Dagstuhl 2010

Proceedings of the Joint Workshop of the German Research Training Groups in Computer Science

Editor: *DFG Research Training Group 1298 "AlgoSyn"*
RWTH Aachen University

Proceedings of the Joint Workshop of the German Research Training Groups in Computer Science

May 31 - June 2, 2010

Proceedings of the Joint Workshop of the German Research Training Groups in Computer Science

May 31 - June 2, 2010

DFG Research Training Group 1298 AlgoSyn
RWTH Aachen University

Bibliografische Informationen der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Proceedings of the Joint Workshop of the
German Research Training Groups in ComputerScience
May 31 - June 2, 2010

1. Auflage Januar 2010

ISBN-10: 3-86130-146-6

Vertrieb:

© Verlagshaus Mainz GmbH Aachen
Süsterfeldstr. 83, 52072 Aachen
Tel. 0241/87 34 34
Fax 0241/87 55 77
www.Verlag-Mainz.de

Herstellung:

Druck und Verlagshaus Mainz GmbH Aachen
Süsterfeldstr. 83, 52072 Aachen
Tel. 0241/87 34 34
Fax 0241/87 55 77
www.Verlag-Mainz.de
www.DruckereiMainz.de
www.Druckservice-Aachen.de

printed in Germany

Preface

For several years the German Research Training Groups in computer science have met at Schloss Dagstuhl - the Leibniz Center for Informatics, the world's premier venue for research seminars in computer science. This joint workshop, funded by the Deutsche Forschungsgemeinschaft (DFG), aims to provide a platform for the exchange of knowledge and the discussion of research findings.

Every year one Research Training Group is responsible for the planning and organization of this event. For the current workshop in 2010, this is done by the Research Training Group AlgoSyn[1] from Aachen. In order to provide as much information as possible about the participating Research Training Groups every scholarship holder was asked to submit an abstract describing briefly his/her research topic. People who give a talk during the workshop got the opportunity to submit extended abstracts. Additionally, the Research Training Groups were invited to shortly introduce themselves.

Sincere thanks to everybody who has contributed to the creation of these proceedings, especially the abstracts' authors. Furthermore, special thanks to the DFG for the yearly funding of the joint workshops.

The AlgoSyn proceedings team

Kai Bollue
Dominique Gückel
Ulrich Loup
Jacob Spönemann
Melanie Winkler

[1] Further information about AlgoSyn ("Algorithmic synthesis of reactive and discrete-continuous systems"), its research topics, and researchers can be found in these proceedings.

Contents

Organizational Information

List of Speakers

Amna Asif	GRK 1076/3 Oldenburg
Vaishak Belle	GRK 643 Aachen
Markus Brückner	GRK 1487 Ilmenau
Ahmad El Maamoun	GRK 1076/3 Oldenburg
Andreas Gaiser	GRK 1480 München
Joanna Geibig	GRK 1324 Berlin
Christian Gerth	IGS Paderborn
David Gassmann	GRK 1424 Rostock
Paul Hänsch	GRK 1298 Aachen
Dominik Haumann	GRK 1362 Darmstadt
Jan Hoffmann	GRK 1480 München
Masha Jenabi	GRK 1042 Konstanz
Aleksandr Karbyshev	GRK 1480 München
Thomas Kemmerich	IGS Paderborn
Philipp Lehsten	GRK 1424 Rostock
Stefan Leye	GRK 1387/1 Rostock
Michael Perscheid	HPI Potsdam
Christian Reinl	GRK 1362 Darmstadt
Liz Ribe-Baumann	GRK 1487 Ilmenau
Mirco Richter	GRK 1042 Konstanz
Christoph Roth	GRK 1194 Karlsruhe
Thorsten Sattler	GRK 643 Aachen
Johannes Schmid	GRK 1194 Karlsruhe
Florian Schmidt	GRK 643 Aachen
Hans-Jörg Schulz	GRK 1387/1 Rostock
Thomas Vogel	HPI Potsdam
Arif Wider	GRK 1324 Berlin
Melanie Winkler	GRK 1298 Aachen

Tentative Agenda

<table>
<tr><th>Time</th><th>Monday</th><th>Tuesday</th><th>Wednesday</th></tr>
<tr><td>9:00</td><td>GRK 1298</td><td>GRK 1480</td><td>GRK 1424</td></tr>
<tr><td>9:45</td><td>GRK 1076/3</td><td>IGS Paderborn</td><td>HPI Potsdam</td></tr>
<tr><td>10:30</td><td>Break</td><td>Break</td><td>Break</td></tr>
<tr><td>11:00</td><td>GRK 1387/1</td><td>GRK 1324</td><td rowspan="2">Tutorials</td></tr>
<tr><td>11:45</td><td>GRK 1487</td><td>GRK 1362</td></tr>
<tr><td>12:30</td><td>Lunch</td><td>Lunch</td><td>Lunch</td></tr>
<tr><td rowspan="3">14:00</td><td rowspan="3">Tutorials</td><td>Talk: DFG</td><td>Departure</td></tr>
<tr><td>Talk: Plagiarism</td><td></td></tr>
<tr><td>Break</td><td></td></tr>
<tr><td>16:00</td><td>Break</td><td rowspan="2">Workshop</td><td></td></tr>
<tr><td>16:30</td><td>GRK 1194</td><td></td></tr>
<tr><td>17:15</td><td>GRK 643</td><td>GRK 1042</td><td></td></tr>
<tr><td>18:00</td><td>Dinner</td><td>Social Event</td><td></td></tr>
</table>

1 GRK 643: Software for Mobile Communication Systems

Prof. Dr. Spaniol
Email: spaniol@informatik.rwth-aachen.de
RWTH Aachen
Internet: `http://www-i4.informatik.rwth-aachen.de/Kolleg/`

The Research Training Group is an interdisciplinary research project at the RWTH Aachen University lead by a group of ten professors representing three different disciplines 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.

1.1 Multi-Agent Only-Knowing Revisited

Vaishak Belle (belle@cs.rwth-aachen.de)
Supervisor: Gerhard Lakemeyer

Levesque's notion of only-knowing is a single agent modal logic that was proposed with the intention of capturing certain types of nonmonotonic reasoning [1]. Levesque already showed that there is a close connection to Moore's [3] autoepistemic logic (AEL). Recently, Lakemeyer and Levesque [2] showed that only-knowing can be adapted to capture default logic as well. The main benefit of using Levesque's logic is that, via simple semantic arguments, nonmonotonic conclusions can be reached without the use of meta-logical notions such as fixpoints. Only-knowing is then naturally of interest in a many agent context, since agents capable of non-trivial nonmonotonic behavior should believe other agents to also be equipped with nonmonotonic mechanisms. For instance, if all that Bob knows is that Tweety is a bird and a default that birds typically fly, then Alice, if she knows all that Bob knows, concludes that Bob believes Tweety can fly.[1] Also, the idea of only-knowing a collection of sentences is useful for modeling the beliefs of a knowledge base (KB), since sentences that are not logically entailed by the KB are taken to be precisely those not believed. If many agents are involved, and suppose Alice has some beliefs on Bob's KB, then she could capitalize on Bob's knowledge to collaborate on tasks, or plan a strategy against him.

As a logic, Levesque's construction is unique in the sense that in addition to a classical epistemic operator for belief, he introduces a modality to denote what is *at most* known. This new modality has a subtle relationship to the belief operator that makes extensions to a many agent case non-trivial. Most extensions so far make use of arbitrary Kripke structures, that already unwittingly discard the simplicity of Levesque's semantics. They also have some undesirable properties, perhaps invoking some caution in their usage [4]. For instance, in a canonical model by Lakemeyer, certain types of epistemic states cannot be constructed. In another Kripke approach by Halpern, the modalities do not seem to interact in an intuitive manner. Although an approach by Halpern and Lakemeyer does indeed successfully model multi-agent only-knowing, it forces us to have the semantic notion of validity directly in the language and has proof-theoretic constructs in the semantics via maximally consistent sets. Precisely for this reason, that proposal is not natural, and it is matched with a proof theory that has a set of new axioms to deal with these new notions. It is also not clear how one can extend their semantics to the first-order case. Lastly, an approach by Waaler avoids such an axiomatization of validity, but the model theory also has problems [5].

