Communications and Networking

Academic and research staff
Professor Vincent W. S. Chan

Visiting scientists and research affiliates
Professor Ada Poon (UIUC), Professor Guevara Noubir (Northeastern University), Professor Ashwin Gumaste (IIT, Mumbai), Professor Perry Shum (NTU Singapore), John Chapin (VANU Inc.), Steven Finn, Alan Kirby (CISCO), Terrence McGarty, Eric Swanson

Graduate students
Patrick Choi, Lillian Dai, Anurupa Ganguly, Bishwaroop Ganguly, James Glettler, Elizabeth Godoy, Kyle Guan, Etty Lee, Justin Pun, Andrew Puryear, Guy Weichenberg, Yonggang Wen

Undergraduate students
Anurupa Ganguly, Justin Pun

Technical and support staff
Susanne Patterson, Pardis Parsa

Overview of the research in the Communications and Networking Group
Research in the group includes topics on networks and communication systems. The work extends to applications in satellite, wireless and optical communication, and data networks. The objective is to develop the scientific base needed to design data communication networks that are efficient, robust and architecturally clean. Wide-area and local-area networks, high-speed networks, and point-to-point and broadcast communication channels are of concern. Topics of current interest include network architectures at all network layers; power control; multiple antenna techniques; media access control protocols; routing in optical, wireless and satellite networks; quality of service control; failure recovery; topological design; and the use of efficient resource allocation for network connectivity and QoS. An important new research frontier that will call for new modeling, analysis and architecture optimization is the next frontier of networks over fiber systems with agile and economical service delivery and fast, dynamically changing communication media, such as mobile wireless communication in the presence of fast fading and changing connectivities, microwave satellite communication and free space optical communication over rapidly changing atmospheric phenomena and dealing with unscheduled, bursty, large-granularity users at the edge of networks. The problems in this direction are important, rich and very challenging. The research usually requires multi-disciplinary techniques and tools to tackle.

A major theme of the group is the pursuit of research that cuts across multiple disciplines. We believe the forefront of this area is at the interface of traditional quantitative communication and network research and adjacent areas such as computer science, physics, device, system design and demonstration, operations research, economics and management. Examples are optical, wireless and satellite networks. Thematic research provides a venue to apply the fruits of basic research to real world problems, act as a reality check to abstractions and models made in fundamental theoretical research and as a source of new and interesting problems. A significant component of our research is the creation of new architectures guided by the understanding of technology and fundamental system limits and the realization and validation of these architectures with hardware and algorithms and system demonstrations. As such, we work closely with industry to extend our research reach and impact beyond the academic boundary.

1 Proactive Wireless Networks

Sponsors
DARPA - HR0011-05-1-0048, HR0011-06-1-0011

Project Staff
Professor Vincent W. S. Chan, Lillian Dai, Elizabeth Godoy, Anurupa Ganguly, Professor Ada Poon (UIUC), Professor Guevara Noubir (Northeastern University), John Chapin (VANU Inc.), Alan Kirby (CISCO), Terrence McGarty

The objective of this program is to develop a Proactive Mobile Wireless Network paradigm for next generation *infrastructureless wireless networking to guarantee critical services to users with time deadline constraints.* In contrast to existing mobile ad hoc wireless networks where frequent network disconnections and greatly degraded services may occur due to fluid changes in the location and composition of wireless devices in combat theatres, a proactive mobile wireless network actively maintains network connections to ensure continuous communication and timely delivery of mission-critical information. Such capabilities will be necessary and crucial for networks that operate under extreme operating scenarios, such as small unit operations, search and rescue, and urban warfare. In the initial phase of our study, we completed the following:

a. Feasibility study of the proposed network architecture.
b. Quantification of network gains.
c. Development of two network topology management algorithms.
d. Impact of signal processing energy and large bandwidth on infrastructureless wireless network routing and scalability.

The proposed Proactive Mobile Wireless Network architecture consists of the following primary components:

a. User location, movement trajectory, and communication channel state estimation for network disconnection prediction.

b. Adaptive deployment and movement command of helper nodes. These are additional wirelessly-enabled devices deployed with the single purpose of network connection maintenance. Examples of helper nodes include navigational robots or vehicles that plan their trajectories to connect user nodes where connections are most needed, or balloons and other aerial vehicles carrying communication relays that hover and drift above a region to provide wider communication coverage.

1.1 Feasibility study of the proposed network architecture

Since the proposed architecture spans several technology areas, we have examined all the building blocks to identify potential technology bottlenecks and enablers. It has been found that, in most cases, current technologies are sufficient for small network operations on the order of tens of nodes. As robots and small autonomous aerial vehicles become more practical, versatile, and robust, these devices can help to connect larger networks. We identify and propose technology enablers in two areas where existing technologies are deficient:

a. Node localization in indoor environments – Signal obscuration and multipath present significant challenges for indoor localization using narrow-band systems. Based on signal measurements from an ultrawide-band radio testbed in an indoor environment, we have compared location estimation error based on time-of-arrival (ToA) and received-signal-strength (RSS) for a variety of wireless beacon placement configurations. The results indicate that sub-meter location estimation accuracy is possible using wide-band ToA ranging. An important factor in this case is that the bandwidth is wide enough so that the multi-paths can be resolved.

b. Channel prediction – In the proposed architecture, it is necessary to estimate future link channel states even if nodes are mobile. Given node locations and predicted trajectories, future channel states may be estimated with the aid of physical and signal attenuation maps of the operating region. Physical maps of outdoor and indoor environments are typically available through overhead intelligence and floor plans. Using a neural network predictor, we have shown that signal attenuation maps can be constructed off-line using prior signal measurement data and quickly updated on-line when more channel state measurements are available.

1.2 Quantification of network gains

The key architectural questions we want to address are: what helper node capabilities are needed in terms of mobility and communication and whether the increased complexity of the proposed architecture is justified. Using a simple random network model, we have quantified the probability of connection, and throughput, delay and energy gains for proactive mobile wireless networks under different helper node deployment schemes (Figures 1.1, 1.2). It is found that for some network scenarios, even a few strategically placed mobile helper nodes with the same communication capabilities as user nodes can significantly improve network connectivity. For networks with sparsely distributed user nodes, aerial helper nodes are better than terrestrial ones and for networks with densely distributed user nodes, helper node mobility does not have significant gains over fixed helper nodes. For networks with moderately distributed user nodes, orders of magnitude savings in the number of helper nodes can be achieved by using mobile helper nodes compared to randomly deploying fixed helper nodes. In all cases, having more helper nodes will further improve throughput, delay, and energy with a more favorable throughput-delay tradeoff compared to networks with user nodes alone. In addition, strategically placing helper nodes can attain the minimum network energy consumption achievable for wireless networks.

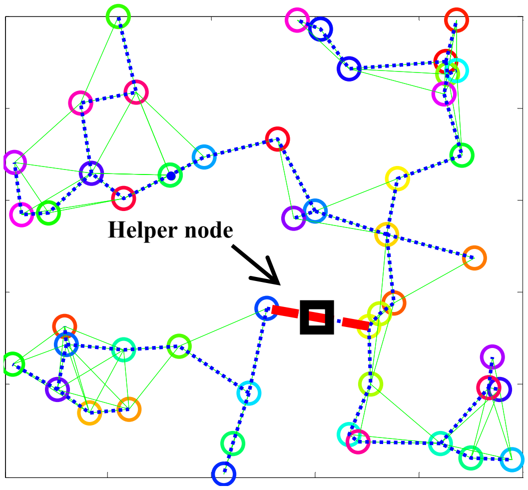


Figure 1.1. Simulation of a 50 node proactive mobile wireless network with one helper node whose trajectory is controlled to provide network connectivity support [1].

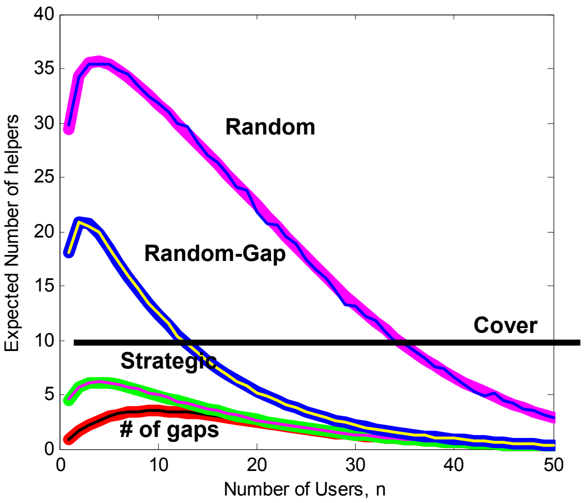


Figure 1.2. Expected number of helper nodes needed under different helper node deployment scenarios for a line network with randomly located user nodes [2].

