Advanced Stochastic Modelling
Applied to Telecommunication Networks and Distributed Systems

Dagstuhl Seminar 9813
30.03 – 03.04.98

Ulrich Herzog, Guy Latouche,
Vaidyanat Ramaswami, Phuoc Tran-Gia (editors)

Advanced Stochastic Modelling Applied to Telecommunication Networks and Distributed Systems

In the field of telecommunications and distributed systems we observe a vivid technological and methodological development; the integration of services and new fields of applications require demands of up to now unknown quality and quantity.

Stochastic modelling methods, necessary for the conception, planning, realisation and operation of such systems, always have to be adapted to this development. In many cases the traditional models are no more sufficient; then totally new ways have to be found. Furthermore, shortened cycles of development require highly efficient and extensively automated support by tool environments.

Two and a half years ago we already organised a workshop bringing together people working in the areas of telecommunications and manufacturing (Dagstuhl Seminar Report 127). The main characteristics was that all had a similar theoretical background but different applications in mind. This time we decided to concentrate on telecommunications and distributed systems. However, now we selected researchers from two different communities: People directly dealing with performance problems in industry or academia, and computer scientists trying to integrate performance modelling into functional specification techniques. In total, the attendees of both groups found this confrontation very stimulating.

Again, we all appreciated very much the atmosphere and hospitality at Schloß Dagstuhl. On behalf of all participants we thank the director and the staff of the conference center and would like to say "Auf Wiedersehen".

The organisers

Ulrich Herzog,
Guy Latouche,
Vaidyanat Ramaswami,
Phuoc Tran-Gia

Abstracts are listed in alphabetical order. In case of multiple authors, speakers are listed first.

A Unified Approach to Bandwidth Management and Traffic Routing in ATM Networks

Ake Arvidsson

We discuss the management of virtual paths and virtual channels in ATM networks as a means of congestion control and performance optimisation. The scope of the control mechanisms is presented in a general framework, after which we review several methods, solutions, results, and open issues in the area.

Window Flow Control in FIFO Networks with Cross–Traffic

Francois Baccelli, INRIA Sophia Antipolis (joint work with T. Bonald)

We focus on window flow control as used in packet-switched communication networks. In the simplest model, the route followed by the data packets and the acknowledgments of the controlled connection is modeled by a series of infinite capacity FIFO queues, each of which receives in addition some cross–traffic represented by an exogenous flow. We investigate the stability region of the system under general stochastic assumptions, that is jointly stationary and ergodic input processes. In particular, we show the existence of a minimal stationary regime and establish tight bounds on the maximal throughput allowed by the flow control. The results are illustrated by some examples, including the case of a non-monotonic, non-convex and fractal stability region.

On Markovian Spatial Processes

Dieter Baum, Universität Trier

Spatial point processes have been used increasingly over the last decade in different areas as, for instance, image processing and pattern recognition, statistical mechanics, and applied mathematics. Recently, the computer science branch of telecommunications has recognized their applicability and benefit for modelling the spread of active users over urban or rural areas with its impact on performance in cellular mobile communication systems. In this paper the chronological evolution of marked spatial Poisson point distributions, being controlled by some Markov random field, is investigated, thereby defining the universal spatial version of N-processes, portrayed in form of a convolutional exponential distribution.

Minimum cost synthesis of VP's in ATM networks

Les Berry

In the report "An ATM network planning model" by Farago, Hai, Cinkler, Fekete and Arato a minimum cost extension algorithm is used as the basis of an iterated VP-based network planning tool which forms part of the PLASMA support tool. The talk will review this method and related literature and then discuss extensions to cover synthesis of VP's within specified hop-limits and non-uniform splitting of traffic demands between origin and destination network nodes.

Approximate Stochastic Modeling of Cellular Mobile Networks

Gunter Bolch, University of Erlangen

A very well known and popular technique for the performance evaluation of current computer, communication and operating systems is analytical modeling based on queueing networks because they are easy to understand and use compared to other modeling techniques. For so called product form queueing networks efficient exact and approximate analysis algorithms exist but most queueing networks of realistic systems do not fulfill the requirements of product form queueing networks. To solve this kind of networks Markovian analysis or discrete event simulation is used. The disadvantage of these techniques is that they are very time consuming to prepare, implement and run, especially when we want to analyze the system on a wide range of parameter values. Problems that appear very often in many fields of computer and communication systems are priorities with and without preemption, mixed priorities and class switching. In this paper we show how to transform a queueing network that cannot be solved because of the above restrictions into a network that can be solved using a standard analysis technique (like MVA, SCAT or Convolution). The applicability of this technique is shown on an example of Cellular Mobile Communications Networks with Multimedia Services.