[1] We use the terms "knowledge" and "belief" interchangeably.

The goal of this work is to show that there is indeed a natural semantics for multi-agent only-knowing for the quantified language with equality. For the propositional subset, there is also a sound and complete axiomatization that faithfully generalizes Levesque's proof theory.[2] We also differ from Halpern and Lakemeyer in that we do not enrich the language any more than necessary (modal operators for each agent), and we do not make use of canonical Kripke models. And while canonical models, in general, are only workable semantically and can not be used in practice, our proposal has a computational appeal to it. We also show that if we do enrich the language with a modal operator for *validity*, but only to establish a common language with [4], then we agree on the set of valid sentences. Finally, we obtain a first-order multi-agent generalization of AEL, defined solely using notions of classical logical entailment and theoremhood.

The main element of our technical construction are epistemic states defined using a notion called k-structures, which keeps separate the possible worlds that Alice believes, from the worlds Alice assumes Bob to believe, to some depth k. We show that k-structures not only capture all our intuitions of how only-knowing should behave in the many agent case, but also simplifies the semantics proofs.

Bibliography

[1] H. J. Levesque. All I know: a study in autoepistemic logic. *Artificial Intelligence*, 42(2-3):263–309, 1990.

[2] G. Lakemeyer and H. J. Levesque. Only-knowing: taking it beyond autoepistemic reasoning. In *AAAI'05*, pages 633–638. AAAI Press, 2005.

[3] H. Levesque and G. Lakemeyer. *The logic of knowledge bases*. The MIT Press, 2001.

[4] J. Y. Halpern and G. Lakemeyer. Multi-agent only knowing. *Journal of Logic and Computation*, 11(1):41–70, 2001.

[5] A. Waaler and B. Solhaug. Semantics for multi-agent only knowing: extended abstract. In *TARK*, pages 109–125, 2005.

[2] The proof theory for a quantified language is well known to be *incomplete* for the single agent case. It is also known that any complete axiomatization cannot be *recursive* [3].

1.2 Multipath Routing in Wireless Mesh Networks

Arnd Hannemann (hannemann@nets.rwth-aachen.de)
Supervisor: Prof. Dr. rer. nat. Otto Spaniol

Over the past years, wireless mesh networks (WMN) have gained more and more popularity in both academic literature and the industry. This is understandable as WMNs, unlike traditional mobile ad hoc networks (MANET), focus primarily on the demands of the mass market by providing *high bandwidth* and *open access* to the Internet. Although the redundant, hierarchical, and layered architecture of wireless mesh networks promises a self-organizing, -healing, and -configuring network, these advantages cannot be fully utilized due to protocol constraints.

Adequate routing protocols are essential for reliable communication with good performance. In today's WMN implementations, mostly single path MANET routing protocols (e.g. OLSR [1]) are employed. While such protocols generally work in WMNs, they fail to exploit the *hierarchical nature* and existing infrastructure. One goal of this thesis is to develop a routing protocol, which supports the heterogeneity of WMNs, including support for legacy clients and existing infrastructure taking into account the special idiosyncrasies found in WMNs. As a part of this work a link quality plugin for an existing OLSR implementation has been developed, which utilizes link layer statistics to calculate a more suitable routing metric.

Moreover, in WMNs where the number of redundant paths by far exceeds those in conventional last-hop wireless or wired networks, multipath routing would improve reliability and performance of end-to-end communication. Thereby, another objective of this thesis is to add multipath capabilities to the developed routing protocol. To retain performance while using a per-packet allocation granularity, multiple path will be advertised to a congestion-aware multipath capable TCP (MPTCP) stack. As the goal is to provide an integrated solution for use in real world networks, current development is performed within the UMIC-Mesh [2] testbed.

Bibliography

[1] T. Clausen and P. Jacquet. The optimized link state routing protocol (OLSR). RFC 3626, Internet Engineering Task Force, Oct. 2003.

[2] A. Zimmermann et al. Performance evaluation of a hybrid testbed for wireless mesh networks. In *Proceedings of the 4th IEEE Int. Conf. on Mobile Adhoc and Sensor Systems (MASS'07)*, pages 1–10. IEEE Computer Society, Oct 2007

1.3 Cluster-based Channel Assignment for Multi-radio Multi-channel Wireless Mesh Networks

Di Li (di.li@nets.rwth-aachen.de)
Supervisor: Prof. Dr. Dr. h.c. Otto Spaniol

Multiple non-overlapping channels provided by IEEE 802.11x standards and increasingly cheaper commodity interfaces make it possible that mesh routers can communicate with other mesh routers via different interfaces and channels concurrently, which can decrease interferences and improve the throughput. There have been many schemes of channel assignment for wireless mesh networks, but many of them are not tailored to WMNs, but suitable for common ad hoc networks Our scheme of channel assignment is tailored for the characteristics of WMNs, there are 3 steps in our scheme.

1. Clustering: All the mesh routers including Gateway nodes will be divided into clusters, every two nodes in one cluster are at most 2 hops away. The node with most links will be elected as the cluster head, which record the information of its member and nearby clusters.

2. Channel assignment: Because the main purpose of WMNs is to provide the last mile connection, the traffic traversed in/out the gateway is the majority of the whole traffic within WMNs, so the links around the gateway nodes are easily to become bottlenecks. Our channel assignment scheme start from the cluster where gateway nodes locate and end in the distant clusters, least used channels will be assigned first.

3. Borrowing channels from other clusters: A problem arises when nodes located in distant clusters want to communicate with brethren. Because the channels assigned to them are heavily loaded and endured serious interferences, so the available quality of certain links can't satisfy some requirements. Our scheme borrow channels from other clusters, if only the total bandwidth occupied by both the lending links and borrowing links do not exceed the threshold, and the valid average bandwidth of the relevant links after borrowing does not increase.

1.4 QoE in Cooperative Wireless Networks

Mónica Alejandra Lora Girón (monica.lora@nets.rwth-aachen.de)
Supervisor: Prof. Dr. Klaus Wehrle

Abstract: Cooperative Wireless Networks (CWN) have become an attractive alternative for providing ubiquitous and inexpensive connectivity for the mobile users. In CWN certain public and popular spaces may facilitate the appearance of the problem of sporadic user congestion; the occurrence of this local congestion adversely impacts the network and the user, degrading the application throughput, therefore the Quality of User Experience (QoE).

The permanent connectivity is a desirable requirement for any user, along with the advantages that this entails such as the availability and have the information at your fingertips, nowadays the users wants to use their mobile devices to become accessible and get interested data according to location and context where they are, from anywhere, at any time, this makes the CWN a low-cost alternative that satisfy in some extent this high degree of connectivity.

Although CWN are a promising solution for the current trends, the lack of a planned growth of the network, and centralized management make the network components are vulnerable to saturation due to the growing demand for this type of access. For example, if one access point is overloaded, it is necessary to distribute the load between different networks components.

When unbalanced load in the wireless network exists, it is desirable looking for neighboring underutilized access points (AP); but the challenging task is hard, because there are unknown factors which hamper the network load balancing and an efficient AP selection such as new traffic patterns in WLAN, type of user applications, number of users, currently load condition of the AP, handover latency, unplanned grow process and mobility patterns of the user. Another issue is finding an efficient and secure way to allow the exchange of information about the network features between APs in different domains, considering an approach without centralized controller.

This work concentrates on how to improve and optimize the user application throughput through a dynamic resources management scheme on the Wireless Community Network. The focus lies on determining the current traffic patterns in order to select the able access point providing the best service according to user requirements, and manage the existing user connections when resources are scarce. To reach this goal, a AP-selection algorithm has to be created and implementation of the standard IEEE 802.11e in the APs in order to allow to the user improve his/her QoE.

