The first part of the work yielded new results on simulated annealing and neural networks while the second part focused on large deviations in high-speed communication networks.
May 19, 1992

Dr. Jagdish Chandra  
Director  
Mathematical Sciences Division  
Department of the Army  
Army Research Office  
P.O. Box 12211  
Research Triangle Park, NC 27709-2211

Dear Dr. Chandra:

Enclosed please find the final technical report for ARO contract DAAL03-89-K-0128, "Multidimensional signals, data and processes", Professors Eugene Wong and Jean Walrand, Principal Investigators.

If you have any questions please call me at 510-642-7201 or Professor Walrand at 510-642-1529.

Sincerely,

Margie Berger  
Grants Administrator

cc: Patsy S. Ashe  
Linden Clausen
Multidimensional signals, data and processes
(Contract DAAL03-89-K-0128)

Department of Electrical Engineering and Computer Sciences
University of California
Berkeley, California 94720

Eugene Wong & Jean Walrand
(510 642 1529; wlr@ee.berkeley.edu)

Final report prepared for the Mathematical and Computer Sciences Division of the Army Research Office

1.0 Abstract

The work performed under this contract has two parts. The first part was performed by Professor Eugene Wong and Mr. George Kesidis. That part of the work yielded new results on simulated annealing and neural networks.

On May 1, 1990, Professor Wong was confirmed as Associate Director for Physical Science and Engineering with the White House Office of Science and Technology Policy. As a result of this appointment, Professor Walrand served as Acting Principal Investigator on this contract. The work performed by Professor Walrand and George Kesidis focused on large deviations in high-speed communication networks.

2.0 Succinct Description:

2.1 Neural Networks and Simulated Annealing

2.1.1 Motivation

A neural network computes the minimum of a function. When the function is suitably chosen, the neural network solves a practical optimization problem. Neural networks have been used to recognize speech, to identify features in pictures, to detect messages...
in noisy signals, to control routing in communication networks, to control robots and machinery, and in a variety of other engineering applications.

Simulated annealing refers to a class of global optimization algorithms. These algorithms are particularly useful when relatively little is known about the structure of the function to be minimized and when that function is likely to have many local minima that would trap the more classical optimization algorithms.

The extent of possible applications of neural networks and simulated annealing and their performance compared to other methods are of considerable interest because of their potential applications.

2.1.2 Technical Objectives

Our objectives are to contribute to the understanding of neural networks and simulated annealing. In this work, we propose a new class of neural networks called diffusion machines. We have designed implementations of these diffusion machines. We have also derived new results on the optimal cooling of simulated annealing algorithms.

The diffusion machine that we propose can be viewed as an improvement of Hopfield networks and of Boltzmann machines.

2.1.3 Approach

The diffusion machine is a stochastic version of a Hopfield network. The diffusion machine can be made to converge to the global minimum of a function when the Hopfield network would be trapped by a local minimum.

Alternatively, the diffusion machine can be viewed as a continuous-time version of a Boltzmann Machine. Consequently, the diffusion machine can be made to converge much faster than a Boltzmann machine.

Finally, we have shown that a diffusion machine is suitable for implementation, either with analog circuitry or with a digital signal processing microprocessor.

2.2 High-Speed Networks

2.2.1 Motivation

The future telephone network will probably be built using the Asynchronous Transfer Mode (ATM) technology. Before transporting information, an ATM network breaks it up into fixed size cells of 424 bits. The network transports cells that belong to the same information stream along the same path, called a virtual-circuit.

ATM networks are designed for loosing only a very small fraction of cells. Both the network parameters, such as buffer size, and the rules of operations of the network, such as the call admission, the routing, and the priority mechanisms, must take this objective into account.
2.2.2 Technical Objectives

New design methodologies and adaptive control techniques must be developed that estimate efficiently and limit the small cell loss probabilities. Our research focuses on the development of such design methodologies and control techniques.

Our work has two main objectives. For network design, we need efficient simulation techniques. For instance, if the target cell loss probability is one per million, then a direct simulation must simulate a few million cells going through the network before it can estimate such a small loss probability. Consequently, such a direct simulation is exceedingly slow. Typically, estimating such a small loss probability by a direct simulation of even a very simple network requires many hours on a high-performance workstation. We have developed quick simulation techniques that achieve the same accuracy in a few minutes.

The second objective is the development of quick estimation techniques for the real-time control of ATM networks. Roughly speaking, to control the cell loss rate, the network must monitor that loss rate and make decisions on the basis of the measurements. If the network takes a long time before it has an accurate estimate of its cell loss rate, then it is not able to control that loss rate efficiently. We have developed quick estimators that speed up the collection of the relevant statistics by a few orders of magnitude.

2.2.3 Approach

Our methods are based on results from the theory of large deviations. This theory indicates the likely behavior of the traffic in an overflowing buffer. We use these results in two ways: to select a change of measure for quick simulation, and to speed up real-time estimations by extrapolations.

The quick simulation method can be described as follows. Say that you want to estimate the number of fish in a lake. Assume that you know that the density of fish in a particular region of the lake is 100 times larger than the average density. An efficient method is then to sample that particular region of the lake and to correct for the average density by dividing by the factor 100. The technique of importance sampling is based on that idea. The difficulty in using the technique is in guessing the dense region of the lake. In the simulation of ATM networks, the lake is the set of possible evolutions of the contents of the buffers in the ATM switches and the fish are the evolutions that lead to cell losses. The dense region of the lake, i.e., the evolutions that are likely to lead to cell losses, is discovered by the theory of large deviations. The correction factor (100 in the fishy analogy) is computed during the simulation by calculating a likelihood ratio. Technically, we use level 3 large deviation results for Markov chains. We have used this method for Markov fluid and Markov modulated Poisson models of ATM traffic. The measured speed up achieved by the quick simulation over the direct simulation is typically from 400 to 600.

The quick estimation method is also based on the theory of large deviations. The idea is simpler than quick simulation. Using the theory of large deviations we can show that the cell loss rate at a buffer of size B in a switch of an ATM network is asymptotically exponential in B. The idea is then to measure the cell loss rate for fictitiously small values of B and to extrapolate using the exponential form of the cell loss rate. In practice, the