On Independent Flow Time Approximation (IFTA) and Correlations in the End-to-End-Delay in Networks

Hans Daduna, Universität Hamburg

We reconsider the exponential Jackson network of multiserver queues with joint queue length process $X = (X(t) = (X_1(t), X_2(t), \cdots, X_J(t)) : t \geq 0)$ in equilibrium. Jackson's result states that in equilibrium for fixed time point t_0 the queue lengths $X(t_0)$ behave as if they are independent. This is remarkable because the nodes in fact interact, the interaction forces being carried by the migrating customer. Due to this, almost nothing is known about distributions, e.g., of vectors $(X(t_1), X(t_2))$ or $(X_i(t_1), X_j(t_2))$. X therefore exhibits a complicated space-time structure.

Individual customers traversing the network experience space-time interaction with other customers – pathwise evaluation of a customer's space-time diagram is almost impossible, if the network has nontrivial structure. Therefore it is not surprising that distributional characterisations of a customer's successive sojourn times on their itinerary through the network are rare.

Hohl and Kühn therefore investigated numerically and by simulation the power of the IFTA-method: Compute the sojourn time distribution at the successive nodes independently and assume that independence holds in between them. (See Hohl, S.D. and Kühn, P.J.: Approximate analysis of flow and cycle times in queueing networks, in: Moraes, L.F.M. de and Souza e Silva, E. de and Soares, L.F.G.(eds.): Proc. of the 3rd Int. Conf. on Data Communication Systems and Their Performance, North–Holland, Amsterdam, 1988, pp.471 – 485.) This procedure is a wellknown approximation scheme, usually applied to approximately solving the steady state problem for non product form networks: it is a disaggregation – aggregation method. Applying this decomposition approximation to computing end–to–end–delay distributions yields results very close to exact values in case of feed–forward networks. Using the IFTA method to compute end–to–end–delay distributions for paths which include cyclic feedback structures is a critical task, as can be seen from specific examples where it works rather badly.

We discuss possible validation of this method by analytical results. These state that under condition of overtake-freeness with respect to the node structure and the network's topology IFTA is exact; for a review of these results see: Boxma, O.J. and Daduna, H.: Sojourn times in queueing networks, in: Takagi, H.(ed.): Stochastic Analysis of Computer and Communication System, North–Holland, Amsterdam, 1990, pp.401 – 450.

Using correlation theory for Markov processes we show that in many cases when IFTA is proved to be not exact, it provides us nevertheless with strict lower bounds for joint sojourn time probabilities. These bounds for complex joint sojourn time probabilities or for passage time probabilities are expressed using probabilities of simple events concerning only the individual queues which are explicitly known. This is joint work with Ryszard Szekli from Wroclaw University. (See: Daduna, H. and Szekli, R.: On the correlation of sojourn times in open networks of exponential multiserver queues, preprint 97–12, Inst. für Mathematische Stochastik, Universität Hamburg, 1997.)

Structured characterization of the Markov chain of phase-type SPN

Susanna Donatelli

This talk presents a characterization of the Markovian state space of a Stochastic Petri Net with phase type distributed transitions as a union of Cartesian products of a set of "components" of the net. The method uses an abstract view of the net based on the vectors of enabling degrees of phase type transitions, as well as on the set of "interrupted clients". Following the decomposition used for the state space characterization, a tensor algebra expression for the infinitesimal generator (actually for its rate matrix) is given, that allows the steady state probability to be computed directly from a set of matrices of the size of the components, without the need of storing the whole infinitesimal generator.

The Polling System with State-Dependent Changeover Times: The Branching Process with State-Dependent Immigration

Martin Eisenberg

We solve the polling system in which the changeover-time distribution at each stage of the server cycle depends on whether the visited queue has customers waiting. This model has many applications in communications and other areas. The solution involves an unusual "two-dimensional" iteration that determines the Laplace-Stieltjes transform of the waiting time distribution at each queue. This problem has a counterpart in the theory of branching processes, specifically the multi-type branching process with state-dependent immigration. We describe the necessary and sufficient conditions for the embedded Markov Chain to be positive recurrent, which are the same conditions for the iteration to converge.