1.5 Integrated Design and Configuration of Versatile Software Documents in Automotive Software Engineering

Cem Mengi (mengi@i3.informatik.rwth-aachen.de)
Supervisor: Prof. Dr.-Ing. Manfred Nagl

Software engineering in the automotive domain has gained more and more importance. Today, about 80% of all innovations are software-based. However, due to the traditional hardware-driven development process, automotive software engineering gets highly *complex*. Moreover, the possibility to select optional fittings, e.g., parking assistant, rain sensor, intelligent light system etc., leads to the situation that an enormous number of *software variants* arise.

To overcome the growing complexity, the automotive industry agrees that there is a need to shift from a hardware-driven to a function-driven process, i.e., to abstract from hardware details such as the underlying ECU topology, deployment decisions, real-time characteristics etc. In GRK 643, we provide new methods, concepts, and tools to support the function-driven design process. We distinguish between four levels of models, the *conceptual model*, *behavioral model*, *implementation model*, and *source code model*. The conceptual model describes the first virtual realization of the static system structure and builds a bridge between requirements and architecture specification. The behavioral model is used for simulation of functional behavior. Typically, there are infinite resources available, the model is abstracted from hardware topology, and from any deployment decision. If deployment decisions are made, the behavioral model is enriched with real-time characteristics, the software structure and data types are optimized, and the model is tied to technical interfaces such as bus systems and hardware resources. We call the result of this task as implementation model. Finally, the source code model is generated by the implementation model so that it could be executed on specific hardware platforms.

In order to handle variability in a function-driven development process we provide an integrated approach to handle variation points in the software documents of the different levels. To capture the points of variation, we provide a so called variability model. Thereby, variation points and their variants are organized in a tree-based structure which is extended with constraints in order to express variability such as mandatory, optional, and exclusive variation points. The variability model is integrated into the four levels of models so that it can be used synchronously. Finally, we provide a mechanism to configure the variability model in order to derive specific variants.

1.6 Anonymous Communication

Andriy Panchenko (panchenko@cs.rwth-aachen.de)
Supervisor: Univ. Prof. Dr. rer. nat. Dr. h. c. Otto Spaniol

With the growth of the digitized world privacy issues get more and more importance. While cryptography can be used to protect integrity and confidentiality of the data part of the packets, everyone along a route of packets can still observe the addresses of the communication parties. This is often enough to uniquely identify a person. Time, duration, and volume of communications can additionally be used to infer further information like, e.g., a social relation between the communicating parties.

Anonymous communication deals with hiding relationships between communicating parties. It is a basic fundamental building block for privacy-friendly web browsing, any viable identity management system, privacy-aware eGovernment, eCommerce, and eHealth technologies. Privacy-friendly communication is also necessary for providing freedom of speech, mind, and the achievement of democratic principles even in those countries that try to filter and censor access to information. Thus, strengthening privacy-enabled communication can be seen as a major goal from a social point of view.

The goal of this research is to enhance exiting and develop new methods of anonymous communication and provide a stable basis for a solid understanding of the topic. Topics of research and interest include, but are not limited to: performance of network layer anonymization, the limits of anonymization techniques, evaluation of attacks and their effectiveness, design of new protocols for anonymity and analysis of their properties, support for law enforcement in anonymous communication infrastructures, etc.

1.7 Knowledge Discovery in Digital Libraries

Manh Cuong Pham (pham@dbis.rwth-aachen.de)
Supervisor: Prof. Dr. Matthias Jarke

Scientific digital libraries play a more and more important role in scientific activities in computer science. They act as knowledge repositories where research communities, research activities, the history and development of research ideas and fields are reflected. With the recently availability of large scale digital libraries such as DBLP, CiteSeerX, ACM Portal, IEEE, Google Scholar, Microsoft Academic Search etc., there are new opportunities to discover the knowledge documented by them and provide useful tools for the discovery process and to help researchers in daily work. The aim of this project is to represent and analyze scientific knowledge of computer science and develop recommendation techniques[1, 2] that support researchers to find conferences and journals, to search for interesting research communities and potential collaborators, and to manage research materials. To realize these objectives, we focus on two aspects of research activities: the social aspect and the knowledge representation. Social relationships are built up via research activities, e.g. publishing papers (co-authoring), referencing to other work (citing), or participating in conferences. Social Network Analysis (SNA) is applied on these sources of information to discover the pattern of interaction between researchers. The pattern then is used in the design and implementation of recommender systems. For the second aspect, we concentrate on how knowledge is organized and how it is transferred between individual researchers, journals, conferences and between research fields. Visual analystics is used to represent and identify research communities and their evolution in term of knowledge diffusion and research collaboration. Combining both aspects is one of the objectives of this project: knowledge representation describes the social aspect of research activities and can be used as the basic for recommendation services.

Bibliography

[1] Adomavicius, G., Tuzhilin, A.: *Toward the next generation of recommender systems: A survey of the stateof-the-art and possible extensions*. IEEE Transactions on Knowledge and Data Engineering, vol. 17(6), pp. 734–749 (2005)

[2] Burke, R.: *Hybrid recommender systems: Survey and experiments*. User Modeling and User-Adapted Interaction, vol. 12(4), pp. 331–370 (2002)

1.8 Localize Me!

Torsten Sattler (tsattler@informatik.rwth-aachen.de)
Supervisor: Prof. Dr. Leif Kobbelt

Due to the recent success and widespread adoption of smart-phones, especially Apple's iPhone and mobile phones running Google's Android operating system, location based services have started to gain popularity. The term 'location based services' refers to services that offer certain information based on the current position of the user. Those services could for example show the user the location of a nearby Italian restaurant on a map centered at the user's position, provide touristic information or offer navigation hints similar to navigation devices for a car.
The most common ways of determining the position of the user are to use either GPS or location information provided by a WLAN connection. While GPS gives rather accurate localization results if the user sees a large part of the sky, the accuracy of GPS drops off in urban areas due to buildings blocking the direct communication with the satellites. Similarly, reflections off the walls of the buildings have a negative impact on the accuracy of localization using a WLAN connection.
Algorithms from the research area of Computer Vision allow a precise localization even in those cases where GPS and WLAN-based localization fail. Besides an internet connection, all that is required is a camera, which is part of the standard equipment of today's mobile phones. The user takes a photo, which is compared to a large set of images stored in a database, each of those is annotated with the position the image was taken from, its so-called geotag. The geotag of the image most similar to the photo taken by the user then gives a very good approximation to the actual position of the user. The 'Localize Me!' application, which has been developed in cooperation with the Arne Schmitz, Pascal Steingrube and Tobias Weyand, uses this technique described above to allow users to determine their position inside the city of Aachen.
The accuracy of the localization result strongly depends on how dense the city of Aachen is sampled with images. By using a textured 3D model of the city, it is possible to further improve the precision of the localization up to a few centimeters. Instead of finding the most similar picture from a set of images, this approach finds common features in both the photo taken by the user and the textures of the 3D model of the city. These matching features are then used to determine the position of the user relative to the 3D model. In this talk, we will present the current state of our research on acquiring a 3D model of the city of Aachen, as well as integrating this model into the 'Localize Me!' application.

1.9 Improving Performance in Corruption-susceptible Networks

Florian Schmidt (Florian.Schmidt@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Klaus Wehrle

Voice over IP (VoIP) has in recent years become a main contender for audio-visual communications, and its widespread use is anticipated to supplant a sizable amount of classical telecommunication via phone lines. Of special importance is the use in wireless devices, and how to improve the quality of communication, from a user point of view in respect to perceived speech or video quality, as well as from a provider point of view in respect to increased communication efficiency.

In current systems for wireless IP transmission of audio-visual signals, residual bit errors on a wireless link are detected by CRC on the physical or link layer, and erroneous packets are discarded. We try in this thesis to increase the quality of communication by devising ways to improve this relatively ineffecient method of ARQ (automatic repeat-request).

