the minimum number of time slots has already been considered for the complete graph [40] and the hypercube [39] virtual topologies. Many other papers consider similar problems [4, 7, 21, 42, 57, 58]. We study this problem for the lightwave networks embedded some virtual topologies, especially de Bruijn graphs and undirected de Bruijn graphs.

In Chapter 2, we develop general graph theoretic results and algorithms and using these build optimal embeddings and optimal transmission schedules for de Bruijn graphs and undirected de Bruijn graphs, assuming certain conditions on the network parameters. We prove our transmission schedules are optimal over all possible embeddings. This is the first work studying the number of the tuning times of tunable devices for transmission schedule.

1.3 Reliability Analysis

In the design of interconnected networks, one of the most important topics is reliability. Fault-tolerant routing is an inevitable issue for supporting network services. People are trying to reduce communication delay even in faulty environments. One example is to provide real-time services for control, command and communication. The fault-tolerant routing becomes critical for making sure that time constraints are satisfied through computer networks. In this study, we will focus on fault-tolerant routing in local computer networks. The same idea can also be applied to wide area networks, such as providing efficient services through internet.

The typical approach to study routing in computer networks is to try to find the shortest path between the sending station and the receiving station. Whenever some stations are faulty on the path between the sending station and the receiving station, the management protocol has to find a way to bypass those faulty stations and set up a new path between them. Similarly, if this new path is disconnected again, a third path needs to be set up if it is possible(the network is still connected or there still exists a path between the sending station and the receiving station).

We study reliability in the all-optical Partitioned Optical Passive Stars (POPS) topology. POPS is a physical architecture to scale the traditional optical passive star couplers and explore the advantages of high noise immunity and single-hop. POPS is a non-hierarchical structure and connects several optical passive star couplers together by the help of some intermediate optical passive star couplers. It remains the property of high noise immunity, single-hop, and no intermediate electronic/optical conversions. POPS is an design without considering the power budget problems, which is flexible to extend the optical passive star networks and keep the simplicity of system.

It was shown in [32] that POPS data channel can be efficiently utilized for random permutation-based communication patterns. In most of applications in network computing and parallel computing, some common communication patterns are widely used. For example, all-to-all personalized communication is the common way to globally exchange information. Global reduction or global broadcasting is the pattern to collect data from all slaves to the master in the master-slave model. In [32], they studied four common communication patterns in TDM POPS networks which shows POPS networks are supportive for distributed/networking computing.

As a scaled-up topology, fault-tolerant routing is very important for POPS. In what situations the whole system is partitioned into several disconnected components? Is there any way to determine and fix the network partitions? There was no result about connectivity of POPS in previous study. Both the station errors and the link errors are discussed with respect to network partitions in Chapter 3. Some worst cases of disco nnected POPS networks due to link errors are analyzed for obtaining the lower bounds on the connectivity. On the other hand, several efficient approaches are proposed to detect if the network is still connected, these approaches can also be used to fix the disconnected system.

1.4 Limited Tuning Range For Tunable Transceivers

The current technology only allows the transceivers to be tunable in a small range, a fact ignored in previous studies. We analyze the effect of the limited tunable range of transceivers. Each channel requires 1-2 nm for wide bandwidth and current technology can only support 3-7 nm for large bandwidth devices. This means the reasonable tunable range can only be 3 to 7 wavelengths with current technology. Even in the near future, the tunable range is not likely to increase significantly because of technical difficulties. We study the effects of the limited tunable range of transceivers on the optical passive star network with an embedded topology.

To make our study more general, we assume that the tunable range is no more than k channels. Current technologies can allow k to be single digits. In the future, k may be larger. Our result can still be applicable for k is a parameter in this study. In an optical passive star network embedded with a virtual topology which is not a complete graph, it is assumed that any two stations connected by a link in the virtual topology should be able to communicate with each other in one hop. In an optical passive star network embedded with a complete graph, it is impossible to have a station communicating with all other stations in one hop unless the total number of wavelengths used is no more than k.

We assume that there are a total of p wavelengths available for the whole system. Each station has a tunable transmitter and a fixed receiver. Let A be a station in the system and its transmitter be tunable in the range of [a, a+k-1]. The wavelength of A's receiver must be in [a, a+k-1]. Only those nodes with the wavelength of their receivers in [a, a+k-1] can receive the message from A directly, i.e. it will take one hop to reach those nodes. For the other stations, it takes multi-hop transmission to receive the message from A. An example is shown in Figure 1.3 to see the effect for assigning wavelengths to embed a mesh into an optical network.

There are many communication protocols for optical networks [25, 18, 20, 25, 35, 43, 51, 54]. But there is no specific communication protocol about lightwave networks with limited tuning ranges for transceivers. We proposed communication protocols based on different assumptions about transmitters and receivers.

In an optical passive star network embedded with a complete graph, we study the relationship between the total number of wavelengths used and the maximum delay without congestion in the system. That means we only consider hop-counts and ignore the potential re-transmissions in our evaluation of delay. The optimal embedding algorithm that minimizes the maximum delay is given. One effect of the limited tunable range is the throughput bottleneck for some wavelengths in the system no matter how the wavelengths are assigned. This bottleneck indicates the complete graph is not suitable as the embedding topology for a system with uniform

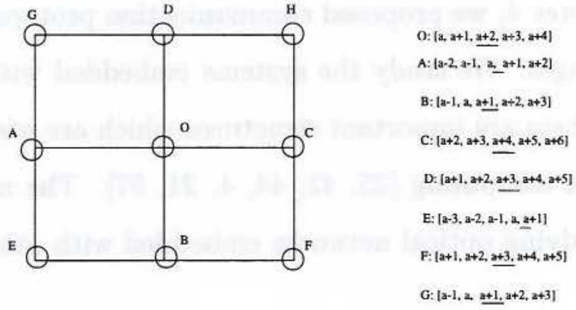


Figure 1.3: An example for wavelength assignment with k=5. The underlined number for each station means the wavelength for its receiver.

communication among all stations. But we show how one-to-all broadcasting in such a system can be efficient.

In an optical passive star network embedded with a topology other than the complete graph, the following questions need to be answered:

- How to create connections between the sending station and the receiving station? How to resolve transmission collisions in each channel?
- How to embed a virtual topology on an optical passive star network satisfying the constraint that neighboring nodes in the virtual topology are one hop away?
- What is the relationship between the topology and the total number of wavelengths which can be used? Are there any tight upper bounds on the total wavelengths which can be used?
- How to embed the topology in the optical passive star network to use as many wavelengths as possible?

The connection and collision problems also exist in the systems embedded with complete graphs. Because of the throughput bottleneck for some wavelength, we recommend that complete graph not be the topology for embedding.

In Chapter 4, we proposed communication protocols which is efficient for limited tuning ranges. We study the systems embedded with rings, meshes and hypercubes. All of them are important structures which are widely used in communication and distributed computing [25, 42, 44, 4, 21, 57]. The methods we use can also be applied for studying optical networks embedded with other virtual topologies.

Chapter 2

Transmission Schedule and Topology Embedding

We consider the problem of embedding a virtual de Bruijn topology, both directed and undirected, in a physical optical passive star time and wavelength division multiplexed (TWDM) network and constructing a schedule to transmit packets along all edges of the virtual topology in the shortest possible time. We develop general graph theoretic results and algorithms and using these build optimal embeddings and optimal transmission schedules, assuming certain conditions on the network parameters. We prove our transmission schedules are optimal over all possible embeddings.

As a general framework we use a model of the passive star network with fixed tuned receivers and tunable transmitters. Our transmission schedules are optimal regardless of the tuning time. Our results are also applicable to models with one or more fixed tuned transmitters per node. We give results that minimize the number of tunings needed. For the directed de Bruijn topology a single fixed tuning of the transmitter suffices. For the undirected de Bruijn topology two tunings per cycle (or two fixed tuned transmitters per node) suffice and we prove this is the minimum possible.

2.1 Introduction

Optical networks present new possibilities for high speed networks, but they also present new network design problems. In this chapter we study transmission schedule problems of embedding a virtual de Bruijn topology in a time and wavelength division multiplexed (TWDM) optical passive star network. The embedding of a virtual topology and the transmission schedule in the lightwave network embedded with the virtual topology are closely related. How to design the transmission schedule depends on how the virtual topology is embedded in the physical lightwave network. Different embeddings of the virtual topology will induce different transmission schedules. In our study, we first show how to embed a de Bruijn graph into an optical passivestar network. Then, based on that embedding, we design the transmission schedule. At last, we show that our transmission schedule is optimal among all the possible transmission schedules on all possible embedding of a de Bruijn graph into an optical passivestar network. This means that our design of embedding of an de Bruijn graph as the virtual topology into a lightwave network is really what we want. The tuning times of the tunables transmitters are also considered, which may be helpful to reduce the requirements for the physical devices and simplify the complexity of the optimal transmission schedules.

The first part of this chapter deals with the design of transmission schedules for a lightwave network embedded with a general topology. Some general results are shown about the low bounds of transmission schedules.

The second part and the third part of this chapter are about the optimal transmission schedule in an optical passive star network embedded an de Bruijn graph(directed or undirected).

De Bruijn graphs are used as a structure for networking computing and parallel computing. There are some well-known network topologies that belong to de Bruijn graphs and some variations of de Bruijn graphs. For example, shuffle exchange graphs are a special case of de Bruijn graphs. Ring topologies are a special case of generalized de Bruijn graphs. The study on de Bruijn can help us understand how to design transmission schedule and topology embedding for a large class of network topologies. The results can be applied to different special cases of de Bruijn graphs and undirected de Bruijn graphs.

The lightwave networks embedded de Bruijn graphs provide the possibility of high performance computing for some large problems. Transmission schedules are important for networking computing and parallel computing for solving different communication patterns. De Bruijn graphs are also used in switching networks design. For example, the shuffle exchange networks are based on de Bruijn graphs. An optical passive star embedded with a de Bruijn graph may be used as a high speed switch or a stage for communication connections. Sivarajan and Ramaswami [53] proposed de Bruijn graphs as logical topologies for multihop lightwave networks. They also proposed de Bruijn graphs as good physical topologies for wavelength routing lightwave networks consisting of all-optical routing nodes interconnected by point-to-point fiber links.

The TWDM optical passive star [10, 30, 26], as used in this chapter, is a physical network architecture with N nodes, each having a single tunable transmitter and a single fixed tuned receiver connected through an optical passive star. See Figure 1.1. The transmissions from each transmitter are routed to all receivers, but in order to communicate both transmitter and receiver must be tuned to the same wavelength. Transmissions of fixed sized packets are scheduled in synchronous time slots; only one transmission can be done on each wavelength in a single time slot. The network supports k different wavelengths, so k transmissions can be done simultaneously. The transmitters take a fixed amount of time, δ, expressed in time slots, to change their tuning from one wavelength to another. If the packet size is

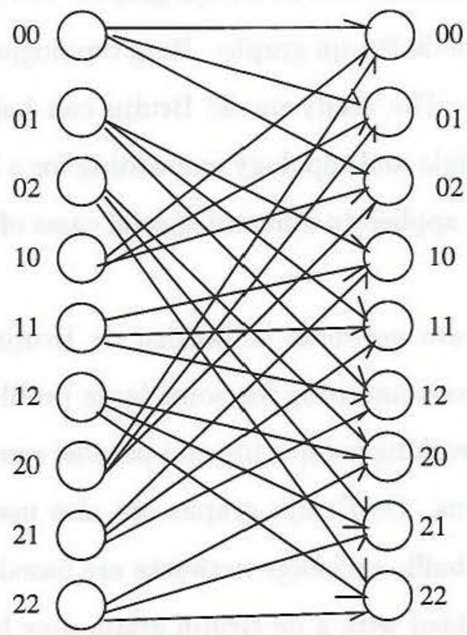


Figure 2.1: An example of de Bruijn Graph

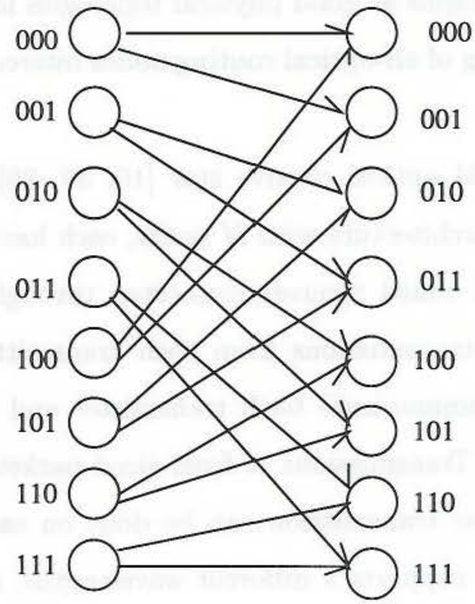


Figure 2.2: Shuffle Exchange Networks: A special case of de Bruijn Graphs

small, as it would be in an ATM model, δ will be large.

We want to embed a virtual network topology, or graph, in this physical network. This entails mapping the virtual vertices of the topology to the physical nodes of the network. Given this embedding, we then want to schedule transmissions in a repeating cycle, so that during each cycle a packet is transmitted along each edge of the virtual topology. We will consider both directed and undirected graphs for the virtual topology; for undirected graphs we need schedule packets in each direction along an edge. *All-to-neighbor transmission* is considered to be the communication pattern for this chapter.

How to design an *all-to-neighbor transmission* is more important for high speed networks, especially for fiber optical networks. If we want to take advantage of the huge bandwidth of fiber channels, the transmission cycle for any kind of communication patterns should be finished as soon as possible. That means the time slots for all-to-neighbor transmission cycle should be as small as possible. This makes bandwidth available for other applications or another round of all-to-neighbor transmission.

The first objective of our study is to find the optimal transmission schedule, i.e. the minimum number of time slots for each cycle. The second objective is to minimize the number of tuning times of tunable transmitters during each cycle. The minimum number of time slots for each cycle means the best utilization of lightwave networks bandwidth. The fewer number of tuning times may reduce the complexity of the transmission schedule. If tuning time is large, fewer tunings give a shorter overall schedule. We can ask for an optimal transmission schedule given an embedding, or even better, we would like an optimal transmission schedule over all possible embeddings. The transmission schedules we give are optimal over all possible embeddings.

Not only do we determine optimal embeddings and schedules, but we build our proofs of optimality on several general graph-theoretic results and algorithms that will be of use in the further study of similar problems.

The model we use in this chapter assumes that the transmitter is tunable to all possible receiver wavelengths. The transmitters take a fixed amount of time, δ, expressed in time slots, to change their tuning from one wavelength to another. It gives a more general framework for our results. Since δ is parameter, our results rely on δ. Our results based on δ are more general. With technologies for optical devices are advancing, the tuning time δ will be getting smaller and smaller. But our results are still applicable for δ is parameter in our study. Another goal for our study is to reduce the number of tuning times for all-to-neighbor transmission cycle. In fact, for the directed de Bruijn graph we tune the transmitters only once before the transmission schedule begins, so we are actually using a fixed tuned transmitter model. For the undirected de Bruijn graph we tune each transmitter to only two wavelengths; this could be implemented as an optical network where each node has two fixed tuned transmitters and one receiver. Many of our results can be interpreted in a model with multiple fixed tuned transmitters, and this model has been successfully implemented experimentally.

There are additional questions about a multiple fixed tuned transmitter model that we do not directly consider. It is possible that slightly shorter transmission schedules could be devised by having a single node transmit on two wavelengths simultaneously; but our schedules are using all possible wavelengths in almost all time slots, so the improvement would only be slight. Another problem for fixed tuned models is determining whether an embedding is feasible given an assignment of wavelengths to the transmitters and receivers, or determining what assignments allow the most flexibility in embedding virtual topologies. Our results simply give an assignment that is optimal for embedding the de Bruijn topologies.

This chapter is organized as follows. In the next section, we give results about optimal embeddings and transmission schedules for arbitrary virtual topologies. Then we design optimal embeddings and scheduling for directed de Bruijn topologies, and finally for undirected de Bruijn topologies. We summarize our work in Section 2.5.

2.2 General Results

An (N, k, δ) *optical passive star network* is a network as shown in Figure 1.1 with N nodes, $0, 1, \ldots, (N-1)$. Each node has a tunable transmitter that can can tuned to any of k wavelengths, $w_0, w_1, \ldots, w_{k-1}$. Each node has a fixed tuned receiver, so that node i is tuned to wavelength $w_{i \bmod k}$. The network transmits packets in synchronous time slots, with only one packet of each wavelength permitted in each time slot. Thus at most k packets can be transmitted in a single time slot. The time needed for a node to tune its transmitter to a new wavelength is δ time slots.

A *network topology*, G, is a directed graph that represents the virtual interconnections among the vertices. If G is an undirected graph, think of each edge as two directed edges, one in each direction.

If all nodes in G have r out-going links and r in-coming links, G is called *regular with degree* r or r *regular graph.* Let V and E be the set of vertices and edges, respectively, of G; and let $v = |V|$ and $e = |E|$ be their sizes. If G is regular, let r be the in degree of the vertices. Then $vr = e$. Regular graphs are one of important classes of graphs. A lot of common network topologies belong to regular graphs, such as ring and hypercube, We will focus on regular graphs in the section. Our general results can be applied to some regular graphs, such as ring and hypercube. We also propose a couple of general approaches to study the transmission schedule for regular graphs and general graphs.

An *embedding* of a network topology G into an optical passive star is a map, f, from the vertices of G to the nodes of the network. This naturally gives a coloring of the vertices of G by the k wavelengths: vertex v is colored by the receiving wavelength of node $f(v)$, that is by wavelength $w_{f(v) \bmod k}$. We denote the coloring function by Receiver-Wavelength(v) and we give only the wavelength index; so here Receiver-Wavelength$(v) = f(v) \bmod k$. This in turn gives a coloring of the directed edges of G, where an edge has the color of the vertex to which it points (the vertex at its head). Each directed edge represents one packet that must be transmitted.

Note that the nodes of a passive star network are distinguished only by the wavelength of their receivers. Hence specifying the receiver wavelength coloring of the vertices of G implicitly defines the embedding: the first vertex of G of color i is assigned to node i in the optical passive star network, the second vertex of color i is assigned to node $i + k$, the third to node $i + 2k$, and so on. As long as we do not use more than N/k of any one wavelength color, the embedding is straightforward. So in what follows we will only specify the function Receiver-Wavelength(v) that defines the coloring by receiver wavelengths, verifying that we do not use more than N/k of any one color. We will not explicitly specify the embedding function, f.

An *optimal transmission schedule* must schedule each packet transmission so that only one transmission occurs in each wavelength in each time slot, and yet so the the total time needed for all transmissions is minimum. The minimum time transmission schedule might well depend on the embedding we choose. We want to construct the embedding so that it allows the minimum transmission schedule out of all possible embeddings, and so that we can efficiently determine the embedding and the optimal transmission schedule. An optimal transmission schedule also depends on the value of δ. For theoretical purposes we will sometimes consider optimal transmission schedules when $\delta = 0$, though in practice δ is never 0. δ has an big effect in designing optimal transmission schedule.

Before we design any transmission schedule for any virtual topology embedded in an optical passive star network, we want to know the low bound for the length of optimal transmission schedules. This low bound tells us the least number of slots which is needed for any transmission schedule. This low bound can be analyzed and determined theoretically. Of course, this may not be an exact bound. After analysis, a transmission schedule will be designed to be as near to the lower bound as possible. If the transmission schedule reaches the low bound, it tells us that an optimal transmission schedule has been proposed and the work has been perfectly done. Otherwise, either the lower bound is not the exact bound for this virtual topology, or the transmission schedule is not optimal. More analysis and design are needed to achieve better performance for such a transmission cycle.

A trivial lower bound on the length of an optimal transmission schedule is $\lceil e/k \rceil$ where e is the number of edges of G, since there are e transmissions and each time slot can accommodate at most k simultaneous transmissions. Clearly, to achieve such a lower bound, we would need to have at most $\lceil e/k \rceil$ edges of each wavelength color. G is regular of degree r, then this lower bound becomes $r\lceil v/k \rceil$, since some color must be used for at least $\lceil v/k \rceil$ vertices, giving at least $r\lceil v/k \rceil$ transmissions on a single wavelength. To achieve this lower bound we would need at most $\lceil v/k \rceil$ vertices of each color. We summarize this in Theorem 2.2.2.

Definition 2.2.1 *A k* uniform vertex (edge) coloring *of a graph is a coloring so that there are at most $\lceil v/k \rceil$ vertices ($\lceil e/k \rceil$ edges) of each color.*

Theorem 2.2.2 *Edge Lower Bound: There is a $\lceil e/k \rceil$ lower bound on an optimal transmission schedule, which can be achieved only if G is uniformly edge colored by the receiver wavelengths. If G is regular of degree r, then there is an $r\lceil v/k \rceil$ lower bound, which can be achieved only if G is uniformly vertex colored.*

Note also that a uniform vertex coloring ensures that as long as $v \leq N$, that is as long as the network topology has fewer vertices than the optical passive star we are embedding it into, then a coloring by receiver wavelengths enduces a valid embedding.

So when G is regular, a necessary condition for an $r\lceil v/k\rceil$ optimal transmission schedule is that G be k uniformly colored by the receiver wavelengths. Surprising, this is also a sufficient condition when $\delta = 0$.

Theorem 2.2.3 *If G is an r regular graph and each receiver wavelength colors at most m vertices and $\delta = 0$, then G has an rm transmission schedule that can be determined in time $O(r^2v^{3/2})$.*

Proof : Look at the adjacency matrix of G; each row and column sums to r. By a well know theorem of Birkhoff [8, 41], the matrix can be written as the sum of r permutation matrices, where each row and column has only a single entry of value 1. This can be done by solving $r-1$ maximum matching problems, which can be done in time $O(r^2v^{3/2})$ [28]. Each entry with value 1 in these permutation matrices represent one directed edge of the graph, and so one transmission that must be scheduled.