Analytical Models of TCP Transport for Wireless Internet Access

Ashok Erramilli, Netmetrix

This paper considers two questions: i) what is the performance of a typical Internet application over a wireless access channel? ii) is it feasible to do any performance modeling of the Internet? The first question is motivated by current industry plans to deploy large scale, wireless, fixed Internet access networks. The second question is motivated by the divergent viewpoints between researchers in the teletraffic community (who believe it is possible to develop simple models of TCP transport) and those in the networking community (who believe highly detailed simulation models are needed). A simple analytical model of TCP transport over lossy wireless channels is developed and analyzed using product form heuristics: mean value analysis is used to study network performance at the IP layer, and the parameters of the IP model (such as window sizes and retransmission timer values) are determined using a mean field analysis of the TCP layer. The model predicts the sensitivity of TCP performance to loss rates in the wireless channel, and shows that a typical application cannot avail itself of the full bandwidth of high-speed, wireless access links because of the short file sizes and overheads in TCP Slow Start. Comparisons with detailed simulations show that while the simple model can predict aspects of TCP performance (such as expected window sizes), more detailed models are needed to accurately estimate download times and throughputs.

Performance of Wavelength Division Multiplexed Optical Networks

David Everitt, The University of Melbourne
Jennifer Yates, Michael Rumsewicz, RMIT, Melbourne, Australia
Bill Henderson, University of Adelaide, Adelaide, Australia

The next generation of high-capacity, long-distance telecommunications networks (so-called "transport networks") will be based on wavelength-division multiplexed optical networks. This talk will discuss some of the traffic engineering problems with such networks, in particular with dynamically-reconfigurable networks. Some of the issues to be addressed include modelling of

wavelength-continuous networks, the use of wavelength conversion, and the tradeoffs between high network performance and the hardware and network management needed to achieve this performance.

Performance Analysis of AAL-2 Carrying CDMA Traffic

Notker Gerlich

Due to cost efficiency, future personal communication networks likely will utilize the Asynchronous Transfer Mode (ATM) technology to connect components of the land network infrastructure. In order to cope with bandwidth limitation on the radio access, voice is transmitted as compressed low bit-rate information. The ATM Adaptation Layer 2 (AAL-2) is designed to handle this type of low bit-rate traffic. This paper discusses the capability of using AAL-2 connections to carry compressed voice between components of a CDMA network. To this end we analyze a discrete-time queuing model of the the AAL-2 mechanism combined with traffic shaping.

Analysis of Non-Markovian Stochastic Petri Nets

Reinhard German

As a modeling formalism, stochastic Petri nets are considered in which the firing time distribution of the transitions may be generally distributed. As a restriction, at most one generally timed transition may be dominant. We first give a survey of the main analysis approaches: the phase method, application of Markov renewal theory, and the method of supplementary variables. A formal relationship of the latter two is then presented. Subsequently, the supplementary variable approach is used as a framework covering a number of cases: stationary as well as transient numerical analysis, computation of state and firing probabilities, different preemption policies, marking-dependent firing time distributions, and the reducible case. Some tools supporting the analysis are discussed: a Mathematica prototype, TimeNET, and SPNL. Finally, a perspective is given for the analysis of models in which generally timed transitions can be concurrently enabled.

A Simple Model for IP over ATM

R.J. Harris, L.T.M. Berry, I. Atov, S. Sharma, Royal Melbourne Institute of Technology

We consider the problem of modelling Internet Protocol (IP) traffic that is being transported over links of an Asynchronous Transfer Mode (ATM) network. A simple flow based model of IP over ATM is presented that can be used to map parameters of the IP service mix to three principal parameters of the Guerin et al model, viz: Peak, mean and burst time.

One-Place Unbounded Stochastic Petri Nets with Applications to Communication Systems Modelling

Boudewijn R. Haverkort, RWTH Aachen

We address a class of stochastic Petri nets (SPNs) for which the underlying Markov chain (CTMC) exhibites a quasi-birth- death structure. Such CTMCs can be solved using the logarithmic reduction method or using spectral expansion. With the software tool SPN2MGM we can easily analyse models with block sizes up to one thousand. We show various examples and present comparisons between the efficiency of the logarithmic reduction and the spectral expansion method.

Approximate Analysis of Finite-Buffer Queueing Networks by Decomposition

Armin Heindl

Decomposition as an approximate solution technique for large queueing networks with generally distributed interarrival and service intervals has been studied intensively for nodes with infinite buffers. The network is partitioned into subsystems which are analyzed in isolation by considering the related input and output traffic processes modelled as renewal processes. In my talk, I presented a new approach to deal with (open) queueing networks, in which all or some of the buffers may be finite. Customers who arrive at a node with a full buffer will be lost without affecting the future behaviour of the network. Further, it is assumed that queues with a finite buffer will only be served by a single server. For networks of this kind, quantities like loss probabilities, throughputs and delays can be approximately obtained in a quasi-prompt answer. The proposed algorithm classifies general finite-buffer queues as pseudo ‘M’/GI/1/K or pseudo GI/‘M’/1/K nodes and exploits a duality between these types of systems together with the principle of maximum entropy to find approximate, but closed-form expressions for different queue length distributions relevant in GI/GI/1/K systems. Experiments have shown that results generally exhibit relative errors of around 10%, often less, when compared with available exact values.