First of all, the general assumption in the field is that packet headers of the used protocols (in the case of VoIP, typically IP, UDP, and RTP) generally have to be error-free, because they contain the information that is necessary to assign packet payloads to the correct communication endpoints (identify the flow). However, the codecs used by the applications can make use of data with bit errors. This can be further enhanced by combining it with so-called soft information, i.e., reliability information from the channel decoder about probabilities of every bit to be correct. Being able to use partly corrupted data is a vital point in enhancing the quality of VoIP communication, especially over wireless links, where bit errors are much more common.

Furthermore, we can try to salvage even those packets where parts of the header have been corrupted. One way to do this are heuristics to recognize which information has been lost by corruption, e.g., by checking against the expected header information for each communication end-point, and assign the packet to the best-matching one. Again, these decisions can make use of reliability information from the channel decoder. This process has to take special care to not assign packets to wrong flows for reasons of data integrity and security. However, this is promising idea, because it can be used not only for headers of packet-switched audio-visual communication, but for every type data sent over such a wireless network (such as 802.11 WLANs).

Finally, a framework is needed to communicate information about data corruption, reliability, etc. across the layers of a standard network stack to facilitate cooperation between the different layers, which in the standard ISO/OSI model are strictly separated from each other. This communication is necessary to, for example, inform the upper layers about bit errors that

were recognized on one of the lower layers, and to hand over reliability information from the channel decoder which resides on the physical layer. Cross-layer communication has been a research topic for some time now; however, solutions that are presented are generally single-purpose for special circumstances, and not reuseable or combinable with different approaches because of potential conflicts between them. What is missing so far is a generic framework that is flexible and easily expandable. We therefore are creating a cross-layer framework that meets these demands, and allows us to exchange information on the level of single packets, as well as packet flows and for generic network-related information. In the scope of this thesis, we will contribute to this framework in the area of packet- and packet-flow-related information and its exchange in a cross-layer fashion.

1.10 Factor-graph Based Channel Estimation for MIMO-OFDM Systems

Xiang Xu (xu@ti.rwth-aachen.de)
Supervisor: Rudolf Mathar

Modern wireless communication systems require high-bit-rate transmission. To achieve higher data rate, one way is to reduce the symbol duration. While the symbol duration is significantly shortened, the multi-path channel will cause severe frequency selectivity. To overcome frequency selectivity, orthogonal frequency-division multiplexing (OFDM) is often used [1]. OFDM is a multicarrier modulation, where the high-bit-rate data stream is divided into several parallel lower bit-rate streams, and modulated to separate orthogonal carriers. Another way to increase data rate is to use multiple antennas to transmit and receive signals simultaneously. This technique is often referred as "Multiple input and Multiple output" (MIMO). The combination of MIMO and OFDM leads to a MIMO-OFDM system, which is very promising in future mobile communications.

To perform coherent demodulation in MIMO-OFDM systems, accurate channel estimation is needed. Due to the time and frequency selective property of the physic channel, the number of parameters to be estimated is very large. Thus, the computational complexity of traditional channel estimation techniques can be overwhelming.

Factor graph is a bipartite graph visualizing the factorization of certain global functions subject to minimization or maximization. It is often helpful in design of low-complexity iterative processing algorithms. By treating channel coefficients as variables as well as data symbols, data detection and channel estimation can be performed over a general factor graph. Differed from the conventional schemes, in which data detection and channel estimation are performed in separate manner, this algorithm does everything in one stage. Currently, static flat fading channels are assumed in most literature of MIMO channel estimation [2]. To get closer to realistic environment, time varying and frequency selective property of the channel should also be considered.

Bibliography

[1] Y. Li and G. L. Stuber. *Orthogonal Frequency Division Multiplexing in Wireless Communications*. Springer, 2006.

[2] T. Wo and C. Liu. *Graph-based Soft Channel and Data Estimation for MIMO Systems with Asymmetric LDPC codes*. IEEE International Conference on Communications (ICC), 2008.

2 GRK 1042: Explorative Analysis and Visualization of Large Information Spaces

Prof. Dr. Dietmar Saupe
Email: Dietmar.Saupe@uni-konstanz.de
University of Konstanz
Internet: `http://www.inf.uni-konstanz.de/gk/`

The program of the Research Training Group focuses on the development of methods for supporting exploration, analysis, and management of large information spaces, especially in the fields of visualization, computer graphics, and user-interaction. These information spaces may of visual nature as well, e.g., multimedia documents or complex geometric structures. A further emphasis of research is the systematic exploration of large data spaces emerging for example in the analysis of software systems. The relevant fields in the Research Training Group are information visualization, computer graphics, digital signal processing, human computer interaction, intelligent data analysis, information retrieval, databases, and information systems. One of the objectives of data mining and explorative 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. Applications that are investigated in the Research Training Group are for example analysis of cell image sequences and graphs in bio-informatics, network analysis of textual corpora, feature engineering for search in multimedia databases, and visualization and analysis of performance measurements from sport- and training-sciences. The Research Training Group implements a structured concept for advising and teaching of its doctoral students. In the first phase, for the duration of two semesters, special courses are provided and special lecture series are held during summer schools. The students are introduced to research tasks in their respective research workgroups and seminars, and acquire soft skills qualification in university wide courses.

2.1 Adaptive Thinning Algorithms of Atmospheric Observations in Data Assimilation for Numerical Weather Prediction

Vladimir Bondarenko (bondaren@inf.uni-konstanz.de)
Supervisor: Prof. Dr. Dietmar Saupe, Prof. Dr. Gerik Scheuermann

Data assimilation is an important part of any forecasting problem. It is a procedure of estimation of the initial state of a dynamical systems from available *indirect* observations. Data thinning (data reduction) - selection of a subset of observations from the complete set - is a necessary preprocessing step if the observations are acquired at the rate exceeding our ability to assimilate them, as in the case of satellite-based observation systems used for numerical weather prediction. The direct "optimal" data reduction methods cannot cope with the everincreasing amount of satellite measurements. We develop new approaches to data thinning, that tradeoff some information losses for processing speed.

One such approach, "adaptive data thinning", has been pursued for the last three years by our group in close cooperation with the German National Weather Service. Two thinning methods were developed: "cluster based thinning" and "estimation error analysis" [Och07]. Moreover, the evaluation framework to assess the quality of different thinning schemes on estimation and forecast error is under development [Bon07].

Bibliography

[Bon07] V. Bondarenko, T. Ochotta, D. Saupe, and W. Wergen. *The interaction between model resolution, observation resolution and observation density in data assimilation: a two-dimensional study.* In 11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, San Antonio, Texas, 2007.

[Och07] T. Ochotta, C. Gebhardt, V. Bondarenko, D. Saupe, and W. Wergen. *On thinning methods for data assimilation of satellite observations.* In 23rd International Conference on IIPS for Meteorology, Oceanography and Hydrology, San Antonio, Texas, 2007.

2.2 Simulation and Optimization of Race-Bike Training on Realistic Tracks

Thorsten Dahmen (thorsten.dahmen@uni-konstanz.de)
Supervisor: Prof. Dr. Dietmar Saupe

The goal of the Powerbike project is to realistically simulate outdoor riding on existing tracks in the lab, extract the relevant information of a multitude of available measurement data and visualize this information for optimal training control and performance prediction.

The internal structure of commercially available bicycle simulators is inaccessable and their resistance control does not simulate cycling on realistic tracks accurately. Therefore, we designed our own simulator based on a Cyclus2 ergometer (RBM Elektronik), [2]. We implemented a resistance control which relies on a validated model that describes the relation between pedalling power P and velocity v on a cycling track.. A display shows a video playback synchronized with the cyclist's current position on the track together with various course and performance parameters. We simulate arbitrary virtual gear transmissions, which the cyclist operates using four electronic pushbuttons.

In order to validate the P-v-model and our simulator, we collected data from a set of rides with varied riders, courses and pacing strategies and compared them with the power/velocity predicted by the P-v-model, [1]. We found that the P-v-model is accurate and that the simulator is able to control the brake force based on the model appropriately.

