D I M A C S

Center for Discrete Mathematics and Theoretical Computer Science

DIMACS SPECIAL YEAR ON NETWORKS: 1999 ANNUAL REPORT

DIMACS Center • Rutgers University • 96 Frelinghuysen Road • Piscataway, NJ 08854-8018 • USA
TEL: 732-445-5928 • FAX: 732-445-5932 • EMAIL: `center@dimacs.rutgers.edu`
WWW: `http://dimacs.rutgers.edu/` • TELNET: `telnet info.rutgers.edu 90`

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs - Research, Bell Labs, NEC Research, and Telcordia Technologies (formerly Bellcore)

Overview:

The Special Year on Networks began in July 1996 and will continue through August 2000 with the help of a three year grant received from the National Security Agency, as it evolves into the follow-up Special Focus on Next Generation Networks Technologies and Applications. The program is stressing postdoctoral training and workshops that reach a broad segment of the dm and tcs research community as well as engage practitioners.

Organizers:

- Stuart Haber, Surety Technologies
- David Johnson, AT&T Labs - Research
- Mihalis Yannakakis, Bell Labs

Scientific Advisory Board:

- Baruch Awerbuch, Johns Hopkins;
- Badri Badrinath, Rutgers;
- Endre Boros, Rutgers;
- Ding-Zhu Du, Minnesota;
- David Goodman, Poly Univ.;
- Jack Holtzman, Rutgers;
- Frank Hsu, Fordham;
- Srinivasan Keshav, Cornell;
- Arjen Lenstra, Citibank;
- Udi Manber, Arizona;
- Ray Miller, Maryland;
- Debasis Mitra, Bell Labs;
- Clyde Monma, Telcordia Technologies;
- Panos Pardalos, Florida;
- K.K. Ramakrishnan, AT&T Labs - Research;
- Greg Shannon, Spanning Tree Technologies;
- Scott Shenker, Xerox;
- Subhash Suri, Washington U. in St. Louis;
- Eva Tardos, Cornell;
- Peter Winkler, Bell Labs; and

- Yacov Yacobi, Microsoft

Sponsors and Collaborating Institutions:

The National Science Foundation
The New Jersey Commission on Science and Technology
The National Security Agency
AT&T Labs - Research
Bell Labs
Institute for Mathematics and its Applications, University of Minnesota
NEC Research Institute
Princeton University
Rutgers University
Telcordia Technologies
WINLAB (Wireless Networking Lab), Rutgers University

Theme of Program:

This special year is intended to bring together researchers and practitioners in a wide variety of fields connected to the common theme of "Networks" and ranging from questions of infrastructure to ones of applications. Special emphasis is placed on Internet-related topics, as this is a rapidly growing area where theoretical computer science and discrete mathematics have a chance to make a real impact.

The special year's topics can be grouped into four overlapping areas:

- **Network Security:** protocols for electronic commerce, public key infrastructure, network security and threat models.
- **Network Applications:** information retrieval, databases and the Web, multimedia transmission and coding, distributed and collaborative computing.
- **Network Control:** routing, congestion and admission control, network protocols, mobile computing.
- **Network Design:** topology and connectivity, switch design, facilities location.

The goal is to engage theoreticians with these problems and foster communication between them and real-world practitioners. Practitioners have much to tell theorists about what the key issues are, and theorists may themselves be able to provide crucial insights and solutions to practitioners.

Workshops, Miniworkshops and Special Events in 1999

- **Large Scale Discrete Optimization in Logistics**
  - Dates: February 8-10, 1999
  - Location: DIMACS Center, CoRE Building, Rutgers University
  - Attendance: 79
  - Organizers: George Nemhauser and Martin Savelsbergh, Georgia Tech
  - Comment: Joint with Special Year on Large Scale Discrete Optimization

An efficient and effective logistics system is the key to being competitive in today's global marketplace. This realization has fueled an enormous interest in new technology for designing and operating logistics systems. Discrete optimization frequently is an integral part of such technology. In particular, in the last decade, the use of large-scale discrete optimization models in transportation and production planning has increased dramatically. This has been made possible by developments in modeling, algorithms, software, and hardware. This workshop brought together leading researchers to present state-of-the-art research activities and to discuss innovative future directions of research for this very active field with significant practical importance. The talks covered a wide range of topics including modeling, algorithms, and applications. Application domains that were covered include network design, production planning, vehicle routing, and crew scheduling. The workshop opened with a talk by William Cook titled "On the Solution of Traveling Salesman Problems". Research on the solution of the traveling salesman problem, the prototypical discrete optimization problem, has stimulated enormous interest in discrete optimization and has helped significantly in the growth of the field. Several other talks focused on methodology, such as "Approximation Algorithms via Linear Programming: Rounding Algorithms" by David Shmoys, "An Overview of Classical and Modern Heuristics for the Vehicle Routing Problem" by Gilbert Laporte, "GRASP: Greedy Randomized Adaptive Search Procedures" by Mauricio Resende, "AI Techniques for Large-Scale Discrete Optimization" by Jimmy Crawford, and "Very Large Scale Neighborhood Search" by Ravi Ahuja. Furthermore, there were several talks related to the main application domains of discrete optimization. In the topic of Airline Optimization, we had "Solving Aircraft Maintenance Scheduling Problems" by Daniel Bienstock, "Optimization Problems in Air Traffic Management" by Mike Ball, and "Airline Integrated Recovery: AIR and SimAIR" by Ellis Johnson. In the topic Manufacturing and Production Planning, we had "BC-PROD and LOTSIZELIB" by Laurence Wolsey, "Optimization in IBM Manufacturing" by Brenda Dietrich, and "Painting Cars and the Traveling Salesman Problem" by Tom Magnanti. In the topic Supply-Chain Management, there was "Optimal and Approximation Algorithms for Large Scale Supply Chain Management Problems" by David Simchi-Levi and "Dynamic Programming Approximations for the Discrete Multi-commodity Network Flow Problem over a Dynamic Graph" by Warren Powell. Finally, in the topic Telecommunication Network Design there was "The Prize Collecting Steiner Tree Problem" by David Johnson. There was also a full group discussion on future directions for the field. The group was very encouraged by the increasing use of discrete optimization methods in industry and the need for the solution of larger more complex models that, for example, include stochastic elements. In conclusion, the workshop provided an excellent environment in which leading researchers in the field could share their ideas not only with each other, but with junior researchers and students as well.

- **Mobile Networks and Computing**
  - Dates: March 25-27, 1999
  - Location: Rutgers University
  - Attendance: 48
  - Organizers: Badri Badrinath, Rutgers University, Frank Hsu, Fordham University, Panos Pardalos, University of Florida, Sanguthevar Rajasekaran, University of Florida

