Discrete Event Simulation – Modelling and Simulation of Computer Networks

Alexander Klein
klein@net.in.tum.de

Nils Kammenhuber
kammenhuber@net.in.tum.de

Georg Carle
carle@net.in.tum.de

Chair for Network Architectures and Services
Department of Computer Science and Mathematics
Technical University of Munich

ABSTRACT

This article introduces the course *Discrete Event Simulation – Modelling and Simulation of Computer Networks* which was recently developed at the Technical University of Munich (TUM). The course is developed as part of the *Tempus IV project – Modernization of Master Program Networks and Communication*. The course has become very popular at TUM since it provides students with a sophisticated theoretical and practical background in statistic fundamentals and simulation which gives them a clear advantage for their MA thesis. The course is divided into 12 to 14 lectures with duration of 90 minutes each. Furthermore, 10 exercises, each with duration of approx. 60 minutes, are also part of the course. In addition, two tutorials are given at the end of the course in which the students learn to use state-of-the-art simulation and evaluation software. The target group of the course is represented by master students. The students get 4 ECTS credits for the course if they successfully pass 6 of the 10 exercises and the oral exam (approx. 20 minutes) at the end of the semester.

INTRODUCTION

The development of the course *Discrete Event Simulation – Modelling and Simulation of Computer Networks* is driven by the fact that many publications show lack of meaningful statistical evaluation. From our point of view, statistical fundamentals and evaluation of simulation results is a mandatory requirement for every scientific researcher. We are confident that we may increase the quality of the MA theses of the students at TU Munich by providing them with a sophisticated background in statistic fundamentals. However, theoretic knowledge is not sufficient when it comes to statistical evaluation of test bed measurements and simulation results. Therefore, we developed exercises that are based on the lectures in order to help students apply their theoretic knowledge on practical examples. Moreover, we focus on real-world examples since students are much more motivated if they can simulate and evaluate everyday situations, e.g. queuing in a super market. Besides these everyday examples, optimization issues are discussed and simulated. Students discuss typical questions from waiting queue theory, e.g. whether a fast system provides a higher performance than several slower systems. They also learn the impact that the variance of the arrival process and the service process has on the overall system performance by comparing the waiting time and the queue length distribution.

CONTENT

The main goal of this course is to provide students with a statistical background and basic knowledge regarding simulation in order to improve the quality of MA theses. The improvements mainly result from two facts. First, the time that students require to get familiar with their research topic is shortened since they already have a good understanding of simulations and the impact of randomness on the outcome of measurements and simulations. Moreover, the gained knowledge enables students to evaluate and interpret their measurement and simulation results in a sophisticated way. In addition, the students clearly benefit from the tutorials with state-of-the-art calculation tools, e.g. Matlab and Octave, such that the visualization of their results is on a much higher level.

The course covers many topics to provide a wide knowledge. It is divided into eight chapters and two tutorials which are supplemented by ten exercises. The exercises are closely related to the topics discussed in each chapter and represent the practical counterpart to the theoretic topics that are introduced and discussed during the lectures. The following table shows a brief outline of the course and its main topics.

Name	Content	Duration
Chapter 1	• Simulation: What it is and when to use it • Types of simulators • Internals of discrete event simulators • Continuations and co-routines	150 minutes
Chapter 2	Statistics Fundamentals: • Introduction to Waiting Queues • Random Variable (RV), Discrete and Continuous RV • Probability Space, Frequency Probability • Distribution(discrete), Distribution Function(continuous) • Probability Density Function, Cumulative Density Function • Definitions: Expectation/Mean, Mode, Standard Deviation, Variance, Coefficient of Variation, p-percentile(quantile), Skewness, Scalability Issues, Covariance, Correlation, Autocorrelation Visualization of Correlation	150 minutes
Chapter 3	Random Numbers: • Generation of Random Variables (RV) Inversion, Composition, Convolution, Accept-Reject • Distributions and their Characteristics Uniform(continuous), Normal, Triangle, Lognormal, Exponential, Erlang-k, Gamma, Uniform(discrete), Bernoulli, Geom, Poisson, General Discrete Random Number Generators: • Linear Congruential Generator(LCG), Shift Register, Generalized Feedback Shift Register, Mersenne Twister • Tests χ^2 Test, Spectral Test, Serial Test	240 minutes
Chapter 4a	Evaluation of Simulation Results: • Estimator Consistent Estimator, Unbiased Estimator, Variance of an	150 minutes