Look at a single permutation matrix. Since each row has only one entry of value 1, each node will only be transmitting to one receiver. Since each column has only one entry of value 1, each receiver will only be receiving from one transmitting node. The wavelength of the transmission is determined by the column in which the entry of value 1 appears. Write this permutation matrix in turn as a sum of at most m $\{0,1\}$-matrices where in each matrix at most one column of each wavelength has an entry of value 1. Since at most m vertices of G have the same receiver wavelength color, this is possible and can be done efficiently. Each such matrix represents the transmissions that occur in a single time slot.

Since we are assuming that $\delta = 0$, the transmission along all of the edges of G can be scheduled in only rm time slots. ∎

Corollary 2.2.4 *If G is an r regular graph and it is k uniformly vertex colored by the receiver wavelengths, and if $\delta = 0$, then it has an $r\lceil v/k\rceil$ optimal transmission schedule that can be determined in time $O(r^2v^{3/2})$.*

Proof : Here that the transmission schedule is optimal by Theorem 2.2.2 for regular graphs. ▮

We can apply this corollary to show that any embedding of the d-ary n-dimensional hypercube that assigns the k receiver wavelengths uniformly has an optimal transmission schedule of length $(d-1)n\lceil d^n/k\rceil$ time slots, when $\delta = 0$.

Another lower bound on the length of an optimal transmission schedule is given by considering the number of transmissions on each wavelength that each node must perform.

Theorem 2.2.5 *Vertex Lower Bound: If G has a vertex of degree at least r and the neighbors of that vertex are colored by at least $\ell > 1$ receiver wavelengths, then there is an $r + \ell\delta$ lower bound on an optimal transmission schedule, using this wavelength coloring.*

Proof : Each the vertex must do r transmissions and change its transmission wavelength ℓ times in each transmission cycle. ▮

It is important to emphasize that this lower bound depends on the wavelength coloring. Note also that if each vertex has neighbors of only one wavelength, $\ell = 1$, then the transmitter can set its wavelength once before any transmissions begin and this lower bound becomes just r.

If δ is large enough then it is easy to find a transmission schedule that meets this lower bound and hence is optimal. If $\delta > e$, for example, then we can schedule only one transmission per time slot and still achieve an optimal transmission schedule by this lower bound. So we will be interested in schedules that are optimal for all values of δ, both large and small.

To reduce this lower bound as much as possible, we would like the neighbors of a vertex in G to be colored by as few different wavelengths as possible. This motivates the next definition.

Definition 2.2.6 *An* ℓ-neighborhood bounded vertex coloring *of a graph is a coloring where each vertex has neighbors colored by at most* ℓ *colors.*

There are limits on how small ℓ can be.

Theorem 2.2.7 *If* G *is undirected and connected, then* G *has a 1-neighborhood bounded coloring with at least two colors if and only if* G *is bipartite.*

Proof : If G is bipartite, then clearly it has a 1-neighborhood bounded coloring with two colors.

Assume G has a 1-neighborhood bounded coloring with at least two colors, and that it is not bipartite. Since G is connected, it must have an edge $\overline{uv}$ with u and v different colors, say u is blue and v is red. Since G is not bipartite, it must have an odd length cycle. Then there must be a path from v to the odd length cycle, but any path from v must alternate colors, since all of the neighbors of v must be blue like u, all vertices length two away must be red, like v, and so on. This means the odd length cycle must be colored alternately by red and blue—but this is impossible. ∎

Theorem 2.2.7 tells us only bipartite graphs have 1-neighborhood bounded coloring. 1-neighborhood bounded coloring is a big plus to reduce the number of tuning times for each node and simplify the transmission schedule. For the transmitter of a station need to tune its transmitter only once to talk to all of its neighboring stations, the transmitter can sit on that channel until the all-to-neighbor transmission cycle is over. The following facts can be known from Theorem 2.2.7:

- Rings have 1-neighborhood bounded colorings
- Meshes have 1-neighborhood bounded colorings
- Hypercubes have 1-neighborhood bounded colorings

We would like to mention one last general technique. When some node needs more than one transmitter wavelength, it is possible to replace two cycles of the transmission schedule by a cycle followed by a reversed cycle. In this way, the tuning of the transmitters at the end of the cycle is unnecessary, since the reversed cycle begins in the same tuning that ends the first cycle. This can reduce the transmission time by as much as δ.

2.3 Directed de Bruijn Graphs

The directed de Bruijn graph, $B(d, n)$, is defined by

$$\begin{aligned} V &= \{(x_1, \ldots, x_n) \mid x_i \in [0, d-1]\} \\ E &= \{(x_1, \ldots, x_n) \to (x_2, \ldots, x_n, \alpha) \mid \alpha \in [0, d-1]\}. \end{aligned}$$

Note this is regular of degree d; however, it has d loops $(x, \ldots, x) \to (x, \ldots, x)$. Since no packets need to be transmitted along these loop edges (a node does not need to transmit to itself), we drop these edges.

Since there are d^n vertices, each of degree d, and since we exclude the d loop edges, we have a total of $d^{n+1} - d$ edges. By the Edge Lower Bound, Theorem 2.2.2, we know an optimal transmission schedule must use at least $\lceil (d^{n+1} - d)/k \rceil$ time slots.

In fact this lower bound can be achieved for some values of k and d. This holds for any δ, since it is possible to embed the directed de Bruijn graph in the optical passive star so that the neighbors of a given node all have the same receiver wavelength—a 1-neighborhood bounded coloring. In such an embedding, the transmitters never need to tune to a different wavelength, so there is no dependence on the tuning time δ.

We will define two different embeddings, the first gives an optimal transmission schedule for values of k and d that satisfy certain simple conditions. The second embedding is more complicated and it gives a good but not necessarily optimal transmission schedule, but without any assumptions on k and d. Actually we must make some assumptions about N, k, and d for the embedding to be feasible; we assume $d^n < N$ so that some embedding is possible, and $d < N/k$ so that a 1-neighborhood bounded coloring is possible.

Here is the first vertex coloring, which by our comments above implicitly defines the embedding.

Embedding A:

$$\text{Receiver-Wavelength}(x_1, \ldots, x_n) = \sum_{j=1}^{n-1} x_j d^{j-1} \bmod k.$$

This coloring can be thought of as taking the number represented in base d by $(x_1 x_2 \ldots x_{n-1})$ and reducing it modulo k to get the color of vertex $(x_1, \ldots, x_n)$. This coloring is not necessarily vertex uniform, but it does use each wavelength color at most $d\lceil d^{n-1}/k \rceil$ times. (If $k|d^{n-1}$ then it is vertex uniform.) This coloring is 1-neighborhood bounded, because the neighbors of $(x_1, \ldots, x_n)$ are all of the form $(x_2, \ldots, x_n, \alpha)$ for some α, and these are all of the same receiving wavelength color.

This gives a $d^2\lceil d^{n-1}/k \rceil$ transmission schedule for the entire directed de Bruijn graph, by Theorem 2.2.3. We would like to improve this to get an optimal $\lceil (d^{n+1} - d)/k \rceil$ transmission schedule for the graph with the loops removed. If $k|d^{n-1}$ and $d < k$ (both reasonable assumptions) then in fact $\lceil (d^{n+1} - d)/k \rceil$ and $d^2\lceil d^{n-1}/k \rceil$ are equal. Hence in this case, we have an optimal transmission schedule.

Next consider the case where $k|d$. First of all, note that the loop edges are uniformly distributed among the k different wavelengths. This follows because the loop $(x,\ldots,x) \to (x,\ldots,x)$ is assigned to wavelength

$$\sum_{j=1}^{n-1} xd^{j-1} \equiv x \pmod{k},$$

since $k|d$. Consequently, the d vertices with loops together have $(d-1)d$ edges uniformly distributed among the k different wavelengths.

We can consider the directed de Bruijn graph as the union of all edges of the form $(x_1, x_2, \ldots, x_n) \to (x_2, \ldots, x_n, x_1)$ and all other edges. Each subgraph is regular, of degree 1 and degree $d-1$ respectively. We can apply Corollary 2.2.4 to each subgraph to get a transmission schedule. Note all loop edges are in the first subgraph, and since they are uniformly colored by the receiving wavelengths, we can remove d/k time slots from this schedule by removing all loop edges. This gives us a schedule for all transmissions in

$$\frac{d^n}{k} - \frac{d}{k} + \frac{(d-1)d^n}{k} = \left\lceil \frac{d^{n+1}-d}{k} \right\rceil$$

time slots, and so it is optimal. We summarize this in a theorem.

Theorem 2.3.1 *If $k|d^{n-1}$ and either $k > d$ or $k|d$, then Embedding A has an optimal transmission schedule, as given above.*

An example of this embedding and an optimal transmission schedule for $B(4,2)$ with 4 wavelengths is given in Figure 2.2. In this example, $d = 4, n = 2$ and $k = 4$. It is a cas of $k|d$. So the number of time slots for optimal schedules of all-to-neighbor transmission Cycle will be no

$$\left\lceil \frac{d^{n+1}-d}{k} \right\rceil = \left\lceil \frac{4^3-4}{4} \right\rceil = 15$$

	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
(0, 0)	L	R	R	R												
(0, 1)					C	R	R	R								
(0, 2)									C	R	R	R				
(0, 3)													C	R	R	R
(1, 0)	R	C	R	R												
(1, 1)					R	L	R	R								
(1, 2)									R	C	R	R				
(1, 3)													R	C	R	R
(2, 0)	R	R	C	R												
(2, 1)					R	R	C	R								
(2, 2)									R	R	L	R				
(2, 3)													R	R	C	R
(3, 0)	R	R	R	C												
(3, 1)					R	R	R	C								
(3, 2)									R	R	R	C				
(3, 3)													R	R	R	L

L = loop edges, C = cycle edges where $(x_1, x_2) \rightarrow (x_2, x_1)$, and R = rest of the edges.

Table 2.1: Adjacency matrix of $B(4,2)$

Figure 2.2 demonstrates the adjacency matrix of B(4,2) and shows how to make an optimal transmission schedule in 15 time slots.

Another example of this embedding and an optimal transmission schedule for $B(2,4)$ with 4 wavelengths is given in Figure 2.3 and Figure 2.4. In this example, $d = 2, n = 4$ and $k = 4$. It is a cas of $k = d^2$. So the number of time slots for optimal schedules of all-to-neighbor transmission Cycle will be no

$$\left\lceil \frac{d^{n+1} - d}{k} \right\rceil = \left\lceil \frac{2^5 - 2}{4} \right\rceil = 8$$

Figure ?? demonstrates the adjacency matrix of B(4,2) and shows how to make an optimal transmission schedule in 8 time slots.

Before we give our second embedding, we will first prove a better lower bound on an optimal transmission schedule. Because the directed de Bruijn graph with loops removed is almost regular, we can get an improved lower bound that falls above the general Edge Lower Bound, but below the Edge Lower Bound for regular graphs.

Theorem 2.3.2 *An optimal transmission schedule on $B(d, n)$ with loops removed*

V\Time	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
(0, 0)				(0, 1)	(0, 2)	(0, 3)									
(0, 1)	(1, 0)			(1, 1)	(1, 2)	(1, 3)									
(0, 2)	(2, 0)			(2, 1)	(2, 2)	(2, 3)									
(0, 3)	(3, 0)			(3, 1)	(3, 2)	(3, 3)									
(1, 0)		(0, 1)					(0, 0)	(0, 2)	(0, 3)						
(1, 1)							(1, 0)	(1, 2)	(1, 3)						
(1, 2)		(2, 1)					(2, 0)	(2, 2)	(2, 3)						
(1, 3)		(3, 1)					(3, 0)	(3, 2)	(3, 3)						
(2, 0)			(0, 2)							(0, 0)	(0, 1)	(0, 3)			
(2, 1)			(1, 2)							(1, 0)	(1, 1)	(1, 3)			
(2, 2)										(2, 0)	(2, 1)	(2, 3)			
(2, 3)			(3, 2)							(3, 0)	(3, 1)	(3, 3)			
(3, 0)	(0, 3)												(0, 0)	(0, 1)	(0, 2)
(3, 1)		(1, 3)											(1, 0)	(1, 1)	(1, 2)
(3, 2)			(2, 3)										(2, 0)	(2, 1)	(2, 2)
(3, 3)													(3, 0)	(3, 1)	(3, 2)

Transmission schedule, showing destination vertex at each time slot.
The wavelength is the first coordinate of the destination vertex.

Table 2.2: Schedule for $B(4, 2)$ with 4 wavelengths using Embedding A.

	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
0000	L	R														
0001			R	R												
0010					R	R										
0011							R	R								
0100									R	R						
0101											R	R				
0110													R	R		
0111															R	R
1000	R	R														
1001			R	R												
1010					R	R										
1011							R	R								
1100									R	R						
1101											R	R				
1110													R	R		
1111															R	L

Adjcency matrix of B(2,4), L=loop edges, R=rest of edges

Figure 2.3: The adjacency matrix of B(2,4)

must use at least

$$d\left\lceil \frac{d^n}{k} \right\rceil - \left\lfloor \frac{d}{r} \right\rfloor$$

time slots, where $d^n = kq + r$ *with* $0 < r \leq k$.

Proof : First assume that some receiving wavelength is used more than $\lceil d^n/k \rceil$ times. Then even if all the loop edges share this same receiving wavelength, we would need at least $d(\lceil d^n/k \rceil + 1) - d$ time slots just to transmit along all the edges of this wavelength, but this is already at least as large as the lower bound of the theorem.

So r different wavelengths must each be used for exactly $\lceil d^n/k \rceil$ vertices. From these r wavelengths, pick the the one that colors the fewest loop edges; it colors at most $\lfloor d/r \rfloor$ loop edges. Therefore, to transmit along all the edges of this wavelength, we must use at least

$$d\left\lceil \frac{d^n}{k} \right\rceil - \left\lfloor \frac{d}{r} \right\rfloor$$

time slots. ∎

Now here is our second embedding. We give it in the form of an algorithm. We need to be sure that the loop edges are colored uniformly; this fact is clear from the algorithm.

Embedding B:

$c \leftarrow 0$

`for` $x = 0$ `to` $d-1$

  `for` $\alpha = 0$ `to` $d-1$

    `Receiver-Wavelength`$(x, x, \ldots, \alpha) \leftarrow c \bmod k$

  $c \leftarrow c + 1$

`for` $(x_1, \ldots, x_{n-1}) = (0, \ldots, 0)$ `to` $(d-1, \ldots, d-1)$

  `if` $x_1 \neq x_2$ `OR` $x_2 \neq x_3$ `OR` $\cdots$ `OR` $x_{n-2} \neq x_{n-1}$ `then`

```
for α = 0 to d − 1
        Receiver-Wavelength(x_1, ..., x_{n−1}, α) ← c mod k
    c ← c + 1
```

Here the colors of the loop edges are uniformly distributed over the k wavelength colors. For the transmission schedule, we can schedule transmission along all edges, including the loop edges, in $d^2\lceil d^{n-1}/k\rceil$ time steps. We can again arrange to have all the loop edges appear together in this schedule, so $\lfloor d/k\rfloor$ time slots will be completely taken up by loop edges and they can be removed from the schedule. This gives a schedule requiring only

$$d^2\left\lceil\frac{d^{n-1}}{k}\right\rceil-\left\lfloor\frac{d}{k}\right\rfloor$$

time slots. This may not be optimal, but it is close to the lower bound established in Theorem 2.3.2. It is easy to see that the difference between the length of this schedule and the lower bound in the theorem is at most d^2 time slots. If $k|d^{n-1}$ then $d^2\lceil d^{n-1}/k\rceil = d\lceil d^n/k\rceil$ and $r = k$ and this schedule meets the lower bound and so it is optimal. We summarize this in the next theorem.

Theorem 2.3.3 *Embedding B has a transmission schedule, as given above, in $d^2\lceil d^{n-1}/k\rceil - \lfloor d/k\rfloor$ time slots. If $k|d^{n-1}$ then it is an optimal transmission schedule. If $k \nmid d^{n-1}$, then the transmission schedule uses at most d^2 more time slots than optimal.*

With Theorem 2.3.1 and Theorem 2.3.3 we get an embedding and an optimal transmission schedule for the directed de Bruijn graph $B(d,n)$ whenever $k|d^{n-1}$, and a near optimal schedule (within d^2 time slots of optimal) in all other cases. This transmission schedule is optimal over all possible embeddings and for all possible values of δ, since it meets the Edge Lower Bound, Theorem 2.2.2, which holds for all embeddings and all δ.

2.4 Undirected de Bruijn Graphs

The undirected de Bruijn graph, $UB(d,n)$, is simply the underlying undirected graph of $B(d,n)$. If $d=1$ the graph has only one vertex; so we assume $d>1$. If $n=1$ the graph is a complete graph on d vertices. The case of the complete graph has been handled in [40], so we will also assume $n>1$. We sometimes need to handle the case $n=2$ specially; we will only mention this case in passing.

We will call the edges $(x_1,x_2,\ldots,x_n)\to(x_2,\ldots,x_n,\alpha)$, left-shift edges, and edges $(x_1,\ldots,x_{n-1},x_n)\to(\alpha,x_1,\ldots,x_{n-1})$, right-shift edges. As in the directed case, we drop the d loop edges. Here there are also multiple edges between $(d^2-d)/2$ pairs of vertices of the form $(a,b,a,b,\ldots)$ and $(b,a,b,a,\ldots)$ where $a\neq b$; there is a left-shift and right-shift edge between these two vertices. We only count one of these two edges. Since we must send messages in both directions along each undirected edge, the total number of transmissions is $2d^{n+1}-d^2-d$.

The first result about undirected de Bruijn graphs is an immediate consequence of Theorem 2.2.7.

Theorem 2.4.1 *If $d>1$ and $n>1$, then $UB(d,n)$ does not have a 1-neighborhood bounded coloring with at least two colors.*

Proof : When $d>1$ and $n>1$ we have the length three cycle

$$(0,0,\ldots,0)\to(0,\ldots,0,1)\to(1,0,\ldots,0)\to(0,0,\ldots,0).$$

Hence, $UB(d,n)$ is not a bipartite graph, and by Theorem 2.2.7 it does not have a 1-neighborhood bounded coloring with at least two colors. ∎

However, a 2-neighborhood bounded coloring is possible, as shown by the following embedding.

Embedding C:

$$\text{Receiver-Wavelength}(x_1, \ldots, x_n) = \sum_{j=2}^{n-1} x_j d^{j-2} \bmod k.$$

Here the left-shift edges of $(x_1, x_2, \ldots, x_n)$ are all colored by $(x_3 \ldots x_n)_d$ (that is, as a base d number) reduced modulo k. The right-shift edges are all colored $(x_1 \ldots x_{n-2})_d \bmod k$. If we are to use all wavelengths we must assume that $d^{n-2} \geq k$.

By the previous theorem, this embedding is optimal for minimizing the total number of the tuning times.

We will consider two special cases: $k|d$ and $k = d^p$ for $2 \leq p \leq n-2$.

2.4.1 Transmission Schedule When $k|d$

When $k|d$, Embedding C takes the form

$$\text{Receiver-Wavelength}(x_1, \ldots, x_n) = x_2 \bmod k.$$

Since each node $x = (x_1, \ldots, x_n)$ needs to transmit a message to $(x_2, \ldots, x_n, \alpha)$, $0 \leq \alpha \leq d-1$, and to $(\beta, x_1, \ldots, x_{n-1})$, $0 \leq \beta \leq d-1$, the wavelength of the transmitter of x need occupy only wavelengths $x_3 \bmod k$ and $x_1 \bmod k$. The wavelength of the receiver of $x = (x_1, x_2, \ldots, x_n)$ is fixed at $x_2 \bmod k$.

Algorithm UB1:

1. **Initialization**: Define

$$V[i,j] = \{x = (x_1, x_2, \ldots, x_n) \mid i \equiv x_3,\ x_1 \equiv x_3 + j \pmod{k}\},$$

where $0 \leq i, j \leq k-1$.

For $x \in V[i,j]$, the wavelength of x's transmitter is switched to i.

We choose an arbitrary order for the vertices in $V[i,j]$ with $j \neq 0$. For the vertices in $V[i,0]$, we divide them into three parts:

$$\begin{aligned} A(i) &= \{x \in V[i,0] \mid x_i = x_{i+1}, 0 \leq i \leq n-1\} \\ B(i) &= \{x \in V[i,0] - A(i) \mid x_i = x_{i+2}, 0 \leq i \leq n-2\} \\ C(i) &= V[i,0] - A(i) - B(i). \end{aligned}$$

Let $B(i) = B(i,1) \cup B(i,2)$ with $B(i,1) \cap B(i,2) = \emptyset$ and $0 \leq |B(i,1)| - |B(i,2)| \leq 1$.

We order the vertices of $V[i,0]$ so that all vertices of $C(i)$ come first, then the vertices of $B(i)$, and finally those of $A(i)$. We pick an arbitrary order within $A(i)$, $B(i)$ and $C(i)$. It is a simple fact that each node in $B(i)$ has a neighbor which is both a left-shift neighbor and a right-shift neighbor, we call this neighbor a common neighbor.