1.3 Development of two network topology management algorithms

Network topology management via helper node trajectory control is a key aspect of the proactive mobile wireless network architecture. In general, there is a tradeoff between the number of helper nodes needed and algorithmic complexity and with the amount of network state information needed. For networks operating in a planar region, determining the minimum number of helper nodes needed and their locations at any instance in time is an NP-complete problem. We have developed two polynomial-time approximation algorithms based on minimum spanning trees with channel link states as weights on the edges of the trees. These algorithms range from centralized with full network state information to partially distributed with local link monitoring and performance within a small numerical multiplier of the optimum.

1.4 Impact of signal processing energy and large bandwidth on infrastructureless wireless network routing and scalability

Throughput scaling and optimal hop-distance of interference-limited wireless networks have been well characterized in literature. For some emerging wireless networks, throughput may be more limited by battery energy rather than by interference. In characterizing throughput scaling and optimal hop-distance of such power-limited networks, prior work have invoked a zero signal processing energy assumption, which led to the belief that “whispering to the nearest neighbor” (WtNN, with the average number of hops per source destination pair increasing with increasing node density) achieves the optimal throughput scaling, which increases with increasing node density. We show that this believe must be modified for power-limited networks when signal processing energy is not an insignificant factor. In fact, for a power-limited network with nodes uniformly randomly distributed in a bounded region, taking $\Theta(1)$ (i.e. does not increase or decrease with increasing node density) number of hops is throughput, energy, and delay optimal, achieving $\Theta(1)$ pairwise throughput, energy per bit, and packet delay under uniform traffic, whereas WtNN is strictly suboptimal, achieving a pairwise throughput of $o\left(\sqrt{\ln N/N}\right)$, which decreases with increasing node density. In addition, we show that a constant characteristic hop-distance of d_{char} simultaneously achieves the pairwise throughput scaling and minimum network energy consumption for random networks, [3,4] (Figure 1.3).

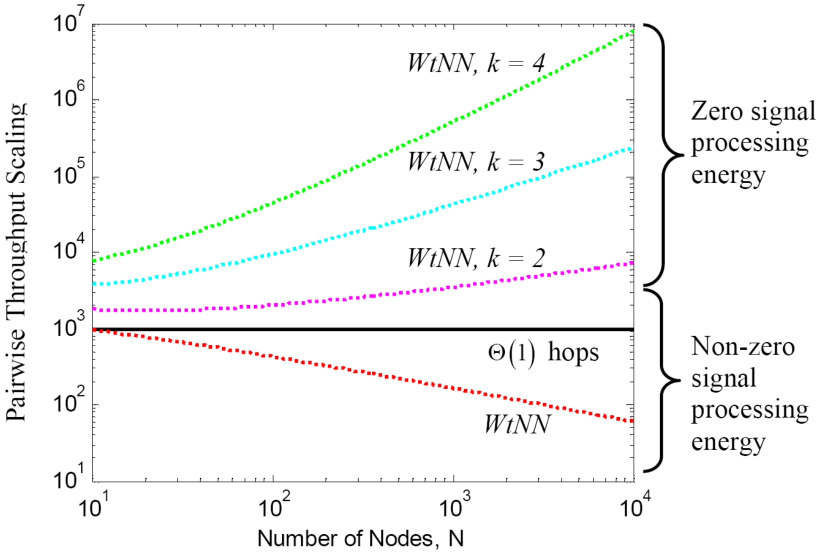


Figure 1.3. Pairwise throughput Scaling and Bounds for Variable Transmission Rate Systems under Zero and Non-zero Signal processing Energy Assumptions (dashed lines = bounds; solid line = exact order)

References

[1] L. Dai and V. Chan, "Helper node trajectory control for connection assurance in proactive mobile wireless networks," *Proceedings of the 16th International Conference on Computer Communications and Networks (ICCCN)*, pp. 882 – 887, August 2007.

[2] L. Dai and V.W.S. Chan, "Proactive Topology Reinforcement of Wireless Networks," *Proceedings of IEEE Wireless Communications & Networking Conference (WCNC 2007)*, pp. 3786 – 3791, March 2007.

[3] V.W.S. Chan, “Some Research Problems for Future Integrated Terrestrial and Satellite Networks,” *Proceedings of the 2007 Military Communications Conference (MILCOM 2007)*, Orlando, FL, Oct. 2007.

[4] L. Dai and V.W.S. Chan, "Throughput of power-limited wireless networks with processing energy considerations," *Proceedings of IEEE Global Telecommunications Conference (Globecom 2006)*, San Francisco, Nov. 2006.

2 Architecture & Applications for Space Information Systems

Sponsor
US Government

Project Staff
Professor Vincent W. S. Chan, Professor Muriel Medard, Professor Robert Gallager, Yongang Wen, Guy Weichenberg, Bishwaroop Ganguly, Lillian Dai, Elizabeth Godoy, John Chapin (VANU Inc.), Steven Finn, Alan Kirby (CISCO), Terrence McGarty, Eric Swanson

This program study advanced satellite network and space information system architectures, joint source and network coding, and transmission scheduling with QOS requirements.

2.1 Satellite network architectures

The satellite network of the future will have the following properties:

a. Streams as well as bursty transactions.
b. Predictable and unpredictable traffic changes with different quality of service requirements.
c. Time varying channel capacities and qualities due to weather and dynamic power and beam allocations.
d. Contending user applications that require dynamic resource allocation and reconfiguration with some rejection.
e. Relatively cheaper space backbone Vs. up and down links.
f. Possibility of shared processing in space, as well as high capacity memories.

In the light of these unusual properties, this research program studied the following network architecture design problems:

a. Dynamic satellite connection topologies and node switching architecture, including reconfiguration of lasercom backbones connections subject to shifting traffic demands and up-and-down link adaptations; node switching architectures to deal with streams as well as bursty transactions.

b. Routing, flow, and congestion control of stream and bursty traffic\Internetworking architecture with terrestrial networks and airborne platforms from Layer 1 to Layer 4 with the possibility of including the Application Layer, with particular attention given to the essential differences between space and other terrestrial and airborne networks, such as Doppler, propagation delays, and channel conditions.

c. Network management and control as a distinct architectural focus that addresses the space environment and end-to-end networking across disparate subnets.

d. Architectural constructs that lend themselves to evolution over time as technology and needs evolve. Especially notable is the revolution in relative capabilities for processing, storage and data transfer as they affect architectural considerations.

If the backbone transport has enough data rates or in the case of analog links enough fidelity in the transport of amplitude and phase, then there is the possibility of long baseline beam forming and interferometry. Figures 2.1 and 2.2 show how conceptually this can be done in a distributed space communication system with space-borne processing at a processing satellite node and a terrestrial wireless network where signals from ‘pico-cells’ are back-hauled to a central office processor for beam forming, nulling and MIMO, multiple-input-multiple-output, communication.

Long base-line interferometry increases receiver dynamic range and sensitivity and in the case of space-borne applications, putting processing in space substantially reduces the downlink capacity requirement. This is a major benefit for space systems where down link resources are very expensive.

For terrestrial wireless applications, dynamic range extension is critical for the support of mobile users that pass through deep multi-path fades. The possibility of large scale frequency re-use is also a huge economic incentive to wireless operators.

This technique is not currently available because it needs either very high data rate digital or very high fidelity analog in the backbone. With fiber networks and the advent of optical cross-links in space, this will be possible in the future.

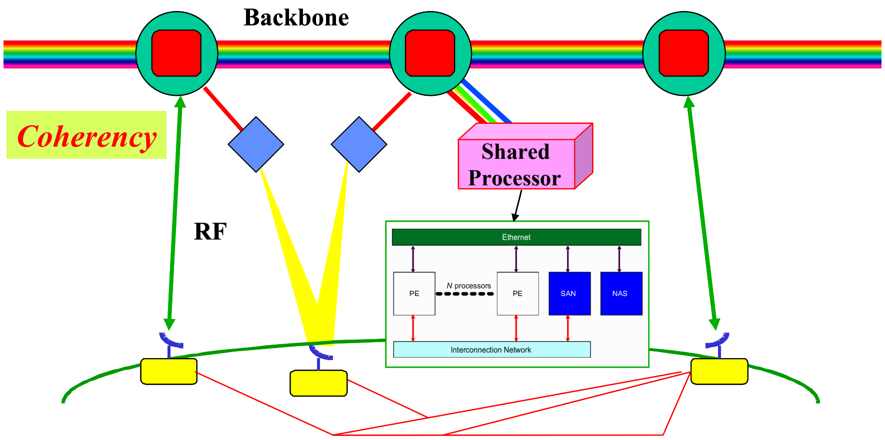


Figure 2.1. Distributed space communication using multiple satellites and long baseline interferometry for frequency re-use, and suppression of interfering signals.

