Inhaltsverzeichnis

GRK 623:

Leistungsgarantien für Rechnersysteme

Saarbrücken

GK 623: "Performance Guarantees for Computer Systems"

Computer Science Department
Saarland University
Speaker: Prof. Raimund Seidel

`http://gk.cs.uni-sb.de`

The research program of this research training group aims at a comprehensive and mathematically rigorous understanding of the concept of "performance guarantee" in the sense of *predictable running time, provable correctness,* and *sufficient quality.*

Provable correctness is the guarantee that a desired functionality is indeed achieved, e.g. the collision avoidance of a moving geometric body in a virtual reality scene, or the achievement of specified properties of a simple reactive system or even of an operating system.

Predictable running time means that the resource usage of a program in terms its computing time, storage requirements, and accesses to the various levels of the memory hierarchy can be a priori quantified as exactly as possible. Beyond the design of efficient algorithms this aspect of performance also implies for instance guaranteed reaction times in real-time applications or anytime user-acceptable response times in information systems.

Sufficient quality is to mean that the achieved functionality is appropriate to the requirements of the application at hand. This aspect of performance guarantee can for instance refer to approximation quality in optimization problems, to accuracy in graphics computations, or to the relevance of the answer documents in a web query.

The research in this research training group draws on the interplay and conflicts between these three main aspects of performance guarantees. Simple procedures are easy to verify, but most of the time they are inefficient. On the other hand, efficient procedures are often complex and very difficult to verify. The comprehensive verifiction of the correctness of a system is often bound to fail due to the horrendous computational expenditure. But the restriction to the certification of some critical properties, e.g. the mutual exclusion of two trains using the same track section, can make the verification problem tractable. Computer science has developed a formidable repertoir of methods for the analysis of the worst case behaviour of single, stand-alone algorithms. But the situation is very different for the case of complex systems in which many algorithms interact and where the efficiency of the individual algorithms depends on environmental parameters such as data distributions, load profiles, or resource contention. Quantitative statements about the reaction or response times of such complex systems are only possible with severe simplifications and with restricted system functionality and quality. Finally, the quality of a computation or a search necessarily depends on the resource expenditure and therefore on the guaranteeable running time.

GK 623: "Leistungsgarantien für Rechnersysteme"

Fachrichtung Informatik
Universität des Saarlandes
Sprecher: Prof. Raimund Seidel

`http://gk.cs.uni-sb.de`

Das Forschungsprogramm dieses Graduiertenkollegs zielt auf einen umfassenden, mathematisch rigorosen Begriff von *Leistungsgarantien*, der die Aspekte der *vorhersagbaren Laufzeit*, *beweisbaren Korrektheit* und *hinreichenden Güte* umfasst.

Beweisbare Korrektheit ist die Garantie, dass eine gewünschte Funktionalität tatsächlich erbracht wird, beispielsweise die Kollisionsfreiheit eines bewegten geometrischen Körpers in einer Virtual-Reality-Szene oder die Erfüllung der spezifizierten Eigenschaften eines reaktiven Systems oder gar eines Betriebssystems.

Vorhersagbare Laufzeit bedeutet, dass der Ressourcenverbrauch eines Programms hinsichtlich Rechenzeit, Speicherplatz und Zugriffen auf Speicherhierarchiestufen a priori, absolut und so genau wie möglich quantifizierbar ist. Über die Entwicklung effizienter Algorithmen hinausgehend impliziert dieser Aspekt der Leistung also beispielsweise garantierte Reaktionszeiten in Echtzeitanwendungen oder jederzeit benutzerakzeptable Antwortzeiten bei Informationssystemen.

Hinreichende Güte bedeutet, dass die erbrachte Funktionalität den Erfordernissen der Anwendung entspricht. Dieser Aspekt einer Leistungsgarantie kann sich beispielsweise auf die Approximationsgüte eines Optimierungsproblems beziehen, oder auf die Berechnungsgenauigkeit bei Graphikanwendungen oder auf die Relevanz der Resultatdokumente einer Web-Anfrage.

Die Forschung im Graduiertenkolleg wird durch die vielfältigen Spannungen und Gegensätze zwischen diesen drei Leitbegriffen bestimmt. Einfache Verfahren sind leichter zu verifizieren, aber häufig ineffizient; umgekehrt sind effiziente Verfahren oft komplexer und damit schwerer zu verifizieren. Da eine umfassende Korrektheitsverifikation von Systemen oft am horrenden Rechenaufwand scheitert, kann die Beschränkung auf den Nachweis bestimmter kritischer Eigenschaften, z.B. den wechselseitigen Ausschluss zweier Züge im selben Gleisabschnitt, den Weg zur Praktikabilität ebnen. Hinsichtlich der Vorhersage von Programmlaufzeiten kann die Informatik auf ein reiches Repertoire an Worst-Case-Analysen von Algorithmen verweisen. Darüber hinausgehende quantitative Aussagen über die Antwort- und Reaktionszeiten komplexer Systeme, in denen viele Algorithmen zusammenspielen und in ihrer Effizienz von Umgebungsparametern wie zum Beispiel Datenverteilungen, Lastprofilen oder Ressourcenkonkurrenz abhängen, sind dagegen nur mit Vereinfachungen und Einschränkungen der Systemfunktionalität und -güte möglich. Die Güte einer Berechnung oder Suche schließlich hängt oft zwingend vom Ressourcenaufwand und damit der garantierbaren Laufzeit ab.

Formal and Pervasive Verification:
From Hardware to Operating Systems

Eyad Alkassar

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
eyad@alkassar.de

In todays world we rely in nearly all security-relevant spheres of our daily life on computer systems. From simple circuits in traffic lights to complex control software of nuclear power plants we *believe* in the infallibility of computers. But who or what ensures that these systems really do what they are intended to do? One way would be to test all possible cases, or at least a set of representative ones. With a complexity of programs increasing in a exponential way, this approach is futile if one wants to ensure that no bugs are left undetected.

The alternative to testing is proving correctness in a mathematical sense, by first formulating accurate models of the real-world system and then verifying formal assertions over these models. This approach was pioneered by, among others, Dijkstra, Flyod, Lamport and Hoare. J. S. Moore, principal researcher of the CLI stack project, declares the formal verification of a practical computing system from transistors to software as a grand challenge problem.

A main goal of the Verisoft project[1] is to bear this challenge. In the academic system, a subproject of the Verisoft project, a general-purpose computer system, covering all layers from the gate-level hardware description to communicating concurrent programs is designed, implemented and verified. The aim is to build accurate formal models without hidden assumptions, that are modular and easily extendable. The verification should be pervasive throughout all layers of abstraction and take advantage of computer-aided verification tools.

Currently we are engaged in the verification of a small page-fault handler, written in a C-like language with inline Assembler code. A page-fault handler realizes process virtualization and hence is an integral part of any modern operating system. The correctness of the page-fault handler depends a.o. on the correctness of of the compiler and the correctness of the underlying hardware, in particular the memory management unit and the interrupt handling of page faults. Sometimes the page fault handler has to swap out data from the main memory to the hard disk. Hence, also the hard disk and its interaction with the processor on gate-level as on specification-level has to be modeled. Furthermore an elementary hard disk driver written in assembler was proven to be correct.

This result will hopefully enable us soon to verify a small, but complete OS kernel.

---

[1] www.verisoft.de

Realtime Ray Tracing Techniques for
Triangular Dynamic Meshes and Volumetric Datasets

Heiko Fridrich

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`friedrich@graphics.cs.uni-sb.de`

Ray tracing is a well known image synthesis technique, for rendering 2D images from 3D scenes. It solves many of todays problems that are inherent in the rendering algorithms implemented in current graphics boards. These problems include optical correct computations of shadows, reflections, and photo realistic images. Unfortunately ray tracing has been considered as too slow for realtime applications and thus only little research has been done. In recent years due to new algorithms and optimized implementations the ideas of ray tracing sparked a new interesest in the research community since realtime rendering has become reality.

1   Realtime Ray Tracing Techniques for Triangular Dynamic Meshes

Ray tracing requires for realtime performance special index structures (IS) that allow for efficient visibility queries. A visibility query answers for a given point and direction what can be seen next. Unfortunately the construction of these data structures is a computational expensive task and all realtime rendering effort is vain endeavor if we have to rebuilt these IS for each new frame when objects deform. In our research we found new IS and algorithms that allow to ray trace animated sequences without the need to rebuilt the IS. The basic idea is to analyse in a preprocess the sequence of the deforming objects, compute a set of cluster with coherent motion, substract the coherent motion from the cluster and finally capture the residual motion in a special IS. Our research in conjunction with other work at the UdS Graphics Chair, MPII AG4, and the University of Utah solves the problem of ray tracing dyamic scenes to a large extend.

2   Realtime Ray Tracing Techniques for Volumetric Datasets

My second research interest focuses on the use of ray tracing techniques for volumetric datasets. Volumetric datasets do not only describe a single surface but a complete volume. Most of these datasets originate from scanners like CT or numerical simulations and consists of points in space that describe a physical property. In our research we focus on the realtime rendering of surfaces and semi-transparent images with just a single IS. In past for each render mode (surface/semi-transparent) a different IS had to be exploited for an efficient rendering. Another problem is that volume datasets tend to be very large in size with up to several gigabytes. A part of our research aims to render such massive datasets on commodity PCs with very compact representations of the necessary data structues and optimized out-of-core memory techniques. Finally we try to minimize the enormous memory-bandwith that is inherent in volume rendering by finding appropriate linear functions of the volumetric space.

Static Analysis of Caches –
Performance, Predictability, Price

Daniel Grund

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`grund@cs.uni-sb.de`

Our work addresses the use of processor caches in hard real-time systems. Embedded systems as they occur in application domains such as automotive, aeronautics, and industrial automation often have to satisfy hard real-time constraints. Off-line guarantees on the worst-case execution time of each task have to be derived using safe methods. Such methods must be conservative, i.e. they must statically overapproximate the dynamic behaviour of a task on all possible inputs and hardware states. Caches, deep pipelines, and all kinds of speculation are increasingly used in such systems to improve average-case performance. But at the same time they increase the variability of execution times of instructions due to the possibility of timing accidents with high penalties: a cache miss may take 100 times as long as a cache hit.

Average-case *performance* of a cache can be measured by its miss-rate, the percentage of accesses where the requested data is not stored in the cache. The notion of *predictability* captures how good a system lends itself to static analyses and how tight the obtained bounds (overapproximations) can be determined. For caches the cache replacement policy (that determines which element to replace upon a cache miss) is dominating the predictability of a cache. And of course the *price* of a policy in terms of hardware implementation costs differs from case to case. Here the number of status bits that are necessary to keep track of the replacement order is a very good measure.

We defined a large class of cache policies (including the most prominent ones) where a policy can be modeled by an ordered set of permutations. Using this representation we are able to compute values for the above metrics and estimates for the performance that depends on a characterization of the executed software.

Computing values for the predictability metrics reduces to *shortest* and *longest path* problems. Deriving the number of necessary status bits essentially is the problem of determining the *group order* of a *permutation group*. To estimate the miss-rates we use a *markov chain model* where the *steady-state* probabilities have to be determined.

All these computations are (or will be) integrated into a *systematic search procedure* to find policies meeting given requirements. The huge number of policies in the class we conider demands several *optimizations*: e.g. determining a bound on the value of a metric even though the set of permutations is only partially known. Or the iterative construction and solution of the markov model that prevents "unnecessary" states to be considered.

Theoretical Analysis of Evolutionary Algorithms

Edda Happ
GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`edda@mpi-inf.mpg.de`

Currently, I work in the field of theoretical analysis of evolutionary algorithms. The idea behind evolutionary algorithms is to use principles inspired by evolution to create algorithms for optimization problems. The principles of evolution are:

1. There exists at any time a number of species on earth.

2. The species change over time, either by mutation or by recombination. Mutation means that through random changes the individual slowly evolves into something different, while recombination (or crossover) is the combination of the genetic material of two different species into a new evolving species.

3. According to how good the species is adapted to live in its environment, it either survives or it dies out. This is Darwin's principle of the "survival of the fittest".

Taking these principles and using them for optimization algorithms translates as follows:

1. Keep a population of candidate solutions.

2. Do either mutation or crossover on these candidate solutions to generate new candidate solutions.

3. According on the *fitness* of the new candidate solutions, they can replace some of the old solutions.

In creating such an evolutionary algorithm, the important tasks are to find a good representation of the candidate solutions, to devise reasonable mutation and crossover operators, and to decide on a selection method appropriate to the problem.

Since evolutionary algorithms are easy to design (even without deep knowledge of the problem to be solved), easy to implement, and highly reusable, they have been implemented and analytically investigated a lot. However, the theoretical analysis of these algorithms is not as sophisticated yet. It involves elaborate probability theory and is still in its infancy. My current research includes a tight analysis of an evolutionary algorithm for the all-pairs shortest paths problem as well as for the single-source shortest path problem.

Hardness of Algebraical Problems

Christian Hoffmann
GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`hoffmann@cs.uni-sb.de`

Algebraic Complexity Theory aims at classifying the hardness of arithmetic or algebraical problems. A common formalization of such problems and the corresponding algorithms are *multivariate polynomials*: in the beginning only the input (indeterminates $X_1, \ldots, X_n$) is available and each arithmetic operation $(+, -, \cdot, \ldots)$ produces a new polynomial from two of the already available polynomials.

We study models computing multivariate polynomials such as arithmetic circuits, arithmetic formulas and also commutative algebraic branching programs (ABPs). An important question is: what size must an arithmetic circuit (arithmetic formula, ABP resp.) have to compute some specific given polynomial? The search for lower bounds of this kind is a driving problem for the whole field. We separated the ABP model from the arithmetic circuit model giving an $\Omega(n^2)$ bound for a polynomial for which arithmetic circuits of size $O(n \log n)$ are known. An outstanding open problem is to give *superpolynomial* lower bounds for any general model.

Polynomial identity testing is another important problem of our research. It asks whether the polynomial computed by some arithmetic circuit is the zero polynomial. Whereas randomized polynomial time algorithms are available for this problem, all known deterministic algorithms take exponential time.

Recent research has revealed a connection between polynomial identity testing and superpolynomial lower bounds in algebraic models: If we had a superpolynomial lower bound on the arithmetic circuit size of some multilinear polynomial, we could deterministically test for polynomial identity in subexponential time. This result is true for the arithmetic circuit model and heavily uses the power of this model. Giving a polynomial size construction for division in the ABP model, we were able to establish a similar result for the weaker ABP model: if we had a superpolynomial lower bound on the ABP size of some multilinear polynomial, we could deterministically in subexponential time solve the polynomial identity testing problem for multilinear polynomials given by ABPs.

Graph polynomials encode much information about the underlying graph. Another topic of our current research is to study how hard it is to evaluate these polynomials. One approach is to find a point which gives a well known hard-to-compute property of the graph (for example the number of 3-colorings) and to reduce the evaluation at other points to the evaluation at this hard-to-evaluate point.

Exploiting User Search Behavior for Information Retrieval

Julia Luxenburger
GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`julialux@mpi-inf.mpg.de`

The analysis of observed user search and browsing behaviour is a valuable information source for many aspects of web search result ranking. From the monitoring of user interactions with a search engine we are able to draw conclusions of different flavor:

- The sequence of queries, a user subsequently poses, allows us to group related queries serving the same information need and learn query reformulation patterns.

- The query-result pages which were clicked on and the ones which were not clicked on after a user saw the summary snippets of the top-10 results, lead us to inferences on the relevance, respectively irrelevance, of result pages to their corresponding queries, as well as to inferences on the general quality of these pages.

- The analysis of complete user search sessions enables us to identify frequent user interaction patterns, as well as deficiencies of state-of-the-art web search[1].

One focus of our work in this area is the incorporation of implicit user feedback into Web link analysis which constitutes an important ranking feature. State-of-the-art authority analysis methods on the Web linkage graph such as the PageRank [2] algorithm are based on the assumption that a web page author endorses a Web page when creating a hyperlink to that page. This kind of intellectual user input can be generalized to a user endorsing a query-result page when visiting that page, and moreover disapproving a result page when prefering a lower-ranked result page.

We study link analysis methods[3] that enhance PageRank by incorporating additional user assessments based on query logs and click streams, including negative feedback when a query-result page does not satisfy the user demand or is even perceived as spam. Our methods use various novel forms of Markov, respectively Markov reward, models whose states correspond to users and queries in addition to Web pages and whose links also reflect the relationships derived from query-result clicks, query refinements, and explicit ratings.

Our experiments, based on real-life query-log and click-stream traces on an excerpt of the English version of the Wikipedia encyclopedia, indicate the potential of our methods.

---

[1]N. Kammenhuber, A. Feldmann, J. Luxenburger, and G. Weikum: **Web Search Clickstreams** IMC(2006).

[2]S. Brin and L. Page: **The Anatomy of a Large-Scale Hypertextual Web Search Engine** WWW(1998).

[3]J. Luxenburger and G. Weikum: **Exploiting Community Behavior for Enhanced Link Analysis and Web Search** WebDB(2006).

Prediction and Alignment of RNA Secondary Structures

Mathias Möhl

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`mmohl@ps.uni-sb.de`

The genetic information of living organisms is encoded in deoxyribonucleic acid (DNA). In order to 'read' this information, the organism builds ribonucleic acids (RNAs) that use the DNA as a template. These RNA molecules are then either translated into proteins or perform their own individual function within the cell. Similar to DNA, an RNA molecule can be abstracted as a sequence of bases. While in the case of DNA this sequence represents the entire information captured in the molecule, the function of RNA does not solely depend on the sequence itself. While DNA molecules form double-stranded helices, the single-stranded RNA molecules form complex secondary structures (so to say 'shapes').

Since the secondary structure determines the function of an RNA molecule to a large extend, there is an active field of research in bioinformatics that is concerned with RNA secondary structures. While determining the base sequence is a well-established method in biology, determining the secondary structure with experiments is much more difficult. Therefore, several algorithms have been developed, that predict for a given RNA sequence likely secondary structures. Another important task is the alignment of RNA structures, that is, the identification of similar regions among a set of given RNAs. Also of practical interest is a combined form of structure prediction and alignment that predicts similar secondary structures for a set of RNA sequences and aligns them at the same time.

Most existing algorithms in this field are limited to structures that do not contain so-called pseudoknots. This restriction allows the use of efficient dynamic programming algorithms since pseudoknot-free structures can recursively be decomposed into smaller fragments. While for arbitrary pseudoknots, structure prediction has been shown to be NP-hard for many cases, with more complex recursion schemes it is also possible to admit certain limited kinds of pseudoknots[1]. As an alternative I explore in my PhD the development of algorithms that are not limited in the kind of pseudoknots, but in the number of pseudoknots that they can handle. This decision is motivated by the fact that no biologically motivated restrictions on the kinds of pseudoknots occurring in nature are known, whereas in general practice, pseudoknots occur rarely. Currently I develop an algorithm for the alignment of RNAs with arbitrary pseudoknots whose worst case complexity is exponential, but that is fixed-parameter tractable for a parameter that only depends on the number and size of the pseudoknots. My long term goal is to develop similar algorithms for structure prediction and the combined version of both problems.

---

[1] AKUTSU, T.: Dynamic programming algorithms for RNA secondary structure prediction with pseudoknots
Discrete Applied Mathematics, 2000, 104, 45-62

Information Extraction for Ontology Learning

Fabian M. Suchanek

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`suchanek@mpi-inf.mpg.de`

Today's search engines are very successful at finding Web pages that contain certain keywords. For example, it is very easy to find a Web page about *London*, the capital of the United Kingdom. However, when it comes to more complex queries, these search engines fail. For example, it is close to impossible to ask Google how many other cities in the world are also called *London*. To answer these queries, we need a huge structure of world knowledge – an *ontology*. Technically, an ontology is a graph, where the nodes are *entities* (e.g. *London/UK* or the concept of a *city*) and the edges are semantic relations between them (e.g. *locatedIn* or *isA*). The goal of my PhD is to create a huge high-quality ontology.

As a first step, we have extracted ontological knowledge from the large online encyclopedia Wikipedia. Unlike previous approaches, our approach exploits the category system of Wikipedia. For example, the page about *London/UK* is in the category *Cities in the UK*. This tells us (1) that *London/UK* is a *city* and (2) that *London/UK* is located in the *United Kingdom*. By sophisticated heuristics, we have been able to combine these ontological data with data from WordNet, the semantic lexicon of the English language. The result is YAGO[1], the largest formal ontology available today. The semantics of YAGO is given by logical axioms. We have shown that the consistency of YAGO is decidable, that its deductive closure is unique and finite and that its canonical base (the smallest equivalent sub-ontology) is also unique and finite. YAGO knows 4 cities called *London*.

As a second step, we focused on extracting knowledge from natural language text documents. Our system, LEILA[2], is given a semantic relationship (such as e.g. *locatedIn*) and a corpus of documents. In a first phase, it finds text patterns that express the semantic relationship. For example, the pattern *X is located in Y* expresses the *locatedIn* relationship. In a second phase, LEILA generalizes these patterns by machine learning techniques. In a third phase, it finds instances of the generalized patterns in the corpus and extracts new pairs of entities that stand in the relationship. Different from previous approaches, LEILA uses deep syntactic patterns instead of surface text patterns, which makes it more robust to variation of the patterns. Thereby, LEILA consistently outperforms previous approaches.

In a third step, LEILA and YAGO shall be combined in a feedback loop: LEILA shall add new facts to YAGO. In turn, YAGO shall help LEILA to extract new knowledge.

---

[1] FABIAN M. SUCHANEK, GJERGJI KASNECI, GERHARD WEIKUM:
YAGO – A Core of Semantic knowledge. WWW 2007.
See the Web interface at `http://www.mpi-inf.mpg.de/yago`
[2] FABIAN M. SUCHANEK, GEORGIANA IFRIM, GERHARD WEIKUM:
Combining linguistic and statistical analysis to extract relations from Web documents. KDD 2006.

Polyhedral Vertex Enumeration

Hans Raj Tiwary

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`hansraj@cs.uni-sb.de`

As part of GK, my research has concentrated primarily on Polyhedral Vertex Enumeration with an occasional detour into unrelated short-term problems. The main object of my research - a polytope - can be described as the convex hull of a minimal finite set $\mathcal{V}$ of points (vertices of the polytope) in $\mathbb{R}^d$ or as intersection of a minimal finite set of halfspaces (facets) $\mathcal{H}$, each described by an inequality. The problem of vertex enumeration asks one to enumerate all the elements of $\mathcal{V}$ given the elements of $\mathcal{H}$. The number of vertices can be anywhere between $\Omega(|\mathcal{H}|^{\frac{2}{d}})$ and $O(|\mathcal{H}|^{\lfloor \frac{d}{2} \rfloor})$. This wide range of output size suggests that we should look for algorithms that are polynomial in the input and output size of the problem. In case of non-degenerate input, *i.e.* polytopes for which none of the vertices are contained in more than d facets or none of the facets contain more than d vertices, output sensitive polynomial algorithms exist but degeneracy in input is hard to handle. Bad examples are known for all the existing algorithms that handle degeneracy.

A problem, somewhat related to Vertex Enumeration, is enumerating the facets of the Minkowski Sum of two polytopes given by their facets. This problem comes up frequently in computational geometry and can be used as a subroutine to enumerate the vertices of a polytope. One of my recent results is that it is not possible to enumerate the facets of the Minkowski sum of two $\mathcal{H}$-polytopes in polynomial time unless $P = NP$. This work has been accepted for publication in the proceedings of 23rd Symposium on Computational Geometry 2007[1]. Currently I am studying certain geometric properties of polytopes and their polar dual. In particular I am studying the sum of solid angles at the vertices of a polytope and how they relate to the sum of solid angles at the vertices of the polar dual.

I have also worked on using algebraic techniques like Fast Fourier Transform (FFT) to solve certain geometric questions like computing the centroid of the vertices of an arrangement of lines. The work has been accepted for publication in Workshop on Algorithms and Data Structures 2007[2]. The use of non-geometric techniques to solve geometric problems is interesting in itself and might find use in solving more interesting problems like 3-SUM where given a set of numbers one would like to quickly (in subquadratic time) determine whether any three numbers sum to zero.

---

[1] HANS RAJ TIWARY: On the Hardness of Minkowski Addition and Related Operations. *Accepted for publication in Proceedings of 23rd Symposium on Computational Geometry, 2007*

[2] DEEPAK AJWANI, SAURABH RAY, RAIMUND SEIDEL, HANS RAJ TIWARY: On Computing the Centroid of the Vertices of an Arrangement and Related Problems. *Accepted in Workshop on Algorithms and Data Structures, 2007*

Approximation and I/O efficient algorithms

Vitaly Osipov
GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`osipov@mpi-inf.mpg.de`

Two research areas which are of particular interest for me are approximation algorithms and algorithms for large data sets.

As for approximation algorithms I continue working on the problem which was partially solved in my master thesis[1], that is a Maximum Weight Planar Subgraph problem. Since this problem is *NP-hard* we are interested in a polynomial approximate solution. A general idea of the algorithm is to construct a subgraph of a given graph, whose cycles (if any) have length at most three, i.e. a *triangular structure*. Observe that such graph is necessary planar. The former approach greedily constructs such a subgraph producing a suboptimal solution. We found a new algorithm that produces such triangular structure of almost maximum weight. We also analyzed our algorithm and proved that in several special cases, most notably in the case when the maximum weight planar subgraph is outerplanar, our algorithm is considerably better then the former approach. The analysis of the algorithm in general case is still open, and of great interest to me.

Algorithms for large data sets are based on a different computational model. Instead of minimizing the running time of the algorithm in terms of the number of instructions as in *RAM* model, we concentrate on minimizing the number of input/output operations between the levels of memory hierarchy, so called *External Memory* model. In this area we designed, implemented and conducted experimental study of I/O efficient *Breadth First Search* algorithm[2,3]. As a part of the system we use *pipelining*, a technique originated from the database community. With the growth of popularity of multicore processors we are interested in extending our sequential approach to the framework of parallel algorithms. The problems one face there include scheduling, high scalability, fast response and dynamic adaptation to the changes of the system. I also hope that the approach we develop in the external memory settings will also have an impact on the original database applications.

---

[1]"A Polynomial Time Randomized Parallel Approximation Algorithm for Finding Heavy Planar Subgraphs" Master thesis under supervision of Prof. Dr. Markus Blaeser, University of Saarland, 2006

[2]Deepak Ajwani, Ulrich Meyer, Vitaly Osipov "Improved external memory BFS implementations" Workshop on Algorithm engineering and experiments (ALENEX 07), New Orleans, USA, 2007. Also accepted at DIMACS implementation challenge on shortest path, Piscataway, NJ, USA, 2006

[3]Deepak Ajwani, Ulrich Meyer, Vitaly Osipov "Breadth first search on massive graphs", to appear in DIMACS Book Series

Shape Recovery from Shading Information

Oliver Vogel

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`vogel@mia.uni-saarland.de`

Computer Vision is a discipline of Computer Science. Roughly, the goal herein is to teach computers to "see". One major area of Computer Vision is retrieving shape information from images, i.e. finding the right 3D model to match an input image.

This is my area of study. Currently I am focused on so-called "shape-from-shading" methods, which rely on the shading information of one single image as input. There are many more possibilities of recovering shape information, such as using multiple images from different viewpoints (Shape from Stereo), different lighting of a scene (Photometric Stereo), or using texture information (Shape from Texture).
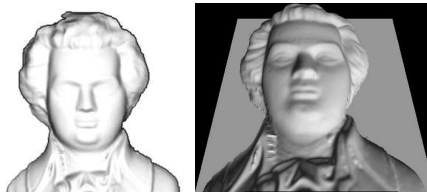

Figure 1: Shape from Shading: The Mozart input image (left) and the corresponding shape (right).

Existing shape-from-shading methods[1][2] suffer from several problems: most of them make simplifying assumptions that are either unrealistic or very limiting, many are very slow and complicatied, and some do not work at all. Current state-of-the-art algorithms[3] are limited to certain lighting conditions and Lambertian reflectance properties of surfaces. I specialize in methods based on the calculus of variations and partial differential equations. In my recent work, I managed to overcome several issues with classic algorithms [4].

My main focus of research lies in improving the accuracy and computation time of algorithms while keeping it as simple as possible. A further step will be to extend the model to handle more natural and general settings. To this end, it might be feasible to combine various shape-from-X techniques to achieve high quality results.

[1] B.K.P. HORN AND M.J. BROOKS: **Shape from Shading.**, MIT Press, 1989

[2] R. ZHANG, P. TSAI, J.E. CRYER, AND M. SHAH: **Shape from Shading: A Survey**, IEEE Transactions on Pattern Analysis and Machine Intelligence 21(8): 690-706 (1999)

[3] E. PRADOS AND O. FAUGERAS: **Shape from Shading: A Well-Posed Problem?**, IEEE Conference on Computer Vision and Pattern Recognition: 870-877 (2005)

[4] O. VOGEL, A. BRUHN, J. WEICKERT, AND S. DIDAS: **Direct Shape-from-Shading with Adaptive Higher Order Regularisation.**, In Scale-Space and Variational Methods in Computer Vision SSVM 2007, Ischia, Italy, May/June 2007 - Lecture Notes in Computer Science, Springer, Berlin, accepted for publication.

Program Behavior Analysis

Andrzej Wasylkowski

GK 623 *"Quality Guarantees for Computer Systems"*
Saarland University
`wasylkowski@cs.uni-sb.de`

Programs usually follow many implicit programming rules or patterns. Programmers maintaining a program are typically not aware of all such patterns and thus introduce defects that have to be later corrected. Making them aware of patterns and places in program code that violate them can help them write correct code and fix already existing defects.

My research focuses on patterns representing sequencing of method calls in JAVA programs, such as "`hasNext()` is called before `next()`"[1]. I discover such patterns by statically analyzing program code and producing so-called *object usage models*. These models are finite state automata with anonymous states and transitions labeled with method calls. An example of such a model for an iterator object is shown in Figure **??**.
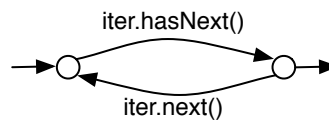

Figure 1: Typical iterator model.

Each model represents one *abstract object*, such as a method parameter or an object created via `new` and so on. Based on the models, we can discover sequencing constraints on method calls, such as the one expressed above. Frequently occurring constraints become patterns and methods that violate them are reported as likely to be defective.

My approach is effective and scales to industrial-sized applications. Mining models from ASPECTJ[2], which is an aspect-oriented extension to the JAVA programming language with over 36,000 methods defined in almost 3,000 classes, produces over 250,000 models in less than 14 minutes. Finding patterns stemming from those models and methods that violate them takes less than 4 minutes. I have applied the method outlined above to five open-source projects and found five previously unknown bugs along with over thirty suggestions that improve code quality.

I expect object usage models to be also helpful in explaining to programmers how to use classes they are new to and plan to perform user studies to verify this expectation. I also want to investigate ways to mine and put to use more complicated patterns that the ones I am currently using.

---

[1] A. WASYLKOWSKI: **Mining Object Usage Models**. Doctoral Symposium, 29th International Conference on Software Engineering (ICSE 2007)

[2] http://www.eclipse.org/aspectj/

DRPU: A Programmable Hardware Architecture for Real-time Ray Tracing of Coherent Dynamic Scenes

Sven Woop

GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`woop@cs.uni-sb.de`

Ray tracing is a rendering technique capable of generating high quality photo-realistic images of three dimensional scenes. Highly accurate reflections and refraction effects can be computed by following the paths of light backward from the camera towards the light sources. Rendering speed of ray tracing algorithms has been an issue for a long period, but recently high performance software implementations have made real-time ray tracing possible. But reaching performance levels comparable to rasterization also requires dedicated hardware solutions.

During my research I developed the DRPU architecture (Dynamic Ray Processing Unit) as the first programmable ray tracing hardware design for coherent dynamic scenes. For programmable shading it contains a shading processor that achieves a high level of efficiency due to SIMD processing of floating point vectors, massive multi-threading, and synchronous execution of packets of threads. A dedicated traversal and intersection unit allow for efficient ray casting even in highly dynamic scenes by using B-KD Trees as spatial index structure - a kind of Bounding Volume Hierarchy. The basic idea here is to build the B-KD Tree only once before the animation. On changed geometry some bounds stored in the data structure are recomputed by a special Update Processor but the structure of the tree is maintained. A Skinning Processor is used to compute dynamic scene changes by applying a matrix based skinning model.

I did an FPGA prototype implementation of the architecture specified in my own hardware description language HWML. The prototype achieves performance levels comparable to commodity CPUs even though clocked at a 50 times lower frequency of 66 MHz. I mapped the prototype to a 130nm CMOS ASIC process that allows precise post layout performance estimates. We also did some extrapolations of the results to a 90nm version with similar hardware complexity to current GPUs. It shows that with a similar amount of hardware resources frame-rates of 80 to 280 frames per second would be possible even with complex shading at 1024x768 resolution. This would form a good basis for game play and other real-time applications.

For more information about my research I refer to my publications at *http://woop.dnsalias.com/publications.html*.

Mining and Predicting Development Activities

Thomas Zimmermann
GK 623 *"Performance Guarantees for Computer Systems"*
Saarland University
`tz@acm.org`

My main research interest is software engineering, with a focus on improving programmer productivity. I develop techniques and tools that make both managers and developers aware of history: *learning from past successes and failures, helps us create better software.* My research activities cover program analysis, aspect-oriented programming, empirical studies, and in particular, the analysis of software archives.

A common theme in my research is *data mining*. Software development produces huge amount of data such as program versions, bug reports, documentation, and electronic discussions. I applied pattern mining to find related locations and method usage patterns; with concept analysis I identified cross-cutting concerns.

The ability to analyze *systems of any scale* is important in software engineering; all of my work is thoroughly evaluated and scales to projects such as Eclipse, one of the largest open source projects with several hundred developers and million lines of code. I analyzed the bug databases of Microsoft Windows Server 2003 to predict failure-prone components.

Mining Change Histories. My eROSE tool learns from history and navigates developers through source code with recommendations such as "Programmers who changed function f() also changed function g()."[1] My DynaMine tool mines *project-specific usage pattern of methods* from version histories: "Developers first call begin(), then insert for several times, and finally end() on objects of class Operation."[2] I contributed to aspect-oriented programming with my HAM tool that reveals *cross-cutting changes*: "A developer inserted calls to lock() and unlock() into 1284 different locations."[3]

Mining Bug Databases. In software development, the resources for quality assurance (QA) are typically limited. A common practice among managers is to allocate most of the QA effort to those parts of a system that are expected to have most failures. In recent empirical studies I developed prediction models to support managers in this task of *resource allocation*.[4]

---

[1] T. ZIMMERMANN, P. WEIŠGERBER, S. DIEHL, AND A. ZELLER: Mining version histories to guide software changes. *Int. Conference on Software Engineering (ICSE), pages 563–572, May 2004.*

[2] V. B. LIVSHITS AND T. ZIMMERMANN: Dynamine: Finding common error patterns by mining software revision histories. *European Software Engineering Conference/Int. Symposium on Foundations of Software Engineering (ESEC/FSE), pages 296–305, September 2005.*

[3] S. BREU AND T. ZIMMERMANN: Mining aspects from version history. *Int. Conference on Automated Software Engineering (ASE), pages 221–230, 2006.*

[4] A. SCHRÖTER, T. ZIMMERMANN, AND A. ZELLER: Predicting component failures at design time. *Int. Symposium on Empirical Software Engineering (ISESE), pages 18–27, September 2006.*

GRK 643:

Software für mobile Kommunikationssysteme

Aachen

Graduiertenkolleg 643:
Software für mobile Kommunikationssysteme

Aachen

Der Informationsaustausch über unterschiedliche Netztypen findet ständig neue Nutzungsmöglichkeiten und -formen. Eine besondere Bedeutung dabei hat die Mobilkommunikation (vorwiegend auf Basis von drahtlosen Netzen) erhalten, die durch höhere Bandbreiten, durch leistungsfähigere Standards und durch Bereitstellung einer wachsenden Vielfalt von Diensten ständig steigende Nutzung erfährt. Neue Technologien und Dienste sind unverzichtbare Grundlage für die Akzeptanz neuer Anwendungen. Zu den wichtigsten Anwendungsbereichen gehören elektronischer Geschäftsverkehr, Verkehrstelematik, neue Arbeitsformen sowie Konzepte wie 'e-learning', 'e-government' oder 'e-home'. Während bisher vorwiegend das Machbare die Entwicklung bestimmte, werden künftig die Anforderungen der Nutzer im Vordergrund stehen. Es zeichnet sich zudem ab, dass die Mobilität der Benutzer zunimmt und dass drahtlose Netze in zunehmendem Umfang als Infrastruktur dienen werden. Die Funktionalität der Anwendungen soll aber weitgehend unabhängig von der Netzinfrastruktur sein, d.h. für drahtlose Netze erwartet der Nutzer eine ähnlich gute Qualität wie für drahtgebundene Netze. Die Verwirklichung dieses Ziels erfordert erheblichen Forschungsaufwand.

Die Arbeiten im Rahmen des Graduiertenkollegs konzentrieren sich auf 'Anwendungen', 'Middleware und neue Dienste' sowie 'Netzwerk-Infrastrukturen'. Neue Anwendungen werden spezifiziert, implementiert und getestet. Middleware-Architekturen und zugehörige neue Dienste werden so entworfen, dass sie einerseits die Anforderungen der Anwender zufrieden stellend erfüllen und sich andererseits an die Fähigkeiten unterschiedlicher Kommunikationsarchitekturen und -infrastrukturen anpassen.

Es werden Anwendungen und Dienstkonzepte betrachtet, für die sich drahtlose Netze besonders eignen. Außerdem werden Verfahren entwickelt und untersucht, die den Datenaustausch über drahtlose Netze effizienter und zuverlässiger gestalten.

New ways of exchanging data through different types of networks are continuously evolving. This includes particularly mobile communication, which has by now become more or less ubiquitous, thanks to an ever increasing level of user acceptance. This, in turn, is largely due to improved standards and a vastly expanded variety of services. Future application developments will more and more be based on user requirements rather than just feasibility. Here, the most important domains include electronic commerce, transport telematics, new forms of working, and other innovative concepts such as e-learning, egovernment, and e-home. For these new applications to be acceptable, new underlying technologies and services are a sine-qua-non. Moreover, the increasing mobility of users will lead to an equally increasing role wireless networks will play in communication infrastructures.

Yet, users expect application functionalities to be independent from the underlying communication network. Considerable additional research will be needed to actually meet this requirement.

Within the framework of the Graduate School research will focus on three areas: 'Applications', 'Middleware and New Services' and 'Network Infrastructure'. New applications will be specified, implemented, and tested. Middleware architectures and associated new services will on the one hand be designed to meet user and application requirements. On the other hand, they will be capable of adapting to different underlying communication architectures and infrastructures.
Work will focus on those applications and services which lend themselves to access via wireless networks. Mechanisms to improve efficiency and reliability will be designed and evaluated.

Alle 14 Stipendien sind im Augenblick (Stand: Mai 07) mit Kandidaten besetzt, die sich durch besondere Leistungen auszeichnen. Weiterhin gehören 13 Kollegiaten dazu, die mit ihren Beiträgen das Graduiertenkolleg erweitern. Unterstüzt wird die Arbeit durch 14 Forschungsstudenten, die sich ebenfalls mit dem Thema „Software für mobile Kommunikationssysteme" beschäftigen. Wir sind stolz, das GK international besetzen zu können; die Mitglieder kommen aus vielen Teilen der Welt, Indien, Mexiko, Türkei, Russland, Pakistan, Thailand.

Zum Betreuerkreis gehören 10 Professoren der Hochschule. In regelmäßigen Abständen werden Gastwissenschaftler zum Erfahrungsaustausch aus dem In- und Ausland eingeladen

Securing Privacy of Mobile eHome Users

Ibrahim Armac

`armac@i3.informatik.rwth-aachen.de`

GRK 643: Software for Mobile Communication Systems

Department of Computer Science, Informatik 3, RWTH Aachen University

Smart environments provide value-added software services to the users by combining the functionality of multiple devices. Smart home environments are also known as *eHomes*. In many cases, the hardware components are available at reasonable costs. However, to offer value-added eHome services that rely on the functionalities of the individual appliances, high level software is needed. The existing eHome solutions in the market are proprietary. Thus, they are not usable in a broad sense. In our group, we try to enable low-cost eHome systems by developing component-based services that can be reused for different eHomes. Instead of developing individual solutions for each eHome, the services are just configured and by this adapted to the specific eHome before installation. This avoids individual eHome-specific implementation of services, which is obviously too cost-intensive.

Considering mobile users visiting multiple environments in daily life, I am working on the migration of context-based services between different smart environments, such as homes, hotels, work places, public places, etc. This makes it possible to enable the visited environments to act in a personalized manner: The environment provides users functionalities like in their home environment. Thereby, I do not consider specific service implementations to be migrated with the user like in agent technology. Instead, the user only needs to indicate which functionalities he would like to take along, such as `Personal Illumination Control` or `Music Follows Person`. Having the wished functionalities, the visited environment can provide individual instances of the corresponding services. In this scenario it is assumed that the visited environment runs (or can install on demand) services that realize the relevant functionalities. Another assumption is that every user possesses a mobile device, such as a PDA or mobile phone. On the one hand, all necessary personalization information, i.e. the *user profile* (wished functionalities, preferences, etc.), can be stored on this device. On the other hand, this device can provide a uniform interaction interface to the user in different environments.

An important issue in supporting mobile users in smart environments is the protection of *user privacy*. My solution is a context- and policy-based profile management that protects user profiles against unauthorized access. The user can define policies which allow the access to only a subset of the information contained in the user profile. This is also known as *identity management*. Its aim is to provide as few information as necessary to the environment that is needed for the realization of the wished functionalities. Closely connected to this issue is the security of the environments. They must be also protected against attackers. This is realized by a two-level ticket-based authentication procedure.

Service Negotiation Architectures in Converged Networks

Juan Miguel Espinosa Carlin

`espinosa@i4.informatik.rwth-aachen.de`

GRK 643: Software for mobile communication systems

Department of Computer Science, Informatik 4, RWTH Aachen University

1 Research activities

Fixed Mobile Convergence (FMC) is the paradigm in the telecommunications industry that allows to interconnect fixed networks (Internet, ISDN, PSTN) with their mobile counterparts (GPRS, UMTS) for the purpose of giving operators the possibility to provide services irrespective of the location, access technology and terminal type of their users.

To implement the delivery of services inside such networks, the 3rd. Generation Partnership Project (3GPP) designed the IP Multimedia Subsystem (IMS), a global access-independent and standard-based IP connectivity and service control architecture that enables various types of multimedia services to end-users using common Internet-based protocols. The intention of the IMS is to aid a form of FMC, allowing the access of multimedia and voice applications across wireless and wireline terminals

The IMS Core Network is composed by several nodes, the most important being the Home Subscriber Servers and Subscriber Location Functions (HSSs and SLFs), the Call/Session Control Functions (CSCFs), the Application Servers (APs), the Media Resource Functions (MRFs), the Breakout Gateway Control Functions (BGCFs), and the PSTN gateways.

Although the procedures and control interfaces for static service delivery in the IMS are well standardized and detailed, the area of service negotiation in multiple IMS environments hasn't been explored yet. In these scenarios, further considerations have to be done regarding the mechanisms used when deciding if a particular service should or should not be provided to a determined user.

From this Inter-IMS service negotiation perspective, a user in a roaming network would have the possibility of either accessing her/his services in her/his home network, accessing the ones in the visited network or using the ones provided by a third party network, based on parameters like availability, cost, and/or QoS. It's important to mention that this service negotiation scheme would have a close interaction with the roaming mechanisms used inside the network.

Currently, my activities are focused on the performance evaluation of available IMS testbeds, the development and testing of native IMS clients on those testbeds and the implementation and deployment of IMS services, all pointing towards the design, implementation, and evaluation of the mentioned Service Negotiation Architecture.

New Communication Primitives for
Wireless Mesh Networks

Tobias Heer

tobias.heer@cs.rwth-aachen.de

GRK 643: Software for Mobile Communication Systems

Department of Computer Science, Distributed Systems Group, RWTH Aachen University

Modern networking applications require extended communication primitives like multicast, anycast, service composition, and delegation. As a matter of fact, providing these communication primitives efficiently surpasses the capabilities of the traditional point-to-point communication paradigm prevalent in today's Internet. While attempts to provide these primitives within the core network structure of the Internet have largely failed, several overlay networks like the Internet Indirection Infrastructure (i3) which enable these communication primitives have emerged. However, i3 has been designed for reliable infrastructure networks like the Internet. This makes it inapplicable in networks that exhibit a more dynamic network behavior, such as mixed ad-hoc and wireless mesh networks.
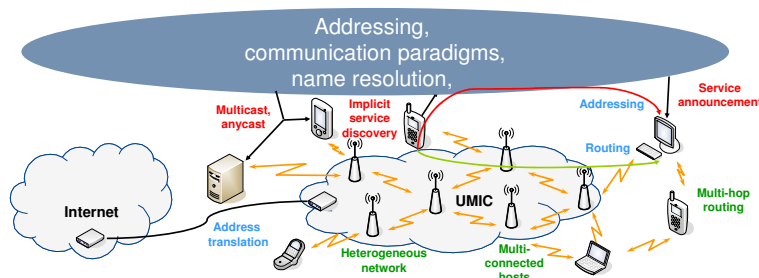

Fig. 1: Communication in UMIC.

Ultra High-Speed Mobile Information and Communication (UMIC) is a cluster of excellence[1] which focuses on providing high-speed communication in wireless networks beyond the limitations of today's WiFi and wireless mesh networks. Figure 1 depicts an exemplary UMIC scenario and some challenges in UMIC. Our contribution to UMIC aims at realizing extended communication primitives in dynamic networks. The integration of these extended primitives into UMIC yields a wide range of challenges. On the one hand, the indirection infrastructure must be tightly integrated into the network to avoid an overhead due to inefficient routing and unnecessary maintenance. On the other hand, the infrastructure must abstract from the actual network topology in order to provide a consistent and clear interface for applications. The peculiarities of wireless routers and mobile hosts, such as limited processing power or battery lifetime, paired with the demand for secure communication also require to rethink the use of CPU intensive authentication and integrity protection. Therefore, alternative authentication schemes are required in order to support CPU restricted mobile devices.