2. **Phase I**: All nodes transmit messages to their left-shift neighbors, except the common neighbors of the nodes in $B(i,2)$ with $0 \leq i \leq k-1$.

```
for j = 1 to k
    Each node in V[i, j mod k] transmits to its left-shift neighbors in
        parallel over 0 ≤ i < k according to its order in V[i, j mod k]
        except the common neighbors of the nodes in B(i,2) with 0 ≤ i < k.
    After completing its transmission, each node switches its transmitter
        wavelength to (i + j) mod k.
```

3. **Phase II**: All nodes transmit messages to their right-shift neighbors, except the common neighbors of the nodes in $B(i,1)$ with $0 \leq i \leq k-1$.

This is similar to Phase I. After completing transmissions nodes switch their transmitters back to wavelength $i \bmod k$.

This algorithm gives a schedule for the time-slot cycle. Note that the schedule returns to the same setup at the end of Phase II, so we can continue another cycle by repeating Phase I and Phase II.

Theorem 2.4.2 *Algorithm UB1 gives an optimal schedule for the time-slot cycle for all δ. If $\delta \leq d^{n+1}/k^2 - d - d/k - \lceil d(d-1)/2k \rceil$, the cycle is $(2d^{n+1} - d^2 - d)/k$ time slots long.*

If $\delta > d^{n+1}/k - d - d/k - \lceil d(d-1)/2k \rceil$, the cycle is $2\delta + 2d$.

Proof : First $|V[i,j]| = d^n/k^2$. Since we keep the same order for all the nodes in $V[i,j]$ in Phase I and Phase II and $x = (x_1, x_2, \ldots, x_n)$ with $x_i = x_{i+1}, 0 \leq i \leq n-1$ has only d-1 left-shift neighbors, there are no wasted time slots if the tuning time from Phase I to Phase II for each node is no more than $dk|V[i,j]| - d - d/k - \lfloor d(d-1)/2k \rfloor$.

Since each node in C of $V[i,0]$ has only $d-1$ right-shift neighbors and each node in B of $V[i,0]$ has $d-1$ right-shift neighbors yet to transmit to (each node in B of $V[i,0]$ has d right-shift neighbors, but one of them is also a left-shift neighbor), there are no wasted time slots if the tuning time from Phase II to Phase I for each node is no more than $dk|V[i,j]| - d - d/k - \lceil d(d-1)/2k \rceil$.

Note that

$$dk|V[i,j]| - d - d/k - \lceil d(d-1)/2k \rceil \leq dk|V[i,j]| - d - d/k - \lfloor d(d-1)/2k \rfloor$$

$$|V[i,j]| * d * k - d - d/k - \lfloor d(d-1)/2k \rfloor \leq |V[i,j]| * d * k - d - d/k - \lceil d(d-1)/2k \rceil + 1$$

Therefore if $\delta \leq d^{n+1}/k^2 - d - d/k - \lceil d(d-1)/2k \rceil$, Algorithm UB1 is optimal; the cycle is $(2d^{n+1} - d^2 - d)/k$ time slots long.

If $\delta > d^{n+1}/k - d - d/k - \lceil d(d-1)/2k \rceil$, the time-slot cycle is

$$(2d^{n+1} - d^2 - d)/k + 2\delta - 2(d^{n+1}/k - d - d/k) - \lfloor d(d-1)/2k \rfloor - \lceil d(d-1)/2k \rceil = 2\delta + 2d$$

∎

For $UB(d,2)$ with $k|d$, it is easy to design a pipeline schedule similar to that given in [40] for the case $n = 1$, the complete graph. An example of Algorithm UB1 for $UB(4,3)$ with 2 wavelengths is given in Figures 2.5 and 2.6. For $UB(4,3)$, $d = 4, n = 3$ and $k = 2$. If $\delta \leq d^{n+1}/k^2 - d - d/k - \lceil d(d-1)/2k \rceil$, the time slots for the optimal transmission schedules are no less than

$$(2d^{n+1} - d^2 - d)/k = (2 * 4^4 - 4^2 - 4)/2 = 246$$

Figures 2.5 and 2.6 tells how to make an optimal schedule in 246 time slots.

An example of Algorithm UB1 for $UB(4,3)$ with 4 wavelengths is given in Figures 2.5 and 2.6. For $UB(4,3)$, $d = 4, n = 3$ and $k = 2$. If $\delta \leq d^{n+1}/k^2 - d - d/k - \lceil d(d-1)/2k \rceil$, the time slots for the optimal transmission schedules are no less than

$$(2d^{n+1} - d^2 - d)/k = (2 * 4^4 - 4^2 - 4)/4 = 123$$

Figures 2.7 tells how to make an optimal schedule in 123 time slots.

2.4.2 Transmission Schedule When $k = d^p$

Now we turn to the case where $k = d^p$ for $2 \leq p \leq n-2$. In this case, we have $n \geq 4$. Here we will schedule transmissions along all edges, even along the loop and multiple edges. This will still give us an optimal transmission schedule as long as $p > 2$. If $p = 2$ the transmission schedule is at most 1 time slot longer than optimal.

	1	2	3	4	5	6	7	8
0000 0001 0010 0011	0100	0001 0101	0010 0110	0011 0111				
0100 0101 0110 0111	1000 1100	1001 1101	1010 1110	1011 1111				
1000 1001 1010 1011					0000 0100	0001 0101	0010 0110	0011 0111
1100 1101 1110 1111					1000 1100	1001 1101	1010 1110	1011 1111

Figure 2.4: Transmission Schedule for B(2,4) with 4 wavelengths using Embedding A

$$
\begin{aligned}
V[0,1] &= \{(1,x_2,0),(3,x_2,0),(1,x_2,2),(3,x_2,2)\} \\
V[0,0] &= \{(0,x_2,0),(2,x_2,0),(0,x_2,2),(2,x_2,2)\} \\
A_1 &= \{(0,0,0),(2,2,2)\} \\
B_1 &= \{(0,x_2,0),(2,x_2,2)\} - A_1 \\
C_1 &= V[0,0] - A_1 - B_1 \\
\\
V[1,1] &= \{(0,x_2,1),(2,x_2,1),(0,x_2,3),(2,x_2,3)\} \\
V[1,0] &= \{((1,x_2,1),(3,x_2,1),(1,x_2,3),(3,x_2,3)\} \\
A_2 &= \{(1,1,1),(3,3,3)\} \\
B_2 &= \{(1,x_2,1),(3,x_2,3)\} - A_2 \\
C_2 &= V[1,0] - A_2 - B_2
\end{aligned}
$$

Figure 2.5: Vertex sets for $UB(4,3)$ with 2 wavelengths using Algorithm UB1.

time slot	1------64	65---96	97--116	117 -120	121----184	185--216	217--240	241-246
V[0,1]	w0				w1			
A_1				w0				w0
B_1			w0				w0	
C_1		w0				w0		
V[1,1]	w1				w0			
A_2				w1				w1
B_2			w1				w1	
C_2		w1				w1		

Figure 2.6: Schedule for $UB(4,3)$ with 2 wavelengths using Algorithm UB1.

V[i,j], $-1 < i, j < k=4$

V[0,0] = { (0, X_2 ,0)}
A(0)={(0, 0, 0)}
B(0)=V[0,0]-A(0)

V[1,0]={ (1, X_2, 1) }
A(1) = { (1, 1, 1) }
B(1) = V[1,1]-A(1)

V[2,0]={ (2, X_2 , 2)}
A(2)={ (2, 2, 2) }
B(2)=V[2, 2]-A(2)

V[3,0]={ (3, X_2 , 3) }
A(3) = { (3, 3, 3) }
B(3) = V[3,0]-A(3)

	1-3	4-12	13-60	61-63	64-75	76-123
A(0)	W_0			W_0		
B(0)		W_0			W_0	
A(1)	W_1			W_1		
B(1)		W_1			W_1	
A(2)	W_2			W_2		
B(2)		W_2			W_2	
A(3)	W_3			W_3		
B(3)		W_3			W_3	
V[*, a] a>0			W0 W_1 W_2 W_3			W_0 W_1 W_2 W_3

Figure 2.7: Schedule for $UB(4,3)$ with 4 wavelengths using Algorithm UB1.

Here is the algorithm.

Algorithm UB2:

1. **Initialization:** Define for $0 \leq j, k < d$

$$
\begin{aligned}
V[j,k,\vec{y}] &= \{(a,b,x_1,\ldots,x_{p-2},a+j,b+k,\vec{y}) \mid 0 \leq a,b,x_i < d\} \\
L[i,j,k,\vec{y}] &= \{(a,b,x_1,\ldots,x_{p-2},a+j,b+k,\vec{y}) \rightarrow \\
&\qquad (b,x_1,\ldots,x_{p-2},a+j,b+k,\vec{y},i) \mid 0 \leq a,b,x_i < d\} \\
R[i,j,k,\vec{y}] &= \{(a,b,x_1,\ldots,x_{p-2},a+j,b+k,\vec{y}) \rightarrow \\
&\qquad (i,a,b,x_1,\ldots,x_{p-2},a+j,b+k,\hat{y}) \mid 0 \leq a,b,x_i < d\}
\end{aligned}
$$

   Note here that $a+j$ and $b+k$ are reduced modulo d. By $\hat{y}$ in the definition of R we mean the vector of coordinates $\vec{y}$ with the last coordinate removed.

2. **Phase I:** All vertices transmit to their left-shift neighbors.

```
for y⃗ = (0,...,0) to (d−1,...,d−1)
    for k = 0 to d−1
        for j = 0 to d−1
            for i = 0 to d−1
                Transmit along the edges of L[i,j,k,y⃗]
                in a single time slot
            Each node in V[j,k,y⃗] tunes its transmitter
            to the wavelength for its right-shift neighbors
```

3. **Phase II:** All vertices transmit to their right-shift neighbors.

   This is the same as phase I, except we transmit along the edges of $R[i,j,k,\vec{y}]$ and then tune the transmitters to the wavelength for the left-shift neighbors.

An example of Algorithm UB2 for $UB(2,5)$ with 8 wavelengths is given in Figure 2.3.

Theorem 2.4.3 *Algorithm UB2 gives a valid transmission schedule, when* $k = d^p$ *and* $2 \leq p \leq n-2$.

Proof : Let us look at a single set of edges $L = L[i, j, k, \vec{y}]$ and show that transmissions along all these edges can be scheduled in one time slot. It is clear that the edges in L all begin in a vertex in $V[j, k, \vec{y}]$. Each such vertex has only one transmission scheduled in L. The receiving vertices in L each receive only one transmission. To see this note that if $(b, x_1, \ldots, x_{p-2}, a+j, b+k, \vec{y}, i)$ is the same as $(b', x'_1, \ldots, x'_{p-2}, a'+j, b'+k, \vec{y}, i)$, then in fact $a = a'$, $b = b'$, and $x_\ell = x'_\ell$ for $1 \leq \ell \leq p-2$; so the source vertices are the same and this is the same transmission. And finally, each wavelength color is used only once, since the color is determined by a, b, and the x_ℓ which are different for each vertex. Hence, these transmissions can be scheduled in one time slot. The same is true about the R sets of edges.

Clearly this schedules transmissions along all edges of $UB(d,n)$, including loop and multiple edges. Therefore Algorithm UB2 gives a valid transmission schedule. ∎

Phase I, excluding tuning time, takes d^{n-p+1} time slots; so the length of this transmission schedule when $\delta = 0$ is $2d^{n-p+1}$. Each set of vertices $V[j, k, \vec{y}]$ transmits to all d of its left-shift neighbors in d consecutive time slots and then tunes its transmitter. This gives the maximal amount of free tuning time in the schedule. If $\delta \leq d^{n-p+1} - d$, then the transmission schedule is still only $2d^{n-p+1}$ time slots long. For larger values of δ, the schedule is

$$2d^{n-p+1} + 2(\delta - (d^{n-p+1} - d)) = 2d + 2\delta$$

$$\begin{aligned}
V[0,0] &= \{(0,0,x,0,0),(0,1,x,0,1),(1,0,x,1,0),(1,1,x,1,1) \mid x=0,1\} \\
V[0,1] &= \{(0,0,x,0,1),(0,1,x,0,0),(1,0,x,1,1),(1,1,x,1,0) \mid x=0,1\} \\
V[1,0] &= \{(0,0,x,1,0),(0,1,x,1,1),(1,0,x,0,0),(1,1,x,0,1) \mid x=0,1\} \\
V[1,1] &= \{(0,0,x,1,1),(0,1,x,1,0),(1,0,x,0,1),(1,1,x,0,0) \mid x=0,1\}
\end{aligned}$$

Set\Time	0–1	2–3	4–5	6–7	8–9	10–11	12–13	14–15
$V[0,0]$	L				R			
$V[0,1]$		L				R		
$V[1,0]$			L				R	
$V[1,1]$				L				R

Vertex\Time	0–1	2–3	4–5	6–7	8–9	10–11	12–13	14–15
(0, 0, 0, 0, 0)	w_0				w_0			
(0, 1, 0, 0, 1)	w_1				w_2			
(1, 0, 0, 1, 0)	w_2				w_4			
(1, 1, 0, 1, 1)	w_3				w_6			
(0, 0, 1, 0, 0)	w_4				w_1			
(0, 1, 1, 0, 1)	w_5				w_3			
(1, 0, 1, 1, 0)	w_6				w_5			
(1, 1, 1, 1, 1)	w_7				w_7			
(0, 0, 0, 0, 1)		w_1				w_0		
(0, 1, 0, 0, 0)		w_0				w_2		
(1, 0, 0, 1, 1)		w_3				w_4		
(1, 1, 0, 1, 0)		w_2				w_6		
(0, 0, 1, 0, 1)		w_5				w_1		
(0, 1, 1, 0, 0)		w_4				w_3		
(1, 0, 1, 1, 1)		w_7				w_5		
(1, 1, 1, 1, 0)		w_6				w_7		
(0, 0, 0, 1, 0)			w_2				w_0	
(0, 1, 0, 1, 1)			w_3				w_2	
(1, 0, 0, 0, 0)			w_0				w_4	
(1, 1, 0, 0, 1)			w_1				w_6	
(0, 0, 1, 1, 0)			w_6				w_1	
(0, 1, 1, 1, 1)			w_7				w_3	
(1, 0, 1, 0, 0)			w_4				w_5	
(1, 1, 1, 0, 1)			w_5				w_7	
(0, 0, 0, 1, 1)				w_3				w_0
(0, 1, 0, 1, 0)				w_2				w_2
(1, 0, 0, 0, 1)				w_1				w_4
(1, 1, 0, 0, 0)				w_0				w_6
(0, 0, 1, 1, 1)				w_7				w_1
(0, 1, 1, 1, 0)				w_6				w_3
(1, 0, 1, 0, 1)				w_5				w_5
(1, 1, 1, 0, 0)				w_4				w_7

Table 2.3: Schedule for $UB(2, 6)$ with 8 wavelengths using Algorithm UB2.

time slots long.

Theorem 2.4.4 *Assume that $k = d^p$ for $2 \leq p \leq n-2$. Algorithm UB2 gives a transmission schedule of length $2d^{n-p+1}$ time slots when $\delta \leq d^{n-p+1} - d$ and of length $2d + 2\delta$ otherwise. If $p > 2$ this schedule is optimal for all δ. If $p = 2$ this schedule is within one time slot of optimal for $\delta \leq d^{n-p+1} - d$ and optimal for larger δ.*

Proof : The calculations for the length of the schedule are given above.

Since $UB(d, n)$ has $2d^{n+1} - d^2 - d$ messages to transmit, by the Edge Lower Bound Theorem 2.2.2 any transmission schedule must have at least $\lceil (2d^{n+1} - d^2 - d)/d^p \rceil$ time slots. When $p > 2$ it is easy to see this is equal to $2d^{n-p+1}$, and when $p = 2$ this is equal to $2d^{n-p+1} - 1$. Therefore our schedule is optimal when $\delta \leq d^{n-p+1} - d$ and $p > 2$ and only one time slot longer than optimal when $p = 2$.

Since $n \geq 4$ we know there must be some vertex with neither loops nor multiple edges. This vertex has $2d$ neighbors to transmit to, and it must change its transmitter wavelength at least twice, by Theorem 2.4.1, no matter what wavelength coloring we use. By the Vertex Lower Bound Theorem 2.2.5, we have a lower bound of $2d + 2\delta$. Hence again the schedule is optimal. ∎

Hence Algorithm UB2 is an optimal schedule for the undirected de Bruijn graph, for all possible embeddings and all possible values of δ, when $k = d^p$ and $2 < p \leq n - 2$.

Using the technique of alternating forward and reverse cycles mentioned in Section 2.2, we can reduce the transmission time for large δ to $d^{n-p+1} + d + \delta$.

2.5 Summary

We have given embeddings of the virtual topologies of the directed and undirected de Bruijn graphs in a TWDM optical passive star network. Along with these embeddings we have given transmission schedules that transmit along all edges of the virtual topology (excluding loop and multiple edges) in a minimal number of time slots. We proved these embeddings minimize the number of tunings needed in each cycle of the schedule and that the schedules are optimal over all possible embeddings.

We have made certain assumptions about the parameters of the virtual and physical networks to get these optimal results, but we also have given near optimal schedules for the directed de Bruijn graph without unnecessary assumptions on the parameters.

Many open questions remain: Can we relax the requirements on the network parameters further and still prove optimality? For what parameters can we find optimal embeddings and schedules for a generalized de Bruijn network? (In a generalized de Bruijn network the edges are defined by $x_i \to x_{(di+\alpha) \bmod n}$ for $0 \leq \alpha < d$.) Can we get lower bounds on the number of tunings needed for other topologies? Most importantly, are there further applications of the general graph theoretic results that we began in this chapter, and further results along these lines that will be useful in solving these embedding and scheduling problems? We think that considering the scheduling problem as a matrix decomposition problem—trying to write the adjacency matrix of the virtual topology as a sum of matrices with certain combinatorial properties—is a particularly useful way to formulate the problem. We hope these ideas have further application.

Chapter 3

Partitioned Optical Passive Star Networks

We study reliability in the all-optical Partitioned Optical Passive Stars (POPS) topology. POPS is a physical architecture to scale the traditional optical passive star couplers and explore the advantages of high noise immunity and single-hop. The reliability and fault tolerance in POPS rely on its connectivity. In this study, we analyze the worst case for network partitions due to either node failures or link errors. For node connectivity, we show that the whole system remains connectivity no matter how many nodes don't work. For link connectivity, we analyzed some worst cases due to link errors and the lower bound for connectivity is demonstrated. Some sufficient approaches are proposed to detect and keep connectivity of the whole system.

3.1 Introduction

The Partitioned Optical Passive Stars (POPS) topology [32, 16, 31] is a physical architecture to scale the traditional optical passive star couplers and explore the advantages of the optical passive star couplers. POPS is a non-hierarchical structure and connects several optical passive star couplers together by the help of some intermediate optical passive star couplers. It remains the property of high noise immunity, single-hop, and no intermediate electronic/optical conversions. POPS is an design without considering the power budget problems, which is flexible to extend the optical passive star networks and keep the simplicity of system.

It was shown in [32] that POPS data channel can be efficiently utilized for random permutatio-based communication patterns. In most of applications in network computing and parallel computing, some common communication patterns are widely used. For example, All-to-all personalized communication is the common way to globally exchange information. Global reduction or global broadcasting is the pattern to collect data from all slaves to the master in the master-slave model. In [32], they studied four common communication patterns in TDM POPS networks which shows POPS networks are supportive for distributed/networking computing.

Reliability is one of important topics for distributed systems and local area networks [13, 14]. we study the connectivity of POPS networks. The networks reliability and fault tolerance rely on the connectivity of networks topologies. There was no result about connectivity of POPS in previous study. Both the station errors and the link errors are discussed with respect to network partitions. Some worst cases of disconnected POPS networks due to station/link errors are analyzed for getting the lower bounds on the connectivity. On the other hand, several efficient approaches are proposed to detect if the network is still connected, these approaches can also be used to fix the disconnected system.

The following section describes the concept of POPS networks and shows a couple of examples. In Section 3.3, connectivity of POPS networks is analyzed for station errors and link errors. Some efficient approaches are proposed to make the whole system connected in Section 3.4. In Section 3.5, we conclude our study and list some topics for future study.

3.2 Concepts for POPS Networks

POPS networks are based on optical passive star networks. An optical passive star coupler is one of the most common ways to interconnect an optical network via optical fibers. Figure 1.1 shows an optical passive star network with d stations. Each station sends its signal to the passive star coupler through its transmitter on a specific channel. In the passive star coupler, all the signals are combined and broadcast to all the stations. The signals in a certain channel can be received by the stations whose receivers occupy the corresponding channel.

A POPS network, POPS[n,d], can be determined by parameters: one is the number of nodes in the system denoted by n, the other is the degree of each coupler denoted by d. Each coupler is a d-station optical passive star. The n nodes are partiti oned into g=n/d groups. An example of POPS networks, POPS[16,8], is shown in Figure 3.1.

Each node has g transmitters and g receivers. There are $c = g^2$ couplers as intermediate couplers. They are denoted by $C_{i,j}, 0 \leq i, j \leq g-1$. The transmitters of nodes in Group j are connected to $C_{i,j}$ for $0 \leq j \leq g-1$. The receivers of nodes in Group i are connected to $C_{i,j}$ for $0 \leq i \leq g-1$.

POPS networks are different from the usual electronic switching networks. A POPS network use dxd passive star coupler as the intermediate connections. There are more research issues in a dxd passive star coupler such as WDM, multiple channel access protocol and transmission collision recovery.