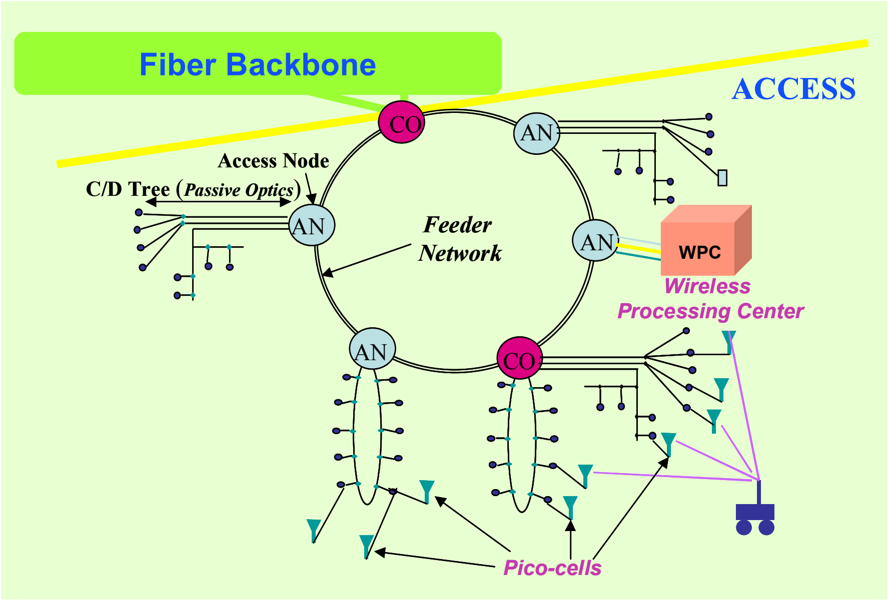


Figure 2.2. Terrestrial wireless networks with pico-cells and fiber network back-haul infrastructure.

Generically, the signal flow and processing architecture is depicted in Figure 2.3. Multiple signals are band-pass-filtered before being brought down to base-band for digitization and the digital samples are sent to a processor for processing. If a high speed digital link is used, the backbone transport is after the A/D converter and before the processor interconnection network. If analog links are used, the backbone transport is between the band-passed signals and the A/D converters at the site of the processor (Figure 2.3).

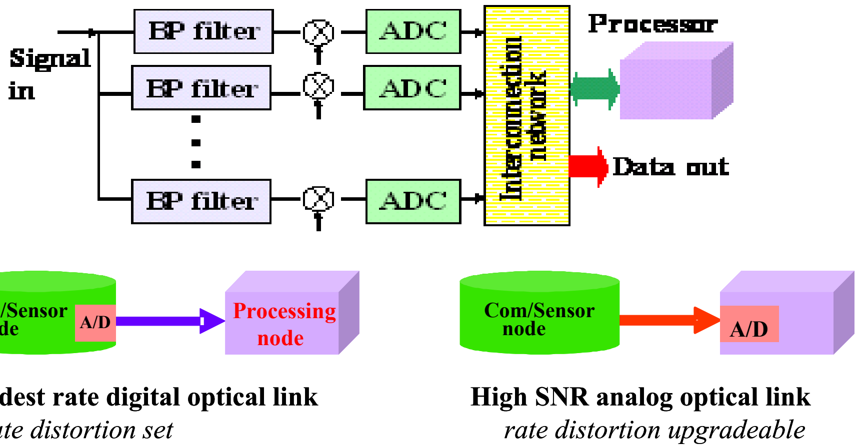


Figure 2.3. Options for the backbone transport network: digital or analog.

When the RF receiving platforms are separated by considerable distances, close in users/interferers can be using the same frequencies and their signals can still be separated (Figure 2.4). A throughput of as large as 60% can be realized for users/interferers within 1/5 of a "beam-width", an increase of x 3 over a non-nulling system. For antennas that are on satellites separated by 1 Km, the two emitters can be as close as 100m. When the number of antenna signals are more than two, the signal processing becomes more complicated due to the highly nonlinear nature of the problem.

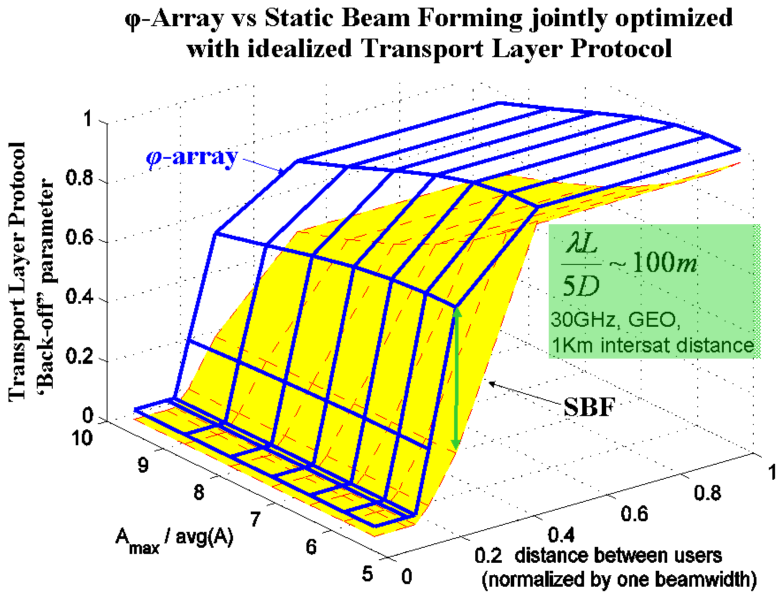


Figure 2.4. Optimum network throughput performance over phase array antenna and idealized transport layer protocol scheduling.

2.3 Joint source and network coding *(Muriel Médard)*

This work explored robust and scalable approaches to the reach-back problem for extracting data from several distributed sources. The approach will sought to establish a parallelizable and distributed approach to pull data from correlated sources under a possibly unknown and varying network topology. We considered network topology changes that are too slow to allow for averaging approaches, such as traditional coding to achieve mean performance for ergodic systems, but too rapid to allow for on-line architectural changes. These changes in topology are congruent with the slow but dramatic changes in topology associated with satellite systems, for instance because of coverage issues. We have also considered joint routing and compression in a topology-independent fashion. In particular, approaches that do not require updating of routing information, which, because of the long delays associated with satellite systems, may severely hamper the operation of traditional networking approaches. Instead, the work uses a distributed approach in which each node performs a mapping from inputs to outputs in a fashion that blurs the line between routing and compression. This distributed approach would require no or little sharing of states among nodes or even overall knowledge of topology. Instead, the net effect of transmission through the network need only be known at the receiving nodes. The obviation of the need to share state among nodes implies that stability issues under network changes are significantly alleviated. Such work begins to address meta data integration of disparate information (something like a web search engine that examines meta data on disparate sources and is able to associate the data to add value in an information sense).

2.3.1 Functional compression via graph coloring

We considered two discrete memoryless sources compressing their information in a distributed manner to be decoded at some central decoder. By distributed, we mean that neither source encoder has information (except by way of distributions) of the other source. The problem thus presented is the classical Slepian-Wolf problem whose solution is well-known [1].

Next, suppose the decoder does not need to recover both sources completely. Instead, suppose that the decoder simply needs to recover some function of those sources. The decoder does not recover both sources in this case. Moreover, if it does not need to recover each source entirely, it stands to reason that compressing at a rate better than that required to recover each source would be sufficient to recover the function. Let us make this clear with an example.

Example: *Consider two sources uniformly and independently producing k-bit integers X and Y ; assume k > 1. We assume independence to bring to focus the compression gains from using knowledge of the function. First suppose f(X,Y) = (X;Y) is to be perfectly reconstructed at the decoder. Then, the rate at which X can encode its information is k bits per symbol (bps); the same holds for Y. Thus, the sum rate is 2k bits per function-value (bpf).*

Next, suppose f(X,Y) = X +Y mod 4. The value of f(X,Y) depends only upon the final two bits of both X and Y. Thus, at most (and in fact, exactly) 2 bps is the encoding rate, for a sum rate of 4 bpf. Note that the rate gain, 2k - 4 is unbounded because we are reducing a possibly huge alphabet to one of size 4.