We selected standard ergometer tests to estimate the endurance time T an athlete can hold as a function of the power P for individual athletes. Together with the P-v-model, we seek the optimal pacing by minimizing the total time, to complete a track, subject to the physiological side constraint that the cyclist is completely exerted when finishing the track.

Bibliography

[1] Thorsten Dahmen, Roman Byshko, Martin RÃűder, Stephan Mantler, Dietmar Saupe. *Modeling, simulation and validation of cycling time trials on real tracks*. Proceedings of the 7th International Symposium on Computer Science in Sport, September, Canberra, Australia, 2009.

[2] Thorsten Dahmen, Dietmar Saupe. *A simulator for race-bike training on real tracks*. Proceedings 14th Annual ECSS Congress, 2009.

2.3 Novel Interaction Techniques for Information Seeking

Mathias Heilig (mathias.heilig@uni-konstanz.de)
Supervisor: Prof. Dr. Harald Reiterer, Prof. Dr. Marc Scholl

Physical libraries were the most important knowledge source in former times. But within the last decades digital libraries and the WWW came up, with a lot of powerful features (everytime and everywhere available, digital search, a lot of metadata, sorting and filtering, visualizations etc.) and threatened the monopoly position of physical libraries as information distributor. However physical libraries are still very important for example as a place for social interaction and offer also implicitly physical knowledge structure. A location of an item in the shelf is not only a storing position, but also an information carrier, which sets the items in different relations and contexts. Users develop unconsciously a spatial literacy to find not only the position and neighborhood of an item in the shelf, but rather perceive additional meta-information like the number of other visitors in the proximity or the look and the age of covers and spines. To maintain these characteristics and additionally enhance the work inside of physical library, it is aimed to blend the advantages of digital libraries into the physical library.
One important activity in physical libraries is searching. Several theoretical models based on decades of empirical work in different domains on how users pass through their seeking process showed that this activity is very complex and multifaceted. The social aspect of seeking or switching between browsing and analytical searching modes are only two examples. At the moment however, people have either to browse the shelves or to use the web frontend of a library to find items. This gap of the single-user web frontend on the one side and the physical library on the other side should be bridged through the blending of digital seeking functionalities into the physical space of the library.

2.4 Cross-Display Input Device for Co-located Collaborative Work

Mahsa Jenabi (mahsa.jenabi@uni-konstanz.de)
Supervisor: Prof. Dr. Harald Reiterer, Prof. Dr. Oliver Deussen

Meeting rooms are nowadays increasingly equipped with many displays having different sizes and a variety of functionalities (e.g., touch sensitivity). While large displays are suitable for presentation, horizontal touch tables, such as Microsoft Surface, may enhance collaborative work. In addition, people often bring their own Tablet PCs and/or mobile phones, which also have integrated displays. This results in coupled displays [1] in the room, introducing new challenges for cross-display interaction. Multi-display environments (MDE) are fundamentally different from traditional single-display workspaces. Therefore superficial adaptations are insufficient in real world applications. It is necessary to pay special attention to their differences. For example, in a presentation meeting, one might want to show data about accidents in one particular year and at the same time showing the geographical locations of these accidents in the Google Earth on another display. Further displays could show a table of accident costs and impacts. In most cases, such a multi visualization option is possible when the displays are all connected to the same computer. Still, the configuration of views is time-consuming. An intuitive interaction method is required in these situations.
Interaction in MDEs can be divided into two types: within-display interaction and between-display interaction. Within-display interaction includes the method of interaction well known from single-display workspaces. Users should be able to move and manipulate objects inside the display. A set of primitive interaction tasks is mentioned in ISO 9241-400 and Foley, et al. [2]. Between-display interaction (also known as cross-display interaction) includes a specific set of interaction tasks that can only be performed, when two or more coupled displays are available. For example transferring an object form one display to another.
Between-display interaction is a new challenge for HCI researchers and interaction designers. Therefore this type of interaction is the focus of this research project. We have identified cross-display interaction primitives, which, so far, have not been investigated in the literature. Interaction designers can use this interaction primitives, while they are prototyping their multi-display space. We implement a practical prototype, using an iPhone, as a mobile input device with integrated display. This incorporates these primitive tasks and solves the problems we recognized in state-of-the-art projects and shall enable users to have a fluid interaction with multiple displays, while collaborating with other team members. Lessons learned from designing and implementing the prototype will show the HCI community and us the advantages of using

mobile input device with integrated display for cross-display interaction.

Bibliography

[Jen09] M. Jenabi. *Multimodal interaction with mobile phones (abstract).* In A. Avanes, D. Fahland, J. Geibig, S. Haschemi, S. Heglmeier, D. A. Sadilek, F. Theisselmann, G. Wachsmuth, and S. Weißleder, eds., des gemeinsamen Workshops der Informatik-Graduiertenkollegs und Forschungskollegs 2009, p. 26, June 2009.

[Jen09a] Jenabi, M., *Multimodal interaction with mobile phones (abstract).* Proceedings des gemeinsamen Workshops der Informatik-Graduiertenkollegs und Forschungskollegs pp. 26, 2009.

[Jen08] Jenabi, M., Reiterer, H., *Finteraction - Finger interaction with mobile phone.* Proceedings Future Mobile Experiences (NordiCHI 2008), Lund, Sweden, 2008.

2.5 Visual Spatiotemporal Analysis of Movement and Event Data

Slava Kisilevich (slaks@dbvis.inf.uni-konstanz.de)
Supervisor: Prof. Dr. Daniel Keim, Prof. Dr. Harald Reiterer

In many applications, spatiotemporal data is generated in rapidly growing amounts, and therefore there is a high demand for scalable analysis methods, which allow a systematic analysis and have a sound theoretical basis. Spatiotemporal data, most importantly movement data, involve geographical space, time, and multidimensional attributes and thereby pose significant challenges for the analysis. We plan to develop theoretical foundations for the analysis of spatiotemporal data, which account for possible variations of the essential properties of the data. We will thereby identify the generic analysis tasks for different types of movement data (quasicontinuous and event-based) and different views of movement (trajectory-oriented and traffic-oriented). The goal is to develop the appropriate analysis methods, which combine visual, interactive, and algorithmic techniques for a scalable analysis. The algorithmic techniques will cluster, aggregate, and summarize trajectories, traffic situations, events, and corresponding multidimensional attributes for providing an overview of the data and extract significant patterns of potential interest. Visual and interactive techniques allow the user to steer the automated algorithms and input his knowledge into the analysis process. Visualizing the transformed data also provides important feedback to the user and allows him to better understand the data. A tight integration of visual and automatic techniques is crucial to the success of spatiotemporal analysis.

Bibliography

[And09a] G. Andrienko, N. Andrienko, P. Bak, S. Kisilevich, and D. Keim. *Analysis of community-contributed space- and time-referenced data (example of Flickr photos)*. In Vision, Modeling, and Visualization-Workshop (VMV), pp. 16Ű18, 2009. Braunschweig, Germany.

[And09b] G. Andrienko, N. Andrienko, P. Bak, S. Kisilevich, and D. Keim. *Analysis of community-contributed space- and time-referenced data (example of Flickr photos)*. In IEEE Symposium of Visual Analytics in Science and Technology (VAST Poster Papers). IEEE Press, 2009. Atlantic City, New Jersey, USA.

[And09c] G. Andrienko, N. Andrienko, P. Bak, S. Kisilevich, and D. Keim. *Demo-Paper: Analysis of community-contributed space- and time-referenced data (by example of Panoramio photos)*. In ACM-GIS Conference. ACM, 2009.

2.6 XML Full-Text Retrieval

Stefan Klinger (klinger@inf.uni-konstanz.de)
Supervisor: Prof. Dr. Marc Scholl, Prof. Dr. Marcel Waldvogel

Mit der Entwicklung von Volltext Anfragesprachen wie NEXI oder XQuery Full-Text wurde der Bedarf an XML Datenbanken mit entsprechenden Fähigkeiten offensichtlich.