---

[1]Exzellenzcluster der DFG Exzellenzinitiative

Wireless Mesh Networks for public Internet access

Sebastian Max

`smx@comnets.rwth-aachen.de`

GRK 643: Software for Mobile Communication Systems

Chair of Communication Networks, Faculty 6, RWTH Aachen University

Wireless Mesh Networks (WMNs) transparently extend the coverage of a portal (e. g. connected to the Internet) via the installation of Mesh Points (MPs) which forward data over multiple hops to the associated stations and back.

One of the most important deployment scenarios for WMNs is the installation of municipal- and community networks. In these public networks, a provider sets up the WMN to enable low-cost communication, either as a competitor to the wired 'last-mile' to a residential neighborhood, or to provide wireless Internet-access for mobile users.

The current research activity by the graduate school member Sebastian Max covers the following topics related to WMNs for public Internet access:

- **Evaluation of Deployment Concepts:** Unlike with traditional radio access networks, it is still unclear how to deploy WMNs in urban/downtown scenarios in an optimal way. The analysis of different deployment concepts, comprising numerous parameters from the topology to transmit power control, allows for a substantiated judgment of the design space.

- **WMNs based on IEEE 802.11:** IEEE 802.11 represents the opportunity for a cost-efficient setup of WMNs: Through its standardization and heavy usage, hardware prices are negligible in comparison to other radio access technologies. Of course, the simplicity of IEEE 802.11 is a drawback as it is not designed for multi-hop operation. Hence, new mechanisms which combine current ammendments (e. g. IEEE 802.11 e,n,s) to optimize the network performance are under research.

- **Multi-channel WMNs**: While common radio access technologies use cells with different fixed frequencies to avoid interference, the structure of a WMN allows for a more flexible frequency usage: Due to the decentralized topology, each link is able to select the appropriate frequency channel and thus lower the interference burden on its neighbors. Combined with multiple radios per mesh device, the optimization space is increased drastically. Hence, algorithms are researched which optimize the channel and radio usage in a decentralized way. Furthermore, the calculation of the resulting spectral efficiency allows for a fair comparison with cell-based topologies.

- **Economical analysis:** As an orthogonal element, the presented points shall not only analyze the performance using the classical throughput and delay measures, but also take into account the economical feasibility of WMNs.

Cooperative MAC for Wireless Multi-Hop Networks

Ulrich Meis

(meis@cs.rwth-aachen.de)

GRK 643: Software for mobile Communication Systems

Communication and Distributed Systems (Informatik 4)

Department of Computer Science, RWTH Aachen University

Wireless networks have become a very popular domain of research. One of the main reasons might be the availability of hardware that anyone can afford.

Networks that span multiples of the average node's signal range, so-called multi-hop networks, originated in research studies funded by the military—which is a very common phenomenon. Therefore, these so-called ad-hoc networks have applications in the military domain. Although researchers have been on the hunt for other usage scenarios for many years, it now seems apparent that there are none. However, ad-hoc networks have evolved into several other network types that have a variety of usage scenarios.

Perhaps the most prominent representative of this class is the wireless mesh network. Contrary to assumptions made in ad-hoc networks, the nodes are relatively stationary and also well connected. Especially this last property makes efficient utilization of the wireless medium a hard problem in these networks. The medium access layer that is nowadays used virtually everywhere (and wireless mesh networks are no exception), IEEE 802.11, shows bad performance in well connected and especially in multi-hop networks. It has not been designed with networks of this kind in mind. In most scenarios, throughput is reduced by about 50 percent for every hop. Paths with more than 5 hops are rarely observed.

IEEE 802.11's bad performance shows that there is much room for improvement. Although a new standard for wireless mesh networks is in the making, IEEE 802.11s, it seems unlikely (judging by its current design) that it will bring much improvement. Since many nodes in a wireless mesh network are stationary (there usually is a fixed backbone) it should be possible to apply more sophisticated techniques on the MAC layer than what is currently envisioned by the IEEE. Problems encountered with the combination of CSMA/CA, high interference, and packet forwarding can be avoided if nodes apply a more cooperative approach to medium access.

Currently, the author investigates the possibility of a TDMA based approach. Coordination between the nodes would have to be carried out in a distributed fashion. Synchronization is the crucial part of this procedure. It is also desirable that the approach can be implemented with todays so-called Soft-MAC interface cards, since this would make its adoption easier, cheaper, and therefore a lot more likely. The UMIC-Mesh testbed located at RWTH Aachen University provides an ideal platform for implementation and measurement.

Context-Aware Smart Cars

Cem Mengi

GRK 643: Software for Mobile Communication Systems
Department of Computer Science 3 - Software Engineering
RWTH Aachen University
mengi@i3.informatik.rwth-aachen.de

The automobile of the future will become more and more to a smart environment. An environment of an automobile thereby comprises every context that supports the automotive system. This includes soft- and hardware components inside the car, user-specific services or the environment beyond the automobile, etc. The management of such a complex system could only be achieved through standardization issues. For example the developments of the AUTOSAR and OSGi standards are the outcomes of such efforts. While AUTOSAR standardizes the software-architecture for each ECU (Electronic Control Unit) and software-interfaces at application-level in an automotive system, the OSGi standard enables the deployment of services over wide area networks to local networks and devices.

During my research activities I have investigated these standards to get background information of future technologies in the automotive area. Furthermore, I will model and implement a simulation environment for smart cars, where the above mentioned context is taken into account. Within this simulation environment the integrated applications and services could be tested and evaluated. Moreover, so called context models play an important role when developing context-aware services. In my work I will evaluate different context models for the domain of smart cars.

The Department of Computer Science 3 at RWTH Aachen University is presently developing a simulation environment for smart homes, so-called eHomes. Especially the integration and configuration of services in a cost-effective way is of primary interest. Further investigations in this area are the composition, personalization and migration of services. These aspects apply also for smart cars but in addition mobility and real-time requirements have to be considered. In addition the communication between a smart home and a smart car is one important task that has to be fulfilled. Therefore in a next phase I will ensure the interoperability of the smart car environment with the smart home environment. Based on this, new services could be adopted, tested and evaluated.

USSA: A Unified Service-oriented System Architecture

Elena Meshkova
eme@mobnets.rwth-aachen.de

GRK 643: Software for Mobile Communication Systems
Department of Wireless Networks, RWTH Aachen University

The need for information fusion across communication layers for cross-layer performance optimization is particularly important in multihop wireless ad hoc and mesh networks. There is an emerging consensus in the networking research community that specifying a "universal" static stack of network protocols for heterogeneous transmission environments is a very difficult, if not possible at all, task. Instead, numerous research projects are defined towards developing an autonomic networking or cognitive resource management functionality. In these approaches atomic networking functionalities with flexible inter-layer interfaces are actively combined by the network nodes at runtime into a protocol suite suitable for a specific communication environment.

We are developing USSA, a Unified Service-oriented System Architecture design. USSA employs a unified way to describe and access information which is used both internally inside an operating system for composition of actual network services and externally for service access by an end user in the service-oriented technology sense. We utilize and extend the experiences from computer and sensor network domains for decomposing complex network services, i.e. routing, into atomic functionalities and present a methodology for dynamic re-configuration of these services. We do not propose a unified structure for a communication protocol stack. A specific set of networking functionalities is combined into a suite that we call an amorphous stack for a particular application level service. The functional content of the amorphous stack can be altered at run-time without physically upgrading the existing software.

The USSA architecture consists of two parts. The static part is invoked during the precompilation phase. It aims to decrease the processing power and memory usage on the devices during the run time. At this stage we use the ontology to check the compatibility of the device components, as well as of the whole network. We also give recommend on the device services composition based on the user requirements. The dynamic part of USSA enables the middleware and cross layer mechanisms that allow run-time performance optimization of the device and the network. Currently our solution is applicable to wireless sensor networks (WSNs). Later we plan to extend our design to the area of the cognitive radios, especially in the relation to Cognitive Resource Manager developed by RWTH Aachen University.

This work forms a part of the doctoral dissertation research, in which we are studying and developing artificial intelligence and ontology based methods to optimize cognitive wireless networks. We will also aim to provide flexible plug-and-play system architecture for cognitive radios.

Graph-based Reengineering of Telecommunication Systems

Christof Mosler
christof.mosler@rwth-aachen.de
GRK 643: Software for Mobile Communication Systems
Department of Computer Science 3 (Software Engineering)
RWTH Aachen University

Eighty percent of programming resources are allocated to maintaining and reengineering existing code. There exist many approaches concerning reengineering of legacy systems, but the majority of these approaches is dealing with systems in the field of business applications. The E-CARES project concerns understanding and restructuring of complex legacy systems from the telecommunication domain. Such systems are real-time embedded systems using the signaling paradigm, thus they pose additional requirements, e.g. regarding the system's performance. Telecommunication experts often think in terms of state machines and protocols which makes it even more difficult to adapt the already achieved results. The E-CARES research project is a cooperation between Ericsson Eurolab Deutschland GmbH (EED) and the Department of Computer Science 3, RWTH Aachen. E-CARES is an acronym for Ericsson Communication ARchitecture for Embedded Systems. The cooperation aims to develop methods, concepts, and tools to support the process of understanding and restructuring complex legacy telecommunication systems. The current system under study is Ericsson's AXE10, a mobile-service switching center (MSC) comprising more than ten million lines of code written in PLEX (Programming Language for EXchanges).

The first phase of the E-CARES reengineering project concerned the reverse engineering of telecommunication software. The project focused on the detection, extraction, and visualization of information on the system's structure and behavior. Extraction from static information (code) and from runtime information (traces) was regarded. The information is used to build a so-called system structure graph. This graph contains information on a system's decomposition into different units in different granularity.

Current work concerns restructuring of legacy telecommunication systems including their re-design and re-implementation. The aim is to extend the reverse engineering tool to a functional reengineering tool, allowing the engineers to interactively modify and improve the software. Several clustering-based algorithms improving the software architecture and performance have been successfully implemented. As telecommunication systems are often planned and modeled using state machines and protocols, it seems reasonable to analyze the systems also on the abstract representation level of state machines. For this reason, we extract the state machines from the source code, compare them with the original specification documents, and take them into account when suggesting appropriate reengineering steps.

Communication Protocols for Wireless Sensor Networks

Kittisak Ormsup

ormsup@i4.informatik.rwth-aachen.de
GRK 643: Software for Mobile Communication Systems
Department of Computer Science , Informatik 4, RWTH Aachen University

Abstract

Challanges in sensor network design nowadays:

Network Lifetime No real test beds can operate longer than 4-9 months. Existing routing protocols for wireless sensor networks don't perform very well in real sensor applications due to poor scalability, poor dynamic and excessive usage of node resources (which causes a short network lifetime). I therefore concentrate on the question:
Is it possible to run an ad hoc wireless sensor network for more than 1 year, using only low-cost motes with 2 AA-batteries?

Cross-layer Design Most of the research works has been done either only on MAC-layer (with static routing) or on the routing-layer (using 802.11-MAC). The results will be completely different using routing protocols for sensor networks or another MAC designed for sensor networks.

This project tries to cope with this problems in 3 steps:

1. **Routing Protocol** - Explore an approach which needs less resources and has a high dynamic, based on Ant Algorithm, where forward and backward agents (packets) are periodically sent between source and destination nodes, thereby leaving pheromone values on the paths, with the help of this values and some calculations, best path can be found without large overheads and maintenance efforts. The goal of the project is to implement the routing protocol on ns-2 network simulator, based on this approach, find the equation that best calculates the path, solve the scalability problem, compare the protocol to existing ones

2. **MAC-Protocol** - Find the MAC-protocol that best suit this routing protocol

3. **Test** - Implement the whole protocols on a real sensor test bed

Results shown that the ant-based routing protocol outperforms Directed Diffusion. Now it's at the stage of finding the appropriate MAC-protocols.

Dynamic Composition of eHome Software

Daniel Retkowitz

GRK 643: Software for Mobile Communication Systems
Department of Computer Science 3 (Software Engineering), RWTH Aachen University
`retkowitz@cs.rwth-aachen.de`

In today's homes, a lot of appliances are available, but in general, these appliances are not interconnected. To facilitate comprehensive services based on multiple appliances, that offer complex functionalities to the users, it is necessary to develop flexible and adaptive software. To achieve this goal at low overall costs, eHome software has to be built from standard components, that are automatically composed and adapted to the user's needs and the individual home environment. The customization of the eHome software is achieved by component composition in a process of specification, configuration, and deployment. We call this process the SCD-process.

An automatic support for the SCD-process is one of the key issues for the application of eHome services. In previous research of our group, support for an automated static configuration has been developed. In this approach, a specification of services, devices, and the eHome environment, and a selection of the desired top-level services is needed. On the basis of this information, a configuration is generated automatically. Up to now, no adapting changes of this configuration are possible later on at runtime. Services are only deployed once in the beginning. This is a strong restriction of the current process, particularly with regard to user mobility.

In my research, an incremental approach is pursued to cope with the dynamics of eHome environments. Whenever changes occur in the eHome environment, the SCD-process has to be re-executed to adapt the software to the new situation. Any change of the user's location or desires or of available devices implies corresponding changes in the specification and hence also the configuration and the deployment. The existing configuration mechanism will be extended to react on specification changes and to support an automated and flexible reconfiguration. For the deployment, capabilities to add services to the runtime environment, to remove others, and to change the execution states of services are needed. To achieve these goals, the present implementation of the SCD-process and the techniques for automatic service composition will be extended.

In certain circumstances, a service cannot provide its full functionality, e.g. if some base services are not available. It may still be possible for the service to provide a reasonable subset of its functionality. Also, it may be possible to replace some service by some other service with similar capabilities. New mechanisms are needed to offer flexibility and adaptability in service composition. This way, less user interaction will be required, since fewer composition mismatches will occur, that require manual resolution. To enable such mechanisms, we develop new concepts for semantically annotated service specifications and the respective composition algorithm.

Real-time Radio Wave Propagation

Arne Schmitz

aschmitz@cs.rwth-aachen.de

GRK 643: Software for Mobile Communication Systems

Computer Graphics Group, RWTH Aachen University

The simulation of the propagation of radio waves is important in many aspects of mobile communications. Both in cellular phone networks and in personal wireless networks the knowledge of radio propagation behavior is essential. Therefore we have implemented a radio wave propagation algorithm based on photon tracing. However computation could not be done in real-time, which would be necessary for interactive analysis of radio networks, antenna placement or usage in a packet-level network simulator.

Thus we developed a solution to use the radio wave propagation simulation at interactive rates. The input of the algorithm is the scene geometry and a sampling of possible radio transmitter positions in the scene. We then compute a simulation of the radio waves for each of the possible positions. This step will take a few minutes but is only necessary once for each scene.

(a)

The result will be a set of up to a few hundred 3D images containing the field strength for each transmitter position. These results can now be viewed with our interactive transmitter placement tool which allows moving the transmitter around. This is achieved by interpolating the results of the precomputed simulations. This is done using programmable graphics hardware allowing for up to 150 frames per second in outdoor (a) and indoor (b) simulations.

(b)

Note that all the simulations run in 3D, as opposed to other works, which only do 2D simulations on maps. The fact that the geometry can be arbitrary is especially important for indoor simulations, where wireless connections between different floors have to be simulated. Also note that complex shadowing effects due to buildings or walls obstructing the transmitter are simulated.

This same approach is also used to get more accurate results when simulating wireless networks in the ns-2 simulation toolkit. We allowed this widely used network simulator to use the precomputed radio wave simulations in order to use a more accurate physical layer model. Before this, radio wave simulations based on ray-tracing techniques were simply too slow to be used in network simulators and less accurate models were used instead.
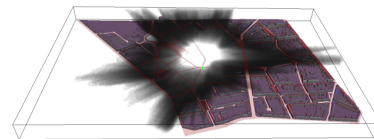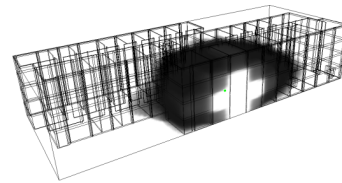
Mobile Web Service Provisioning

Satish Narayana Srirama

GRK 643: Software for Mobile Communication Systems
Department of Computer Science, Informatik 5
RWTH Aachen University
srirama@cs.rwth-aachen.de

The increased processing, storage and communication capabilities of mobile devices, in today's world of pervasive computing, combined with the latest developments in the area of Service Oriented Architecture lead to mobile web services domain, where the smart phones are being used as both web service clients and providers ('Mobile Host') [SJP06a]. While service delivery and management from Mobile Host are technically feasible, as proved from our performance analysis of Mobile Host, the ability to provide proper Quality of Service (QoS), especially in terms of security and scalability, is observed to be very critical.

Mobile Host opens up a new set of applications and it finds its use in many domains like mobile community support, collaborative learning, social systems etc. Primarily, the smart phone can act as a multi-user device without additional manual effort on part of the mobile carrier. While the applications possible with mobile web services are quite welcoming, the huge number of web services possible, with each Mobile Host providing some services in the wireless network, makes the discovery of these services quite complex. Proper discovery mechanisms are required for successful adoption of mobile web services into commercial environments.

So my current research focuses at checking application scope of Mobile Host, providing proper QoS for mobile web services and identifying alternate discovery mechanisms for mobile web services using P2P networks. The analysis also identified the need for a mediation framework, integrating all these domains. I am trying to realize an Enterprise Service Bus based Mobile Web Services Mediation Framework [SJP06b], acting as an intermediary between the mobile web service clients and the Mobile Hosts.

References

[SJP06a] Srirama, S.; Jarke, M.; Prinz, W.: Mobile Web Service Provisioning. In: Int. Conf. on Internet and Web Applications and Services (ICIW06), IEEE Computer Society, 2006; pp. 120-125.
[SJP06b] Srirama, S.; Jarke, M.; Prinz, W.: A Mediation Framework for Mobile Web Service Provisioning, In: 10th IEEE International Enterprise Distributed Object Computing Conference Workshops (EDOCW'06), 2006.

An Adaptive Checkpoint Placement
for Providing Reliability on Battery-Power Devices

Thitinan Tantidham

thitinan@i4.informatik.rwth-aachen.de

GRK 643: Software for Mobile Communication Systems

Department of Computer Science, Informatik 4, RWTH Aachen University

Abstract: This research focuses on session persistence for personal mobility computing which is based upon the former FAST Integration project (Framework Architecture Supporting Telematics Integration). Since the rapid advance of mobile computing and communication technologies, the ratio of price to performance for battery-power devices like PDA, smart phone and computer laptop as well as peripheral storages has decreased. It becomes popular that people on the move can suspend and restore services from one device to another by capability of wireless communication ports (e.g. WLAN, GSM or UMTS) for accessing their remote storage, or by a USB port for accessing a portable storage. However, a finite battery capacity is a major concern that users are often faced because of battery exhaustion and inadvertently session interruption. In order to keep service continuation, basic requirements are portable checkpointing and migration techniques in user level for a heterogeneous environment. In recent researches, a checkpointing protocol for a mobile computing environment has been developed on the basis of local storage limitations. This protocol employs a proxy server or a mobile station system for locating a persistent storage and functioning as a checkpoint manager in order to save and maintain the checkpoints on behalf of user devices. However, in real life due to high access cost and inadequate network coverage, while traveling, users are mostly in disconnection mode. To cope with this problem, a self-checkpoint engine with the capability of storing the checkpoint on the peripheral storage is introduced.

The research goal is to find an optimum checkpoint interval which is optimized between user application time and failure recovery time under the boundaries of battery lifetime. In this case, system lifetime is represented by battery lifetime. Therefore, a checkpoint is placed at least once before battery exhaustion. The checkpoint interval for a given battery lifetime is investigated on the basis of energy consumption, Mean Time to Failure (MTTF) and failure distribution.

In order to study battery characteristics and energy consumption, we made an experimental setup based on several laptops with different ages of lithium-ion batteries, USB portable storages, and a remote storage server. In the reality of usage, we assume that users do not know the number of discharge cycles and physicochemical properties as well as electrical measurements. Therefore, we use ACPI (Advanced Configuration and Power Interface) enabled by operating systems for monitoring battery capacity. As the results show, the computation of remaining battery life time from ACPI which is based only on discharge rate takes some minutes to predict battery life and has high prediction errors in some cases. We propose a methodology based on learning techniques with a combination of some measurable effects such as CPU utilization, CPU frequency, and temperature over time for more accuracy and faster prediction. The expected result is an optimal checkpoint placement scheme as an adaptation to battery lifetime, workload, and failures which is evaluated by the simulation.

Realistic Simulation of wireless multi-hop networks

Martin Wenig

`wenig@i4.informatik.rwth-aachen.de`

GRK 643: Software for mobile communication systems

Department of Computer Science, Informatik 4, RWTH Aachen University

Simulation of multi-hop wireless networks is still the main tool to develop new protocols and techniques and asses their performance. A quick overview of the published papers of the well-known MobiCom conference reveals that more than 2/3 of all papers report results based on simulation. Many factors determine the reliability of the generated results. Surely the statistical evaluation of the results is of great importance, but the focus of my research is the accuracy of the primary simulation results.

Two key components of each simulation model for mobile multi-hop networks are the radio wave propagation model and the mobility model on which the user behaviour is based on. These two parts are actually strongly interconnected. For example, the well-known Manhattan mobility model suffers from the fact that communication between the lanes is still possible because of missing obstacles. On the other hand, models which consider obstacles mostly restrict the mobility too strongly.

The goal of my research was to build a integrated modeling framework for accurate radio wave propagation and realistic mobility. The radio wave propagation part was solved using a ray-tracing approach which delivers the needed accuracy. It takes into account the most important phenomenons like reflection, scattering, and attenuation of walls. However, since the runtime of the simulation would be severely prolonged if the ray-tracing process would be started on-line, the developed solution works with precomputed energy density maps which are used to interpolate the needed field strength values. Hence, the simulation time is affected only very little (an increase of around 20% compared to reported factors of 100 with other solutions).

The mobility model works by dividing the simulation area into smaller regions. Regions can model streets, offices, hallways, or similar small entities of the simulation scenario. Each region is modeled by one of the well-known mobility models: Random Waypoint, Random Direction, Manhattan Grid, and the like. All modeling parameters of each region are independent of the other regions. The important feature is that if regions overlap, mobile nodes can travel from one region to the other, further on behaving like determined by the new mobility model. The exit probability is set for each region separately, thereby defining a directed and weighted graph on the set of regions.

The proposed framework can not only be used for simulation, but it can also be applied to other modeling tasks: my current research activities comprise the numerical analysis of possible performance limits in multi-hop networks and network planning for wireless mesh networks.

TCP Performance in Wireless Multi-Hop Networks

Alexander Zimmermann
`zimmermann@cs.rwth-aachen.de`

Department of Computer Science, Informatik 4, RWTH Aachen University

As various wireless networks evolve into the next generation to provide better services, a key technology has emerged recently, *Wireless Mesh Networks (WMN)*. WMNs are multi-hop networks with the main focus on the mass market's demands: providing *high bandwidth* and access to the *Internet*. They construct a permanent, wireless multi-hop and hierarchical infrastructure in order to connect the clients to the Internet.

A challenging task in WMNs is the reliable and efficient transfer of user data between nodes. Typically, the TCP is used for this purpose, which offers a reliable byte stream service. However, the TCP cannot be efficiently used in wireless networks. Since it was originally designed for wired network, it does not take into account the special properties of wireless networks. Especially in wireless multi-hop networks like WMNs, TCP suffers from several distinct problems.

One of the well-known reasons for the TCP performance degradation in wireless networks is that TCP assumes that all packet loss is due to congestion. However, in wireless networks a non negligible amount of loss is caused by wireless transmission errors, route failures or disconnections. In these cases TCP's error recovery mechanism can not distinguish congestion and non-congestion loss. As a consequence TCP performs a reduction of the congestion window and slow start threshold, resulting in an unnecessary performance degradation.

TCP is critically dependent on acknowledgment, so its performance can be severely impacted by network asymmetry which is defined as the situation where the forward direction of a network is different from the reverse direction in terms of bandwidth, loss rate, and latency. In WMNs, data and ACK packets may take different paths, and thus experience different packet loss rate, latency, or bandwidth. Even if the same path is taken by data and ACK packets, they still face network asymmetry problem, since the channel condition on the path varies from time to time.

In WMNs, mesh routers are connected in a wireless multi-hop fashion, so dynamic changes of the routing path is common. Considering client mobility, variable link quality, traffic load, and other factors, the change may be frequent and may cause large variations in RTT. This will degrade the TCP performance, since the normal operation of TCP relies on a smooth measurement of RTT.

The focus of my research is to improve the performance of TCP in wireless multi-hop networks. On the one hand TCP has to be adjusted or improved to work more efficiently in wireless networks. On the other hand the improved version of TCP must be compatible to the old protocol because of the high utilization of TCP in classical, wired networks.

OFDMA Testbed Implementation in Software Defined Radio

Milan Zivkovic

GRK 643: Software for Mobile Communication Systems
Institute of Theoretical Information Technology
RWTH Aachen University
Sommerfeldstraße 24
52074 Aachen
zivkovic@ti.rwth-aachen.de

OFDM (Orthogonal Frequency-Division Multiplexing) transmission technology presents an effective implementation of a multi-carrier modulation principle where serial high-rate data stream is converted into multiple parallel low-rate streams, each modulated on another sub-carrier. The orthogonality of sub-carriers offers high bandwidth savings allowing high efficiency data transmission in wireless dispersive multipath environments. Due to its good performance, high flexibility, and simple implementation, OFDM presents underlaying technology for current wireless transmission standards as DAB (Digital Audio Broadcasting), DVB-T (Digital Terrestrial Video Broadcasting), WLAN (Wireless Local Area Network) IEEE 802.11 a/g and WiMAX (Worldwide Interoperability for Microwave Access) IEEE 802.16 standard for fixed and mobile broadband wireless access. Also, in multiuser environment, OFDMA (OFDM Access) presents an effective solution for multiple access.

Since simulation can not include all environmental effects there is need for a hardware testbed where proposed algorithms and solutions can be evaluated. SDR (Software Defined Radio) platform is possible efficient solution for such testbed due to its easy reconfigurability. In SDR architecture, transmitter/receiver algorithms and functions which describe transmission standards are realized as programs running on suitable processors or reprogrammable hardware components. My current research activity is focused at implementation of GNU Radio architecture as SDR solution for an OFDMA system testbed. GNU radio is a free software toolkit licensed under GPL for building and deploying some transmission standards. USRP (Universal Software Radio Peripheral) boards are used as corresponding general purpose integrated hardware components for "over-the-air" communication. According to present channel state conditions, adaptive modulation and coding schemes can be easily realized through the corresponding software blocks and integrated in OFDM symbol transmission. This allows great flexibility for evaluation of resource allocation algorithms in multiuser OFDMA environment.

GRK 1042:

Explorative Analysis and Visualization of Large Information Spaces

Konstanz

Graduiertenkolleg Explorative Analysis and Visualization of Large Information Spaces

University Konstanz
July 2004 – Dec 2008

Im Zentrum des Forschungsprogramms des Graduiertenkollegs steht die

Entwicklung von Methoden insbesondere im Rahmen von Visualisierung und Computergrafik zur Unterstützung von Exploration, Analyse und Management großer Datenräume. Dabei können diese Datenräume auch selber visueller Natur sein, z.B. in Form von Multimediadokumenten oder komplexen geometrischen Strukturen. Die im Graduiertenkolleg besonders relevanten Fachrichtungen sind Informationsvisualisierung, Computergrafik, Human Computer Interaction, Intelligente Datenanalyse, Information Retrieval, Datenbanken und Informationssysteme, Softwareanalyse sowie digitale Kommunikation. Ein Ziel der Datenexploration und -analyse ist es, neue a-priori unbekannte, aber für den Anwender nützliche Informationen zu finden. Die Forschung zielt darauf ab, existierende Verfahren effektiver und effizienter zu machen sowie neue Verfahren der Exploration und Analyse zu entwickeln, die den speziellen Erfordernissen der z.B. im Internet gespeicherten und zu übertragenden Informationen gerecht werden. Dabei sollen die Informationen analysiert und gruppiert (geclustert) sowie ihre Qualität bewertet werden. Für diesen Prozeß werden zunächst Informations-Repräsentations-Methoden der Informationswissenschaft sowie Datenmodellierungsmethoden aus dem Bereich Datenbanksysteme benötigt. Die Informationsmodellierung ist dann die

Ausgangsbasis für die eigentliche Exploration der Daten, die durch eine Kombination von automatischen Clusteranalyseverfahren, einer wissensbasierten semantischen Analyse sowie einer interaktiven Visualisierung der Daten erfolgen soll. Ein gewichtiger Anwendungsbereich der im Graduiertenkolleg zu entwickelnden Verfahren liegt in der explorativen Analyse von großen Beständen an Bioinformatikdaten.

Central to the research programs of the Graduate College is the development of methods, specially within the framework of visualization and computer graphics, in support of data mining,data analysis, and the management of large information spaces, whereby the information spaces themselves may be visual in nature, i.e., in form of multimedia documents or complex geometric structures. The subject areas most relevant within the program are information visualization, computer graphics, human computer interaction, intelligent data analysis, information retrieval, database and information systems, softareanalysis as well as digital communication. One of the objectives of data mining and data analysis is to find new, previously unknown, yet useful information. The research aims at perfecting existing procedures to be more effective and more efficient, and at the same time it seeks to develop new procedures with regards to exploration and analysis, which serve more adequately special requirements, such as the vast information stored and transferred in the internet. The information must first be analyzed and clustered, as well as qualified. To complete this process, methods of knowledge representation within the area of information science, as well as data modeling methods within the area of database systems are needed. Hence, the information modeling is the starting point for the actual exploration of the data. The latter should be worked through using a combination of automatic cluster analysis procedures, knowledge based semantic analysis, and an interactive visualization of the data. An important range of application lies within the explorative analysis of vast amount of bioinformation data.

Liste der Betreuer/Advisors List:

1. Prof. Dr. Dietmar Saupe (Sprecher/Chair) (Multimedia Signal Processing)
2. Prof. Dr. Michael Bertold (Intelligent Data Analysis);
3. Prof. Dr. Ulrik Brandes (Network Text Analysis and Visualization);
4. Prof. Dr. Oliver Deussen (Computer Graphics);
5. Prof. Dr. Raouf Hamzoui (Multimedia Signal Processing);
6. Prof. Dr. Daniel Keim (Information Visualization and Visual Data Mining);
7. Prof. Dr. Stefan Leue (Software Analysis and Verification);
8. Prof. Dr. Harald Reiterer (Human Computer Interaction);
9. Prof. Dr. Marc Scholl (Database and Information Systems).

Doctorate Students:

Shakeel Ahmad
Vladimir Bondarenko
Joachim Böttger
Thorsten Büring
Nicolas Cebron
Ioan Cleju
Martin Hoefer
Stefan Klinger
Florian Mansmann
Evghenia Stegantova
Hartmut Ziegler

Associated doctorate students:

Husain Aljazzar
Michael Balzer
Thomas Gabriel
Jens Gerken
Christian Grün
Alexander Holupirek
Werner König
Johannes Kopf
Thomas Luft
Svetlana Mansmann (geb. Vinnik)
Tilo Ochotta
Daniela Oelke
Christian Pich
Martin Röder
Mauro Ruggeri
Wei Wei
Bernd Wiswedel

Efficient transmission of multimedia data over unreliable channels

Shakeel Ahmad

Real-time transmission and visualization of multimedia data over Internet imposes a challenge because Internet provides transportation service to the realtime data only in a best-effort manner and does not provide any guarantee to deliver the data packets in-time. In particular we focus on efficient transmission of high resolution three dimensional (3D) models and live video data. Data packet which are lost or arrive at the receiver beyond their presentation deadline impact the perceptual quality of the received multimedia significantly. Error-resilience against packet losses can be achieved through forward error control (FEC), automatic repeat request (ARQ), error resilient coding, error-concealment or any combination thereof. Usually ARQ type solutions have unbounded end-to-end delay because they are based on retransmissions. Hence ARQ based solutions may not be well suited for real-time transportation of multimedia data as the time constraints may not be met. In our work, we concentrate on FEC based solutions. The basic principle of FEC is to add redundant information to the compressed source data so that the original source data can be reconstructed at the receiver in case of packet losses. We also investigate the potential benefits of recently proposed rateless or digital fountain codes for live video streaming applications.

Advisors: 1. Prof. Dr. Raouf Hamzaoui 2. Prof. Dr. Dietmar Saupe

Publications:

Ahmad, S., Hamzaoui, R., Optimal error protection of progressively compressed 3D meshes, Proceedings IEEE International Conference on Multimedia and Expo (ICME'06), Toronto, Canada, July 2006.

Adaptive thinning algorithms of atmospheric observations in data assimilation for numerical weather prediction

Vladimir Bondarenko

Data assimilation is a procedure of production of an accurate image of the true state of the atmosphere at a given time, which is used as the initial state for a numerical weather prediction (NWP). In this procedure, measurements of various observation systems are combined with an a priory estimate of the model state, called background information. Current and future observation systems, in particular satellite instruments, produce huge amounts of measurements with high spatial and temporal density. Such data sets significantly increase the computational costs of the assimilation and, moreover, can violate the assumption of spatially independent observation errors and more complex observation error statistics would be needed leading to additional increase in the computational costs. Thinning of observations is an efficient way to reduce the size and the effective error-correlation of the observation data sets. Among of other NWP centers, the German National Meteorological Service (Deutscher Wetter Dienst, DWD), applies some simple uniform thinning strategy, while more involved adaptive thinning algorithms are a subject of current research.

The main goal of this thesis consists in development of thinning algorithms similar to methods for mesh simplification in computer graphics and data compression techniques in digital signal and image processing. This implies a statistical modeling of the data assimilation procedure as well as a verification of the thinning algorithms on the real observations using the experimental system of DWD. Furthermore, approaches for visualization of the multidimensional data sets of atmospheric variables are of strong interests to this work. This thesis is to be conducted in cooperation with the German National Meteorological Service.

Advisors: 1. Prof. Dr. Dietmar Saupe 2. Prof. Dr. Gerik Scheuermann(Leipzig)

Publications:

Ochotta, T., Gebhardt, C., **Bondarenko, V.**, Saupe, D., Wergen, W., On thinning methods for data assimilation of satellite observations, Proceedings 23rd International Conference on IIPS for Meteorology, Oceanography and Hydrology, San Antonio, Texas, 2007

Bondarenko, V., Ochotta, T., Saupe, D., Wergen, W., The interaction between model resolution, observation resolution and observation density in data assimilation: a two-dimensional study, Proceedings 11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, San Antonio, Texas, 2007.

Visualization of Orders of Magnitude

Joachim Böttger

Exploring data is a challenge, if it contains lots of details that are very small in respect to the size of the whole dataset. One example for such data is information about the surface of the earth: While the whole planet spans several thousands of kilometers, details like trees and houses are several orders of magnitude smaller.

The traditional technique to explore such large information spaces is to zoom in on the details, however, the information about their context is then lost. Another possibility is to show the details in their context: Several solutions exist, which magnify the details the viewer is interested in, while simultaneously shrinking the surrounding parts of the data. Most of these techniques, however, suffer from introducing distortions such as compression or shear in parts of the resulting images, and hence make details in these areas hard to recognize.

On the other hand analytic complex mappings, which are well known from cartography, leave the details in an image locally undistorted while often scaling different parts of the data very differently. One of these mappings, the complex logarithm, is approximately employed in the mapping from the retina to the visual cortex in the brain. The same method allows for showing in one seamless visualization individual houses in the context of the continents that they are built on. Disadvantages are that the mapping introduces a cut in the data, and rotates parts of the image in ways that we are not used to.

One open question is how interaction can help users to explore data with mappings like the complex logarithm. Another interesting field of research is the use of such mappings for data with more dimensions. The resulting techniques will be useful for objects other than satellite data like maps, electron microscopy data, computer chips, and biological objects as well as for abstract data like filesystems, electronic circuits, and various graphs.

Advisors: 1. Prof. Dr. Oliver Deussen 2. Prof. Dr. Daniel Keim

Publications:

Böttger, J., Balzer, M., Deussen, O., Complex logarithmic views for small details in large contexts, IEEE Transactions on Visualization and Computer Graphics, Vol. 12, No. 5, pp. 845-852, Sept. 2006.

Mobile Information Visualization

Thorsten Büring

Due to improved connectivity and performance features, personal digital assistants (PDAs) are increasingly used to administrate, explore and search large information spaces. Unfortunately, most of todays handheld software relies on interfaces that present information textually on subsequently displayed screens. Being forced to jump back and forth, the user is not able to form a consistent mental model of the underlying data space and thus to accomplish tasks efficiently.

The problem of how to present information on computer screens more appropriately has been subject to extensive research in the field of information visualization (IV). The latter is commonly defined as the use of computer-supported, interactive, visual representations of abstract data to amplify cognition [Card et al. 1999]. Well-known IV applications like FilmFinder, SmartMoney and Peets Coffee Selector have already shown that by using visualizations, human computer interaction can be improved significantly. This perception is also reflected by the fact that an increasing number of IV ideas are turned into commercial products (eg. Spotfire, Inxsight, Human IT) or are incorporated into already existing software packages such as SPSS, Adobe Photoshop and Microsoft Outlook. A drawback however remains: most of the disciplines implementations have been designed for desktop computers featuring large-scale screens and while portions of the design approaches can be leveraged to PDAs, many of the PC interaction techniques and design philosophies simply do not apply. The authors thesis will therefore deal with the development of visualizations specifically targeted for small devices.

Advisors: 1. Prof. Dr. Harald Reiterer 2. Prof. Dr. Daniel Keim

Publications:

Büring, T., Gerken, J., Reiterer, H., Dynamic text filtering for improving the usability of alphasliders on small screens, , submitted, 2006;

Büring, T., Interaktionsstrategien für Punktdiagramm-Visualisierungen auf kleinen Bildschirmen, i-com, Zeitschrift für interaktive und kooperative Medien, Heft 2/2006.

Büring, T., Gerken, J., Reiterer, H., User interaction with scatterplots on small screens - A comparative evaluation of geometric-semantic zoom and fisheye distortion, IEEE Transactions on Visualization and Computer Graphics, Vol. 12, No. 5, Proceedings Visualization / Information Visualization 2006, Sept.-Oct. 2006.

Büring, T., Gerken, J., Reiterer, H., Usability of overview-supported zooming on small screens with regard to individual differences in spatial ability, Proceedings Advanced visual interfaces (AVI'06), Venice, Italy, May 2006.

Büring, T., Reiterer, H., ZuiScat - Querying and visualizing information spaces on personal digital assistants, Proceedings MobileHCI 2005 Human Computer Interaction with Mobile Devices and Services, Salzburg, Austria, Sept. 2005.

Data Mining in Bioinformatics

Nicolas Cebron

The development of high-throughput imaging instruments - e.g. flourescence microscope cameras - resulted in them becoming the major tool to study the effect of agents on different cell types. These devices are able to produce about 55,000 images per day; until recently, visual inspection by a domain expert was the only way to distinguish between 'active' and 'nonactive' cells.

The aim of this project is to design classifiers that are able to learn the differences between cell types. As we are dealing with a large amount of unlabeled data, the expert should label only a small subset to train the classifier. Choosing randomly drawn examples from the dataset would render the classifier biased towards the underlying distribution of the different kinds of cells.

Therefore, we try to apply the concept of 'active learning' to this task, where our learning algorithm has control over which parts of the input domain it receives information about. This concept is very similar to the human form of learning, whereby problem domains are examined in an active manner.

The objective of this thesis is to develop new concepts and algorithms based on the idea of active learning for miscellaneous data-mining algorithms in order to build stable classifiers in the field of bioinformatics.

Advisors: 1. Prof. Dr. Michael Berthold 2. Prof. Dr. Ulrik Brandes

Publications:

Berthold, M., **Cebron, N.**, Dill, F., Di Fatta, G., Gabriel, T., Georg, F., Meinl, T., Ohl, P., Sieb, C., Wiswedel, B., Knime: the Konstanz Information Miner, Proceedings International Workshop on Multi-Agent Systems and Simulation (MAS&S), 4th Annual Industrial Simulation Conference (ISC 2006), Palermo, Italy, June 2006, pp. 58-61;

Cebron, N., Berthold, M., Adaptive fuzzy clustering, Proceedings North American Fuzzy Information Processing Society (NAFIPS 2006), 2006, Montreal, QC, Canada;

Cebron, N., Berthold, M., Adaptive active classification of cell assay images, Proceedings Proc. European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), Springer Berlin / Heidelberg, 2006, pp. 79-90, vol. 4213;

Cebron, N., Berthold, M., Mining of cell assay images using active semi-supervised clustering, Proceedings Workshop on Computational Intelligence in Data Mining at ICDM, Houston, 2005, pp. 63-70.

Compression Methods for Point-Based 3D Models

Ioan Cleju

The 3D models have become more important among the media data types. As the volume of data will further increase, the rendering complexity of the traditional 3D models based on polygonal primitives becomes prohibitive. An alternative trend in computer graphics considers using the point-based models. This approach is further encouraged by the use of 3D scanning systems that generate point based 3D models of the objects. Each point contains information about position, color, normal and radius. In dealing with such large data, that contains millions of points, efficient compression is vital. The goal of this PhD thesis is to study and develop new compression methods for the point-based 3D models. The geometrical data compression tries to exploit the fact that the point samples lay on a surface, while the color data compression tries to adapt the 2D image compression methods that are currently used. Additional demands are possible that require a maximal error or certain data size, or the possibility for progressive decoding.

Advisors: 1. Prof. Dr. Dietmar Saupe 2. Prof. Dr. Ulrik Brandes

Publications:

Cleju, I., Saupe, D., Evaluation of supra-threshold perceptual metrics for 3D models, Proceedings Symposium on Applied Perception in Graphics and Visualization APGV06, Boston, USA, July 2006, pp. 41-44;

Cleju, I., Fränti, P., Wu, X., Clustering by principal curve with tree structure, Proceedings International Symposium on Signals, Circuits and Systems (ISSCS'05), vol.2, July 2005, pp. 617-620;

Cleju, I., Fränti, P., Wu, X., Clustering based on principal curve, Proceedings Scandinavian Conference on Image Analysis (SCIA'05), LNCS vol. 3540, June 2005, pp. 872-881.

Clustering and optimization under distributed competition

Martin Hoefer

Combinatorial optimization problems have long been studied in the area of efficient algorithms. The emergence of the internet and the ever-growing size of the instances poses new distributed and competitive approaches. They must capture the economic and selfish nature of independent agents that build, maintain and operate the internet. The consideration of these aspects poses new questions about cooperation and coordination of distributed rational agents. Under which conditions are they motivated to contribute towards a globally efficient feasible solution? How efficient is a solution that independent agents can agree upon? How hard is it to calculate such stable outcomes?

Mathematically rigorous treatments of questions like these involve standard algorithmic techniques in combination with notions from game theory. The thesis will focus on characterizations of efficiency and stability and their trade-offs in distributed competitive variants of network design and combinatorial optimization problems. These variants serve to capture fundamental aspects of the development of the internet and e-commerce. Furthermore, game-theoretic models in graph-theoretic optimization, clustering and facility location are studied. They are used to model problems arising in distributed network design, data analysis and the development of social networks.

Advisors: 1. Prof. Dr. Ulrik Brandes 2. Prof. Dr. Dietmar Saupe

Publications:

Brandes, U., **Hoefer, M**., Lerner, J., WordSpace - Visual summary of text corpora, Proceedings SPIE-IS&T Electronic Imaging, SPIE Vol. 6060 (VDA'06), 2006;

Brandes, U., **Hoefer, M**., Pich, C., Affiliation dynamics with an application to movie-actor biographies, Proceedings 8th Eurographics/IEEE-VGTC Symposium on Visualization (EUROVIS 2006), May 2006, pp. 179-186;

Hoefer, M., Non-cooperative tree creation, Proceedings 31st International Symposium on Mathematical Foundations of Computer Science (MFCS 2006), LNCS 4162, Springer-Verlag, 2006, pp. 517-527;

Cardinal, J., **Hoefer, M**., Selfish service installation in networks, Proceedings 2nd International Workshop on Internet & Network Economics (WINE 2006), 2006, pp. 174-185, LNCS 4286;

Hoefer, M., Non-cooperative facility location and covering games, Proceedings 17th International Symposium on Algorithms and Computation (ISAAC 2006), 2006, pp. 369-378, LNCS 4288;

Hoefer, M., Krysta, P., Geometric network design with selfish agents, Proceedings 11th International Computing and Combinatorics Conference (COCOON 2005), Lecture Notes in Computer Science 3595, 2005, pp. 167-178.

Extending an XQuery compiler with Full-Text capabilities

Stefan Klinger

Transfer of IR-techniques to the world of semi-structured data (XML) is an ongoing issue. Although there are many ideas about how to do IR on trees, these focus on improving IR quality, not on optimizations for speed or memory consumption.
Can IR techniques be integrated into existing XML-DBMSs?

IR-systems developed stand-alone in general do not profit from available high performance DBMSs. Currently I focus on injecting the Score Region Algebra IR technology, introduced by the CIRQUID Project (from Twente University, EWI Database Group), into the Pathfinder XQuery compiler (developed at TU München, Lehrstuhl für Datenbanksysteme).
Future issues of this research project are:

- the developement of optimizations like, e.g., specialized index structures, and
- the evaluation of their value and portability to other platforms.

Advisors: 1. Prof. Dr. Marc Scholl 2. Prof. Dr. Marcel Waldvogel

Visual Analysis of Network Communication

Florian Mansmann

Network communication has become indispensable in education, government, and business. Along with the pervasive use of the internet, the problems of information overload as well as large-scale misuse of the network communication infrastructure arise. Especially misuse for destructive purposes, such as spreading malicious code, compromising remote hosts, or damaging data through unauthorized access, has grown immensely in recent years. In this talk, I will exemplify these two problems and show how visualization technology can help to solve them.