Finally, suppose f(X,Y) = X +Y mod 4 as before, but the decoder is allowed to recover f up to some distortion. We consider the Hamming distortion function on f. Consider recovering f up to a 1-bit Hamming distortion. One possible coding scheme would simply encode the single least significant bit for both X and Y. Then one could recover the least significant bit of f(X;Y), thus achieving an encoding rate of 1 bps per source or 2 bpf.

The problem in the above example is relatively simple because the function is separable. The function need not always be separable: consider f(X,Y) = |X - Y|, for example. Further, in the more general setting where N sources are separately compressed and a function on those N sources is to be decoded at a decoder with side information, we need a more general framework.

Here, we allude to the applications of this problem in sensor networks where bandwidth constraints encourage better compression techniques. Whenever the central node's intent is to perform some specific task on the source data (such as in a privacy-preserving database, or a camera network, or Blue Force Tracking discussed in Section V), our framework will provide compression gains.

The main idea of this research is that of using knowledge of the decoder's final objective, to achieve the best possible compression rates. We provide a general framework that allows us to solve both the problem of finding the best possible rates as well as finding coding schemes that allow for approximations of these rates.

Consider two discrete memoryless sources X and Y. In such a setup, we considered three problems roughly labeled as the functional side information problem, the distributed functional compression problem, and the functional rate distortion problem. We provide an overview of previous results here while deferring the technical details until later in the paper.

2.3.2 Functional side information

We considered the functional side information problem. In this problem, one source Y is available at a decoder. The decoder also has access to an encoded version of the second source X. The goal is for the decoder to compute f(X,Y) given the encoding from X at a minimal rate.

We described an optimal coding scheme for coding X such that f(X,Y) can be computed at the receiver. The functional side information problem can also be considered a special case of the distributed functional compression problem when one source is compressed at entropy-rate thus allowing for reconstruction at the decoder.

Wyner and Ziv [2] gave the best encoding rate for X in order to recover X at the decoder (i.e. f(X) = X) under a distortion condition. This was extended by Yamamoto [3] who considered the computation of a function f(X,Y) at the receiver; he also gave the rate-distortion function for encoder X. This was further extended by Feng, Effros, and Savari [4] who considered the case where only noisy versions of X and Y were available at the encoders; they also gave a rate-distortion function. All of the above rate-distortion functions are given in terms of an auxiliary random variable, W.

For zero distortion, Orlitsky and Roche [5] gave a new interpretation to W. They showed that W could be defined over independent sets of the characteristic graph with respect to the sources and the function. Nevertheless, the coding scheme does not lend itself to simple application.

The rate-distortion function for the functional side information problem was given by Orlitsky and Roche in the zero-distortion case. For this problem, our goal is to provide a new interpretation for that rate through a simple algorithm that can be approximated easily. We provide an alternate interpretation of W. This interpretation leads to an achievability scheme for the Orlitsky-Roche rate that is simple; it can be extended to and motivates all of our other schemes.

Specifically, we considered a modular approach to the compression problem. We take as given a fixed-block-length encoder, which can encode a source X at rate H(X|Y) (or later, Slepian-Wolf encoders that separately encode at a joint rate in the relevant rate region). Indeed, there is abundant literature suggesting near-optimal codes. The input to this module is a precoded version of the data.
The data is "colored." Specifically, we constructed a graph, the characteristic graph, whose vertices are the source values, and whose edges are given according to the conflict between vertices. These vertices are then colored according to a "minimum entropy coloring." Finding such a coloring is difficult (in fact, NP-hard). Nevertheless, there is abundant literature here suggesting several heuristics to the given problem.

We prove the coding scheme given by these two modules (first color, then send colored data with an entropy code) is optimal, in the sense of minimal rate. This gives us hope that such a modularization of the code would be optimal in the other problems as well.

2.3.3 Functional distributed source coding

We also considered the distributed functional source-coding problem. This is the same as the previous problem except instead of Y available at the decoder, only an encoded version of Y is available. The problem now is of determining the set of all rate pairs at which the decoder can compute the function f(X,Y), the rate region. We assume the sources do not communicate with each other.

Slepian and Wolf [1] solved the problem when f(X,Y) = (X,Y). The solution for a general deterministic function f is unknown. We extend the Slepian-Wolf rate region to general (deterministic) function f(X,Y). We show that we can precode the data using the same graph techniques as in the side information problem, and then use a Slepian-Wolf code over the precoded data, in order to recover a function at the receiver.

In other words, we have found a way to modularize the compression to give a scheme that always achieves better rates than Slepian and Wolf, and under a certain mild condition, the scheme achieves the optimal rates. This modular scheme uses the same modules as the side-information scheme: coloring and a Slepian-Wolf code. Thus, we provide a complete rate region under a mild condition and an inner bound for the general case.

2.3.4 Functional rate distortion

Finally, we considered the functional rate distortion problem in the side information case. This is the same as the functional side information problem except the function must be computed within some nonzero distortion.

Here, Yamamoto has given a single-letter characterization. However, as stated earlier, this is in terms of the auxiliary random variable W. For the general nonzero distortion case, an interpretation for W that lends itself to a simple scheme has not been established.
We give a new interpretation to Yamamoto's rate-distortion function for nonzero distortion. Our formulation of the rate distortion function leads to a coding scheme that is an extension of the schemes for the other problems.

Further, we give a simple achievability scheme that achieves compression rates that are certainly better than the Slepian-Wolf rates, but also better than our previous rates, which is to be expected because of the distortion allowance. This uses a specific small graph over which to apply the previous results. This graph is necessarily colored at a lower rate than in the side information problem.

References

[1] Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.
[2] Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
[3] Yamamoto, "Wyner-Ziv theory for a general function of the correlated sources," *IEEE Trans. Inform. Theory*, vol. 28, no. 5, pp. 803–807, Sept. 1982.
[4] Feng, M. Effros, and S. Savari, "Functional source coding for networks with receiver side information," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Sept. 2004, pp. 1419–1427.
[5] Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

3 Future Optical Network Architecture

Sponsors
DARPA - HR0011-08-1-0008
National Science Foundation (NSF) – CNS-0626800

Project Staff
Vincent W. S. Chan, Anurupa Ganguly, Bishwaroop Ganguly, Ori Gerstel (Cisco Systems), Kyle Guan, Ashwin Gumaste (Indian Institute of Technology), Terrence McGarty, Perry Shum (Nanyang Technological University), Eric Swanson, Guy Weichenberg, Yonggang Wen

Optical networking technology is poised to make significant contributions in next-generation data network architecture because of: (i) the enormous viable data bandwidth of optical fiber, and (ii) the ability to perform simple operations (e.g., switching) in the optical domain “in bulk”, leading to high-rate services which are independent of bit rate, format, and protocol. With the ubiquitous deployment of WDM technology in core networks in recent years – capitalizing on property (i) – long-haul transmission has become cheaper than routing and switching at core node routers. To continue reaping the economic benefits of optical networking technology, property (ii) of optical networking technology must be exploited via optical access and some form of optical switching and routing for large transactions. This would amount to a significant network architecture change – not mere substitution of optical components for electronic ones. The consequence of such a shift towards optical networking technology is that most architectural elements of networks – from the physical layer to higher layers, as well as network management and control (NMC) – must be rethought at the most fundamental level. Since NMC expenses can constitute a significant portion of the cost of a network, its cost-optimization is critical to ensure that it does not negate the other economic benefits of employing optical networking technology.

The FONA architecture is geared towards initial adoption for defense applications. The fast flow switching architecture is needed for on-demand high-data rates, distributed-sensor tasking, as well as networking and processing with short time deadlines. The efficient optical network’s state diagnostic architecture, based on graph-theoretic and information-theoretic techniques, will ensure that the network is available at the short time scales at which services must be provided. Our architecture will support this dynamic service with far fewer network resources, and will also allow idle network resources to be used by other lower-priority services in the absence of such higher-priority services.

In the following, we report on our progress in developing the FONA architectural concepts, which are geared towards the creation of optical networks that are high-rate, low-cost, and very agile for user requirements – especially for those in the defense sector – with high assurance and availability.

3.1 Performance-cost tradeoffs of optical transport architectures

We have studied the optical flow switching (OFS) architecture in detail [3] and compared it to other optical transport network architectures – namely, electronic packet switching (EPS), optical burst switching (OBS) of which Tell-and-Go (TaG) is a special case, and GMPLS [1,2,4]. OFS provides dynamic and fast changing capacity to meet bursty high-volume user demands. With OFS, an (off-band) signaling protocol will be employed by users to request lightpaths for their transaction durations, after which the network sets up an end-to-end lightpath for the duration of the transfer (>100 msec). When the transaction is over, the network resources are then relinquished for use by other users. Our study was motivated by our hypothesis – which was corroborated – that, for large transactions, optical flow switching is the right choice of architecture.