Given POPS[n,d], the number of intermediate couplers is $c = n^2/d^2$. To reduce the cost of such a POPS network and fully utilized the resource, it is reasonable to assume that the number of intermediate couplers be no more than the number of nodes in the system. This is equivalent to $d \geq n^{1/2}$. In a WDM POPS[n,d], some new constraints on n and d will be needed to reduce the number of intermediate couplers. In order to reduce the cost of construction of a POPS network, people have to consider the tradeoff of the cost of larger degree passive star couplers and the number of intermediate passive star couplers. The large degree passive star coupler is more expensive, but the POPS will use less number of optical passive couplers. On

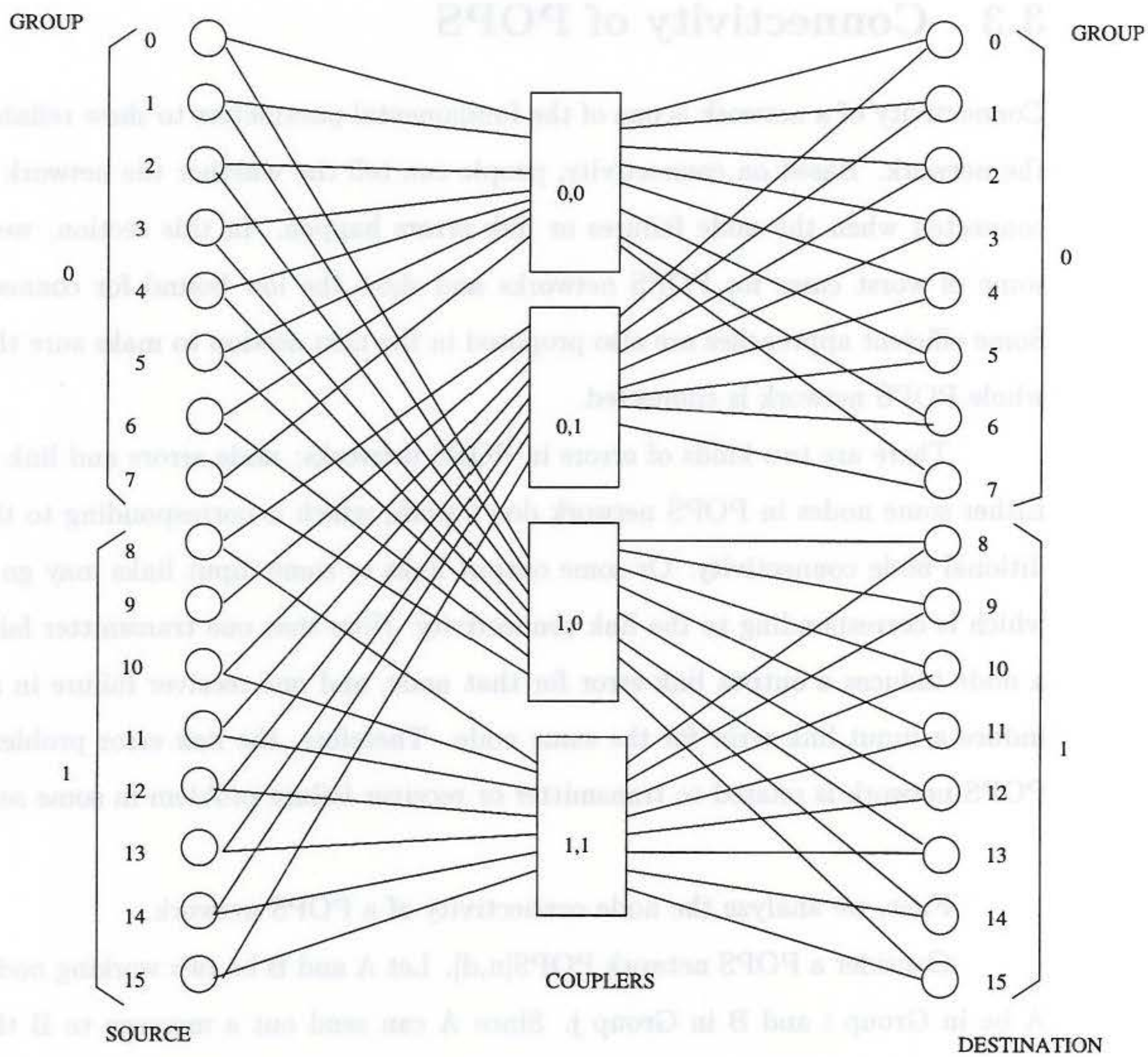


Figure 3.1: A POPS network with $n = 16, d = 8$, and $g = 2$.

the other hand, the small degree passive star coupler is cheaper, but the total number of couplers will increase to support the same number of nodes in the system.

3.3 Connectivity of POPS

Connectivity of a network is one of the fundamental parameters to show reliability of the network. Based on connectivity, people can tell the whether the network is still connected when the node failures or link errors happen. In this section, we study some of worst cases for POPS networks and show the low bound for connectivity. Some efficient approaches are also proposed in the next section to make sure that the whole POPS network is connected.

There are two kinds of errors in POPS networks: node errors and link errors. Either some nodes in POPS network don't work, which is corresponding to the traditional node connectivity. Or some output links or some input links may go down, which is corresponding to the link connectivity. Note that one transmitter failure in a node induces a output link error for that node, and one receiver failure in a node induce a input link error for the same node. Therefore, the link error problem in a POPS network is related to transmitter or receiver failure problem in some sense.

First, we analyze the node connectivity of a POPS network.

Consider a POPS network POPS[n,d]. Let A and B be two working nodes and A be in Group i and B in Group j. Since A can send out a message to B through A's transmitter to $C_{j,i}$ and B can receive the message from B's receiver from $C_{j,i}$, the connection from A to B and be set up. Similarly, the connection from B to A can be set up. This shows that the connection between any two working nodes can always be set up. We put this result in the following theorem.

Group	ouput links			
	1	2	...	s
0	(0,1)	(0,2)	...	(0,s)
1	(1,2)	(1,3)	...	(1,s+1)
2	(2,3)	(2,4)	...	(2,s+2)
...	...	...	...	...
g-2	(g-2,g-1)	(g-2,0)	...	(g-2,s-2)
g-1	(g-1,0)	(g-1,1)	...	(g-1,s-1)

Group	input links			
	1	2	...	s
0	(0,0)	(1,0)	...	(s-1,0)
1	(1,1)	(2,1)	...	(s,1)
2	(2,2)	(3,2)	...	(s+1,2)
...	...	...	...	...
g-2	(g-2,g-2)	(g-1, g-2)		(s-3, g-2)
g-1	(g-1,g-1)	(0,g-1)		(s-2,g-1)

Figure 3.2: An example with s input links and s output links

Theorem 3.3.1 *The POPS network is connected no matter how many nodes don't work.*

Theorem 3.3.1 tells us that POPS networks are very reliable with respect to node failures. The working nodes are always connected with each other. There is no need to worry about node fault tolerance issues in the remaining working system.

In the rest of this section, we discuss the link connectivity of a POPS network. It turns out the link connectivity is much more complicated than its node connectivity. We proposed some efficient approaches to guarantee that the whole system is still connected in the next section.

Consider the following example Figure 3.2 of a POPS network POPS[n,d]. Each node has up to s output link working and up to s input link working, and $s \leq g/2$. Let (i,j) denote the link to $C_{i,j}$ from transceivers.

We Claim that the whole system is totally disconnected for $g > 2$, i.e. no two nodes can communicate with each other. No sending message can be received by any nodes in this situation. This is the worst case happening to a system for network partitions.

Let $G > 2$. Let A and B be two nodes. A is in Group i and B in Group j. Suppose that there is a hop such that A's message can be received by B. There must exist p and q such that

$$(i, i+p) \equiv (j+q, j)(mod\ g), 1 \leq p \leq s\ and\ 0 \leq q \leq s-1.$$

That is equivalent to

$$i \equiv j+p\ and\ i+p \equiv j(mod\ g)$$

We have $i+j+p+q \equiv i+j(mod\ g)$, i.e. $p+q \equiv 0(mod\ g)$. Since $1 \leq p+q \leq 2s-1$ and $0 \leq 2s \leq g$, this is impossible.

Thus the sending message message can't be received by B. Because A and B can be any two station in the POPS network, we can say the whole network is totally disconnected in this situation.

Based on this example, we generalize the result to have

Theorem 3.3.2 *In POPS[n,d], each node has up to x input links working, and y output link working. If $x + y \leq g$ and $g > 2$, the whole network may be totally disconnected. No sending message could be received by any nodes in the network.*

Property 3.3.3 *In POPS[n,d], if each node has up to s input links working and s output link working, and $s \leq g/2 + 1$. Then the whole network may be partitioned into $g/2$ connected components.*

Based on the previous example, we add one more input link and one more output link for each node.

If $g = 2s$ and $g > 2$,

	(s+1)th IN	(s+1)th OUT
Group 0:	(0,0)	(s,0)
Group 1:	(1,1)	(s+1,1)
Group 2:	(2,2)	(s+2,2)
Group 3:	(3,3)	(s+3,3)
....		
Group s-1:	(s-1, s-1)	(2s-1,s-1)
Group s:	(s,s)	(0,s)
Group s+1:	(s+1,s+1)	(s,s+1)
....		
Group 2s-2:	(2s-2,2s-2)	(s-2,2s-2)
Group 2s-1:	(2s-1,2s-1)	(s-1, 2s-1)

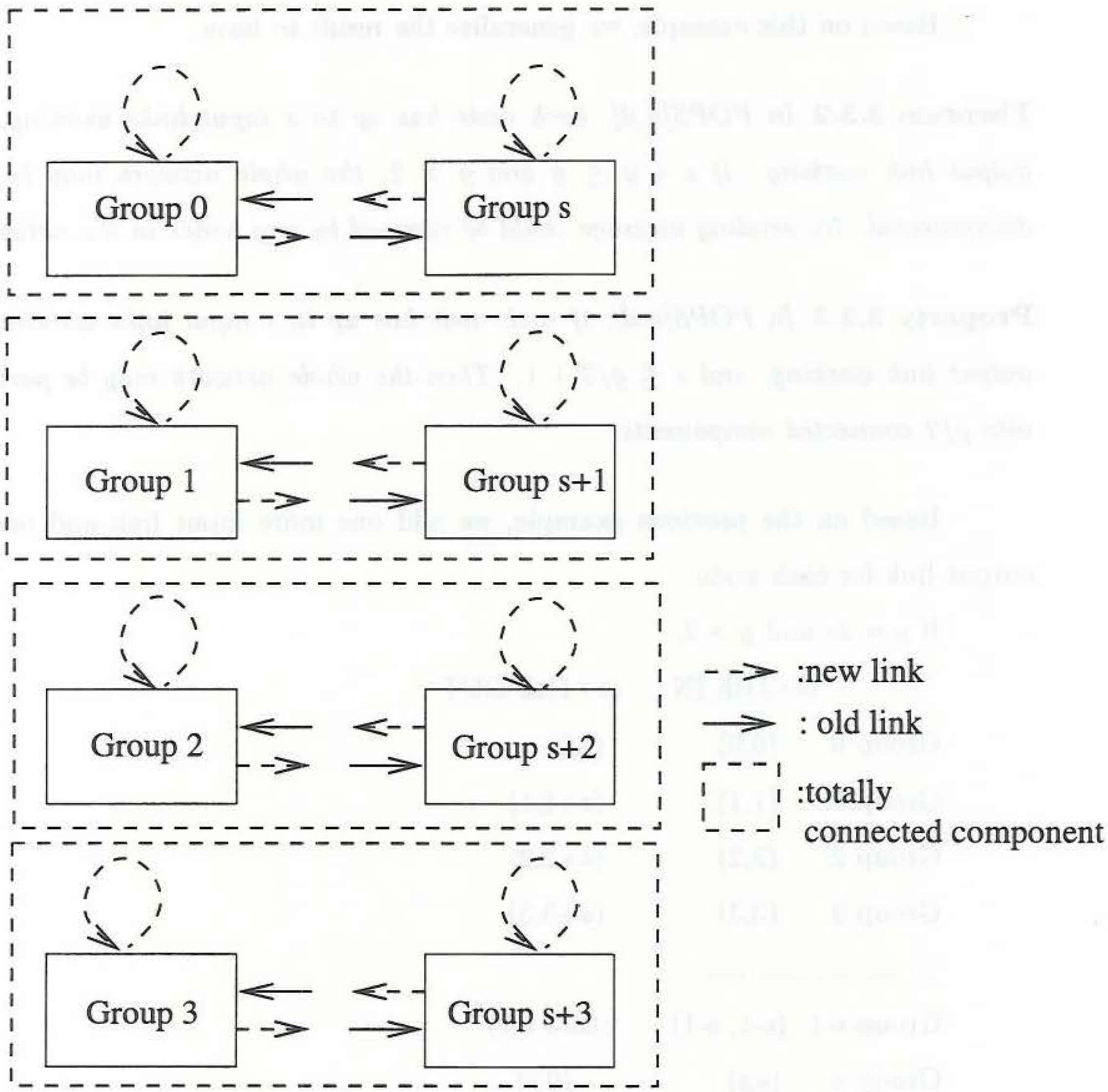


Figure 3.3: Connectivity for (s+1)th input/output links

Since we already know that there is connection between any old input links and any old output links, we only need to consider the effects of new input links and new output links. It is obvious all the new input links and all new output links are used to set up a path from one node to another. From Figure 3.3, we have that all the nodes in Group i and Group $i + g/2$ are fully connected with one another. There are $g/2$ connected component in this POPS network:

Component 1: Group 0 and Group $g/2$

Component 2: Group 1 and Group $g/2 + 1$

......

Component g/2: Group $g/2 - 1$ and Group $g - 1$

Property 3.3.4 *In POPS[n,d], if each node has up to s input links working and s output link working, and* $s \leq g/2 + 2$. *Then the whole network may be partitioned into* $g/4$ *connected components.*

We continue to add one more input link and one more output link for each node. Similarly to the above analysis, we only need to consider the effects of new input links and new output links. It is obvious all the new input links and all new output links are used to set up a path from one node to another.

Let $p = g/2$. From Figure 3.5, there are g/4 connected component in this POPS network:

Component 1: Group 0, 1, p, $p + 1$

component 2: Group 2,3, $p + 2, p + 3$

......

Component g/4: Group $p - 2, p - 1, g - 2, g - 1$

Property 3.3.5 *In POPS[n,d], if each node has up to s input links working and s output link working, and* $s \leq g/2 + 4$. *Then the whole network may be partitioned into g/8 connected components.*

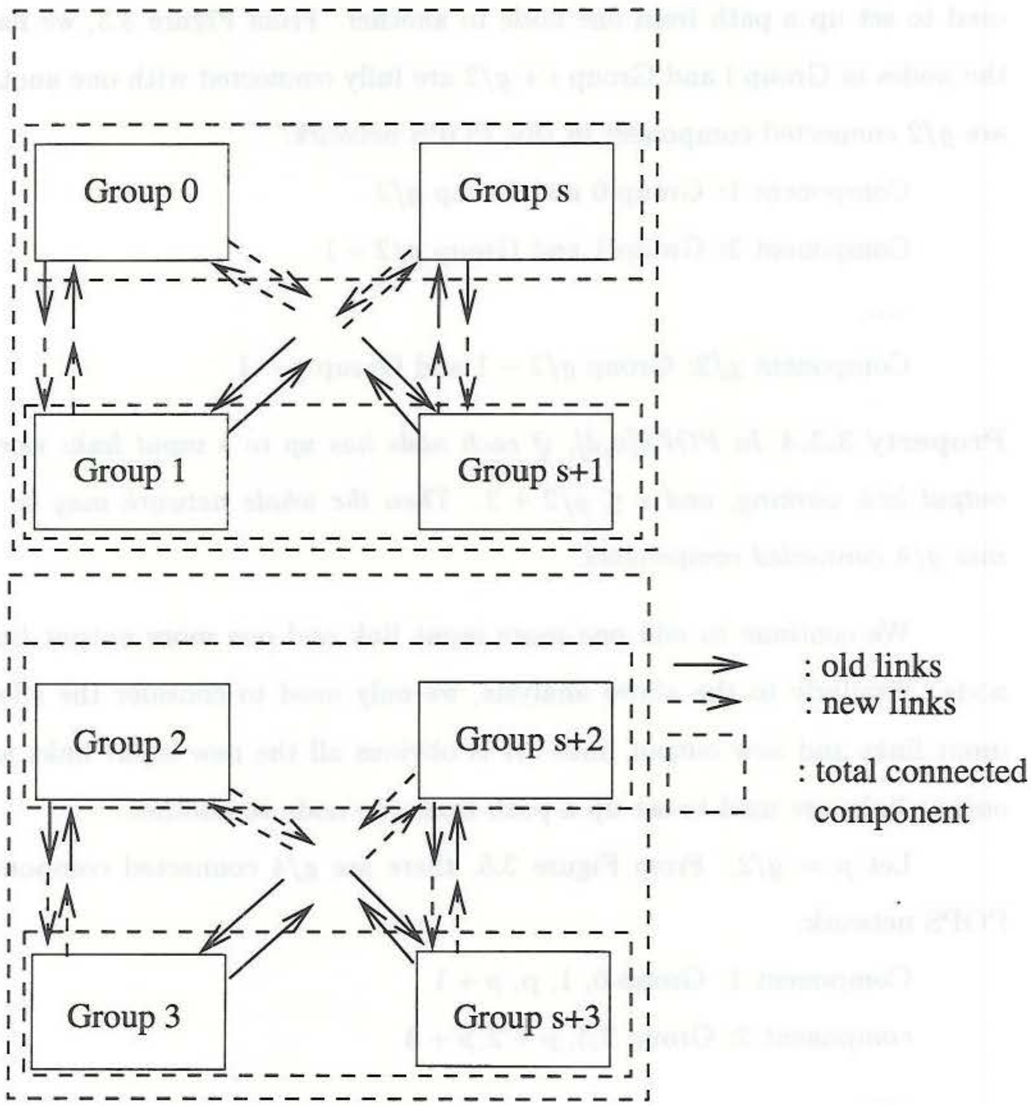


Figure 3.4: A demonstration for (s+2)th input/output links.

This time we add two more input links and two more output links for each node. Similarly to the above analysis, we only need to consider the effects of new input links and new output links. It is obvious all the new input links and all new output links are used to set up a path from one node to another.

Let $p = g/2$. From Figure 3.5, there are g/8 connected component in this POPS network:

Component 1: Group $0, 1, 2, 3, p, p+1, p+2, p+3$

......

Component g/8: Group $p-4, p-3, p-2, p-1, g-4, g-3, g-2, g-1$

This process can be continued until we get an lower bound for the link connectivity for a POPS network.

Theorem 3.3.6 *In POPS[n,d], even if each node has up to* 75 *percent input links working and up to* 75 *percent output link working, the whole network may be still disconnected.*

More specially, if each node has up to s input links working and s output link working, and $s \leq \frac{3}{4}g$*, there may be no less than two network connected component in the system.*

The total number of new input links or new output links for each node during the process:

$$1 + 1 + 2 + 4 + ... + g/8 = g/4$$

The total number of input links or output links after this process:

$$g/4 + g/2 = \frac{3}{4}g$$

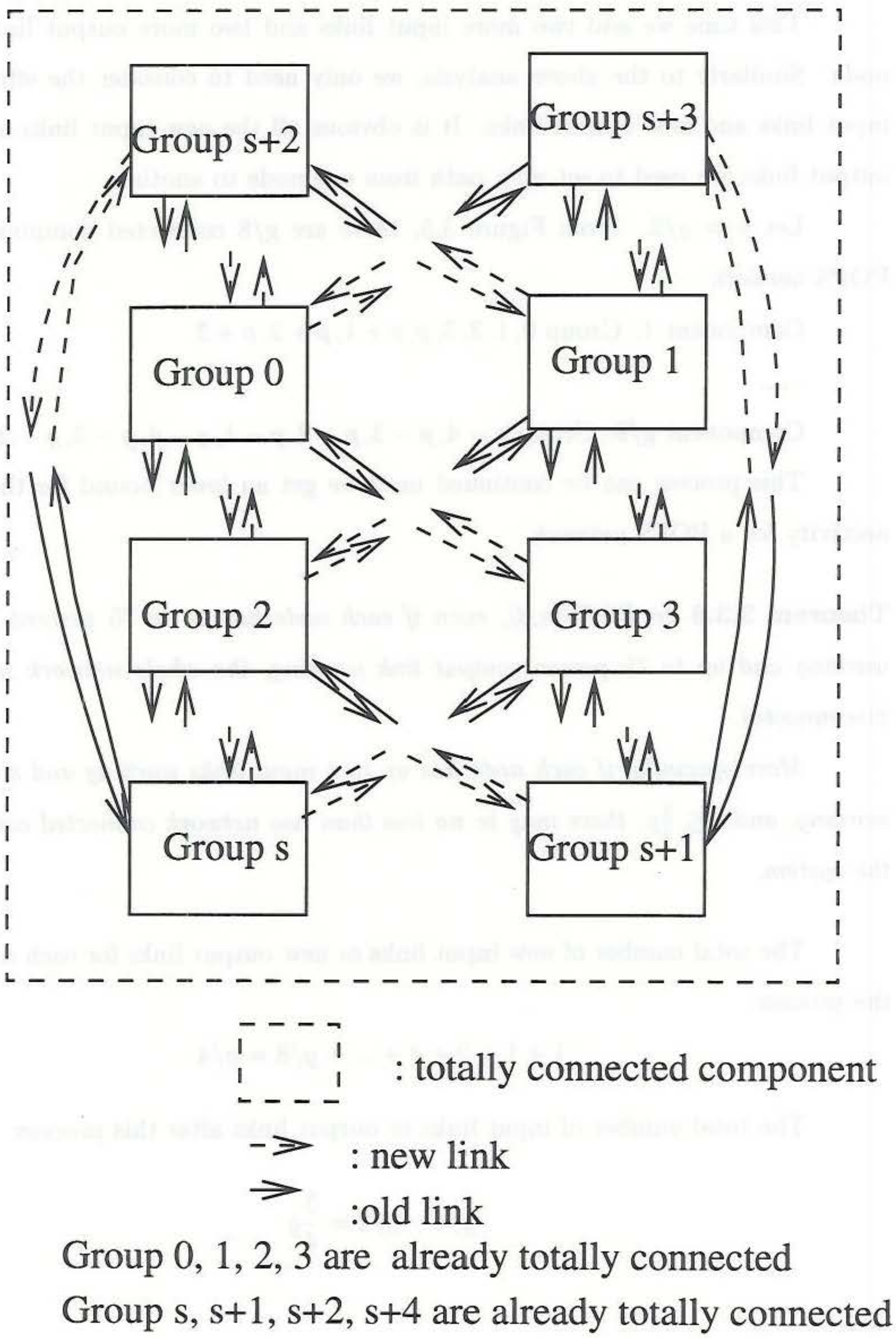


Figure 3.5: A demonstration for (s+3)th and (s+4)th input/output links.