Advances in the technologies of networking, wireless communications, and miniaturization of computers lead to the rapid development in mobile communication infrastructure and have engendered a new paradigm of computing. Users carrying portable devices can freely move around while still being connected to the network. This provides flexibility in accessing information anywhere and at any time. On the other hand, the price of this flexibility has been the introduction of new levels of complexity that were not encountered in software and protocol design in wired networking. The new challenges in designing software systems for mobile networks include location and mobility management, channel allocation, power conservation, information dissemination, and sensor networks, among others. This workshop focused at bringing together researchers from academia as well as industry who are working on various aspects of mobile computing. One of the key successes of the workshop was that collaboration started among researchers from different institutes. In addition, several graduate students participated (and some presented papers) in the workshop. A refereed volume with papers presented at the Workshop will be published early next year by the American Mathematical Society. Among the presentations were discussions of DataSpace - querying and monitoring deeply networked collections of physical objects by Tomasz Imielinski; of improving the performance of TCP over wireless links, by Zygmunt J. Haas and A. Warkhedi; software implementation strategies for power-conscious systems by David Wei; and multicriteria optimization frequency assignment by Robert A. Murphy, Panos M. Pardalos, and Eduardo Pasiliao. In addition, Jean-Claude Bermond talked about routing in satellite constellations, Jeng-Ping Lin, Sy-Yen Kuo, and Yennun Huang spoke about a cluster-based checkpointing scheme for mobile computing on wide area networks, Stephen Olariu spoke about efficient initialization and leader election in mobile radio networks, and Ibrahim Korpeoglu, Pravin Bhagwat, Chatschik Bisdikian, and Mahmoud Naghshineh presented a talk about the challenges of mobile and wireless networking support for palmtop devices. Ravi Prakash and Mukesh Singhal made a presentation on the impact of unidirectional links in wireless ad-hoc networks, Javier Gomez-Castellanos and Andrew T. Campbell talked about supporting integrated-QoS in wireless networks, Sanguthevar Rajasekaran dealt with frequency assignment in cellular networks, and Sumi Helal, Jin Jing, and Ahmed Elmagarmid spoke about supporting application-level handoff in overlay networks. Brigitte Jaumard, Christophe Meyer, and Tsevi Vovor presented a paper entitled "How to Combine a Column and Row Generation Method with a Column or Row Elimination Procedure – Application to a Channel Assignment Problem", Anup K. Talukdar, B.R. Badrinath, and Arup Acharya spoke about "Rate Adaptation Schemes in Networks with Mobile Hosts", Sajal K. Das and Sanjoy K. Sen spoke about quality of service provisioning in next generation wireless networks, and Kshirasagar Naik presented a talk about energy-conserving software design for mobile computers. A talk on agent-based architectures for securing mobile IP was given by Xun Yi, one on adaptive detection of multi-level modulation formats for DC/CDMA systems in a slowly fading channel by Ali F. Almutairi, Scott L. Miller, and Haniph Latchman, and one entitled "Online Protocols for Channel Assignment in Wireless Networks" was presented by Sunil Shende. The meeting concluded with talks by Sandeep K.S. Gupta and Pradip K. Srimani on using self-stabilization to design adaptive multicast protocols for mobile ad hoc networks, Anupam Joshi on mobility and agents, Xin Liu, Panos M. Pardalos, and Sanguthevar Rajasekaran on frequency assignment problems in mobile networks, and Julio C. Navas and Tomasz Imielinski on the topic "Geographic Routing: Improving Routing through Approximation"

- **On-Line Decision Making**
  - Dates: July 12 - 15, 1999
  - Location: Rutgers University
  - Attendance: 89

- Organizers: Yoav Freund, AT&T Labs - Research, and Rakesh Vohra, Northwestern University
- Comment: This was both a Special Year Networks workshop and an exploratory workshop, investigating a potential new special focus.

In recent years there has been increasing interest in the analysis of algorithms for making repeated decisions within an unknown environment. It has been realized that interesting performance guarantees can be obtained without assuming that the environment's behavior is generated by a stochastic process. Work along these lines has been going on in several different fields under a variety of names, for example: information theory: Universal coding of individual sequences, universal prediction, universal portfolios ...; game theory and decision theory: Adaptive game playing, learning in games ...; machine learning: Aggregating strategies, using expert advice, prediction in the worst case ...; statistics: Prediction in the worst case, universal calibration, prophet inequalities ...; computer science: Competitive analysis of Online Algorithms. The goal of the workshop was to bring together researchers from these different areas to identify common questions as well as explore different perspectives on the same questions. The talks were (by design) mostly survey talks of various areas with a minority of talks on more focused or specialized issues. We have not created a printed proceedings for this workshop; instead, a web page which summarizes the meeting was created and is accessible at http://www.research.att.com/~yoav/DIMACS.workshop/index.html. Rather than offer a talk by talk description of what transpired it might be more useful to provide an overview of what seem to be the important issues. The basic problem that everyone was focused on is what we call "the online decision problem." This problem considers a scenario in which an agent performs some kind of repeated interaction with its environment. On each iteration of this interaction the agent chooses one out of many actions whose consequences/costs depend on the unknown environment. The action must be chosen *before* its cost is known but can depend on the outcomes of previous actions. Unlike other frameworks for studying this problem, we do *not* assume that the environment is an oblivious stochastic process. Rather, we allow the environment's behaviour to be arbitrarily complex; in particular, it can depend on the actions of the agent. There is general agreement that 'good' in this context means relative to some benchmark collection of procedures, but differences arise as to how the benchmarks should be chosen. These differences stem, in large part, from the context in which the online decision problem arises. To illustrate, we summarize a debate that took place at the workshop. In several fields it is common to use a benchmark based on "regret". A list of procedures (also called "experts", "models" or "strategies") is given a-priori. One does not know in advance which procedure will generate the lowest time average cost. However, one *can* describe adaptive methods that are guaranteed to do as well as the best procedure in hindsight. That is, one can, *after* using the adaptive method, go back and calculate the total loss (or gain) of each expert (by itself) against the sequence of outcomes one faced in the past. The regret is defined as the difference between the total gain of the adaptive algorithm and the total gain of the best procedure. (It is called regret because it quantifies the feeling of regret one suffer when one considers their total loss had they followed the recommendations of the best procedure). Methods developed in various fields have provable a-priori guarantees on the maximal value of the regret. However, such guarantees have no meaning in some contexts. The assumption implicit in the guarantee is that nature does not respond to the choice of procedure used. Specifically one assumes that had one used a procedure different from the one used initially, the sequence of outcomes stays the same. This does not make sense in a game theory context. Regret-type benchmarks have been popular in both computer science and game theory and a number of tools have been developed to analyze them. Two that were highlighted in the talks were Bregmann distances and Blackwell's approachability theorem. A tantalizing possibility is that Bregmann distances are related to the potentials associated with Blackwell's Approachability theorem. A number of the more focused talks discussed applications of these tools as well as generalizations. When one moves beyond regret-type measures of performance to measures that assume that nature chooses the outcomes as a function of the procedure one uses, guarantees appear impossible to come by. This was a point made in a series of talks on impossibility theorems. More importantly they highlight an important weakness of the Bayesian approach to decision making. Interestingly, this was a topic that that was discussed in a minor way among the participants and continued via email after the workshop.

- **Scaling Phenomena in Communication Networks**
  - Dates: October 22-24, 1999
  - Location: Institute for Mathematics and its Applications (IMA), University of Minnesota
  - Attendance: 90
  - Organizers: Ashok Erramilli, Netmetrix Inc., Vern Paxson, Lawrence Berkeley National Laboratory, Iraj Saniee, Lucent Technologies, Walter Willinger, AT&T Labs - Research
  - Co-sponsor: Institute for Mathematics and its Applications

The study of scaling phenomena in modern communication networks is another realization of Mandelbrot's vision of order in physical and social phenomena that are characterized by scaling laws. This exciting new multi-disciplinary field has attracted the attention of researchers from networking, mathematicians with interest in fractal geometry, physicists experienced in dealing with scaling laws, and computer scientists, economists and control theorists concerned with robustness and scaling issues associated with complex large-scale interacting systems. Moreover, developments in this field have been accompanied by the availability of extended, high quality data sets of network traffic measurements that are unprecedented in other disciplines. This 3-day workshop was intended to bring together the leading researchers in this emerging area, representing its various constituencies. Its main objective was to foster the exchange of ideas between leading networking experts and researchers in other fields, matching problem areas with solution methods. The workshop was structured around three fundamental aspects of the study of scaling phenomena in networks: description (e.g., empirical evidence, physical understanding, dynamical systems); analysis (e.g., network performance with fractal traffic flows, renormalization group techniques for large scale distributed systems, mean-field theory approaches for full-service networks); and control (e.g., self-organization; pattern formation, evolution and adaptation in spatially extended non-equilibrium systems). Participants were expected to contribute to this effort by giving a talk and/or actively engaging in the proceedings. In contrast to many other fields where scaling phenomena have had a long history but have not moved beyond the descriptive stage, this area shows great potential to apply the theory to analyze and control complex, large-scale networks such as the Internet. The workshop was intended to advance the study of scaling phenomena in networks from a descriptive theory to a prescriptive reality. A high point of the meeting was an IMA Public Lecture at the end of the first day by Benoit Mandelbrot: "Fractals/Multifractals in Finance, the Internet and Other 'Wild' Aspects of Man's Works." This talk (preceded by a reception for workshop participants) attracted an outside audience of about 300, including high school students and teachers and college students. The format of the workshop involved one-hour and half-hour talks interspersed with discussion, summary panel discussions that pointed out the importance of modeling and scaling issues for the future development of communications networks, and time off for informal interaction. In the first session on Networks as Large-Scale Complex Systems, Vern Paxson spoke about why analyzing the Internet is painfully hard, Mark Handley about architectural principles of the Internet, and John Doyle about HOT power laws in Networks and other complex systems. A second session on the same topic featured talks by Sally Floyd on the evolution of end-to-end congestion control in the Internet, Andrew Odlyzko on economics and its implications for the Internet, and Steven Strogatz on complex networks from the point of view of nonlinear dynamics. There was a session on scaling phenomena in the Internet, in which Scott Shenker talked about asymptotics in the Internet, Mark Crovella about self-similarity and powers laws in the web, and Don Towsley on network tomography. This was followed by a session on scaling phenomena in other branches of science. Here, Ed Waymire spoke about statistical estimation of multiscaling exponents in multiplicative cascades, Larry Liebovitch about scaling phenomena in biological and medical systems, Donald Turcotte on mechanisms for generating fractional noises and walks, and Didier Sornette on generic mechanism for hierarchies. In the session on Scaling Analysis and Wavelets, Patrice Abry, Patrick Flandrin, and Darryl Veitch presented a talk on wavelet tools for scale invariance, Anna Gilbert spoke about network inferences from scaling analysis, and Jim Roberts about flow aware networking for effective quality of service control. In the session on Network Performance Analysis and Scaling, Robert Savit gave a talk with the title "That Place is so Crowded, No one Goes there Anymore," which dealt with adaptive competition for scarce resources. Then Onuttom Narayan spoke about scaling and network performance. There was also a work in progress session for students and new researchers. The full program schedule, copies of the transparencies, PDF files and Power Point presentations for many of the talks, and the audio recording of the IMA Public Lecture are available online at http://www.ima.umn.edu/reactive/fall/networks.html The IMA provided office space, computer accounts and secretarial support to all registered participants. The event proved so popular that registration had to be closed at 90 persons in order to preserve the "workshop nature" of the meeting and maximize informal interaction.