In our initial study of the different optical network architectures, we characterized their achievable throughput (i.e. capacity region) when used solely as a transport mechanism in the WAN [1]. Our results, notionally illustrated below in Figure 3.1, indicate that EPS maximizes capacity followed by OFS and GMPLS, followed finally by TaG and OBS.

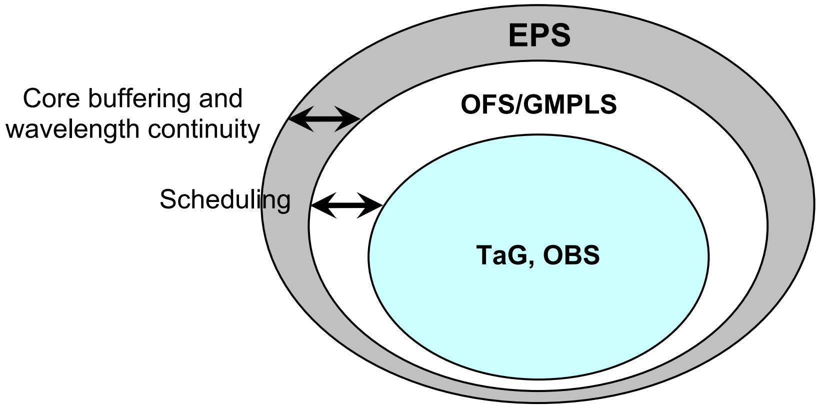
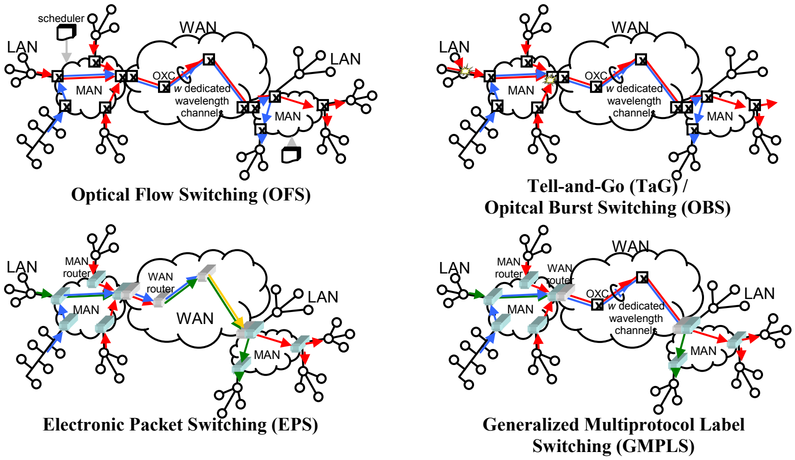


Figure 3.1. Capacity regions in the WAN

These results, while telling, are limited in impact for several reasons, including 1) the absence of a concrete cost metric, and 2) that the local- and metro-area network were suppressed. Our next studies addressed these two shortcomings [2,4].

We conducted a throughput-cost comparison of a simple, yet scalable, multi-tiered optical network comprising two groups of users, each in a distinct metropolitan-area network (MAN), which wish to communicate over a wide-area network (WAN), as drawn below. Our network cost model focused on initial capital expenditure: transceiver, switching, routing, and amplification costs. Our network model, though simple in that it only considers the communication of two sets of users across a WAN, is a building block for more complex network topologies, and, more importantly, captures the essence of the throughput-cost tradeoffs of these more complex networks.



Optical Flow Switching (OFS)

Tell-and-Go (TaG) / Opitcal Burst Switching (OBS)

Electronic Packet Switching (EPS)

Generalized Multiprotocol Label Switching (GMPLS)

Figure 3.2: Candidate architectures in the simple two MAN network setting.

Our throughput-cost comparison of the four network architectures, using cost and architectural parameters which reflect the state of present-day networks, indicates that each architecture is optimum for a range of user rates [2,4]. Figure 3.2 illustrates the normalized network cost for each of the four possible homogeneous architectures, given a particular network size; and Figure 3.3 illustrates the cost-optimal homogeneous network architecture as a function of the number of users per MAN and the average user data rate. EPS and EPS/GMPLS dominate for lower rates because a small amount of electronic equipment is necessary to support the aggregate traffic; whereas, for OFS and TaG, expensive tunable, long-haul transceivers are always required at each end user. On the other hand, at high data rates, regardless of the number of users, OFS always dominates, implying that OFS is the most scalable architecture of all. In the high user data rate regime, aggregate traffic is always high, so requiring electronic equipment to support this traffic in the network – even if only in the MAN – is expensive. We note that there does not exist a regime of optimality for TaG since the low cost of scheduling in OFS yields great performance benefit relative to the otherwise identical TaG architecture.

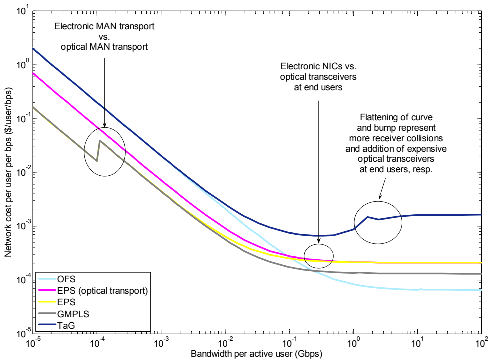


Figure 3.3: This plot assumes that 10% of users are active at any time, and that routers run at 30% utilization.

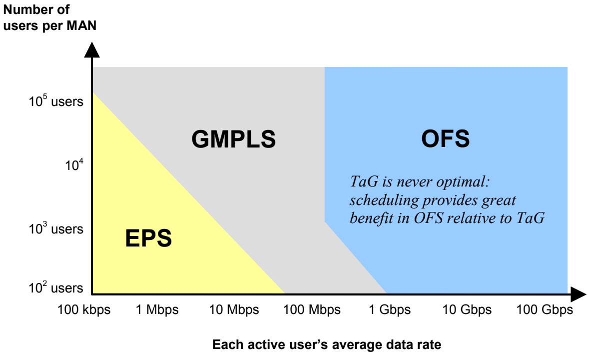


Figure 3.4: Cost-optimal homogenous architecture vs. network size and data rate when a homogenous user population is assumed.

Figure 3.4 suggests that a hybrid network architecture is likely the most sensible design of all, especially when user demands are heterogeneous. For example, a network in which high data rate users employ EPS or GMPLS may perform better than a network in which only one of the aforementioned architectures is employed. Indeed, Figure 3.5 illustrates that if user demands are heterogeneous, then a hybrid network architecture can be more cost-efficient than a homogeneous architecture [2]. Figure 3.6 illustrates the normalized cost for the different components of the cost-optimal architecture. Note that, when employed, OFS has the lowest normalized cost, followed by GMPLS and then EPS. Finally, Figures 3.7 and 3.8 illustrate the proportion of bandwidth and users, respectively, served by the different components of the cost-optimal architecture.

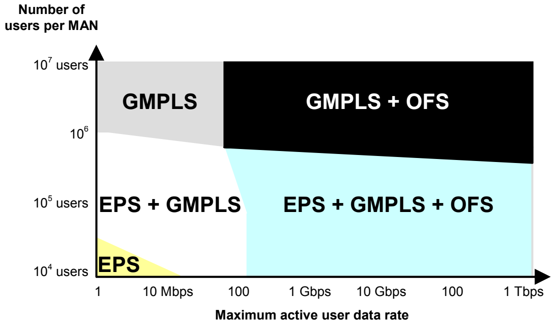


Figure 3.5: Cost-optimal hybrid architecture vs. network size and data rate for a heterogeneous (truncated heavy-tail distributed) user population.

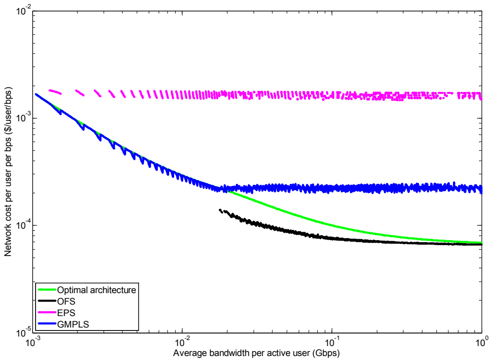


Figure 3.6: Normalized cost by component architecture. In this figure, it is assumed that user data rates have a heavy tailed distribution (with exponent -1.5) from a minimum of 10^{6} bps to a maximum value plotted on the x-axis.