3.4 Detection Approaches

Theorem 3.3.6 tells us that POPS networks are not reliable in link errors or transceiver errors. Even there are 75 percent links working in the system, we can't guarantee that the network is connected. So some efficient methods are needed to protect the network as a whole system without any partition.

Approach 3.4.1 *Each group is connected locally. If each node has no less than $g/2+1$ input link working and no less than $g/2+1$ output link working. A output link to $C_{i,j}$ from Group j if and only if there exist an input link to Group i working.*

This approach guarantees the network is connected. For any two nodes A and B, A in Group i and B in Group j (i and j may be equal). Let (i, p_t) be the $g/2+1$ inputs to $A, 0 \leq t \leq g/2$. Let (q_t, j) be the $g/2+1$ outputs from B.

By the Pigeon Principle, there exist α and δ such that

$$p_\alpha = q_\delta, 0 \leq \alpha \leq g/2 \ \ and \ \ 0 \leq \delta \leq g/2.$$

A has a input link from C_{i,p_α}. By the approach, there must be a output link from a node in Group p_α to C_{i,p_α} . B has a output link to $C_{q_\delta,j}$. By the approach, there must be a input link from a node in Group p_α to $C_{q_\delta,i}$. Since all the nodes in Group p_α are connected, we find a path from B to A through Group p_α.

Similarly, there is a path from A to B by path of a certain group. Therefore, the whole network is connected.

This is efficient approach with respect to each intermediate optical passive star couplers. By checking and fixing up links to each intermediate star coupler to satisfy this condition, the whole system is connected.

Approach 3.4.2 *(1)For each node in Group i, check to make sure its input link and output link to $C_{i,i}$ always work.*

(2)For Group i, $0 \leq i \leq g-1$, check to make sure there is at least one output link to $C_{i+1,i}$ and one input link from $C_{i,i-1}$.

This approach guarantee that the POPS network is connected.

Each group is locally connected for each node can send and receive message from $C_{i,i}$ by (1).

The links in (2) form a ring among all the groups. So any two groups can talk to each other based on this ring.

This approach asks the special treatment for some links. The advantage is the total number of links to guarantee the connectivity of the whole network is greatly reduced for checking these special links.

The total number of links checked: $2n + 2g$. The total number of links in the network: $2ng$. So we only check $(\frac{1}{n} + \frac{1}{g})$ of links to guarantee that there is no network partition.

3.5 Summary

We consider relability in the all-optical Partitioned Optical Passive Stars (POPS) topology. The reliability and fault tolerance in POPS rely on its connectivity. In this study, we analyze the worst case for network partitions due to either node failures or link errors. For node connectivity, we show that the whole system remains connectivity no matter how many nodes don't work. For link connectivity, we analyzed some worst cases due to link errors and the lower bound for connectivity is demonstrated. Some sufficient approaches are proposed to detect and keep connectivity of the whole system.

Chapter 4

Limit Tuning Range For Tunable Transceivers

Wavelength Division Multiplexing (WDM) has been widely used for studying the performance of optical networks, especially those employing optical passive star couplers. Many models have been proposed for WDM on an optical passive star coupler, such as each station equipped with a single tunable transmitter and a single fixed wavelength receiver, and each station with multiple tunable transmitters and multiple tunable receivers. The current technology only allows the transceivers to be tunable in a small range, a fact ignored in previous studies. In this chapter, we focus on the design of WDM optical passive star networks with tunable transmitters of limited tuning range and fixed wavelength receivers. The limited tuning range has effects on the maximum delay, the total number of wavelengths which can be used, and the topological embedding. Complete graphs, rings, meshes and hypercubes are the four topologies studied in this chapter. The relationship between the total number of wavelengths which can be utilized and the embedded topology is established. The bound for the maximum delay is analyzed. The optimal embedding algorithms are given for the systems embedded with one of the four topologies.

4.1 Introduction

The transceivers can be either tunable or fixed at a channel. In previous studies, some combinations of the transmitters and receivers have been considered. When the transceivers are tunable, the schedule of wavelengths for all transceivers becomes very important for achieving high performance of the system. Lee and etc. [39, 40] studied the transmission scheduling for the systems embedded with hypercubes and complete graphs, Cao and Borchers [11] gave the optimal schedules for the systems embedded with de Bruijn graphs.

One issue missing in previous research is the limited tunable range of transceivers. Each channel requires 1-2 nm for wide bandwidth and current technology can only support 3-7 nm for large bandwidth devices. This means the reasonable tunable range can only be 3 to 7 wavelengths with current technology. Even in the near future, the tunable range is not likely to increase significantly because of technical difficulties. In this chapter, we study the effects of the limited tunable range of transceivers on the optical passive star network with an embedded topology.

We assume that each station has one tunable transmitter and one fixed wavelength receiver. The range of the tunable transmitter is limited to k channels (k is a small number). There are two reasons for picking a fixed wavelength receiver over a tunable one. First, the cost of the system can be kept lower. Tunable devices (especially receivers) are considerably more expensive than fixed wavelength devices. Second, with a fixed wavelength receiver, a transmitter wishing to send a message only needs to tune its wavelength to that of the receivers fixed wavelength. There is no pre-transmission coordination required between the two to determine which wavelength to employ for communication as would be the case if a tunable receiver was to be used. We assume the receiver at each station occupies a wavelength within the range of its tunable transmitter. This permits easy design of transceivers and wavelength switching in the optical domain. Other requirements may become more clear when we discuss the details of the network.

Consider other models for the stations. In a system where each station with one fixed wavelength transmitter and one fixed wavelength receiver, it becomes a model of time multiplexing division for the whole system can only use one wavelength. In a system where each station has one tunable transmitter and several fixed wavelength receivers, the wavelengths of those receivers should be in the tuning range of its transmitter. Otherwise, the receiving signals can not be sent out. The system cost will increase for the number of fixed-wavelength receivers and the synchronization devices among them. Transmission schedule needs to be considered even within one station. In a system where each station has several tunable transmitters and one fixed wavelength receiver, let A be a station with channel a occupied by its receiver. Then channel a must be within the range of all its transmitters. Since the tuning range is limited to k channels, the union of the tuning range of all of A's transmitters is at most in the range [a-k+1, a+k-1]. This is akin to a system where each station has a fixed wavelength receiver and a tunable transmitter with range 2k-1 instead of k. For a system with multiple tunable transmitters and multiple fixed wavelength receivers at each station, we can decompose the system into several subsystems, with each subsystem having only one tunable transmitter and one fixed receiver. Thus our assumption is reasonable and provides the basis for other types of systems as well.

Topology embedding is very important for communication and high performance parallel computing. With a limited tunable range for each transmitter, one station can only communicate with those stations whose receivers are in its tunable range, in one hop. So the topological embedding determines the nature of communication between any two stations. We can determine a path between any two stations and the maximal network delay only when the topological embedding is given. For scientific computing, some special topologies may need to be embedded into the system to more accurately reflect the computing model used.

We assume that there are a total of p wavelengths available for the whole system, and the tunable range is no more than k channels and each station has one tunable transmitter and one fixed receiver. Each station has a tunable transmitter and a fixed receiver. Let A be a station in the system and its transmitter be tunable in the range of [a, a+k-1]. The wavelength of A's receiver must be in [a, a+k-1]. Only those nodes with the wavelength of their receivers in [a, a+k-1] can receive the message from A directly, i.e. it will take one hop to reach those nodes. For other stations, it takes multi-hop transmission to receive the message

This Chapter is organized as follows. In Section 4.2, the basic network protocol is proposed for setting up the connections and solving the transmission collisions. In Section 4.3, we consider the optical passive star network embedded with the complete graph, i.e. each station is connected to all stations in the system. Since it is a multi-hop system, the bounds for the maximum delay in the virtual topology embedding without congestion are studied and embedding algorithms designed to be optimal in terms of the maximum delay are proposed. The transmission schedule is discussed for common parallel communications. In Section 4.5, meshes, an important topology for computing, are considered as the embedded virtual topology. The relationship between the structure of the mesh and the maximum number of wavelengths which can be used is analyzed. The algorithm for topological embedding is designed to maximize the use of the available channels. In Section 4.6, hypercubes, another important topology for parallel architectures, are studied. The upper bound for the maximum number of wavelengths used is shown. Dynamic programming algorithms are designed for the topological embedding. In Section 4.7, some general results are presented for another assumption on receiver's wavelengths. In Section 4.8, we conclude this chapter and discuss the possibility of extending our work to other models and other topological embeddings.

4.2 Basic Network Protocol

In a WDM optical passive star network, a network protocol is needed to create a connection between the sending station and the receiving station and to handle transmission collisions in some channels.

We consider systems embedded with topologies other than complete graphs. Under our assumption, each station can communicate with its neighboring stations directly. Let A be a station with [a, a+k-1] being the tuning range of its transmitter. The wavelengths occupied by the receivers of A and all its neighboring stations must be in [a, a+k-1]. There are big advantages for this constraint. First, the virtual topology is actually embedded into the system. Each link in the topology can be implemented with one hop in the system. The properties of the topology are kept in the system. Without this assumption, some links may correspond to multiple hops, which destroys the characteristics of the topology. Second, routing and network control become easier. Based on the topology, it is not difficult to find the shortest path between any two stations. The communication feedback can be also implemented easily.

For each station, a routing table is created based on its position in the embedded topology. Using Bell-Ford algorithm , it is easy to find all the shortest paths from a station to all other stations. The routing table is defined as follows:

Based on our assumption, every link in the embedded topology can be implemented by one hop in the system. Not only is the routing table well-defined, but the routing table also is easy to create and guarantees the shortest path between any two stations. Each station also has a table for the wavelength occupied by the receivers of its neighboring stations.

We use a variation of IEEE 802.3 as the transmission protocol. Carrier Sense

Multiple Access with Collision Detection (CSMA/CD) is used on each channel to solve the problem of transmission collision.

We will first discuss several ways of implementing a network protocol based on CSMA/CD in the proposed optical network environment. The easiest WDM optical passive star network for implementing CSMA/CD requires that each station has a fixed wavelength receiver, one tunable transmitter and one tunable receiver with the same tuning range. The fixed wavelength receiver receives the data from other stations. The tunable transmitter sends the data to its neighboring stations. The tunable receiver acts as a carrier sensor for CSMA/CD. The tunable receiver always tunes its wavelength to that of the tunable transmitter and listens on that channel. When the tunable receiver finds that channel is available, it will inform the tunable transmitter to begin the transmission. When the tunable receiver finds there is a collision on that channel, it will inform the tunable transmitter to stop the current transmission. When the tunable receiver finds that channel is busy, it will inform the tunable transmitter to wait until it finds that channel is available. One difference between this system and the normal CSMA/CD is that the wavelength used by the tunable transmitter and the tunable receiver must be different from the wavelength occupied by the fixed-wavelength receiver. Otherwise, collisions on that channel are

Destination Station	Next Station	Channel to next station
0	i_0	j_0
......	......	......
n-1	i_{n-1}	j_{n-1}

Figure 4.1: The Routing Table Format for One Station

inevitable.

Now we consider how to implement the variation of IEEE 802.3 on the system in which each station has a tunable transmitter and a fixed-wavelength receiver. We propose Time Division Multiplexing on the WDM optical passive star network. i.e. the system is TWDM.

We assign some time slots for acknowledgement only. This is the only way to avoid the collisions between the acknowledgements and the data and to guarantee that acknowledgements are received. Since the acknowledgement packet is much smaller than the data packet on each time slot, we further divide the time slot for the acknowledgement into subslots, where each subslot for the acknowledgement stands for a previous time slot for data packets.

For the sending station, it sends its data packet in the time slots for data packets. In the time slot for the acknowledgement, it listens to its receiver's channel. Note that there is only one receiving channel for each station. If there is no collision in the time slot during its transmission, it receives the acknowledgements from the receiving station. In the next slot for the data packet, it continues the data transmission if it has more to transmit. If there is a collision in the time slot for the data packet, it does not receive any acknowledgement in the following time slot for acknowledgement. In one of the next few time slots for data packets, it resends its data with a certain possibility to avoid the collision according to CSMA/CD protocol. Therefore, this is based on a positive acknowledgement scheme. That is, an acknowledgement is received if there is no collision. However, there will be no acknowledgement if the data packet is collided with other transmissions.

For the receiving station, it listens to its receiver's channel and checks if there is any collision in each slot for data packets. In each time slot for the acknowledgement corresponding to a successful transmission, it tunes its transmitter to the wavelength occupied by the receiver of the sending station by the routing table and the header of the data packet, and sends the acknowledgement. In each time slot for acknowledgement with a collision in the previous time slot, it doesn't send out anything.

If a station is both a sending station and a receiving station, it will simultaneously send data and receive data in the time slots for data packets, and simultaneously listen to the acknowledgements and send the acknowledgement in the time slots for acknowledgement.

In Figure 4.3 and Figure 4.2, we show an example of this protocol. Every time slot for data packe ts is followed by a time slot for the corresponding acknowledgement. Another approach is to allocate a dedicate time slot for acknowledging a bunch of previous time slots for data packets instead of acknowledging only one time slot for data packets at a time.

Another way to solve collisions is to assign a fixed round-robin scheme for the transmission of each station. When station A wants to talk to station B which is one of A's neighboring stations, it will wait for the corresponding slot for B according to the scheme. In the subslots for acknowledgement, all the neighboring stations of A also send back the acknowledgements according to the round-robin scheme. This approach is much simpler for finding the collision and reducing the possibility of collision than the previous approach, but it is very wasteful for each time slot of data packets to be dedicated to one station and they can not be borrowed (used) by other stations. For a heavily-loaded system, we suggest using the round-robin scheme for the transmission of each station. For a lightly-loaded system, the previous approach will be better.

The format for data packet can be designed as follows:

The format for acknowledge packet can be designed as follows:

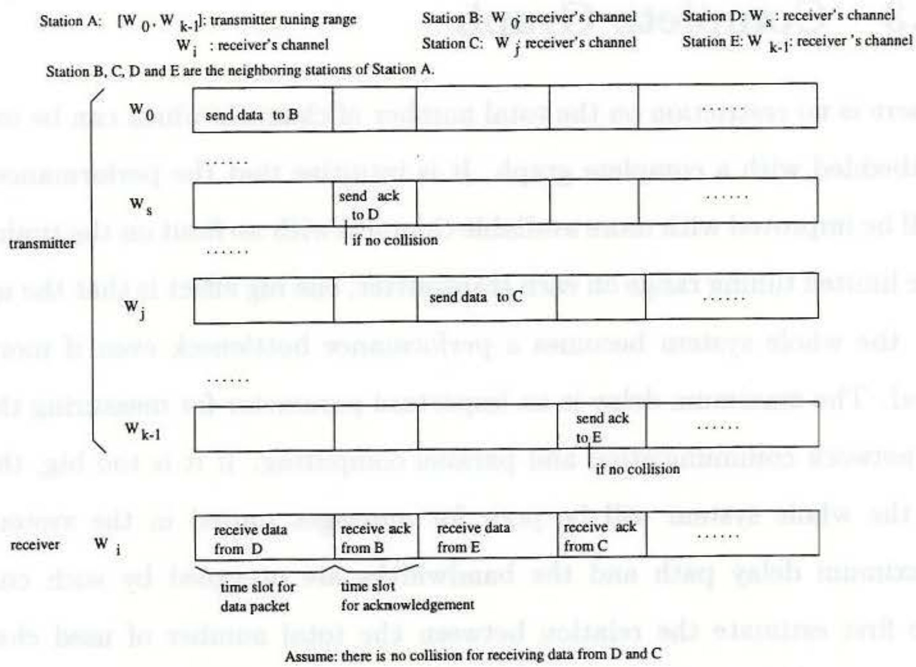


Figure 4.2: The Sending and Receiving for Station A

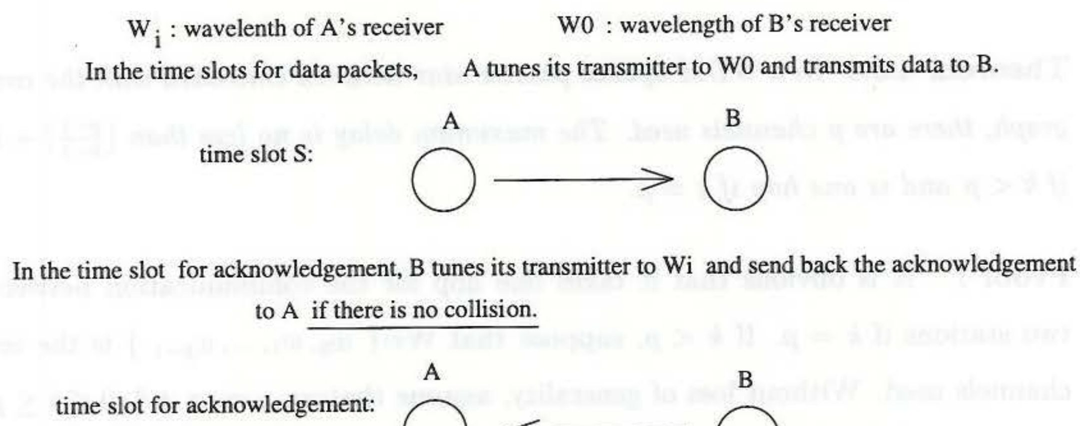


Figure 4.3: Two Stations in Sending and Receiving

source station	destination station	sending station	data

Figure 4.4: The Data Packet Format On One Channel

4.3 Complete Graph

There is no restriction on the total number of channels which can be used in a system embedded with a complete graph. It is intuitive that the performance of the system will be improved with more available channels with no limit on the tuning range. With the limited tuning range on each transmitter, one big effect is that the maximum delay for the whole system becomes a performance bottleneck even if more channels are used. The maximum delay is an important parameter for measuring the performance of network communication and parallel computing. If it is too big, the performance of the whole system will be poor for messages routed in the system through the maximum delay path and the bandwidths are occupied by such communications. We first estimate the relation between the total number of used channels and the maximum delay to help us understand the effect of bounded tunable range for the transmitters.

Theorem 4.3.1 *In a WDM optical passive star network embedded with the complete graph, there are p channels used. The maximum delay is no less than* $\lfloor \frac{p-1}{k-1} \rfloor + 1$ *hops if* $k < p$ *and is one hop if* $k = p$.

Proof : It is obvious that it takes one hop for the communication between any two stations if $k = p$. If $k < p$, suppose that W=$\{ w_0, w_1, ..., w_{p-1} \}$ is the set of p channels used. Without loss of generality, assume that $w_{i+1} = w_i + 1, 0 \leq i \leq p-2$.

Case 1: (k-1) | (p-1). Consider the communication between a station A whose receiver occupies w_0 and a station B whose receiver occupies w_{p-1}. Suppose that

subslot number	sending station	receiving station	collision information	

Figure 4.5: The Acknowledge Packet Format for One Subslot On one Channel

k	3	4	$p^{\frac{1}{3}}$	$p^{\frac{1}{2}}$	$\frac{1}{2}p$	p
maximum delay	$\lfloor\frac{p-1}{2}\rfloor+1$	$\lfloor\frac{p-1}{3}\rfloor$	$\lfloor\frac{p-1}{p^{\frac{1}{3}}-1}\rfloor+1$	$\lfloor p^{\frac{1}{2}}\rfloor+2$	3	1

Table 4.1: Lower bounds for the maximum delay

the maximum delay is no more than $\frac{p-1}{k-1}$ hops, there exists $\frac{p-1}{k-1}-1$ stations C_i, $1 \leq i \leq \frac{p-1}{k-1}-1$, such that C_i's receiver occupies $w_{i*(k-1)}$ and C_i's transmitter is tunable in $[i*(k-1),(i+1)*(k-1)]$. The communication from A to B passes through all C_i. Now consider the communication from $C_{\frac{p-1}{k-1}-1}$ to A. Since the tuning range of $C_{\frac{p-1}{k-1}-1}$ is $[p-k,p-1]$, $C_{\frac{p-1}{k-1}-1}$ has to send its message to one station D whose receiver occupies one channel in $[p-k,p-1]$. Then D forwards $C_{\frac{p-1}{k-1}-1}$'s message to A.

If there exists i such that C_i is the only one station whose receiver occupies $w_{i*(k-1)}$, it will take at least $\frac{p-1}{k-1}$ hops from D to A. Thus the maximum delay is no less than $\frac{p-1}{k-1}+1$. Otherwise, the maximum delay is no less than $\frac{p-1}{k-1}$.

Case 2: k-1 is not a factor of p-1. Consider the communication between a station A whose receiver occupies w_0 and a station B whose receiver occupies w_{p-1}. It is easy to see that it needs at least $\lfloor\frac{p-1}{k-1}\rfloor+1$ to send a message from A to B. ▮

The following table demonstrates the lower bound of the maximum delay in the above theorem.