In the context of personal information management, I research novel visual approaches to interact with large collections of emails, namely the space-filling geographical distortion technique HistoMap, a calendar-based recursive pattern to show temporal distributions, and a self-organizing map to retrieve similar emails. In the context of network security, it is promising to complement automatic intrusion detection mechanisms through visual analysis methods because novel attacks are hard to identify with automatic methods and pure statistics about the number of alerts rarely lead to the essential insight. To gain insight into the distribution of diverse traffic loads and large quantities of security alerts over the internet, I will introduce the Hierarchical Network Map, which is designed as a spatial, hierarchical, and space-filling visualization technique of the IPv4 address space. Different visual approaches to analyze application port as well as host statistics will be discussed.

Advisors: 1. Prof. Dr. Daniel Keim 2. Prof. Dr. Marcel Waldvogel

Publications:
Mansmann, S., **Mansmann, F.,** Scholl, M., Keim, D., Hierarchy-driven exploration of multidimensional data cubes, Proceedings 12. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW 2007), Aachen, Germany, to appear, 2007.

Mansmann, F., Vinnik, S., Interactive exploration of data traffic with hierarchical network maps, Transactions on Visualization and Computer Graphics, Vol. 12, No. 6, Nov.-Dec. 2006;
Vinnik, S., **Mansmann, F.,** From analysis to exploration: Building enhanced visual hierarchies from OLAP cubes, Proceedings 10. International Conference on Extending Database Technology (EDBT), Munich, Germany, Springer-Verlag, March 2006, pp. 496-514;

Schreck, T., Keim, D., **Mansmann, F.,** Regular TreeMap layouts for visual analysis of hierarchical data, Proceedings Spring Conference on Computer Graphics (SCCG'2006), Casta Papiernicka, Slovak Republic, Acm Siggraph, April 2006;

Keim, D., **Mansmann, F.,** Schneidewind, J., Ziegler, H., Challenges in visual data analysis, Proceedings Information Visualization (IV 2006), London, United Kingdom, IEEE Press, July 2006;

Keim, D., **Mansmann, F.,** Schneidewind, J., Schreck, T., Monitoring network traffic with radial

traffic analyzer, Proceedings IEEE Symposium on Visual Analytics Science and Technology 2006 (VAST 2006), 2006;

Janetzko, H., Keim, D., Kramis, M., **Mansmann, F.**, Waldvogel, M., Interactive poster: exploring block access patterns of native XML storage, Proceedings IEEE Symposium on Information Visualization (InfoVis 2006), 2006.

Mansmann, S., Scholl, M., Keim, D., **Mansmann**, **F.**, Interactive poster: exploring OLAP aggregates with hierarchical visualization techniques, Proceedings IEEE Symposium on Information Visualization (InfoVis 2006), Oct. 2006.

Keim, D., **Mansmann, F.**, Panse, C., Schneidewind, J., Sips, M., Mail Explorer - Spatial and temporal exploration of electronic mail, Proceedings EuroVis 2005: Eurographics/IEEE-VGTC Symposium on Visualization, June 2005.

Keim, D., **Mansmann, F.,** Schreck, T., Analyzing electronic mail using temporal, spatial, and content-based visualization techniques, Proceedings GI Jahrestagung (1), 2005, pp. 434-438.

Keim, D., **Mansmann, F.**, Schreck, T., MailSOM - Exploration of electronic mail archives using self-organizing maps, Proceedings Conference on Email and Anti-Spam 2005, 2005.

Data Mining Techniques in Software Fault Localization

Evghenia Stegantova

Software is a key point in most system development projects. Modern software systems are large and highly complex. As the complexity of the systems grows, it is becoming more and more complicated to understand the structure of the software system and to localize and analyse the errors . Nowadays it is an increasingly challenging task to ensure the robustness and reliability of a software system. In program analysis software faults can be classified into two categories: crashing faults ,i.e. core dump or segmentation fault, and noncrashing faults, i.e. errors that do not incur crashes, logical errors. One could apply debugging techniques to locate the cause of a crashing fault. In case of noncrashing faults it may require a great deal of a human effort to find the cause of an error, since no crashing point, hence no backtrace is available.

The aim of this project is to automate the process of locating the noncrashing faults using data mining techniques. Tracing data generated during program execution may disclose important patterns and outliers that may help discover software faults.

Therefore, we try to use this approach and apply data mining methods to the data gathered during program runs. As a result we expect an analysis of the program execution that would point out the possible source of the error.

Advisors: 1. Prof. Dr. Stefan Leue 2. Prof. Dr. D.Saupe

Publications:

Di Fatta, G., Leue, S., **Stegantova, E.**, Discriminative pattern mining in software fault detection, Proceedings Third International Workshop on Software Quality Assurance SOQUA'06, 14th ACM Symposium on Foundations of Software Engineering (ACM SIGSOFT), Portland, USA, November 2006, pp. 62-69;

Visual Exploration of Network Data

Hartmut Ziegler

The large amounts of data on the financial market today pose manycomputational challenges. Currently, companies like Reuters deliver around 50.000 data updates per second of financial stock market data. Such amounts of data can be analyzed by data mining algorithms, but for understanding market mechanisms and in order to evaluate the characteristics of assets, innovative visualization techniques offer many advantages over numeric techniques with regard to knowledge discovery in financial markets as one image can communicate more insight than ten thousand numbers.

The main objective of this project is the development of new concepts and algorithms for pixel-based visualization techniques to overcome the shortcomings of traditional techniques, and which significantly improve the insight into the behaviour of todays financial markets. The work focuses on detailed performance analysis of assets with new focus-and-context techniques ,inter-asset comparison of all assets on the market, pattern and trend recognition in different market sectors, portfolio analysis, and also includes the development of scalable techniques for high-resolution financial data analysis on large displays walls (powerwalls).

Advisors: 1. Prof. Dr. Daniel A. Keim

Publications:

Keim, D., Nietzschmann, T., **Ziegler, H.**, Relevance driven visualization of financial performance measures, Proceedings IEEE-VGTC Symposium on Visualization, Norköping, Sweden (EuroVis 2007: Eurographics), to appear, May 2007;

Keim, D., Nietzschmann, T., Schelwies, N., Schneidewind, J., Schreck, T., **Ziegler, H.,** A spectral visualization system for analyzing financial time series data, Proceedings EuroVis 2006: Eurographics/IEEE-VGTC Symposium on Visualization, Lisbon, Portugal, May 2006;

Keim, D., Mansmann, F., Schneidewind, J.**, Ziegler, H.**, Challenges in visual data analysis, Proceedings Information Visualization (IV 2006), London, United Kingdom, IEEE Press, July 2006, Invited Paper.

DiRePro: Directed Model Checking in the Analysis of Reactive and Probabilistic Systems

Husain Aljazzar

The complexity of embedded software systems calls for efficient automated methods that ensure that these systems satisfy correctness criteria ensuring safe operation. Important classes of properties that these systems have to meet relate to their real-time behavior, and to the probability of performing certain services within a given time period. In this project we contribute to the development of automated verification methods that allow these properties to be checked for a given software design model. The particular challenge is to ensure that these methods scale to the huge size of the sate spaces that concurrent embedded systems entail. We address this challenge by reconciling model checking techniques for real-time and probabilistic systems with heuristics directed, intelligent state space traversal techniques, also known as directed model checking. In particular, we use heuristic techniques to increase the efficiency of the abstraction refinement loop for real-time model checking, and to elicit meaningful error traces for probabilistic system models.

Advisors: 1. Prof.Dr. Stefan Leue

Publications:

Aljazzar, H., Leue, S., Extended directed search for probabilistic timed reachability, Proceedings 4th International Conference on Formal Modelling and Analysis of Timed Systems FORMATS '06, LNCS, Springer-Verlag, 2006, pp. 33-51;

Aljazzar, H., Hermanns, H., Leue, S., Counterexamples for timed probabilistic reachability, Proceedings 3rd International Conference on Formal Modelling and Analysis of Timed Systems FORMATS'05, LNCS 3829, Springer-Verlag, 2005, pp. 177-195.

Data Exploration and Visualization using Hierarchical Fuzzy Rule Systems

Thomas Gabriel

Rule systems have failed to attract much interest in large data analysis problems because they tend to be too simplistic to be useful or consist of too many rules for human interpretation. We propose to extend a rule induction algorithm that constructs a hierarchical rule system, with only a small number of rules at each stage of the hierarchy. The underlying rule algorithm's structure is based on a probabilistic neural network that can be employed by various activation functions, such as radial and fuzzy-membership basis functions. Lower levels in this hierarchy focus on outliers or areas of the feature space where only weak evidence for a rule was found in the data. Rules further up, at higher levels of the hierarchy, describe increasingly general and strongly supported aspects of the data. The rule hierarchy allows building much smaller rule systems and the model—especially at higher levels of the hierarchy—remains interpretable. Each level in the hierarchy can be visualized independently using standard methods. Furthermore, one challenge is to visualize certain levels of the hierarchy in parallel.

Advisors: 1. Prof. Dr. Michael Berthold

Publications:

Berthold, M., Cebron, N., Dill, F., Di Fatta, G., **Gabriel, T.**, Georg, F., Meinl, T., Ohl, P., Sieb, C., Wiswedel, B., Knime: the Konstanz Information Miner, Proceedings International Workshop on Multi-Agent Systems and Simulation (MAS&S), 4th Annual Industrial Simulation Conference (ISC 2006), Palermo, Italy, June 2006, pp. 58-61;

Gabriel, T., Thiel, K., Berthold, M., Rule visualization based on multi-dimensional scaling, Proceedings IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, 2006;

Gabriel, T., Pintilie, A. S., Berthold, M., Exploring hierarchical rule systems in parallel coordinates, Proceedings 6th International Symposium on Intelligent Data Analysis (IDA 2005), LNCS vol. 3646, 2005, pp. 97-108.

Gabriel, T., Berthold, M., Missing values in fuzzy rule induction, Proceedings IEEE Conference on Systems, Man and Cybernetics, 2005.

BaseX - Processing and Visualizing large XML instances

Christian Grün

As XML has been proven to be successful as a standardized format for textual data exchange, it is increasingly used as a representation format for large databases, stretching over several gigabytes. Yet, most XML processors are still specialized on storing and querying pretty small XML documents. In the scope of our research, we analyze the storage, query, update and visualization of large XML instances. BaseX is our native XML database prototype. It features compact storage structures, a very efficient XPath implementation and a graphical user interface, facilitating visual access to large XML documents.

Advisors: 1. Prof. Dr. Marc H. Scholl

Publications:

Melting Pot XML - Bringing File Systems and Databases One Step Closer. Alexander Holupirek, **C. Grün,** and Marc H. Scholl. In 12. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW 2007), pages 309--323, Aachen, Ger., March 2007.

Visually Exploring and Querying XML with BaseX.

 C. Grün, Alexander Holupirek, and Marc H. Scholl. In 12. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW 2007), Aachen, Ger. March 2007.

Pushing Xpath Accelerator to its Limits. **C. Grün,** Alexander Holupirek, Marc Kramis, Marc H. Scholl, and Marcel Waldvogel. In Proc. of the First Int'l Workshop on Experimental and Performance Evaluation of Data Management Systems, ACM, June 2006.

Pushing XML Main Memory Databases to their Limits. **C. Grün.** In Proc. 18th GI Workshop Foundations of Database Systems, Lutherstadt Wittenberg, Ger. June 2006.

MedioVis – a User-Centred Library Metadata Browser. **C. Grün,** Jens Gerken, Hans-Christian Jetter, Werner König, and Harald Reiterer. In Research and Advanced Technology for Digital Libraries, Proc. 9th Eur. Conf. ECDL, in Lecture Notes in Computer Science (LNCS), pages 174--185, Springer Verlag, Vienna, Austria, Sept. 2005.

HyperGrid –Accessing Complex Information Spaces. Hans-Christian Jetter, Jens Gerken, Werner König, **C. Grün**, and Harald Reiterer. In People and Computers XIX - The Bigger Picture, Proc. of HCI 2005, in Lecture Notes in Computer Science (LNCS), Springer Verlag, Edinburgh, United Kingdom, Sept. 2005.

Zoomtechniken zur Exploration komplexer Informationsräume am Beispiel "HyperGrid". Harald Reiterer, Hans-Christian Jetter, Werner König, Jens Gerken, and **C. Grün.** In Mensch & Computer 2005: Kunst und Wissenschaft - Grenzüberschreitungen der interaktiven ART, in Lecture Notes in Computer Science (LNCS), Oldenbourg Verlag, Linz, Austria, Sept.2005.

DeepFS - Querying the file system

Alexander Holupirek

The long term perspective of DeepFS is to find synergies between file system and semi-structured database techniques. While file systems provide an easy and well-understood interface to the data, they lack important and demanded features like, for example, the ability to query the data.

DeepFS breaks with the long tradition to consider a file merely as a sequences of bytes. It unseals the black-box and lets the classical file hierarchy emerge into the files itself. The consideration of content and structure opens the door for query languages that operate on semi-structured data.

The project is tightly coupled to BaseX, an emerging database for semi-structured data, and Idefix, a block-oriented persistent storage layer for semi-structured data. Currently we are following the concept of a joint storage for both, database and (a user level implementation) of DeepFS.

Finally DeepFS will provide both, proven and stable access to the data leveraging file system techniques and query support for all stored files.

Advisors: 1. Marc H. Scholl

Publications:

Holupirek, A., Grün, C., Scholl, M., Melting pot XML - Bringing file systems and databases one step closer, Proceedings GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW 2007), Aachen, Germany, March 2007, pp. 309—323;

Grün, C., **Holupirek, A.**, Scholl, M., Visually exploring and querying XML with BaseX, Proceedings 12. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW 2007), Aachen, Germany, March 2007;

Grün, C., **Holupirek, A**., Kramis, M., Scholl, M., Waldvogel, M., Pushing XPath Accelerator to its limits, Proceedings First International Workshop on Performance and Evaluation of Data Management Systems (ExpDB06), ACM, 2006;

Holupirek, A., Implementing file systems by XML-aware databases, University of Konstanz, July 2006.

Interaction Techniques for High Resolution Displays

Werner König

In account of a typical display size of more than 200 inches and a resolution of more than seven megapixels, large, high-resolution displays (LHRD) are capable of visualizing complex and large data without sacrificing context or detail. They open new perspectives for collaborative work in multi user / multiple devices environments and offer great opportunities for pinboard-like spatial information management. With the new capabilities the users get confronted with so far unknown information amount and density on interactive displays. LHRDs introduce a number of interaction challenges not addressed by traditional input devices and conventional graphical user interface metaphors.

The purpose of this thesis as part of the research project "information at your fingertips - interactive visualization for gigapixel displays" funded by Förderprogramm Informationstechnik Baden-Württemberg (BW-FIT) is the analysis of existing and conception of new interaction and visualization techniques for high resolution displays in consideration of human capabilities and restrictions. In particular Zoomable User Interfaces approaches with Semantic Zooming are investigated regarding their applicability on LHRDs in a consistent and user oriented concept. Also the comparison of recently proposed input mechanisms based on tracking systems with unconventional input devices for LHRDs like PDAs, UMPCs (Ultra Mobile PCs) and game controllers are in the scope of the project. Thereby PDAs are considered as very promising, since they can offer complex and parallel user interaction with additional visual feedback and orientation possibilities due to their built-in display.

Furthermore the conception and implementation of a generic interaction framework for high resolution displays based upon the preceding analysis is a part of the project. Besides integrating PDA-Interaction, 3D-Flysticks, Hand-Gestures, mobile Eye-Trackers and further flexible input devices into the interaction framework, we particularly focus on Laserpointer-Tracking, which allows very intuitive and direct interaction with large, high resolution displays without restricting user's position and distance to screen.

A key issue with regard to Semantic Zooming and evaluations is tracking of user's position, orientation and her attention in single or multi user environments. Primarily fundamental question regarding screen capturing, user tracking and interaction logging should be answered. Upcoming empirical evaluations will focus on interaction and visualization techniques, diverse input devices and the designed framework itself.

Advisors: 1.Prof. Dr. Harald Reiterer

Publications:

Laserpointer-Interaktion für große, hochauflösende Displays, **König, Werner** A.; Bieg, Hans-Joachim; Reiterer, Harald Proceedings of Mensch & Computer 2007: Interaktion im Plural, Oldenbourg Verlag, Sep 2007;

Jetter, H. C., **König, W.**, Gerken, J., Reiterer, H., MedioVis: Neue Formen der Navigation in Webdatenbanken, Proceedings Usability Day IV: Information nutzbar machen. FH Vorarlberg, Dornbirn, Pabst Publishers, June 2006, pp. 126-129

Reiterer, H., Jetter, H. C., **König, W.**, Gerken, J., Grün, C., Zoomtechniken zur Exploration komplexer Informationsräume am Beispiel HyperGrid, Proceedings Mensch & Computer 2005: Kunst und Wissenschaft - Grenzberschreitungen der interaktiven ART, LNCS, Oldenbourg-Verlag, Sept. 2005.

Jetter, H. C., Gerken, J., **König, W**., Grün, C., Reiterer, H., HyperGrid - Accessing complex information spaces, Proceedings HCI 2005, LNCS, Springer-Verlag, Sept. 2005.

Extending the OLAP Technology to Support Non-Conventional Application Domains

Svetlana Mansmann(geb. Vinnik)

Comprehensive data analysis has become indispensable in a variety of real-world applications. However, business-oriented OLAP (On-Lin eAnalytical Processing) systems tend to perform poorly or even fail when applied to complex data scenarios found in novel application domains, such as government, science, medicine etc. The causes of the deficiencies are manyfold, from the rigidness of the underlying multidimensional data model, the conventional logical schemata, and the data transformation techniques to the established metadata modeling standards and end-user frontend issues, such as visual navigation and query specification, available visualization and interaction techniques, and user friendliness.

Our research is concerned with overcoming the above mentioned limitations of the relational OLAP systems by analyzing the requirements of non-conventional application scenarios towards such systems and proposing the modeling extensions at the conceptual, logical and metadata levels as well as their implementation in a visual exploration framework. The proposed modeling extensions are captured accordingly at the metadata level and can be handled using the standard database query language SQL.

We verify our solutions by implementing them in a prototype OLAP system "DataCube Explorer". The frontend offers a powerful context-aware schema-based navigation for multidimansional data cubes, with support for user-defined measures and parallel exploration of multiple fact tables (Multicube Join). Query results are presented in form of a classical pivot table or using more elaborate visualization techniques.Our contribution in the field of visual exploration is focused on designing a class of hierarchical visualization techniques called Enhanced Decomposition Trees, which arrange iteratively computed aggregates into a decomposition hierarchy. Various tree layouts and embedded visualization techniques are tested to account for a wide range of analysis tasks. The feasibility of the proposed enhancements is verified using realistic study cases from the fields of academic administration and medical research.

Advisors: 1. Prof. Dr. Marc H. Scholl

Publications:

Mansmann, S., Scholl, M., Decision support system for managing educational capacity utilization, IEEE Transactions on Education, to appear, 2007.

Mansmann, S., Scholl, M., Empowering the OLAP technology to support complex dimension hierarchies, International Journal of Data Warehousing and Mining, to appear, Oct. 2007.

Mansmann, F., **Vinnik, S.,** Interactive exploration of data traffic with hierarchical network maps, Transactions on Visualization and Computer Graphics, Vol. 12, No. 6, Nov.-Dec. 2006;

Mansmann, S., Mansmann, F., Scholl, M., Keim, D., Hierarchy-driven exploration of multidimensional data cubes, Proceedings 12. GI-Fachtagung für Datenbanksysteme in Business,

Technologie und Web (BTW 2007), Aachen, Germany, 2007, pp. 96 – 111;

Mansmann, S., Scholl, M., Exploring OLAP aggregates with hierarchical visualization techniques, Proceedings 22nd Annual ACM Symposium on Applied Computing (SAC 2007), Multimedia & Visualization Track, Seoul, Korea, March 2007, pp. 1067 – 1073;

Mansmann, S., Scholl, M., Extending visual OLAP for handling irregular dimensional hierarchies, Proceedings 8th International Conference on Data Warehousing and Knowledge Discovery (DaWaK'06), Krakow, Poland, Springer-Verlag, September 2006, pp. 95-105;

Mansmann, S., Scholl, M., Keim, D., Mansmann, F., Interactive poster: exploring OLAP aggregates with hierarchical visualization techniques, Proceedings IEEE Symposium on Information Visualization (InfoVis 2006), Oct. 2006;

Vinnik, S., Mansmann, F., From analysis to exploration: Building enhanced visual hierarchies from OLAP cubes, Proceedings 10. International Conference on Extending Database Technology (EDBT), Munich, Germany, Springer-Verlag, March 2006, pp. 496-514;

Vinnik, S., Scholl, M., UNICAP: Efficient decision support for academic resource and capacity management, Proceedings TED Conference on e-Government (TCGOV2005), Bozen-Bolzano, Italy, 2005.

Vinnik, S., Scholl, M., Decision support system for managing educational capacity utilization in universities, Proceedings International Conference on Engineering and Computer Education (ICECE'05), Madrid, Spain, 2005.

Automatic and visual text analysis

Daniela Oelke

Due to the ever growing data volume, today, an automatic analysis is indispensible. However, most of the existing approaches focus on structured data. This excludes the huge amount of information that is represented in textual form. One of the reasons why an automatic analysis of text is challenging is that to properly understand a text not only the words that are used are important but also the context they are used in, and understanding them in the context it difficult to achieve algorithmically. There are different aspects of text that can be considered in an analysis such as the phonological level, the morphological level, the syntactical level, the semantical level and - if the text has been edited by the author - the structural level. Extracting features that reflect the characteristics of a text on all levels is the first step towards an automatic analysis, whereas some aspects, such as the semantics of a text, are very difficult to grasp computationally. The goal of this thesis is to improve and apply automatic text analysis methods for specific applications, such as authorship attribution, translation criticism and language evolution. We also plan to automatically enrich text with additional features to support the manual text analysis. To achieve this, new visual analytics methods will be developedwhich are based upon sophisticated text feature analysis methods and use novelvisualization techniques to allow an improved text understanding. The visual methods enable the user to understand the development of the feature values over the text and to create new features. Key for the success is the combination of automated methods with visualization-based user steering.

Advisors: 1. Prof. Dr. Daniel Keim

Publications:

Keim, D., **Oelke, D.**, Truman, R., Neuhaus, K., Finding correlations in functionally equivalent proteins by integrating automated and visual data exploration, Proceedings 6th IEEE Symposium on BioInformatics and BioEngineering (BIBE 20

06), 2006;

Belle, S. K., **Oelke, D.,** Oettl, S., Sips, M., Exploration of local distribution of major ethnic groups in the USA, Proceedings IEEE Symposium on Information Visualization (InfoVis 2006), October 29 - November 03, Baltimore, Maryland, USA, 2006, poster paper.

Oelke, D., Datenanalyse in der Molekularbiologie: Suche nach Abhängigkeiten in alignierten Sequenzen funktional gleicher Proteine, University of Konstanz, March 2006.

Applications of Multidimensional Scaling to Graph Drawing

Christian Pich

Multidimensional Scaling (MDS) is a family of techniques for analysis and visualization of complex data. Objects in a data set are represented as points in a geometric space; distance in this space represents proximity or similarity among objects. We study MDS in the context of network visualization and develop efficient algorithms that are scalable to very large instances. The usability of our approaches is illustrated by several applications.

Advisors: 1. Prof. Dr. Ulrik Brandes

Publications:

Brandes, U., **Pich, C.,** Centrality estimation in large networks, International Journal of Bifurcation and Chaos, special issue on , to appear, 2007;

Fleischer, D., **Pich, C.,** Positioning and virtual coordinates, Proceedings Dorothea Wagner and Roger Wattenhofer (Eds.): Algorithms for Sensor and Ad-Hoc Networks, to appear, 2006, within the GI-Dagstuhl research seminar.

Brandes, U., Hoefer, M., **Pich, C.,** Affiliation dynamics with an application to movie-actor biographies, Proceedings 8th Eurographics/IEEE-VGTC Symposium on Visualization (EUROVIS 2006), May 2006, pp. 179-186.

Brandes, U., **Pich, C.,** Eigensolver methods for progressive multidimensional scaling of large data, Proceedings 14th Intl. Symp. Graph Drawing (GD '06), 2006, pp. 42-53.

Brandes, U., Lerner, J., **Pich, C.,** GXL to GraphML and vice versa with XSLT, Proceedings 2nd Intl. Workshop Graph-Based Tools (GraBaTs '04), Elsevier Science, 2005, pp. 113-125.

Efficient Rate-Distortion Optimized Media Streaming

Martin Röder

We consider the problem of finding optimal transmission strategies in rate-distortion optimized media streaming over a packet erasure channel with random delay. A compressed media stream, such as an audio or video stream, that is split into a set of usually interdependent data units, is to be sent to a receiver for instant playback. The data unit dependencies are modeled as a directed acyclic graph. The goal is to find transmission strategies that minimize the expected playback distortion for a given transmission rate and the loss and delay characteristics of the channel. The problem can be solved in two steps. The first step is to find all optimal transmission policies for a single data unit, the second step is to combine these optimal transmission policies into optimal transmission strategies for a group of interdependent data units. We develop algorithms for both steps.

Advisors: 1. Prof. Dr. Raouf Hamzaoui 2. Prof. Dr. D.Saupe

Publications:

Röder, M., Hamzaoui, R., Fast tree-trellis list Viterbi decoding, IEEE Transactions on Communications, Vol. 54, pp. 453-461, March 2006.

Röder, M., Cardinal, J., Hamzaoui, R., Branch and bound algorithms for rate-distortion optimized media streaming, IEEE Transactions on Multimedia, Vol. 8, pp. 170-178, Feb. 2006.

Röder, M., Cardinal, J., Hamzaoui, R., Efficient rate-distortion optimized media streaming for tree-reducible packet dependencies, Proceedings SPIE vol. 6071, MMCN'06, San Jose, Ca., Jan. 2006.

Röder, M., Wu, X., Dumitrescu, S., Error resilient multiple description compression of vector graphics, Proceedings IEEE International Conference on Multimedia and Expo, ICME'06, Toronto, Canada, July 2006.

Cardinal, J., Hamzaoui, R., **Röder, M.**, Constructing dependency trees for rate-distortion optimized media streaming, Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'06, Toulouse, May 2006.

Röder, M., Cardinal, J., Hamzaoui, R., Dynamic programming algorithm for rate-distortion optimized media streaming, Proceedings IEEE International Conference on Image Processing (ICIP'05), Genoa, Sept. 2005

Clustering in Parallel Universes

Bernd Wiswedel

Classical data analysis methods typically assume that all objects of a data set are described in a single feature space. This feature space is assumed to comprise all necessary information to classify an object. However, in many real-world applications there are numerous ways to describe complex objects. An example are musical songs, i.e. audio streams, which can be represented based on dynamics, melody, and key or - as a different representation - based on rhythm and harmony. A third representation may be more descriptive, such as interpreter, position in music charts, length, and so on. Further examples of complex objects are images, 3D objects or molecules in drug discovery. With regard to learning, such as clustering or building classification models, it is often unclear, which of the available descriptors are optimal for any given task. Clustering in Parallel Universes is a new research field that deals with such multiply described data sets. It aims at identifying interesting patterns in data, e.g. groups of objects that cluster well in one (or few) universe(s).

This thesis focuses on novel techniques to learn in Parallel Universes.It begins by defining the concept of Parallel Universes nd drawing differences to groups of machine learning techniques that also take special properties of the underlying feature space(s) into account.These include subspace clustering, multi-view or multi-instance algorithms, to name but a few.

To illustrate the concept of Parallel Universes, this thesis presents two new clustering algorithms: The first method is a supervised clustering technique, which focuses on a model for a particular (minor) class of interest by constructing local neighborhood histograms, so-called Neighborgrams for each object of interest in each universe. The algorithm assigns a quality value to each Neighborgram and greedily includes the best Neighborgram, no matter from which universe it stems, in the global prediction model. Objects that are covered by this Neighborgram are finally removed from consideration in a sequential covering manner. This process is repeated until the global model has sufficient predictive power. We will also demonstrate how the one-dimensional visualization of Neighborgrams can be used to interactively explore a dataset and help the user to manually or semi-automatically construct an interpretable model of the data.

Secondly, we will present an unsupervised clustering method based on the fuzzy c-means algorithm. The method uses fuzzy membership values to encode degrees of clustering contributions for each object to each universe. The algorithm iteratively learns these newly introduced values and - similar to the classical c-means - partitioning values in each universe. The outcome of the algorithm are clusters spread across different parallel universes, each cluster modeling only a small subset of the data. We will furthermore discuss two extensions to this basic algorithm to overcome the negative influence of noise and/or outliers (caused by the partitioning property of the underlying algorithm) and the problem of specifying the number of clusters in advance. Firstly, we will show how incorporating an additional term to the global objective reduces the influence of noise. This term represents a virtual noise universe with one single cluster, attracting objects that do not contribute to any of the clusters and hence reducing their impact on the cluster formation. The second

extension discards the side constraint requiring that universe memberships need to sum to one, allowing potential overlaps of clusters or empty clusters. It is therefore possible to overestimate the number of clusters as meaningless clusters can be easily identified.

Advisors: 1.Prof. Dr. Michael Bertold

Publications:

Wiswedel, B., Berthold, M., Fuzzy clustering in parallel universes, International Journal of Approximate Reasoning, submitted, 2006;

Berthold, M., Cebron, N., Dill, F., Di Fatta, G., Gabriel, T., Georg, F., Meinl, T., Ohl, P., Sieb, C., **Wiswedel, B.**, Knime: the Konstanz Information Miner, Proceedings International Workshop on Multi-Agent Systems and Simulation (MAS&S), 4th Annual Industrial Simulation Conference (ISC 2006), Palermo, Italy, June 2006, pp. 58-61.

Wiswedel, B., Berthold, M., Fuzzy clustering in parallel universes, Proceedings Conference North American Fuzzy Information Processing Society (NAFIPS 2005), 2005, pp. 567-572;

Wiswedel, B., Berthold, M., Fuzzy clustering in parallel universes with noise detection, Proceedings IEEE International Conference on Data Mining, Workshop Computational Intelligence in Data Mining (ICDM 05), Houston, TX, USA, 2005, pp. 29-37.

GRK 1076:

TrustSoft

Vertrauenswürdige Software-Systeme

Oldenburg

Graduate School TrustSoft:
Trustworthy Software Systems

Wilhelm Hasselbring
hasselbring@informatik.uni-oldenburg.de,
Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany
http://trustsoft.uni-oldenburg.de/

Software increasingly influences our daily life, as we depend on an raising number of technical systems controlled by software. Additionally, the ubiquity of Internet-based applications increases our dependency on the availability of those software systems. Exemplarily consider complex embedded software control systems in the automotive domain, or IT systems for eGovernment and eHealth.

Fortunately, the rise of the software industry creates jobs for academically trained professionals and generates an increasing proportion of the national creation of value. However, this increased dependency on software systems intensifies the consequences of software failures. Therefore, the successful deployment of software systems depends on the extent we can trust these systems. This relevance of trust is gaining awareness in industry. Several software vendor consortia plan to develop so-called *Trusted Computing* platforms. These current initiatives primarily focus on security, while trust is a much border concept. In fact, trust is given by several properties, such as safety, correctness, reliability, availability, privacy, performance, and certification.

Therefore, the graduate school will contribute to this comprehensive view on trusted software systems by bundling the Oldenburg computing science competences with those of computer law. Our research in computer science and computer law will advance new interdisciplinary scientific methods for creating trustworthy software systems and to educate highly qualified graduates who transfer into practice the required skills for constructing and certifying trustworthy software systems.

From a technical point of view, the research programme of the graduate school builds on and advances the paradigm of component-based software engineering. Besides the industrial relevance of components, components also constitute a more general paradigm employed successfully in the areas of formal verification (compositional reasoning), the prediction of quality properties, and the certification of software systems. The scientific methods to be developed in the graduate school vary according to the aspects of trust under investigation. For example, correctness is demonstrated by mathematical proofs while quantifiable quality properties, such as availability, reliability, and performance require analytical prediction models, which need additional empirical studies for calibration and validation.

Graduiertenkolleg TrustSoft:
Vertrauenswürdige Software-Systeme

Wilhelm Hasselbring
hasselbring@informatik.uni-oldenburg.de,
Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany
http://trustsoft.uni-oldenburg.de/

Unter dem Begriff Vertrauenswürdigkeit von Software werden in der Regel verschiedene Qualitätsattribute zusammengefasst: Korrektheit (engl. correctness), Zuverlässigkeit (engl. reliability), Verfügbarkeit (engl. availability), Performanz (engl. performance), Sicherheit (engl. safety und security) und Einhaltung von Datenschutzrichtlinien (engl. privacy).

Der Aspekt der Vertrauenswürdigkeit von Software-Systemen hat international an Bedeutung gewonnen, sowohl in der Forschung als auch in vielen Anwendungsgebieten. Aus den folgenden Gründen ist davon auszugehen, dass sich dieser Trend in der Zukunft weiter verstärken wird:

- Die Einsatzmöglichkeiten von Software erweitern sich ständig. Software-Systeme werden sich aber nur dann in vielen neuen Anwendungsgebieten etablieren können, wenn sie sich von vornherein als vertrauenswürdig demonstrieren lassen. Beispiele sind vernetzte, personalisierte Informationssysteme im Internet und eingebettete Systeme in der Automobiltechnik.

- Die rechtliche Praxis, Software-Systeme prinzipiell nicht nach den in den Ingenieurwissenschaften üblichen Maßstäben zu zertifizieren und Hersteller dieser Systeme weitgehend aus der Haftung zu entlassen, wird auf Dauer nicht haltbar sein. Vielmehr werden Anbieter, die eine ingenieurmäßige Zertifizierung nachweisen und die Übernahme von Haftungsverpflichtungen garantieren können, einen massiven Wettbewerbsvorteil erringen.

Im Forschungsprogramm des Graduiertenkollegs soll insbesondere der Ansatz der komponentenbasierten Software-Entwicklung weiterentwickelt werden. Dabei wird der Begriff der Komponentenbasierung bewusst paradigmatisch weit gefasst, d.h., darunter verstehen wir hier neben der Komposition von Software aus bestehenden Software-Komponenten auch Verfahren zum kompositionellen Schließen sowie analytische Vorhersagemodelle für Qualitätseigenschaften. Ein wesentlicher Vorteil dieses Ansatzes besteht auch darin, dass komplexe Software-Systeme so schrittweise aus Komponenten konstruiert und bewertet werden können. Komponenten können rekursiv wieder aus Komponenten bestehen. Ebenso gehören zur Komponentenbasierung auch juristische Fragen des Haftungsanspruches bei Software, die aus Komponenten evtl. verschiedener Hersteller gefertigt wird und daraus sich ergebende spezifische Anforderungen an die Zertifizierung von Komponenten und komponentenbasierter Systeme.

Model-driven Performance Measurement and Estimation with Relational Traces
TrustSoft Research Training Group, University of Oldenburg

Marko Bošković

marko.boskovic@informatik.uni-oldenburg.de

Model-driven Engineering (MDE) is a software development approach which tries to move the focus of software development from the implementation level, in to the problem domain. Model-driven Architecture (MDA) is one of the approaches for MDE in which core is the Unified Modeling Language (UML).

Non-functional properties, such as performance, are becoming more and more important characteristic of current software systems. Performance is defined as the degree to which a software system or a component meets its objectives for timeliness. It essentially is timing behavior of an application.

Temporal databases are databases which introduce a notion of time in to relational data model. This enables time dependent analysis of stored data.

At the present moment, in the MDE, most of research is dedicated to the performance prediction with a simulation and analytical models. Nevertheless, predictions have to be empirically estimated. It is done by inserting additional code for data collection during testing, and later computation of metrics. Techniques for at the code level, they are suitable for MDE.

The main challenge is to introduce the approach for empirical measurement and estimation of software performance on the model level. This implies that data are collected and computed in the terms of modeling constructs. Furthermore, the specification of data to be collected and metrics to be computed should be specified at the model level.

The approach presented in this thesis suggests a declarative specification of performance metrics, and employment of temporal databases for data storage and analysis. Declarative specification liberate developer from specification of code for data collection and metrics computation. Temporal databases provide temporal algebras for metrics computation.

The main contribution of this thesis is the integration of the approach in the form of a UML Profile for Instrumentation and Performance Estimation (PIPE). Furthermore, to support the profile in the database management systems independent way, we define our own algebra with the set of aggregates. Finally, we define transformations from the profile to the algebra. The profile will be implemented using MagicDraw tool for UML Modeling and the MySQL as a database management system. It will be applied to Class diagrams and Statecharts. Simultaneously with the code for the basic functionality, code for data collection and metrics computation will be generated.

An Algebraic Framework for the Composition of Self-Stabilizing Distributed Algorithms

Abhishek Dhama
Department of Computer Science,
Carl von Ossietzky University of Oldenburg
`abhishek.dhama@informatik.uni-oldenburg.de`

Ever increasing integration of distributed systems in daily lives for safety-critical operations has put high premium on their dependability. High dependability of such mission-critical distributed systems can be ensured by endowing them with *fault tolerance*. *Stabilization* techniques are used to introduce *non-masking* fault tolerance. A system **S** is self-stabilizing, if and only if, irrespective of its starting state it is *guaranteed* to reach a *legal* set of states within finite number of steps and does not leave this set voluntarily. The legal set of states is defined with the help of a predicate $\mathcal{P}$. A self-stabilizing system is able to "recover" from bursts of *transient* faults independent of the fault span.

Even though self-stabilization is a desirable property of a distributed algorithm (that is executed by a distributed system), its verification is not trivial. The importance of self-stabilization entails a *holistic* and systematized approach for the design and verification of the self-stabilizing algorithms. Thus, the primary aim of this dissertation is to build an algebraic framework for the analysis and synthesis of the self-stabilizing algorithms. In our framework a "property-endowed" distributed algorithm is the basic operand. We define three basic composition operations on the operands, namely *symmetric* composition, *feedback* composition and *interdependent* composition. In order to make the algebraic framework more powerful we would also like to investigate the properties of the basic operators. We would like to determine the strongest algebraic structure that the proposed algebraic framework forms. In order to provide simpler interpretations to the complex algebraic expressions, we would like to determine whether we can "compose" operators themselves from the basic operations. We plan to evaluate the proposed algebraic framework by utilizing its theorems for various case studies. We have divided our case studies into three major categories. 1) We plan to use the theorems of the proposed algebra to construct large self-stabilizing algorithms as part of *synthetic* case studies. 2) We would like to use inverse operations and heuristics for *analytic* case studies. 3) We would use the theorems to determine the unknown potentially self-stabilizing components to build a self-stabilizing system with a given specification and known components. Due to the non-static nature of sensor networks and their increasing prevalence, we believe that they form suitable testbed for the planned case studies.

Performance Analysis
of Concurrent, Component-based Software Systems

Jens Happe

Graduate School TrustSoft (GRK 1076)
Carl-von-Ossietzsky-Universität Oldenburg
Uhlhornsweg, 26111 Oldenburg
jens.happe@informatik.uni-oldenburg.de

In software development, extra-functional requirements are often not considered until late development stages, i.e. during testing. This puts the success of a project at unnecessary high risk. If extra-functional requirements cannot be met, necessary changes can be expensive and time consuming. To minimise such risks, model-based performance prediction methods offer means to evaluate the response time, throughput, and utilisation of a software system based on its architectural description. To derive the performance metrics of interest, design models, e.g. UML diagrams with performance annotations, are transformed to analytical models.

The most established classes of a large variety of analytical methods are queueing network models (QNM), stochastic process algebras (SPA), and stochastic Petri nets (SPN). Most of these formalisms base on Markov chains, which imply a set of strict mathematical assumptions. Unfortunately, these assumptions do not hold in reality for most systems. On the other hand, simulation-based methods impose weaker assumptions allowing more realistic predictions. However, simulation is time consuming, especially if high confidence in the results is required.

Concurrency amplifies this conflict, since analytical methods require stronger assumptions while simulation becomes more time consuming. With the rise of symmetric multi-core processors in the area of common desktop and server systems, concurrency becomes an important factor for developing efficient applications. Thus, it has to be considered for performance predictions.

The contribution of the thesis is twofold. First, we develop a new process algebra based on SPADES, an SPA for discrete event simulation. Our algebra employs a hybrid approach of analysis and simulation methods. Furthermore, it considers influences of resource contention, scheduling policies, and multi-core systems on performance. Second, we introduce a set of model abstractions for concurrent behaviour based on concurrency design patterns. These abstractions capture only the performance relevant aspects of a pattern. The approach has been validated with a set of smaller case studies. Predictions and measurements deviate less than 5%, if the model assumptions hold. The case studies pointed to some additional performance relevant factors, which need to be considered. We also plan a larger case study to show the benefit of our prediction method.

Energie Optimisation of Embedded Hardware/Software Systems

Henrik Lipskoch

henrik.lipskoch@informatik.uni-oldenburg.de

Graduate School TrustSoft, University of Oldenburg, Germany

Specialised Software running on tailored and optimised Hardware is the field of Embedded Systems. As examples this application specific combination of Hardware and Software can turn out as mobile camera sensors for object idenfication or personal digital assistants (PDAs). Mobile in this context is understood as constraints on size and weight as well as independency of the power line.

Out of the function the system has to provide several requirements are derived. For example some parts of the software have to fulfill certain time constraints such that their computational results can be used by other following software tasks and do not have to be discarded or lead to bad circumstances. If there would be an object identification programme discovering an obstacle for a vehicle too late, the brake assistant software probably would not have enough time to do a smooth slowdown or even worse it would come to a collision. A requirement is called hard real-time if it is guaranteed to be fulfilled after a certain deadline or be discarded otherwise so that it is either delivered and in time or not delivered.

Another requirement follows out of the independence of the power line. To provide mobility embedded systems carry batteries or even fuel cells. Since the weight and size of the system has limits the capacity of the energy source is bounded. On the other hand the energy in the source needs to be high enough to guarantee the complete function processing (capacity) and the source must be able to deliver the required energy within a certain time span (power). Because of the limits in the dimensions it is likely that other battery technologies are required as the ones planned. As an example Nickel-Cadmium cells have a two-third less energy density than Lithium-Ion types.

On providing real-time requirements the event-stream methodology has showed to be very efficient and appropriate to provide fast and accurate feasibility tests. Other tests on real-time feasibilty exist but the event-stream model can cover the same timing behaviour of the software tasks with less complexity, because it abstracts from the set of all possible schedules to some kind of worst-case scenario. And if the worst-case scenario is feasible so are the other schedules.

The same methodology can be applied to energy behaviour by replacing the execution times with the energy values. The resulting questions and applications are the focus of this research.

A Petri Net Semantics for Pi-Calculus Verification

Roland Meyer

Graduate School on Trustworthy Software Systems

Roland.Meyer@informatik.uni-oldenburg.de

In dynamically reconfigurable systems the interconnection structure between system parts is not static but evolves over time. Prominent examples of this system class are service-oriented architectures where required functionality is plugged in during runtime, ad-hoc networks of devices that enter, share, and leave a wireless network, or object-oriented programs where object instances are created dynamically and references are passed.

An essential factor of success for every software and hardware system is that it correctly provides the functionality the user expects. Automated verification has become widely accepted to ensure the correctness of hardware systems and distributed algorithms and currently wins recognition in the verification of software. In dynamically reconfigurable systems, besides all problems that come with the distributed nature, bugs arise due to the unforseen collaboration of incompatible parts. We believe that automated verification methods will help to ensure the correct functioning for these systems.

For modelling dynamically reconfigurable systems, we use the Pi-Calculus [Mil99, SW01], a well established process algebra. It relies on a small number of operators and rules which makes it easy to understand but also theoretically appealing. One may hope that verification approaches, when developed for this basic model, can be extended to more complex ones in the future.

For automatic verification purposes, we translate Pi-Calculus processes into place/transition Petri nets. Finite Petri nets come with efficient verification techniques that exploit the graph structure for reasoning, [Sta90], compared to uninformed state space enumeration techniques. Furthermore, finite Petri nets are not Turing complete. Thus, for several verification problems exact algorithms are known, [Esp97].

As mentioned, our approach to Pi-Calculus verification is based on a semantical mapping that translates Pi-Calculus processes into place/transition Petri nets. Different from known approaches (cf. [MP95, BG95, Eng96, AM02, FGMP03, DKK06]) that split the given processes along seqential subprocesses, we decompose along common substructures. We therefore derive finite Petri net representations for processes where neither the number of sequential subprocesses nor the number of restricted names is bounded.

Since the Pi-Calculus is Turing complete and finite Petri nets are not, our translation may lead to infinite nets. For applying standard Petri net verification techniques, finiteness is a prerequisite. A main result of our work is the characterisation of processes that are

translated into finite nets. Having understood the sources of infinity, we identify a rich syntactic class of processes with finite Petri net semantics. With case studies we provide evidence that, using our translation, different kinds of properties can be verified by efficient techniques for various systems.

References

[AM02]   R. M. Amadio and C. Meyssonnier. On Decidability of the Control Reachability Problem in the Asynchronous π-Calculus. *Nordic Journal of Computing*, 9(1):70–101, 2002.

[BG95]   N. Busi and R. Gorrieri. A Petri Net Semantics for π-Calculus. In *Proceedings of the 6th International Conference on Concurrency Theory, CONCUR 1995*, volume 962 of *LNCS*, pages 145–159. Springer-Verlag, 1995.

[DKK06]   R. Devillers, H. Klaudel, and M. Koutny. A Petri Net Translation of π-Calculus Terms. In *Proceedings of the 3rd International Colloquium on Theoretical Aspects of Computing, ICTAC 2006*, volume 4281 of *LNCS*, pages 138–152. Springer-Verlag, 2006.

[Eng96]   J. Engelfriet. A Multiset Semantics for the pi-Calculus with Replication. *Theoretical Computer Science*, 153(1-2):65–94, 1996.

[Esp97]   J. Esparza. Decidability of Model Checking for Infinite-State Concurrent Systems. *Acta Informatica*, 34(2):85–107, 1997.

[FGMP03]   G.-L. Ferrari, S. Gnesi, U. Montanari, and M. Pistore. A Model-Checking Verification Environment for Mobile Processes. *ACM Transactions on Software Engineering and Methodology*, 12(4):440–473, 2003.

[Mil99]   R. Milner. *Communicating and Mobile Systems: the π-Calculus*. Cambridge University Press, 1999.

[MP95]   U. Montanari and M. Pistore. Concurrent Semantics for the π-calculus. *Electronic Notes in Theoretical Computer Science*, 1:411–429, 1995.