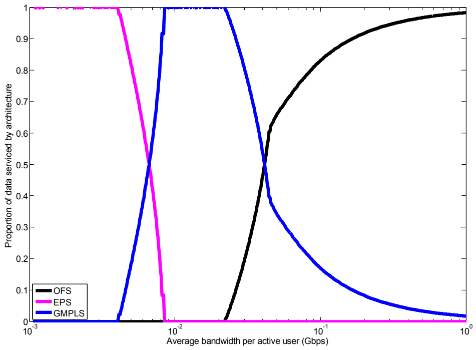


Figure 3.7: Proportion of bandwidth served by the component architectures

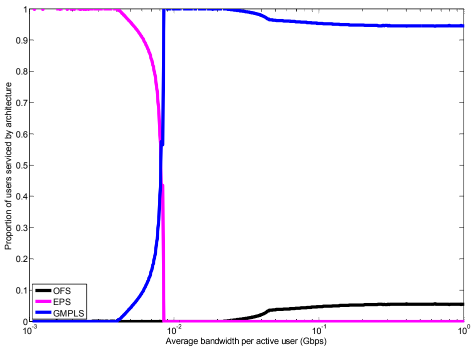


Figure 3.8: Proportion of users served by the component architectures

References

[1] "On the Capacity of Optical Networks: A Framework for Comparing Different Transport Architectures," G. Weichenberg, V. W. S. Chan, and M. Medard, *Proceedings of IEEE Infocom, Barcelona*, April 2006.

[2] "Cost-Efficient Optical Network Architectures", G. Weichenberg, V. W. S. Chan, and M. Medard, *Proceedings of ECOC*, September 2006.

[3] "Optical Flow Switching", V. W. S. Chan, G. Weichenberg, and M. Medard, *Proceedings of Workshop on Optical Burst Switching*, October 2006.

[4] "On the throughput-cost tradeoff of multi-tiered optical network architectures", G. Weichenberg, V. W. S. Chan, and M. Medard, *Proceedings of Globecom 2006*, San Francisco, November 2006.

3.2 Cost-efficient physical architecture for OXC-switched WDM mesh networks

In current telecom environment, carriers have deployed huge capacity in the long-haul networks. At the meantime, end users' access to higher data rates is still costly. To bridge the gap between the bandwidth glut at the backbone and the high access cost for the end-users, the architecture of next generation optical MAN will be an important contributor to the reduction of access network cost. The objective is to design networks that not only require a low installation cost, but also have good scalability -- a decreasing cost-per-node-per-unit-traffic as the number of users and transaction size increase. This architecture feature is essential for any commercially deployed network to attract serious providers and investors to commit to the venture as part of a sensible business.

The central theme of this thesis is to identify scalable network architectures over the possibilities of optical networks as allowed by the technology: fiber connection topologies, switching technology selection and dimensioning, as well as routing and wavelengths assignment, with emphasis on exploring the benefit of optimizing over fiber connection topologies. Due to the intrinsic complexity of such an optimization problem and because of our interest in gaining insights into how the cost structures drive architectural tradeoffs, for the first part of this thesis, we take an analytical approach by concentrating on networks with regular topologies and (static and random) uniform all-to-all traffic. These assumptions are idealizations, nonetheless they keep the analysis tractable and provide us insights into more complex problems and they act as points of departures for the analysis of more realistic scenarios (such as irregular networks and non-uniform traffic) in the later part of this thesis.

The search for the scalable fiber connection architecture hinges on analyzing the tradeoffs among expensive network resources. In our parametric, first-order, and homogeneous cost model, the constituent parts, which are closely related to fiber topology, are fiber cost and switching cost. To build a network, one can support lightpaths by laying down direct fiber connections among all source-destination nodes. This design obviously requires minimal switching resources but maximal amount of fibers. Another way to establish lightpaths is by hopping through one or more nodes. Such a design requires less fiber connections but more switching resources. As such, the optimal connectivity of a fiber connection topology is determined by the fiber-to-switching cost ratio. Further, we show that the amount of switching resources used at nodes is proportional to the average minimum hop distance (for regular topologies and uniform traffic)[1]. To support the same set of (uniform) demands, we show that regular topologies with the smallest average minimum hop distances have lowest fraction of pass-through traffic and thus require less switching ports.

These provide guidelines for rigorous studies of cost-effective network architectures. For a few representative classes of regular networks (e.g., Δ-nearest Neighbors and Generalized Moore Graphs, whose constructions are detailed in [1]), we first derive or approximate the closed-form expressions for important parameters, such as average minimum hop distance, switch size, and network cost. We then set up the corresponding optimization formulations. We have found that for regular networks and uniform traffic, the joint design problems of fiber connection topology, dimensioning, and routing can be solved optimally and analytically for a special class of regular graphs -- Generalized Moore Graphs [2]. That is, we prove that with minimum hop routing, Generalized Moore Graphs, whose average minimum hop distances scale favorably as $\log_{\Delta}^{N}$, achieve the lower bound on network cost and are good reference topologies. We also show that topologies with structures close to Generalized Moore Graphs can achieve close-to-minimum cost. The investigation of the cost scalability further demonstrates the advantage of the Generalized Moore Graphs and their close relatives as benchmark topologies: with linear switching cost, the minimal normalized cost per unit traffic decreases with increasing network size (Figure 3.9). In comparison, for less efficient fiber topologies (e.g., Δ-nearest Neighbors) and switching cost structures (e.g., quadratic cost), the minimal normalized cost per unit traffic plateaus or even increases with increasing network size. Our study has also revealed other attractive properties of Generalized Moore Graphs in conjunction with minimum hop routing. When minimum hop routing is employed for uniform traffic in Generalized Moore Graphs, the aggregate network load is evenly distributed over each fiber. Further, to support a given uniform traffic demand, Generalized Moore Graphs require the minimum (or close to the minimum) number of wavelengths, which directly affects the complexity and the cost of dispersive elements and filters in the network. This is the first time that Generalized Moore Graphs are identified as optimum architectures in the context of network cost efficiency. These architectures are very different from the currently used ones in MANs, such as rings, interconnected rings, or non-optimized mesh networks (Figure 3.9).

With the parametric network cost structure, closed-form solutions of the optimal degree and cost as functions of various network design parameters (such as network sizes and wavelengths of traffic between node pairs) are obtained. These results show that for a MAN of moderate size (a few tens to a few hundred nodes), under certain cost structures, neither rings nor fully connected mesh networks are optimal topologies. The optimal network connectivity is in the range of 0.03*N* to 0.1*N* (when the fiber-to-switch cost ratio α / β_1, α and β_1 denote the cost per fiber connection and the cost per port for linear switching cost structure, respectively). The advantage of analytical approaches is self-evident: they provide valuable references on how the optimal network connectivity scales as the design parameters change. More importantly, the results demonstrate that switching technologies have a tremendous impact on the final topological architectures. The optimal topologies connecting the same set of nodes can differ significantly when different switching fabrics are used, even when these topologies are designed to serve the same traffic demand. Among all-optical technologies currently available, for smaller networks (a few to a dozen nodes) and light traffic, quadratic switching cost structures (e.g., 2-D switching fabrics) have cost advantage. However, as the size of the network and

---

[1] For example, in [1] we show that the size of an OXC, K_{o}, equals to $(N-1)t[H_{\min}(N,\Delta)+1]$, with $H_{\min}$, *N*, and *t* denoting the average minimum hop distance, network size, and unit of traffic between a node pair, respectively.

[2] As elaborated in [1] [3], the assumptions of uniform traffic and symmetric topology, as well as the special construction of Moore Graphs make this joint problem solvable. For non-uniform traffic and other classes of graphs, the joint problem remains difficult to solve.

the demand among node pairs increase, linear switching cost structures (e.g., 3-D switching fabrics) have the best scalability. Thus, the cost benefit of deploying 3-D switching technology for the future network is apparent. Moreover, a comparison of the cost benefit between OXC and OEO switches shows that at low data rate (e.g., below 1Gb/s for every source-destination pair for a 50-node network, as shown in Fig. 3.9), it is economical to use OEO switches; at high data rate (e.g., above 1Gb/s for every source-destination pair for a 50-node network, as shown in Fig. 3.10), it is more cost-advantageous to use OXC switches.