In the next step, we design an algorithm, *Complete-Graph-Embedding*, for the topological embedding. It has the optimal maximum delay. That means it is best in such a multihop system. Without loss of generality, we can assume that $N \geq p$.

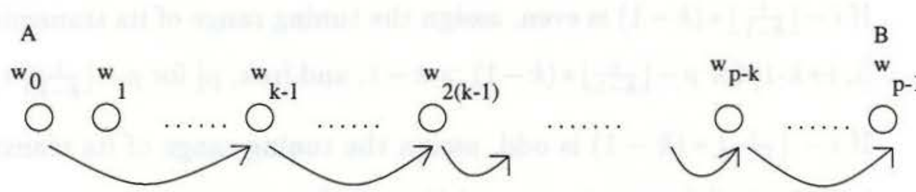


Figure 4.6: The path from A to B

If $N \geq 2 * p$, it is possible that each channel are occupied by two fixed wavelength receivers. So we consider $N \geq 2 * p$ and $N \leq 2 * p$ seperately.

Algorithm *Complete-Graph-Embedding* for $N \geq 2 * p$

1. **Receivers' wavelength assignment**

   Divide N stations into p disjoint sets, N_i with $|N_i| \geq 2$ $(0 \leq i \leq p-1)$. Assign wavelength w_i to the receivers of the stations in N_i. Divide N_i into two disjoint sets $L[i,1]$ and $L[i,2]$ such that $||L[i,1]| - |L[i,2]|| \leq 1$.

2. **Transmitters' wavelength assignment**

   For each station in $L[i,1]$, assign its tuning range of its transmitter to be $[i-k+1, i]$ for $i \geq k-1$ and $[0, k-1]$ for $i < k-1$. For each station in $L[i,2]$, assign its tuning range of its transmitter to be $[i, i+k-1]$ for $i \leq p-k$ and $[p-k, p-1]$ for $i > p-k$.

Algorithm *Complete-Graph-Embedding* for $N < 2 * p$

1. **Receivers' wavelength assignment** Divide N stations into p disjoint sets, N_i with $|N_i| \geq 1$ $(0 \leq i \leq p-1)$. Assign wavelength w_i to the receivers of the stations in N_i.

2. **Transmitters' wavelength assignment**

   If $i - \lfloor \frac{i}{k-1} \rfloor * (k-1)$ is even, assign the tuning range of its transmitter in N_i to be [i, i+k-1] for $p - \lfloor \frac{i}{k-1} \rfloor * (k-1) > k-1$, and [p-k, p] for $p - \lfloor \frac{i}{k-1} \rfloor * (k-1) \leq k-1$.

   If $i - \lfloor \frac{i}{k-1} \rfloor * (k-1)$ is odd, assign the tuning range of its transmitter in N_i to be [i-k+1, i] for $i \geq k-1$, and [0, k-1] if $0 \leq i \leq k-1$.

In the stage of Receivers' wavelength assignment, the grouping of N_i can be quite convenient. The only constraint here is that the size of each set should be no less than two. In the stage of Transmitters' wavelength assignment, the wavelengths of the transmitters are uniformly distributed. An observation of the system is as following:

Theorem 4.3.2 *The maximum delay in the system generated by Complete-Graph-Embedding is no more than $\lfloor \frac{p-1}{k-1} \rfloor + 1$. This shows its has the optimal maximum delay.*

Proof : We first consider $N \geq 2 * p$.

Since $|N_i| \geq 2$ for $0 \leq i \leq p-1$, there is at least one station in $L[i,1]$ and $L[i,2]$ respectively. Let A and B be two stations in the system with their receivers' occupying w_i and w_j respectively with $i \leq j$. Consider the communication between A and B.

If A $\in$ L[i, 1], A can send its message to another station C which is in L[i,2] by tuning its transmitter to w_i. C sends that message from A to B by pass those stations which is in L[l, 2] for $l = i + p * (k-1), 1 \leq p \leq \lfloor \frac{j-i}{k-1} \rfloor$. So the communication from A to B is no more than $\lfloor \frac{p-1}{k-1} \rfloor + 1$ hops.

Similarly, we can route the message from B to A with no more than $\lfloor \frac{p-1}{k-1} \rfloor + 1$ hops.

For $N \leq 2 * p$, we can route the message between A and B in the way as the proof in Theorem 4.3.1 . By the above theorem, we know that *Complete-Graph-Embedding* provides the optimal maximum delay. ∎

By *Complete-Graph-Embedding*, the whole system has the optimal maximum delay. For example, the system with N=22, p=7 and k=3. The channel graph generated by Algorithm *Complete-Graph-Embedding* is:

$N_i = \{3i, 3i+1, 3i+2\}, L[i,1] = \{3i\}$, and $L[i,2] = \{3i+1, 3i+2\}$ for $0 \leq i \leq 5$. $N_6 = \{18, 19, 20, 21\}, L[6,1] = \{18, 19\}$ and $L[6,2] = \{20, 21\}$.

Another example is for N=12, p=7 and k=3. Let N_i be nonempty for $0 \leq i \leq 6$.

The optimal maximum delay can be achieved by our algorithm for the channel assignment. One important issue for the system embedded with the complete graph is the heavy load on some channels for uniform communication among the stations. Those channels can be the bottlenecks of the whole system with heavy or bursty loads. But this phenomenon is one of the key characteristics for a system of limited tuning range for each transmitter embedded with the complete graph . No matter how the channels are assigned, the bottlenecks are always there. The main reason is that the tunable range is bounded by k and the communications between the stations with low numbered channels and those with high numbered channels have to traverse the system along the intermediate channels.

Theorem 4.3.3 *In a system with limited tuning range for each transmitter embedded with the complete graph, there always exist k consecutive channels which are the bottleneck for communication if the communication is uniformly distributed.*

Proof : Suppose that $W = \{w_0, w_1, ..., w_{p-1}\}$ is the set of used channels with $w_{i+1} = w_i + 1$. Let $[w_i, w_{i+k-1}]$ be the range which divides the whole used channels into three parts: X is the set of stations whose receivers occupy a channel in $[0, w_{i-1}]$, Y the set of stations whose receivers occupy a channel in $[w_i, w_{i+k-1}]$, and Z the set of stations whose receivers occupy a channel $[w_{i+k-1}, w_{p-1}]$ with $||X| - |Z||$ as small as possible. Define $x = |X|, y = |Y|, z = |Z|$, and N=x+y+z.

Since the communication is uniformly distributed, we consider the receivers workload in $[w_i, w_{i+k-1}]$ for all-to-all communication in such a environment.

	w_0	w_1	w_2	w_3	w_4	w_5	w_6
L[0,1]	X Y	Y	Y				
L[0,2]	X Y	Y	Y				
L[1,1]	Y	XY	Y				
L[1,2]		XY	Y	Y			
L[2,1]	Y	Y	X Y				
L[2,2]			X Y	Y	Y		
L[3,1]		Y	Y	X Y			
L[3,2]				X Y	Y	Y	
L[4,1]			Y	Y	X Y		
L[4,2]					X Y	Y	Y
L[5,1]				Y	Y	X Y	
L[5,2]					Y	X Y	Y
L[6,1]					Y	Y	XY
L[6,2]					Y	Y	X Y

X: the wavelength for receivers, Y: the wavelength for tunable transmitters

Figure 4.7: Wavelength assignment for the system with N=22, p=7 and k=3.

	w_0	w_1	w_2	w_3	w_4	w_5	w_6
N_0	X Y	Y	Y				
N_1	Y	X Y	Y				
N_2			XY	Y	Y		
N_3		Y	Y	X Y			
N_4					X Y	Y	Y
N_5				Y	Y	X Y	
N_6					Y	Y	X Y

X: the wavelength for receivers, Y: the wavelength for tunable transmitters

Figure 4.8: Wavelength assignment for the system with N=12, p=7 and k=3.

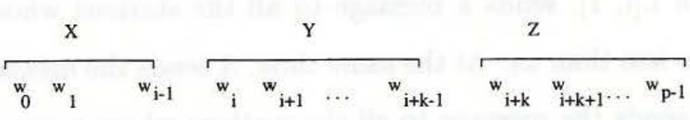


Figure 4.9: The partition of the stations

$$
\begin{aligned}
workload &= x(y+z) + z(x+y) + y(y-1) \\
&= y(x+z) + 2xz + y(y-1) \\
&= y(N-y) + 2xz + y(y-1) \\
&= O(Ny - y + 2 * (\frac{N-y}{2})^2) \\
&= O(\frac{1}{2}(y^2 - 2y + N^2)) \\
&= O(\frac{1}{2}(y-1)^2 + \frac{1}{2}N^2 - \frac{1}{2}) \\
&= O(N^2)
\end{aligned}
$$

The workload on $[w_i, w_{i+k-1}]$ is in the order of N^2. It will take $O(N^2/k)$ for the transmission of this workload. We have specified that k is much smaller than p. This means that the wavelength range from w_i to w_{i+k-1} will be quite congested and overloaded. Now matter how the complete graph is embedded, such a tuning range always can be found according to the above construction. Therefore there always exist k consecutive channels which are the bottleneck of communication if the communication is uniformly distributed.

∎

We discuss how to implement one-to-all broadcasting for parallel computing by using Algorithm *Complete-Graph-Embedding* for $N \geq 2 * p$. One-to-all broadcasting is one of the most important operations in parallel computing. Let A be a station in N_i. A in L[i, 1], sends a message to all the stations whose receivers occupy the wavelengths less than w_i. At the same time, A sends the message to a station B in L[i, 2]. B then sends the message to all the stations whose receivers occupy wavelengths greater than w_i.

This shows that this algorithm is flexible and useful for future Parallel Virtual Machines (PVM) on such a system embedded with the complete graph.

4.4 Rings

Rings are one of the simplest network topologies. Rings are used widely in local area computer networks for its simple structure and easies to be maintained [1, 43]. for example, token rings and FDDI are based on rings for communication, operation and maintain.

In this section, we study how to embed rings in an optical passive star coupler and how to maximize the number of wavelengths for such a system. This will give us some hints for the following sections about meshes and hypercubes.

Before exploring the embedding of a ring into an optical passive star, we demonstrate how to embed a line segment.

A line L[n] has n nodes, $\{0, 1, ..., n-1\}$, and node i is connected to i-1 and i+1 $(1 \le i \le n-2)$. An example is shown in Figure 4.10.

Lemma 4.4.1 *If $n > 2$ is even, the total number of wavelengths for L[n] which can be used is no more than $\frac{n}{2}(k-1) + max\{k-2, 0\} + 1$.*

If $n > 2$ is odd, the total number of wavelengths for L[n] which can be used is no more than $\frac{n+1}{2}(k-1) + 1$.

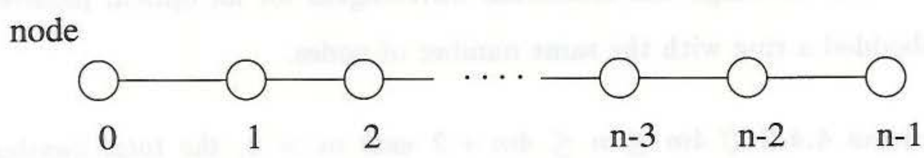


Figure 4.10: An example of L[n]

We just show how to proof the above lemma for $n = 4$ and $n = 5$ in Figure 4.19. For $n > 5$, the proof is similar. If $n = 4$, we consider the receiver's wavelength of each node. Let $r(i)$ denote the receive's wavelength of node i. Assume $r(0) = a$, we must have that $|r(2) - r(0)| \leq k - 1$ for node 2 and 0 are both neighbors of node 1. The transmitter of node 1 should cover both $r(0)$ and $r(1)$. Without loss of generality, assume $r(2) = a + k - 1$. Similarly, $|r(3) - r(1)| \leq k - 1$. The total wavelengths for transmitters are no more than

$$[r(2)+k-1]-[r(1)-(k-1)] = 2(k-1)+[r(2)-r(1)] \leq \frac{n}{2}(k-1)+max\{k-2,0\}$$

.

If $n = 5$, we consider the receiver's wavelength of each node. Let $r(i)$ denote the receive's wavelength of node i. Assume $r(0) = a$, we must have that $|r(2) - r(0)| \leq k - 1$ for node 2 and 0 are both neighbors of node 1. The transmitter of node 1 should cover both $r(0)$ and $r(1)$. Without loss of generality, assume $r(2) = a + k - 1$. Similarly, $|r(4) - r(2)| \leq k - 1$ and let $r(4) = a + 2(k - 1)$. The total wavelengths for transmitters are no more than

$$[r(3) + k - 1)] - [r(1) - (k - 1)] = 2(k - 1) + [r(3) - r(1)] \leq \frac{n+1}{2}(k - 1)$$

Now we discuss the total range of receiver's wavelengths of all nodes in the system for a ring with n nodes. Figure 4.12 shows $n = 5, 6, 7$, and 8. They represent four cases to assign the maximum wavelengths for an optical passive star network embedded a ring with the same number of nodes.

Lemma 4.4.2 *If $4m \leq n \leq 4m + 2$ with $m > 0$, the total number of receiver's wavelengths of all nodes is no more than $m(k - 1) + 1$.*

If $n = 4m + 3$ with $m > 0$, the total number of receiver's wavelengths of all *nodes is no more than* $(m+1)(k-1)+1$.

The receiver's wavelength assignments are displayed in Figure 4.12. In each case, we only consider the odd nodes. Because the increasement or decreasement of wavelengths from node i to i+2 is k-1, the receiver's wavelength of node i must fall in the interval of wavelengths of node i and i+2. That means the range of all receiver's wavelengths of odd nodes are enough for our study.

4.5 Meshes

A mesh M[c,d], is a set of nodes $V(M[c,d]) = \{(x,y) \mid 0 \le x \le c-1, 0 \le y \le d-1\}$, and two nodes, (x_1, y_1) and (x_2, y_2), are connected by an edge iff $|x_1 - x_2| + |y_1 - y_2| = 1$. It is easy to see that diameter of M[n,m] is $n+m-2$. So we have

Lemma 4.5.1 *The maximum delay in a system embedded with M[n,m] is* $n+m-2$.

One effect of the bounded range of the tunable transmitters is the total number of channels which can be used in the whole system. Because of the presence of tunable transmitters of a limited range, the total number of channels will be restricted to a certain range. Thus the channel assignment becomes more difficult than before where tunable transmitters of unlimited range were considered. With unlimited tuning range, one station can tune its transmitter's channel to any channel. The following theorem tells us the upper bound for the number of channels which can be used for the system.

Theorem 4.5.2 *The total number of wavelengths which can be used in a system with the embedded M[n,m] is no more than* $(2k-2) * \lceil \frac{n+m}{4} \rceil + k$.

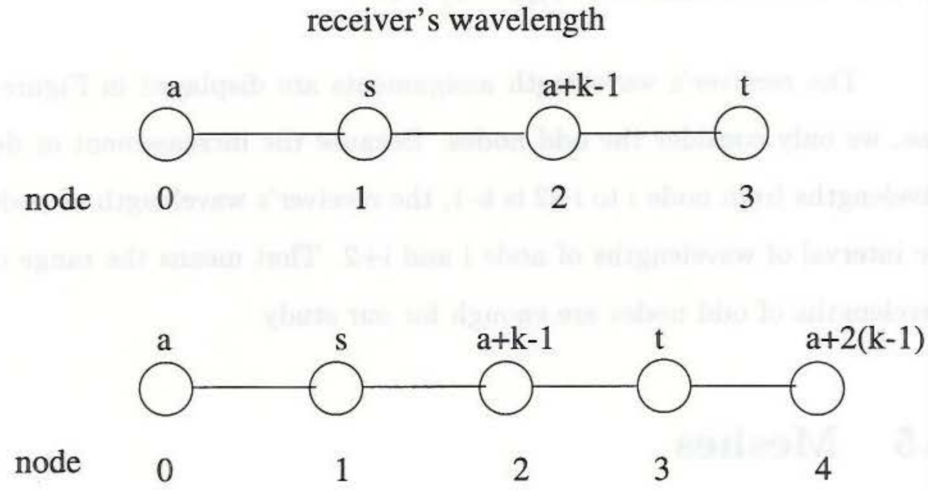


Figure 4.11: Proof for L[4] and L[5]

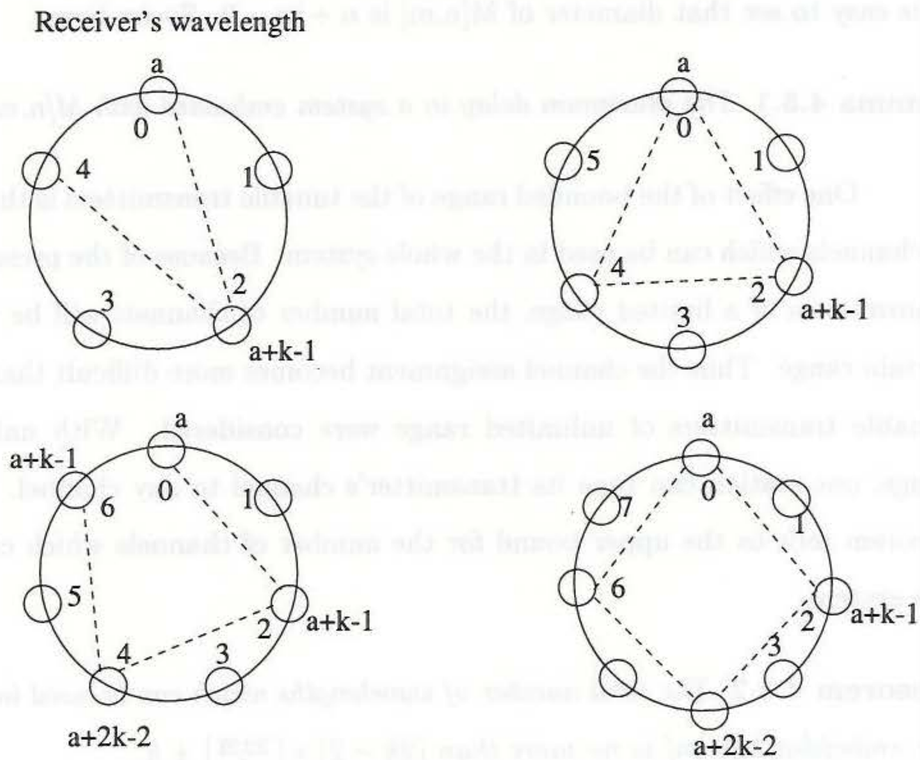


Figure 4.12: Four cases for rings

Proof : Let A= $(\lfloor \frac{n}{2} \rfloor, \lfloor \frac{m}{2} \rfloor)$ = (a, b) be the center point of M[a, b]. Define $L_i = \{(x,y) : |x-a|+|y-b| = 2i\}$, $1 \le i \le \lceil \frac{n+m}{4} \rceil$. Let w be the channel occupied by A's receiver.

For any station B in L_1, there exists a station C such that A and B must be the two neighboring stations of C. That means the set of channels occupied by the receivers in L_1 must be of the form $[s_1, t_1]$ with $w \in [s_1, t_1]$ and $|t_1 - s_1| \le 2k - 1$.

Similarly, we can determine the set of channels occupied by the receivers in L_2. It must be of the form $[s_2, t_2]$ with $[s_1, t_1] \subseteq [s_2, t_2]$ and $t_2 - s_2 \le 4k - 3$.

Thus the set of channels occupied by the receivers in $L_{\frac{n+m}{4}}$ must be of the form $[s_{\frac{n+m}{4}}, t_{\frac{n+m}{4}}]$ with $[s_{\frac{n+m}{4}-1}, t_{\frac{n+m}{4}-1}] \subseteq [s_{\frac{n+m}{4}}, t_{\frac{n+m}{4}}]$ and

$$t_{\frac{n+m}{4}} - s_{\frac{n+m}{4}} \le 1 + (2k-2) * \lfloor \frac{n+m}{4} \rfloor$$

Let $A' = (\lfloor \frac{n}{2} \rfloor, \lfloor \frac{m}{2} \rfloor + 1) = (a', b')$. Define $L_i' = \{(x,y) : |x-a'|+|y-b'| = 2i\}$, $1 \le i \le \lceil \frac{n+m}{4} \rceil$. Let w' be the channel occupied by A''s receiver. Since A' is adjacent to A, $|w' - w| \le k - 1$. The same argument can apply to L_i'. Because of the relation between w' and w, the wavelengths used by L_i' differentiate with those of L_i with at most k-1 different channels. Note that the union of L_i and L_i' is the vertex set of M[n,m].

So the total number of wavelengths are no more than

$$1 + (2k-2) * \lceil \frac{n+m}{4} \rceil + (k-1) = (2k-2) * \lceil \frac{n+m}{4} \rceil + k$$

∎

A table demonstrates the upper bound of the number of channels in the above theorem.