	Estimator, Bessel's Correction, Efficient Calculation • Confidence Interval (CI) Chebyshev CI, Central Limit Theorem, t-Distribution CI • Evaluation and comparison of Simulation Results Replicate-Delete Method, Batch Means Method, Stationarity	
Chapter 4b	How to Lie with Statistics: • Lessons for Authors and Readers • Examples and Discussion	90 minutes
Chapter 4c	Model Validation: • Calibration, Overfitting • Structural Change, Parameter Change • Comparison of Confidence Intervals: Welsh, Law & Kelton	90 minutes
Chapter 5	Experiment Planning: • Hypothesis Testing • Linear Regression • Variance Analysis (ANOVA) • Factorial Design	90 minutes
Chapter 6	Parallel Simulation: • Conservative approach: Deadlock avoidance, deadlock detection and recovery • Optimistic approach: Time Warp • Alternatives to parallel simulation	180 minutes
Chapter 7	Mobility: • Mobility in General • Human Mobility Pattern • Visualization: Density, Speed Histograms, Bouncing Rule, Obstacles • Characteristics of Mobility Pattern: Link Duration, Transient Phase, Node Distribution, Speed Distribution, Correlated Movement • Synthetic Mobility Models: Random Waypoint, Random Direction, Random Walk, Levi-Flight, Brownian Motion, Group Mobility	90 minutes
Chapter 8	Point Fields: • Generation of Point Fields • Homogeneous and Inhomogeneous Point Fields • Poisson Field, Clusterfields, Matern Cluster Field Random Graphs: • Graph Definition • Generation of Random Graphs • Probabilistic Model, Waxman Model • Random Graphs with Predefined Characteristics • Scale-free Graphs, Social-networks	90 minutes
Total: 8 Chapters		1320 minutes = 22 hours

Table 1: Discrete Event Simulation - Content of the lectures

In order to supplement the lecture with its theoretic content, tutorials and exercises are given to deepen the understanding of the students. The content of the exercises is based on the topics of the previous lecture. Their structure is cumulative such that the students are implementing a discrete event simulator (waiting queues), a random number generator and a statistic evaluation tool. At the beginning of the course Java or C++ is chosen as programming language depending on the programming skills of the students. Besides the typical topics, e.g. waiting queues, the students implement point fields, random graphs and synthetic mobility models. Table 2 shows the topics of the exercises and gives an estimation of the expected processing time.

Exercise	Topic	Processing Time	Duration
Exercise 1	Implementation of a GI / GI / 1 - 1 queuing system	180 minutes	60 minutes
Exercise 2	Evaluation of waiting queues Evaluation of medium access procedures	180 minutes	60 minutes
Exercise 3	Evaluation of waiting queues Implementation of random variables	180 minutes	60 minutes
Exercise 4	Implementation of histogram Evaluation of system performance	180 minutes	60 minutes
Exercise 5	Generation of random variates Evaluation of samples	180 minutes	60 minutes
Exercise 6	Implementation of random number generators Evaluation of random numbers	180 minutes	60 minutes
Exercise 7	Implementation of a state dependent service unit Intelligent system initialization	120 minutes	60 minutes
Exercise 8	Implementation of a 2D point field generator Random Graphs	120 minutes	60 minutes
Exercise 9	Random Graphs Comparison of confidence intervals	120 minutes	60 minutes
Exercise 10	Matlab/Octave OPNET Modeler	120 minutes	60 minutes
Total: 10 Exercises		1560 minutes	600 minutes

Table 2: Discrete Event Simulation – Overview of the Exercises

COURSE

First of all, we introduce the basic types of simulations in chapter 1 of the course. We highlight design choices, e.g. continuous time-driven or discrete event-driven simulations. In addition, the advantages and disadvantages of both approaches are discussed. An introduction to waiting queues represents the starting point for chapter 2 which concentrates on statistic fundamentals. We decided to include this chapter in the course since the subsequent chapters require this knowledge in order to understand their more advanced topics. However, the majority of the students already have some knowledge of statistics. Thus, the lecturer can go through the slides more quickly compared to the slides of the other chapters. Nonetheless, we strongly recommend not skipping this chapter since it turned out that students often think that they have more knowledge in statistics than they truly have. The topics discussed in the first two lectures provide the basis for the first two exercises.