Selected Research Results

Estimating Distances in Internet Applications Yuval Shavitt (PM, BL), Xiadong Sun (GS, RU), Avishai Wool (PM, BL) and Bulent Yener (PM, BL) have found a new algorithm for distance estimation in Internet applications. Distance estimation is important to a WWW client that needs to select a server among several potential candidates. To aid in estimating distances in the Internet, past research suggested placing tracer stations in key locations to measure distances among themselves and supply client applications with an approximated map of the Internet. The new algorithm extracts more data from the same measurements, namely, obtaining good estimates to locations where tracers are not placed, obtaining data from the same measurements as if the number of tracers were several-fold larger, and obtaining good estimates for distance between locations where we cannot place tracers. This is extremely important when unidirectional measurements are conducted using specialized equipment such as GPS that cannot be placed everywhere.

Special foci at DIMACS cross-fertilize each other. The following two results connect up our special years on Networks and Massive Data Sets.

Maximum Clique Problems in Very Large Graphs The proliferation of massive data sets brings with it a series of special computational challenges. Many of these data sets can be modeled as very large multidigraphs with a special set of edge attributes that represent special characteristics of the application at hand. Understanding the structure of the underlying digraph is essential for storage organization and information retrieval. Such problems arise for example from telephone billing records. One builds a multidigraph with vertices being telephone numbers and edges calls made from one number to another. James Abello (PM, ATTL) sudied the evolution of such a graph over one 20-day period and found that it grew to have 290 million vertices and 4 billion edges. In practice, this multidigraph is treated as if it were undirected and multiple edges are collapsed into a single edge. Even a computer with six gigabytes of main memory cannot hold the full call graphs arising in practice. Cliques in the call graph represent important structural information. Maximum cliques are important in such practical applications as project selection, classification theory, fault tolerance, coding theory, computer vision, economics, information retrieval, signal transmission theory, and molecular biology. Yet, identifying a maximum clique in a graph is computationally difficult even in a graph of modest size. Abello and collaborators Panos Pardalos (V), Mauricio Resende (PM, ATTL), and Sandra Sudarsky (GS, V) developed an approach to finding cliques in very large multidigraphs. They developed several graph decomposition schemes for breaking up the problem into manageable pieces and a semi-external greedy randomized adaptive search procedure for finding approximate solutions to the maximum clique problem and the maximum quasi-clique problem in very large sparse graphs. They experimented with their heuristic on real data, for instance a one-day call graph with 53,767,087 vertices and 170 million edges, finding large cliques of as many as 30 vertices.

Maintaining Hierarchical Graph Views In connection with a problem in visualizing large graphs, Adam Buchsbaum (PM, ATTL) and Jeff Westbrook (PM, ATTL) considered a hierarchy defined on the vertices, and how one would visualize aggregations of the graph by contracting vertices with respect to the hierarchy. For example, in the graph induced by telephone customers (vertices) and calls (edges), one can contract all the vertices in a single area code together. The problem then becomes one of maintaining this induced, contracted graph under expansion (area code into exchanges, for example) and contraction (the inverse) operations. They formalized the problem and provided data structures that allow navigation in optimal time (linear in the size of the change to the induced graph), although they use superlinear space.

Combinatorial Design of Multi-Ring Networks with Combined Routing and Flow Control Avishai Wool (PM, BL) and Bulent Yener (PM, BL) developed a novel design technique, and combined routing and flow control algorithms, for congestion-free packet switched networks. The design is based on the construction of multiple virtual rings, under the constraint that the path between any two nodes is either confined to a single ring, or traverses exactly two rings (passing through a single bridge node). Their best designs were constructed by using Finite Generalized Quadrangles of combinatorial design theory, together with a scaling algorithm for realizing networks of arbitrary size. The target topology was obtained by taking the edge union of the multiple virtual rings. Capitalizing on the underlying topological properties, they designed routing, flow and access control protocols. They proved that the proposed network architecture, coupled with their protocols, ensures that (i) no loss due to congestion occurs inside a network, under arbitrary traffic patterns; (ii) all the packets reach their destinations within bounded time; and (iii) the bandwidth is allocated fairly and no host is starved. Moreover, they achieved these desirable properties much more efficiently than earlier proposals.

Dynamic Behavior of Differential Pricing and Quality of Service Options for the Internet The simple model on which the Internet has operated, with all packets treated equally, and charges only for access links to the network, has contributed to its explosive growth. However, there is wide dissatisfaction with the delays and losses in current transmission. Further, new services such as packet telephony require assurance of considerably better service. These factors have stimulated the development of methods for providing Quality of Service (QoS), and this will make the Internet more complicated. Differential quality will also force differential pricing, and this will further increase the complexity of the system. The solution of simply putting in more capacity is widely regarded as impractical. However, it appears that we are about to enter a period of rapidly declining transmission costs. The implications of such an environment were explored by Peter Fishburn (PM, ATTL) and Andrew Odlyzko (PM, ATTL), considering models with two types of demands for data transport, differing in sensitivity to congestion. Three network configurations were considered: (1) with separate networks for the two types of traffic, (2) with a single network that provides uniformly high QoS, and (3) with a single physical network that provides differential QoS. The best solution depends on the assumptions made about demand and technological progress. However, they showed that the provision of uniformly high QoS to all traffic may well be best in the long run. Even when it is not the least expensive, the additional costs it imposes are usually not large. Their results concluded that in a dynamic environment of rapid growth in traffic and decreasing prices, these costs may well be worth paying to attain the simplicity of a single network that treats all packets equally and has a simple charging mechanism.