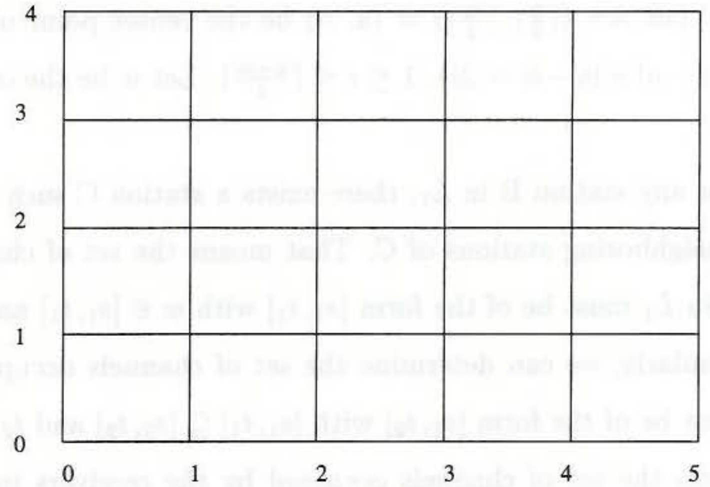


Figure 4.13: An example of meshes: M[6,5]

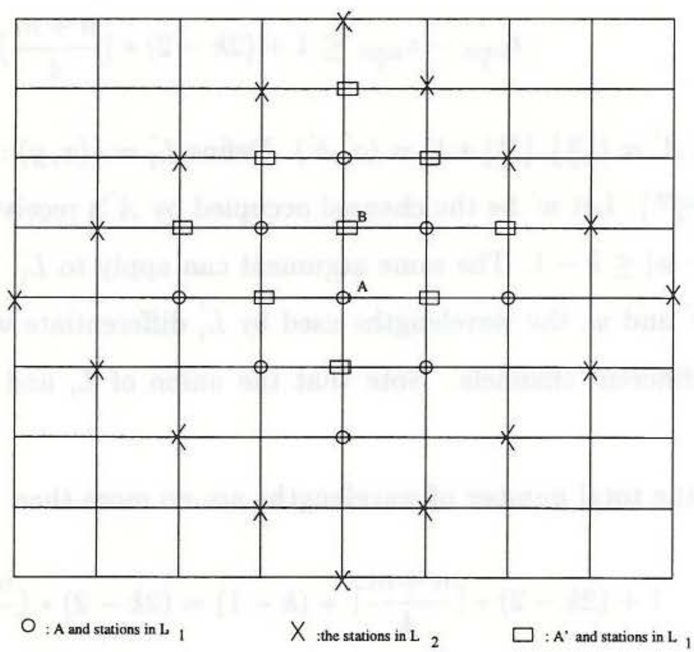


Figure 4.14: A, B, U_i and U_i'

k	O(1)	$O(n^{\frac{1}{2}})$	$O(m^{\frac{1}{2}})$	O(n)	O(m)
maximum channel	$O(n+m)$	$O(n^{\frac{1}{2}}(n+m))$	$O(m^{\frac{1}{2}}(n+m))$	O(n(n+m))	O(m(n+m))

Table 4.2: Upper bound for wavelengths on meshes

The channel assignment algorithm is important for this system. The reason is that this system tries to use as many channels as possible while any two adjacent stations can communicate with each other directly.

We design an algorithm for channel allocation for general k. p is the number of channels available for the system. This algorithm can attain the upper bound for the maximum number of channels used in the system.

Define $a\ (rmod\ p) = a \quad if\ 0 \leq a \leq p-1,\ a - \lfloor \frac{a}{p} \rfloor * p \quad if\ \lfloor \frac{a}{p} \rfloor\ is\ even,\ p - a + \lfloor \frac{a}{p} \rfloor * p \quad if\ \lfloor \frac{a}{p} \rfloor\ is\ odd.$

Algorithm Mesh-Embedding

1. **Initialization**: For any station in the system, pick up its coordinate in the mesh. Divide the mesh into two parts: $X = \{(i,j) \in M[n,m] : i+j\ is\ even\}$ and $Y = \{(i,j) \in M[n,m] : i+j\ is\ odd\}$. Divide X into disjoint subsets: $X_i = \{(a,b) \in X : a+b = i\}$, i is even and $0 \leq i \leq n+m$. Divide Y into disjoint subsets: $Y_i = \{(a,b) \in Y : a+b = i\}$, i is odd and $0 \leq i \leq n+m$.

2. **Channel Assignment**: For each station in X_i and $\frac{i}{2}$ even, assign wavelengths in the range $[\ (\frac{i}{2}-1)*(k-1)(rmod\ p), (\frac{i}{2}-1)*(k-1)+\lfloor\frac{k}{2}\rfloor(rmod\ p)]$ to its receiver and $[\ (\frac{i}{2}-1)*(k-1)-\lfloor\frac{k-1}{2}\rfloor(rmod\ p), \frac{i}{2}-1)*(k-1)+\lfloor\frac{k-1}{2}\rfloor(rmod\ p)\ ]$ as the tuning range of its transmitter.

   For each station in X_i with $\frac{i}{2}$ odd, assign wavelengths in the range $[\ (\frac{i-1}{2}-1)*(k-1)+\lfloor\frac{k-1}{2}\rfloor+1(rmod\ p), \frac{i-1}{2}*(k-1)-1(rmod\ p)]$ and $[\ \frac{i-1}{2}-1)*(k-1)(rmod\ p), \frac{i-1}{2}*(k-1)-1(rmod\ p)]$ as the tuning range of its transmitter.

   For each station in Y_i, assign the wavelengths to its transmitter and receiver in same manner as described above for stations in X_{i-1}.

Note that the maximal channels which can be used in the system embedded with M[n,m] by Mesh-Embedding is $\lceil \frac{n+m}{4} \rceil * (k-1)$, which is in the order of O(k*(n+m)). Therefore we say this algorithm provides an optimal way to use as many wavelengths as possible.

We show how to assign wavelengths to a system with k=3 and 4 in m[7,6]. These values are the ones most employed in current research.

4.6 Hypercubes

Hypercubes are widely used in parallel computing. The same questions raised in the previous section also need to be answered in the systems embedded with hypercubes. A hypercube H[2,n], is a set of nodes $V = \{X = x_0x_1...x_{n-1} : x_i = 0 \text{ } or \text{ } 1\}$, and two nodes $X = x_0x_1...x_{n-1}$ and $Y = y_0y_1...y_{n-1}$ are connected by an edge iff $\sum_{i=0}^{n-1} |x_i - y_i| = 1$. It is easy to see the diameter of H[n] is n. So we have the following:

Lemma 4.6.1 *The maximum delay in a system embedded with H[2,n] is n.*

We want to determine the maximum number of wavelengths which can be used in a system with tunable transmitters of the limited tuning range and the embedded hypercube topology.

Theorem 4.6.2 *The total number of wavelengths which can be used in a system with the embedded Hypercube[2,n] is no more than*

$$2 + (k-1) * n$$

Proof : Let $X = x_0x_1...x_{n-1}$ with $x_i = 0$, $i = 0, 1, ..., n-1$ and w be the wavelength of X's receiver. X has n adjacent stations Y_i, denoted by $U_1 = \{Y_i : 0 \leq i \leq n-1\}$.

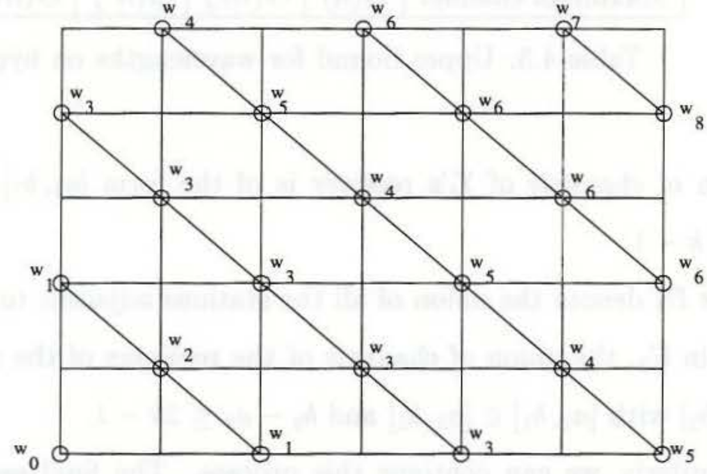


Figure 4.15: An example of Mesh-Embedding for M[7,6] with k=3

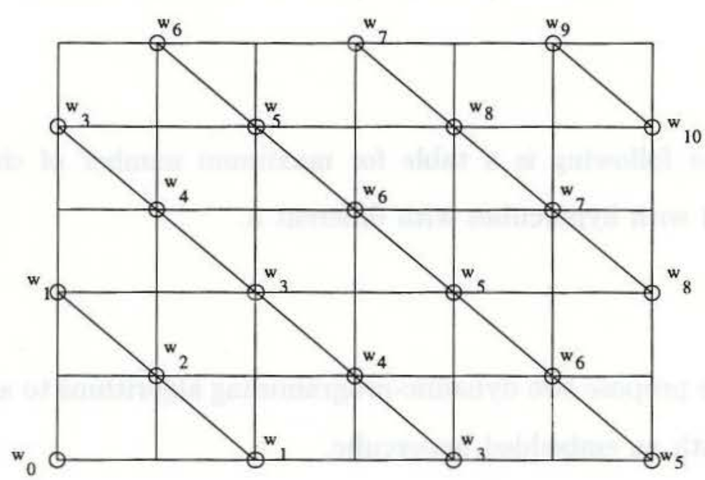


Figure 4.16: An example of Mesh-Embedding for M[7,6] with k=4

k	O(1)	$O(n^{\frac{1}{2}})$	$O(n)$	$O(n^2)$	$O(2^n)$
maximum channel	$O(n)$	$O(n^{\frac{3}{2}})$	$O(n^2)$	$O(n^{\frac{5}{2}})$	$O(2^n)$

Table 4.3: Upper bound for wavelengths on hypercubes

The union of channels of Y_i's receiver is of the form $[a_1, b_1]$ with $w \in [a_1, b_1]$ and $b_1 - a_1 \leq k - 1$.

Let U_2 denote the union of all the stations adjacent to any node in U_1. Since X is also in U_2, the union of channels of the receivers of the stations in U_2 is of the form $[a_2, b_2]$ with $[a_1, b_1] \in [a_2, b_2]$ and $b_2 - a_2 \leq 2k - 1$.

Similarly, we can continue this process. The final set is U_n. The union of channels occupied by the stations which are in U_n is of the form $[a_n, b_n]$, $[a_{n-1}, b_{n-1}] \subseteq [a_n, b_n]$, i.e. $[a_n, b_n]$ contains all the channels which are possibly used in the system. Therefore

$$b_n - a_n \leq 1 + (k-1) + (k-1) + ... + (k-1) = 1 + n * (k-1)$$

∎

The following is a table for maximum number of channels for the system embedded with hypercubes with different k.

We propose two dynamic-programming algorithms to allocate channels for the system with an embedded hypercube.

Algorithm Hypercube-Embedding-One:

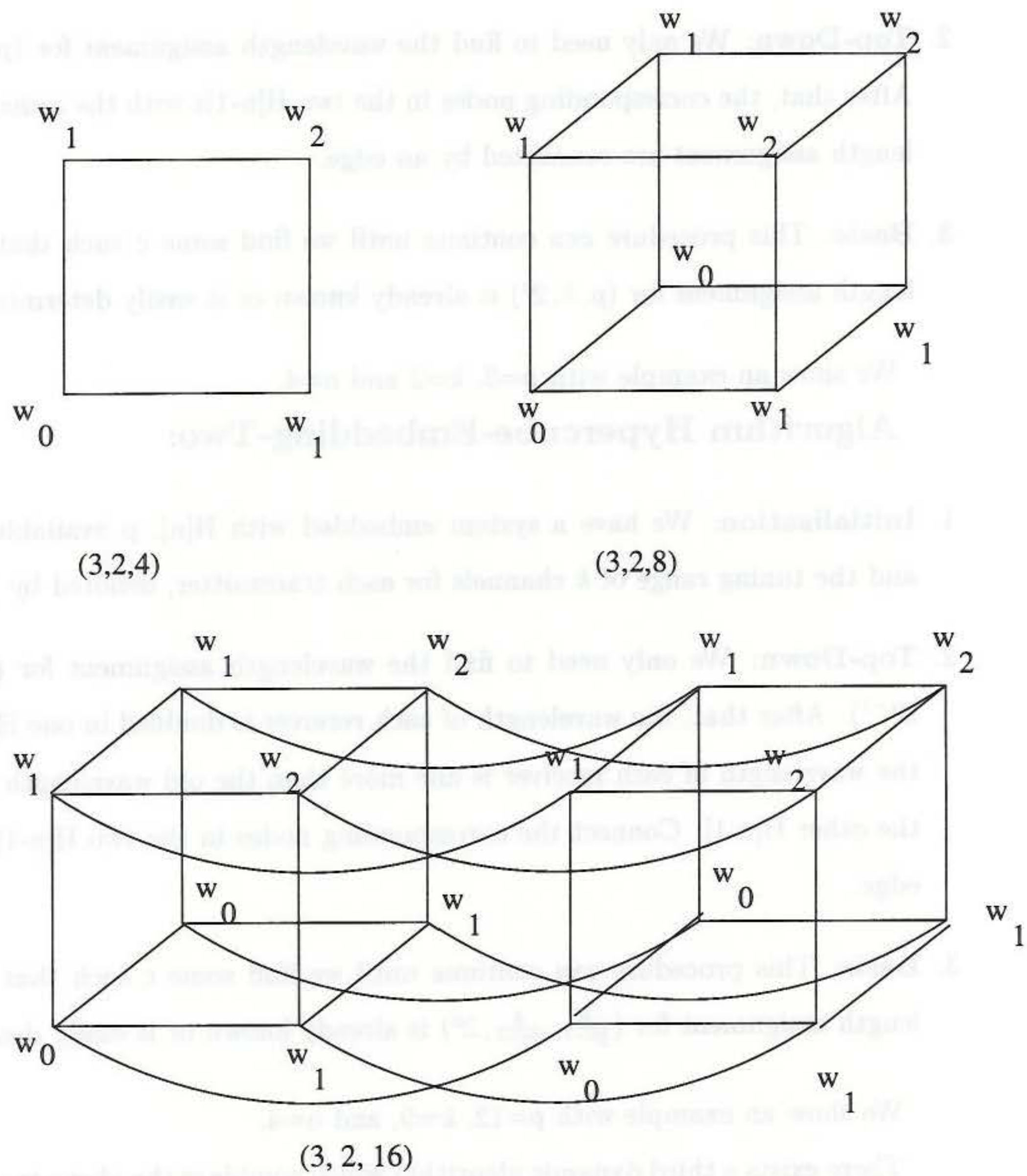


Figure 4.17: Algorithm Hypercube-Embedding-One for (3,2,16)

1. **Initialization**: We have a system embedded with H[n], p available channels and the tuning range of k channels for each transmitter, denoted by (p, k, 2^n)

2. **Top-Down**: We only need to find the wavelength assignment for (p, k, 2^{n-1}). After that, the corresponding nodes in the two H[n-1]s with the same the wavelength assignment are connected by an edge.

3. **Basis**: This procedure can continue until we find some c such that an wavelength assignment for $(p, k, 2^c)$ is already known or is easily determined.

We show an example with p=3, k=2 and n=4.

Algorithm Hypercube-Embedding-Two:

1. **Initialization**: We have a system embedded with H[n], p available channels and the tuning range of k channels for each transmitter, denoted by (p, k, 2^n)

2. **Top-Down**: We only need to find the wavelength assignment for (p/2, k/2, 2^{n-1}). After that, the wavelength of each receiver is doubled in one H[n-1], and the wavelength of each receiver is one more than the old wavelength double in the other H[n-1]. Connect the corresponding nodes in the two H[n-1]s with an edge.

3. **Basis**: This procedure can continue until we find some c such that an wavelength assignment for $(\frac{p}{2^{n-c}}, \frac{k}{2^{n-c}}, 2^c)$ is already known or is easily determined.

We show an example with p=12, k=9, and n=4.

There exists a third dynamic algorithm which combines the above two dynamic algorithms into a new dynamic algorithm for the hypercube embedding.

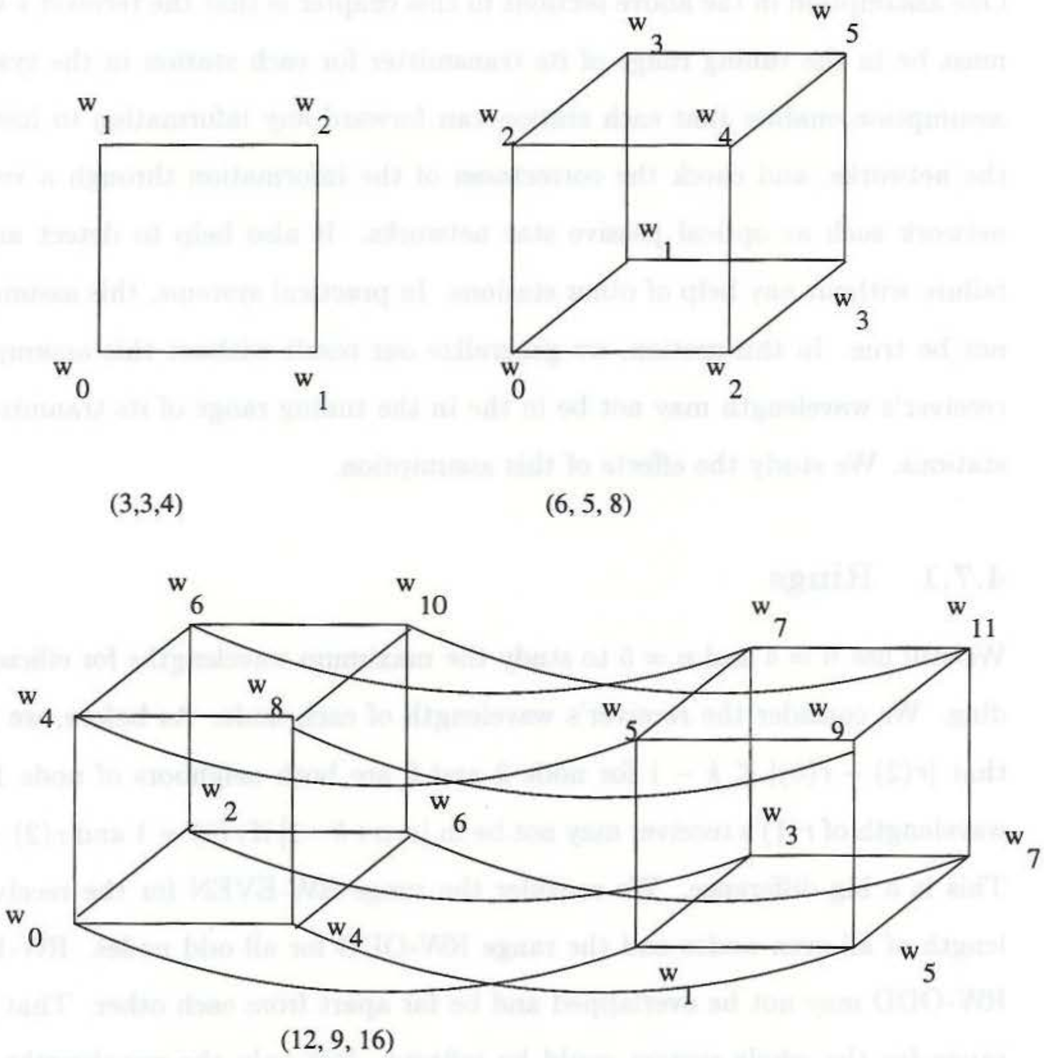


Figure 4.18: Algorithm Hypercube-Embedding-Two for (12,9,16)

4.7 Limited Tuning Range Under Another Assumption

One assumption in the above sections in this chapter is that the receiver's wavelength must be in the tuning range of its transmitter for each station in the system. This assumption enables that each station can forward any information to itself through the networks, and check the correctness of the information through a very reliable network such as optical passive star networks. It also help to detect any receiver failure without any help of other stations. In practical systems, this assumption may not be true. In this section, we generalize our result without this assumption. The receiver's wavelength may not be in the in the tuning range of its tranmitter for any stations. We study the effects of this assumption.

4.7.1 Rings

We still use $n = 4$ and $n = 5$ to study the maximum wavelengths for efficient embedding. We consider the receiver's wavelength of each node. As before, we must have that $|r(2) - r(0)| \leq k - 1$ for node 2 and 0 are both neighbors of node 1. But the wavelength of $r(1)$'s receiver may not be in $[a, a+k-1]$ if $r(0) = 1$ and $r(2) = a+k-1$. This is a big difference. We consider the range RW-EVEN for the receiver's wavelength of all even nodes and the range RW-ODD for all odd nodes. RW-EVEN and RW-ODD may not be overlapped and be far apart from each other. That makes the range for the whole system could be infinite. But only the wavelengths which are occupied by any receiver will be really utilized for communication. In the rest of this chapter, we emphasize the total number of wavelengths *which can be utilized.*

Lemma 4.7.1 *If $n > 2$, the total number of wavelengths for $L[n]$ which can be used is no more than $n(k-1) + 2$.*

Figure 4.20 shows how to assign wavelengths for all different cases for receiver wavelength assignment. If n is odd, one cycle is formed to contained all the nodes in the system.

Lemma 4.7.2 *If $n = 4m$ or $4m + 2$ with $m > 0$, the total number of receiver's wavelengths of all nodes is no more than $2m(k-1)+2$.*

If $n = 4m + 1$ ith $m > 0$, the total number of receiver's wavelengths of all nodes is no more than $2m(k-1)+1$.

If $n = 4m + 3$ ith $m > 0$, the total number of receiver's wavelengths of all nodes is no more than $(2m+1)(k-1)+1$.

4.7.2 Meshes

The upper bound for the total number of wavelengths for systems embedded with meshes are similar to Theorem 4.5.2. The only difference is that the upper bound is doubled than that of Theorem 4.5.2. $L_i \subseteq L_{i+1}$ is not true any more without the previous assumption. There are two sets of receiver's wavelengths. One is the unions of all receiver's wavelengths in $\bigcup_{i:odd} L_i$, the other is the unions of all receiver's wavelengths in $\bigcup_{i:even} L_i$. This makes the range for receiver's wavelengths double.