Compositional Analysis of an Ordinary Telephony System

Holger Hermanns, Universität Erlangen-Nürnberg

Stochastic Process Algebras support compositional generation and analysis of complex performance models. In this talk I will introduce a new variant of this general approach, ’Interactive Markov Chains’. Significantly different from other approaches, this algebra contains ordinary process algebra and and continuous time Markov Chains as proper subalgebras. By means of a nontrivial case study, I will show that especially this property allows one to tackle performance models of an immense size. An ordinary telephony system, with more than 10 million states will be analysed compositionally. We will also discuss the central algorithms used to make the state space tractable.

Compositional Solution of Stochastic Process Algebra Models

Jane Hillston, University of Edinburgh

Using a stochastic process algebra (SPA) offers the performance modeller the opportunity to develop models in a formalism which is familiar and accessible to designers of telecommunication networks and distributed systems. However, as with all state-based approaches, solving SPA models of realistic systems is fraught with problems of state space explosion. This talk will describe attempts to overcome these problems by taking advantage of the structure which is inherent within all process algebra models.

After a brief introduction to the Markovian process algebra, PEPA, I will explain how syntactic conditions on the modelling language, which can be observed during model construction, may be used to identify models which are amenable to exact decomposed solution. In particular, I will outline recent results which characterise in PEPA the product form Markov processes identified by Boucherie in 1993.

URL: http://www.dcs.ed.ac.uk/publications/lfcsreps.html

Internet Traffic Models – Queueing Performance Comparison

Frank Huebner, AT&T Labs

There are a fast-growing number of studies in the literature that deal with capturing the statistical properties of today's data traffic. While this is a good development and a lot of progress has been made in terms of understanding the properties of this traffic, little progress has been made in regard to the question of how traffic characteristics like long range dependence/self-similarity or heavy tailed distributions of interarrival times impact the engineering of data networks.

A number of traffic models have been shown to capture certain aspects of the statistical properties of measured data traffic well, but it was not studied how well those models can be used to predict queueing performance for engineering purposes. We compare in this paper the performance prediction of a wide range of traffic models for a number of performance measures that are common in network engineering. In particular, we investigate how accurately traffic models like Poisson, 2-state MMPP, AR(1), Weibull, Pareto, and FBM predict queueing performance like loss probabilities and queue lengths (delays) when compared to results from measured Internet traffic.

The numerical results lead to the conclusion that traffic models that are based on heavy-tailed distributions like the Pareto distribution are needed when dealing with small buffer sizes. For larger buffer sizes, the effect of long range dependence that is present in the data traffic influences the queueing behavior considerably and therefore traffic models that exhibit long range dependence like FBM are needed. It is interesting to see how the long range dependent traffic models fail to predict the queueing performance for small buffer sizes. This indicates that the choice of the traffic model depends not only on the type of traffic but also on the considered application/device in terms of buffering.

The non-Markovian Stochastic Process Algebra ♤

Joost-Pieter Katoen, Universität Erlangen-Nürnberg
(joint work with Pedro D'Argenio and Ed Brinksma)

Stochastic process algebras (SPAs) like PEPA, TIPP and EMPA allow the user to specify the occurrence of actions after a random delay. These delays are determined by exponential distributions. This facilitates a formal interpretation in terms of standard transition systems and establishes a close link to continuous-time Markov chains. To put it in a nutshell, these process algebras allow the compositional specification (and analysis) of continuous-time Markov chains.

In this talk we discuss the extension of SPAs with general distributions. We present an extension of transition systems with clocks that are basically random variables of continuous or discrete nature. This model, referred to as stochastic automata, is strongly related to generalised semi-Markov processes (GSMPs). We present a non-Markovian SPA called ♤ (SPADES, SPA for Discrete-Event Simulation) and discuss its compositional semantics using stochastic automata. Besides we present some equivalence relations, and an axiomatisation. In particular we show how an expansion law, an important law for reducing parallel composition into more elementary operations, is obtained. As a result, the process algebra ♤ allows the compositional specification (and analysis) of time-homogenenous, mono-rated GSMPs with deterministic branching.