It is clear that the implementation of a discrete event simulator and tools for statistic evaluation represent challenging tasks which has to be split into several exercises. The basic idea of the exercises is to encourage students to build a discrete event simulator step by step from scratch. Individual errors and implementations are discussed in the group during the exercises which helps the students to get familiar with different kinds of programming styles. Due to the fact that an efficient implementation of the simulator requires deeper knowledge than the students have at this point, we guide them through the first exercises by providing them with additional background information. Furthermore, a sample solution is introduced and discussed after each exercise to allow students to catch up if they were not able or had to many problems in solving the exercise.

The focus of the first exercises lies on the implementation of waiting queues since the evaluation of waiting queues represents an ideal starting point. The students get familiar with the impact of high variance on the overall system performance under different utilizations. Moreover, the influence of correlated random numbers can be shown which also represents the motivation for chapter 3 of the course.

Chapter 3 starts with an introduction of how random variables can be generated. The four most popular mechanisms, inversion, composition, convolution and accept-reject are introduced. In addition, the efficiency of these mechanisms is discussed, e.g. the low precision of the inversion method due to the fact that values between [0;1] are used to generate the values of the random variable which can be in the range [-∞;∞]. Also the complexity of the mechanisms is discussed which can become a major problem if a large number of random numbers is required for the simulation. Thus, a typical question in the oral exam is which random variable generation mechanism is the best choice depending on the number of required random numbers and the probability distribution of the random variable? This type of questions allows a better estimation of the students' knowledge since the students have to anticipate what they have learned during the lectures. The second section of chapter 3 introduces a large number of discrete and continuous probability distributions. The students learn the characteristics of the distributions and how to generate random numbers according to these distributions. Most of the introduced distributions are also part of the exercises in which students have to integrate the distributions in the discrete event simulator.

The introduced mechanisms for random variable generation require uniform distributed pseudo random numbers. Therefore, various random number generators are discussed, e.g. linear congruential generator, shift register, general feedback shift register, and the very popular mersenne twister with its outstanding statistical properties. As part of the exercise, the students implement and evaluate statistic properties of the linear congruential generator and the shift register, e.g. autocorrelation and period length depending on the chosen seed and configuration. The quality of the introduced random number generators is evaluated in the lecture by applying the chi-square, spectral and serial tests. However, the implementation of these tests is not part of the exercise.

Chapter 4 is divided into three major parts: evaluation of simulation results, how to lie with statistics, and model validation. The first part, evaluation of simulation results, describes the characteristics of different estimators in terms of variance, consistency and bias. The estimators provide the basis for the introduction and calculation of confidence intervals. In this context, the central limit theorem and the t-distribution are discussed in order to outline the importance of doing multiple simulation runs for a single setup to generate meaningful results. Moreover, simulation techniques, like replicate-delete and batch means, are introduced. The second part of chapter 4 is motivated by the fact that statistical results are often (willingly or unwillingly) visualized with the intention to mislead the reader. Thus, the different examples are discussed with the students to increase their knowledge in reading and comparing statistics. The focus of the last section of chapter 4 lies on the validation and calibration of simulation models. The intention is to make students aware of the impact of structural and parameter changes. Furthermore, the problem of overfitting is discussed which is often neglected by authors of papers that are using simulation. Another very important topic is the comparison of different systems or configurations. During the development of a simulation model, the impact of structural or parameter changes have to be statistically evaluated. Therefore, the confidence intervals of the outcome have to be compared to ensure that the changes have an impact on the outcome with a certain statistical significance level. In this case, it is usually possible to apply a method introduced by Welsh which allows a comparison of the confidence intervals if the samples have the same variance and are of the same size. Especially the latter assumption is often not the case if simulation results are to be compared with measurements from test beds. Thus, a different comparison method has to be applied. For this reason, the more robust method from Law and Kelton is introduced which has proven to provide good results in most cases.