[Sta90]   P. H. Starke. *Analyse von Petri-Netz-Modellen*. B. G. Teubner, 1990.

[SW01]   D. Sangiorgi and D. Walker. *The π-calculus: a Theory of Mobile Processes*. Cambridge University Press, 2001.

Liability for Defective Open Source Software

Alexandra Gudrun Paul

TrustSoft
Graduate School on Trustworthy Software Systems
University of Oldenburg
Adviser: Prof. Dr. Juergen Taeger
Alexandra.Paul@offis.de

With regard to the distribution of open source software there are two contracts to be distinguished: First the acquirer of a copy concludes a contract with the distributor. The type of this contract depends on the specific conditions. In most cases it is a donation or a contract of sale. Together with the copy the acquirer receives a legal license to run the program; i.e. as long as the acquirer only runs the program he does not need to conclude a license agreement with the copyright holder. But as soon as the acquirer wants to copy, modify, distribute or make available the program he has to conclude an Open Source License with the copyright holder. With conclusion of this agreement the acquirer receives the right to undertake acts of use which exceed the act of running the program.

The contractual liability depends on the question which contract is impaired. In principle material defects impair the donation or contract of sale while defects in title impair the license agreement. Apart from the fact that the disclaimers of liability contained in Open Source Licenses are invalid according to German civil law, these disclaimers are not able to affect the relationship between distributor and acquirer because they are part of the license agreement between licensor and licensee. The contractual liability for software defects complys with the agreement between distributor and acquirer or, if there is no explicit agreement, with the regulations of the Civil Law Code. According to the Civil Law Code the giver of a gift is only liable in cases of intent or gross negligence while the seller is responsible to provide an object free of faults. If the software is defective the seller has to repair and he is liable for every damage caused by negligence or intent.

Concerning non-contractual liability the main problems are product liability and manufacturer liability. According to the Product Liability Law the producer is - without consideration of fault! - liable for any damages caused by his product. Among lawyers it is widely accepted that software is a product. Hence the Product Liability Law is applicable to software. A product is defective if it does not provide that degree of safety which can be justifiably expected having regard to the circumstances. Producer - and therefore liable – can be a developer community as well as a single software engineer or a committee which decides on the components of a new release. In case a community is liable the plaintiff can claim the whole damage from a single community member. In the context of manufacturer liability it has to be clarified which specific safety measures the manufacturer has to meet to avoid damages.

Development of Correct Graph Transformation Systems and Programs

Karl-Heinz Pennemann

Graduate school TrustSoft

University of Oldenburg

Pennemann@Informatik.Uni-Oldenburg.de

Graph transformation has many application areas in computer science, such as software engineering or the design of concurrent and distributed systems. It is a visual modeling technique and is expected to play a decisive role in the development of growingly larger and complex systems. However, the use of visual modeling techniques alone does not guarantee the correctness of a design. In context of rising standards for trustworthy systems, there is a growing need for the verification of graph transformation systems and programs. The research of appropriate methods for this purpose is the topic of this thesis.

In the considered computation model, applications of transformation rules form the elementary steps. Intuitively spoken, transformation rules consist of a mapping between two graphs (premise and conclusion) and describe local changes on graphs and graph-like structures. Graph programs are defined by transformations rules, nondeterministic choice, sequential composition and iteration. Programs are able to model transactions that deal with an unbounded number of elements and are computationally complete. As language for the specification of state properties, graphical conditions are investigated and used as they are convenient to describe system requirements as well as suited to infer knowledge about system behavior. Graphical conditions may be seen as a tree of graph mappings equipped with logical symbols and provide an intuitive formalism for structural properties.

According to Dijkstra, the correctness of a program with respect to a pre- and a post-condition can be shown in a classical way by constructing a weakest precondition of the program relative to the postcondition and checking whether the precondition implies the weakest precondition. Hence the problem, if a given system or program is correct with respect to a given specification, can be reduced to the implication problem of conditions. In a first part of my work, it has been shown how to construct weakest preconditions for graph programs and (graphical) postconditions. A second part of my work is to consider fragments of conditions, for which the implication problem is decidable and investigate an approximative solution in the general case. A third part of my work is to approximate invariants of graph programs. All research is done within the framework of adhesive high-level replacement categories. Therefore, the results will be applicable to different kinds of graph replacement systems and petri nets, providing theoretical fundamentals and general concepts for the development of correct rule-based systems and programs.

Differential Logic for Hybrid System Verification
Reasoning about Interacting Discrete and Continuous Change

André Platzer

University of Oldenburg, Department of Computing Science, Germany
TrustSoft Graduate School on Trustworthy Software Systems
`platzer@informatik.uni-oldenburg.de`

Abstract. The behaviour of safety-critical systems typically depends on both the state of a discrete controller and continuous physical quantities. These *hybrid systems* combine continuous evolution along *flows* and discrete jumps. In addition to the challenges of purely continuous dynamical systems, discrete control behaviour needs to be taken into account for verifying hybrid systems. The primary challenge in verification is that hybrid dynamics give rise to complicated quantified mathematical formulas that can neither be handled by purely continuous reasoning (because of the discontinuities caused by the discrete transitions) nor by considering the discrete change in isolation. Frequently, the dynamics of a system further depends on the values of external parameters and on a correct adjustment of internal control parameters, which need to be handled during the analysis.

With the ambition of developing techniques for verifying parametric hybrid systems, we enrich dynamic logic by differential equations to cover systems with interacting discrete and continuous transitions. We introduce differential logics as a family of extensions of dynamic logic for hybrid systems. Differential logics can be used to specify and verify safety statements of hybrid systems with free parameters. As a means for verification, we introduce sequent calculi for differential logic. Our calculi uses an interaction of deductive and algebraic reasoning following the principle of deduction modulo quantifier elimination to verify the interacting discrete and continuous behaviour of hybrid systems. Using our calculus, we prove a parametric inductive safety property of cooperating traffic agents in the European Train Control System. See, e.g., [Pla07a, Pla07b] for details on our approach.

References

[Pla07a] André Platzer. Differential Dynamic Logic for Verifying Parametric Hybrid Systems. In Nicola Olivetti, editor, *Automated Reasoning with Analytic Tableaux and Related Methods, TABLEAUX 2007, Proceedings*, volume 4548 of *LNCS*, pages 216–232. Springer, 2007.

[Pla07b] André Platzer. A Temporal Dynamic Logic for Verifying Hybrid System Invariants. In Sergei Artemov and Anil Nerode, editors, *Logical Foundations of Computer Science, LFCS 2007, Proceedings*, volume 4514 of *LNCS*, pages 457–471. Springer, 2007.

Probabilistic Fault Diagnosis for Software Exceptions in Grid Middleware

Jan Ploski

TrustSoft Graduate School on Trustworthy Software Systems

Jan.Ploski@offis.de

To improve the availability of systems, software engineering research has traditionally focused on constructing correct systems and fault tolerance. In comparison, the related topic of fault diagnosis—troubleshooting existing system installations—has received little attention. Software fault diagnosis consists of explaining the reasons of encountered software errors in order to identify and support the required repair actions.

In contrast with other domains, such as medicine and electronics, in which dedicated tool-supported diagnosis methods are available, diagnosis of software problems is often a manual task which requires considerable effort and expertise. This is true even for errors that are not caused by unknown software defects, but also for ones which occur due to misconfigurations, insufficient resources and improper user interactions. The success of diagnosis thus depends on the skill level of individual troubleshooters and their ability to collect and evaluate data from disparate sources rather than on a systematic approach. Tool support is limited to general-purpose tools such as text-based search engines or very specialized tools such as debuggers, which only execute in software development environments and require substantial expertise.

The available exception handling techniques focus on reporting errors and on the recovery to a consistent program state in case of an error. However, they do not support processing of the generated error reports, or executing tests and querying system state to obtain additional information needed for diagnosis. Most importantly, they do not guide the process of information collection, making it difficult to determine which data is or is not relevant for a particular diagnosis.

The goal of the PhD thesis is to supplement traditional exception techniques by introducing a method of automated fault diagnosis based on the accumulation of diagnostic rules and probabilistic reasoning. The main assumption is that software faults are activated multiple times across different software installations. Thus, the cause-effect relationships observed in initial (manual) diagnoses can be captured and reused during subsequent ones, particularly to drive the collection of just the relevant information from the diagnosed systems. The evaluation of the proposed method occurs in context of Grid middleware used in the ongoing research project WISENT (http://wisent.d-grid.de) and the German Grid initiative (http://www.d-grid.de).

Slicing Petri Nets

Astrid Rakow

astrid.rakow@informatik.uni-oldenburg.de

Graduate School TrustSoft

We are examining an approach for reducing the state space for model checking temporal logic formulas on Petri Nets. As a method that explores the state space, model checking suffers under the state explosion problem: the number of states grows exponentially in the number of variables and processes of the system.

We consider systems specified as Peri nets. Petri nets are a mathematical formalism with an intuitive graphical representation. A Petri net consists of places, which are denoted as circles, and transitions, denoted as boxes. Places can have tokens, denoted as black dots. A transition t may fire if its input places -the places from which an arc leads to t- have sufficiently many tokens. If t fires, it consumes tokens from its input places and produces tokens on its output places, i.e. the places to which an arc from t leads.

For model checking, the formula often is rather short, so that only a few nodes of the Petri net are referenced by a formula ϕ. Since only little of the net is mentioned in ϕ, it should be possible to omit parts of the net without changing the satisfiability of ϕ. Our approach to define such Petri net reduction rules is inspired by program slicing, c.f. e.g. [1]. A program slice contains at best only the statements that affect the so called slicing criterion, a variable at a given program line. We examine a Petri net by starting from the atomic propositions of ϕ. We focus on structures in the net that divide the net into subnets with limited interaction. Given such a subnet is free of atomic propositions, in certain cases it is possible to find a smaller net to replace it. We implemented an algorithm based on reading transitions [2]. The algorithm reduces certain strongly connected nets and shows good experimental results especially for nets that are not strongly connected. The reduction preserves satisfiability of LTL without next-time and CTL without next-time. Currently we are investigating the generalisation of reading transitions to reading subnets that may temporarily remove tokens. Another research direction is to find restrictions on the net that allow to identify subnets that play the role of data and control flow, so that the concept of relevant variables, as defined in [1], is applicable.

Literatur

[1] Weiser, M.: Program slicing. In: Proceedings of the 5th international conference on Software engineering, IEEE Press Piscataway, NJ, USA (1981) 439–449

[2] Rakow, A.: Slicing Petri Nets. http://parsys.informatik.uni-oldenburg.de/pub/index.html Technical Report, 2007

Timing Behavior Anomaly Detection for Automatic Failure Detection and Diagnosis

Matthias Rohr

rohr@informatik.uni-oldenburg.de

Graduate School TrustSoft, University of Oldenburg, Germany

Business-critical enterprise-scale software systems have to satisfy high availability requirements. The availability of a system is determined by its reliability (probability of failure free operation over a period of time) and the time to repair after a failure has occurred. In practice, it is not feasible to develop complex software systems that do not contain faults. After development, availability may be improved by reducing the time to repair with automatic failure detection and diagnosis. In complex software systems, manual failure diagnosis and detection can be time-consuming or error-prone. Anomaly detection is part of many failure detection and diagnosis approaches. Anomaly detectors compare system behavior observations against some reference model of "normal" system behavior. Deviations from normal system behavior are used as indicators in further diagnosis steps that evaluate patterns of anomalies.

In software applications, several runtime behavior metrics, such as service response times, memory utilization, and parameter values are candidates for anomaly detection. Our approach to anomaly detection focuses on the evaluation of response times of internal operations in the software. Many faults cause timing behavior anomalies that may be detected earlier than actual (non-functional or functional) violations of correct system operation. However in particular for timing behavior, it is a challenge to decide whether variations are the "normal" result of changes in system usage and resource utilization. Existing approaches do not explicitly model system usage variations in the evaluation of timing behavior.

It is hypothesized that the quality of timing behavior anomaly detection is improved by explicitly modeling system usage. Two major influences to timing behavior are workload intensity and service demand characteristics. We explicitly include these two influences into a timing behavior evaluation model that is similar to classical performance prediction models. More precisely, we quantify workload intensity by the amount of active jobs sharing a subset of system resources. We reflect the service demand characteristics of single user requests with operation execution sequences (traces). Our approach to timing behavior anomaly detection targets distributed, multi-user, component-based software systems.

Currently, we employ empirical evaluation in case studies in the lab and in the field to confirm the applicability and to estimate the effectiveness of our approach. Workload drivers and fault injectors are used to generate scenarios to evaluate the timing behavior evaluation and failure diagnosis capabilities. Additionally, these experiments explore the general dependencies between software faults and timing behavior in enterprise information systems.

Adaptive Dynamic Replication Schemes

Christian Storm

christian.storm@informatik.uni-oldenburg.de
Graduate School TrustSoft, University of Oldenburg, Germany

Replication techniques are commonly used to improve the availability of critical data objects like files or databases. A copy of the data object is placed on several independently operating and interconnected nodes of a network in order to allow the data access operation to succeed despite of some nodes being crashed or unavailable. The used replication scheme defines the modus operandi of the data access operation, i.e. which and how many replicated data objects are involved in an operation, and hence bounds the number of crashed or unavailable nodes that can be tolerated until the data access operation finally fails. Many application scenarios require the data access operations to exhibit the same behavior for a replicated data object as for a non-replicated data object such that the execution of data access operations on a replicated data object must be equivalent to the execution of the same sequence of operations on a non-replicated data object. This property is known as One-Copy-Serializability (1SR) and most replication schemes use a so-called quorum system for enforcing 1SR due to its ability and flexibility of not just tolerating node crashes and node unavailability but also because of being able to cope with network partitions. Once a particular replication scheme is chosen it is used throughout the entire lifespan of the data object. It was shown that the data access operation availabilities of these static replication schemes is bounded. Allowing the quorum sets to be changed over time offers the potential of dynamically reacting to changing network characteristics. However, most dynamic replication schemes impose an upper bound on the number of replicated data objects they can manage. At design-time it has to be clear what the projected maximal number will be. Moreover, dynamic replication schemes exhibit a special notion of dynamics because they manage all replicated data objects up to the static upper bound in a scheme-specific, and in this sense, homogeneous way: the same strategy for deriving the quorum sets is used for every number of replicated data objects. Since there is no strategy that is superior for every number of replicated data objects to be managed, using an inhomogeneous replication scheme, i.e. one that allows the usage of multiple strategies for deriving the quorum sets, can be beneficial. Analogously, there is no replication scheme that is best-suited for every application scenario. Until now, a particularly well-suited replication scheme has to be identified and applied manually for each application scenario anew. If the application scenario requirements vary over time then each time a cost- and time-intensive manual intervention to adapt the replication scheme accordingly is required which renders the system unavailable during this maintenance phase.

This gives reason to the need for an adaptive dynamic replication scheme that adapts itself at run-time in an automatic manner to best-fit varying application scenario requirements and network characteristics without the need for administrative downtimes.

Symbolic Decision Procedures for Robustness Analysis of Timed Systems

Mani Swaminathan

mani.swaminathan@informatik.uni-oldenburg.de

GK Trustsoft, Universität Oldenburg

Abstract: Real-time systems, which have strict timing requirements, have emerged as an enabling technology for several important application domains such as air traffic control, telecommunications, and medicine, to name a few. Such systems are becoming increasingly pervasive, and hence rigorous methods and techniques to ensure their correct functioning are of utmost importance. Timed Automata (TA) have been extensively studied as a formalism for modelling real-time systems. TA extend ω-automata by augmenting them with "clock" variables based on a dense-time model, which quantitatively capture the behaviour of the system with time.

TA model-checkers such as UPPAAL are now available and have been successfully used in industrial case studies. The core of such TA model-checkers is a zone-based Forward reachability Analysis algorithm, which is primarily used to test safety properties in timed systems. However, this assumes that the clocks of the TA are perfectly synchronous.

This dissertation focusses on developing symbolic algorithms for timed systems modelled as timed automata and extensions thereof, under the notion of robustness w.r.t small imprecisions in the model parameters. We have developed an efficient algorithm using a "neighbourhood" operator on zones for deciding safety and deadlock freedom of timed automata . In contrast to other approaches in this direction, our algorithm is zone-based, fully forward, is guaranteed to terminate, and requires no special treatment of cycles in TA. The overhead imposed by our algorithm w.r.t the standard non-robust analysis used in tools such as UPPAAL is minimal. A first prototype of our algorithm has been implemented in the model-checker UPPAAL, and has been tested on simple examples for safety and deadlock freedom under robustness. Our notion of robustness is weaker than the classical definition, but is nonetheless realistic in the sense that it can be applied to all systems that have a bounded life-time.

We intend to extend this technique to more general models of timed systems, such as linear priced timed automata, stopwatch automata, among others, and apply them to large-sized case-studies.

SAT-Modulo-Theory based Analysis of Probabilistic Hybrid Systems

Tino Teige

`tino.teige@informatik.uni-oldenburg.de`

Graduate School TrustSoft, University of Oldenburg, Germany

The framework of hybrid discrete-continuous systems becomes increasingly popular for modeling and verifying large real-world, often safety-critical, embedded systems, e.g. in the car and aviation industry. In contrast to purely discrete models, like Kripke structures, hybrid systems additionally show continuous behavior usually modeled by differential equations. Therefore, the state space of a hybrid system is infinite and given by the Cartesian product of the system modes and the discrete and continuous system variable domains. The system can change its current mode by a potentially non-deterministic choice of a transition to another mode, while during residing in a mode the variables continuously evolve over time. This rich formalism allows a direct and expressive modelling of real-world systems. On the other hand, the complexity of model checking such a system against requirements grows with increasing expressiveness of the modeling framework. While the state reachability problem of very restricted subclasses (timed automata) is decidable, it becomes undecidable in general when adding some small extra functionality (e.g., stopwatch automata). Nevertheless, industrial needs motivate the development of efficient model-checking algorithms for richer classes of hybrid systems involving arbitrary differential equations and arithmetic.

Common model-checking approaches are able to disprove the safety of a system in the sense of deciding reachability of a failure mode under certain conditions. However, the question rarely is whether a failure mode is reachable, but whether the failure occurs with non-negligible probability. Failures in a system may happen in extreme situations caused by the environment, e.g. power blackout combined with a defect of the emergency power generator. The probability of such uncertainties could be estimated and integrated within the model description. These requirements lead to the extended notion of probabilistic hybrid systems. Here, the non-deterministic selection of a transition is enriched by a probabilistic choice according to a distribution over variants of the transition. More formally, each transition carries a (discrete) probabilistic distribution. Each probabilistic choice within such a distribution leads to a potentially different successor mode while performing some discrete actions.

We are especially interested in k-bounded model checking problems, i.e. we want to prove or disprove whether a given property P is satisfied with probability greater or equal p in a probabilistic hybrid system $\mathcal{H}$ along all its traces of length up to k. We will use a symbolic model checking approach. Therefore, we represent the non-deterministic, stochastic, and non-linear behavior of the system $\mathcal{H}$ (up to the specified length k) by a formula φ of size

linear in $\mathcal{H}$, P, and k. More precisely, the formula φ is a Boolean combination of non-linear arithmetic expressions involving stochastic variables. The construction of φ ensures that φ is satisfiable with probability $\geq p$ if and only if the system $\mathcal{H}$ (restricted to traces of length k) fulfills property P with probability $\geq p$. Hence, we can reduce the k-bounded probabilistic satisfaction problem of probabilistic hybrid systems to the stochastic satisfiability problem of mixed arithmetic and Boolean formulas entailing stochastic quantifiers, generalizing the concept of stochastic quantification in stochastic Boolean SAT solving (SSAT, [Maj04]).

For solving such a formula φ, we will develop an algorithm which tightly integrates, on the one hand, approaches from stochastic Boolean SAT solving and stochastic constraint programming (SCP, [TMW06]) and, on the other hand, a generalization of the Davis-Putnam-Loveland-Logemann (DPLL) algorithm to non-linear arithmetic over the reals (iSAT, [FHR$^+$07]). SSAT and SCP algorithms provide powerful decision procedures for stochastic Boolean and finite-domain problems, resp., yet fail to address infinite domains and continuous arithmetic. The iSAT algorithm efficiently solves non-linear arithmetic formulas over the reals with complex Boolean structure. iSAT itself is a tight integration of interval constraint propagation (ICP) and the DPLL algorithm, exploiting techniques like conflict-driven learning and non-chronological backtracking. To the best of our knowledge, there is just one other tool, called ABSOLVER (cf. [BPT07]), which solves (large) mixed non-linear arithmetic and Boolean formulas. We showed that the classical lazy theorem proving approach of ABSOLVER is outperformed by orders of magnitude by our tool iSAT (cf. [FHR$^+$07]).

For developing an efficient solver for the stochastic non-linear arithmetic and Boolean satisfiability problem, we aim at an extension of the iSAT algorithm supporting stochastic variables within the solver core. We prefer this approach over a loose integration of an SSAT algorithm with iSAT as a subordinate solver for efficiency reasons. This work will enhance the state of the art concerning stochastic satisfiability by solving problems over infinite domains and undecidable theories, which is of prime importance to the analysis of probabilistic hybrid systems. To evaluate our novel algorithm, we will apply it in the context of bounded model-checking of probabilistic hybrid systems.

References

[BPT07]   A. Bauer, M. Pister, and M. Tautschnig. Tool-support for the analysis of hybrid systems and models. In *Proceedings of the 2007 Conference on Design, Automation and Test in Europe (DATE'07)*, Los Alamitos, CA, April 2007. IEEE Computer Society.

[FHR$^+$07]   M. Fränzle, C. Herde, S. Ratschan, T. Schubert, and T. Teige. Efficient Solving of Large Non-linear Arithmetic Constraint Systems with Complex Boolean Structure. *JSAT Special Issue on Constraint Programming and SAT*, 2007. Accepted for publication.

[Maj04]   S.M. Majercik. Nonchronological Backtracking in Stochastic Boolean Satisfiability. *ICTAI*, 00:498–507, 2004.

[TMW06]   S.A. Tarim, S. Manandhar, and T. Walsh. Stochastic Constraint Programming: A Scenario-Based Approach. *Constraints*, 11(1):53–80, 2006.

Structural Failure Models for Fault-tolerant Distributed Computing in Asynchronous Environments

Timo Warns

`timo.warns@informatik.uni-oldenburg.de`

Graduate School TrustSoft, University of Oldenburg, Germany

The development and evaluation of fault-tolerant distributed systems require the specification of assumptions on expected component failures. Such assumptions are given by a failure model that characterises all failures that may occur. We distinguish between functional and structural failure models in the context of process failure models. A *functional failure model* specifies how individual processing entities may fail, that is, the *failure mode* of components. A *structural failure model* specifies the potential scope of failures within a system, for example, how many of these entities may fail simultaneously.

In fault-tolerant computing, *threshold models* are often used as structural failure models. Threshold models characterise the sets of failed components by *t-out-of-n assumptions*, that is, at most t of n components may fail. The major advantage of these models is their simplicity that eases proofs of correctness. However, it is well-known today that threshold models have significant limitations with respect to their expressiveness such that they are inapplicable in many real-world situations. In particular, they are based on the implicit assumption that failures are stochastically independent.

The contribution of the proposed thesis is to formally define the notion of structural failure models in the context of asynchronous distributed systems and to develop new structural failure model classes that are more expressive than previous classes. In particular, the new classes address important aspects of real-world situations (e.g., stochastic dependencies or failure propagation) to improve the assumption coverage of solutions.

The thesis evaluates the new classes in terms of their expressiveness to show the gained assumption coverage and, exemplarily, evaluates the resilience and the quality of service (QoS) of solutions under the new classes. In particular, the new classes allow to generalise results obtained with previous classes. For the evaluation, the thesis focuses on three well-known problems to be solved under different models: failure detectors, consensus, and quorum-based replication. If a class is more expressive than another, solutions can be found that are more resilient and efficient than solutions based on less expressive classes. However, on the downside, a more expressive class complicates formal reasoning such that the class is impractical. The newly developed solutions will transform known solutions (if possible) to illustrate that the respective class is simple enough for uncomplicated formal reasoning.

Specification and Analysis of Dynamic Topology Systems

Bernd Westphal,
westphal@informatik.uni-oldenburg.de,
Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

Formal methods for embedded systems currently mainly focus on single components or fixed configurations of finitely many components. Examples for the former case are open finite-state systems as models of single discrete controllers in a discrete environment. Examples for the latter comprise models of clusters of controllers with fixed inter-connection, that is, no controllers are added or removed at runtime and the inter-connection is not re-configured, e.g., the numerous controllers in a car for airbags, braking assistance, etc.

These concepts are not sufficient when many autonomous systems interact. A characteristic example is the *car platooning* application as studied by the California PATH project. The intention of car platooning is to reduce fuel consumption by dynamically merging cars into platoons where they drive with significantly reduced safety distance and hence benefit from slipstream. In order to remain safe in case of braking manoeuvres, the first car in a platoon plays the role of a leader which notifies the other cars, the followers, about upcoming braking manoeuvres thereby minimising reaction time. To faithfully model car platooning, there have to be means to describe (i) unbounded appearance and disappearance of cars within the system "highway", (ii) topologies, that is, selective connections between cars like between leader and follower, and (iii) (asynchronous) communication.

We propose to extend the particular finitary abstraction *Data Type Reduction* (McMillan, 2001) known for parameterised systems to class of *dynamic topology systems* as characterised by (i)–(iii). To this end, we employ labelled transition systems where states are labelled with graphs. This allows us to model nasty, critical effects like dangling links, that is, connections to already disappeared processes. Furthermore, we introduce a first-order extension of classical temporal logic. It is process-oriented in the sense that quantified variables range over processes and follow their evolution over time. That is, we can express properties requiring that, for instance, *the particular* car, which initiated a merge, will finally complete the merge. The semantics of this logic also has to treat pre-mature disappearance of processes. By re-stating the finitary abstraction in terms of the graph-labelled transition system allows us to gain insight into the potentials and limitations of this technique; beforehand, it has only been described in terms of a construction procedure. Individual-oriented properties, which are easily lost in many other abstractions, are essentially preserved by following what we call the *spotlight principle* (Wachter & Westphal, 2007). Finally, we demonstrate the applicability of this approach by sketching a translation from a relevant fragment of UML and of the DCS language (Bauer, Schaefer, Toben & Westphal, 2006) into graph-labelled transition system, the latter allowed us to establish safety and liveness properties for the car platooning case-study.

GRK 1194:

Selbstorganisierende Sensor-Aktor-Netzwerke

Karlsruhe

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke
RTG 1194 Self-organizing Sensor-Actuator-Networks

Fakultät für Informatik
Fakultät für Elektrotechnik und Informationstechnik
Universität Karlsruhe (TH)

Forschungsprogramm (Deutsch)

Verteilte Netzwerke, welche aus einer großen Anzahl von miniaturisierten und autonomen Sensor-Aktor-Systemen bestehen, bieten völlig neue Möglichkeiten, die Umwelt zu beobachten. Statt des passiven Blicks erlauben sie eine Durchdringung verschiedenster Phänomene mit einer durch die Knotendichte wählbaren Auflösung. Die integrierte Aktorik ermöglicht zudem die gezielte Beeinflussung der Sicht und eine Interaktion mit der Umgebung. Zudem kann durch die große Anzahl von Systemen auch bei einer geringen Zuverlässigkeit und Verfügbarkeit der individuellen Sensor-Aktor-Systeme die notwendige Toleranz gegenüber Fehlern und Ausfällen erreicht werden. Da die Knoten aber typischerweise energieautonom arbeiten müssen, ergeben sich enge Limitierungen für die Messgenauigkeit, die Rechenleistung und die Kommunikationskapazität. Außerdem kommt es durch die bewusste Selbstabschaltung individueller Knoten zum Zweck der Energieeinsparung, möglichen umgebungsbedingten Ausfällen und der inhärenten Mobilität zu regelmäßigen Topologieänderungen. Damit ergeben sich für derartige Sensor-Aktor-Netzwerke neuartige Anforderungen bzgl. der Kommunikationssysteme, die auf deren spezielle Charakteristika zugeschnitten sind und der Informationsverarbeitung, da den einzelnen Sensor-Aktor-Knoten jeweils nur Teilinformationen des zu untersuchenden Phänomens zur Verfügung stehen, so dass zusammen mit den beschränkten Kommunikationsmöglichkeiten eine dezentrale Informationsverarbeitung innerhalb des Netzwerks erforderlich ist.

Auch wenn Sensor-Aktor-Netzwerke in den letzten Jahren viel Aufmerksamkeit in der Forschung erzielt haben, sind die verfügbaren Theorien und Methoden nur für eine geringe Anzahl von Knoten praktikabel und lassen sich bisher nicht auf eine große Anzahl gleichberechtigter Knoten skalieren. Die entstehende Komplexität kann nur durch die Selbstorganisation dieser Netzwerke auf allen Ebenen beherrscht werden. Außerdem müssen zur Erhöhung der Lebensdauer des Netzwerks eine verlustleistungsminimierte Architektur und ein energieeffizienter Betrieb aller Knoten und deren Komponenten gewährleistet werden. Die Entwicklung von Sensor-Aktor-Netzwerken macht demnach neue bereichsübergreifende Designprinzipien in den drei Hauptgebieten Hardware/Software-Systemintegration, Kommunikation und Informationsverarbeitung notwendig. Hier setzen die Forschungsaktivitäten des Graduiertenkollegs an, bei denen eine Bündelung unterschiedlicher

Expertisen in diesen drei Bereichen vorgenommen wird. Im Vordergrund stehen dabei vor allem eine integrierte Betrachtungsweise und die gemeinsame Entwicklung systematischer Vorgehens- und Referenzmodelle für den Entwurf und die Konstruktion von selbstorganisierenden Sensor-Aktor-Netzwerken. Besonderer Wert wird auf eine enge Verzahnung zwischen den drei Hauptgebieten gelegt, um Schnittstellen und Wechselwirkungen aufzeigen zu können. Aus den komplexen Randbedingungen solcher Netzwerke (Unzuverlässigkeit, verteilte Verarbeitung, begrenztes Energiebudget) und weiteren Anforderungen (Selbstorganisation, Privatheit, Rekonfigurierbarkeit) ergeben sich wichtige Fragestellungen, die im Rahmen des Graduiertenkollegs bearbeitet werden.

Research Program (Englisch)

Sensor-actuator networks consisting of a large number of miniaturized and autonomous nodes offer novel possibilities to collaboratively observe and control distributed phenomena. Such networks can be used in industrial, medical, urban, and many other applications. In practical implementations, the individual sensor-actuator nodes are normally densely deployed either inside the phenomenon or very close to it. Thus, the network provides a good spatial resolution, which can be adapted depending on the dynamics of the distributed phenomena to be observed. Besides the observation task, the individual nodes are able to interact with the phenomena itself by means of integrated actuators. Thanks to the large number of nodes, the fault tolerance and robustness of the overall system can be significantly increased, even for low reliability and availability of the individual nodes. However, the energy constraints resulting from the required autonomy nodes impose severe performance limits with respect to communication bandwidth and processing power and, hence, provide a challenge for architectural design and operation of this type of networks. Due to the intrinsic mobility, possible failure, and selective switch-off for energy reasons, the topology of the network is regularly changing. This leads to novel requirements regarding communication and information processing. Furthermore, it is often desirable to reduce communication activities between the individual nodes to a minimum and to reduce the heavy computational burden in signal processing. To tackle these challenges, a decentralized estimation approach is desirable, which implies that just parts of the state vector are manipulated at each processing step.

Within the framework of this Research Training Group, the fundamental problems in the area of sensor-actuator networks will be addressed. This includes the whole spectrum from communication and signal processing to hardware/software system integration. The research activities aim to provide a unifying view to the closely coupled problem areas that have so far only been addressed in isolation. These research activities include the joint examination of query processing, content-based addressing, energy efficiency, and self-organizing middleware, which will be achieved by a close cooperation of experts from different disciplines. An integrated treatment and the development of a systematic approach to sensor-actuator network design represent our key interests.

Anwendungsspezifische Architekturen für Energieautarke Sensor-Aktor-Knoten

Ashok-Kumar Chandra-Sekaran

`ashok@itiv.uni-karlsruhe.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke
Fakultät für Elektrotechnik und Informationstechnik,
Institut für Technik der Informationsverarbeitung, Universität Karlsruhe (TH)

Aim of the Subproject: The usage of sensor actuator networks for concrete applications arises many challenges: longer battery lifetime of sensor-actuator nodes, to economically produce complex sensor nodes in large number, to form a heterogeneous network specific to architecture. The sensor-actuator nodes have to be made energy self-sufficient using energy efficient techniques at the globalised level. Methods of power production and storage like energy scavenging should be used. The optimal resource distribution in sensor actuator networks has to be regarded too.

Current work: The following features of the wireless sensor network, which arise from the objectives of the project, were considered for the selection of an application: Heterogeneous network, large number of nodes, reduction of manual management, dynamic /moving nodes, energy consumption restrictions, Multi-hop communication, no alternative technology (infrastructure), ad-hoc networks, localization. Based on the above features the Industry automation and Disaster management applications were potentially interesting.

The Disaster Management scenario is interesting because the disaster site is infrastructure less and requires a dynamic WSN. As part of the field study a disaster simulation drill is attended to understand the disaster response process followed in Germany called „Massenanfall von verletzten (MANV)". A wireless disaster aid network (DAN) to provide emergency mass casualty response is formulated.

The DAN comprises of the patient bracelet, doctor's PDA, doctor's bracelet and the Monitor station for Org.L. (Organisationsleiter), LNA (Leitender Notarzt), TEL (Technischer Einsatzleiter). The patient bracelet contains: electronic triage, sensors (vital sign, physical activity), Actuators (Emergency indication), RFID tag (patient identification).

The main functionalities of this disaster aid network are: Patient monitoring; context aware data access-patient information recorded by emergency doctor's PDA; efficient resource management which includes: tracking patient/emergency doctors, triage information of patient, on request- patient vital signs and activity, estimate resource further required and dispose available resource efficiently. Since the nodes are battery powered the disaster aid network should be energy efficient and hence energy efficient techniques and the low power Zigbee technology are to be used. A demonstrator implementing all the functiona-

lities of the DAN has to be developed.

The first version of this demonstrator that consists of a ZigBee mesh network of around 50 nodes is currently being developed. This mesh network consists of the acceleration sensor node(Fig. 1), routers and the ZigBee dongle enabled Laptop as coordinator. The acceleration sensor node hardware is based on the TI CC2431 Zigbee SOC. The functionalities that this network can realise are: to evaluate the CC2431 SOC hardware based localisation technique; to estimate resource further required and dispose available resource efficiently using the 'care zone algorithm' based simulator ; to indicate the activity of the patient (ex: standing, running).
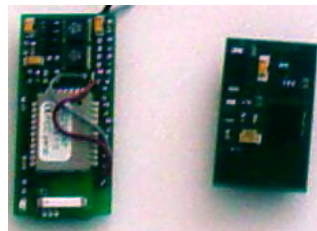

Abbildung 1: Zigbee sensor node

Dezentrale Aufgabenbearbeitung mittels Kooperation und Interaktion

Heiko Hamann

hamann@ira.uka.de

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Prozessrechentechnik, Automation und Robotik,
Universität Karlsruhe (TH)

Um die Implementierung großer Netzwerke zu ermöglichen, müssen die einzelnen Knoten einen gewissen Grad an Autonomie besitzen. Dies impliziert, dass es keine zentrale Steuereinheit gibt und die gestellte Aufgabe durch Kooperation der Knoten untereinander bearbeitet wird. Hier sollen vor allem die Methoden der Selbstorganisation und Emergenz zur Anwendung gebracht werden, die ein Maximum an Dezentralität, Robustheit und Skalierbarkeit garantieren. Es werden Knoten betrachtet, die ihre Bewegungsrichtung selbst beeinflussen können mehrheitlich sog. Schwarmroboter-Systeme, d.h. große Gruppen von Robotern mit kleiner Baugröße und stark eingeschränkten Fähigkeiten in Bezug auf die Reichweiten der Sensorik und die Leistungsstärke des datenverarbeitenden Systems, die kollektiv eine gemeinsame Aufgabe lösen. Der effiziente Entwurf von Steueralgorithmen für selbstorganisierende Roboterschwärme ist eine klare wissenschaftliche Herausforderung. Mit bestehenden Methoden ist der schnelle und erfolgreiche Entwurf derartiger Steueralgorithmen nicht durchführbar. Dabei ist die mangelhafte Fähigkeit des Menschen, das Verhalten von komplexen Systemen mit einer Vielzahl an Interaktionen intuitiv zu verstehen und korrekt vorhersagen zu können, das Hauptproblem. Die Lösung könnten Modelle sein, die dem Enwickler schnell eine korrekte Vorhersage des Schwarmverhaltens geben, das sich aus dem momentanen betrachteten Steueralgorithmus ergibt.

Im Rahmen dieses GK wurde ein Modell zur Vorhersage des Verhaltens eines Roboterschwarms entwickelt. Dabei wurde im Gegensatz zu vorhanden Ansätzen besonders Wert auf eine möglichst genaue Modellierung des Raumes gelegt. Ein weiterer Vorteil ist die physikalisch motivierte Modellierung der Bewegung, die die Verteilung der Roboter im Raum beschreibt. Der sog. Micro-Macro-Link wird dabei direkt durch traditionelle Methoden der Physik abgeleitet. Das Modell ist ein System partieller Differentialgleichungen, welche die Wahrscheinlichkeitsdichte der Roboterverteilung im Raum und einen relevanten Umweltaspekt modellieren. Erste Ergebnisse wurden gegen realitätsnahe Simulationen validiert und Möglichkeiten, wie das Modell den Entwurf von Steueralgorithmen unterstützen kann, wurden aufgezeigt.

Literatur: [HW07] H. Hamann und H.Wörn. A Space- and Time-Continuous Model of Self-Organizing Robot Swarms for Design Support. In First IEEE International Conference on Self-Adaptive and Self-Organizing Systems, Boston, USA, July 2007.

Energieeffiziente Lastverteilung zwischen Kommunikation und lokalen Rechenressourcen

Dominic Hillenbrand

`hillenbrand@informatik.uni-karlsruhe.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Technische Informatik, Universität Karlsruhe (TH)

Sensor-Aktor-Knoten sind in ihre Umwelt eingebettete Systeme, die kooperativ verteilte Phänomene erfassen und beeinflussen. Durch eine entsprechende Anzahl Sensor-Aktor-Knoten kann eine hohe Ortsauflösung und Fehlertoleranz erreicht werden. Da Sensor-Aktor-Knoten ihre Energie aus (wiederaufladbaren) Batterien und/oder aus ihrer Umwelt (z.B. Sonnenenergie) beziehen, ist der Energieverbrauch ein entscheidender Faktor für den Nutzen des Netzwerkes.

Ziel des Forschungsprojekts ist es, typische Abläufe in Sensor-Aktor-Netzwerken energieeffizient zu gestalten. Dazu gehören unter anderem: Datenakquise, Signalverarbeitung, Kommunikation und Selbstorganisation. Insbesondere Kommunikation und Informationsverarbeitung werden gemeinsam betrachtet, um durch geschickte Abwägungen die Energie-Effizienz des Netzwerkes zu optimieren.

Ein Einsparpotential ist in der drahtlosen Kommunikation zu finden. Dort gilt es, Abwägungen zwischen der Berechnung von unterschiedlich aufwendigen Fehler- und Kompressionscodes, sowie der zu übertragenen Datenmenge zu finden. In diesem Zusammenhang ist auch die Umsetzung der gewählten Strategien in Hard- und Softwaremodule mitentscheidend, da Hardware und Software sich in Hinsicht auf Leistung, Energieverbrauch und Flexibilität unterscheiden.

Die Abbildung 1 zeigt den am Lehrstuhl CES entwickelten Hyperion-Sensorknoten. Der Sensorknoten verfügt über vielfältige Möglichkeiten zur Energiemessung und wird dank seiner flexiblen Hardware (u.a. FPGA basiert) als Evaluierungsplattform für unterschiedliche HW/SW-Architekturen verwendet.
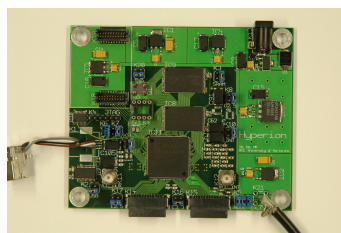

Abbildung 1: Hyperion-Sensorknoten (FPGA basiert)

Sensoreinsatzplanung und Routing in Sensor-Aktor-Netzwerken

Marco Huber

`mhuber@ira.uka.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke
Fakultät für Informatik, Institut für Technische Informatik, Universität Karlsruhe (TH)

Die Bereitstellung der Informationen, auf deren Verarbeitung sich das „In-Network-Processing" stützt, stellt die zentrale Aufgabe der verteilten Sensorik des Sensor-Aktor-Netzwerks dar. Hierbei erforderliche Messaufgaben bedürfen in der Regel eines Großteils der vorhandenen, limitierten Energieressourcen. Um eine lange Lebensdauer des Netzwerks zu gewährleisten, ist neben der Abwägung zwischen Kommunikation und Informationsverarbeitung, auch der gezielte Einsatz der Sensorik vonnöten. Teilprojekt I4 behandelt aus diesem Grund die Einsatzplanung der Sensoren zur Reduktion des Energiebedarfs. Mit einer optimalen Datenakquisition reduziert sich desweiteren die Menge der im Netzwerk zu verarbeitenden Daten, welche in Teilprojekt I1 zur Rekonstruktion benötigt werden. Hierbei ist der bloße Informationsgewinn durch die Messung mit den anfallenden Kommunikationskosten in Relation zu setzen, weshalb eine gemeinsame Betrachtung von Einsatzplanung und Routing naheliegend erscheint.

Stochastische Modellierung: In diesem Teilprojekt sollen Verfahren zur gezielten Einsatzplanung der Sensorik und Aktorik eines Sensor-Aktor-Netzwerks erforscht und entwickelt werden. Als Grundlage dienen hierfür die Techniken des „Stochastic Optimal Control", also der stochastischen, optimalen Steuerungs- und Regelungstheorie. Im Gegensatz zur klassischen Regelungstechnik finden hierbei Störeinflüsse, welche nur teilweise bekannt sind, als stochastische Größen ihre Berücksichtigung. Hierdurch wird eine stochastische Modellierung des zu beobachtenden Phänomens und des Sensor-Aktor-Netzwerks erforderlich. Bei dieser Modellierung ist auf eine möglichst hohe Allgemeingültigkeit zu achten. Gegenwärtig ist das Problem der Sensoreinsatzplanung erst für wenige, eher unrealistische Spezialfälle unter der Annahme fehlerfreier Kommunikation hinreichend untersucht. Die notwendige Verallgemeinerung bedingt unter anderem das Aufstellen von nichtlinearen Sensor-Aktor-Modellen, was zudem die Entwicklung und den Einsatz von nichtlinearen Schätzverfahren erfordert.

Integration der Aktorik: Vielfach verfügen die Netzwerkknoten auch über eine integrierte Aktorik, wodurch eine aktive Fortbewegung oder die Beeinflussung des verteilten Phänomens möglich ist. Neben dem Fällen von Planungsentscheidungen zum Einsatz der Sensorik, lassen sich die entwickelten stochastischen Regelungsmethoden auch zur Ansteuerung der Aktorik einsetzen.

Kombination von Routing und Sensoreinsatzplanung: Zur Bewertung des möglichen Informationsgewinns ist die Formulierung von Kriterien unter den Randbedingungen drahtlo-

ser Kommunikation und begrenzter Energieressourcen vonnöten. Insbesondere die draht-lose Kommunikation erfordert die Einbeziehung der Routingentscheidung in die Sensorauswahlplanung, da maßgebliche Kosten auf die Bereitstellung einer zuverlässigen Kommunikation entfallen. Langfristiges Ziel ist daher die integrierte, dezentrale Kommunikationswege- und Sensoreinsatzplanung.

Literatur: [HH07] M. F. Huber und U. D. Hanebeck. The Hybrid Density Filter for Nonlinear Estimation based on Hybrid Conditional Density Approximation. The 10th International Conference on Information Fusion (Fusion), Québec, Kanada, Juli 2007.

Algorithmische Aspekte der Clusterbildung, Topologiekontrolle und Lokalisierung

Bastian Katz

`bastian.katz@kit.edu`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Theoretische Informatik, Universität Karlsruhe (TH)

Ziel des Teilprojektes K2 ist die Untersuchung algorithmischer Aspekte von Sensor-Aktor-Netzwerken, das heißt eine Betrachtung der in Sensor-Aktor-Netzwerken anfallenden Probleme aus algorithmischen Sicht. Dies beinhaltet wesentlich die Abstraktion und Modellierung von Problemen, sowie Komplexitätsbetrachtungen, aber auch den Entwurf und die Analyse von Algorithmen. Beispielhafte Probleme sind Clusterbildung, Topologiekontrolle und Lokalisierung:

Clusterbildung: Sensor-Aktor-Netzwerke können auf keine vorgegebene Infrastruktur aufbauen, können aber durch die Auswahl von Substrukturen selbst Infrastrukturen ausprägen, die für das Verteilen, Einsammeln oder auch das gezielte Versenden von Nachrichten ausreichen. An diesen Strukturen sollen möglichst nicht alle Knoten beteiligt sein, um einerseits Energie einzusparen und andererseits gegenseitige Störungen zu minimieren. Die Bedingungen und Optimierungskriterien für solche Substrukturen lassen sich oft sehr direkt auf prominente graphentheoretische Probleme zurückführen, zum Beispiel *minimum (connected) dominating sets* oder *maximum independent sets*. Besondere Herausforderungen stellen sich hier unter anderem durch die notwendig verteilte Berechnung.

Topologiekontrolle Sobald in Sensor-Aktor-Netzwerken Knoten ihre Sendeleistung beeinflussen können, stellt sich die Frage nach einer optimalen Wahl der Sendestärke. Möglichst geringe Sendeleistungen reduzieren dabei nicht nur den Energieverbrauch und tragen somit zur Lebensdauer des Netzwerks bei – sie verringern darüber hinaus auch Störungen bei gleichzeitigem Senden mehrerer (naher) Knoten. Auf der anderen Seite sinkt mit der Sendestärke eines Knotens auch die erreichbare Nachbarschaft an Knoten, deshalb müssen Sendestärken groß genug gewählt werden, um wichtige Eigenschaften des Netzwerkes (wie eine vertretbare Verlängerung von kürzesten Wegen) zuzusichern.