Die Ansätze der Volltext Anfragesprachen unterscheiden sich dabei stark: NEXI geht z.B. nicht von einer definierten Semantik aus, sondern möchte diese zur Laufzeit "'entdecken"'. Auf der anderen Seite findet man XQuery Full-Text, eine Spezifikation die versucht IR Konzepte möglichst ohne Brüche in die XQuery Sprache einzufügen, und dabei nur einig spezielle Konzepte undefiniert lässt. Da dieser letzte Ansatz die Beziehung zwischen angefragter und gefundener XML Struktur strikt beachtet passt er besser zur Welt der Datenbanken, und bildet den Schwerpunkt meiner Forschung.

Das Pathfinder Projekt [Gru04] hat den Aufbau eines rein relationalen XQuery Systems zum Ziel. Unser Projekt untersucht nun, wie sich der Pathfinder Compiler um die für XQuery Full-Text nötige Scoring-Infrastruktur erweitern lässt. Anstatt eine spezielle XQuery Full-Text Implementierung zu schaffen welche die verwendeten IR- und Scoring-Modelle festlegt, werden Schnittstellen definiert, über die sich verschiedene solcher Modelle einbinden lassen.

XQuery stellt von sich aus nur Sequenzen (heterogene Listen) und Bäume (XML) zur Verfügung, jedoch keine Record-Typen. Für XQuery Full-Text musste eine Paarung von XQuery- und Score-Werten gefunden werden, die effizient und implizit neben XQuery Ausdrücken verarbeitet werden kann.

Unser Prototyp zeigt, daß sich der Pathfinder Compiler elegant um eine transparente Scoring-Infrastruktur erweitern läßt. Allerdings zeigt sich auch, daß eine implizite Verarbeitung von Scores nicht unbedingt mit den expliziten Operationen verträglich sein muss [Hie08].

Bibliography

[Gru04] Torsten Grust, Jens Teubner. *Relational Algebra: Mother Tongue - XQuery: Fluent.* Twente Data Management Workshop on XML Databases and Information Retrieval (TDM 2004), Invited Paper, Enschede, The Netherlands, June 2004.

[Hie08] D. Hiemstra, S. Klinger, J. Flokstra, and P. Apers. *Sound ranking algorithms for XML search.* In Workshop on Focused Retrieval (SIGIR 2008), Singapore, July 24 2008.

2.7 Visual Analytics Methods for Large-Scale Analysis of Information Dynamics in News Streams

Miloš Krstajić (milos.krstajic@uni-konstanz.de)
Supervisor: Prof. Dr. Daniel Keim, Prof. Dr. Oliver Deussen

Analysis of news data requires automated and interactive visual methods that will facilitate processing of the constantly growing amount of the acquired information for the analysts in different domains. News data items in online information streams are characterized by textual and temporal attributes, and analyzing the relationships between these attributes and across data presents one of the fundamental problems in text data mining. The topics that are appearing in the news are characterized by high temporary changes, hierarchical structure and inter-topic relationships. The change in amount, frequency and content describes dynamics of this type of data and, additionally, data stream perspective augments temporal dimension of the problem. Although streaming data is similar to time-series data, its large-scale and unbounded characteristics make regular temporal data visualization techniques ineffective.

In my research the goal is to develop a visual analytics framework for the analysis of online information streams. I will identify the tasks for the analysis of news data, such as monitoring of online information streams and exploration of large text collections. The goal is to develop algorithmic methods, which will perform incremental clustering and aggregation across sources, languages and categories appearing in the news data streams. Interactive visualization techniques will be developed to allow the user to explore and gain meaningful insight from the processed data.

Bibliography

[Krs10a] M. Krstajic, P. Bak, D. Oelke, M. Atkinson, W. Ribarsky, and D. Keim. Applied visual exploration on real-time news feeds using polarity and geo-spatial analysis. In 6th International Conference on Web Information Systems and Technologies (WEBIST 2010), 2010. Valencia, Spain.

[Krs10b] M. Krstajic, F. Mansmann, A. Stoffel, M. Atkinson, and D. Keim. Processing online news streams for large-scale semantic analysis. In DESWeb Workshop 2010, 26th International Conference on Data Engineering, ICDE 2010, March 1-6 2010. USA.

2.8 Network Visualization: Dynamics, Modes, and Collections

Martin Mader (martin.mader@uni-konstanz.de)
Supervisor: Prof. Dr. Ulrik Brandes, Prof. Dr. Michael Berthold

As complex networks receive increasing interest by network analysts, there is a strong need for appropriate network visualizations that allow for visual analysis and exploration. One task in network visualization is to obtain graph layouts for a *dynamic network* comprising a sequence of graphs. The goal here is to find coherent representations of successive networks that respect qualitative criteria of the drawing, while at the same time preserving the *mental map* the viewer has built from a previous drawing in the sequence.

In a recently submitted work we proposed a modification of the well-known, energy-based *stress-minimization* layout algorithm for this purpose: In a first step, a reference layout is computed that incorporates the information of all graphs in the sequence, i.e., all vertices and edges that ever occur in the sequence. Secondly, this reference layout serves as initialization for a modified stress-minimization process on the individual networks in the sequence. Besides reducing the stress of the layout by means of the stress function – and therefore optimizing structural criteria, like good vertex distribution and faithful representation of graph distances in the individual layouts – the modified version incorporates a term that also penalizes movement away from positions in the reference layout (*anchoring*), thus working towards the mental map given by the reference layout. Although the anchoring approach gave satisfactory results, there are other strategies, known and new ones, to be investigated by means of formulating them as modifications of stress-minimization.

In a further step, these different fundamental strategies need to be systematically compared with respect to their ability to trade off between displaying structural properties and complying with the mental map. The maintenance of the mental map can be measured by various difference metrics, such as deviation in the absolute or relative positions of successive network layouts, or the number of vertices leaving their Voronoi-regions. The bachelor thesis of a student that I supervised shows a preliminary attempt for this comparison, that is planned to be extended and improved.

Other interesting classes posing constraints on the layout algorithm are *multi-modal* networks or network *collections*. Concepts similar to those for dynamic graphs are likely adaptable for the task of proper visualization, i.e., finding layouts that preserve the mental map constituted by modes of network objects or relations between networks in a collection, respectively.

2.9 Analysis of Network Ensembles

Uwe Nagel (Uwe.Nagel@uni-konstanz.de)
Supervisor: Prof. Dr. Ulrik Brandes, Prof. Dr. Michael Berthold

Attributed networks provide a powerful modeling framework used in fields as biology, chemistry, social sciences and other. Given a collection - an ensemble - of such networks it is often crucial to analyze it in terms of underlying structural trends. For the description of structural trends in networks random graph models can be employed. My work led to the development of two approaches that describe ensembles of networks as a mixture of differently parametrized random models. Given an ensemble, these methods cluster the graphs by their structural similarity.

In a first approach the spectrum of the graphs in the ensemble was used to derive a distance capable of distinguishing graphs by their underlying planted partition model. The distance between graphs was defined using the eigenvalue distribution of the adjacency matrices. We could show that this distance is for graphs of sufficient size significantly larger if the graphs stem from different models in contrast to the case where they emerged from identical models.

This approach cannot be applied to small graphs. Therefore of a model driven approach that additionally takes node attributes into consideration, was developed. Here, the node set of the ensemble is partitioned using the node attributes and graphs are projected to this partion. From the projections, feature vectors can be derived that describe structural properties of the projected graphs. Using these feature vectors it is possible to cluster the graphs of the ensemble by structural similarity. An application of this method was presented at the ASNA'09 in Zürich. We were able to match proposed acculturation strategies of migrants with concrete data and thereby give empirical support for the proposed thesis.

Bibliography

[Ber09] M. Berthold, U. Brandes, T. Kötter, M. Mader, U. Nagel, and K. Thiel. Pure spreading activation is pointless. In 18th Conf. on Information and Knowledge Management (CIKM 2009), pp. 1915Ű1919, November 2009.

[Bra10b] U. Brandes, J. Lerner, U. Nagel, and B. Nick. Structural trends in network ensembles. In Intl. Workshop on Complex Networks (CompleNet 2009), pp. 83Ű97, 2010.