Another important topic is represented by experiment planning. Complex simulation models are typically affected by a large number of parameters which lead to a huge configuration space. However, linear regression, analysis of variance (ANOVA) and factorial design can help to mitigate the problem caused by the huge configuration space. Linear regression is introduced in the lecture since it is very popular due to the fact that it allows modelling the relationship between a variable and set of other variables which significantly reduces the configuration space. ANOVA is a collection of statistical models and their associated procedures which allow a partitioning of the variance of the observed variable into components of different sources of variation. It provides statistical tests of whether a group of means are all equal which helps to evaluate the impact of different variables. Factorial design is a strategy which helps focusing on the relevant simulation parameters. Typically a low value and a high value are chosen for each simulation parameter which reduces the configuration space to 2^k if k variable are evaluated. Therefore, it represents a very practical method for students at the beginning of their master thesis when they do not have great knowledge of the system or protocol that they are evaluating as part of their thesis.

Chapter 6 introduces the more advanced topic of parallel simulation. Different methods of deadlock avoidance and deadlock detection, as well as deadlock recovery, are discussed in the lecture. Moreover, the advantages and disadvantages of the optimistic time warp approach are part of the course. The limits of parallel simulation are discussed and alternatives are introduced. Due to the complexity of the topic, no exercise is planned for this chapter since it is considered to be too time consuming for a four ECTS course.

The evaluation and generation of synthetic mobility models represent the major topics of chapter 7. First of all, the philosophical question, what random mobility is, is discussed which has proven to be a very motivating starting point for this topic. After introducing typical human mobility pattern, the focus is shifted towards simulation issues for mobility models, e.g. shape of the simulation plane, bouncing rule, and obstacles. In the second part of chapter 7, characteristics of mobility pattern, e.g. duration of the transient phase, spatial node distribution, speed distribution, and link duration, are further discussed. Finally, the most popular synthetic mobility models are described and evaluated.

In chapter 8 point fields and random graphs are described. The focus of the first part is represented by point fields and their generation. Moreover, characteristics of point fields depending on the applied generation method and probability distributions are outlined. Random graphs are introduced in the second part of the chapter. Besides basic graph theoretic topics, characteristics of different random graph generation methods are introduced and evaluated. In this context, the small world phenomenon is introduced since it is a typical characteristic of scale-free graphs. The number of lectures that is required can vary depending on the number of students and the number of questions which are asked during the lecture. Note that we encourage students to frequently ask questions since it has proven that all students benefit from this kind of lecture.

TUTORIALS

Depending on the preferences of the students, a basic tutorial with Matlab or Octave is given. Both tools almost use the same syntax. Therefore, many scripts can be used in both programs with only minor changes. The idea of these tutorials is to make students familiar with state-of-the-art software in order to help them to visualize their results in a scientific way. Moreover, it has shown that the practical character of the tutorial helps students to understand complex statistical issues.

Besides the Matlab and Octave tutorials, which are intended to teach students how to visualize simulation results or measurements from test beds, we offer students the possibility to do a tutorial with the OPNET Modeler software. OPNET Modeler is a commercial discrete-event simulator which is optimized for analyzing and designing communication networks and protocols. However, the software is free for universities if they use it for research and teaching only. The software is able to simulate the whole ISO/OSI stack including a simplified physical layer. The software uses a hierarchically structure to modify or build a network. The top layer is represented by the network. On this layer it is possible to drag and drop existing objects, e.g. workstations, routers, switches, and sensor nodes, into the simulation scenario. Each object is described by a node model which consists of several process models. A process model is responsible for a certain task and may interact with other process models. The behaviour of a process model is defined by a finite state machine. The finite state machine consists of one or more states and transitions which can be used in a very intuitive way. The focus of the OPNET tutorial is based on the simulation of waiting queues on the one hand since students are already familiar with the functionality and the behaviour of waiting queues depending on the variance of the distributions of the arrival process and the service process. On the other hand, more complex scenarios, e.g. wireless communication, are simulated. The students have to implement a simple carrier-sense multiple access (CSMA) protocol which often leads to fruitful discussions regarding the simulation of wireless signal propagation and issues of parallel simulations.

CONCLUSION

Theoretical topics, like statistics, are usually not the primary choice for students since most of them prefer more practical topics where they get a direct feedback. Therefore, it is necessary to introduce theoretical topics in a practical context in order to motivate students. This becomes especially essential at a large university as the Technical University of Munich where lectures are in competition for drawing the students' attentions to a particular course due to the large number of courses. The combination of theoretical topics with practical exercises has proven to be very attractive to students.