Orientations of Annular Networks and Communication in Interconnection Networks Orientations of undirected graphs arise in a variety of applications, including traffic flow in cities and communication in interconnection networks. If G is a graph, the minimal diameter among all orientations is useful in such applications. For example, gossiping (also called total exchange or all–to–all communication) in interconnection networks is the process in which initially each processor has an item of information that must be distributed to every other processor of the system. Gossiping arises in a large class of parallel computation problems, such as linear system solving, matrix manipulation, and sorting, where both input and output data are required to be distributed across the network. The process is accomplished by means of a sequence of *synchronous calls* between processors. Furthermore we suppose that during each call a processor can communicate with all its neighbors. If both adjacent processors can communicate, this communication model is usually referred to as *Full–Duplex* Δ*–Port* (F_*). In that case the gossiping time (minimum number of steps needed to complete the protocol) is equal to the diameter of the graph G representing the interconnection network. If we suppose that simultaneous exchange of a message on the same link is not authorized, we have another popular communication model called *Half–Duplex* Δ*–Port* (H_*). In that case, for a given orientation, we can consider the greedy protocol which at each step sends all the information known by a vertex to all of its out-neighbors according to the orientation; so we have the diameter of the orientation as an upper bound of the gossiping time. Since a lower bound is the diameter of G, for any network G that admits an orientation whose diameter is close to the diameter of G, the performance of the greedy protocol is provably close to an optimal performance. Motivated by this type of application, Jean-Claude Bermond (V), Johny Bond, Carole Martin, Aleksandar Pekec (V), and Fred Roberts (PM, RU) studied annular networks, graphs that can be represented as a 2-dimensional grid consisting of concentric circles around the center and straight lines crossing all the circles. They found bounds on the diameter of orientations of such networks and orientations that show that in most cases, these bounds are tight.

Deadlock-free Routing in Networks Jie Wu and Li Sheng (V) proposed a deadlock-free routing in irregular networks using prefix routing. Prefix routing is a special type of routing with a compact routing table associated with each node (processor). Basically, each outgoing channel of a node is assigned a special label and an outgoing channel is selected if its label is a prefix of the label of the destination node. Node and channel labeling in an irregular network is done through constructing a spanning tree. The routing process follows a two-phase process of going up and then down along the spanning tree, with a possible cross channel (shortcut) between two branches of the tree between two phases. Wu and Sheng showed that the proposed routing scheme is deadlock- and livelock- free. They also studied a broad class of deadlock-free wormhole routing algorithms for interconnection networks and found a simpler proof of a lower bound of Libeskind-Hadas on the number of channels required by such routing algorithms.

Characterizations of Consistent Marked Graphs A *marked graph* is a graph with a + sign or a - sign on each vertex and it is called *consistent* if each cycle has an even number of - signs. This concept is motivated by problems of communication networks and social networks. For example, a communication network is consistent if whenever nodes x and y have the same sign, then if a binary message originating at x goes to y by two node-disjoint paths and if it is reversed at - nodes, the same message will arrive at y no matter which of the two paths is used. Fred Roberts (PM, RU) and Shaoji Xu (GS, RU) found characterizations of consistent marked graphs and quadratic algorithms to recognize such graphs.

An Active Network Approach to Efficient Network Management Active networks are a framework where network elements, primarily routers and switches, are programmable. Programs that are injected into the network are executed by the network elements to achieve higher flexibility and to present new capabilities. Danny Raz (PM, BL) and Yuval Shavitt (PM, BL) described a novel active network architecture which primarily addresses the management challenges of modern complex networks. Its primary component is an active engine that is attached to any IP router to form an active node. The active engine, designed and implemented, executes programs that arrive from the network and monitors and controls the router actions. The design is based on standards (Java, SNMP, ANEP over UDP) and can be easily deployed in today's IP networks. The contribution of Raz and Shavitt was the introduction of novel architectural features such as: isolation of the active mechanism, the session concept, the ability of active session to control non-active packets, and blind addressing. Implementing these ideas, Raz and Shavitt built a system that enables the safe execution and rapid deployment of new distributed management applications in the network layer. This system can be gradually integrated in today's IP networks, and allows smooth migration from IP to active networking.

The following result was obtained as part of our project on simulation of the global Internet described in Section **??**.

Efficient Radio Wave Propagation for Wireless Radio Signal Prediction Due to the ever increasing complexity of wireless network design, wireless network simulation engines have become a necessary tool for wireless network engineers. A signal-strength prediction system lies at the foundation of every wireless simulator, making signal-strength prediction for microcells and picocells an important research topic in the radio field. Efficient and accurate methods are needed for use in current wireless network design and future wireless network research. Greg Perkins (GS, RU) and Diane Souvaine (PM, RU) developed a system, PREDICT, designed for use in wireless radio signal prediction research. PREDICT tracks the expansion of a transmitter's wavefront by modeling the wavefront as an ever-expanding sphere centered upon the transmitter. The sphere's expansion is tracked by discretizing the sphere's surface with a triangulation and then propagating each triangular wavefront through the environment. In the propagation prediction of these triangular wave fronts, PREDICT supports reflection, transmission and diffraction based upon the geometric optics (GO) model of electromagnetic waves. Signal prediction is performed through a special summation of the predicted multi-path at a receiver. PREDICT's algorithms were specifically designed for use in radio signal prediction. The current environments supported by PREDICT are urban with flat or varied terrain and a single floor of a building, where all obstacles in the environment have vertical or horizontal walls. The system design is flexible, allowing PREDICT to be expanded for use in many different environments. Modularity allows for easy modification to the GO model or implementation of alternate radio modeling methods (one can easily modify PREDICT, or use the results generated by PREDICT, with one's own radio wave modeling methods). PREDICT is publicly available for any non-commercial use.

Average-Case Analysis of the Sum-of-Squares Bin Packing Algorithm In a "discrete distribution" we are given a bin capacity B and a probability distribution over the item sizes from 1 to B-1. Such distributions model many real-world applications of bin packing, where our goal is typically to minimize the total number of bins used or, equivalently, to minimize the wasted space in the non-empty bins of the packing. If N is the number of items generated, it is known that for any discrete distribution, the total expected waste in the optimal packing must grow either as $O(N), O(\sqrt{N})$, or $O(1)$, depending on the distribution. Standard bin packing algorithms such as Best Fit can be far from optimum, for instance yielding linear waste under distributions where the optimal expected waste is $O(1)$. This is true even for offline algorithms such as Best Fit Decreasing. Janos Csirik (V), David Johnson (PM, ATTL), Claire Kenyon, Jim Orlin, Peter Shor (PM, ATTL), and Richard Weber described a remarkably simple $O(nB)$-time online algorithm, the "Sum-of-Squares" algorithm (SS), for which the expected wasted space is provably $O(\sqrt{N}$ whenever the optimal expected waste is sublinear. Based on this, they found an $O(nBlogB)$ algorithm that essentially matches the optimal expected waste growth rate for all discrete distributions. The asymptotic worst-case ratio for SS lies somewhere between 1.54 and 3, but SS can be incorporated into a randomized offline $O(nBlogB)$ algorithm whose worst-case expected ratio is asymptotically 1.

Low Discrepancy Sets Yield Approximate Min-Wise Independent Permutation Families Motivated by a problem of filtering near-duplicate Web documents, Broder, Charikar, Frieze and Mitzenmacher defined the following notion of *ϵ-approximate min-wise independent permutation families.* A multiset $\mathcal{F}$ of permutations of $\{0, 1, \ldots, n-1\}$ is such a family if for all $K \subseteq \{0, 1, \ldots, n-1\}$ and any $x \in K$, a permutation π chosen uniformly at random from $\mathcal{F}$ satisfies

$$|\ \Pr[\min\{\pi(K)\} = \pi(x)] - \frac{1}{|K|}\ | \leq \frac{\epsilon}{|K|}.$$

Michael Saks (PM, RU), Aravind Srinivasan, Shiyu Zhou (PM, BL) and David Zuckerman (V) showed connections of such families with *low discrepancy sets for geometric rectangles*, and gave explicit constructions of such families $\mathcal{F}$ of size $n^{O(\sqrt{\log n})}$ for $\epsilon = 1/n^{\Theta(1)}$, improving upon the previously best-known bound of Indyk. They also presented polynomial-size constructions when the min-wise condition is required only for $|K| \leq 2^{O(\log^{2/3} n)}$, with $\epsilon \geq 2^{-O(\log^{2/3} n)}$.

A Minimax Game of Power Control in a Wireless Network under Incomplete Information Tiina Heikkinen (GS, RU) studied optimal transmit power control and capacity of a wireless multimedia network, where there is incomplete information about the link gain coefficients. Previously, Perron-Frobenius theory had been applied to the study of optimal power control under complete information about the link gain matrix and the interference matrix. Heikkinen derived a characterization of the optimal power vector and capacity in terms of an incomplete-information-matrix-game against nature choosing the elements of the link gain and interference matrices. A game-theoretical rationale for not using power control was obtained.