URL: http://www7.informatik.uni-erlangen.de/ katoen/procomet98.ps.gz

Bottleneck Analysis in Multiclass Closed Queueing Networks and Its Application to Bandwidth Dimensioning

Yaakov Kogan and Arthur Berger, AT&T Labs

This paper is motivated by a new application of closed queueing networks (CQN) with a large number of customers. The application is the dimensioning bandwidth for elastic data sources subject to feedback control in packet-switched communication networks, such as Internet Protocol (IP) or Asynchronous Transfer Mode (ATM) networks when available bandwidth at the servers is shared between all active sources. Elastic data sources can adapt to time-varying available bandwidth via a feedback control such as the Transmission Control Protocol (TCP) or the Available Bit Rate (ABR) transfer capability in ATM. Typical elastic data sources generate transfers supporting e-mail or the world wide web.

In a CQN, data sources are modeled by an infinite server (IS) station, and network nodes are modeled by processor sharing (PS) stations. It is known that the steady state queue length distribution in such a CQN has a product form that is defined explicitly up to the normalization constant. The distinguished property of the new application is that this CQN model is valid only if one or more PS stations form a bottleneck. The bottleneck station is defined asymptotically as the station where the number of customers grows proportionally to the total number of customers in the network as the latter increases simultaneously with service rates at PS stations.

Asymptotic behavior of queues is studied for large closed multi-class queueing networks consisting of one IS station with K classes and M PS stations. A simple numerical procedure is derived that allows to identify all bottleneck PS stations for which the queue lengths, normalized by the large parameter, converge to positive deterministic limits. For the case when $K = M = 2$, the set of

network parameters is identified that correspond to each of the three possible types of behavior in heavy traffic: both PS stations are bottlenecks, only one PS station is a bottleneck and a group of two PS stations is a bottleneck. In the last case both PS stations are equally loaded by each customer class and their individual queue lengths, normalized by the large parameter, converge to uniformly distributed random variables. These results are directly generalized for arbitrary $K = M$. Generalizations for $K \neq M$ are also indicated.

The case of two bottlenecks is illustrated by its application to the problem of dimensioning bandwidth. An engineering rule is provided for determining the link rates such that a service commitment on a per-class throughput is satisfied.

Admission Control and Routing for Connection-oriented Multirate Traffic

K. R. Krishnan (Bellcore) and F. Huebner-Szabo de Bucs (AT&T Labs)

It is NOT always advantageous to combine different classes of traffic on the links of a network, without a suitable admission control scheme. The problem of devising such an admission policy is studied here by means of Markov decision theory.

For n traffic classes, with Poisson arrivals and exponential holding-times, the Markov decision process for even the single-link admission problem has an n-dimensional state-vector, defined by the numbers of calls of the n classes in progress on the link. In this original formulation, the single-link problem itself becomes numerically intractable, even for small values of n. We propose an approximation which allows us to work with the scalar state variable consisting of the total bandwidth occupied by all the calls in progress, regardless of the number of traffic classes. As a result, the decision process for the multirate admission-control problem on a link is solved by a single system of linear equations of size $(C + 1)$, where C is the number of "trunks" (integer multiples of a basic bandwidth unit) in the link, regardless of the number of traffic classes. The call-admission criterion derived from this single-link analysis at once generalizes to a state-dependent rule for alternate routing and flow-control for multirate connection-based traffic in a network.

We have applied our link-admission and routing rules both to examples of connection-based CBR traffic and VBR traffic; in the latter case, we assign suitable effective bandwidths to the VBR connections and treat them like CBR connections. In the simulations to demonstrate our results, we use the Fractional Brownian Motion (FBM) model for VBR traffic with long-range dependence, and compare the performance of the proposed admission-rule with that of the complete-sharing rule. The examples show that our method offers a computationally feasible algorithm for admission control and alternate routing for connection-oriented multi-rate traffic, whether CBR or VBR.

Stochastic Models Derived from Formal Descriptions of Distributed Systems

Pieter S. Kritzinger, University of Cape Town, South Africa

Predicting the performance of a communication protocol from a formal description was first proposed about a decade ago. Such performance prediction involves three issues: the semantics to give time, the analysis method to solve the resultant model and, depending on the latter, estimates of the

values of the input parameters of the model. In this paper we address mainly the first issue: The semantics to give time and how to introduce it into the model. The approach we propose does not affect the syntax of the formal description technique and does not depend on the FDT used. We point out that the problem of determining the values of the appropriate parameters in those cases remain however. We illustrate our ideas using SDL and derive performance measures from a simulation or execution of the specification and briefly describe a software tool which implements our proposals and an example to illustrate their application in practice.