We also take steps to broaden the scope of this work to address more realistic design scenarios from two facets. We look into irregular network topologies and (static) non-uniform traffic, which represent most existing networks. We show that if the switching cost is linear with port counts, minimum hop routing is still optimum. The results of Generalized Moore Graphs can be used to provide useful estimates for the cost of irregular networks - a Generalized Moore Graph with a node degree that equals to the minimal node degree of an irregular network provides lower bound on network cost. Also Generalized Moore Graphs can be exploited to suggest improvements for irregular fiber connection topologies. We show examples that connecting the same set of nodes via a Generalized Moore Graph results in savings on the number of switching ports[3]. For irregular topology under arbitrary traffic, we provide the lower bound on network cost using the concept of minimum flow tree, which is based on the unique construction of Generalized Moore Graphs (each of its nodes has a full (or almost full) Δ-ary routing spanning tree) and an (idealized, not always realizable) permutation of traffic matrix (the Δ destinations with the largest traffic are connected by one hop paths, the next Δ^2 destinations in descending order of traffic are connected by two-hop paths, and so on). More importantly, the construction of the minimum flow tree, though idealized, yields a general yet crucial design guideline: a cost-effective physical topology should minimize the propagation of large traffic values.

We also investigate designing networks that are robust to demand uncertainties, which are caused by diversification of services, changing of usage patterns, and data-dominated traffic in the metro environment, etc. We present a framework to assist network designers to dimension optical MAN, incorporating uncertainties in demands. In this framework the interplay among topology design, resource provisioning, and routing are analyzed based on two stochastic optimization models that use probability distributions of demands as inputs. In one model, the weighted sum of network installation cost and expected penalty cost for unsatisfied traffic is minimized. In another model, the network installation cost is minimized subject to certain service level requirements. The optimization results enable us: (1) to identify the Generalized Moore Graphs as the physical architectures that are most robust (in cost) to demand uncertainties among rich classes of regular topologies, assuming the (random) uniform all-to-all traffic, and (2) to provide analytical references on how optimal dimensioning, network connectivity, and network costs change as functions of the designer's level of risk aversion, service level requirements, and probability distributions of the demand.

References

[1] C. Guan and V. W. S. Chan, “Topology Design of OXC-Switched WDM Networks”, *IEEE Journal for Selected Area in Communication*, vol. 23, no. 8, pp.1670-1686, Aug. 2005.
[2] C. Guan and V. W. S. Chan, "Topology Design and Resource Dimensioning for Optical Mesh Networks under Demand Uncertainties”, *Globecom 2005*, St. Louis, Nov. 2005.
[3] C. Guan and V. W. S. Chan, “Cost-Efficient Physical Architecture for OXC-switched WDM Mesh Networks – Generalized Moore Graphs and Their Close Relatives”, *Globecom 2006,* San Francisco, November 2006.

---

[3] In these examples, every improved network uses the same number of fiber connections as the original one does, as illustrated in [3].

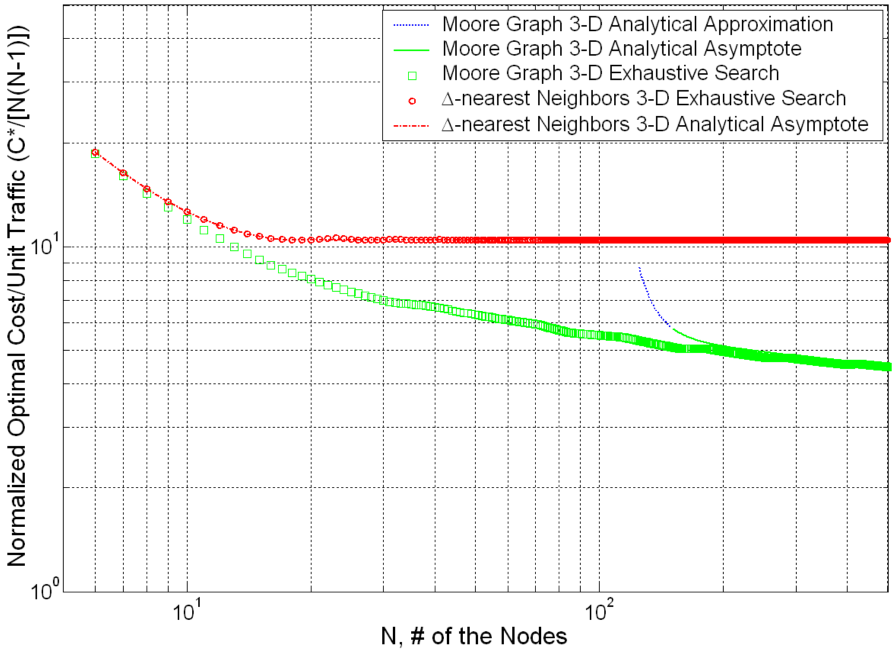


Figure 3.9: Optimal normalized network cost per unit traffic $C^*/[N(N-1)]$ as a function of network size N for the Δ-nearest Neighbors and the Moore Graphs. The points represent the results of exhaustive searches; while the lines represent the analytical asymptote. The switching fabric is 3-D with $F_1(K_o)=\beta_1 K_o$. The fiber-to-switching cost ratio α/β1=40

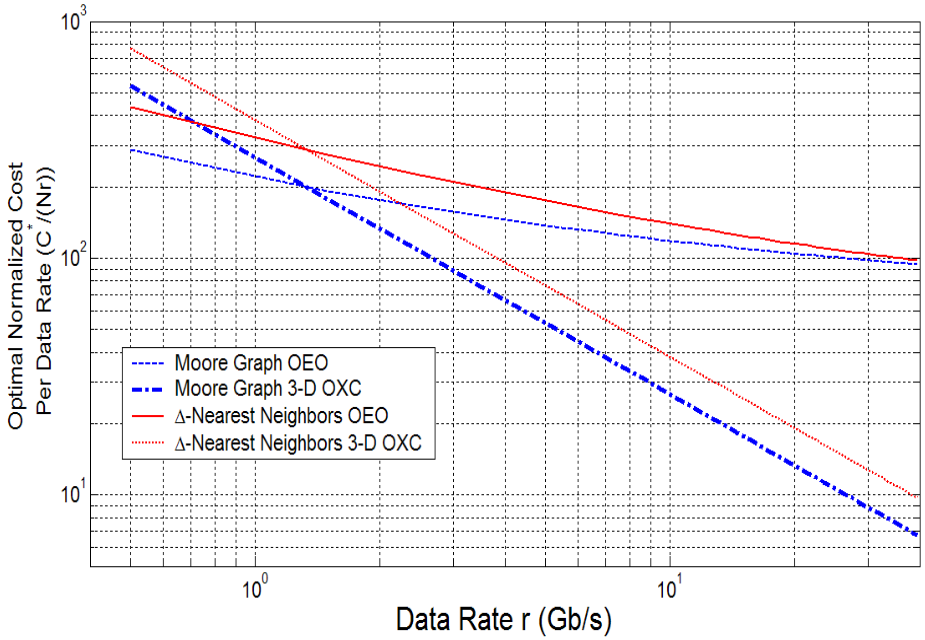


Figure 3.10: Minimal normalized network cost per data rate $C^*/(Nr)$ as a function of data rate per wavelength r for combinations of two classes of network topologies (Δ-Nearest Neighbors and Moore Graphs) and two types of switching fabrics (OEO switch and 3-D OXC). *N*=50, $\alpha=20$, $\beta_1=1$, and $\beta_e=7.5$.

3.3 Cost-effective fault diagnosis for all-optical networks

All-optical networks [1, 2], where data traverse lightpaths without any optical-to-electronic conversion at intermediate nodes, promise significant cost benefits. The significant cost savings are due to optical switching of high data-rate lightpaths at intermediate network nodes, thereby reducing electronic processing costs. As a result, broadband network services can potentially be delivered to large populations at much lower cost than today's technologies. However, without the electronic processing capability at intermediate nodes, network architects have to develop new technologies to address problems whose solutions previously rely on the electronic processing capability at intermediate nodes. In particular, all-optical networks are susceptible to various physical failures, e.g., fiber cuts, switch node failures, transmitter/receiver breakdowns, and optical amplifier breakdowns. These failures can result in the disruption of communication, and can be costly to detect and localize within the current management framework (e.g. SONET/G.709). Since all-optical networks lack parity checks at intermediate nodes as SONET/G.709 does, either optical signal is tapped out at each intermediate node for parity check or new mechanisms are needed to diagnose link/node failures. If tapping is indeed done, a lot of cost gains of all-optical networks will be mitigated. Therefore, we aim to develop new fault diagnosis approach to keep the network operating cost low.