High Performance Distributed Protocols on a Cluster of PCs Distributed Shared Memory (DSM) protocols involve a lot of communication and this communication often proves to be a bottleneck in the performance of such protocols due to decreased processor utilization. Liviu Iftode (RU) and Murali Rangarajan (GS, RU) developed a 2-pronged approach to improve the performance: Hide the communication latency by using multi-threading, and further improve the performance by using appropriate scheduling policies; evaluate the use of a co-processor on the network interface to perform protocol actions, thus freeing the host processor from being interrupted to perform these actions. They did this by using one of the processors in an SMP machine to emulate the coprocessor in the network interface, freeing the other processor from being interrupted for protocol actions. Iftode and Rangarajan developed a multi-threaded implementation of Home-based Lazy Release Consistency (HLRC) protocol, which is a DSM protocol. With an 8 node cluster of SMPs (two PII 300 MHz processors) connected by a high speed Myrinet (Gigabit) network, they obtained results to indicate performance gains from multi-threading by a significant reduction in the idle time, leading to increased processor utilization.

EmailValet: Learning User Preferences for Wireless Email Sofus Macskassy, Aynur Dayanik (GS RU), and Haym Hirsh (PM, RU) studied EmailValet, a system developed at RU that learns users' email-reading preferences on email-capable wireless platforms – specifically, on two-way pagers with small "qwerty" keyboards and an 8-line 30-character display. They gathered data on email-reading preferences over more than 8900 email messages received by their people for about three months. They compared the ability of different learning methods to form models that can predict whether a given message should be forwarded to the user's wireless device. Their results showed that the best performance of one method, over a range of established learning methods developed by the information retrieval and machine learning communities, was able to achieve a break-even point of over 53% for one user that had received over 5000 messages. They also found that, in general, all methods are able to achieve better performance than what would be achieved by a baseline of simply forwarding all messages to the wireless device, and that many methods are able to find procedures that, although they forward only a small fraction of the messages that a user would want, achieve 100% precision on those messages that they do actually choose to forward.

Exploiting Regularities in Web Traffic Patterns for Cache Replacement Caching web pages at proxies and in web servers' memories can greatly enhance performance. Proxy caching is known to reduce network load and both proxy and server caching can significantly decrease latency. Web caching problems have different properties than traditional operating systems paging, and cache replacement can benefit by recognizing and exploiting these differences. Edith Cohen (PM, ATTL) and Haim Kaplan (PD) addressed two aspects of the predictability of traffic patterns: the overall load experienced by large proxy and web servers, and the distinct access patterns of individual pages. They formalized the notion of "cache load" under various replacement policies and demonstrated that the trace of a large proxy server exhibits regular load. Predictable load allows for improved design, analysis, and experimental evaluation of replacement policies. They provided a simple and (near)-optimal replacement policy when each page request has an associated distribution function on the next request time for the page. Without the predictable load assumption, no such online policy is possible and it is known that even obtaining an offline optimum is hard. For experiments, predictable load enables comparing and evaluating cache replacement policies using *partial traces*, containing requests made to only a subset of the pages. Their results are based on considering a simpler *interval caching* model for cache replacement. They related traditional and interval-caching policies under predictable load, and derived (near)-optimal replacement policies from their optimal interval-caching counterparts.

Connection Caching Communication between clients and servers in the Web is performed using TCP (Transmission Control Protocol). TCP first opens a connection between the two nodes and then sends data packets via this open connection. The time needed to open a connection contributes to the latency of servicing a request. Hence, latency can be reduced if open connections are kept idle and re-used for future requests. Indeed, new versions of HTTP support that. Web servers and proxies typically maintain many open connections, many of them idle. The number of connections each node can keep open is limited. Hence, connection management is increasingly perceived and handled as a caching problem. A fundamental difference between *connection-caching* and standard caching is that connection caching involves multiple caches that affect each other. Each open connection is cached at two nodes. If one of these nodes decides to terminate the connection, the other node is forced to accept this decision. Edith Cohen (PM, ATTL), Haim Kaplan (PD), and Uri Zwick proposed a theoretical model for connection caching and studied the connection-caching problem in this model. While the ordinary off-line caching problem is easily solved in polynomial time, they showed that the off-line connection caching problem is NP-complete, and in fact also APX-complete, i.e., it is even hard to approximate. They complemented this result by showing that a distributed version of Belady's algorithm, which is optimal for a single cache, is a 2-approximation algorithm for off-line connection-caching. They provided examples that show that the analysis of this algorithm is tight.

Prefetching the Means for Document Transfer: A New Approach for Reducing Web Latency User-perceived latency is recognized as the central performance problem in the Web. Haim Kaplan (PD) and Edith Cohen (PM, ATTL) systematically measured factors contributing to this latency, across several locations. Their study revealed that DNS query times, TCP connection establishment, and start-of-session delays at HTTP servers, more so than transmission time, are major causes of long waits. Wait due to these factors also afflicts high-bandwidth users and has detrimental effect on perceived performance. They proposed simple techniques that address these factors: (i) *pre-resolving* host-names (pre-performing DNS lookup), (ii) *pre-connecting* (prefetching TCP connections prior to issuance of HTTP request), and (iii) *pre-warming* (sending a "dummy" HTTP `HEAD` request to Web servers). Trace-based simulations demonstrated a potential to reduce perceived latency dramatically. Their techniques surpassed document prefetching in performance improvement per bandwidth used and can be used with non-prefetchable URLs. Deployment of these techniques at Web browsers or proxies does not require protocol modifications or the cooperation of other entities. Applicable servers can be identified, for example, by analyzing hyperlinks. Bandwidth overhead is minimal, and so is processing overhead at the user's browser. Kaplan and Cohen proposed scalable deployment solutions to control the potential overhead to proxies and particularly to Web servers.

The following six results connect up two special years, Large Scale Discrete Optimization and Networks.

The Element Connectivity Problem in Survivable Network Design The element connectivity problem falls in the category of survivable network design problems – it is intermediate to the versions that ask for edge-disjoint and vertex-disjoint paths. The edge version is by now well understood from the view-point of approximation algorithms, but very little is known about the vertex version. Vijay Vazirani (V) considered the problem where vertices are partitioned into two sets: terminals and non-terminals. Only edges and non-terminals can fail – we refer to them as *elements* – and only pairs of terminals have connectivity requirements, specifying the number of element-disjoint paths required. Vazirani found an algorithm that achieves an approximation guarantee of factor $2H_k$, where k is the largest requirement and $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$. Besides providing possible insights for solving the vertex-disjoint paths version, the element connectivity problem is of independent interest, since it models a realistic situation.

Primitive Ring Generation and Ring Grooming in Design of Optimal Networks A central problem in the design of optical networks is the placement of the ADM (Add/Drop Multiplexes) devices. Many restricted versions of this general design problem have been studied in the literature. The general version can be formulated as a discrete optimization problem. Due to the complexity of the problem, the current methodology recommends dividing the general version of the design problem into two phases, called Primitive Ring Generation and Ring Grooming. The General Design Problem, Primitive Ring Generation, and Ring Grooming are NP-Hard. We have obtained theoretical results on the quality of approximation algorithms. Primitive Ring Generation has two variants, depending on whether or not traffic demands can be split. Gruia Calinescu (PD) and Peng-Jun Wan designed new algorithms for Primitive Ring Generation guaranteed to produce an output of cost at most 25% bigger than optimum if traffic demands can be split, and at most 50% bigger than optimum if traffic demands cannot be split. All previously known algorithms either have bigger worst-case guarantees or are provably not better than the new algorithms. For Ring Grooming, the traffic granularity parameter seems to influence the quality of the approximation algorithms. Gruia Calinescu (PD), Li-Wu Liu and Peng-Jun Wan provided a nontrivial analysis of a natural algorithm, proving performance guarantees of 1.5 and 2.5 respectively, for the important case when traffic granularity is 4 and 8 respectively.