Theorem 4.7.3 *The total number of wavelengths which can be used in a system with the embedded M[n,m] is no more than $2 * [(2k-2) * \lceil \frac{n+m}{4} \rceil + k]$.*

Algorithm Mesh-Embedding shows how to assign the receiver's wavelengths to $X_i = \{(a,b) \in X : a+b = i\}$ with i is odd. For $Y_i = \{(a,b) \in X : a+b = i\}$ i is even, we can copy the receiver's wavelengths from neighboring stations in X_i.

We still apply Algorithm Mesh-Embedding to X_i. For Y_i, we apply the methods for X_i to Y_i also with the increasement on wavelengths for Y_i.

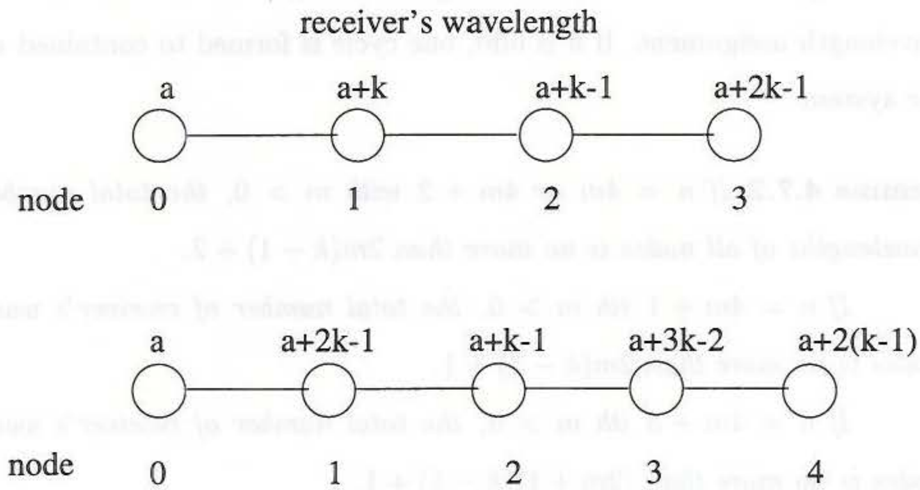


Figure 4.19: An optimal assignment for L[4] and L[5]

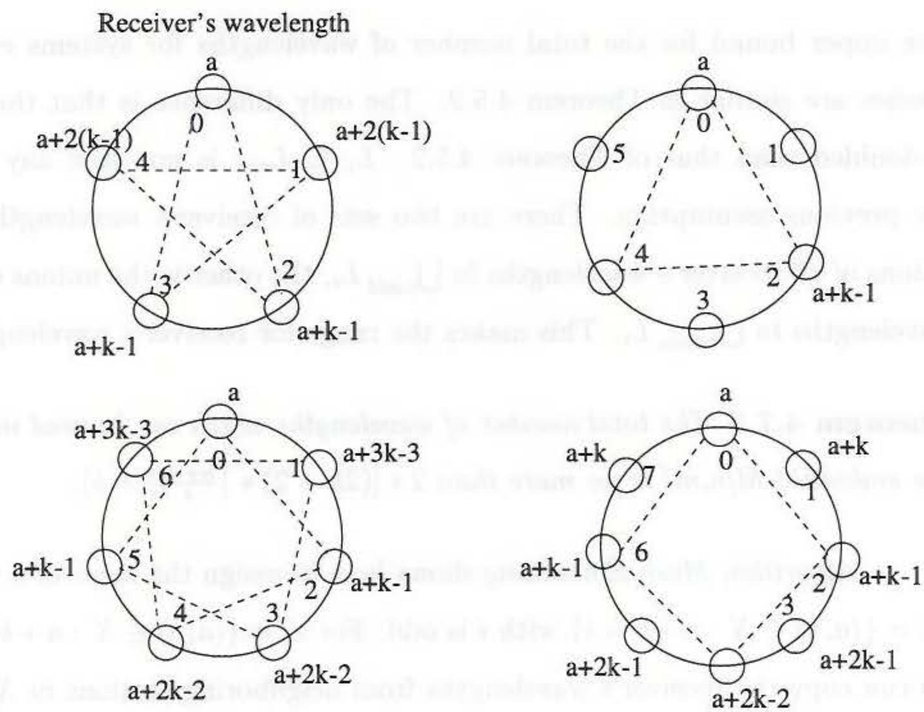


Figure 4.20: An example for four cases for rings

Figure 4.22 shows how to assign the receiver's wavelength to M[7,6] with $k = 8$.

4.7.3 Hypercubes

We show in Theorem ?? the maximum number of wavelengths which can be used in a system with tunable transmitters of the limited tuning range and the embedded hypercube topology.

Theorem 4.7.4 *The total number of wavelengths which can be used be a system with the embedded Hypercube[2,n] is no more than*

$$2 * \lceil \frac{n-1}{2} \rceil (2k-2) - k + 3$$

Proof : Let $X = x_0 x_1 ... x_{n-1}$ with $x_i = 0$, $i = 0, 1, ..., n-1$ and w be the wavelength of X's receiver. X has n adjacent stations Y_i, denoted by $U_1 = \{Y_i : 0 \leq i \leq n-1\}$. The union of channels of Y_i's receiver is of the form $[a_1, b_1]$ with $w \in [a_1, b_1]$ and $b_1 - a_1 \leq k - 1$.

Let U_2 denote the union of all the stations adjacent to any node in U_1. Since X is also in U_2 and X is a common neighbor for all nodes in U_1, the union of channels of the receivers of the stations in U_2 is of the form $[a_2, b_2]$ and $b_2 - a_2 \leq 2k - 1$.

Let U_3 denote the union of all the stations adjacent to any node in U_2. Let U_4 denote the union of all the stations adjacent to any node in U_3. Similarly, the union of channels of the receivers of the stations in U_4 is of the form $[a_4, b_4]$ with $[a_2, b_2] \in [a_4, b_4]$ and $b_4 - a_4 \leq 4k - 3$.

Consider $\bigcup_{i:even} U_i$. The total number of receiver's wavelengths is no more than $\lceil \frac{n-1}{2} \rceil (2k-2) + 1$.

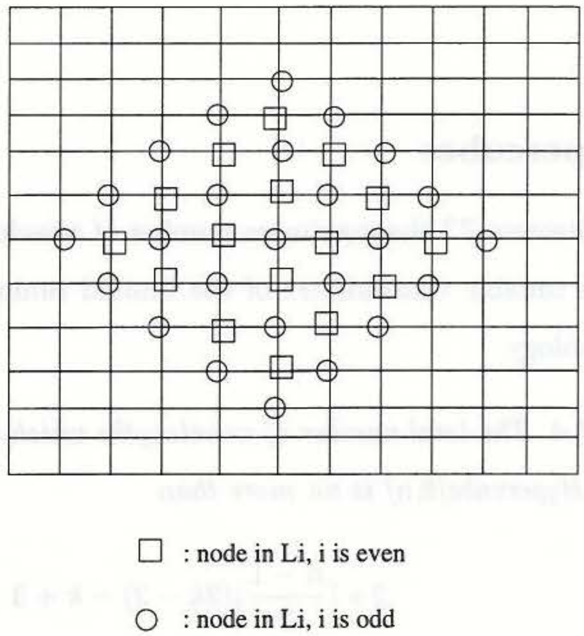


Figure 4.21: Two different sets of receiver's wavelengths

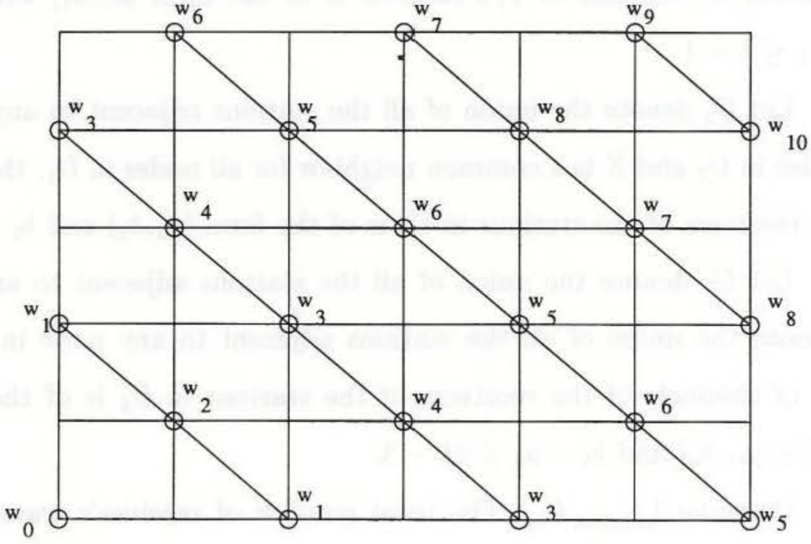


Figure 4.22: An example of Mesh-Embedding for M[7,6] with k=4

Consider $\bigcup_{i:odd} U_i$. The total number of receiver's wavelengths is no more than $max(\lceil \frac{n-3}{2} \rceil, 0) * (2k-2) + k$.

Thus the total number of wavelengths which can be used are no more than

$$\lceil \frac{n-1}{2} \rceil (2k-2) + 1 + max(\lceil \frac{n-3}{2} \rceil, 0) * (2k-2) + k = 2 * \lceil \frac{n-1}{2} \rceil (2k-2) - k + 3$$

∎

4.8 Conclusion

We specify the problems introduced by the limited tuning range of transmitters for topological embedding, which was ignored in previous studies. After formulating the problems, we study the optical passive star network embedded with the complete graph, meshes and hypercubes. These three topologies are among the most widely used structures in both network communication and parallel computing. For the system embedded with the complete graph, the bounds for the maximum delay are studied and the embedding algorithm is designed which is optimal in terms of the maximum delay. The bottleneck phenomenon is analyzed in such a environment. One-to-all broadcasting can be implemented for the common parallel communications by using our embedding algorithm. For both meshes and hypercubes, the relation between the structure of the topologies and the maximum number of wavelengths which can be used are analyzed. The algorithm for the topological embedding is designed to show how to maximize the use of available channels. Examples are given to show the efficiency of our algorithms for the topological embedding.

The problems we proposed in Section 1.4 also need to be answered for the systems embedded with other topologies. The methods we used for these three structures can also be helpful in designing new embedding algorithms and analyzing the rela tions among different system parameters. Our future work will be based on othe r topologies, and the impact of embedding multiple topologies in the same system .

Bibliography

[1] ANSI/IEEE Standard 802.5-1985, Token Ring Access Method and Physical Layer Specifications, 1985

[2] Anthony S. Acampora, An Introduction to Broadband Networks, *Plenum Press*, New York, 1994.

[3] Nima Ahmadvand and Terence Todd, Dual-Hop LANs Using Station Wavelength Routing, *IEEE ICCCN'95*, September, 1995.

[4] J. D. Attaway and J. Tan, HyperFast: Hypercube Time Slot Allocation in a TWDM Network, preprint, 1994.

[5] Subrata Banerjee and Biswanath Mukherjee, Fairnet: A WDM-based Multiple Channel Lightwave Network with Adaptive and Fair Scheduling Policy, *Journal of Lightwave Technology*, pp. 1104-1112, 5/6, 1993.

[6] J. A. Bannister and L. Fratta and M. Gerla, Topological design of the wavelength-division optical network, *IEEE Infocom'90*, pp. 1005-1013, 1990.

[7] S. Bhattacharya, D. H. Du, and A. Pavan, A Network Architecture for Distributed High Performance Heterogeneous Computing, *Proceedings of the Heterogeneous Computing Workshop*, pp. 110–115, 1994.

[8] G. Birkhoff, Tres observaciones sobre el algebra lineal, *Univ. Nac. Tucumán, Rev.*, Ser. A, 5, pp. 147–151, 1946.

[9] Michael S. Borella and Biswanath Mukherjee, Efficient Scheduling of Nonuniform Packet Traffic in a WDM/TDM Local Lightwave Network with Arbitrary Transceiver Tuning Latencies, *IEEE Infocom'95*, pp. 129-137, 1995.

[10] C. A. Brackett, Dense Wavelength Division Multiplexing Networks: Principles and Applications, *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 6, pp. 948–964, 1990.

[11] F. Cao and A. Borchers, "Optimal Transmission Schedules for Embeddings of the De Bruijn Graphs in an Optical Passive Star Network", *Proceedings of IEEE 5th International Conference on Computer Communication and Networks (IC-CCN'96), October 1996, Maryland, USA.*

[12] F. Cao, D. H. C. Du and A. Pavan, Design of WDM Optical Passive Star Networks with Tunable Transceivers of Limited Tuning Range, *University of Minnesota Technical Report TR 96-071*, submitted to IEEE Transactions on Computers.

[13] F. Cao, D.Z. Du, D.F. Hsu and P. Wan, Fault-tolerant routing in Butterfly Networks, *DIMACS workshop on network design, April, Princeton, 1997. Technical report TR 95-073, Department of Computer Science, University of Minnesota, 1995*

[14] F. Cao, D.Z. Du and D.F Hsu, On the fault-tolerant diameters and containers of large bipartite digraphs, *International Symposium On Combinatorial Algorithm, Tianjin, China, June, 1996.*

[15] Ming Chen and Nicolas D. Georganas, Multiconnection Over Multichannels, *IEEE Infocom'95*, pp. 1037-1043, 1995

[16] D. Chiarulli, S. Levitan, R. Melhem, J. Teza and G. Gravenstreter, Multiprocessor interconnection networks using partitioned optical passive star topologies and distributed control, *Proceedings of the 1st International Workshop on Massively Parallel Processing Using Interconnections*, 70-80, IEEE, April 1994

[17] K. W. Cheung, Scalable, Fault Tolerant 1-Hop Wavelength Routing, *GlobeCom*, pp. 1240-1244, 1991.

[18] R. Chipalkatti and Z. Zhang and A. S. Acampora, High Speed Communication Protocols for optical star coupler ing WDM, *IEEE Infocom*, pp. 2124-2133, 1992.

[19] S.D. Cusworth and J.M. Senior and A. Ryley, Wavelength Division Multiple Access on a High-speed Optical Fibre LAN, *Computer Networks and ISDN Systems*, pp. 323-333, 1989/90.

[20] Patrick W. Dowd, Random Access Protocols for High-Speed Interprocessor Communication Based On an Optical Passive Star Topology, *Journal Of Lightwave Technology*, pp. 799-808, vol. 9, 1991.

[21] P. W. Dowd, Wavelength Division Multiple Access Channel Hypercube Processor Interconnection, *IEEE Transactions on Computers,*, vol. 41, no. 10, pp. 1223–1241, 1992.

[22] D. H. C. Du and R. J. Vetter, *Advanced Computer Networks*, preprint.

[23] C. Ersoy and S. P. Pamwar, Topological design of interconnected LAN-MAN networks, *IEEE Infocom*, pp. 2260-2269, 1992.

[24] A. Ganz, B. Li and L. Zenou, Reconfigurability of multi-star based lightwave LANs, *GlobeCom'91*, volume 3, 1906-1910, 1992

[25] Mario Gerla and B. Kannan and P. Palnati, Protocols for an Optical Star Interconnect for High Speed Mesh Networks, *IEEE Infocom'95*, pp. 146-153, 1995

[26] G. R. Green, *Fiber Optic Networks*, Englewood Cliffs, New Jersey, Prentice Hall, 1993.

[27] Optical Communications Systems, *Prentice-Hall*, 2nd Edition, 1993.

[28] J. E. Hopcroft and R. M. Karp, "An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs," *SIAM Journal on Computing*, 2 (4), pp. 225-231, 1973.

[29] Thomas Inukai, An Efficient SS/TDMA Time Slot Assignment Algorithm, *IEEE Transactions On Communications*, pp. 1449-1455, vol. COM-27, 1979

[30] A. Ganz and Y. Gao, "Time-Wavelength Assignment Algorithms for High Performance WDM Star Based Systems," *IEEE Transactions on Communi cations*, vol. 42, no. 2,3,4, pp. 1827–1836, 1994.

[31] G. Gravenstreter, R. Melhem, D. Chiarulli, S. Levitan and J. Teza, The partitioned optical passive stars(POPS) topology, *Proceedings of the Ninth International Parallel Processing Symposium*, April 1995

[32] G. Gravenstreter and R.G. Melhem, Realizing Common Communication Patterns in Partitioned Optical Passive Stars(POPS) Networks, *Proceedings of the 2nd Int. Conf. on Massively Parallel Processing Using Optical Interconnections*, San Antonio, TX, 1995

[33] IEEE Std 902.3-1985, Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, *IEEE ISDN 0-471-892749-5*, 1985.

[34] M. Janoska and T.D. Todd, A Simplified Optical Star Network Using Distributed Channel Controllers, *Submitted for publication*, 1993

[35] B. Kannan and Shivi Fotedar and Maria Gerla, A Protocol for WDM Star Coupler Networks, *IEEE Infocom'94*, pp. 1536-1543, June, 1994.

[36] Mohsen Kavehrad and Ganti Sudhakar and Nicolas Georganas, Slotted Aloha and Reservation Aloha Protocols For Very High-Speed Optical Fiber Local Area Networks Using Passive Star Topology, *Journal Of Lightwave Technology*, pp. 1411-1422, vol. 8, 1991.

[37] Milan Kovacevic and Mario Gerla, On the performance of shared-channel multihop lightwave networks, *IEEE Infocom'94*, pp. 5440-5510, 1994

[38] J. P. Labourdette and A. S. Acampora, Wavelength agility in multihop lightwave networks, *IEEE Infocom'90*, pp. 1022-1029, 1990.

[39] Sang-Kyu Lee, A. Duksu Oh, and Hyeong-Ah Choi, Transmission Schedules for Hypercube Interconnection in TWDM Optical Passive Star Networks, Department of Electrical Engineering and Computer Science, George Washington University Technical Report GWU-IIST-95-07, 1995.

[40] Sang-Kyu Lee, A. Duksu Oh, Hongsik Choi, and Hyeong-Ah Choi, Optimal Transmission Schedules in TWDM Optical Passive Star Networks, Department of Electrical Engineering and Computer Science, George Washington University Technical Report GWU-IIST-95-03.

[41] J. H. van Lint, and R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 1992.

[42] G. Liu, K. Y. Lee, and H. Jordan, TDM Hypercube and TWDM Mesh Optical Interconnections, *Proceedings of IEEE GLOBECOM*, pp. 1953–1957, 1994.

[43] M. Marsan and A. Bianco and E. Leonardi and M. Meo and F. Neri, On the Capacity of MAC Protocols for All-Optical WDM Multi- Rings with Tunable Transmitters and Fixed Receivers, *submitted to IEEE Infocom'96*, 1996.

[44] N. F. Maxemchuk, Regular Mesh Topologies in Local and Metropolitan Area Networks, *AT&T Technical Journal*, pp. 1659-1683, vol. 64, 1985.

[45] R. Melhem and G. Hwang, Embedding rectangular grids into square grids with dilation two, *IEEE Transactions on Computers*, 1446-1455, December 1990

[46] B. Mukherjee, WDM-Based local lightwave networks, Part I: Single-Hop systems, *IEEE Networks*, 12-27, May 1992

[47] A. Pavan, S. Bhattacharya and D. H. C. Du, Reverse Channel Augmented Multihope Lightwave Networks, Proceedings of the IEEE Infocom, San Fransico, March, 1993.

[48] A. Pavan, P. J. Wan, S. R. Tong and D. H. C. Du, A New Multihop Lightwave Network Based on the Generalized De-Bruijn Graph, *the Proceeding of the IEEE Local Networks Conference*, Minneapolis, October 1996.

[49] R. Ramaswami, Multiwavelength lightwave networks for computer communication, *IEEE Communications Magazine*, pp. 78-88, vol. 31, 1993.

[50] G. N. Rouskas and M. H. Ammar, Analysis and Optimization of transmission schedules for single-hop WDM networks, *IEEE Infocom'93*, pp. 1342-1349, 1993.

[51] A. Ryley and S.D. Cusworth and J.M. Senior, Piggybacked Token Passing Access Protocol for Multichannel Optical Fibre LANs, *Comput. Comm.*, pp. 213-222, vol. 12, 1989.

[52] J. Sharony and T. E. Stern and K. W. Cheung, Extension of Classical Rearrangable and Non-blocking Networks to the Wavelength Dimension, *IEEE GlobeCom*, pp. 1901-1905, 1992.

[53] K. N. Sivarajan and R. Ramaswami, "Lightwave Networks Based on de Bruijn Graphs," *IEEE Transactions on Networking*, vol. 2, no. 1, pp. 70-79, 1994.

[54] G. N. M. Sudhakar and N. D. Georgana and M. Kavehrad, Slotted Aloha and Reservation Alha Protocols for Very High Speed Optical Fiber Local Area Networks Using Passive Star Topology, *Journal of Lightwave Tech*, vol. 9, 1991.

[55] Terence D. Todd and Adrian M. Grah and Oliver Barkovic, Optical Local Area Networks (LANS) Using Wavelength-Selective Couplers, *IEEE Infocom'95*, pp. 916-923, 1995.

[56] S. R. Tong, "Efficient Designs for High Speed Network Architectures", Ph.D. Thesis, University of Minnesota, February, 1994

[57] K. A. Williams and D. H. C. Du, Efficient Embedding of a Hypercube in an Irregular WDM Network, Technical Report, Department of Computer Science, University of Minnesota, 1991.

[58] K. A. Williams and D. H. C. Du, Time and Wavelength Division Multiplexed Architectures for Optical Passive Star Networks, Technical Report, Department of Computer Science, University of Minnesota, 1991.