Instead of the passive paradigm based on parity check in SONET/G.709, we have proposed a proactive fault diagnosis paradigm in [3-5]: optical probing signals are sent along some lightpaths to test the health of the network, and *probe syndromes* (i.e., results of the probes) are used to differentiate failure patterns. The design of proactive fault diagnosis schemes for all-optical networks bears two key objectives: (i) detecting faults quickly, and (ii) keeping the diagnosis cost low. The importance of objective (i) stems from the current SONET standard, in which the 50-ms restoration time leaves little room for fault detection and localization. This will probably be reduced further in future all-optical networks to avoid large amount of data loss during a short period of communication disruption. Hence, when parts of a network are malfunctioning, it is critical to locate and identify these failures as soon as possible. At the same time, the cost of fault diagnosis has to be kept low such that the cost advantage of all-optical networks, compared to traditional optical networks, can materialize.

We believe that the two design objectives could be tightly related to two parameters of proactive fault diagnosis schemes (i.e., the number of probes and the number of probing steps[4]). First, the number of probes could serve as the manifestation of fault management effort. In particular, each probe requires certain amount of effort in both network management/control plane (e.g., signaling) and data plane (e.g., transmission and detection) that otherwise could be used to generate revenue. In addition, each probe results in one bit of management information, whose transportation, storage and processing consumes additional network resources. Second, under the assumption that each step takes approximately equal amount of time, the number of probing steps indicates how fast the fault pattern could be identified. In this research, we exploit two alternative designs for choosing probes (i.e, *adaptive* probing, and *non-adaptive* probing) to balance these two objectives.

3.3.1 Adaptive fault diagnosis schemes

In adaptive fault diagnosis schemes [3-5], probing signals are sequentially sent to probe the health of the network until the failure pattern is identified. Owing to its sequential nature, successive probes can be chosen according to previous probe syndromes, and thus the number of probes required is usually quite small. However, the number of probing steps might be quite large for some network failure patterns and/or in some large networks. Indeed, the design objective is to minimize the average number of probes to identify any failure pattern (Figure 3.11).

Under the probabilistic link failure model [3], we have established a mathematical mapping between the fault diagnosis problem in network management and the source coding problem in Information Theory, under the constraint that only probes along lightpaths in the network are permissible. This

---

[4] One probing step corresponds to a set of parallel probes.

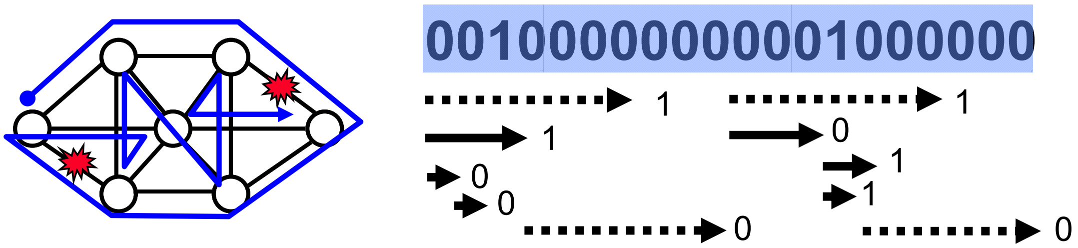


Fig. 3.11. Demonstration of run-length probing scheme over a network. It contains a sequence of concatenations of two phases: the failure detection phase (dotted lines) and the failure localization phase (solid lines).

mapping has both theoretical and practical implications in designing efficient adaptive fault diagnosis schemes for all-optical networks.

Theoretically, this mapping first suggests that the average number of probes is lower bounded by the information entropy of network states, i.e.,

$$L \geq mH_b(p), \tag{1}$$

where L is the average number of probes, m is the number of links in the network, p is the link failure probability, and $H_b(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the Shannon information entropy function. This lower bound can be understood intuitively as follows. First, the total amount of network state information contained in the failure pattern is the information entropy $mH_b(p)$. Second, each probe can provide at most 1 bit of network state information. Therefore, the total number of probes is at least the information entropy of the failure pattern.

Practically, this mapping also suggests an effective approach to designing efficient adaptive fault diagnosis schemes via translating existing source coding algorithms, under the probe feasibility constraint. Using this information insight, we have developed, for Eulerian networks, the run-length probing scheme, whose probe syndrome is a concatenation of run-length codewords, as illustrated in Figure 3.11. We have also characterized its performance as

$$\overline{L}_\infty(p) \approx p \cdot \left(\lfloor \log_2 K \rfloor + 1 + \frac{(1-p)^k}{1-(1-p)^K} \right), \tag{2}$$

where $\overline{L}_\infty(p)$ is the average number of probes per link, $K = \lceil -\log_q(1+q) \rceil$ and $k = 2^{\lfloor \log_2 K \rfloor + 1} - K$. Analytical and numerical investigation suggests that the average number of probes is always within 5% of the entropy lower bound, and approaches the entropy lower bound as the link failure probability decreases.

With a deep understanding of the run-length probing scheme for Eulerian network with probabilistic link failures, we have extended its applications to more practical situations such as non-Eulerian networks with link failures [4] and networks with link/node failures [5]. For non-Eulerian networks with probabilistic link failures, we suggested two alternative approaches to deploy the run-length probing scheme: (1) the disjoint trail decomposition approach that decomposes the network into a set of link-disjoint trails for diagnosis, and (2) the path augmentation approach that replicates a minimum set of links to make the network Eulerian.

Our analytical and numerical investigations reveal a *guideline* for efficient fault diagnosis schemes: *each probe should provide approximately 1-bit of information and the total number of probes required is approximately equal to the entropy of the state of the network.* This result provides an insightful guideline to reduce the overhead cost of fault management for all-optical networks and can further the understanding of the relationship between information entropy and network management.

3.3.2 Non-adaptive fault diagnosis schemes

In non-adaptive fault diagnosis schemes [6], instead of sending optical probing signals sequentially, a pre-determined set of probing signals are sent in parallel to probe the network state of health. so that the number of probing steps is always one. In addition, compared to the probabilistic failure model (i.e., each link fails independently and no upper bound on the number of failures) used in our previous work, we also assume a worst-case failure model in that the number of simultaneous failures is upper bounded by a constant. Under such a framework, the design objective is to minimize the number of parallel probes for non-adaptive fault diagnosis schemes, so as to keep the fault diagnosis cost low.

Our fault detection methods are based on techniques from the field of *combinatorial group testing (CGT)* [7], where defected samples are identified through a set of parallel testing on different combinations of unknown samples. In our work, we propose a variant of classical CGT in which the valid tests are determined by the structure of a graph. In the all-optical network context, this graph corresponds to the network topology, and the constraint on valid tests is due to the fact that lightpaths can only traverse a set of interconnected edges. We formally analyze the number of tests needed for certain interesting classes of graphs (e.g., ring, bus, tree, 2-D grid, and complete graph), and even arbitrary graphs. We find out that the number of probes needed depends on the topology. In some cases, we can give matching upper- and lower-bounds on the number of tests needed. Our fault diagnosis schemes have a common theme, which suggests a practical rule-of-thumb for efficient fault diagnosis schemes: *a fault-free sub-graph in the network topology should be identified, and used as a "hub" to diagnose other failures in the network*.

References

[1] V. W. S. Chan, et al, "A pre-competitive consortium on wide-band all-optical networks," *IEEE/LEOS Journal of Lightwave Technology*, Vol. 11, No. 5/6, May/June, 1993, pp. 714-735.

[2] V. W. S. Chan, "All-optical networks," *Scientific American*, No. 3, September 1995, pp. 56-59. Y.G. Wen, V. W. S. Chan and L. Z. Zheng, "Efficient Fault Diagnosis Algorithms for All-Optical WDM Networks with Probabilistic Link Failures (*invited paper*)," *IEEE/OSA Journal of Lightwave Technology*, Vol. 23, No. 10, October 2005, pp.3358-3371.

[3] Y. G. Wen, V. W. S. Chan, and L. Z. Zheng, "Efficient Fault Diagnosis Algorithms for All-Optical Networks: An Information Theoretical Approach," *2006 IEEE International Symposium on Information Theory (ISIT),* Seattle, WA, July 2006.

[4] Y. G. Wen, V. W. S. Chan, and L. Z. Zheng, "Efficient Fault Detection and Localization for All-Optical Networks," *2006 IEEE Global Communications Conference (GlobeCom),* San Francisco, December 2006.

[5] N. Harvey, M. Pătraşcu, Y. G. Wen, S. Yekhanin and V. W. S. Chan, "Non-Adaptive Fault Diagnosis for All-Optical Networks via Combinatorial Group Testing on Graphs," *2007 IEEE Conference on Computer Communications (Infocom),* Anchorage, Alaska, May 2007.

[6] D. Du and F. Hwang, Combinatorial group testing and its applications, *World Scientific*, 2000.