Directed Multicut vs. Multicommodity Flow Given a graph $G = (V, E)$ and k pairs of vertices $(s_1, t_1), \ldots, (s_k, t_k)$, a *multicut* is a subset of edges $F \subseteq E$ such that if all edges in F are deleted, none of the pairs (s_i, t_i) are in the same connected component in the remaining graph $G' = (V, E \setminus F)$. In the minimum multicut problem, we are also given a positive cost $c(e)$ for each edge $e \in E$ and we wish to find the multicut of minimal total cost. This problem is known to be NP-complete for $k \geq 3$, and therefore it is interesting to consider approximation algorithms for it. One way to do it is to state the multicut problem as an integer programming problem, relax it to a linear programming problem, and solve this problem or it's dual - the *maximum multicommodity flow problem.* The maximal multicommodity flow will always be at most as big as cost of the minimal multicut. On the other hand, it is easy to see that there are multicuts which cost at most k times the value of the maximal flow. Thus, the described approach gives a factor-k approximation algorithm for a minimal multicut. It turns out that when the underlying graph is *undirected* more can be said, and in fact the cut/flow ratio in this case can be at most as large as $O(\log(k))$. On the positive side, an algorithm of Garg, Vazirani and Yannakakis exists, which takes a maximal flow and constructs a multicut which costs at most as much as $O(\log(k))$ times the flow cost. Michael Saks (PM, RU), Alex Samorodnitsky (PD), and Leonid Zosin (PD) showed that in the more general *directed* case, the cut/flow ratio could be as large as the trivial upper bound of k. It seems, therefore, that to find a better than a factor-k approximation, one must look for other approaches than the simple linear relaxation technique.

A Unified Approach to Approximating Resource Allocation and Scheduling Amotz Bar-Noy (ATTL), Reuven Bar-Yehuda (V), Ari Freund, Joseph (Seffi) Naor, and Baruch Schieber developed a general framework for resource allocation and scheduling problems. Given a resource of fixed size, they developed algorithms that approximate the maximum throughput or the minimum loss by a constant factor. Their approximation factors apply to many prominent problems, among others: (i) Real-time scheduling of jobs on parallel machines. (ii) Bandwidth allocation for sessions between two endpoints. (iii) General caching. (iv) Dynamic storage allocation. (v) Bandwidth allocation on optical line and ring topologies. For many of the problems, this is the first constant factor approximation algorithm. The algorithms are extremely simple and efficient. They use the local-ratio technique but they can also be equivalently interpreted within the primal-dual schema.

Load Balancing in a Distributed Computing Environment Endre Boros (PM, RU) and Ewald Speckenmeyer (V) studied load balancing in a distributed computing environment. They showed that the minimum message length to re-balance the load of n processors can be computed in linear $O(n)$ time, even though the actual balancing may require $O(n^2)$ messages exchanged between the processors.

A Bottleneck Load Balancing Minimization Problem Consider a processor network with n overloaded and m underloaded processors. The overloaded and the underloaded nodes are connected by a completely bipartite network $K(n, m)$. Ewald Speckenmeyer (V) considered the problem of achieving a uniform load-distribution among all processors via sending appropriate amount of load from overloaded to underloaded nodes so that the maximum of load sent via a link is minimized. He showed that this problem can be solved in time $O(mn)$.

The following three results are joint with the Special Focus on Computational Intractability.

Randomizing Low-Degree Polynomials and Round-Efficient Secure Computation Yuval Ishai (PD) and Eyal Kushilevitz introduced a notion of *randomizing low-degree polynomials* as a powerful and conceptually simple paradigm for obtaining round-efficient security. "Standard" low-degree multivariate polynomials can be securely evaluated in a small number of rounds by virtually any "off-the-shelf" secure multi-party computation protocol. However, most Boolean functions f on n input bits cannot be evaluated by a polynomial whose degree is smaller than n. Ishai and Kushilevitz got around this barrier by relaxing the requirement of evaluating f into a weaker requirement of *randomizing* f; that is, mapping the inputs for f along with independent random inputs into several outputs, whose *joint distribution* depends only on the value of f. They showed that degree-3 polynomials are sufficient to randomize *any* function f, where the number of outputs is at most quadratic in the branching program size of f, and where both input variables and random variables count towards the degree. As an immediate application, the secure evaluation of any function f reduces to a secure evaluation of degree-3 polynomials; this reduction is applicable in any reasonable setting for secure computation. One corollary is that two (respectively, three) communication rounds are sufficient to compute any n-argument function f with perfect information-theoretic $\lfloor \frac{n-1}{3} \rfloor$-privacy (resp., $\lfloor \frac{n-1}{2} \rfloor$-privacy), probabilistic correctness, and communication complexity which is at most quadratic in the branching program size of f. This improves over a previous constant-round protocol of Bar-Ilan and Beaver, which in a similar setting requires more communication rounds and a higher communication complexity.

Adaptive vs. Non-Adaptive Security In a *non-adaptive* attack against a cryptographic protocol, the adversary must choose in advance the set of parties it corrupts. A stronger type of attack, called an *adaptive* attack, allows the adversary to dynamically choose which parties to corrupt as the protocol proceeds. R. Canetti, I. Damgard, Dziembowski, Y. Ishai (PD), and T. Malkin clarified the relation between adaptive and non-adaptive security. They showed that under some previously studied settings for secure computation, such as perfect security against a passive adversary or a restricted type of an active adversary, any protocol which is non-adaptively secure must also be adaptively secure. On the other hand, they provided provable separations between non-adaptive and adaptive security in other settings.

Correlated Pseudorandomness and Compressing Cryptographic Resources A private-key cryptosystem may be viewed as a means by which a trusted dealer privately conveys a large, shared pseudorandom object to a pair of players, using little communication. Alternatively, the messages distributed by the dealer may be viewed as a secure *compression* of a pair of large identical random pads (or random functions) into a shorter shared "key" or "seed". Yuval Ishai (PD) extended this compression problem to more general correlation patterns among *several* players. Unlike the simple case of identical pads, where the main security concern is with respect to external eavesdroppers, in the case of general correlations participants also have to be protected from each other. That is, collusions of computationally-bounded players should gain no additional knowledge about the joint pads of the remaining players from the compressed messages they receive, other than what follows from the pads they generate and from knowing the joint distribution of all pads. Ishai developed a modular replication-based approach for securely compressing any *linear* correlation pattern, in which the random pads jointly form a random vector in a linear space. Specifically, the compressed messages received by the players consist of *independent* random seeds, where each seed may be replicated among several players. This approach allows one to use pseudo-random generators or pseudo-random functions in a black-box manner.

Mixing in Small Batches Markus Jakobsson (PM, BL) and Ari Juels developed Millimix, a mix network that is highly efficient on small input batches. The construction is conceptually simple, and both robust and private in the face of collusion by any minority set of malicious players. Additionally, Millimix has the feature of being publicly verifiable. In other words, the mixing operation yields a transcript that demonstrates correctness to a third party that the mix proceeded correctly.

Approximate Solution to the Directed Steiner Tree Problem The directed Steiner tree problem is the following: given a directed graph $G = (V, E)$ with weights on the edges, a set of terminals $S \subseteq V$, and a root vertex r, find a minimum weight branching T rooted at r, such that all vertices in S are included in T. This problem is known to be NP-hard. There are many papers that give approximation algorithms for this problem starting with the work by Zelikovsky, and more recently by Charikar, Chekuri, Cheung, Dai, Goel, Guha and Li. In these papers a polynomial time algorithm is developed with an approximation ratio of $O(k^\epsilon)$ for any fixed $\epsilon > 0$. This is extended to a poly-log approximation algorithm, but the running time is not polynomial. Samir Khuller and Leonid Zosin (PD) proved that the gap between a standard fractional formulation of this problem and the integral minimum Steiner tree is $O(\sqrt{k})$. Seffi Naor and Zosin found a new fractional formulation for this problem and proved that the gap for this formulation is at most $O(\log^3 k)$.

Testing Shared Memories Sequential consistency is the most widely used correctness condition for multiprocessor memory systems. Phil Gibbons (PM, BL) and Ephraim Korach (V) studied the problem of testing shared-memory multiprocessors to determine if they are indeed providing a sequentially consistent memory. Their work represented the first formal study of this problem, which has applications to testing new memory system designs and realizations, providing run-time fault tolerance, and detecting bugs in parallel programs. A series of results were obtained for testing an execution of a shared memory under various scenarios, comparing sequential consistency with linearizability, another well-known correctness condition. Linearizability imposes additional restrictions on the shared memory, beyond that of sequential consistency; these restrictions were shown to be useful in testing such memories. (Gibbons and Korach received the Yossi Levy Prize in Operations Research and Managerial Economics for the year 1999 from the Operations Research Society of Israel for this work.)