2.10 Real-Time Rendering and Modeling of Vegetation

Sören Pirk (soeren.pirk@uni-konstanz.de)
Supervisor: Prof. Dr. Oliver Deussen, Prof. Dr. Harald Reiterer

The visual complexity of virtual worlds and environments increased up to an enormous level in the past decade. A higher level of photorealism strives to immerse the user more and more each generation of graphics hardware arising. As frequent objects in our daily life, vegetation is part of almost all virtual sceneries, ranging from single plants and trees to huge outdoor landscapes.
As part of computer graphics, rendering and modeling of vegetation has been content of research for many years. Researchers as well strive to increase the richness of virtual representations of botanical objects but also to render them with hard time constraints in real-time. Even though a lot of work has already been done in this field, many problems are content of current and future research.
The detailed visual representation of trees and plants implies accurate modeling of the underlaying geometry which requires a large number of primitives.

Outdoor scenes easily have up to 100M Vertices when modeling even the smallest details. But not only the tremendous amount of geometric data outrages the possibilities of todays graphics hardware. The inhomogenous visual appearance of different parts of natural vegetation requires complex shading of the artificial counterpart when targeting photorealism. Physical interaction, the effect of different tropisms, simulating growth and even lower levels of animation would increase the realism but to the costs of not being able to meet the hard requirements set by real-time rendering, when applied to the afore mentioned intricacy.

My current research is focused on real-time rendering of large botanical scenes with hundreds or thousands of botanical objects. Experimenting with different rendering techniques like Buffer Objects, Deferred Shading and Level-of-Detail methods like Stochastic Pruning, we try to find an optimized way through the graphics pipeline as well as to cleverly reduce the amount of geometry.

2.11 High-Dimensional Methods for Dementia Analysis

Mirco Richter (mirco.richter@uni-konstanz.de)
Supervisor: Jun. Prof. Dr. Dorit Merhof

Due to the population's demographic development, dementia research has gained importance over the last years. Medical imaging has potential in identifying dementia in-vivo and is therefore of particular interest. Diagnosis as well as classification of different types of dementia are challenging topics of current research and require sophisticated methods from pattern recognition and visualization.

There are about 50 different causes provoking about 10 different types of dementia, the most common types being Alzheimer's disease (AD), frontotemporal dementia (FTD) and vascular dementia (VD). Each type of dementia results in changes to tissue in certain areas of the human brain. These biomarkers are obvious in late stages of the disease, but require reliable and robust identification in the early stage when treatment has the best possible effect.

The first part of the research project adresses the extraction of different biomarkers from medical imaging data such as cortical thickness, ventricular shape, degeneration of white matter tracts and hippocampus volume. In a next step, these biomarkers will be combined into a multidimensional classifier which will be evaluated on patients suffering from mild cognitive impairment (MCI), AD and other types of dementia. The patients take part at a standardized long term study at the Memory-Clinic Basel. For this classifier, different methods from pattern recognition and machine learning will be explored, such as Fisher discriminant analysis, support vector machines, clustering approaches, neuronal networks and genetic algorithms. The visualization of each biomarker and the multidimensional classification will be a further aspect of the research project. The final goal of the research project is to determine the best set of biomarkers and algorithms for the automated classification of different types of dementia.

2.12 Document Structure Analysis for Large Document Collections

Andreas Stoffel (andreas.stoffel@uni-konstanz.de)
Supervisor: Prof. Dr. Daniel Keim, Prof. Dr. Oliver Deussen

Informationen über die Strukturierung von Dokumenten lassen sich vielfältig verwenden und sind für viele Anwendungen im Bereich der automatischen Dokumentenverarbeitung eine Grundvoraussetzung.

Meine Forschungsarbeit beschäftigt sich mit der Fragestellung eine allgemeine Strukturerkennung zu entwickeln, die sich für die Digitalisierung von Archiven mit vielen unterschiedlichen Dokumentarten eignet. Bisherige Ansätze zur Strukturerkennung von Dokumenten eigenen sich nur für die Verarbeiten von einzelnen Arten von Dokumenten, wie beispielsweise Rechnungen, lassen sich aber nicht auf grosse Archive mit vielen unterschiedlichen Dokumentarten anwenden.

Ein allgemeiner Ansatz zur Strukturerkennung konnte durch ein Kombination von maschinellen Lernverfahren mit Visualisierungstechniken implementiert werden. Es konnte gezeigt werden, dass dieser Ansatz vergleichbare oder bessere Ergebnisse auf mehreren unterschiedlichen Dokumentarten erreicht, wie die entsprechenden spezialisierten Algorithmen.

Durch die Kombination von Visualisierung mit maschinellen Verfahren erhält ein Benutzer die Möglichkeit die Fähigkeiten und Probleme der verwendeten Modelle zu untersuchen. Er kann Problemfälle erkennen, bewerten und geeignete Massnahmen zur Verbesserung der Erkennungsergebnisse ergreifen.

Die Strukturanalyse lässt sich neben der automatisierten Verarbeitung von Dokumenten auch für die Inhaltliche Analyse nutzen. So wurde für die "Document Card" Visualisierung von Dokumenten, die für die Berechnung von repräsentativen Termen die Struktureranylse verwendet.

Bibliography

[Sto10] A. Stoffel, D. Spretke, H. Kinnemann, and D. Keim. Enhancing document structure analysis using visual analytics. In ACM Symposium on Applied Computing 2010, 2010. Sierre, Switzerland.

[Str09] H. Strobelt, D. Oelke, C. Rohrdantz, A. Stoffel, D. Keim, and O. Deussen. Document cards: A top trumps visualization for documents. In IEEE Symposium on Information Visualization (InfoVis 2009), 2009.

2.13 Advanced Visualizations of Large Document Collections

Hendrik Strobelt (hendrik.strobelt@uni-konstanz.de)
Supervisor: Prof. Dr. Oliver Deussen, Prof. Dr. Daniel Keim

Nowadays, large document collections, such as research paper corpora and news feeds, grow at high rate. Many of these documents contain text and images for describing facts, methods, or telling stories. It is an exhaustive task for a user to get an overview of of a larger collection. So far, search engines allow only a search-term-centric view on documents on bas of either small texts or figures. For the tasks of overviewing and individually arranging the document collection, a compact visualization of each document is eligible. Furthermore this visualization should make use of the concept-describing feature of texts and the instance-describing feature of images.

In cooperation with Daniele Oelke, Christian Rohrdantz, and Andreas Stoffel we addressed the problem and developed a technique which represents the documentts key semantic as a mixture of images and important key terms[Str09b].

A more sufficient positioning of terms and images achieving closeness of related items was investigated (cooperation with Iris Adae and Martin Mader) by extending the popular technology of creating text clouds. Working on the base of news articles we addressed the tasks of finding good terms, finding good distance measures between terms, and finally finding good projection technologies of high dimensionally distance data into 2D positions.

Finally the layout and the arrangement of DocumentCards will profit form the acquired insights of the text cloud project. A more general approach of combined image and text representation of documents (NewsCards, BookCards,...) will be proposed.

Bibliography

[Str09a] H. Strobelt. *Advanced text visualizations.* In A. Avanes, D. Fahland, J. Geibig, S. Haschemi, S. Heglmeier, D. A. Sadilek, F. Theisselmann, G.Wachsmuth, and S.Weißleder, eds., des gemeinsamenWorkshops der Informatik-Graduiertenkollegs und Forschungskollegs 2009, p. 28, June 2009.

[Str09b] H. Strobelt, D. Oelke, C. Rohrdantz, A. Stoffel, D. Keim, and O. Deussen. *Document cards: A top trumps visualization for documents.* In IEEE Symposium on Information Visualization (InfoVis 2009), 2009.