A D,MAP/D1,D2/1 queue with regular preemptions in ATM networks

Fumiaki Machihara, Tokyo Denki University

It is necessary to study a D,MAP/D1,D2/1 queue with regular preemptions for the performance evaluation of the statistical multiplexer with an interface based on the synchronous digital hierarchy (SDH). Since it is generally very difficult on numerical calculations to analyze the regularly preempted model, we propose an approximate model MAP/SM/1 queue with some semi-Markov services. In addition, using two modified models with the semi-Markov services which have the exceptional services for the first customers of the busy periods(a set-up queue or something like that), we derive upperbounds and lowerbounds for some performance measures of the regularly preempted model and the approximate semi-Markov service model.

Evaluation of Connection Admission Control Methods in ATM Networks

Michel Mandjes, Hans van den Berg, KPN Research, the Netherlands

In ATM, the Call Admission Control (CAC) mechanism has to decide whether a new connection can be accepted without violating the QoS commitments. In order to make CAC safe, one usually assumes that the traffic on the connections is of “worst-case” type. Among all possible traffic profiles that are compliant to the traffic contract parameters PCR, SCR and MBS, worst-case traffic essentially maximizes the cell loss probability. These traffic stream are often modelled as periodic on-off, with peak rate, on-time and off-time in agreement with the contract parameters. A number of CAC methods based on worst-case traffic have appeared in the literature. A neat mapping from traffic contract parameters (SCR, PCR, MBS) on the network resources, exclusively dedicated to the connection, is due to Elwalid, Mitra, and Wentworth. They provide algorithms for the lossless case as well as the situation in which small loss probabilities are allowed. In the latter case statistical multiplexing can be exploited. A further improvement is made by Lo Presti, Zhang, Towsley, and Kurose. In order to evaluate and compare these CAC methods (w.r.t. safety and bandwidth efficiency) one would like to have a tool that is able to calculate, for a given combination of periodic on-off sources, the loss probability. Unfortunately, exact analysis for the corresponding queueing model is not available and estimation of the small cell loss probabilities by simulation would take a lot of time. For that reason we rely on large deviations asymptotics, as developed by Simonian and Guibert, Courcoubetis and Weber, and Botvich and Duffield. We adapted their analysis and verified the quality of the approximation by extensive simulation experiments. Application of the resulting “CAC evaluation tool” shows that the above mentioned CAC methods may perform rather poor.

We also pay attention to CAC in the case that real-time and non-real-time traffic streams are multiplexed on a single link. Cells originating from the real-time traffic are served with higher priority than the non-real-time cells. Assuming a small buffer for the real-time traffic, it is shown that the cell loss for each traffic type can essentially be analysed as in the case of a single traffic type. CAC methods are adapted and evaluated.

Offered Load Models for Telephony, Wireless, and ATM

William A. Massey, Bell Laboratories of Lucent Technologies

Many features of modern telecommunications are not easily modelled by traditional queueing theory. Using stochastic integration over nonhomogeneous Poisson processes, we can create new queueing models that allow us to paradoxically analyze finite server (bandwidth) systems through an exact analysis of infinite server systems. These "offered load" models incorporate more realistic features such as time varying arrival rates and non-exponential holding times.

The Isotropic Phase Planar Point Process

Marie-Ange Remiche, Université Libre de Bruxelles

In this talk, we first recall the definition of the family of phase-type planar point processes (as defined by Latouche and Ramaswami, see Seminarbuch September '95 - April '96, page 80). We then especially put emphasis on the process we call in short IPhP^3 for Isotropic Phase Planar Point Process. Secondly, we propose to highlight the tractability of the IPhP^3 by establishing the two first moments of the number of points repectively in a polar rectangle and a disk. Those two distribution are of particular importance when evaluating performance in a wireless network.

A Memory Markov Chain Model For VBR Traffic With Strong Positive Correlations

Oliver Rose, University of Würzburg

We propose a new modeling approach for variable bit rate (VBR) traffic in packet networks based on a Markov chain with memory. The model is simple, comprehensive, and facilitates the modeling of strong positive correlations over a considerable range of lags. The model can easily be adapted to measured traffic sequences, e.g. MPEG video or packet interarrival time sequences. Due to its simplicity and its close relationship to Markov chains the model is valuable to both simulation and analysis of currently considered types of traffic.