Placing Web Proxies on Tree-Broadcasting Networks Mordecai Golin (V) and Lixin Gao (V) designed an algorithm for placing a given number of proxies (copies) of a web page on nodes in a tree-topology network so as to minimize the average time any user would need to access a copy of the page. In graph-theoretic terms, this is a k-median problem.

New Upper and Lower Bounds on the Channel Capacity of Read/Write Isolated Memory In general, we may regard the set of all permissible binary memory configurations as a channel alphabet which obeys some restrictions, where, for each alphabet symbol, the rewriting restrictions (if they exist) dictate which symbols may succeed it. In this way, the rewritable memory can be viewed (over its lifetime) as a noiseless communication channel. Shannon defined the capacity of memory, in bits per rewrite, as $C_k = \lim_{r\to\infty} \frac{1}{r} \log_2 N(k, r)$, where k is the size of the memory in binary symbols, r is the liftime of the memory in rewrite cycles, and $N(k, r)$ is the number of distinct sequences of r characters. Shannon showed that $C_k = \log_2 \lambda_k$ where λ_k is the largest eigenvalue of the channel graph of order k, i.e., of its associated adjacency matrix. Xuerong Yong (GS, V), Yuanping Zhang (GS, V), Mordecai Golin (V), and Li Sheng (V) found some new recursive properties of λ_k and new lower and upper bounds on C_k.

Watermarking Maps: Hiding Information in Structured Data The storage of data in digital form makes reproduction of valuable data, and thus piracy, easier than ever. Watermarking techniques address this problem by embedding hidden information into the data, which either asserts a copyright or allows tracing of copies. Sanjeev Khanna (PM, BL) and Francis Zane (PD) examined the problem of protecting map data from piracy. Consider the following scenario: MapQuest commissions resources to produce an accurate map of the US, and sells modified, or watermarked, copies to web sites like Yahoo!. Each web site then uses this map to answer queries (e.g, what is the shortest route from A to B?). If a new map web site appears, MapQuest wants to know two things: Is this site using a pirated MapQuest map? If so, which copy is it using and who purchased that copy? To do this, they hide some information in each copy of the map (say, small changes in the distances), allowing them to distinguish the copies. However, these small changes can easily accumulate over long paths to yield unacceptably distorted distances. Khanna and Zane presented a method for making these changes which encodes the identities of many users while introducing a very small distortion. In addition, a web site with a pirated copy will naturally attempt to disguise its copy with misleading answers to queries. Under some natural assumptions (e.g., the pirate's answers cannot distort the graph too much), they presented a scheme which detects even such evasive pirates with high probability.

Associated Projects on Network Modeling and Parallel Simulations of Networks

In addition to the research work arising from its Networks special year, DIMACS has run two multi-disciplinary, multi-institutional research projects in network modeling and parallel simulations, each funded separately, but closely related to our Special Year in Networks. These projects relate to some very important problems in networking research, aiming to provide the methodology and software tools for creating software descriptions of very large networks – such as the global Internet – which could then be simulated on multiprocessor machines or on multiple networked machines, to speed up the simulations. Parallel discrete event simulations of telecommunications systems involve very difficult problems, and so far have not been successfully implemented in commercial systems despite a great desire for such tools from industry and the military.

The first of these two research efforts was funded by NSF and the second, followup project by DARPA. We describe these two projects together in this section. At Rutgers, this project involves not only faculty, postdocs, and graduate students from DIMACS, but also a strong group from WINLAB (a leading NSF/industry/university research center on Wireless Networks). In 1998, WINLAB received an NSF Instrumentation Grant for a 14-processor SUN ultrasparc that is heavily used in these simulation projects by both local and remote project members. These projects have also led to a small subcontract from the Ultra Corporation for research on a global network simulator.

Simulations of Integrated Communications Systems

This project was funded by the National Science Foundation for a three year period over 1995 - 1998 with a one year extension through 1999, for a total of $1.5M.

Participating Scientists:

The original PI team was composed of Sandeep Bhatt (PM, BC/TC and RU), Richard Fujimoto (Georgia Tech), Albert Greenberg (PM, ATTL), Jim Kurose (University of Massachusetts), David Nicol (William and Mary / Dartmouth), Andrew Ogielski (PM, BC and RU), Diane Souvaine (PM, RU), Don Towsley (University of Massachusetts), and Paul Wright (PM, ATTL). During the lifetime of the project, Andrew Ogielski left Bellcore and joined Rutgers full-time, taking the leadership of the project. Nicol moved from William and Mary to Dartmouth.

The Centerpiece of the Project:

The centerpiece of the project has been the development of and experimentation with a network-centric simulation meta-language called TeD (Telecommunications Description Language). TeD provides language-independent constructs that may be embedded in a general computer language; in practice we have used C++. The TeD compiler is a source-to-source translator that transforms TeD code into pure C++ that is compiled and executed by the GTW (Georgia Time Warp) parallel simulation engine. An initial version of TeD was developed very early in the project, thus allowing us to quickly begin the important cycle of model building and tool refinement. The official TeD home page http://www.cc.gatech.edu/computing/pads/teddoc.html contains information about TeD and instructions about how to obtain and install the core software.

Collaborations:

This project brought together experts in network modeling, design and analysis, distributed algorithms, and high performance discrete-event simulation kernels. Perhaps the most important contribution of this effort has been unfettered research and software design experimentation on most efficient, parsimonius and powerful methods to develop a user-friendly approach to harnessing parallel processing for large network simulation problems, a very difficult problem that has eluded the parallel simulation community for years. This required cross-disciplinary collaboration on implementation of a diverse suite of large and nontrivially detailed protocols and networks (traditionally not of interest in the community of parallel simulator software researchers), and very detailed parallel performance analyses of realistic network models (normally not of interest to network researchers).

Specific areas of network engineering of interest to us have included various aspects of multicast protocols, ATM networks, PNNI routing, and radio resource management in mobile wireless networks. Each of these areas developed substantial additional network modeling components and software. This project included a large number of graduate students whose educational benefits included exposure to industrial scale problems and internships.

Scalable Self-Organizing Simulations (S3)

This project was funded by DARPA for a three year period over 1996-1999, with an extension of the SSF sub-project (see below) through 2000, for a total of $3.1M. It represents a highly successful transition from the NSF-supported basic research on network modeling techniques for efficient parallel simulations to a very large network modeling software development effort.

Participating Scientists:

Andy Ogielski (PM, RU, leading PI), Narayan Mandayam (RU), Roy Yates (RU), Mark Crovella (Boston University), James Cowie (Cooperating Systems Corporation), Richard Fujimoto (Georgia Tech), Kalyan Perumalla (Georgia Tech), Ellen Zegura (Georgia Tech), Don Leskiw (Ultra Corporation), David Nicol (Dartmouth), with about thirty graduate students, postdocs, and part-time scientists at various times during the project lifetime.

Collaborations:

The complexity of modern protocols and networking technologies on the one hand, and of the large software systems on the other, demands that a project of this sort be carried out by an interdisciplinary team of experts in parallel simulations, networking, and industrial-strength software engineering. The latter in particular benefited greatly from the participation of Cooperating Systems, a small company specializing in high quality scientific software.

Objectives:

21st century global communications architectures will emerge from internetworking of millions of networks, as wireline, wireless, and satellite technologies interconnect hundreds of millions of diverse nodes. Verification of the design, operation, and management of these future architectures will rely on fast, accurate simulations. PITAC and NGI (Next Generation Internet) milestones called for a capability of simulations of 100,000 node, multiprotocol networks in year 2000, and simulations of 10,000,000 node networks in year 2001. At the beginning of this project the available network modeling tools and simulation techniques did not even approach systems of such complexity, falling short by many orders of magnitude.