Lokalisierung Für die Verarbeitung von in Sensor-Aktor-Netzwerken anfallenden Daten ist die Kenntnis von Knotenpositionen notwendig; für viele weitere Aufgaben, z.B. das Routing ist sie zumindest sehr hilfreich. Hardware oder Infrastruktur, die es einzelnen Knoten erlaubt, ihre Positionen zu bestimmen, ist aufwendig, teuer und darüber hinaus nicht überall einsetzbar. Deshalb gibt es ein großes Interesse an Verfahren, die möglichst verteilt eine Rekonstruktion der Knotenpositionen aus Messungen ermöglichen, die die Knoten selbst lokal vornehmen können, also zum Beispiel Schätzungen der Abstände kommunizierender Knoten.

In diesen und weiteren Bereichen lassen sich zwei Arten von Fragestellungen unterscheiden: Zum einen gilt es, Problemstellungen zu modellieren und in den vorhandenen Modellen zu analysieren. Das betrifft insbesondere die Lokalisierung, deren Ausprägungen längst nicht alle untersucht wurden. Hier stehen die Frage nach der Komplexität solcher Probleme und nach (verteilten) algorithmischen Lösungen im Vordergrund. Zum anderen sind relevante graphentheoretische Probleme zum Beispiel der Clusterbildung bereits ausgiebig in vorhandenen Modellen des *distributed computing* untersucht worden. Hier gilt das wesentliche Interesse der Beschreibung von Modellen, die die Charakteristika von Sensor-Aktor-Netzwerken besser abbilden, zusammen mit einem Transfer der bekannten Probleme in diese Modelle.

ServiceCast — eine Architektur zur dienstorientierten Kommunikation in selbstorganisierenden Sensor-Aktor-Netzen

Andreas Kuntz

`akuntz@tm.uka.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Telematik, Universität Karlsruhe (TH)

Während die Architekur selbstorganisierender Sensor-Aktor-Netze (SSAN) derzeit noch stark auf die jeweilige Zielapplikation hin optimiert ist, bietet das Konzept der Dienste einen Abstraktionsmechanismus um die Flexibilität, Robustheit und Wiederverwendbarkeit einmal geschaffener Lösungen zu erhöhen. ServiceCast behandelt wichtige Aspekte, um Dienstorientierung in SSAN umzusetzen. Ein zentrales Merkmal ist *Dienste* statt Knoten als Grundlage des Routings zu betrachten.

Aufgrund der Annahme, dass Dienste durch mehrere Instanzen redundant vertreten sind, kann ein Dienstbezeichner einen Kommunikationspartner nur unzureichend charakterisieren. Daher fährt ServiceCast Kontexte ein, welche Sensor-Aktor-Knoten ihrer Umgebung zuordnen. Kontexte erlauben, das „wo" in einer vom Benutzer beeinflussbaren Granularität zu definieren, während die Dienstklasse das „was" definiert. Damit abstrahieren Kontexte vom Begriff der Lokation eines Sensor-Aktor-Knotens.

ServiceCast adressiert Kommunikationspartner durch die Kombination von Ziel-Kontextbereich und Ziel-Dienstklasse. So werden durch dieselbe Adresse potentiell alle gleichwertigen Dienstinstanzen innerhalb des angegebenen Kontextbereiches angesprochen. Die Instanzauswahl wird transparent vom Netz vorgenommen, was die Robustheit der Kommunikation verbessert.

Routing in den Zielkontext   Aktive Instanz auswählen (eine / mehrere / alle)
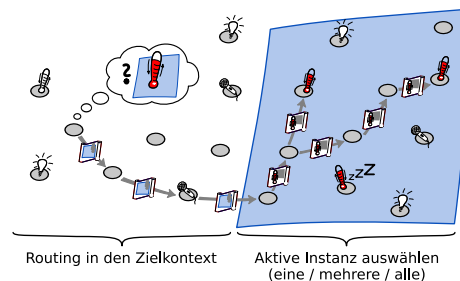
Nachrichten werden in zwei Schritten zur Ziel-Dienstinstanz geleitet: Zunächst wird die Nachricht in den Zielkontext geroutet, wobei allein die Kontextinformation relevant ist (Abb. links). Innerhalb des Zielkontextes wird die Ziel-Dienstklasse ausgewertet und die Nachricht zu den aktiven Instanzen des Zieldienstes geroutet (Abb. rechts). So kann z.B. die temporäre Nicht-Verfügbarkeit eines Dienstes, verursacht durch einen schlafenden Knoten, kompensiert werden. Ein Parameter in der Nachricht gibt an, wieviele der Instanzen die Nachricht erhalten sollen. Derzeit sind „genau eine", „mindestens eine" und „alle" als mögliche Werte vorgesehen.

Neben Mobilität sind Reprogrammierung, Kommunikationsstörungen und temporäre Un-

erreichbarkeit durch Energiesparmaßnahmen wichtige Ursachen für Dynamik in SSAN. Allen Ursachen gemeinsam ist eine Veränderung der Topologie des Netzes, welche u. A. Auswirkungen auf das Routing und das Auffinden von Diensten haben kann. Durch die mit ServiceCast eingeführte abstrakte Adressierung durch Dienstklassen und Kontexte, lassen sich Auswirkungen von Dynamik transparent kompensieren. Da ServiceCast Dienste als Grundlage des Routings verwendet, wird kein explizites Verzeichnis benötigt, welches Dienstbezeichner auf die IDs der diensterbringenden Knoten abbildet, wie dies heute in vielen dienstorientierten Architekturen üblich ist.

Weiter sind für ServiceCast zur Zeit Methoden zur Verbindungsverwaltung, verteilten Konfiguration und Instanziierung von Diensten, sowie ein dienstbasiertes Sicherheitskonzept in Arbeit.

Selbstorganisation im Bereich verlustleistungsminimierter Hardware

Oliver Sander

`sander@itiv.uni-karlsruhe.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Elektrotechnik und Informationstechnik,

Institut für Technik der Informationsverarbeitung, Universität Karlsruhe (TH)

Die herausragendste Eigenschaft von Sensor-Aktor-Netzwerken ist die sehr eingeschränkte Verfügbarkeit von Energie in den einzelnen Netzwerkknoten. Das Ziel sind Laufzeiten, die ohne weiteres die Dauer eines Jahres überschreiten können, betrieben durch Batterien in Knopfzellengröße. Neben der Möglichkeit der Energierückgewinnung durch beispielsweise Solarzellen liegt bei der Hardware der Fokus dadurch insbesondere in dem äußerst sparsamen Umgang mit der verfügbaren Energie. Zum Einsatz kommen in der Regel Ultra-Low-Power Mikrocontroller deren Rechenleistung äußerst begrenzt ist.

Auf der anderen Seite gibt es Dienste, die nahezu in jedem Sensor-Aktor-Netzwerk verwendet werden. Typische Vertreter sind das Routing (Multi-Hop-Kommunikation), Lokalisierung und Verschlüsselung/Komprimierung der Daten. Darüber hinaus rücken die als Self-X bezeichneten Eigenschaften immer mehr in den Fokus der Betrachtungen. Dem liegt zugrunde, daß eine Wartung einzelner Knoten innerhalb eines Netzwerks aufgrund der hohen Knotenanzahl nicht wünschenswert ist. Vielmehr soll ein Netz sich auf neue Anforderungen, Aufgaben und Umgebunsbedingungen autark und selbstständig anpassen. Dies sind insbesondere Selbstadministration, Selbstkonfiguration, Selbstoptimierung und Selbstheilung. Hinzu kommt der Wunsch die im Netz verfügbare verteilte Rechenleistung optimal zu nutzen.

Die beiden vorherigen Abschnitte gegenübergestellt, sieht man deutlich die Problematik, die der Realisierung eines Sensor-Aktor-Netzwerks entgegensteht. Die Erfüllung der Anforderungen allein für die Basisdienste ist mit den momentan eingesetzten Architekturen nicht möglich. Hier bietet der Einsatz rekonfigurierbarer Hardware Vorteile. Sie ermöglicht eine hohe Performanz bei niedrigen Taktraten durch die Parallelisierung von Aufgaben. Damit ist eine deutlich effizientere und auch getrennte Verarbeitung einzelner Tasks möglich. Denkbar ist beispielsweise eine Trennung von Routing und Verarbeitung der Sensorwerte. Die noch relativ hohe Verlustleistung von heute verfügbaren FPGAs wird in zukünftigen Generationen reduziert werden. Dennoch ist es notwendig weitere Methoden die zum effizienten Einsatz der Architekturen beitragen anzuwenden. Neben Clock Gating und Clock Scaling trägt die partiell dynamische Rekonfigurierbarkeit insbesondere zur Reduzierung von Chipfläche und damit Verlustleistung bei. Die Austauschbarkeit einzelner Module (HW Beschleuniger) erhöht deutlich die Adaptivität der Sensor-Aktor-Knoten. Auf eine neue Anforderung kann der Knoten sich durch Konfigurieren eines ge-

eigneten Moduls deutlich besser anpassen. Ist die Granularität geeignet gewählt können mehrere dieser Module in neuer Form für die Lösung einer Aufgabe zusammmengefügt werden, was ebenfalls deutlich zur Adaptivität der Knoten beiträgt. Die Fragestellungen dieses Teilprojekts befassen sich mit der Untersuchung der Einsetzbarkeit von rekonfigurierbarer Hardware als Sensor-Aktor-Knoten einschließlich Anwendung neuartiger Methoden zur Reduzierung der Verlustleistung. Darüber hinaus ist es von Interesse inwieweit die für Self-X notwendigen Eigenschaften bereits innerhalb der Hardwarestrukturen, also auf niedrigem Abstraktionsniveau, realisiert bzw. unterstützt werden können. Zur Realisierung von HW Beschleunigern ist ebenfalls die Untersuchung geeigneter Toolflows notwendig. Schließlich soll ein Demonstrator aufgebaut werden, der die Leistungsfähigkeit der rekonfigurierbaren Hardware demonstriert.

Dezentrale Rekonstruktion verteilter kontinuierlicher Phänomene aus orts- und zeitdiskreten Messungen

Felix Sawo

`sawo@ira.uka.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke
Fakultät für Informatik, Institut für Technische Informatik, Universität Karlsruhe (TH)

Eine der Hauptanwendungen von Sensor-Aktor-Netzwerken ist die Beobachtung von räumlich verteilten physikalischen Phänomenen, wie beispielsweise einer Temperatur- oder Schadstoffverteilung. Dabei nehmen in der Umwelt verteilte Sensor-Aktor-Knoten allgegenwärtige Daten auf, verarbeiten diese und führen entsprechende Aktionen aus.

Das Ziel des Teilprojektes I1 ist es, aus den zeit-, orts- und wertdiskreten Messungen sowohl das gesamte Phänomen effizient zu rekonstruieren als auch nicht direkt messbare Größen, z. B. Position einer Schadstoffquelle oder Diffusionskoeffizient einer Wärmeausbreitung, zu identifizieren. Für die Rekonstruktion der gewünschten Information aus den messbaren Daten werden modellbasierte Verfahren untersucht und entwickelt. Durch die Einbeziehung von Hintergrundwissen in Form eines physikalischen Modells wird die Qualität der Schätzergebnisse verbessert und nur eine geringe Anzahl an Sensoren benötigt. Damit ist es möglich, die charakteristischen Größen zu bestimmen, welche das verteilte Phänomen an jedem Ort und zu jedem Zeitpunkt vollständig beschreiben. Dadurch können auch Aussagen über das Phänomen an Nichtmesspunkten getroffen werden. In der Regel treten Unsicherheiten sowohl bei der Modellierung des physikalischen Phänomens, z. B. durch nicht modellierbare Effekte oder Störungen, als auch bei der eigentlichen Messung auf. Die in diesem Teilprojekt zu entwickelnden Verfahren und Methoden zeichnen sich durch eine systematische und integrierte Betrachtung *aller* auftretenden Unsicherheiten aus, d. h. ausgehend von dem Modellierungsprozess bis hin zur eigentlichen Fusion der Messungen.

Bei großen Sensor-Aktor-Netzwerken wird aus verschiedenen Gründen eine dezentrale Verarbeitung der Informationen in den einzelnen Knoten einer Verarbeitung in einem Zentralrechner vorgezogen. Dabei werden die von den Sensor-Aktor-Knoten gesammelten Informationen lokal, d.h. vor Ort, vorverarbeitet, anschließend durch das Sensor-Aktor-Netzwerk propagiert und mit den Informationen anderer Sensoren zur Verbesserung der Schätzung entsprechend fusioniert. Durch diese Fusion und das Propagieren der Information durch das Sensor-Aktor-Netzwerk entstehen zwischen den Teilschätzungen auf den einzelnen Sensor-Aktor-Knoten stochastische Abhängigkeiten. Da aufgrund begrenzter Kommunikationskapazitäten und Speicherressourcen diese Abhängigkeiten zwischen den auf unterschiedlichen Sensor-Aktor-Knoten gespeicherten Teilschätzungen nicht aufrechterhalten werden können, gehen diese verloren. Allerdings ist bei der Fusion von zwei Teilschätzungen deren Abhängigkeitsstruktur, also im Allgemeinen die Verbund-

wahrscheinlichkeitsdichte, erforderlich. Deshalb sollen Verfahren entwickelt werden, welche die teilweise *unbekannte* Abhängigkeitsstruktur systematisch vernachlässigen, aber trotzdem konsistente und vor allem konservative Schätzergebnisse liefern. Die entwickelten Schätzverfahren sollen dann für eine *dezentralen* Rekonstruktion und Identifikation von verteilten Phänomenen angewandt werden. Somit ist das vorrangige Ziel des Teilprojektes I1 die Entwicklung einer systematischen Methodik zur dezentralen Rekonstruktion verteilter Phänomene, welche die Randbedingungen von Sensor-Aktor-Netzwerken erfüllt, d. h. Berücksichtigung aller auftretenden Unsicherheiten, dezentrale Informationsverarbeitung und eine autonome Anpassung an das zu untersuchende Phänomen.

Kapazitätsbetrachtungen, Protokolle und Architekturen

Felix Schmidt-Eisenlohr

`fschmidt@tm.uni-karlsruhe.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Telematik, Universität Karlsruhe (TH)

Im Rahmen des Teilgebietes K1 werden die Fähigkeiten, Einschränkungen und Möglichkeiten, die der drahtlose Übertragungskanal dem Gesamtsystem bietet, untersucht. Als Fallbeispiel dient die Inter-Fahrzeug-Kommunikation, bei der zur Erhöhung der Verkehrssicherheit die Kommunikation zwischen den Fahrzeugen erforscht wird, die in diesem Szenario die Sensor-Aktor-Knoten darstellen. Es gibt zwei typische unterschiedliche Nachrichtentypen, so genannte Beacon-Nachrichten, die periodisch von allen Knoten versandt werden (z.B. zur Positions- und Geschwindigkeitsbestimmung der Fahrzeuge), und ereignisbasierte Nachrichten, die bei Eintritt spezieller Ereignisse verwendet werden, beispielsweise bei aktuen Gefahren wie Stau oder Glatteis.

Für diese Art der Nutzung des drahtlosen Netzwerkes werden die typischen Charakteristika des Kanals und der möglichen Datenübertragung unter verschiedenen Lastsituationen untersucht. Hierbei sind insbesondere unterschiedliche verkehrliche Situationen zu unterscheiden. Für die einzelnen Szenarien werden optimale Strategien und Parameter für die Aussendung von Beacon- und Ereignisnachrichten bestimmt. Die Ergebnisse zeigen die Kapazitätsgrenzen des drahtlosen Mediums auf, erlauben aber ebenfalls die Anpassung der Strategien an diese Gegebenheiten zum Erreichen eines bestmöglichen Ergebnisses. Eine wesentliche Rolle kommt den verwendeten Modellen für Kanal, Mobilität und den einzelnen Schichten der Netzwerkknoten zu.

Für die zur Untersuchung notwendigen Simulationsstudien wird der Netzwerksimulator ns-2 verwendet. Im Rahmen der Arbeiten wurden akkuratere, realitätsnahe Modelle des Kanals, der physikalische Schicht sowie der Medienzugriffsschicht entworfen, implementiert und validiert, so dass weitaus realistischere Ergebnisse erzielt werden können als dies ansonsten möglich wäre.

Die bisherigen Ergebnisse zeigen, dass abhängig von den jeweils vorhandenen Bedingungen spezifische Parameter wie Sendestärke, Contention Window, Beaconrate oder verwendetes Modulationsschema unterschiedlich festzulegen sind, um ein optimales Ergebnis zu erreichen. Die Ergebnisse zeigen ebenfalls, wie unerlässlich eine akkurate Modellierung der Szenarien ist, da ansonsten massiv abweichende Ergebnisse erzielt werden. Ziel der Arbeit ist es somit, zum einen für gegebene Bedingungen einen Satz an optimalen Parametern angeben zu können, der ein bestmögliches Verhalten des Gesamtsystems erlaubt. Zum anderen sollen die Charakteristika des Systems in Abhängigkeit von Bedingungen und gewählten Parametern bestimmt werden können, und der Einfluss der einzelnen Parametern aufgezeigt werden.

Selbstorganisation im Bereich Middleware-Architekturen

Sebastian Schuster

sschu@ira.uka.de

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Prozessrechentechnik, Automation und Robotik,

Universität Karlsruhe (TH)

Thema des Teilprojekts H3 „Selbstorganisation im Bereich Middleware-Architekturen" ist die Untersuchung der Anwendungsentwicklung für Sensor-Aktor-Netzwerke (SAN). Charakteristisch für SANs ist ihre Einbettung in die reale Welt. Dabei arbeiten sie oft unüberwacht, da aufgrund der Einsatzszenarien und der Anzahl der Knoten eine manuelle Unterstützung durch den Menschen unmöglich ist. Daher ist es für SANs notwendig, sich autonom an wechselnde Umweltbedingungen anzupassen.

Um produktiv Software entwickeln zu können, müssen diese Algorithmen nach ihrer Implementierung zur Wiederverwendung durch den Entwickler bereitgestellt werden, statt sie für jede Anwendung von Neuem zu implementieren. Für verteilte Systeme typisch ist die Nutzung von Middleware als Schicht über dem Betriebssystem der beteiligten Knoten. Sie bietet generische Lösungen für oft auftretende Probleme verteilter Systeme. So können beispielsweise die Heterogenität des Netzes und Kommunikationsfehler vor dem Entwickler verborgen werden, indem entsprechende Algorithmen und Protokolle in der Middleware implementiert sind. Der Entwickler kann auf höherem Abstraktionsniveau und damit produktiver Software entwickeln, ohne immer wieder ähnliche Probleme lösen zu müssen.

Da sich die verfügbaren Ressourcen in einem SAN stark unterscheiden und im Falle von Sensor-Aktor-Knoten sehr beschränkt sind, ist eine einfache Übertragung des Middleware-Ansatzes nicht ohne weiteres möglich. Nötige Anpassungen oder alternative Ansätze zu entwickeln, ist daher das Hauptziel dieses Teilprojekts.

Algorithmen zur Wiederverwendung zur Verfügung zu Stellen ist nur dann sinnvoll, wenn sie auch mehrfach bei der Entwicklung verschiedener Anwendungen genutzt werden. Die Vision des Ubiquitous Computing beschreibt eine Klasse von allgegenwärtigen Anwendungen mit charakteristischen Problemen, bei denen, aufgrund ihrer Verbreitung, eine Wiederverwendung von Lösungen von großem Vorteil wäre. Ubiquitäre Anwendungen unterstützen den Menschen in ihrem täglichen Leben unaufdringlich, indem sie wahrnehmen was der Mensch tut und darauf reagieren. Sensor-Aktor-Netzwerke bilden damit die Infrastruktur für solche Anwendungen.

Typische Zielumgebungen ubiquitärer Anwendungen unterscheiden sich jedoch sehr stark hinsichtlich der dort vorhandenen Geräte – beispielsweise gleicht keine Wohnung einer anderen exakt. Gleichzeitig sind die verfügbaren Geräte sehr heterogen, sie reichen von Sensor-Aktor-Knoten über PDAs und Computer bis zu eingebetten Geräten. Zum einen

kennt der Entwickler die Zielinfrastruktur also nicht genau, so dass Middleware ihm eine möglichst abstrakte Sicht auf die Zielumgebung bieten muss. Zum anderen machen die unterschiedlichen und teilweise stark beschränkten Ressourcen in der Zielumgebung den Einsatz einer klassischen Middleware, die einen hohen Abstraktionsgrad bieten könnte, unmöglich.

Daher wird in diesem Teilprojekt der Einsatz von generativen Verfahren, die Code passend für verschiedene Zielumgebungen erzeugen können, untersucht. Der Entwickler soll auf hohem Abstraktionsniveau spezifisch für die Zieldomäne ubiquitärer Anwendungen entwickeln können – daraus werden maßgeschneiderte Lösungen per Transformation, ähnlich den modellgetriebenen Verfahren, entwickelt. Schwerpunktprobleme bilden hierbei die Heterogenität der Zielgeräte und die Dynamik in ubiquitären Umgebungen. Diese Aspekte sollen im Rahmen des Transformationsprozesses berücksichtigt werden, so dass sich der Entwickler auf die Lösung der Kernprobleme seiner Anwendungen konzentrieren kann.

Anfragebearbeitung in Sensornetzen

Mirco Stern

`Mirco.Stern@ipd.uni-karlsruhe.de`

GRK 1194 Selbstorganisierende Sensor-Aktor-Netzwerke

Fakultät für Informatik, Institut für Programmstrukturen und Datenorganisation,

Universität Karlsruhe (TH)

Dieses Teilprojekt beschäftigt sich mit Problemen der Bearbeitung deklarativer Anfragen auf Sensornetzen mit dem Ziel, deklarative Anfragen beispielsweise mittels einer SQL-ähnlichen Anfragesprache bereitzustellen. Das zugrunde liegende technische Problem ist die Umsetzung einer solchen Anfrage in einen ausführbaren Plan. Bei der Abbildung deklarativer Anfragen auf Ausführungspläne steht vor allem die Anforderung der Energie-effizienz im Vordergrund. Dabei werden im Rahmen dieses Projekts vor allem zwei Optimierungsziele betrachtet. Zum einen wird untersucht, wie zu verfahren ist, um den Energieaufwand insgesamt zu minimieren und zum anderen wie eine möglichst gleichmäßige Verteilung des Energieaufwands erreicht werden kann, um den vorzeitigen Ausfall einzelner Knoten zu vermeiden. Mit Blick auf das Gesamtziel der Abbildung deklarativer Anfragen auf Ausführungspläne ist das Projekt in zwei Schwerpunkte untergliedert:

(1) Bereitstellung von effizienten Methoden für die in Anfragen auftretenden Operatoren
(2) Erstellung eines optimalen Ausführungsplans basierend auf diesen Operatoren.

Bereitstellung effizienter Methoden für Operatoren. Bisher wurden vor allem Anfragen nach Werten einzelner Sensoren sowie Aggregaten darüber betrachtet, d.h. es existieren effiziente Methoden zur Bearbeitung von Selektions-, Projektions- sowie Aggregationsoperatoren. Offen ist, wie Anfragen nach korrelierten Werten unterstützt werden können. Insbesondere eine effiziente Bearbeitung des Verbund-Operators (Join) ist derzeit ein offenes Problem. Dies bedeutet eine besondere Einschränkung der Ausdrucksmächtigkeit von Anfragen, da der Join der klassische Datenbankoperator ist, mit dessen Hilfe Korrelationen in den Daten ausgedrückt werden. In Folge dessen nimmt der Join in diesem Projekt eine besondere Stellung ein. Die Methodik zum Entwurf effizienter Methoden ist hier die Verlagerung von Teilen der Ergebnisberechnung ins Netz („In-network Processing"), um möglichst frühzeitig das zu übertragende Datenvolumen zu reduzieren.

Erstellung eines optimalen Ausführungsplans. Ein wesentliches Merkmal der Verbund-Berechnung ist, dass sich je nach aktuell gemessenen Daten unterschiedliche Methoden als am effizientesten erweisen. Hier ergibt sich das Problem der Wahl der momentan besten Methoden, welches einen Teil des klassischen Optimierungsproblems im Datenbankbereich bildet. Zu dieser Entscheidung benötigt man Wissen über die vorliegende Situation im Netz. Hier ergibt sich die Frage, welches Wissen in Sensornetzen benötigt wird und wie dieses Wissen bereitzustellen ist. Eine Herausforderung im Vergleich zu herkömmlichen Datenbanken ist, dass sich die von den Sensoren gemessenen Werte zum Teil sehr schnell

ändern. Die Bereitstellung der notwendigen Informationen sowie deren Nutzung zur Wahl eines optimalen Ausführungsplans bilden den zweiten Schwerpunkt dieses Projekts.

GRK 1298:

AlgoSyn

Algorithmische Synthese reaktiver und diskret-kontinuierlicher Systeme

Aachen

1 Thema

Methoden der Softwarevalidierung und -verifikation sind heute gut etabliert, abgesichert durch adäquate formale Modelle und auch erprobt in der praktischen Anwendung. Dagegen ist der anspruchsvollere Ansatz der automatischen Synthese von Software (und Hardware) erst rudimentär entwickelt, trotz der Tatsache, dass es ein Potential für den Einsatz von Syntheseverfahren in zahlreichen Bereichen gibt, vor allem und in Leit- und Steuerungssystemen und in (Multi-Agenten-) Systemen mit geringer Datenkomplexität. Sowohl in der theoretischen Informatik als auch in der Praxis der Ingenieurdisziplinen wird die Frage der Synthese verfolgt, allerdings aus ganz unterschiedlichen Blickwinkeln. Das Graduiertenkolleg AlgoSyn will eine Integration dieser Forschungen erreichen und ein Instrumentarium von formalen Modellen und algorithmisch realisierbaren Verfahren entwickeln, welches auf die Anforderungen einschlägiger Ingenieurdisziplinen passt. Zentrale Forschungsprobleme in diesem Feld sind einerseits die Etablierung von System-Modellen, welche algorithmische Lösungen des Syntheseproblems unterstützen, andererseits (ähnlich wie in der Verifikation) die Kombination von diskreten und kontinuierlichen Parametern (hybride Systeme) und schließlich die Integration der Theorie-Forschung mit Modellen und Sprachen, die sich in verschiedenartigen Anwendungsbereichen herausgebildet haben. Im Graduierten-kolleg AlgoSyn wird diesen Herausforderungen durch eine Bündelung der Expertise aus Informatik, Mathematik und vier Ingenieurdisziplinen (Prozessorarchitekturen, Regelungstechnik, Prozessleittechnik sowie Schienenverkehrswesen) begegnet. Sowohl ein forschungsnahes Lehrprogramm als auch die interdisziplinären Themen der Dissertationsvorhaben sollen die nötige Methodenintegration vorantreiben. Die Forschungen erfolgen in vier Themenbereichen. Zwei Bereiche („Algorithmik für agentenbasierte, probabilistische und hybride Systeme" sowie „Formale Modelle reaktiver Systeme und spieltheoretische Methoden") haben die Weiterentwicklung der theoretischen Grundlagen zum Ziel. Der Bereich „Softwaretechnische Einbettung und Modellierungssprachen" bildet die Schnittstelle zu den Ingenieurfächern, und im Bereich „Anwendungen und Demonstratoren" wird die Systementwicklung in unterschiedlichen Feldern wie Chipdesign, Prozessleittechnik und Verkehrswesen verfolgt.

2 Topic

Methods of software validation and verification are by now well established, based on adequate formal models and successful in practical applications. On the other hand, the more ambitious approach of automatic synthesis of software (and hardware) is as yet only developed in quite rudimentary form, despite the fact that there is a potential for applying methods of synthesis in several domains, such as control systems or multi-agent systems of low data complexity. In theoretical computer science as well as in engineering disciplines the issue of synthesis is pursued, however from quite different viewpoints and with different methods. The research training group AlgoSyn aims at an integration of this research and at the development of a stock of models and algorithmic procedures that fits the needs of the engineering disciplines involved. Central issues in the area are the development of system models which allow an algorithmic solution of the synthesis problem, the combination of discrete and continuous parameters in hybrid systems (as this is also familiar from verification), and the exploration of the potential of applications. The research training group meets this challenge by a combination of expertise from computer science, mathematics, and four engineering subjects (processor architectures, automatic control, process control engineering, train traffic systems). Both a research oriented program of seminars and courses and interdisciplinary research topics in the dissertation projects should push forward the desired integration of methods. The research will be carried out in four subject areas. Two of them ("Algorithmics for agent-based probabilistic and hybrid systems", "Formal methods of reactive systems and game-theoretic methods") are concerned with theoretical foundations, the area "Software development and modelling languages" serves as methodological interface to engineering, and the latter is represented by the area "Applications and demonstrators" with an emphasis on chip design, process control engineering, and train traffic control systems.

Steuerungssynthese für diskret-kontinuierliche Systeme mit Hilfe von Petri-Netzen

Kai Bollue

Institut für Regelungstechnik
RWTH Aachen
Steinbachstraße 54
52074 Aachen
kai.bollue@irt.rwth-aachen.de

Abstract: Ziel ist die Erweiterung bestehender Theorien zur Steuerungssynthese für diskrete Systeme auf hybride Systeme unter Berücksichtigung erweiterter Anforderungen. Hierbei liegt der Schwerpunkt auf der Anwendung von Petri-Netzen.

Ziele

Es besteht bereits einige Theorie über die Steuerungssynthese für diskrete Systeme. In zahlreichen Anwendungsfällen ist dies jedoch nicht ausreichend. Daher sollen die Möglichkeiten untersucht werden, eine Steuerungssynthese auch für hybride Systeme zu realisieren, d.h. die Eigenschaften kontinuierlicher Komponenten zu berücksichtigen. Hierzu soll nach Möglichkeit von der existierenden Theorie zu Petri-Netzen profitiert werden. Neben den „klassischen" Anforderungen an eine Steuerung sind auch weitere Merkmale wie etwa Robustheit gegenüber unvorhergesehenen Störungen denkbar.

Anwendungsbeispiele

Mögliche Anwendungsbeispiele ergeben sich im Rahmen der Modellfabrik des Instituts für Regelungstechnik. Denkbar wäre hier die Steuerung des Greifers eines Hochregallagers unter Berücksichtigung zeitlicher Anforderungen. Hierbei sind die Anfahrzeiten einzelner Positionen korrekt im Modell abzubilden und bei der Steuerungssynthese zu berücksichtigen.

Algorithmic Synthesis of Hybrid Control Systems

Gerlind Herberich

Lehrstuhl für Informatik 11, Software für eingebettete Systeme, Prof. Dr.-Ing. Stefan Kowalewski
RWTH Aachen
Ahornstraße 55
52074 Aachen
herberich@informatik.rwth-aachen.de

Methods for the synthesis of control systems have been developed for purely continuous and for purely discrete systems. However, real control systems are often hybrid, that is to say that discrete as well as continuous dynamics have to be considered. In many cases, a large amount of the functionality is determined by the discrete behaviour of the system. The modelling of such hybrid systems can then be realized by combining discrete transition systems with piecewise linear continuous transfer elements [AC95].

Algorithms for the synthesis of this class of systems only exist in quite rudimentary form. Based on the results that have been obtained concerning the synthesis of purely discrete systems such as game-theoretical approaches or the supervisory control theory [RW89], a synthesis approach for hybrid systems has to be developed as well as to be explored in terms of application potential for industrial use, especially in embedded systems.

Literaturverzeichnis

[RW89] Ramadge, P.; Wonham, W.: The control of discrete event systems. In: Proc. of the IEEE, Vol. 77, No. 1, 1989.

[AC95] Alur, R.; Courcoubetis, C. et. al.: The algorithmic analysis of hybrid systems. In: Theoretical Computer Science, Vol. 138, No. 1, p. 3-34, 1995.

Complexity of Winning Strategies in Infinite Games

Michael Holtmann

Lehrstuhl für Informatik 7
RWTH Aachen, 52056 Aachen, Germany
holtmann@i7.informatik.rwth-aachen.de

Infinite Games. Given a (finite) directed graph $G = (Q, E)$ with no dead ends any infinite path through G is considered a play ρ of the game $\Gamma = (G, \varphi)$. It is played by Player 0 and Player 1 each of them having his own set of vertices, i.e. $Q = Q_0 \,\dot\cup\, Q_1$. The set $\varphi \subseteq Q^\omega$ of winning plays for Player 0 is determined by a winning condition from one of the following four categories: A subset $F \subseteq Q$, a family $\mathcal{F} = \{F_1, \ldots, F_k\}$ of subsets, a family $\mathcal{F}_{pair} = \{(E_1, F_1), \ldots, (E_k, F_k)\}$ of pairs of subsets, a coloring $c : Q \to \{0, \ldots, k\}$ ($k \in \mathbb{N}$). I am working on complexity issues for computing and formalizing winning strategies in infinite games. A strategy can be categorized according to its complexity which is basically the minimal amount of memory that is needed to implement it. Let $Occ(\rho)$ be the set of all vertices visited at least once in ρ, $Inf(\rho)$ the set of vertices visited infinitely often in ρ and $n := |Q|$.

Game	Acc.	Winning plays of Player 0	Memory
Reachability	F	$Occ(\rho) \cap F \neq \emptyset$	None
Büchi	F	$Inf(\rho) \cap F \neq \emptyset$	None
parity	c	$\max(Inf(c(\rho)))$ even	None
Staiger-Wagner	$\mathcal{F}$	$Occ(\rho) \in \mathcal{F}$	2^n
Muller	$\mathcal{F}$	$Inf(\rho) \in \mathcal{F}$	$n! \cdot n$
Streett	$\mathcal{F}_{pair}$	$\forall i(Inf(\rho) \cap F_i \neq \emptyset \implies Inf(\rho) \cap E_i \neq \emptyset)$	$k! \cdot k^2$
Request-Response	$\mathcal{F}_{pair}$	$\forall i \forall r(\rho(r) \in E_i \implies \exists s \geq r : \rho(s) \in F_i)$	$2^{k+1} \cdot k$

Strategies and Memory. The easiest kind of a strategy is called "positional" (or "memoryless"). Having a positional winning strategy means that from a given vertex a player can always choose the same edge no matter what happens in the history of the play. Such a strategy can be implemented by a finite automaton with output (Mealy machine) with only one state, i.e. no memory is required. For example, in a reachability game both players have positional winning strategies.

A "finite-state" strategy depends on the history of a play and, hence, is more complex. To implement it we need a (finite) memory. For the most common winning conditions lower bounds for the required memory are known. These bounds are merely reached in exotic examples but, usually, less memory is needed. In my diploma thesis I have developed an algorithm that reduces the memory necessary for implementing winning strategies in infinite games. It has been applied to Staiger-Wagner and Request-Response games and, very recently, to Muller and Streett games. Later in my research I will focus on the problem of how to find winning strategies that are optimal. There, optimality of a strategy has several meanings:

- What are possible ways of representing the components of a strategy (memory, specification by a program,...) and which one has better/optimal complexity for computation/formalization?

- How can we search effectively for an optimal winning strategy? Which parameters of known solution algorithms do have which impact on the computed strategy?

Games, Reductions and Model Checking

Łukasz Kaiser

Mathematische Grundlagen der Informatik, RWTH Aachen, Germany
kaiser@informatik.rwth-aachen.de

Infinite ω-regular games are an active topic of interest in computer science as they are used in verification and synthesis of reactive systems. The possibility to use games for that purpose comes from the fundamental theorem [1, 2] stating that parity games are exactly model checking games for the μ-calculus. Another important result is the possibility to use latest appearance record as memory and so reduce general Muller games to parity games. Moving to games as a model for checking formulas of μ-calculus turned out to be very fruitful. Not only is the game easier to understand intuitively than the formula, but efficient algorithms, e.g. strategy improvement [3], can be used to solve the game.

The success of games for μ-calculus motivates our work to look for other game classes that can be used for model checking interesting logic and to describe reductions between game classes. We are especially looking at first-order logic on finitely presented structures and we define model-checking games with finite arenas for such structures. These are multi-player games with imperfect information and we study the relation between information exchange in games and quantifiers in logic in more detail [4].

Even though we defined model checking games on finite arenas, it is sometimes useful to look at games on infinite arenas. Especially when doing game reductions it is often not clear whether the reduction method works for infinite or only for finite arenas. We study reductions of infinitary games and provide a memory structure called finite appearance record that is a generalisation of latest appearance record to games on infinite arenas [5].

References

[1] E. A. Emerson and C. S. Jutla, Tree automata, mu-calculus and determinacy. In *Proc. 5th IEEE Foundations of Computer Science*, pages 368–377, 1991.

[2] A. W. Mostowski, Games with Forbidden Positions. In *Technical Report 78*, Instytut Matematyki, Uniwersytet Gdański, 1991.

[3] J. Vöge and M. Jurdziński, A discrete strategy improvement algorithm for solving parity games. In *CAV 2000*, volume 1855 of LNCS, pages 202–215, Springer-Verlag, 2000.

[4] Ł. Kaiser, Game Quantification on Automatic Structures and Hierarchical Model Checking Games. In *CSL 2006*, volume 4207 of LNCS, pages 411–425, Springer-Verlag, 2006.

[5] E. Grädel and Ł. Kaiser, What kind of memory is needed to win infinitary Muller games? In *Proc. th Augustus de Morgan Workshop on Interactive Logic, Games and Social Software*, 2007.

Modelle zur Kapazitätsberechnung großer Eisenbahnnetze

Sebastian Klabes

GK AlgoSyn
RWTH Aachen
Mies-von-der-Rohe-Str. 1
52074 Aachen
klabes@via.rwth-aachen.de

1 Motivation

Mit der geplanten Liberalisierung im Schienenverkehr treten neue Fragestellungen in den Vordergrund. Während in der Vergangenheit das ganze Schienenverkehrssystem Deutschlands von einem Unternehmen (ehemals Staatsbetrieb) betrieben wurde, sollen in Zukunft das Schienennetz und der Fahrbetrieb getrennt bewirtschaftet werden. Zusätzlich zu dieser grundsätzlichen Trennung soll es in Zukunft auch unterschiedlichen Akteuren ermöglicht werden, Fahrbetrieb auf der Schieneninfrastruktur zu betreiben.

2 Problemstellung

Der Betrieb des schienengebundenen Verkehrs stellt wesentlich höhere Sicherheitsansprüche als beispielsweise der Straßenverkehr. Das macht eine exakte Planung des Betriebs notwendig.

Bei dieser Planung ist zu beachten, dass die Akteure unterschiedliche Anforderungen an die Infrastruktur haben und der Betreiber der Infrastruktur unter Umständen auch Prioritätsunteschiede bei der Trassenvergabe zu beachten hat. Dieses Betriebsplanungsproblem stellt sich somit als sehr komplex dar und soll unter Anwendung spieltheoretischer Ansätze gelöst werden.

3 Forschungsvorhaben

Um die Kapazität bzw. Auslastung einer gegebenen Infrastruktur zu berechnen, muss nicht nur die Topologie der Infrastruktur, sondern auch der darauf zu betreibende Verkehr bekannt sein. Zunächst müssen also die Trassenwünsche der unterschiedlichen Akteure berücksichtigt und eingeplant werden, um das zu erwartenden Verkehrsaufkommen zu bestimmen. Die Trassenwünsche beinhalten in der Regel Konflikte, die es gilt mit Hilfe von spieltheoretischen Ansätzen diskriminierungsfrei zu lösen. Im Rahmen dieses Forschungsprojektes soll zunächst ein Mehrpersonenspiel spezifiziert und implementiert werden, das diesen Anforderungen genügt.

Agent-Based Resource Management

Lars Olbrich

Algorithms & Complexity
RWTH Aachen
Ahornstr. 55
52064 Aachen
lars@cs.rwth-aachen.de

Abstract: Consider Wardrop's traffic model: an infinite set of agents each of which controls an infinitesimal amount of traffic allocate a path between a source/destination pair in a given network. The cost of a resource (edge) is a function of the congestion. We are interested in efficient approximation algorithms for fair and stable solutions to this problem.

Optimisation and Game Theory

In today's large-scale networks, most prominent the Internet, no authority is capable of enforcing a centralized data management. In that situation, both game theory and optimization come into play.

In optimization, one seeks for an optimal solution among a set of feasible solutions usually specified in a concise way, e.g., the set of feasible solutions is the set of paths between a specified source/destination pair in a given network and the objective is to find a shortest path. In game theory, several players seek at maximizing their personal payoff. Each player can choose among different strategies and its payoff depends on the choices of all players.

Wardrop's Traffic Model

We concentrate on the well known game theoretic traffic model due to Wardrop, which has been studied in the 1950's in the context of road traffic systems.

In this model, we are given a network equipped with load-dependent latency (negative payoff) functions on the edges, a fixed traffic demand is to be routed from its source to its destination. The traffic is splittable into arbitrarily small pieces, i.e., traffic is modelled as flow. From a game-theoretic perspective, this establishes a game with infinitely many users, each carrying an infinitesimal amount of traffic from its source to its destination. As each user acts selfishly, it aims at minimising its personal cost, which is defined to be the sum of the edge latencies on a path connecting its source with its destination.

In this environment the Wardrop-equilibrium captures the idea of rational user behaviour: all paths used between a given source-destination pair have equal latency. Since equilibrium flows are solutions to a convex optimization program, they can be computed in polynomial time. Nevertheless non-trivial upper bounds on the complexity of computing (approximate) equilibria are still to establish.

Recent Work

We recently considered adaptive routing policies in a dynamic extension of this model. In a given network with several source-destination pairs and corresponding flow demands, assume that every source-destination pair is controlled by an agent, which aims at balancing its flow. In a round-based fashion each of the agents may simultaneously redistribute its flow once in a round based on latency observation at the beginning of the round. Using our distributed algorithm, one quickly approaches approximate Wardrop equilibria.

Further Research

A weakness of our notion of approximate equilibria, is the fact that the average latency may be far away from the minimum latency. As an alternative, one could also consider deviations from the minimum latency rather from the average latency. One promising approach in that direction might be to adapt techniques designed to approximate multi-commodity flow ([AR], [GK]). Another line of research focuses on efficient balancing algorithms for dynamical networks.

Literaturverzeichnis

[AR]   Awerbuch; Rao: *Distributed Algorithms for Multicommodity Flow Problems via Approximate Steepest Descent Framework*, In Proc. of the ACM-SIAM Symposium on Discrete Algorithms (SODA), 2007.

[GK]   Garg, Könemann: *Faster and Simpler Algorithms for Multicommodity Flow and other Fractional Packing Problems*, In Proc. of the 39th Annual IEEE Computer Society Conference on Foundations of Computer Science, 1998.

[Wa]   Wardrop: *Some theoretical aspects of road traffic research*, In Proceedings of the Institute of Civil Engineers, Pt. II, volume 1, pages 325–378, 1952.

Decentralised Analysis and Safety Assurance of Product Flow Paths

AlgoSyn - Algorithmic Synthesis of Reactive and Discrete-Continuous Systems

Gustavo Quirós

Lehrstuhl für Prozessleittechnik, RWTH Aachen

Gustavo.Quiros@plt.rwth-aachen.de

Production and processing plants commonly consist of a network of interconnected components through which products flow during operation. The routes or paths that these products take depend not only on the structure of the plant, but also on the dynamic state of some types of components, for instance valves. The precise and up-to-date knowledge about the product paths in a plant is of great benefit to plant automation systems, which may use this information to fulfill control tasks such as product tracking, process safety assurance, diagnostics, computation of physical balances, handling of device failures and optimisation of plant usage.

Given the ever-growing complexity of modern-day processing plants, decentralised or component-based automation systems offer greater flexibility, scalability and maintainability than their centralised or monolithic counterparts. Therefore, modern solutions to automation problems present an added benefit when they are implemented in a decentralised fashion.

This work concentrates on determining decentralised techniques for the analysis and safety assurance of product flow paths. The former consists of determining which product flow paths are present in a plant at a given plant state. The latter uses this information to determine and enforce safe working conditions for the use of such paths. In both cases, the corresponding algorithms must adapt fully not only to state changes of the components of the plant, but also to changes in the structure of the plant itself. In this way, system synthesis is achieved in a dynamic manner by means of adaptive self-transformation.

# Signaltechnisch sichere Systeme im Eisenbahnwesen

Dipl.-Ing. Dipl.-Kfm. Bernd D. Schmidt

Abstract: Ein Drittel der Kosten im Bereich der Vorhaltung von Eisenbahninfrastruktur wird durch Signal- und Kommunikationstechnik generiert. Ziel dieses Projektes ist dabei unter Verwendung moderner Mikrocontroller ein modulares, den bisherigen Systemen vergleichbares System zu entwickeln und durch Massenproduktion die Kosten zu senken.

Das Projekt *Signaltechnisch sichere Systeme im Eisenbahnwesen* beschäftigt sich im Rahmen des Graduiertenkollegs Algosyn mit der Nutzung moderner eingebetteter System in der Eisenbahnsicherungstechnik.

1 Motivation

Der Verkehrsträger Eisenbahn soll seinen Fahrweg im Gegensatz zum Konkurrenten Straße selber finanzieren. Dies führt z. Z. zu einem Anteil der Infrastrukturkosten von ca. 40 % an den Gesamtkosten einer Zugfahrt. Bei näherer Betrachtung setzt sich dieser Anteil wie folgt zusammen:
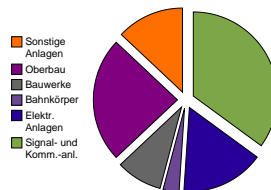
Auffällig ist, daß ein Drittel der Kosten durch Signal- und Kommunikationstechnik generiert wird. Diese ist somit ebenso teuer, wie der eigentliche Fahrweg. Bei der Suche nach der Ursache könnten die folgenden Aspekte relevant sein:

- monopolistische Marktstrukturen seitens der Zulieferer,

- kaum Standardisierung von Komponenten und Schnittstellen,

- Sicherheitsstandards (CENELEC-Normen etc.) fordern immer Eigenschaften des Gesamtsystems, womit ein System aus Komponenten unterschiedlicher Zulieferer durch kostspielige Sicherheitsanalysen inzwischen wirksam verhindert wird,

- kein Baukastensystem, die Sicherheitslogik jedes Bahnhofs ist eine Einzelanfertigung,

- steigende Kosten durch Obsoleszenz elektronischer Bauteile,

- veraltete Schnittstellenkonzepte für die Außenanlagen mit immensen Tiefbau- und Verkabelungskosten,

- moderne Stellwerke inkompatibel zu Altsystemen und teuer bei Sonderanforderungen.