Asymptotics of Stochastic Networks with Subexponential Service Times

Volker Schmidt, University of Ulm
François Baccelli, Sabine Schlegel, INRIA Sophia Antipolis

We analyse the tail behaviour of stationary state variables in $(\max, +)$-linear systems. For a K-station tandem network of single server queues with infinite buffer capacity at each station, we show that if the tail of service times of one server, say server $i_0 \in \{1, \ldots, K\}$, is subexponential and heavier than those of the other servers, then the stationary distribution of the waiting or response time until the beginning of service at server $j \geq i_0$ asymptotically behaves like the stationary waiting time distribution in an isolated single-server queue with server i_0. A similar result is then derived for the asymptotics of $(\max, +)$-linear systems with i.i.d. matrices. As a direct application, we give the asymptotics of stationary response times in stochastic event graphs with independent firing times, whenever the tail of the firing times of one of the transitions is subexponential and heavier than those of the others.

Symbolic Representation and Manipulation of Stochastic Transition Systems

Markus Siegle, Universität Erlangen-Nürnberg

A BDD (Binary Decision Diagram) is a compact canonical representation of a Boolean function. While BDDs are considered key technologies in the area of hardware verification and model checking, their use for the purpose of performance analysis is a new idea.

We use BDDs to represent labelled transition systems which arise from higher-level model specifications such as stochastic process algebras or structured stochastic Petri nets. BDDs offer a compact representation of transition systems with very large state space. They are therefore promising candidates for alleviating the problem of state space explosion.

We offer a new approach to the question of how to incorporate stochastic information into the BDD, concentrating on Markovian models. A new data structure, Decision-node BDD (DNBDD), is introduced and used to represent stochastic transition systems. A DNBDD is structurally identical with the corresponding ordinary BDD, but some of its nodes — which we call decision nodes — carry additional information.

Generation and manipulation algorithms for DNBDDs are discussed. We show how the parallel composition of two stochastic transition systems can be performed in the DNBDD context. Reachability analysis can also be carried out on the DNBDD representation. DNBDD-based bisimulation minimisation (which corresponds to Markov chain lumpability) is also discussed. All algorithms are supported by a prototype tool.

URL: http://www7.informatik.uni-erlangen.de/ siegle/own.html

Continuous Petri Nets: On the Analysis of Autonomous Models

Manuel Silva, University of Zaragoza

Performance evaluation of DEDS models with a very large state space may become unfeasible, unless closed solutions exist. In some cases, continuization (fluidification) of the state space may be applied. This traditional approach is frequently used, for example, when considering population dynamics, like in predator-prey models (in the most simple case, the well-known Volterra-Locka equation is obtained).

Continuization of DEDS models described with Petri Nets has been done by R. David and H. Alla, introducing the so called Continuous (and later Hybrid) PNs, and by K. Trivedi and his group, introducing Fluid Stocastic PNs (in fact some kind of hybrid models in which one or more places can hold fluid rather than discrete tokens).

The motivation for our work is the consideration of Petri Nets as a paradigm usable all along the diverse phases of the life cycle of DEDS. We investigate epistemological and pragmatical aspects, essentially related to the coherence between discrete and continuous PN models and their respective analysis theories. More precisely, even before giving any time intrepretation, we concentrate on some basic qualitative properties of continuous PN models, like reachability, boundedness, deadlock-freeness, liveness, reversibility, implicit places, …and their relationships wrt the corresponding DEDS models.

Structured Markov Chain Models using Transition Classes

Christoph Strelen, Universität Bonn

Transition classes are proposed as a high-level specification method for Markov chain models. They define Markov chains, classify the state transitions, and provide structured and compact models which are well suited for immediate application of solvers, including aggregation/ disaggregation techniques, and simulation. Other high- level models like stochastic Petri nets or queueing networks can be transformed automatically into transition class models. Transition classes have close similarities with stochastic process algebras.

Policies for Spectrum Sharing in Mobile Systems

P.G. Taylor and M.A. Stewart, University of Adelaide

Consider the situation where a number of operating companies have paid for a block of channels which make up some part of the spectrum in a cellular mobile system. For operating company i, users arrive in a Poisson process with parameter λ_i and remain in the system for an exponentially distributed length of time with parameter μ. If a user arrives to find that it cannot be accomodated, then the user is rejected and the operating company loses revenue.

The owner of the spectrum has the problem of managing it so that the operating companies receive a share which is proportional to the amount that they have paid. It wishes to do this in the best way possible so that revenue is maximised.

One method of managing the spectrum would be to partition it so that each operating company receives sole access to a portion of the bandwidth. However it may be more efficient to allow all companies access to all of the bandwidth and employ a control which ensures that users from company i get a grade of service which is related to the amount that company i has paid for use of the resource.