In 1999 the network modeling software developed in this project, a direct descendant of the NSF-funded basic research described above, exceeded the year 2000 DARPA goal of modeling and simulations of 100,000-node multiprotocol Internet models. The project has developed and distributes a suite of detailed models of the major Internet protocols (IP, OSPF, BGP, TCP, plus sockets and assorted application level models), core modeling platforms, and other software subsystems, under the name of SSFNET (Scalable Simulation Framework). In June 1999 the SSF Research Network web site with scalable Internet modeling software and extensive documentation and tutorials went on-line at http://www.ssfnet.org.

Recent Accomplishments:

Our parallel network simulators have demonstrated execution rates and entity counts on realistic models well in excess of those achievable using the best tools available elsewhere. Using commodity multiprocessor machines (like SUN or SGI) the following results were established:

- November 1998: DaSSF simulations of the multi-domain Internet on a SUN Enterprise server deliver scalable performance for modeling networks of 100,000 multi-protocol hosts and routers with UDP Pareto-distributed traffic sources, or for multi-million-node networks of simpler elements, at the speeds of a million of IP packet hops per second.

- May 1999: DaSSF simulations on SGI Origin 2000 reached 100,000 concurrently active TCP connections (gigabit links, no losses); Java SSF simulations on SUN Enterprise servers (JDK 2) were closing in on this target. DaSSF multiprotocol Internet simulations ran faster than real time for several thousand hosts and routers.

These results exceeded the year 2000 goals of modeling and simulations of multiprotocol Internet models. By inherent, demonstrated scalability of SSF-based network models, and of the SSF parallel simulators, the SSF software can run unchanged on larger machines (as well as follow Moore's law), and reach many millions of hosts and routers in multi-domain, multiprotocol Internet models within the next couple of years or sooner.

The SSF event processing parallel kernels support high-speed interactive operation with real-time events and can be embedded in network monitoring and network management systems running on commodity multiprocessors (Linux, Solaris, SGI Irix), as well as in human-in-the-loop interactive applications.

Design and implementation of SSFNET, the first collection of SSF-based models for simulating Internet protocols and networks, was completed. Available in C++ and Java forms, the SSFNET models are open source software, distributed under the GNU General Public License. SSFNET models are self-configuring; that is, each SSFNET class instance can autonomously configure by querying a parameter database, which may be locally resident or available over the Web.

Development of a suite of core Internet protocol models for SSFNET was completed. They include IP, UDP, TCP, OSPFv2, BGP-4. Additional pseudo-protocols include network interface cards and queue management, bsd sockets, client and server applications. All protocols are new implementations, conforming to the SSF.OS "plug-and-play" self-configuring design. The protocols' conformance to IETF RFC requirements is rigorously tested.

Development of the first network design automation tools for SSFNET was completed. These include: automated IP address assignment following the CIDR guidelines; automatic verification of the network configuration syntax prior to model instantiation; verification of IP routing tables constructed by OSPF and BGP-4.

The SSF Research Network has been established, with the web site at http://www.ssfnet.org. It is an open consortium of SSF users, whose goal is to promote, develop, and validate SSF-based Internet models. This site contains substantial SSFNET documentation and tutorials, and will support automatic submission, search, and retrieval of a digital library of validated and experimental model components developed by the SSFNET community.

SSFNET 0.8, a major revision adding hierarchical Net specification, has been released.

DML 0.8, a major revision to the model configuration database code to support object-oriented attribute extension and import capabilities, has been released. Model database format was designed as a target for automated network design and validation tools.

Java SSF version 0.5 parallel simulator, a major revision supporting the full SSF 1.0 API, has been released.

Dartmouth SSF DaSSF 1.0.5 parallel simulator, a major revision supporting the full SSF 1.0 API was released.

The Surge workload generator was expanded to encompass a wider range of traffic types (HTTP 1.0 and HTTP 1.1), and both closed-loop and open-loop client-server interactions. We used an expanded workload generator to compare the performance of HTTP 1.0 vs. HTTP 1.1 over a wide range of system configurations. SURGE can be used for server benchmarking, and traffic generation both for networks and simulators.

Mobile wireless network modeling classes (WiPPET) were ported from TeD to SSF. Major new additions included: Wideband radio channel models (mobility; multipath; Rayleigh, shadow and distance fading; interference); wideband CDMA receiver signal processing (NTT DoCoMo / ARIB transceiver design and framing structure); multiple cells with antenna diversity and SIR based power control; soft hand-off protocol and the I/Q multiplexed framing structure. The granularity of these SSF models is at the baseband level (DS-CDMA chips).

We developed a new release of the TeD network description language compiler (predecessor of SSF) and the GTW parallel simulation engine, along with documentation, for public release. We developed a new release of the ATM Private Network to Network protocol (PNNI) models, along with documentation. We demonstrated near-linear speedup up to 16 processors clocking well over one million events per second, by using the novel "reverse computing" optimistic synchronization technique.

The next phase of SSF-based network modeling research and development will focus on a suite of scalable, multi-resolution tools for network design, validation, and analysis: Graphical editors for generation and validation of network topologies, scalable network data collection and multiresolution display facilities, statistical multiresolution tools for sensitivity analysis, etc.

We are planning the integration of SSF based Internet models (SSFNET) and physical/link wireless SSF models (WiPPET) for the first ever multi-scale baseband-to-IP wireless Internet modeling systems.

Technology Transition:

In August 1999, there was a public release of the SSFNET 0.8 scalable Internet modeling and simulation environment, comprising the parallel simulation kernel (JSSF), the model configuration database management system (SSF.DML), and the network hardware and Internet protocol models (SSF.Net, SSF.OS). Extensive documentation and tutorials are available on the SSFNET community web site http://www.ssfnet.org. Early adopters of SSFNET include:

- Internet traffic analysis (AT&T Labs - Research),
- Design of network security protocols and traffic engineering models (NIST)
- Design of Simulator of Air Force C4ISR aerospace operations within the Global Grid (Ultra Corporation)
- multicell GPRS (data for mobile GSM phones) models (University of Barcelona)

and a number of other projects in the USA and abroad are beginning.

In June 1999, there was a public release of Dartmouth SSF DaSSF 1.0.5 parallel simulator (all SSF simulators implement the same modeling API). Dartmouth SSF is ultra fast and demonstrably capable of simulating extremely large network models. See http://www.cs.dartmouth.edu/ jasonliu/projects/ssf/.

A portable version of Surge (web server workload generator) has been developed for public distribution. The Surge generator has been distributed to over forty sites, including:

- high-tech startups (Infolibria, Gigabit Wireless, Arrowpoint)
- telecom research labs (Lucent Bell Labs, AT&T Research, Telecom Italia, British Telecom, GTE Labs)
- industrial research labs (Xerox PARC, HP Labs, IBM T. J. Watson Labs, Tandem, Qualcomm, Motorola)
- over 20 universities, including MIT, CMU, Berkeley, UT Austin, and Cornell.

A professionally supported version of the C++ SSF parallel simulator was developed by Cooperating Systems Corp. The pre-release version was used for modeling and performance evaluation of the 3rd generation wireless networks (Wideband CDMA for IMT-2000) in a joint project of WINLAB and NTT DoCoMo. Modeling at the baseband signal level of detail ranges from multipath radio propagation with fading and interference, to radio receiver signal processing, to channel coding and link-layer protocols, and includes multi-cell mobility and signaling protocols such as power control and soft handoff. See

http://www.winlab.rutgers.edu/pub/projects/Spring99/p013.html

WINLAB adopted SSF as the platform for all wireless network projects. Several other wireless networking companies began negotiations to license this package for implementation of other receiver models (e.g., EDGE) and protocol suites (e.g., GPRS).

There was a new release of TeD-PNNI Models of ATM PNNI protocols for large-scale, high-fidelity simulation, comprising the protocol and network models, documentation, and the parallel simulation package Jane/TeD/GTW. The package is being actively used in networking projects in over 20 corporations and universities, including Lucent Bell Labs, Telcordia Technologies, Newbridge Networks, Traveller Multimedia Networks. See

http://www.cc.gatech.edu/computing/pads/pnnidoc.html

C41SR Global Network Simulator

In 1999, we received an award of $28,500 for our work on “C41SR Global Network Simulator.” The award is a subcontract from the Ultra Corporation, one of the small companies that we have partnered with on the earlier DARPA project on telecommunications networks simulations, and comes from an Air Force contract aimed at design of a simulation of Air Force aerospace operations within the global grid.