Folge aus diesem Wildwuchs ist ein erheblicher Wettbewerbsnachteil für den Verkehrsträger Eisenbahn. Aus vergleichbaren industriellen Anforderungen unter Wettbewerbsbedingungen kann ein Anteil von ca. 10 % der Infrastrukturkosten als ein realistischer Sollwert für die Eisenbahnsignaltechnik angesehen werden.

Ziele und Methoden

- Entwicklung eines Konzeptes zum automatisierten Entwurf von Stellwerken

- Anwendung von Verfahren der algorithmischen Synthese in der Eisenbahnsicherungstechnik

- Sicherheitsnachweis über probabilistische Risikobetrachtungen

    - Festlegung von Risikoakzeptanzkriterien (z. B. ''Nachweis der mindestens gleichen Sicherheit" in Deutschland)
    - Vollständig quantitative Betrachtung der neuen Technologie
    - Retrospektive quantitative Sicherheitsanalyse für die Referenztechnologie

- Reduktion von Komplexität und Größe des Zustandsraumes mittels geeigneter Methoden zur sicheren Seite hin

- Pessimistischen Schätzungen der Sicherheit dürfen keine Kostensteigerungen verursachen

Literatur

[1] Braband, J.: Erfahrungen mit Risikoanalysen in der Eisenbahnsignaltechnik. In *Eisenbahntechnische Rundschau*, Jg. 51 (2002) 9, S. 575–584.

[2] Braband, J.: Risikoakzeptanzkriterien und -bewertungsmethoden. Ein systematischer Vergleich. In *Signal+Draht*, Jg. 96 (2004)4, S. 6–9.

[3] Meyer zu Hörste, M.: Methodische Analyse und generische Modellbildung von Eisenbahnleit- und -sicherungssystemen. In *VDI Fortschritt-Berichte*, Reihe 12, Nr. 571, Düsseldorf (2004).

Optimizing Compilers for Heterogeneous Multi-Processor Systems

Weihua Sheng

Institute for Software for Systems on Silicon
RWTH Aachen University
Templergraben 55
52056 Aachen
sheng@iss.rwth-aachen.de

With increasing demands for high performance and complex embedded systems in mobile and multimedia applications, *Multiprocessor System-on-Chips* (MPSoCs) are quickly becoming the next trend in VLSI technology [1]. A modern MPSoC (figure 1) usually combines dozens or even hundreds of specialized programmable processors, which provides powerful processing power and flexibility at the same time. Driven by the heterogeneous and programmable natures of MPSoC platforms, the efficient and optimal way to allocate and deploy application tasks on an MPSoC platform is very much needed. "The processor is the NAND gate of the future" but "how to program them" is yet to be investigated. The goal of this research topic is to create a new MPSoC application programming environment, which captures the coarse-grained tasks from the application (given most in C) and optimally partitions/allocates the tasks spatially and temporally onto MPSoC platforms [2]. The extraction of parallelism and optimization techniques of partitioning/allocation/scheduling are key challenges as well as the underlying hardware architectural considerations like communications networks.
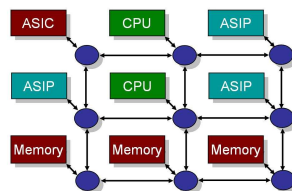

Figure 1: An example of MPSoC platform

References

[1]   Isshiki, T.; Urfianto, M. Z.; Khan, A. U.; Li, D. & Kunieda, H.: Tightly-Coupled-Thread Model: A New Design Framework for Multiprocessor System-on-Chips. IP-SOC 2006, Design & Reuse, S.A., 2006

 [2]   Kempf, T., Dörper, M., Leupers, R., Ascheid, G., H. Meyr, Kogel, T. and B. Vanthournout: A Modular Simulation Framework for Spatial and Temporal Task Mapping onto Multi-Processor SoC Platforms. In Proceedings of the Conference on Design, Automation & Test in Europe (DATE), Munich, Germany, March 2005

Formale Methoden für Signalisierungssysteme

Stephan Swiderski

swiderski@informatik.rwth-aachen.de

Das Ziel dieses Projekts ist die Entwicklung eines Signalisierungssystems für Eisenbahnsysteme. Die Hardware basiert auf handelsüblichen Micro-Controllern. Für diese Micro-Controller soll durch neue Synthese-Techniken automatisch Assembler-Code synthetisiert werden, wobei diese Synthese durch eine formale Spezifikation des Signalisierungssystems gesteuert werden soll. Wir werden das neue System mit Hilfe von probabilistischem Model-Checking mit aktuellen Systemen in den Bereichen Zuverlässigkeit und Durchsatzrate vergleichen. Zusätzlich werden Techniken der Verifikation und der Terminierungsanalyse verwendet, da der Assembler-Code einige gesetztliche Bestimmungen für den Betrieb von Eisenbahnanlagen erfüllen muss. Als Basis für diese Techniken werden schon bekannte Verfahren aus dem Bereich der Termersetzungssysteme verwendet, welche entsprechend adaptiert werden müssen. Außerdem sollen diese Techniken lesbare formale Beweise liefern, welche den zuständigen Behörden vorgelegt werden, um eine Betriebserlaubnis für das Gesamtsystem zu erhalten.

Algorithmic Theory of Multiplayer Games

Michael Ummels

Mathematische Grundlagen der Informatik, RWTH Aachen, Germany

`ummels@informatik.rwth-aachen.de`

Infinite games have been successfully applied in the verification and synthesis of reactive systems. Such a system is usually modelled as a game between the system and its environment where the environment's objective is the complement of the system's objective, so the environment is considered hostile. Therefore, traditionally, the research in this area has mostly looked at two-player games where each play is won by precisely one of the two players, so-called *two-player zero-sum games*. However, motivated by the modelling of distributed systems, interest in the general case has increased in recent years.

Different *solution concepts* have been proposed to model rational behaviour in games. The classical solution concept offered by game theory is the one of a *Nash equilibrium*. In a Nash equilibrium no player can receive a better payoff by unilaterally changing her strategy. For games with stochastic transitions and players taking turns, Chatterjee et al. [1] have shown that Nash equilibria always exist for most winning conditions. However, if players are allowed to choose actions simultaneously, the existence of Nash equilibria has only been shown for relatively simple winning conditions like reachability.

A problem with Nash equilibria is that they ignore the sequential structure of an infinite game where players can change their behaviour at any point of time. This allows for Nash equilibria with *incredible threats*. A refinement of the Nash equilibrium concept that deals with this problem is the notion of a *subgame perfect equilibrium (SPE)*. For games with players taking turns and no stochastic transitions, we [2] have shown that SPEs always exists for most winning conditions, and we are currently working on extending this result to games with stochastic transitions.

On the algorithmic side, there are mainly two problems we consider: Computing some equilibrium and one that meets certain requirements on the payoff. Whereas an algorithm for the former problem can, in general, be deduced from the existence proof, to find an algorithm for the latter problem is more challenging. For instance, regarding SPEs, we [2] made use of tree automata to give such an algorithm.

References

[1] K. Chatterjee, M. Jurdziński and R. Majumdar. On Nash equilibria in stochastic games. In *Proc. CSL '04*, vol. 3210 of *LNCS*, pages 26–40. Springer-Verlag, 2004.

[2] M. Ummels. Rational behaviour and strategy construction in infinite multiplayer games. In *Proc. FSTTCS '06*, vol. 4337 of *LNCS*, pages 212-223. Springer-Verlag, 2006.

Konstruktion zeitkontinuierlicher Markov-Entscheidungsprozesse

Daniel Willems

daniel.willems@cs.rwth-aachen.de

Traditionelle Verifikationsverfahren sind fokussiert auf absolute Korrektheit, wobei in der Praxis solche strengen Konzepte kaum, manchmal auch überhaupt nicht garantiert werden können. Phänomene stochastischer Art wie Fehler, unbekannte Kontexte, und zufällige Verzögerungen spielen jedoch häufig eine prominente Rolle, weswegen Korrektheit nicht immer absolut gefordert werden kann. Stochastisches Model-Checking ist eine Technik, die dazu geeignet ist, statt qualitativer Aussagen wie "das System gerät nie in einen Deadlock", quantitative Eigenschaften wie "die Wahrscheinlichkeit, dass das System innerhalb der nächsten 12 Stunden in einen Deadlock gerät, ist niedriger als 0.0001" automatisch zu überprüfen. Die derzeit verfügbaren Software-Tools und Algorithmen konzentrieren sich auf diskret- oder zeit-kontinuierliche Prozesse. Für eingebettete Systeme aber sind Einhaltung von Zeitschranken und Interaktion mit der (unzuverlässigen) Umgebung wesentlich. Stochastische Modelle für solche Systeme sollten deswegen sowohl zeit-kontinuierlich sein als auch offen, d.h. das Verhalten der Umgebung soll nicht vollständig spezifiziert sein. Nichtdeterminismus ist die natürliche Modellierungstechnik dafür. Im Rahmen dieses Themas werden zeit-kontinuierliche Markov Entscheidungsprozesse (CTMDPs) im Vordergrund stehen. Diese Prozesse haben einen breiten Anwendungsbereich wie z.B. stochastic scheduling und dynamisches Energieverbrauchsmanagement.

Auch für Abstraktion von zeit-kontinuierlichen Markov Ketten sind CTMDPs von Bedeutung. Eine auf probabilistischer Simulation basierende Abstraktionstechnik mit der radikale Reduktionen des Zustandsraums einer Markov-Kette möglich sind - dafür aber auch unbestimmte Ergebnisse beim Model Checking liefern können - kann auf CTMDPs abgebildet werden. Falls das Model Checking Ergebnis unbestimmt sein sollte bedeutet dies, dass durch die Abstraktion zu viele Informationen (durch Nichtdeterminismus) verloren gegangen sind. In diesem Fall bedarf es einer Analyse des abstrakten nicht-deterministischen Modells um eine geeignete Verfeinerung der Abstraktion zu bestimmen. Dieses Themengebiet wird insbesondere bearbeitet werden.

GRK 1324:

Metrik

Modellbasierte Entwicklung von Technologien für selbstorganisierende dezentrale Informationssysteme im Katastrophenmanagement

Berlin

Model-Based Development of Self-Organizing Information-Systems for Disaster Management

GK METRIK: Humboldt-Universität zu Berlin, GFZ Potsdam, Fraunhofer FIRST, HPI Potsdam, ZIB Berlin

Each year natural disasters cause a loss of life and property all over the world. This is why the United Nations designated the 1990's as the International Decade for Natural Disaster Reduction (IDNDR) which became the International Strategy for Disaster Reduction (ISDR). The goal is to use scientific and engineering know-how to reduce disasters within Disaster Management (DM) measures.

DM encompasses all measures taken to reduce the damages caused by disasters. These measures take place in phases: collecting knowledge to prepare for a disaster, establishing prophylactic, preparative and preventive measures based on this knowledge, response-activities during an event are followed by reconstruction measures. Information and communication technology (ICT) can greatly help in collecting, preparing and disseminating information as well establishing reliable communication. Enhancing the overall process of disaster management is the goal of our research in METRIK.

We focus on self-organizing wireless networks (SWN) within ICT. Their comparably small costs and short deployment times make them a good candidate to collect data through attached sensors and provide reliable, dedicated communication and information services in the disaster area. METRIK investigates the design of self-organizing protocols and applications that consider the special properties of the technology. A core application in this context are self-organizing information systems (SIS) to gather, store, prepare and disseminate the information as they are needed.

Research topics include routing, replication, effective gathering and processing of decentralized data, dependable services, automated deployment and update of software components at runtime, the use of workflow technology for supporting processes and workload balancing among terminal devices with limited resources.

The development of SIS with SWN faces a number of challenges: The system relies on domain specific knowledge from different fields of geo sciences, government agencies, and computer sciences. It must be correct and robust. Ongoing research in SWN and information systems suggests an integrated development spanning all technological levels to satisfy resource constraints. We therefore research the development and use of domain specific languages for modeling different aspects of the system by experts. Complementing with model-coupling, verification and testing techniques, METRIK aims on providing a development methodology for disaster management systems.

Modellbasierte Entwicklung selbst-organisierender Informationssysteme für das Katastrophenmanagement

GK METRIK: Humboldt-Universität zu Berlin, GFZ Potsdam,
Fraunhofer FIRST, HPI Potsdam, ZIB Berlin

Jedes Jahr sind weltweit Naturkatastrophen Ursache für Zerstörung und den Verlust von Menschenleben. Deshalb hat sich aus der „International Decade for Natural Disaster Reduction" (IDNDR), 1990-1999, der Vereinten Nationen die „International Strategy for Disaster Reduction" (ISDR) gebildet. Ihr Ziel ist es Katastrophen mit Hilfe wissenschaftlicher Methoden und Ingenieurswissen im Katastrophen-Management (KM) zu verringern.

Das KM umfasst alle Maßnahmen, die geeignet sind Schäden durch Katastrophen zu verringern. Dazu gehören: Wissenerhebung zur Vorbereitung auf eine Katastrophe, vorbeugende, vorbereitende und präventive Maßnahmen auf Basis dieses Wissens, reagierende Maßnahmen während eines Ereignisses und der Wiederaufbau im Anschluss. Informations- und Kommunikationstechnologie (IKT) kann helfen, Informationen zu sammeln, aufzubereiten und zu verteilen sowie verlässliche Kommunikationsdienste anzubieten. Die Forschung im GK METRIK unterstützt den gesamten Prozess des KM.

Im Bereich der IKT erforschen wir besonders selbst-organisierende drahtlose Netzwerke (SDN). Wegen ihrer vergleichsweise geringen Kosten und kurzen Aufbauzeiten sind sie geeignet im Katastrophengebiet über angeschlossene Sensoren Daten zu sammeln sowie dort Kommunikations- und Informationsdienste anzubieten. METRIK erforscht den Entwurf von Selbst-Organisations-Protokollen und Anwendungen, die die speziellen Eigenschaften der Technologie berücksichtigen. Eine Kernanwendung sind selbst-organisierende Informationssysteme (SIS) die Informationen sammeln, speichern, aufbereiten und dort bereitstellen, wo sie benötigt werden.

Forschungsthemen sind u.a. Routing, Replikation und effizientes Sammeln und Verarbeiten von Daten, verlässliche Dienste, automatisiertes Verteilen und Aktualisieren von Software-Komponenten zur Laufzeit, der Einsatz von Workflow-Technologie zur Unterstützung der Prozesse sowie Lastverteilung auf Endgeräten mit begrenzten Ressourcen.

Die Entwicklung von SIS und SDN für das KM ist anspruchsvoll: Ein effektives System beruht auf spezifischem Wissen aus den Geowissenschaften, Behörden und der Informatik. Es muss korrekt und robust arbeiten. Aktuelle Forschung für SDN und SIS legt einen integrierten Entwicklungsprozess über alle technologischen Stufen hinweg nahe, um Ressourcen-Anforderungen zu erfüllen. Wir erforschen daher die Entwicklung und den Einsatz domänenspezifischer Sprachen zur Modellierung der verschiedenen Aspekte des Systems durch die jeweiligen Experten. Durch Ergänzung mit Modellintegrations-, Verifikations- und Testverfahren wollen wir eine Entwurfsmethode für Katastrophen-Management-Systeme bereitstellen.

Dependable, Service-Based Infrastructure for Disaster Management

Stefan Brüning
GK METRIK
bruening@gk-metrik.de

Computer systems for use in disaster management are based on very heterogeneous infrastructures. Powerful servers are the backbone of those systems which are used for workflow management, task force coordination and communication. These servers process data generated by humans and sensors. Sensor data can come from a variety of different sensor types mounted to individual sensor nodes. The complexity of these sensors ranges from sophisticated physical systems to simple nodes as used in sensor networks that have only very limited resources.

All parts of the system have to work effectively and efficiently. A special focus lies on dependability. A break down of some parts must not lead to a failure of the whole system but instead must be compensated by the remaining parts. The system has to self-adapt to the changed infrastructure and still allow successful management of the disaster.

Service-oriented architectures (SOAs) have been established as a standard for distributed systems in recent years. The possibility to dynamically discover and bind to new services allows a loosely coupled and adaptive infrastructure. The dependable integration of powerful servers and resource limited sensor nodes into one disaster management system is a challenging task.

Servers for use in workflow systems can be embedded in a service- based architecture based on web services. Several approaches for increasing dependability for web services exist. Nevertheless, there are no exact measurements yet whether those increases in dependability can really be reached.

In my dissertation I will examine to which extend certain technologies and methods are increasing dependability. For this purpose, a SOA will be deployed in our laboratory which will consist of several test services on different computers. Artificial failures will be simulated using fault injection technologies. The results will be evaluated based on metrics for service dependability. Additionally a fault taxonomy will be developed as a basis for fault injection.

Modeling and Verifying Declarative Workflows

Dirk Fahland
GK METRIK
fahland@gk-metrik.de

The support of disaster management by self-organizing information systems and networks cannot succeed without support of the intermeshed *work procedures* emerging in this context: Administrative processes that are carried out when disaster strikes and that cover the overall situation need to fit to emergency and rescue procedures for locally limited actions. Both will interfere with procedures and behavior of the supporting self-organizing information system. A failure-free and robust concurrence, that is *correctness* of the involved procedures, is prerequisite for successful disaster management.

Workflows have been established as a mathematically founded, operational model for work procedures in different variants. Thus questions on the correctness of one or several joint processes can be formulated and answered on a principle basis. As any model, established workflow models have been built under *assumptions* like availability of resource, communication infrastructure and continuity of cooperating partners. The dynamics of disaster management and self-organizing systems violates these assumptions: Workflows for disaster management need to be *flexible* to allow the execution of a process in various circumstances. Unpredictable events require the need for (runtime) *adaption* of a workflow. At the same time, a flexible and adaptive workflow has to be correct in order to be a reliable tool in disaster management. A property of interest will be the *self-stabilization* of a workflow that encountered a faulty situation.

We identified two ideas that may effectively allow the modeling of flexible and adaptive workflows such that these models can be analyzed for properties of interest. We step back from classical operational models like *Petri nets* and introduce declarative and scenario-based elements: *Temporal logics*, for instance linear-time temporal logic, allow to precisely model critical aspects like communication or resource access while giving only a loose, but sound, characterization of ordering tasks of workflow. A scenario-based approach as exercised in Harels *Life-Sequence Charts* breaks behaviorally complex models into human-conceivable parts that can be modified as the (changing) situation requires.

This approach requires the definition of a feasible (declarative and/or scenario-based) *modeling language* for workflows that meets the ideas and concepts of processes in disaster management. Having models in such a language, the models have to be made effective, for instance by translating them into an *operational model* (and vice versa). Finally, we are interested in a (temporal-)logic-based characterization of relevant *properties* like self-stabilization and feasible *verification procedures* to verify the correctness of a workflow.

Data Management for Wireless Sensor Networks

Timo Mika Gler
timo.glaesser@gk-metrik.de

Graduierten Kolleg METRIK

The Graduierten Kolleg METRIK focuses on self-organizing systems for use in disaster management before, during and after natural disasters like earthquakes, tsunamies or volcanic eruptions. We will try to address the various problems that arise with a) data gathering and b) event-detection using wireless sensor networks in the above mentioned scenarios.

In fact wireless sensor networks play a central role in the envisioned systems for disaster management. Wireless sensor networks consist of small, independent sensor devices with highly incomplete and local knowledge that can measure physical properties of their environment and communicate with other sensor devices in their area using broadcast wireless communication. These sensor devices are equipped with small batteries as their only power-source. However each sampling operation, each computation and communication operation on the sensor devices consumes energy. Hence respecting the energy constraints of these wireless sensor devices is critical to keep them operational and optimizing for lower energy consumption may increase device and network lifetime.

We will try to provide a declarative publish/subscribe-style interface to the sensor network that enables users to specify queries like "Report all locations of sensors in the area around Istanbul where the shaking exceeds a threshold τ." or "Report all clusters of nodes and the average shaking in the clusters in the Istanbul area where at least 80% of all nodes reported an earthquake of magnitude greater than τ."

When a user hands such a query to our system, the query will automatically be decomposed into smaller operations arranged in a dataflow graph and be spread across the wireless sensor network. The operations in the dataflow graph can be executed in different orders yielding different energy usages at both the individual sensor devices and the network level. Our system will continuously try to optimize the cost of multiple such queries to extend the wireless sensor network's lifetime. Additionally as routing is of utmost importance in multi-hop wireless networks we provide novel ideas of including routing decisions in the evaluation and optimization process of queries.

Prior research in the area does either not perform in network query processing - all data is routed to a gateway node and processed there - or considers single-queries only. Additionally our approach differ from others by not only allowing simple aggregations like MIN, MAX, AVG but also supporting the clustering of nodes based on the network topology, e.g. hop-distance.

QoS in 802.11 Multihop Networks

Kai Köhne
GK METRIK
koehne@gk-metrik.de

Wireless local area networks using the IEEE 802.11 standard are popular both in home and office use. Most often, the technology is used for the last hop, letting the user access the (wired) core network via an access point. However, projects such as MIT Roofnet and Berlin Roofnet demonstrate that even a larger metropolitan network can rely solely on wireless links. In these networks, every participating node does not only process its own data, but also relays traffic for neighbouring nodes, so that every node in the network can be reached by multi-hop connections. These techniques have also been studied and examined for sensor networks. In the EDIM and SAFER projects we create an earthquake early warning system based on 802.11 multi-hop technologies.

Implementing demanding real-time applications such as earthquake early warning based on wireless multi-hop networks require high network reliability. However, each link in a 802.11 network is highly volatile, suffering from physical phenomena like fading, but also from interferences with other devices. Moreover, in cities often 10 or more 802.11 networks operate in parallel, effectively limiting the available bandwidth for each of them.

One way to achieve a higher reliability for time critical applications is service differentiation. The recent addition 802.11e provides new medium access methods that share the medium either in a fully distributed (EDCA) or centralized (HCCA) way. EDCA is of special importance for multi-hop networks: High priority packets have a higher chance to access the medium in the case of conflicts.

In my work I evaluate the advantages and problems of using EDCA in wireless multi-hop networks. One problem largely ignored in the literature is spectrum competition between different 802.11 networks. Networks are able to optimize their EDCA parameters so that they can gain a higher share of the bandwidth. If every network does use the most aggressive settings, the differentiation mechanism is effectively undermined. We therefore develop a new differentiation protocol which builds upon EDCA. In these protocol, every data message has the same EDCA priority. Service differentiation is achieved using a separate control channel. This avoids the mangling of intra- and inter-network priorities, and resolves unfairness issues due to interference between data and control packets.

Using metaprogramming and metamodeling for prototyping middleware abstractions in wireless sensor networks

Daniel A. Sadilek
GK METRIK
sadilek@gk-metrik.de

In METRIK, wireless, self-organizing sensor networks (WSNs) are supposed to be the basis of complex disaster management systems. Software development for such systems is difficult because of numerous reasons: applications on WSNs are distributed and show the usual synchronization and communication problems; they have to respect the limited resources of the sensor nodes; deployment and debugging is very costly. Middleware could help with these problems but existing middleware technologies are inappropriate for WSNs—and it is not yet clear which services, which programming model and which abstraction level a middleware technology for WSNs should offer. Furthermore, for WSNs the term "middleware" is used in a much broader meaning than traditionally: there are operating systems, languages and pure libraries all named "middleware".

I am looking at middleware from the perspective of domain specific languages (DSLs), which may enable domain experts to actively develop parts of a WSN application. Because services, programming model and abstraction level are not yet clear, I am working on technologies that allow the prototyping and simulation of a DSL's domain concepts and its semantics. The definition and usage of DSLs is routine since decades in metaprogrammable languages like Lisp and Smalltalk. However, these languages lack support for purpose-built concrete syntax. On the other hand, technologies from the field of model-driven software engineering (MDSE) provide such support. However, they miss a thing that metaprogrammable languages provide: direct executability—because metamodels alone don't have operational semantics. The semantics is usually provided by a transformation, in which the knowledge how to map domain specific concepts to concepts of the target platform is encoded.

My approach is to use flexible metaprogrammed domain abstractions in which fixed object-oriented metamodeling layers are introduced. I hope to gain synergies from this combination that allow a great part of the language semantics done for the simulation to be reused for the target platform. Currently, I am doing a proof of concept implementation based on the Eclipse modeling technologies and Scheme. I am working on a development process specification, the simulation of multiple nodes, support for communication primitives, configuration of runtime parameters, support for mixed graphical/textual DSLs, and compilation for the target platform. As first evaluation project, I am developing a stream-based language, which can be used for the development of an earthquake early warning system with algorithms for earthquake detection, distributed warning determination, and warning dissemination.

Language Modelling

Markus Scheidgen
GK METRIK
scheidge@gk-metrik.de

Computer languages are in the center of software engineering; they are the tools to communicate problems and solutions to both humans and computers. Computer languages have to allow unambiguous and effective expressions, independent of the specific language nature, its level of abstraction, purpose, or area of application. Language sentences have to be proofed and tested. The artefacts written in computer languages have meaning; these semantics is manifested in compilers, simulators, model transformations, or employing calculi and formalisms. In this sense, a computer language is much more than a set of sentences; it is more a set of tools that provide for all these aspects.

Language modelling is based on the hypothesis that languages are pieces of software and that they should be developed like software in a model driven fashion. We use models to describe languages, their concepts, notations, and semantics. Language models, known as meta-models, are artefacts written in several meta-languages. Each of these languages can be used to describe a single language aspect. These aspects include, for example, abstract language structure, static semantics rules, textual or graphical notations, operational semantics or code generation. The semantics of a meta-language is manifested in a generic tool which uses a language description written in that meta-language. The ultimate goal is to provide a complete framework of meta-languages that allow to describe a language in all necessary aspects. Whereby, these descriptions can be used by humans to understand the language and by machines to provide automated tool support for the language.

We contribute several meta-languages and generic tool support for the modelling and use of computer languages. We developed the meta-modelling framework *A MOF 2 for Java* based on the MOF 2.0 standard. This framework, in contrast to others, supports enhanced refinement and specialisation features that allow for more flexible and reusable structure models. It also uses a new Java mapping which utilises *generics* and *variant* return types for safer programming with models. We developed the Textual Editing Framework, which provides a meta-language for textual notation modelling. The framework includes an generic editor that provides semantic rich editing with syntax highlighting, code completion, error annotations, and much more. A third contribution provides meta-language and generic tool support for execution semantics. Based on structure definitions for abstract language syntax and runtime data, the language engineer can specify behaviour at the meta-level. The generic model simulator can then execute language instances solely based on such descriptions. We are evaluating on the Specification and Description Language (SDL). This proofs the general applicability (and also the weaknesses) of our contributions and is our basis to reason about the hypothesis: *languages, as software, can be modelled*.

Model Driven Engineering (MDE) for Modeling and Simulation in Disaster Management

Falko Theisselmann
GK METRIK
theissel@informatik.hu-berlin.de

Scientific models and simulations of environmental systems are major information sources in disaster management. Various frameworks and domain specific languages have been developed to ease modeling and simulation tasks and provide access to new modeling and simulation technologies to non-experts.

These frameworks usually stick to certain modeling formalisms (i.e. box models, state machines, cellular automata etc.) and modeling languages. As a result, models are usually framework and language specific. Using frameworks requires knowledge about frameworks, their concepts, languages, and capabilities. Changing the framework usually requires expensive, manual reimplementation of models. In environmental modeling practice, this is an obstacle for the use of frameworks, the reuse of models, and their integration in coupled multi-model models.

We investigate the use of MDE to tackle this problem. Our approach is conceptually based on the Object Management Group's Model Driven Architecture (MDA). Usual modeling formalisms in the disaster management domain are platforms in the sense of MDA. Domain specific models (DSMs) are mapped onto formalism specific models (FMSMs). In the following steps, these are mapped onto framework specific models (FWSM) that allow for the generation of framework specific executable code.

One DSM maybe transformed into several FMSM. Note that one FSMS maybe transformed into other FMSMs and into several FWSMs. This yields the possibility for the extensive reuse of models, if the formalisms allow for that. Also, several DSMs maybe translated into a single FMSM, which facilitates model integration/coupling.

To formulate DSMs, we use DSLs, where the concepts and notational elements are tailored to the concepts and cognitive spaces of the domain experts. The logic defined by the transformations ensures the correct application of formalisms and concepts which otherwise would be the responsibility of the domain expert.

The goal of our research is to evaluate and show how MDE may enhance model reuse, integration and combination of different domain models and promote the sensible use of frameworks. Moreover, we want to investigate how to represent domain knowledge adequately and support collaborative, multidisciplinary modeling in the field of disaster management.

Graduiertenkolleg 1324 – METRIK
Metamodel-based Language Engineering

Guido Wachsmuth

guwac@gk-metrik.de

In METRIK, we follow a model-based approach to develop disaster management systems. We want to provide domain experts with modelling means that are intuitive, concise, and semantically precise. To fullfill these requirements, we propose a combination of several domain-specific languages (DSLs). Language Engineering brings Software Engineering to languages. It is concerned with language design, language maintenance, language extension, language recovery, translation, generation, interpretation, etc. Metamodels are our choosen tool to specify languages. Like other software artefacts, metamodels and metamodel-based language descriptions evolve over time. Metamodel evolution is usually performed manually by stepwise adaptation. Furthermore, models need to co-evolve in order to remain compliant with the metamodel. Without co-evolution, these artefacts become invalid. Like metamodel evolution, co-evolution is typically performed manually. This error-prone task leads to inconsistencies between the metamodel and related artefacts.

In our work, we explore transformational metamodel adaptation as a foundation for *Metamodel-based Language Engineering*. As a first result, we provide a theoretical basis to study the effects of metamodel evolution in terms of metamodel relations. We employ well-defined evolutionary steps for metamodels compliant to OMG's Meta Object Facility (MOF). The steps are specified as transformations in QVT Relations, the relational part of OMG's Query-View-Transformation language. Each step forms a *metamodel adaptation* and is classified according to its semantics- and instance-preservation properties. Automatic co-evolution steps are deduced from these well-defined evolution steps. This *co-adaptation* prevents inconsistencies and metamodel erosion. Starting from our theoretical results, we develop an *Adaptation Browser* for MOF compliant metamodels. The tool is built upon the Eclipse Language Toolkit. This provides us with undo/redo support, adaptation history, and scripting facilities.

We now examine our approach in two case studies. The first study is concerned with the design of a new DSL. The transformational approach facilitates a well-defined stepwise metamodel design. Starting from basic features, new features are introduced by construction. We hope that extensive usage of this principle leads to an agile process. The second study is concerned with the extraction of a language out of a natural language document. The document describes disaster management processes semi-formally. Language recovery is concerned with the derivation of a formal language specification from such sources. For grammar recovery, a transformational approach already proved to be valuable. In a similar way, we employ our approach to assist metamodel recovery.

Model Coupling and Test

Stephan Weißleder
GK METRIK
stephan.weissleder@gmail.com

Models are a fundamental element of model-driven software development (MDSD). By using model transformations and code generation patterns, available tools are already able to generate a large part of the source code automatically. However, while effort has been invested in source code generation, the automatic generation of test code has been widely neglected. This test code should reveal errors, which are occurances from abnormal program terminations to small deviations from the specified behavior. Normally, the late detection of such errors results in high costs for their removal and for the compensation of their consequences. So, early and extensive testing is most important. Nevertheless, the priority of testing is often still considered lower than the priority of delivering products early, which is mainly due to high costs for test suite development.
All in all, models and tests both aim at improving the quality of the developed system. Our current work combines methods from these two fields.

The intention of this combination is to gain extra benefit for software testing by using different models together. Therefore, we use several models describing different aspects of a system, combine them, and derive enough information for automatic test case generation. Several former approaches only dealt with single models or with models that used just a very small aspect of another model. Likewise, constructive but random approaches for automatic test input data generation have been neglected.

Our current state of affairs includes the automatic and model-based generation of a whole test suite based on state machines and class diagrams of UML. We combine both models via their OCL constraints and construct partitions for the corresponding test input values. Furthermore, we operate boundary testing on these partitions to generate test inputs for positive and negative test cases. Additionally, our current work combines structural and behavioral models to benefit from their particular relationships. For instance, behavioral specifications (e.g. state machines) can be reused in several static contexts (e.g. classes) via structural relationships like inheritance. We already implemented our main ideas in *SMOTEG* - a tool based on the Eclipse modeling framework.

The identified challenges of our current approach are the following: the identification of retraceable constraints so that the influence of input parameters becomes clear; the combination of several behavioral specifications for one class (multiple inheritance); and the introduction of concurrency, which is important for distributed systems like sensor nets. In the context of METRIK, we aim at adapting our approach to special tasks of geo information systems (GIS) or meta-modeling. Furthermore, we need to identify significant case studies to show our approach's benefit.

GRK 1362:

Mixed Mode Environments

Cooperative, Adaptive and Responsive Monitoring in Mixed Mode Environments

Darmstadt

Graduiertenkolleg 1362
Mixed Mode Environments

Darmstadt

This Research Training Group which is funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) combines two exciting and challenging research areas:

- navigation and coordination of multiple autonomous vehicles to perform a common task possibly together with a human mission manager; and

- monitoring in mixed mode environments that are characterized by the heterogeneity of their components in terms of resources, capabilities, and connectivity.

We apply these technologies in the domain of disaster recovery, for example, after an earthquake or an explosion in a chemical plant where the environment might not be safe for human beings.

We consider heterogeneous sensors that may range from low end temperature sensors to high end video sensors, and which may be stationary or mobile. The autonomous vehicles are for once carrier platforms for the sensors but they also exploit sensor information for their own navigation and coordination.

We identify many exciting research problems, ranging from recognition of human shapes and dynamic terrain mapping after a disaster, to the fundamental problems of how to provide adequate quality of service for communications and reliability in the middleware when the underlying platform is inherently unstable and mobile, and the coordination of multiple autonomous vehicles combining to perform a common task.

This Research Training Group addresses significant problems that will yield cutting edge research topics for 9 years. We propose an innovative supervision concept that will provide both international exposure and goal oriented research.

The Research Training Group has an additional distinguishing feature which is the cooperation with the Virginia Polytechnic Institute and State University (Virginia Tech). The researchers at TU Darmstadt will benefit from this cooperation through the access to Virginia Tech's leading competence in unmanned ground, aerial and marine vehicles; the researchers at Virginia Tech will benefit from TU Darmstadt's expertise in communications and middleware in highly unstructured environments, and its experience with robot coordination strategies and shape recognition. This results in a win-win situation from which both institutions benefit.

Agent architecture for a mixed mode network

Arthur Herzog

Database and Distributed Systems Group
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt
aherzog@dvs1.informatik.tu-darmstadt.de

Abstract: In my research I want to apply multi agents to monitoring in mixed mode environments.

Motivation for multi agent aproach

In a mixed mode environment we must deal with different hardware like small sensors, unmanned vehicles, mission control stations, etc. All those devices have their specific capabilities and constraints and a wide range of available resources. We want these different nodes to cooperate to achieve some goal. This should still work if part of the nodes fail. This means that the nodes shall be able to adapt to the changes in the environment and reorganise themselves. These are only some of the qualities that correspond to the qualities of multi agent systems. Each of the nodes in our mixed mode environment can be seen as an agent who knows its capabilities and constraints. Depending on his capabilities an agent can have tasks he tries to fulfil or can offer services to other agents to help them fulfil their tasks. The complexity of agents running for example on a small sensor node and on an unmanned vehicle (UV) can vary considerably. So a sensor-agent might be only capable to measure the current temperature and to send it to someone who is interested in this data, whereas the UV-agent can move through the environment, collect the data form the sensor-agents, aggregate and evaluate the collected data and use the gathered information further in his decision making process.

To make this possible the different kinds of agents must be able to communicate with each other independent of hardware or software they use. This means we need a platform- independent communication language. The language must also reflect the various interaction patterns that are typical for event-based systems. In the area of multi agent research there exist numerous solutions, but most of those are rather heavy weighted and not applicable for the small, resource-constrained sensor nodes.

Wireless Sensor Networks in motion

Daniel Jacobi

GK Cooperative, Adaptive and Responsive Monitoring
in Mixed Mode Environments
Technische Universität Darmstadt
Hochschulstraße 1
64289 Darmstadt
jacobi@dvs1.informatik.tu-darmstadt.de

Abstract: In short my research interests are mobility in middleware for wireless sensor networks (WSN) and topology control. My idea at the moment is a middleware that can respond to events in a way, that the network itself changes its topology to get the best performance.

1 Middleware in WSN's

Wireless sensor networks are expected to have a huge size of several hundred or thousand of nodes after deployed in the field. Every node is equiped with some sensors that apply to the specific task they have to solve and they are all running on batteries.

Middleware in wireless sensor networks has to cope with much more constrains than the 'big brothers' in a regular network. Security and quality of service are problems that both systems in a real environment have. In wireless sensor networks and especially in the scenario of our GK heterogenity and dynamic network organization are big issues. WSN have additional issues like limited resources, scalability (in the sense of number of nodes) or data aggregation.

2 Controlling the network

On the one hand introducing mobility in WSN's enhances the complexety for routing algorithms and the maintenance of neighborhood information. On the other hand a network that can control the mobility of it's nodes can enable many interesting abilities. E.g. a self-healing network, that can alter its network topology to regain coverage after some nodes ran out of energy, or the inspection of events, that occure on the edge of the actual network.

A drawback of this approach is the raised energy consumption. Every movement costs energy and in this sense lifetime of the network. But with changing the topology the network can regain a nearly fully operational state.
What abilities are possible and which of them are feasible for a WSN, are some of the next questions. Also how to integrate these abilities in a middleware and what constrains the middleware has to fullfill is not answered, yet.

Alternative Communication Techniques for Mobile Platforms

Kalina Karova

kkarova@tu-darmstadt.de

Darmstadt University of Technology
Institute for Automatic Control
Control Theory and Robotics Lab
Landgraf-Georg-Str. 4
D-64283 Darmstadt

Communication is the key point for the cooperative, adaptive and responsive behavior of robot teams in a mixed mode environment. The traditional way of information exchange between robots is the global radio communication, which might be not available depending on the scenario and the environmental condition. Thus, in such cases the centralized coordination of the robot swarm is not possible any longer and an autonomous robot behavior and alternative local communication methods are required. The research and the development of alternative techniques can provide a solution in those cases. Furthermore a novel method can support the existing traditional ones and can increase the QoS and the robustness of the information exchange between mobile platforms.

A first step in searching for this kind of techniques can be to investigate biological systems concerning communication and their social coordination between individuals. Biologic systems like bats, wales, bees, ants and mots provide valuable examples for alternative information exchange. Bees combine for example local communication at centralized places and global pheromone communication in order to distribute information. The concept of stigmergy is also used by ants to solve collectively highly complex tasks.

This research work concentrates on these mechanisms inspired by biological systems. A main focus will be on stigmergy, a mechanism well known in biology where individual parts communicate with one another by modifying their local environment. It includes scent communication but also other methods like human road signing. The proposed novel communication techniques enables the exploration, navigation and action planing in an unknown dynamic environment and can be applied in a large variety of possible scenarios. For example it allows an effective signing of an area and local communication between mobile heterogeneous robots, able to perceive and act specific to their set of sensors and actuators. In this sense it can coordinate the exploration and manipulation of large unknown areas without the need for global communication.

To verify the developed methods extensive simulations of different scenarios will finalize this work and demonstrate its applicability.

Dynamic Formation Control of Multi-Robot-Systems

Kim Listmann

klistman@rtr.tu-darmstadt.de

Darmstadt University of Technology
Institute for Automatic Control
Control Theory and Robotics Lab
Landgraf-Georg-Str. 4
D-64283 Darmstadt

Controlling a formation of mobile sensor nodes could be useful for many applications. The topology of a network has a decent effect on the reliability and the reachability of every node regarding the communicative needs. Moreover by changing its topology energy-efficiency and spatial distribution of the network can be optimized. Regarding the nodes as mobile robots, either ground, air or underwater vehicles, formations can be used to traverse known parts of an unknown environment quickly or to collect huge amounts of data from unknown environments at a time.

One problem regarding the nodes as mobile robots is the requirement of a mathematical model of the group dynamics respecting the specific motion and sensing constraints of every vehicle. As vehicles posses nonlinear dynamics, the group dynamics will be nonlinear as well. Hence, effective control methods must be used to apply a formation control on a group of robots having only limited computation capacity. To this end, methods from graph theory and control theory will be combined to obtain a feasible result. Firstly, a consensus must be found to establish a formation and it has to be clarified if this formation will actually be stable. To build up a formation the multi-robot system can be modeled as a graph and formation control means switching the topology of this graph. Once the formation is set up, the group dynamics can be described as a system evolving on Lie groups and therefore, very efficient control algorithms can be derived.

A further step will be to evolve a motion planning algorithm for the formation as a whole applying a virtual leader approach. By this and the applied control scheme formation tracking will be facilitated. Meaning that formations may evolve dynamically, so that formation merging or splitting will be possible depending on the environment. This will lead to a reactive formation control where obstacles/collisions will be avoided and exploration schemes may be enhanced.

Finally, these ideas will be applied to existing robots and tested in a suitable scenario.

Combining Two Extremes of Replication

Matthias Majuntke

majuntke@informatik.tu-darmstadt.de

To increase the dependability of a system replication is the most fundamental principle. In the literature there are basically two extreme approaches concerning replication. On the one hand there are quorum systems that provide the abstraction of an *atomic read/write register* [Lam86], on the other hand there are distributed systems implementing a *replicated state machine* [Sch90].

Distributed storage systems may be efficiently handled by quorum systems. Clients directly write to or read from a quorum of servers. As two quorum sets intersect, consistency is guaranteed without inter-replica communication. A drawback of quorum systems is that they are not suitable for requests with stronger semantics than read/write — e.g. for general requests necessary to implement a replicated state machine. Abd-El-Malek *et al.* [AEMGG$^+$05] developed a state machine abstraction based on quorum systems but at the cost of $5t + 1$ instead of $3t + 1$ replicas to tolerate t Byzantine processes.

To guarantee consistency among the replicas of a replicated state machine all requests have to be processes in exactly the same order by all replicas. This property is provided by an *atomic broadcast* primitive [Lam98]. A replicated state machine may implement any service but at the cost of an expensive inter-replica communication.

In general, requests can be classified into three types:

- *read.* requests that do not change the current state of the server

- *write.* requests that change the state of the server, but the state change is not dependent on the current state.

- *modify.* requests that change the state of the server dependent on the server's current state

Quorum systems can process only *read* and *write* requests efficiently; for *modify* operations a replicated state machine is necessary. Using a replicated state machine for read/write would not be efficient. Thus, it is obvious two combine both approaches into one protocol that enables efficient processing for any type of request without violating consistency.

References

[AEMGG⁺05] Michael Abd-El-Malek, Gregory R. Ganger, Garth R. Goodson, Michael K. Reiter, and Jay J. Wylie. Fault-scalable Byzantine fault-tolerant services. In *SOSP '05: Proceedings of the twentieth ACM symposium on Operating systems principles*, pages 59–74, New York, NY, USA, 2005. ACM Press.

[Lam86]   Leslie Lamport. On Interprocess Communication. *Distributed Computing*, 1(2):77–85, 1986.

[Lam98]   Leslie Lamport. The Part-Time Parliament. *ACM Transactions on Computer Systems*, 16(2):133–169, 1998.

[Sch90]   Fred B. Schneider. Implementing Fault-Tolerant Services Using the State Machine Approach: A Tutorial. *ACM Computing Surveys*, 22(4):299–319, 1990.

Conditional Random Fields for Detection of Visual Object Classes

Paul Schnitzspan
MIS Group
TU Darmstadt

Object detection is one of the key tasks in computer vision. Therefore, locations of different object instances have to be highlighted within one image. Conditional random fields gained increasing attention and were applied to tasks of computer vision. In my work I am focussing on these models to increase the discriminative nature of conditional random fields and thereby achieve better detection results.

In the context of Cooperative, Adaptive and Responsive Monitoring in Mixed Mode Environments my goal is to infer a meaningful interpretation of one given scene (e.g. victim detection and estimation of risks in an earthquake scenario). In this setting I will face the challenges of multiscale, multiview and heavily occluded object instances.

Dynamic map building for optimal trajectory planning by using multiple UVs

Armin Strobel

Institute of Flight Systems and Automatic Control
TU Darmstadt
Petersenstr. 30
D-64287 Darmstadt
strobel@fsr.tu-darmstadt.de

For unmanned vehicles no matter which kind they are of, it is a well known problem to conduct there missions in an unknown or even changing environment, e.g. a catastrophe scenario after a hurricane or an earthquake. Even if there has been a map of the region before the disaster, it is likely that it is more or less useless at the moment, when the unmanned vehicles are used to explore the region. Besides missing or new objects, additional challenges an unmanned vehicle has to cope with are moving obstacles like other vehicles or persons.

To be able to navigate efficiently through a disaster area, an autonomous robot has to have a good knowledge of its environment. A good starting point is a map dated before the disaster but it is much more important to update this map while operating in the area.

The aim of this research is to get to an updated map of a disaster scenario much faster by using multiple heterogeneous UVs exploring the area from all sides and sharing their experiences with all other vehicles in the swarm. One approach could be to have every robot equipped with a single camera, sending the data to a ground station, where all pictures are used to calculate a virtual 3D representation of the environment (distributed photogrammetry). A little swarm of heterogenous robots will be build to implement the algorithms and evaluate them in real scenarios.

GRK 1387:

dIEM oSiRiS

**Integrative Development of Modeling and SImulation
Methods for RegeneratIve Systems**

Rostock

Integrative Development of Modelling and Simulation Methods for Regenerative Systems

University of Rostock
Graduate School dIEM oSIRIS
Adelinde Uhrmacher
lin@informatik.uni-rostock.de

The ability of regeneration characterizes cell biological systems and is increasingly required for computer science systems as well. Regenerative systems are able to overcome significant perturbations, and maintain autonomously their functionality in dynamic and uncertain environments.