An example of such a control is *trunk reservation* in which a user from operating company i is accepted only if there is room for at least r_i further calls from customer of the other companies. Our first result is that the numbers r_i, called trunk reservation parameters, can be chosen so that calls from each of the operating companies receive grades of service no worse than if the spectrum is partitioned. Moreover we present numerical evidence to show that, using trunk reservation, it is possible to give each of the operating companies exactly the grade of service that they would achieve under partitioning of the spectrum, using an amount of spectrum which is less than the total available.

There is, however, a problem with the practical implementation of trunk reservation in a shared spectrum environment. In order to perform call admission control, each operating company needs to know the total number of calls, which requires that they know something about their competitors' business. In today's deregulated telecommunication environment this is unlikely to be acceptable. This difficulty can be partially overcome implementing a call admission procedure in which operating company i searches the spectrum and allocates a call to the $(r_i + 1)^{st}$ free channel that it observes. This, however, is likely to introduce a significant delay at call set-up stage, which may not be satisfactory. Also the companies would still be able to infer a reasonable amount of information about each others' business.

Another call admission control, which might be expected to behave in a similar way to trunk reservation, could be constructed by allowing operating company i to allocate a call to the first free channel that it finds, but restricting its search to a portion of the spectrum. This may perform almost as well as trunk reservation but would not have the same implementation difficulties.

In this talk we discuss the fine details of how these different strategies might be implemented, and make some recommendations about which is likely to provide the best performance.

Stochastic and Deterministic Server Behaviour – A Simple Case Study

Erwin Thurner, Siemens AG

It is well-known that systems containing deterministic servers lead to other performance measures than systems with stochastic servers. Considering a simple real-world server model, we show on the one hand that modeling the wrong server behaviour may cause performance measure deviations in both directions. On the other hand, this model is big enough that simple approaches to solve this problem analytically do not fit right. Because analytical solution methods have significant advantages compared to simulation, this modeling example can be used to study the usefulness of analytical methods for deterministic problems.

Teletraffic Models in Wireless Network Planning

Phuoc Tran Gia, University of Würzburg

The planning task of modern wireless communication networks strongly influence the overall network capacity and related performance. As mobile networks have been deployed, the customer number increases and their teletraffic behavior is rapidly changing, the need to develop appropriate mobile network planning methods becomes crucial to guarantee the desired quality of service.

In this talk we take the example of CDMA (Code Division Multiple Access) wireless networks to show teletraffic and related modeling problems arising in network planning. The first part deals with traffic estimation and description, where the cluster traffic and elasticity of the cell coverage area are analyzed. The second model example takes the power control mechanism into account, where analytical expressions for the mobile station transmit power based on a discrete time description are presented, which allow to characterize the dynamic behavior of the power control loop in a realistic customer population distribution.

Monte Carlo simulations: Gibbs sampling and variance reduction by conditional expectations

Jorma Virtamo, Helsinki University of Technology

Product form systems, such as multirate loss systems, are theoretically simple and well understood. Still, the calculation of e.g. call blocking probabilities in real networks with a very great number of traffic classes is a challenging task. Monte Carlo simulations provide one approach, though even then the state space explosion poses a problem. In this talk, two aspects of the Monte Carlo method are addressed.

First, the Gibbs sampling for the generation of the sample points is reviewed. In this method an artificial Markov chain is constructed such that the chain has the same stationary distribution as the original system. In this chain the system makes alternatingly transitions in different co-ordinate directions with a transition matrix where the probability of the next state is given by the conditional probability of the stationary distribution in a one-dimensional subspace which the current state belongs to.

Second, as a natural extension of the Gibbs sampling, an improved statistics collection method is introduced. The method is based on an elementary identity of conditional expectations. Its use in the Monte Carlo context seems to have escaped attention, in spite of its simplicity and the significant variance reduction that can be obtained. Basically, the method utilises known exact conditional expectations, which can easily be precalculated in the case of product form systems. Numerical examples on the Monte Carlo simulation of blocking probabilities are given. In this particular case, the conditional expectations represent the conditional blocking probabilities in one-dimensional subspaces and are given by the ordinary Erlang loss function. Thus, instead of sampling the values (0 or 1) of an indicator function, indicating whether the current state is a blocking state or not, one samples the values of the Erlang function in the range (0,1).

There is more to network traffic than self-similarity

Walter Willinger, AT&T Labs-Research

I will discuss three recent developments in the area of understanding data networks and data network traffic: structural modeling of self-similar network traffic, the emergence of wavelet techniques for analyzing network measurements, and the multifractal nature of wide-area traffic.