To analyse or develop these types of systems modelling and simulation plays a crucial role. However, due to being large scale and embracing many heterogeneously acting and interacting sub-systems with variable patterns of behaviour, interaction and composition, they require the development of new modelling and simulation methodologies to support a flexible modelling at different levels of organization, space, and time, and an efficient execution of experiments.

The development of methods will be driven by different Computer Science disciplines. E.g. they require specific data base concepts for storing, retrieving, and matching of model components. Particularly, the integration of state of the art visualization methods promises new and more effective modelling and simulation methods. Visualization shall support designing models, analyzing processes and data, and supporting the discussion between computer scientists and biologists. Thus, visualization will be used for exploration, confirmation, and presentation.

The analysis of characteristics and requirements of regenerative systems and the evaluation of the developed concepts shall be based on a concrete biological application: the exploration of the Wnt signalling pathway which plays a significant role in the differentiation of neuronal cells. Therefore, the wet-lab experiments will utilize diverse methods. In addition to gel electrophoresis, blotting and chromatography, for the investigation of spatial processes also confocal fluorescence microscopy is applied.

The Graduate College brings together researchers from Medicine, Biology and Computer Science and will contribute towards achieving new insights into the functioning of biological cell systems, establishing modelling and simulation as an experimental methodology in Biology, and developing innovative modelling and simulation methods and tools from which the understanding as well as the design of regenerative systems in general will benefit.

Die integrative Entwicklung von Modellierungs- und Simulationsmethoden für regenerative Systeme

University of Rostock
Graduate School dIEM oSIRIS
Adelinde Uhrmacher
lin@informatik.uni-rostock.de

Regenerative Systeme besitzen die Fähigkeit, signifikante Störungen aus eigener Kraft zu überwinden, und Mechanismen, welche ein langfristiges Funktionieren von Systemen auch in a priori unbekannten Umgebungen ermöglichen. Diese Eigenschaft ist charakteristisch für zellbiologische Systeme und wird auch zunehmend für Informatiksysteme gefordert.

Für die Untersuchung oder Entwicklung regenerativer Systeme spielen Modellierung und Simulation eine zentrale Rolle. Regenerative Systeme stellen spezielle Anforderungen an die zu entwickelnden Methoden, denen existierende Modellierungs- und Simulationsmethoden nicht gerecht werden. Es gilt daher, Modellierungs- und Simulationsmethoden zu entwickeln und diese integrativ aufeinander abzustimmen. Sie sollen es ermöglichen, eine Vielzahl von interagierenden, heterogenen Subsystemen mit variablen Kompositions-, Interaktions- und Verhaltensmustern auf unterschiedlichen Organisations- und Abstraktionsebenen effektiv und effizient zu modellieren und zu simulieren.

Die Methodenentwicklung wird dabei durch unterschiedliche Informatikdisziplinen getragen. Beispielsweise müssen spezielle Konzepte zur Speicherung, Matching und Retrieval von Modellkomponenten in Datenbanken entwickelt werden. Insbesondere durch die Kombination von Visualisierung, Modellierung und Simulation werden neue, effektivere Modellierungs- und Simulationsmethoden entwickelt. Visuelle Unterstützung erleichtert dabei nicht nur den direkten Entwurf von Modellen, sondern fördert insbesondere auch die Prozess- und Datenanalyse sowie die Kommunikation zwischen Biologen und Informatikern. Die Visualisierung wird damit sowohl für die explorative Analyse, die konfirmative Analyse als auch für die Präsentation genutzt.

Die Auseinandersetzung mit Charakteristika und Anforderungen regenerativer Systeme und die Evaluierung der entwickelten Modellierungs- und Simulationskonzepte soll anhand einer konkreten biologischen Anwendung erfolgen: der Untersuchung des wnt-Signalweges, der in der Differenzierung neuronaler Zellen eine wichtige Rolle spielt. Die Datenerfassung im Wet-Lab nutzt unterschiedliche Verfahren. Neben molekular-biologischen Standardmethoden wie Gel-Elektrophorese, Blotting und Chromatographie, wird zur Untersuchung von räumlichen Prozessen auch konfokale Fluoreszenzmikroskopie eingesetzt.

Das Graduiertenkolleg ist stark interdisziplinär angelegt und führt Experten aus den Bereichen der Informatik, der Medizin und der Biologie zusammen. Es wird dazu beitragen,

neue Erkenntnisse im Bereich biologischer Systeme zu gewinnen, Modellierung und Simulation als experimentelle Methodik in der Biologie zu etablieren und innovative Modellierungs- und Simulationsmethoden zu entwickeln, die auch in weiteren Anwendungsbereichen, die durch ähnliche Eigenschaften gekennzeichnet sind, fruchtbar eingesetzt werden können.

Dagstuhl Conference 2007

The role of Wnt pathway-mediated changes in the cytoskeleton-motor enzyme system during differentiation of neural progenitor cells

Benjamin M. Bader, Dieter G. Weiss

University of Rostock
Institute for Biological Sciences
Albert-Einstein-St. 3
D-18059 Rostock
Germany

benjamin.bader@uni-rostock.de
dieter.weiss@uni-rostock.de

The canonical Wnt-pathway is involved in the differentiation of neural progenitor cells of different origins (regions) and species [1]. Differentiation of neural progenitor cells is accompanied by a total reorganization of cell morphology. The development of neural extensions and branches such as axons and dendrites is mediated by a reorganization of the cytoskeleton and an increase of organelle transport activity, based on both, microtubules and actin filaments. Here cytoskeleton-associated proteins such as MAPs (microtubule associated proteins) or ABP (actin binding proteins) play a major role in stabilizing the filaments. These proteins are activated by phosphorylation which results in binding to their target molecules. Through mechanisms of the canonical Wnt-pathway MAP-1B, APC, CRMP-2 and other microtubule stabilizing factors are influenced in their phosphorylation status by kinases like GSK-3 [2, 3]. Changes in the cytoskeleton-motor enzyme system will most likely also influence the distribution of protein complexes of the Wnt-pathway. Therefore, the question is studied, how the cytoskeleton-motor enzyme system is influenced by the Wnt-pathway and *vice versa* and how this influences the differentiation of neural progenitor cells.

A human neural progenitor cell line (ReNcell VM197, ReNeuron, Guildford UK) derived from 10 week old foetal ventral mesencephalon [4] is used to answer these questions. Confocal laser scanning microscopy is used to analyse the spatial and temporal correlation of the cytoskeleton with Wnt-pathway proteins such as ß-catenin, Axin, APC, GSK-3ß or Dvl. Colocalisation analyses and life cell imaging data will be used to develop an improved mathematical model of the pathway. The combination of the two approaches namely experimental work, such as in vivo exploration, and computational modelling [5] is expected to give the possibility to gain new insights in Wnt pathway-mediated regulation and in subcellular protein trafficking and localization.
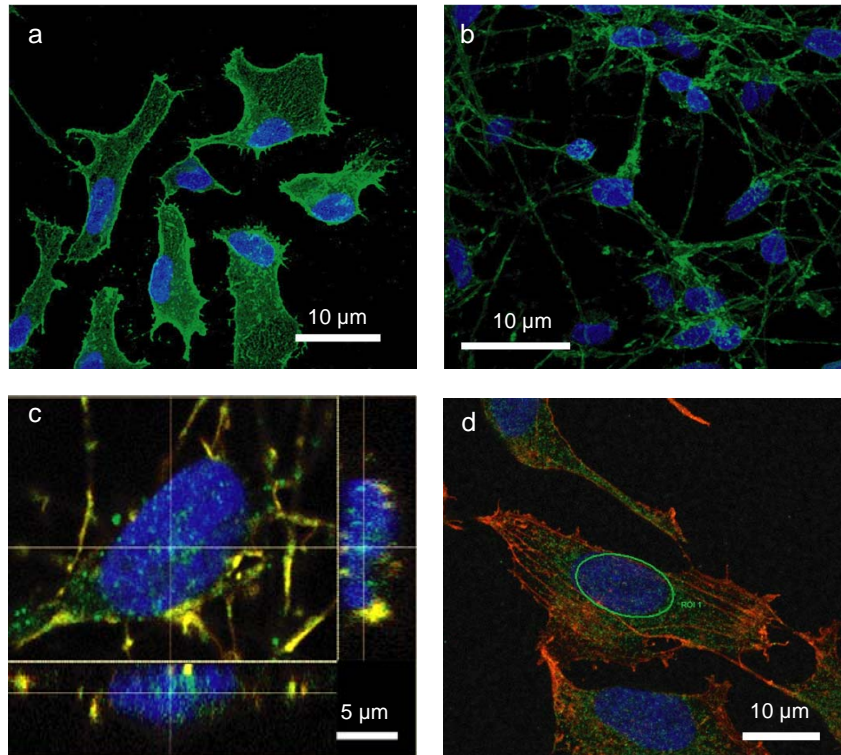
Fig. 1: Confocal microscopy analysis of proliferating and differentiating VM197 cells. (a) Proliferating and (b) differentiating cells immunostained for ß-catenin (green). (c) Differentiating cells stained for Dvl-1 (green). Confocal cross section shows nuclear location of Dvl-1. (d) Proliferating cells stained for GSK-3ß (green), with regions of interest (ROIs). Quantitative information about fluorescence signals in each section of confocal stacks is used to analyse changes in subcellular protein location. (c-d) Additional staining of f-actin (red). (a-d) Nuclear staining (blue).

References:
[1] Castelo-Branco et al. 2004, J Cell Sci 117: 5731-7
[2] Ciani et al. 2004, J. Cell Biol 164: 243-53
[3] Krylova et al. 2000, J. Cell Biol 151: 83-94
[4] Hoffrogge et al. 2006, Proteomics 6: 1833-47
[5] Lee et al. 2003, PLoS Biol 1: 116-32

Influence of Wnt proteins on differentiation of human neural progenitor cells

Universität of Rostock
Graduate School dIEM oSIRIS
Rayk Hübner
rayk.huebner@med.uni-rostock.de

Introduction

Wnts are secreted lipid-modified glycoproteins found in all metozoans examined to date. During development, they are used by cells to influence the fate or behavior of neighboring cells. It is well established that WNT proteins are important mediators of intercellular communication, and that signalling by members of the WNT family of molecules is crucial for normal embryo development in various systems, including the nervous system. WNTs have been shown to regulate diverse cellular processes, including cell proliferation and fate, cell polarity and movement, and programmed cell death. In the mature organism, Wnts are implicated in maintaining stemcell- like fates in the intestinal epithelium, skin and hemopoietic cells. Inappropriate activation of Wnt signaling contributes to numerous human cancers and reduced Wnt signaling has been implicated in osteoporosis. WNT proteins can influence tissue organization and growth by functioning locally and also act at a distance, by generating a gradient across a tissue.
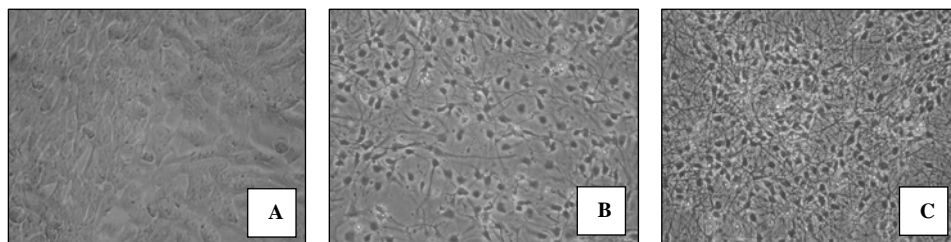

Fig.1 Images of (A) untreated life cells prior to removal of growth factors, (B) 4 days and (C) 7 days after differentiation. Proliferating cells show the typical „paving stone" morphology, after removal of growth factors they differentiate in neurons and astroglia.

Aims of this work

In this work I want to elucidate the influence of WNT proteins on the differentiation of the fetal human ventral-midbrain derived neural stem cell line ReN VM. Studies will include overexpression and specific knock-down of WNT proteins. Treated cells will be analyzed by immunofluorescence using antibodies specific for neuronal markers. Levels and phosphorylation status of relevant proteins will be analyzed by western blot technique. WNT dependent transcription is investigated by quantitative RT-PCR as well as TCF assays.

References:

[1]   Cadigan KM, Liu YI. Wnt signaling: complexity at the surface. *J Cell Sci.* 2006 1;119:395-402.

[2]   Ciani L WNTs in the vertebrate nervous system: from patterning to neuronal connectivity. *Nat Rev Neurosci.* 2005;6(5):351-62.

[3]   Hoffrogge R et al. 2-DE proteomeanalysis of a proliferating and differentiating human neuronal stem cell line (ReNcell VM). *Proteomics 2006* 6, 1833-1847.

Characterization of Apoptosis-pathways

in human neuronal progenitor cells

Universität Rostock
Graduate School dIEM oSIRIS
Alexandra Jaeger
alexandra.jaeger@uni-rostock.de

Apoptosis, or programmed cell death, is an essential component of the ontogenetic development of neuronal cells [1]. Proliferation and differentiation of neuronal progenitor cells are closely coupled to apoptotic processes and cell cycle regulation [2].

Aim of our project is to study the apoptotic processes during wnt-driven differentiation in the human neuronal progenitor cell line ReNcell VM197 (ReNeuron, Guildford UK, [3]). We will characterize the interactions and localization pattern of apoptosis relevant proteins and describe the underlying apoptosis pathways at different stages of differentiation employing protein biochemistry and light microscopy.

The canonical Wnt-signaling pathway is physiologically involved in neuronal differentiation. Effects of this pathway on gene expression of apoptosis-related proteins will be investigated at different time points of differentiation.

Based on the experimental data, computational models and visualization methods will be developed for the description of the involved pathways and the predictive power of the models will be evaluated in further *in-vitro* studies. This approach will allow a better understanding of cell death mechanisms during differentiation of neuronal progenitor cells and will be of major importance for clinical studies of neurodegenerative diseases.
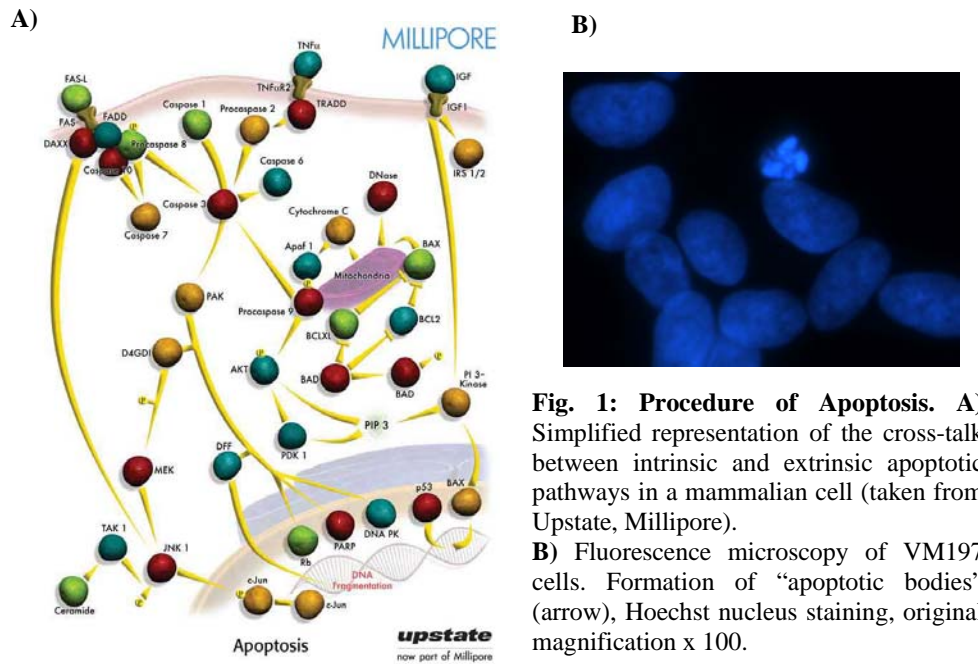
A)

B)

Fig. 1: Procedure of Apoptosis. A) Simplified representation of the cross-talk between intrinsic and extrinsic apoptotic pathways in a mammalian cell (taken from Upstate, Millipore).
B) Fluorescence microscopy of VM197 cells. Formation of "apoptotic bodies" (arrow), Hoechst nucleus staining, original magnification x 100.

References:

[1] W. Yeo and J. Gautier. Early neural cell death: dying to become neurons. *Dev Biol.*274(2): 233-44, Oct **2004**.

[2] L.Sommer and M. Rao. Neuronal stem cells and regulation of cell number. *Prog Neurobiol.* 66(1): 1-18, Jan
    2002.

[3] Pollock et al. A conditionally immortal clonal stem cell line from human cortical neuroepithelium for the
    treatment of ischemic stroke. *Exp Neorol.* 199: 143-55 May **2006**.

Flexible and Efficient Simulation of Biological Systems

University of Rostock
Graduate School dIEM oSIRIS
Matthias Jeschke
mj@informatik.uni-rostock.de

Models for biological systems become increasingly complex, e.g. by including spatial information or being composed of different sub-models, and require flexible and efficient simulation systems.

The use of formalisms for modeling (e.g. Petri Nets, StateCharts, DEVS, and π-calculus) can simplify the creation of models. As a consequence, a simulation system shall provide simulators for models described in different formalisms. An example for such a simulation system is JAMES II, a plugin based, easy extendable modeling and simulation framework, written entirely in Java [HU06].

Sometimes biologists might be interested only in specific parts of a model and need a high resolution simulation for these sub-models. Thus, a high precision simulator is required only for those critical parts while maintaining a lower simulation resolution for the rest of the model and therefore saving computational resources. One example are stiff reaction networks with slow and fast reactions [RPCG03].

Due to the high complexity of models, distributed, resource sharing simulation systems seem to be a promising approach. For example, the Aurora Parallel and Distributed Simulation System tries to fill the gap between distributed computing and parallel discrete event simulations using the master/worker paradigm[PF06].

My work is focused on the development, implementation and evaluation of efficient simulators that can operate on a wide range of biological models, including spatial, multi-level and multi-formalism models. Furthermore, I want to explore the potentials of various distributed computing approaches (e.g. Grid computing) in the field of discrete event simulation.

References

[HU06]   Jan Himmelspach and Adelinde Uhrmacher. Plug'n simulate. In *Proceedings of the 40th Annual Simulation Symposium*, 2006.

[PF06]   Alfred Park and Richard M. Fujimoto. Aurora: An Approach to High Throughput Parallel Simulation. In *PADS'06*, 2006.

[RPCG03]   M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *Journal of Chemical Physics*, 119:12784–12794, December 2003.

Spatial Modeling of Molecular Biological Systems

University of Rostock
Graduate School dIEM oSIRIS
Mathias John
mj069@informatik.uni-rostock.de

With the progress of wet-lab techniques spatial information becomes available that emphasizes the central role of space in inter and intra cellular dynamics. Same as crucial phases in the mitosis are distinguished by the spatial distribution of key actors in the cell so does the functioning of signaling pathways depend on the location of its components [Kho06]. Hence, modeling that does not take space into account abstracts away a major aspect of inter- and intracellular phenomena.

Existing approaches for modeling space cover population based as well as individual based methods. As discussed in [TLN06] population based approaches that account for space like e.g. partial differential equations have crucial limitations in Systems Biology (state explosion, lack of appropriate spatial resolution). Individual based spatial modeling approaches avoid these limitations by incorporating two major views. One is to start with the components and defining space in a relative manner by grouping them. Formalisms that support this view are e.g. BetaBinders [PQ05]. The other is to associate the components with individual, absolute locations and motions as done e.g. in case of simulation systems, e.g. [LAB05]. However, there is a lack of modeling formalisms that combine an individual-based approach toward spatial modeling with absolute space.

The goal of my work is to investigate existing approaches in order to create a formal toolbox for the modeling of spatial phenomena in bio-molecular systems. An emphasis shall be put on the definition of appropriate formalisms, because only formal approaches feature advantages that support next-generation modeling as it is proposed in [Pag94].

References

[Kho06]   B.N. Kholodenko. Cell-signalling dynamics in time and space. *Nature Reviews Molecular Cell Biology*, 7(3):165–176, 2006.

[LAB05]   K. Lipkow, S. S. Andrews, and D. Bray. Simulated diffusion of phosphorylated CheY through the cytoplasm of Escherichia coli. *J Bacteriol*, 187(1):45–53, January 2005.

[Pag94]   Ernest H. Page. *Simulation modeling methodology: Principles and etiology of decision support*. PhD thesis, Department of Computer Science, Virginia Tech, Blacksburg, 1994.

[PQ05]   Corrado Priami and Paola Quaglia. *Beta Binders for Biological Interactions*. 2005.

[TLN06]   Dominic P. Tolle and Nicolas Le Novère. Particle-Based Stochastic Simulation in Systems Biology. *Current Bioinformatics*, 1(3):1–6, 2006.

Definition, Storage, and Query of XML-based Model Components

University of Rostock
Graduate School dIEM oSIRIS
Dagmar Köhn
dk103@informatik.uni-rostock.de

Computer-aided analysis of signalling pathways is one of the growing fields in systems biology. A big challenge therefore is the development of models for cells, which then can be simulated in order to support biologists with assumptions about potential experiments. Those models have to be stored within a database and as well have to be available for later querying, which has to be efficient and as easy as possible.

The *Systems Biology Markup Language* (SBML [FH03]) is one of the most important XML-based standards. It is under constant development since 2000 and has been designed especially for the description of models for bio-chemical reactions. An example for the description of such a reaction in SBML is shown in figure .[1]

```
%{Code snipped from the SBML XML Schema for the EGFR model}{list:sbmlDoc}
<reaction name="EGFR_autophosphorylation" id="REACT_9388_1">
 <listOfReactants>
  <speciesReference species="REACT_2812_1" />
  <speciesReference species="REACT_9820_1" />
 </listOfReactants>
  ...
```

Figure 1: Extract of an SBML model description for the EGFR signalling pathway

The main question of this phd work is how SBML models – and XML models in general – can be stored into a database efficiently. Furthermore, it is a research question to find a way of identifying model components in an XML model automatically and exchange them with other models so that the model will still be executable and holds the same semantics.

References

[FH03]   Andrew Finney and Michael Hucka. Systems Biology Markup Language (SBML) Level 2: Structures and Facilities for Model Definitions, 2003.

---

[1]Code snippet from the "Epidermal Growth Factor Receptor" signalling pathway (EGFR), taken from the Reactome Datenbank [JTGV+05]

[JTGV⁺05] G Joshi-Tope, M Gillespie, I Vastrik, P D'Eustachio, et al. Reactome: a knowledge-base of biological pathways. *Nucleic Acids Research*, 33(D428-D432), 2005. PMID: 15608231.

Exploration of the canonical Wnt signaling pathway

University of Rostock
Graduate School dIEM oSIRIS
Orianne Mazemondet
orianne.mazemondet@med.uni-rostock.de

Wnt molecules have a key role in the differentiation of neural cells. Their major signaling pathway is mediated by beta-catenin which is the core of a complex protein network. As dopaminergic neurons are involved in Parkinson's disease, the exploration of Wnt signaling pathway in human neural stem cells is relevant for regenerative medicine.

I am working with human progenitor stem cells (ReNcell VM197, ReNeuron, Guildford UK) derived from 10 week old foetal ventral mesencephalon. These cells differentiate after growth factor withdraw into different neural cell types. I am specifically interesting in dopaminergic neurons as they are involved in Parkinson's disease.
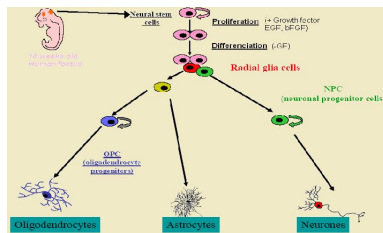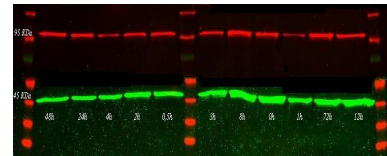

Figure 1: Differentiation of neural stem cells

Figure 2: Time kinetic of beta-catenin

Wnt pathway network is investigated at different time points through the differentiation process of the cells. I fractionate the cell lysate to obtain each subcellular compartment independently (cytoplasm, nucleus, membranes). Western blot analysis is used to explore the major proteins involved in the Wnt canonical pathway, such as βcatenin, Axin, APC or Frizzled/LRP6 receptors. Quantitative data is obtained from Infrared fluorescence system (Odyssey scanner, LI-COR) and analyzed in order to create a mathematical model of the pathway. The system biology approach opens new hypotheses for βcatenin mediated regulation and protein transport through subcellular compartments.

Systems biology of the canonical Wnt pathway

University of Rostock
Graduate School dIEM oSIRIS
Yvonne Schmitz
schmitz@informatik.uni-rostock.de

Processes in biological cells, e.g. signal transduction, metabolism and proliferation, are affected by complex networks of biochemical reactions. Systems biology aims to identify relations between structure and function of biochemical networks by mathematical modeling and computer simulations. The mathematical models, which describe temporal changes of protein concentrations, are established on the basis of the interaction between the different network proteins. A widely used approach is based nonlinear ordinary differential equations. The model properties are investigated by formal and numerical methods, for example bifurcation analysis. The Wetlab directly benefits from this theoretical analysis, as testable predictions of critical parameters can be used to optimise refined experiments.
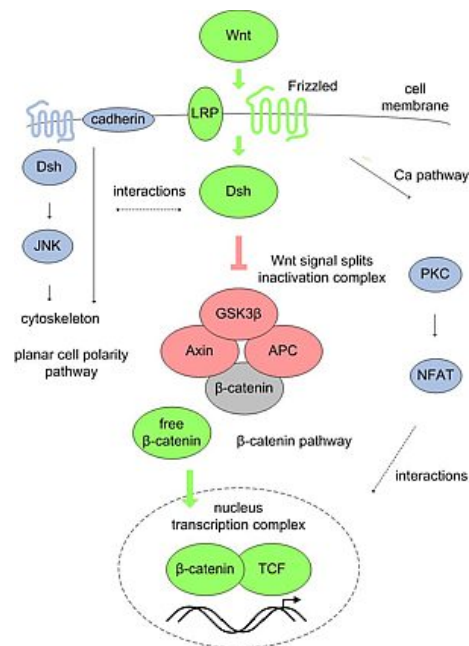
Figure 1: The canonical Wnt signalling pathway

The Wnt signalling pathway plays an important role in both, the embryonic development and the biology of adult stem cells, including cancer [LN04]. From all existing Wnt signalling pathways, the canonical Wnt signalling pathway is the one most widely studied. The reaction scheme of this pathway is shown in Fig. 1. Via Wnt signalling, the concentration of β-catenin is stabilised, which acts as a transcription cofactor.

[LSK+03] have developed a mathematical model that describes important biochemical reactions of the canonical Wnt signalling pathway in *Xenopus* egg extracts. Analogously, we want to investigate the role of the canonical Wnt signalling pathway during the differentiation of neural progenitor cells. On the basis of [LSK+03] and in discussion with experimentalists from biology and neurology, we are aiming to develop and analyse an appropriate mathematical model describing the pathway in this process. The model parameters will be calculated from experimental time series of the pathway proteins. By computer simulations we study the influence of spatial aspects and feedback loops on the dynamics of intercellular and intracellular signal transmission.

References

[LN04]   C.Y. Logan and R. Nusse. The Wnt-Signaling Pathway in Development and Disease. *Ann. Rev. Cell Dev. Biology*, 20:781–810, 2004.

[LSK+03] E. Lee, A. Salic, R. Krüger, R. Heinrich, and M.W. Kirschner. The Role of APC and Axin Derived from Experimental and Theoretical Analysis of the Wnt Pathway. *PLOS Biology*, 1:116–132, 2003.

GRK dIEM oSiRiS: Visual Graph Mining

Hans-Jörg Schulz

hjschulz@informatik.uni-rostock.de

The analysis of networks or graphs is a current field of research in many application domains ranging from Homeland Security to Biochemistry. Its aim is to further the understanding of given networks by extracting important properties and prominent patterns. One aspect that lately receives much attention is the combination of graph analysis with information visualization techniques – the Visual Graph Mining. Ideally the visualization can be utilized to communicate results and to steer the analysis process. But in practice, it is hard to integrate the interactivity of the visualization on one side and the often long-running analysis methods on the other side in a useable way. While there already exist a few approaches that bridge this gap, in most cases they are tailored to specific application domains and thus to specific data characteristics and analysis tasks. In the dIEM oSiRiS graduate school, the need for an intuitive way to analyze graphs arises in many different fields:

- in medicine, when evaluating interaction partners in protein complexes

- in biology, when identifying biochemical pathways

- in databases, when examining database schemata

- in modeling and simulation, when exploring model structures

This variety of different applications requires a broader approach to Visual Graph Mining, whose conception is the aim of my dissertation. A first step towards such a more generic approach is the systematization and classification of the most prevalent methods from network analysis and graph theory, as well as from graph visualization and graph drawing. A possible classification of graph visualization and drawing techniques has already been published, a publication of a systematization for graph analysis techniques is in preparation. In addition to this formal approach, a prototype of a software framework for Visual Graph Mining has been implemented and published as well, that is used as a testbed for new ideas. It also serves as a basis for developing special purpose tools to support the specific graph analysis problems within the scope of dIEM oSiRiS. As some of these problems make extensive use of non-standard graph-models like multi graphs, hypergraphs or compound graphs, these real world use cases will provide challenging touchstones to prove the flexibility and validity of the general Visual Graph Mining concept in different domains.

GRK dIEM oSiRiS: Visual support for the development of methods and models for regenerative systems

Andrea Unger

aunger@informatik.uni-rostock.de

In many scientific disciplines, huge volumes of data are generated and stored today. One focus of visualization is the development of suitable interactive methods for the analysis of such data by taking advantage of the human capability to comprehend visually represented data.

One field where very large and heterogeneous data sets need to be analyzed is the modeling, simulation, and evaluation of regenerative systems. The field can benefit to a great extend from visualization if visual methods are utilized to explore the data or confirm assumptions about the data. However, it is not sufficient to apply already existing visualization techniques. The appropriateness of a visualization method always depends on the data to be visualized, the goals of the analysis, and the visual representations which are commonly accepted by domain experts. Therefore, suitable concepts from visualization must be adapted to the requirements of the application domain to derive tailored visualization techniques. As a first step, a methodology to integrate both fields has been derived to investigate the potential of visualization techniques for the development of regenerative systems systematically. Hence, it can be identified which requirements are not met by currently existing visualization techniques and, thus, further research is necessary.

One important subject identified throughout the process of developing regenerative systems is the presence of uncertainty. Uncertainty is a general term referring to problems in data quality. Different concepts are included, such as error, imprecision, lineage, and subjectivity. The presence of uncertainty in data must be communicated to allow correct decisions which are based on the data. This is notably true for the application domain of developing regenerative systems. They are build from experimental data of limited reliability. This uncertainty continues in the modeling and simulation process. But data quality problems do not solely result from data acquisition and modeling. They can also arise in all steps of the visualization pipeline, which includes filtering, mapping of data to visual attributes, and the rendering process. Although some specific methods to visualize uncertainty information have already been proposed, there is still a need to develop intuitive concepts to represent and integrate uncertainty information into data visualization. In particular, the handling of large volumes of uncertainty in addition to large volumes of data is a great challenge. To avoid visual cluttering, decisions must be made about quantity and type of displayed uncertainty values, depending on the visualization context.

This work will focus on new concepts to visualize uncertainty, with special consideration of the application background of regenerative systems.

GRK 1424:

MuSAMA

Multimodal Smart Appliance Ensembles for Mobile Applications

Rostock

Multimodal Smart Appliance Ensembles for Mobile Applications

Universität Rostock
Graduiertenkolleg MuSAMA

MuSAMA liegt die These zugrunde, dass die zukünftige ubiquitäre Informationstechnik in der Regel von dynamischen Appliance Ensembles gebildet wird – lokale Ansammlungen voneinander unabhängiger, stationärer und mobiler Artefakte, deren Zusammensetzung sich unvorhersehbar ändern kann. Die Mitglieder eines solchen Ensembles – etwa ein Mobiltelefon und ein Fahrkartenautomat – müssen in der Lage sein, spontan und ohne Hilfestellung sinnvoll miteinander zu kooperieren.

Die zentrale Innovation von MuSAMA ist die Verwirklichung von kohärenter Assistenz des Menschen durch ein Geräteensemble, ohne globales Wissen über die Struktur und die Fähigkeiten des Ensembles an irgendeiner Stelle erforderlich zu machen. Daraus ergibt sich auch das Projektakronym, das für multimodale, smart mobile Appliance Ensembles steht und damit das emergente intelligente Verhalten zum Ziel hat, das mobile, ad hoc verbundene Appliances aufweisen sollen. Aus Sicht der Benutzerinteraktion stehen multimodale Kommunikationsformen mit dem Anwender im Vordergrund.

Um die gestellten Anforderungen erfüllen zu können, sind Verfahren erforderlich, die es solchen „Smart Appliances"ermöglichen spontan zu kooperieren, um sowohl die Intentionen des in der Umgebung handelnden Nutzers zu erkennen, als auch Strategien für geeignete Aktionen bzw. Reaktionen zu bestimmen und diese gemeinschaftlich auszuführen. Dabei muss berücksichtigt werden, dass jede Einzelkomponente letztendlich lediglich Kenntnis über die eigenen Fähigkeiten und über die lokale Kooperation mit anderen Komponenten verfügen kann: bei einem hinreichend großen Spektrum unterschiedlicher Appliances ist die Aufzählung der „interessanten"Ensembles (und das Festlegen des globalen Verhaltens dieser Ensembles) nicht mehr sinnvoll möglich. Globales Verhalten - synergetische Kooperation mehrerer Appliances um ein Ziel zu erreichen, das außerhalb des Begriffsrahmens jeder einzelnen Appliance liegt – muss daher dezentral realisiert werden. Wir bezeichnen diese Fähigkeit folgend als emergente kooperative Assistenz.

Innerhalb von MuSAMA werden wir Konzepte und Verfahren für die Bereitstellung emergenter kooperativer Assistenz auf der Basis spontaner Ensembles entwickeln. Diese Verfahren sollen es einem Ensemble ermöglichen, aus aktuellen Sensordaten, Interaktionsereignissen und – möglichst lokalen – Vermutungen über die Präferenzen und Handlungsziele des/der Nutzer(s) eine sinnvolle Hypothese über den aktuellen Unterstützungsbedarf zu entwickeln.

Catchwords like „Ubiquitous Computing ", „Pervasive Computing ", or „Ambient Intelligence "paraphrase the vision of a world, in which we are surrounded by smart, intuitively operated devices that help us to organize, structure, and master our everyday life. They share the notion of a smart, personal environment which characterizes a new paradigm for the interaction between a person and his everyday surroundings: Smart environments enable these surroundings to become aware of the human that interacts with it, his goals and needs. So it is possible to assist the human proactively in performing his activities and reaching his goals.

Hitherto, it is the user's responsibility to manage his personal environment, to operate and control the various appliances and devices available for his support. But, the more technology is available and the more options there are, the greater is the challenge to master your everyday environment, the challenge not to get lost in an abundance of possibilities. Failing to address this challenge adequately results in technology becoming inoperable, effectively useless. The goal of smart environments is to take over this mechanic and monotonous control task from the user and manage appliance activities on his behalf. By this, the environment's full assistive potential can be mobilized for the user, tailored to his individual goals and needs.

MuSAMA is based on the hypothesis that ubiqutous machine intelligence, envisioned for our future everyday environments, will be provided by dynamic ensembles: Local agglomerations of smart appliances, whose composition is prone to frequent, unforeseeable, and substantial changes. Members of such ensembles need to be able to cooperate spontaneously and without human guidance in order to achieve their joint goal of assisting the user. Achieving this capability of autonomous cooperative assistance poses new challenges for the research on ubiquitous and ambient information technology.

Work in MuSAMA therefore concentrates on the investigation of models and algorithms that allow dynamic, ad-hoc ensembles to deliver the assistive power of Smart Environments independently of external or global knowledge. Globally coherent ensemble behavior with respect to a user's need emerges from local interaction of individual appliances. The scenario of a „pervasive university "provides the required confrontation with reality for empirical analysis and evaluation.

Confidentiality Models in Smart Appliance Ensembles

University of Rostock
DFG Graduate School MuSAMA
Christian Bünnig*
christian.buennig@uni-rostock.de

Smart appliance ensembles (SAE) consist of several devices embedded into a common context. This context is derived by detection algorithms which operate on a combined storage that gets filled by all participating appliances. Preconfigured security with use of centralized security techniques like *PKI* or mutual trust relations like *Web of Trust* are not applicable to such ensembles because of their ad-hoc nature. In order to enable a comprehensive cooperation between nodes from the ensemble and in order to make detection of context as precise as possible, every node initially trusts any other node. A node which needs to derive some common context must be allowed to gain access to any data stored in the combined storage and to write new context information into that storage. This situation raises some security related problems, especially in concerns of privacy.

Access to all available context information implies access to detailed information about a person's behavior and activities. This information may get misused for purposes originally not intended by persons acting in an SAE, e.g. location tracking or behavior profiling. The open structure of SAE further makes it possible for malicious devices to join such ensembles. These devices are another critical part which may misuse the information stored in the combined storage of the ensemble. It is crucial to protect information derived from a person's activities. On the other hand, if societal conventions request it, in certain circumstances it may be necessary to make this information accessable to authorized third parties.

This thesis is going to tackle these problems by collecting typical scenarios in SAE and analyze them regarding the flow of confidential information. Based on these investigations, possible fraudulent use of confidential information will be examined. Once potential threats to privacy has been specified, it is planned to develop concepts and protocols which make participation in SAE guarded against these threats. The basic idea is to leverage adequate encryption mechanisms to protect identity information of persons acting in an SAE. On the one hand these mechanisms shall prevent a global tracking and profiling of a person's activities within one or more SAE. On the other hand it shall still be possible to use authorization techniques and trust systems with encrypted identity information. The privacy protection mechanisms are supposed to get integrated in context dependent security policies.

Modelle und Deskriptoren für Verhaltenserkennung in physischen Umgebungen

Universität Rostock
Graduiertenkolleg MuSAMA
Christoph Burghardt
Christoph.Burghardt@uni-rostock.de

Aktuelle Forschungen im Bereich Ubiqitous Computing versuchen mit Hilfe von sequentiellen Monte-Carlo Methoden und wahrscheinlichkeitstheoretischen Verfahren die Intention eines Nutzers in bestimmten Situationen vorhersagen. Einige Projekte zeigen viel versprechende Ansätze. Die eingesetzten Modelle des Nutzers haben jedoch nur beschränkte Gültigkeit für bestimmte Bereiche oder eine kurze Zeitdauer. Deshalb ist es nötig, diese Modelle dynamisch an die Umwelt und die verfügbaren Sensoren / Effektoren anzupassen.

Die Anpassung eines Modells an eine bestimmte Situation übernimmt ein Entwickler, der versucht die möglichen Kombinationen vorab zu berücksichtigen. Die Synthese eines Modells für ein bestimmtes Szenario aus Vorwissen ist sehr aufwendig und fehleranfällig. Es wird eine Ontologie benötigt, welche das mögliche Nutzerverhalten beschreibt. Der Aufwand steigt mit der Anzahl an Situationen, die im Modell berücksichtigt werden sollen. Deshalb müssen Lösungsstrategien entwickelt werden, welche die Modelle automatisch an das jeweilige Szenario anpassen können.

Neben der Dynamik der Situation muss auch der Veränderlichkeit des Geräte-Ensembles berücksichtigt werden. Durch Informationen über die vorhandenen Ensemble-Mitglieder lassen sich die Interaktionsmöglichkeiten des Menschen mit den Ensemble vorhersagen und eingrenzen. Dies erfordert ein einheitliches Basisvokabular, mit dessen Hilfe die verschiedenen Appliance-Mitglieder im System Wissen über den Umgebungszustand (Position und Zustand der physischen Objekte, Positionen und aktuelle Tätigkeit von Personen, Lautstärke, Helligkeit, . . .) und die Fähigkeiten der verfügbaren Effektoren austauschen können. Hierfür müssen entsprechende Basisontologien bereitgestellt werden, die dieses Vokabular und die zugrunde liegenden Modelle definieren.

Ziel der Arbeit ist es, für dynamische Situationen und Ensembleverhalten ein jeweils optimales Modell für die Intentionsanalyse zu synthetisieren, welches automatisch den aktuellen Umgebungszustand berücksichtigt. Das Modell ist die Grundlage für eine deklarative Beschreibung von Nutzerzielen (als Menge von Bedingungen an den Umgebungszustand). Es bildet die Schnittstelle für die weitere Ensembleplanung.

Models and descriptions for behaviour recognition in physical environments

University of Rostock
Graduate school MuSAMA
Christoph Burghardt
Christoph.Burghardt@uni-rostock.de

Current research in ubiquitous computing use sequential Monte-Carlo methods to make predictions about the intention of a user within certain situations. Some projects show promising results. The problems so far are the underlying user-models, which are generally restricted to limited areas or a limited duration. Therefore it is necessary to adapt these models dynamically with respect to the changing environment and the number of sensors and effectors.

The adaption of a model to a new situation is the responsibility of a developer, who tries to imagine the possible intentions and actions of a user. The development from a-priory knowledge is complicated and error prone. An ontology is needed to describe the user behaviour. The complexity increases with the number of situations the model needs to handle. New solutions must be developed, which can automatically adapt the model to a new scenario.

Beside the dynamic nature of the environment, the changing structure of the device ensemble needs to be taken into account. By incorporating information from other ensemble members the number of possible human interactions with the devices can be calculated. Therefore the complexity of the model can be reduced. For a practical implementation, a unique basic ontology is needed. With the help of the ontology ensemble members can exchange information about the state of the environment (position and state of physical objects, positions and activities of persons, loudness, luminance, ...) and the capabilities of the involved actuators. For this purpose a basic ontology needs to be developed, which defines the vocabulary and the underlying models.

The goal of the thesis is to synthesize an optimal model for intention analysis, depending on the environment and the device ensemble. The model gives a declarative description of user goals (as a quantity of environment states). Thus it forms the foundation for the further ensemble planning.

Zuverlässige kontextbezogene Geräte-Sensor-Netzwerke in horizontalen und vertikalen drahtlosen Kommunikationsarchitekturen

Universität Rostock
Graduiertenkolleg MuSAMA
Enrico Dressler
Enrico.Dressler@uni-rostock.de

Im Rahmen dieses Vorhabens wird die Bereitstellung einer flexiblen Gerätekommunikation angestrebt, die einen zuverlässigen und kontextorientierten Datenaustausch durch die Integration von funk- und kabelgebundenen Technologien jeglicher Art ermöglicht. Dabei werden Qualitätsanforderungen beim Datenaustausch zwischen den mobilen Geräten und den in der Umgebung befindlichen Sensoren sowie der Kontext einer Anwendung zur Selektion geeigneter Kommunikationskanäle in einer hochgradig heterogenen Netzwerkinfrastruktur (Ethernet, Bluetooth, WLAN, WiMAX, UMTS, GPRS,...) berücksichtigt.

In diesem Forschungsvorhaben werden automatisch konfigurierbare horizontale und vertikale Kommunikationsstrukturen entwickelt, mit denen der ressourcenschonende Informationsaustausch zwischen den Sensoren, den mobilen Systemen sowie den durch Festnetz angeschlossenen zentralen System- und Applikationsservern organisiert wird. Horizontale Strukturen werden dabei als eine Kommunikationsplattform innerhalb einer Zelle von Knoten mit gleichen Netzwerkschnittstellen gesehen. So können zum Beispiel ein PDA, ein Tablet PC und andere Bluetooth-fähige Geräte ein gemeinsames lokal beschränktes Scatternetz aufbauen. WLAN-fähige Geräte bilden ein Ad-hoc-Netzwerk, das eine eigene horizontale Netzwerkstruktur darstellt. Geräte, die über Ethernet-Schnittstellen verfügen, können Teilnehmer eines Infrastrukturnetzes sein und über zentralisierte Mechanismen kommunizieren. Geräte mit Schnittstellen in verschiedenen horizontalen Strukturen verbinden diese durch den Aufbau vertikaler Strukturen. Einzelne horizontale Kommunikationsstrukturen werden so zu einem heterogenen Ensemble erweitert, und die Mobilität der Geräte wird über horizontale Strukturen hinaus ermöglicht.

Dieses Vorhaben liefert das Konzept einer neuen, allgegenwärtigen Kommunikationsverbindung zum kontextbezogenen Austausch von Daten zwischen den mobilen Sensoren und Geräten, in Ad-hoc- und Infrastrukturnetzen. Hierbei werden alle verfügbaren Kommunikationsverbindungen in Betracht gezogen und abhängig vom Kontext bzw. von Randbedingungen im Netzwerk zur dynamischen Selektion geeigneter Kanäle genutzt. Dabei sind Faktoren wie Verfügbarkeit, Energiebedarf, Auslastung der Geräte/Übertragungskanäle sowie Informationen über den Kontext in dem sich ein Gerät oder eine Applikation befindet, zu berücksichtigen. Auch Aspekte wie Authentifizierung, Autorisierung, Accounting und Sicherheit werden zur Gewährleistung von Vertrauen und Sicherheitseigenschaften zwischen Kommunikationsknoten berücksichtigt.

Precise Estimation of Position and Orientation for Ambient Intelligence

University of Rostock
Research Training Group MuSAMA
Dipl.-Ing. Dominik Lieckfeldt
dominik.lieckfeldt@uni-rostock.de

Using modern radio technologies provides the possibility of determining position and orientation of things. Position and orientation are fundamental information and form a base for models and algorithms which will be used in a smart ensemble. Especially, mobility of members of a smart ensemble requires continuous adaptation of their position and orientation. Central aspects of this Ph.D. thesis are 1) evaluation of quality, robustness and applicability of methods of positional and directional inference for mobile devices and 2) to develop new methods for precise estimation of position and orientation. In particular, methods following an approach using transponder technology shall be investigated (RFID-transponder). Regarding position, different configurations of stationary reader/mobile transponders as well as mobile reader/stationary transponders shall be considered. Concerning orientation, possibilities of applying multiple transponders to an object shall be explored. Thereby, I will try to exploit the distinct feature of RFID systems, i.e. identification of RFID transponders depends on the alignment of reader and transponder antennas.

Initially, the Cramer-Rao-bound on the estimation of orientation using RFID-transponders will be investigated. The results will provide information about the achievable accuracy of estimation with respect to orientation and will be the base for further investigations. Typically, in order to estimate the orientation of real world objects several transponders are needed. The total number will largely depend on the rotational movement to be expected from one particular object. A division into classes can provide a base for future modeling and prediction of rotational movement. The objects in each class shall show a preferably similar behavior concerning their temporal rotational patterns. Such a classification will reduce the state space for estimation of orientation and, therefore, will on one hand support a precise estimation and on the other hand reduce the number of transponders per object needed for some classes. During the last years, methods for estimating positions which rely on radio signal strength attracted increasing attention in the research community. As radio signal strength is used for both positional and directional inference, the benefits of jointly estimation of position and orientation using radio signal strength will be investigated.

Synthesis of Assistant Services in Adhoc Multi Media Environments

University of Rostock
Graduate School MuSAMA
Florian Marquardt
fm@informatik.uni-rostock.de

In the context of project MuSAMA adhoc ensembles are understood as spontaneous conglomerations of various multi media devices in one a spatially confined area, like a room for instance. These devices offers different services. Besides the multi media devices several sensors, capable of measuring the environment, are located is this ensemble. Using the sensor data intentions of persons located in the room can be determined.

Based on the intentions and the potentialities of the devices, it shall be possible to precisely offer complex functions of the ensemble to the users. In order to reach this, available services have to be composed.

In this dissertation different ways of composition will be explored and compared. The research should focus on a planning approach, whereas hierarchical task network (HTN) planning will be considered in detail.

Recent approaches manly considered tight coupled ensembles, wherein the involved devices are well known from the beginning. Thus one main focus of the work is the examination of requirements and consequences of enhanced dynamic behavior. Particular it is about unpredictable events like joining of a new device or a devices malfunction. Besides dynamic behavior, interoperablilty of the devices arises as a further essential issue in this connexion. Real dynamic behavior in matters of components exchangeability requires semantics and structures, formulated in common and universally valid formats and standards.

Services composition and orchestration should be realized in a mobile broker servce. Therefore a central coordination of the adhoc ensemble is assumed in this dissertation.

The dissertation shall furthermore determine whether semantic web standards and techniques are appropriate to reach the described and required dynamic behavior in adhoc ensembles.

To be able to create plans based on intentions of users, it becomes necessary to formalize possible user intentions. Since there is no universal valid language to formalize such intentions, it must be considerd within the scope of the dissertation.

Synthese von Assistenzdiensten in Adhoc Multi-Media Umgebungen

Universität Rostock
Graduiertenkolleg MuSAMA
Florian Marquardt
fm@informatik.uni-rostock.de

Unter Adhoc-Ensembles werdem im Projekte MuSAMA spontane Ansammlungen von Multi-Media Geräten verstanden, die sich in einem räumlich begrenztem Gebiet, wie z.B. einem Raum befinden. Diese Geräte bieten verschiedenste Dienste an. Neben den Multi-Media Geräten befinden sich in einem solchen Ensemble Sensoren, die es ermöglichen, die Umgebung zu erfassen. Mit Hilfe der Sensordaten können dann Intentionen der im Raum befindlichen Personen ermittelt werden.

Anhand der Intentionen und den Möglichkeiten der Geräte, soll es möglich werden, dem Nutzer zielgerichtet komplexe Funktionen des Ensembles anzubieten. Um das zu erreichen, müssen die vorhandenen Dienste miteinander komponiert werden.

In dieser Dissertation sollen verschiedene Kompositionsmöglichkeiten betrachtet und abgewägt werden. Dabei wird sich auf einen Planungsansatz konzentriert, in dessen Mittelpunkt das Konzept des Hierarchical Task Network (HTN) Planning stehen soll.

Bisherige Ansätze betrachteten vor allem eng aneinander gekopplte Ensembles, in denen die beteiligten Geräte bekannt waren. Ein Schwerpunkt der Arbeit ist daher die Betrachtung der Anforderungen und Konsequenzen erweiterter Dynamik. Im konkreten Fall geht es darum, was z.B. geschieht, wenn ein neues Gerät mit neuen Funktionen in das Ensemble kommt, Geräte ausfallen oder sonstige unvorhergesehen Ereignisse eintreten. Neben der Dynamik, lässt sich hier die notwendige Interoperatbilität der Geräte als weiterer wesentlicher Punkt identifizieren. Nur wenn die erforderlichen Semantiken und Strukturen in allgemein gültigen und verbreiteten Formaten und Standards formuliert werden, kann echte Dynamik auch in Bezug auf die Austauschbarkeit der Komponenten erreicht werden.

Komposition und Orchestrierung der Dienste sollen in einem mobilen Broker-Dienst realisiert werden. Es wird also in dieser Dissertation von einer zentralen Koordinierung des Andoc-Ensembles ausgegangen.

Mit dem Ziel, die beschriebene eforderliche Dynamik zu erreichen, soll in der Dissertation untersucht werden, wie die Techniken und Standards des Semantic Web in Adhoc Ensembles genutzt werden können.

Um auf Grundlage von Nutzerintentionen automatisiert planen zu können, wird es weiterhin nötig werden, eine Formalisierung von möglichen Nutzerintentionen zu ermöglichen. Da bis jetzt keine allgemeine Sprache zur Formalisierung solcher Intentionen existiert muss sich im Rahmen der Dissertation damit beschäftigt werden.

Global Cooperation through Self-Organization, Learning, and Artificial Intelligence

Universität Rostock
Graduiertenkolleg MuSAMA
Ulf Ochsenfahrt
ulf.ochsenfahrt@uni-rostock.de

This dissertation investigates cooperation strategies based on (1) minimal modeling, (2) automatic learning of ontologies, and (3) evolution. This approach can best be illustrated by a short example. Consider a laboratory room containing light sources, window blinds, and brightness sensors. How should the light sources and window blinds be controlled to achieve a particular room lighting goal?

1. *Minimal Modeling*

Previous approaches often build elaborate ontologies to model all the events happening in the room. In this case, this would involve a complete physical room and light distribution model. Computer graphics can provide the second, e.g. with a radiosity algorithm, but experience shows that predictions according to these are computationally rather expensive, possibly prohibitively expensive due to lack of computational resources in the room.

2. *Automatic Learning of Ontologies*

If this problem is solved with minimal modeling, the reality is still complex. The software thus might have an implicit world model, but this is formed through self-organization and learning. If so, it might be possible to infer or learn an explicit world model from the implicit, learned model.

3. *Cooperation through Evolution*

In the real world, cooperation can be observed in various situations and circumstances. The field of bionics is concerned with copying these strategies and bringing them into the technological toolkit of the engineer. One such strategy is simulated evolution. A straight-forward application of simulated evolution can solve the example described above. An initial lamp configuration is guessed. Then, iteratively, the lamp configuration is changed, and, if the new configuration yields brightness values closer to the goal, the new configuration is kept.

Although evolution can solve the laboratory room example quite robustly, it also has problems. In particular, the cost of guessing a particular room configuration is high, because it can significantly interfere with the current use of the room, such as a meeting. It is a part of this dissertation to identify and formalize these problems and find ways to solve or circumvent them.

Usability Tests in Mobile Environments

Universität Rostock
Graduiertenkolleg MuSAMA
Stefan Propp
stefan.propp@informatik.uni-rostock.de

Usability Tests are typically carried out in a static experimental environment, where the test user is located at a fixed position. In a smart environment as established within the MuSAMA Project the user changes the position leading to some difficulties for the observation. To overcome these obstacles the planned work deploys the technique of Remote Usability Testing in combination with a model-based approach.

The vision of smart appliance ensembles combines everyday appliances and environments to form an ensemble. Individual capabilities of the appliances are composed to provide the user with more complex features. The user intention is predicted and the according action is accomplished automatically. As a result the user can move freely in a mobile fashion, which may lead to a change of the context. An example is a movement of a user from the beamer surface to the whiteboard, which expresses a change in the user intention. Furthermore while the user is carrying out the intended task, he might switch the utilized device to accomplish the task more comfortable.

This illustrates roughly the obstacles introduced in a mobile environment. Remote usability testing can cope with it, capturing the user interaction at the specific devices to transmit the data to the device of the usability expert. Hence the test user and the expert may act at different place and different time. For the analysis the captured data is lifted to a higher level of abstraction, by the employment of task models, which describe the user task. A commonly used notation are the ConcurTaskTrees (CTT) suggested by Paterno. The interaction of the user is reflected by a path through the corresponding task model.

The research effort is planned to focus on both, the analysis of such task paths and to elaborate whether a visualization can support the usability analysis. The analysis of a task path indicates usability problems, according to aspects of time and structure. The analysis comprises for instance the investigation of differences in the duration of tasks, typical usage patterns and implications for the user interface. A visualization of the task paths may be comfortable for usability experts. A further approach could emphasize on the needs of end users to ensure a better user acceptance based on the subjective sense to keep the control in the users hands. The overall effort focusses on concepts that provide a better acceptance for the mobile environment.

Ad-hoc cooperation of multimedia appliances

University of Rostock
Graduate School MuSAMA
Christiane Reisse
christiane.reisse@uni-rostock.de

The aim of research in the field of smart environments is to explore techniques which allow us to build ensembles of devices in such a fashion that the user can interact with them in a natural way, not feeling disturbed by having to manually control all the devices. This is a challenge, especially when it comes to ad-hoc ensembles which exist only for a limited timespan and where appliances can be added or removed dynamically.

There are several projects concerned with the development of solutions for smart environments, for instance Microsoft's Easy Living project [BMK$^+$00] or the MavHome project [CD04]. However, the systems developed in these projects either rely on pre-defined rules for the behaviour of the ensemble or they learn by observing the users' actions. Therefore, none of these approaches is feasible for truly dynamic environments as the ensemble members as well as the users may never have worked together before.

Clearly, in dynamic ensembles several requirements must be fulfilled. The system must be able to produce coherent behaviour with no training, it has to be robust and flexible and, as we have no prior knowledge about how many members will join the ensemble, it should be scalable. Due to possibly limited processing capabilities of the appliances, we cannot rely on a central device which will be capable of doing all the coordination work. Therefore the system should be distributed and algorithms should be efficient. Another important issue is the limited bandwidth of communication channels between the ensemble members. Thus, the devices should avoid any unnecessary communication.

One approach to tackle those issues might be to view the devices as agents which have goals and act autonomously, only engaging in communication with other agents if they cannot get by on their own. This way, intelligent behaviour may emerge automatically. In current research, strategies like generalized partial global planning, goal-based cooperation and amorphous computing (e.g. swarm intelligence) are being explored. This is the point where investigations for this thesis will start.

References

[BMK$^+$00] Barry Brumitt, Brian Meyers, John Krumm, Amanda Kern, and Steven A. Shafer. EasyLiving: Technologies for Intelligent Environments. In *HUC*, pages 12–29, 2000.

[CD04] Diane Cook and Sajal Das. *Smart Environments: Technology, Protocols and Applications (Wiley Series on Parallel and Distributed Computing)*. Wiley-Interscience, 2004.

Intelligent and Distributed Content Addressable Memory

University of Rostock
DFG Graduate School MuSAMA
Henry Ristau[*]
henry.ristau@uni-rostock.de

Distributed systems have to manage much data. The central question of this PhD thesis is how this data can be provided to the right subsystems at the right time in the right format even if this very data is not yet existing in the system but must be processed from different existing data elements first. To solve this question, a combination of data distribution and data processing where the data is processed while it is communicated inside the system will be subject of this thesis.

Parallel data distribution and processing is interesting for a variety of distributed systems like sensor networks, that have to deal with large amounts of very simple input data that has to be processed to gain complex knowledge about the environment and control actuators based on these information. Another example are workflow systems where stepwise information distribution and processing at the same time according to a well defined workflow is the main task of the system.

The state of the art provides many different means of data distribution. These include routing algorithms in fixed and ad-hoc networks, distributed databases and a large variety of peer-to-peer networks as well as concepts like virtual shared memory, tuple spaces and publish/subscribe. Also methods for distributed data processing are available like cluster- and grid computing. But in current research the problems of data distribution and data processing have only been considered seperately as two independent layers in distributed system design. This lead to the availability of very reliable and very effective concepts and algorithms in both areas connected through inferfaces. This layered structure anticipates the modeling of data processing while the data is communicated through the system.

This thesis provides a new architectural metaphor for combining data distribution and data processing, focussing on dynamic decisions on where data shall be processed and along which paths it will be communicated. Furthermore a taxonomie for distributed systems has been established, that enables the grouping of distributed systems according to their employment of data distribution and processing.

To better understand these groups of distributed systems, requirements and criteria will be identified that enable the matching of different algorithms to the special needs of each group of distributed systems and allow the comparison of such algorithms to one another. To allow this kind of comparison and analysis, the development of a simulation environment has been started and provided first results.

Visuelle Oberflächen in Smart Ensembles

Universität Rostock
Graduiertenkolleg MuSAMA
Conrad Thiede
conrad.thiede@uni-rostock.de

Die Generierung intuitiver visueller Schnittstellen ist in starkem Maße kontextabhängig. Verschiedene Einflussfaktoren entscheiden darüber, ob das kommunikative Ziel, das mit der visuellen Repräsentation verbunden ist, auch tatsächlich erreicht werden kann. In multimedialen adhoc-Umgebungen verändern sich diese Einflussfaktoren dynamisch und verändern zudem ihre Relevanz. In der Literatur liegt der Schwerpunkt auf den Einflussfaktoren, die sich aus den Zielen und dem Gegenstand der visuellen Repräsentation ergeben. Einflussfaktoren die dem Kontext einer visuellen Repräsentation zugeordnet werden können wurden bisher nur am Rande behandelt.

Ein Schwerpunkt dieser Arbeit liegt auf den Einflussfaktoren, die sich durch die Ressourcen ergeben, die in multimedialen adhoc-Umgebungen bereit stehen. Ziel ist es, ein Modell zur Generierung angepasster visueller Schnittstellen in heterogenen, adhoc-Umgebungen zu entwickeln. Auf der Basis des Modells soll eine dynamische Auswahl von relevanten Informationen vorgenommen sowie eine geeignete Darstellungstechnik adaptiv ausgewählt werden. Weiterhin soll es möglich sein, dass die als relevant eingestuften Einflussfaktoren die Verteilung der visuellen Repräsentation steuern.

In der Umsetzung sollen die Darstellungstechniken durch adaptierbare, verteilbare Dienste realisiert werden. Dabei übernimmt ein Mediator die intelligenten Entscheidungen. Dadurch können die Dienste kompakt und einfach gehalten werden.

Specification and Design of Context-Sensitive and Cooperative Work in Dynamic Environments

University of Rostock
Graduate School MuSAMA
Maik Wurdel
maik.wurdel@informatik.uni-rostock.de

Dynamic environments are in principle complex systems, since diverse actors (technical devices and users) can dynamically appear and disappear. Thus the analysis and design phase of such an environment is highly important to assure an expedient behavior of the complete system. One way to specify the behavior of actors are task models, which describe the tasks and activities an actor has to execute to achieve a certain goal. In recent years task models have been widely used for model-based UI development, since they are able to specify the activities of users using an interactive system to accomplish a certain task in early stages of development.

We propose to use this kind of model in dynamic environments to design the possible behavior of components within the environment. One of the major requirements of the design of dynamic environments is the capability of how to cope with cooperative work. We already demonstrated that task models are able to specify cooperative work. The simulation of the environments and their actors specified by task models will be examined as well, because such a simulation can expose and visualize design mistake made in early stages of development with low costs.

By using this approach to design dynamic environments model transformations are an essential tool. Weaving of models can be done in two different ways:

1. Top-Down. During the design one task model is created which describes the behavior of the complete system. The solution we are looking for is an appropriate assignment of tasks of the task model to actors within the environment to achieve a sound behavior.

2. Bottom-Up. Task models are designed which express the features and capabilities of the actors in the environment. The aim is one task model which specifies the behavior of the complete system.

These different approaches would allow us to specify the system in a flexible manner starting either with one task model or with a set of model fragments. A hybrid approach is considerable as well.

Another field of research in this area is the interface between task models as one way to specify such a system and bayesian inference models as execution models. Our group already made some effort in transforming annotated task models into markov chains. Further examination of this field is needed to prove the soundness of this approach.

Context-based Resource Management

University of Rostock
Research Training Group MuSAMA
Jiaxi You
jiaxi.you@uni-rostock.de

The goal of the context-based resource management is the efficient monitoring, arbitration and allocation of smart environment resources with consideration of the context information, which indicates the participants of the ensemble together with the surrounding infrastructure. Such ensemble is to work on a quantity of tasks according to the user requirements using available resources, whose range is constrained to device deployment, summable resources such as bandwidth, energy consumption or computational costs for this work. Resources are in this context, at least temporarily, a scarce property, and stand critical for the performance of an ensemble.

Wireless Sensor Networks (WSNs) will be the primary option for observing the physical aspects of an environment, and thus the main target that our resource management mechanism serves. Due to the dynamic natures of their deployment and the limited power that sensor nodes can store or harvest, various optimization mechanisms have been developed for the robustness, security and longevity of WSNs. In the first working phase, existing resource optimization strategies on device level are to be examined. Afterwards an ensemble-view will be emphasized to integrate and coordinate promising strategies with the given context of an environment, in order to identify competitive strategies of devices, and adapt them when necessary. An efficient use of resources involves (1) trade-off between cost and usage of resources, particularly by shifting workloads from scarce resources to non-scarce ones when possible, (2) avoidance of temporary and local maximum and (3) resolution of conflicting aims between tasks of ensemble during resource allocation. The basis of the context-based resource-efficiency is a precise resource monitoring, which helps to supply information about the available resources of an ensemble at any time. Definition of the global ideal condition of an ensemble, as well as suitable representation techniques also support this work, and will be studied in the MuSAMA project together with other related topics. The following sub-questions shall be investigated in the thesis: (1) How do mobile devices recognize the resources of a smart ensemble, and to reliably predict the demands of resources for a particular task? (2) How to balance the trade-offs between competitive optimization strategies, and to which criteria an efficient treatment of a task is determined? (3) How can parasitic behaviors of individual ensemble participants be recognized, contained and prevented?

Kontextorientierte Gerätekommunikation

Universität Rostock
Graduiertenkolleg MuSAMA
Raphael Zender
raphael.zender@uni-rostock.de

Kontextbezogene Anwendungen müssen in die Lage versetzt werden, auf Änderungen in der Umgebung zu reagieren. Im Gegensatz zu einer sender- oder empfängerinitiierten Kommunikation werden die Sender und Empfänger durch das Ereignis bestimmt, d.h. es liegt eine ereignisinitiierte Kommunikation vor. Um die Ereignisbeobachtung unabhängig von den konkreten Sensorsystemen und der Verarbeitung von Sensordaten in einer heterogenen Netzumgebung gestalten zu können, sind adäquate Abstraktionen zur Erfassung von Ereignissen notwendig.

Im vorliegenden Forschungsvorhaben werden die Fähigkeiten und Ressourcen von Geräten innerhalb von Geräte-Ensembles anderen Geräten in Form von Diensten angeboten. Durch das aus Dienstekonzept (Service Oriented Architecture, SOA) wird von den konkreten Geräten abstrahiert.

Anwendungen, die Sensordaten nutzen, müssen somit nicht an konkrete Sensorsysteme angepasst werden, sofern diese ihre Messungen in Form von Sensordiensten anbieten. Anhand der von der Anwendung benötigten Sensordaten wie beispielsweise der Helligkeit, der Lautstärke oder der Identifizierung der Personen in einem Raum werden Sensordienste transparent ausgewählt und genutzt. Aufbauend auf diesem Abstraktionskonzept werden verschiedene Kommunikationsstrukturen für die dienstebasierte Unterstützung kontextbezogener Anwendungen untersucht.

Beispielsweise können Anwendungen sich über den Sensordienst registrieren, um beim Eintreten bestimmter Sensorwerte (Ereignisse) benachrichtigt zu werden. Das Auffffinden eines des Sensordienstes und die Registrierung erfolgen über das SOA-Konzept. Sobald eine Anwendung sich bei dem Dienst registriert hat, stehen Sensor und Anwendung in einem Publisher-Subsriber-Verhältnis. Dadurch wird der Flaschenhals eines zentralen Brokers umgangen, sobald die Registierung durchgeführt wurde. Die Ableitung von Situationen aus den Sensordaten ist in diesem Fall Aufgabe der Anwendung. Eine Middleware, die die Auswahl von Sensordiensten, die Registrierung bei diesen und die Ableitung von Situationen für die Anwendung übernimmt, vereinfacht dadurch die Entwicklung kontextbezogener Anwendungen. Die Konzeption und Implementierung einer derartigen Middleware ist das Ziel dieses Forschungsvorhabens.

Aspekte wie die Semantik von Ereignissen, die Heterogenität der Netzinfrastruktur eines Geräte-Ensembles und Quality of Service (QoS)-Anforderungen beim Nutzen von Diensten werden explizit berücksichtigt. Außerdem wird ein Konzept entwickelt, um Dienste auch außerhalb des lokalen Geräte-Ensembles zur Verfügung zu stellen.

Forschungskolleg

Service-oriented Systems Engineering

Berlin

Research College "Service-Oriented System Engineering" at Hasso Plattner Institute

Design and Implementation of service-oriented architectures imposes a huge number of research questions from the fields of software engineering, system analysis and modeling, adaptability, and application integration. Component orientation and web services are two approaches for design and realization of complex web-based system. Both approaches allow for dynamic application adaptation as well as integration of enterprise application.

Service-Oriented Systems Engineering represents a symbiosis of best practices in object-orientation, component-based development, distributed computing, and business process management. It provides integration of business and IT concerns.

Hasso Plattner Institute (HPI) for Software Engineering is a privately financed An-Institute at University Potsdam. Since October 2005, HPI hosts the research college on "Service-Oriented Systems Engineering", which is modeled after the DFG graduate schools. The HPI research college currently has 10 PhD students and one postdoc.

The Professors of the HPI with their research group are supporting pillars for our PhD school. With its interdisciplinary structure, the research college on "Service-Oriented Systems Engineering" interconnects the HPI research groups and fosters close and fruitful collaborations.

"Service-Oriented Systems Engineering" denotes a current research topic in the field of IT systems engineering with high potential in academic research as well as in industrial application. Supported by an internationally renowned grant, PhD students at our college participate in joint activities such as lectures, seminars, winter schools and workshops.

In context of the research college, the different groups at HPI work on the following topics:

- Service-Oriented Geovisualization Systems (Prof. Dr. Jrgen Dllner)
- Tools and Methods for Software Engineering in Service-Oriented Systems (Prof. Dr. Robert Hirschfeld)
- Security Engineering of Service-Based IT Systems (Prof. Dr. Christoph Meinel)
- Formal Semantics of Service-Oriented Systems (Prof. Dr. Felix Naumann)
- Evolutionary Transition of Enterprise Applications to Service-Orientation (Prof. Dr. h.c. Hasso Plattner)
- Operating System Abstractions for Service-Oriented Computing (Prof. Dr. Andreas Polze)
- Services Specification, Composition, and Enactment (Prof. Dr. Mathias Weske)
- Quantitative Evaluation and Optimization of Service-Oriented Systems (Prof. Dr. Werner Zorn)

Der Entwurf und die Realisierung dienstbasierender Architekturen wirft eine Vielzahl von Forschungsfragestellungen aus den Gebieten der Softwaretechnik, der Systemmodellierung und -analyse, sowie der Adaptierbarkeit und Integration von Applikationen auf. Komponentenorientierung und WebServices sind zwei Anstze fr den effizienten Entwurf und die Realisierung komplexer Web-basierender Systeme. Sie ermglichen die Reaktion auf wechselnde Anforderungen ebenso, wie die Integration groer komplexer Softwaresysteme.

Service-Oriented Systems Engineering reprsentiert die Symbiose bewhrter Praktiken aus den Gebieten der Objektorientierung, der Komponentenprogrammierung, des verteilten Rechnen sowie der Geschftsprozesse und bercksichtigt auch die Integration von Geschftsanliegen und Informationstechnologien.

Das Hasso-Plattner-Institut (HPI) fr Softwaresystemtechnik ist ein stiftungsfinanziertes An-Institut an der Universitt Potsdam und betreibt seit Oktober 2005 mit dem Forschungskolleg "Service-Oriented Systems Engineering" eine Doktorandenschule nach Vorbild der DFG-Graduiertenkollegs. Momentan sind 10 Doktoranden und ein PostDoc Im Forschungskolleg aktiv.

Tragende Sulen des Forschungskollegs bilden die Professoren des HPIs mit ihren Arbeitsgruppen. Mit seiner Querschnittsstruktur verbindet das Forschungskolleg "Service-Oriented Systems Engineering" die Fachgebiete des HPI und sorgt fr eine enge und fruchtbare interdisziplinre Kooperation.

Mit dem Thema "Service-Oriented Systems Engineering" widmet sich das Forschungskolleg einer aktuellen Thematik aus dem IT Systems Engineering, mit einem besonderen Potenzial sowohl in der akademischen Forschung als auch in der industriellen Anwendung. Mit einem international vergebenen Stipendium ausgestattet, nehmen die Doktoranden teil an regelmigen Kolleg-Veranstaltungen wie Vorlesungen, Seminaren, Winterschulen und Workshops.

Die einzelnen Fachgebiete beteiligen sich am Forschungskolleg mit folgenden Themen:

- Service-Oriented Geovisualization Systems
  (Prof. Dr. Jrgen Dllner)
- Tools and Methods for Software Engineering in Service-Oriented Systems
  (Prof. Dr. Robert Hirschfeld)
- Security Engineering of Service-Based IT Systems
  (Prof. Dr. Christoph Meinel)
- Formal Semantics of Service-Oriented Systems
  (Prof. Dr. Felix Naumann)
- Evolutionary Transition of Enterprise Applications to Service-Orientation
  (Prof. Dr. h.c. Hasso Plattner)
- Operating System Abstractions for Service-Oriented Computing
  (Prof. Dr. Andreas Polze)
- Services Specification, Composition, and Enactment
  (Prof. Dr. Mathias Weske)
- Quantitative Evaluation and Optimization of Service-Oriented Systems
  (Prof. Dr. Werner Zorn)

Speaker of the research college on "Service-Oriented Systems Engineering" is:

Prof. Dr. Andreas Polze
Operating Systems and Middleware Group
Hasso-Plattner-Institut for Software Engineering
Postfach 900460
D-14440 Potsdam

andreas.polze@hpi.uni-potsdam.de
http://kolleg.hpi.uni-potsdam.de

Adaptive Transport Architecture for Service-oriented Systems

Flavius Copaciu

Research College "Service-Oriented System Engineering"

at Hasso Plattner Institute

The research of Flavius Copaciu focuses on relevant aspects of the communication infrastructure in service-oriented systems, under the supervision of Prof. Dr.-Ing. Werner Zorn, head of the "Communication Systems Group".

1 Vision

Web services, as implementation of service-oriented architectures, are transport agnostic although most of the them run over HTTP. The transport independence of web services enables the use of optimized transport protocol, taking into account the state of the network, service consumer and service provider, and makes them the ideal candidate in the efforts for making better use of the existing infrastructure by means of a context-aware, adaptive architecture.

2 Current Research

The research activity consists of two main parts: (1) design and implementation of an Adaptive Transport Architecture (ATrA) for web services, based on the relevant characteristics of the existing transports and (2) developing a methodology for performance prediction in order to support the adaptive nature of the solution.

Adaptive Transport Architecture: One way of improving service-oriented systems is to make better use of the available communication infrastructure. This can be done by using an architecture capable to adapt to specific network conditions as well as to the current state of the systems where the service consumer and service provider are running.

This architecture enables providers and consumers to make context aware decision regarding their selection of the transport stack used for web service invocations and to define and implement highly abstracted profiles to govern their interaction (ex. fastest possible invocation, minimum server load, etc.)

Performance prediction: In order to support the adaptive nature of the solution, work is currently undergoing to implement the performance estimation mechanism. FMC-QE - a new approach for quantitative evaluation, has been employed and the Axis2 framework for web services, has been modeled using it. FMC-QE has proved well suited to capture the hierarchical and modular structure of Axis2.

Service-based Virtual 3D City Models

Benjamin Hagedorn
Research College "Service-Oriented System Engineering"
at Hasso Plattner Institute

Benjamin Hagedorn deals with service-based provision, processing, and visualization of geoinformation at the "Computer Graphics and Visualization" research group of Prof. Dr. Jürgen Döllner.

1 Vision

Geoinformation is a good which is becoming more and more important for a variety of application domains such as disaster management, urban security, and logistics and telematics. The combination of virtual 3D city models and service-orientation enables the integration of distributed geoinformation, and the provision of high-level geoinformation related capabilities to a human user or for the integration into business processes.

2 Current Research

Current work is investigating the applicability of virtual 3D city models as a platform for the integration of complex geoinformation and the question of how to provide this high-level capability in a service-based manner.

Service-based Integration of Complex Geoinformation: In the context of the Open Geospatial Consortium Web Service Initiative Phase 4 an OpenGIS web service enabled 3D viewer client was contributed which is capable of integrating geoinformation from different domains (GIS, CAD, BIM) and so supporting new insight into space related information. Here, integration is done at data level by the usage of low-level geoservices (WMS, WFS) and CityGML as an exchange format and at the visualization level by the combination of different types of geoinformation (e.g., terrain data, aerial image, and different building models) in one virtual 3D city model.

Construction of High-Level Geoinformation Services: Going one step further, the provision of high-level geoservices on the basis of virtual 3D city models is investigated. High-level geoservices (1) integrate complex geoinformation, (2) add value to geoinformation, (3) provide specific business functionality, (4) provide high-quality geovisualization , (5) support user-interaction, or (6) support user-, task-, and device-related context awareness. These high-level tasks can be mapped to geovisualization instruments which are views (on data), styles (e.g., sketchiness), rendering constraints (e.g., display size), or interaction capabilities (e.g., stylus-based input) and their consideration in the high-level geoservices information and behavior model.

Ontology-based Matching for Service-oriented Systems

Frank Kaufer

Research College "Service-Oriented System Engineering"

at Hasso Plattner Institute

Frank Kaufer works on formal semantics of service-oriented systems and especially on matching service requests and advertisements based on formal knowledge models called ontologies. Furthermore he is member of the research group *Information Systems* headed by Prof. Dr. Felix Naumann, where he investigates the employment of services and ontologies for information integration.

1 Vision

Service-oriented systems constitute a paradigm that aims for the development of software systems which are no longer rigid and platform-dependent programs, but rather consist of loosely coupled and distributed components. Such components are called services, because they are considered independent from how or where they are implemented, but only in terms of their provided service (functionality) to other components or the user. To enable the loose and dynamic coupling of services, overcoming the initial heterogeneity, the service providing components have to be specified in a homogeneous way.

One approach addressing this is to use so called ontologies which are formal specifications of a shared conceptualisation among a certain community and can be considerd as high-level knowledge models. In general variants of first-order logic like Description Logics or Logic Programming languages are employed for the formalisation.

2 Current Research

Service Matching: A central task in service-oriented architectures is to connect a service provider which has advertised a service with a requester demanding an equivalent or similar service. Software agents with such mediation capabilities are called matchmakers. The underlying matching of software specifications is a quite old research topic with roots in Hoare logic and a large research corpus on program verification and software component retrieval. Nevertheless, the matching of service specifications on base of ontologies promises more feasibility and automation due to the abstraction aspect which is inherent to both, services and ontologies.

The Logic Progamming based hybrid matchmaker WSMO-MX is a recent research result which combines ontology-based matching and text information retrieval.

Quantitative Modeling and Analysis of Service-oriented Architectures

Stephan Kluth

Research College "Service-Oriented System Engineering"
at Hasso Plattner Institute

Stephan Kluth's main research interest is the quantitative modeling and analysis of systems, with a focus on Service-oriented Architectures. Within the Hasso-Plattner-Institute, he works in the "Communication Systems Group" of Prof. Dr.-Ing. Werner Zorn.

1   Vision

In the research of Stephan Kluth and the Communication Systems Group a new modeling approach, FMC-QE, is under development. Starting from a hierarchical service request, the systems are modeled in three dimensions, using Entity-Relationship-Diagrams to model the service requests, Block Diagrams to model static and Petri Nets to model dynamic aspects. The performance values are calculated in a hierarchical manner. The fundamentals of FMC-QE are the results and laws of the Queueing Theory and the Time Augmented Petri Nets.

2   Current Research

Closed vs. Open Systems: In FMC-QE there is no distinction between closed and open systems, these two views are combined in a hybrid approach, where the arrival rate and the population could be defined as free parameters. This approach is under research and evaluation.

Case Studies: In order to evaluate and enhance the whole modeling approach, case studies are done. Modeled products and scenarios are:

- a service based SAP NetWeaver prototype application;
- Find.IT - a service based integration scenario;
- Apache Axis2 - a web services framework;
- fundamental research questions in the area of quantitative modeling, e.g. central server or semaphore synchronization.

Security Engineering of Service-Based IT Systems

Michael Menzel

Research College "Service-Oriented System Engineering"
at Hasso Plattner Institute

Michael Menzel deals with security related aspects of service-based systems at the "Internet Systems and Technology" research group of Prof. Dr. Christoph Meinel.

1 Vision

Service Oriented Architectures promise a seamless communication between applications, an improved reusability of components and an increased responsiveness to changing business requirements. However, these aspects are only advantageous if scalable security solutions preventing misuse are introduced as well.

Organizations providing Web Services have to securely identify service consumers across different trust domains to apply authorization decisions. These trust domains can have a different understanding of identities and trust. Federated identity management is the solution to these problems by promoting the idea of authentication and authorization as a service. This solution decouples the authentication and authorization mechanisms from applications and services by transferring the authentication/authorization decisions in a common format across different trust domains, which can be accepted by all services. The research work focuses on using federated identity mechanisms to facilitate secure service compositions in a service-oriented architecture across multiple trust domains.

2 Current Research

Research on Federated Authentication for Web Services: Research work on federated authentication deals with approaches to provide a variable choice of authentication means. The advanced approach is to replace concrete authentication requirements by a method-independent numerical value indicating the minimum trust that has to be established between collaborative partners. Of special interest is multi-factor-authentication that comprises a model to calculate the combined effect of two authentication methods taking into account the similarity of the single mechanisms.

Research on Authorisation Mechanisms in cross-domain Service Invocation: Federated authentication provides a standardize way to communicate identity information to service providers. Each participating trust domain applies some common policies to enforce ac-

cess control based on the identity information. However, the meaning of this information such as the user's role can be tied to the issuing domain and is therefore not useful for collaborative partners. The research work focuses on methods to communicate security policies and authorization enforcement in a collaborative context, without exposing sensitive information to collaborative partners that are not authorized.

Semantics of Services and Service Compositions

Harald Meyer

Research College "Service-Oriented System Engineering"
at Hasso Plattner Institute

Harald Meyer investigates semantics of services and service compositions at the "Business Process Technology" research group of Prof. Dr. Mathias Weske.

1 Vision

Service-orientation allows for the alignment of the IT to the business processes of a company. Services provide a clearly defined business functionality. Composing them enables the provision of services previously unavailable. The focus of my research is on finding suitable services and composing them as easily as possible.

The semantics specify the functionality formally and hence support the automated discovery. This is especially crucial in large service landscapes with hundreds or thousands of services. My research in this area is mainly focussed on how to get to the semantics of services and service compositions.

Research on the semantics of individual services is currently focussed on formal, logic-based approaches like OWL-S and WSML. My approach is to use an informal approach based on collaboration [MW06b]. How to automatically derive service composition semantics from the semantics of contained individual services, is mostly ignored at the moment. This is rather strange as it allows for the automated publication of service compositions and the verification of composition results. In a recent publication [Mey07], I showed how the semantics of compositions represented as Workflow nets can be derived.

2 Current Research

Automated Service Composition: Automated service composition allows for the dynamic, run-time creation of compositions based on service consumer requests. This allows for the adaption to changing consumer needs as well as changes in the service landscape. Several approaches to do automated composition exist [ZBL$^+$03, PBB$^+$04, SPW$^+$04, BCGM05]. My approach [MW06a] is based on enforced hill-climbing, a heuristic search algorithm [HN01]. It allows for the creation of parallel and alternative control flow and the dynamic creation of objects.

Semi-automated Composition: While automated composition showed interesting research results, the requirement for complete service specifications as well as possible legal issues, let to a shift of research activity to semi-automation. In [SM06, SMT06, SMW07] we presented the elements of our approach for semi-automated service composition. It is based on three features: filter services, verify compositions, and suggest partial plan. All three are implemented using results from automated service composition and the semantics of service compositions.

References

[BCGM05] Daniela Berardi, Diego Calvanese, Giuseppe De Giacomo, and Massimo Mecella. Composition of Services with Nondeterministic Observable Behaviour. In *Proceedings of the Third International Conference on Service-Oriented Computing*, volume 3826 of *Lecture Notes In Computer Science*, pages 520–526, Heidelberg, 2005.

[HN01] Joerg Hoffmann and Bernhard Nebel. The FF Planning System: Fast Plan Generation Through Heuristic Search. *Journal of Artificial Intelligence Research*, 14:253 – 302, 2001.

[Mey07] Harald Meyer. On the Semantics of Service Compositions. In *Proceedings of The First International Conference on Web Reasoning and Rule Systems (RR 2007)*, 2007.

[MW06a] Harald Meyer and Mathias Weske. Automated Service Composition using Heuristic Search. In Schahram Dustdar, JosÂ Luiz Fiadeiro, and Amit Sheth, editors, *Business Process Management (BPM 2006)*, volume 4102 of *Lecture Notes In Computer Science*, pages 81–96, Heidelberg, 2006. Springer.

[MW06b] Harald Meyer and Mathias Weske. Light-Weight Semantic Service Annotations through Tagging. In Asit Dan and Winfried Lamersdorf, editors, *Service-Oriented Computing - ICSOC 2006*, volume 4294 of *Lecture Notes In Computer Science*, pages 465–470, Heidelberg, 2006. Springer.

[PBB$^+$04] Marco Pistore, F. Barbon, Piergiorgio Bertoli, D. Shaparau, and Paolo Traverso. Planning and Monitoring Web Service Composition. In *Workshop on Planning and Scheduling for Web and Grid Services (held in conjunction with The 14th International Conference on Automated Planning and Scheduling*, pages 70 – 71, 2004.

[SM06] Jan Schaffner and Harald Meyer. Mixed Initiative Use Cases for Semi-Automated Service Composition: A Survey. In *Proceedings of the International Workshop on Service Oriented Software Engineering*, pages 6–12, New York, NY, USA, 2006. ACM Press.

[SMT06] Jan Schaffner, Harald Meyer, and Cafer Tosun. A Semi-automated Orchestration Tool for Service-based Business Processes. In *Proceedings of the 2nd International Workshop on Engineering Service-Oriented Applications: Design and Composition*, pages 54–65, 2006.

[SMW07] Jan Schaffner, Harald Meyer, and Mathias Weske. A Formal Model for Mixed Initiative Service Composition. In *Proceedings of The IEEE International Conference on Services Computing (SCC 2007)*, 2007.

[SPW$^+$04] Evren Sirin, Bijan Parsia, Dan Wu, James Hendler, and Dana Nau. HTN Planning for Web Service Composition Using SHOP2. *Journal of Web Semantics*, 1(4):377–396, 2004.

[ZBL⁺03]   Liangzhao Zeng, Boualem Benatallah, Hui Lei, Anne Ngu, David Flaxer, and Henry Chang. Flexible Composition of Enterprise Web Services. *Electronic Markets – Web Services*, 13:141–152, 2003.

Resource-Oriented Architecture

Hagen Overdick

Research College "Service-Oriented System Engineering"

at Hasso Plattner Institute

Hagen Overdick deals with resource orientation as a viable subset of service orientation at the "Business Process Technology" research group of Prof. Dr. Mathias Weske.

1 Vision

Service orientation puts a high focus upon processes as services are designed to be assembled or "mashed up" upon user demand. Instead of monolithic architectures with implicit business logic, fine-grained services are being assembled dynamically into applications by the means of processes.

Simultaneously, Web-based applications are more and more replacing classical desktop applications. The software itself becomes a service. This is made possible by modern browser technologies like AJAX. This particularly eases the deployment of applications, since the newest version is 'just a bookmark away'.

The thesis aims for a unified model of tree-structured data (XML), graph-structured knowledge (RDF), and process-orientated interaction (π-calculus) to build a foundation for next-generation application design. The utopia aimed at is to build a *processpedia*, enabling community-driven management of process knowledge, as *wikipedia* does for encyclopic content today.

2 Current Research

The research activity is currently focused on implementing a prototype of the current research before continuing on advanced topics.

Xenodot: Xenodot is a persistent Super-DOM for one's view onto the Internet. Rooting in a search engine optimization project is has turned into an application builder's tool. Xenodot deals with tree-structured data and graph-structured knowledge concurrently, providing search capabilities on both content and structure at the same time.

AReSS and Oryx: A bachelors' project is building a concrete use case scenario upon Xenodot. Special attention is put on the implications of resource orientation in process-oriented applications. At the same time, a browser-based process editor is build by a second bachelors' project. Together, they form the basis to validating the current research results

Operating System Support for Service-based Systems

Alexander Schmidt

Research College "Service-Oriented System Engineering"
at Hasso Plattner Institute

Alexander Schmidt investigates operating system (OS) related aspects in service-oriented computing infrastructures at the Operating Systems and Middleware group of Prof. Dr. Andreas Polze.

1 Vision

Server consolidation and virtualization are two of the most central approaches used in data centers worldwide that form the foundation for tomorrows world of services and service-oriented architectures (SOA). Service-oriented computing (SOC) often relies on the assumption that services are independent with respect to resource usage and data dependencies. Although this assumption might hold for widely distributed scenarios (i.e., the Grid) it will not hold for service computing in the data center. Within these environments, the likelihood of having several—or even all—services ending up on the same server machine is quite significant. To cope with the interdependencies among services, awareness for service computing on the OS level is badly needed. The research focuses on the following topics:

2 Current Research

Monitoring: Monitoring services inside the OS allows monitoring the service's execution flow on a very fine level of granularity. The challenge here is (1) how to gather OS events without decreasing service's performance to a significant extent and (2) to map these events to an execution flow of a service.

Distributed Debugging: Monitoring events only on a local machine is insufficient for today's world of distributed applications. Thus, research concentrates on the aggregation of such local events monitored on several machines. The aim is to assist developers to evaluate their applications and to identify performance bottlenecks and the reason thereof.

Adaptive Resource Management: Basing on the results of distributed debugging or additional provided meta-information, the OS might adapt its resource scheduling policies for the CPU, the memory, communication bandwidth, etc., in order to enhance the performance of a particular service.

Operating System Resource Management for Service-based Systems

Michael Schïel

Research College "Service-Oriented System Engineering"
at Hasso Plattner Institute

Michael Schïel deals with operating system related aspects of service-based systems at the "Operating Systems and Middleware" research group of Prof. Dr. Andreas Polze.

1   Vision

From an operating system point of view, service-based computing can be seen as the execution of cooperating tasks or processes. Special properties or (meta-) information concerning the structure of such service-based systems are ignored on the operating system level most of the time.

More appropriate operating system abstractions can lead to optimized server implementations by exploiting special properties and knowledge about the structure of service-based systems. Optimizations are possible in the areas of scheduling, memory management or virtualization concepts.

2   Current Research

The research activity is split in two parts: (1) identifying special properties of service-based systems and (2) exploiting these properties on operating system level.

Monitoring Kernel: Monitoring (server) applications from an operating system point of view show how high-level concepts like workflows or service composition can be mapped to low-level execution flows. Communication and synchronization activities are particularly interesting.

A special operating system kernel which can monitor arbitrary applications was developed. Additionally, tools were created to visualize and analyze the results.

Resource Partitioning: For optimized execution of concurrent activities (e.g. inside a server application) fine-grained resource provisioning can be useful. Also, certain service-level agreements can be ensured by using guaranteed resource shares for specific activities.

Operating system kernel extensions to support low-level resource partitioning were developed in context of the "Windows Research Kernel" (WRK).

Workflow Support for User-Centered Design

Matthias Uflacker

Research College "Service-Oriented System Engineering"

at Hasso Plattner Institute

Matthias Uflacker investigates user-centered software design processes at the "Enterprise Platform and Integration Concepts" group of Prof. Dr. h.c. Hasso Plattner.

1 Vision

A rich user experience adapted to the needs and goals of end-users has become a differentiating factor for complex enterprise applications [Ufl07]. Integrated into the development process, user-centered design (UCD) provides methods to create value through usability and end-user satisfaction. While simplicity is the ultimate goal of UCD [UB07], early user involvement, contextual research, prototyping, and constant evaluation are the core facets.

Human communication and collaboration in knowledge-intense UCD processes introduces additional complexity in the development of interactive applications. Considerable amounts of user data, insights, and requirements must be aggregated and communicated across a multi-disciplinary design team for further refinement. Easy and effective coordination of collaborative design activities is critical for the UCD outcome. Knowledge management and workflow systems are required, which can adapt to the flexible and variable nature of creative design activities.

2 Current Research

Methods for information aggregation and syndication appropriate to distributed, large-scale software design environments are being investigated. In the focus of this research are resource-oriented architectures, providing a well-suited platform for flexible knowledge accumulation and representation in design teams. Coordinated by personal "ad-hoc" workflows (cases), a resource-oriented design information framework is envisioned, that presents context-relevant information to process participants such as designers, programmers, management, and end-users [OU07].

References

[OU07] Hagen Overdick and Matthias Uflacker. A REST-ful Approach to Workflow Support in Collaborative Design Communities. In review. 2007.

[UB07] Matthias Uflacker and Daniela Busse. Complexity in Enterprise Applications vs. Simplicity in User Experience. In *HCII'07: Proceedings of the 2007 Human-Computer Interaction Conference*, 2007.

[Ufl07] Matthias Uflacker. The Role of Design in Enterprise Application Development: Challenges and Chances Imposed Through Service-Orientation. Technical report, Hasso-Plattner-Institut, 2007.