3 GRK 1076/3: TrustSoft

Prof. Dr. Ernst-Rüdiger Olderog
Email: ernst.ruediger.olderog@informatik.uni-oldenburg.de
Carl von Ossietzky University of Oldenburg
Internet: `http://www.uni-oldenburg.de/trustsoft/en/`

Software increasingly influences our daily life, as we depend on a 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 broader 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. 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.

3.1 Multimodal User Interfaces for a Car Navigation System

Amna Asif (amna.asif@informatik.uni-oldenburg.de)
Supervisor: Susanne Boll

An activity of driving is multi-tasking and complex. A modern car is not only made-up of primary driving controls but also contains information systems, warning systems, and entertainment systems. A car navigation system is among one of the information systems. The visual and auditory interfaces are available in the current car navigation systems. The visual modalities are superior in a noisy environment but driving itself mainly requires a visual attention [2].The driver gets visual distraction from the primary task of driving while performing secondary tasks. It is difficult for driver to perceive information from a number of auditory resources simultaneously. The auditory modality can only be utilized for one task while driving. A cognitive workload is the *amount of mental effort required by a human operator in order to perform a task or tasks* [3]. The available channels are limited to deal with the complexity of tasks in automotive environment along with interacting with the car navigation system and results in the cognitive workload. In year 2007 around two million accidents happened in Germany according to statistics.

The driver can face problems of time sharing [8] and workload due to availability of a limited number of displays for many tasks in automobiles. Time sharing can be *a performing multiple tasks by using available modalities in a unit time*. According to MRT (Multiple Resource Theory) [7] the cognitive workload can be reduced by adding a new modality to the existing modalities. The tactile devices interact with the nerve ending in the skin [6]. We have proposed the tactile interface in the car navigation systems as an extension of the previous work of our research group [4]. The tactile display can be used as an alternative interface for the car navigation system to reduce the mental workload and the distraction. Our objective is to present direction and distance information with the vibro-tactile cues augmented with the non-speech interface.

We have conducted a user evaluation in a controlled environment to investigate the acceptability of tactile belts in the car navigation systems. The results show that tactile belts are acceptable by the drivers for perceiving directional information [1]. We have aimed to present direction and distance information for four scenarios that present different types of crossings. The different scenarios are composed of the crossing with a roundabout, cardinal directions and ordinal directions. We will evaluate the presentation of the direction and distance information with the help of vibro-tactile signals on the given scenarios. The car navigation will display the turn by turn spatial

information to the driver with the tactile signals. In our research process, we will proceed from a simple to the complex scenario. We will follow sequential steps of context analysis, design and encoding, evaluation, and analysis of the results for each of the scenario.
We are expecting the usable interface in the car navigation system that helps the driver to follow the route without imposing visual distraction and the cognitive workload.

Bibliography

[1] Amna Asif, Wilko Heuten, and Susanne Boll. Acceptance of tactile belts for directional information representation in automobiles. In 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI2009)in-cooperation with ACM SIGCHI Essen, 2009.

[2] Yujia Cao, Sandro Castronovo, Angela Mahr, and Christian Mller. On timing and modality choice with local danger warnings for drivers. In Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 2009.

[3] Grech, M. R.; Horberry, T. & Koester, T. Human Factors in the Maritime Domain CRC Press Inc, 2008

[4] Wilko Heuten, Niels Henze, Susanne Boll, and Martin Pielot. Tactile Wayfinder: A non-visual support system for wayfinding. In-NordiCHI'08, 2008.

[5] Martin Pielot, Dirk Ahlers, Amna Asif, Wilko Heuten, and Susanne Boll. Applying tactile displays to automotive user interfaces. In Proceedings of the Workshop on Automotive User Interfaces and Interactive Applications (AUIIA), 2008.

[6] Christopher M. Smith. Human factors in haptic interfaces. Crossroads The ACM Student Magazine, Special issue on human computer interaction, 3 (3):14-16, 1997.

[7] Christopher. D. Wickens. Processing resources in attention. In Processing resource in attention. London: Academic, 1984.

[8] Christopher. D. Wickens. Multiple resources and mental workload. Human Factors: The Journal of the Human Factors and Ergonomics Society, 50(3):449-455, 2008.

3.2 Achieving Fault Tolerance for Arithmetic Circuits with a Mutable RNS Monitor

Ahmad El Maamoun (a.maamoun@informatik.uni-oldenburg.de)
Supervisor: Prof. Dr. Martin Fraenzle, Prof. Dr. Wolfgang Nebel

Arithmetic circuits are widely used in signal processing, where they are predominantly used to compute polynomials which are composed of a succession of additions and multiplications. Like any other type of circuit, arithmetic circuits are subject to failures which can be caused by some physical transformations, such as the deterioration of the digital components. These failures are, if not tolerable, usually overcome by the introduction of redundancy. One famous type of hardware redundancy, triple-modular-redundancy, consists of including three copies of the system that operate in parallel which entail high energy and space consumption, which are two factors that can become highly inadequate in some limited power circuits. Another type of hardware redundancy can be realized by including some extra hardware in the system that would perform partial calculations (parity bits etc.) that would verify the main processor's results.

On the other hand, Residue Number Systems are systems that can represent large numbers by using a vector of smaller integers. These systems operate on the remainders of large numbers with respect to some Base numbers which should be co-prime amongst each other. The advantages of RNS lie in the reduction of the time needed to perform arithmetic operations, since the individual arithmetic operators will deal with smaller numbers.

Having said that, fault tolerance for arithmetic circuits can, in principle, be performed by executing the arithmetic operations in a residue number system, in parallel with the original arithmetic circuit, and then the issued result would be changed back to normal value and compared with the circuit's output for validation. The disadvantages in this method lie first of all, in the space requirements of a full RNS circuit running in parallel and, second, in the backward conversion from normal to residue numbering systems which consumes time and space. As for forward conversion, it can be achieved with a few small width adders, making it affordable in contrast to the backward conversion which is considered to be a limiting factor for residue number systems.

Furthermore, mere forward coding would permit to use only a single channel of the RNS vector, would abandon the need of a fully redundant RNS implementation on the chip. In addition, if this channel would be able to sequentially change its respective base then a close to 100% detection of all permanent errors can be achieved with low area overhead.

Arithmetic circuits and Residue Number Systems are chosen to be used together because the Arithmetic circuits involve the use of additions and

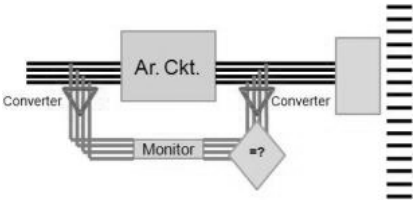

Figure 3.1: A modular representation of the RNS monitor

multiplications and rarely use division, which cannot be easily achieved in RNS. Our approach aims to provide fault tolerance for Arithmetic circuits, by having a RNS circuit running in parallel with the Arithmetic circuit monitoring for correctness the results issued by the computations. This monitoring RNS circuit will sequentially shift between a modulo-set of the form "2^n -1, 2^n, 2^n +1, 2^{n+1} -1" because this special set is proven to have easier forward conversion and easier arithmetic circuits that provide addition and multiplication. The choice of n depends on the arithmetic system at hand, in a way that the product of the moduli has to be greater than or equal to the largest number processed by the arithmetic circuit. As for the use of four moduli, it will ensure a complete fault coverage that might occur in the circuit, since some faults can be hidden from one base but visible to some other base.

The monitor will have a forward RNS converter that supports all of the four previously mentioned bases, which will be controlled by a "base switcher" that provides the monitor with the appropriate base in a definable cycle. Thus, the converter will receive the system's input and convert it to RNS mode and feed the result into the arithmetic sequence. The selected base will also be used by the arithmetic modulo operators, to perform the moduli operations.

Since this monitor uses one base at a time, and since all the bases are needed at one time to shift back to normal representation, this monitor will not be able to convert its output back to normal representation. It will instead convert the output of the Arithmetic circuit to RNS, and then compare the outputs generated from both systems, thus removing the need for a backward converter which is a major performance bottleneck for RNS.

Arithmetic components are the main building blocks of the Arithmetic circuit, but there are other components that will also need to be checked for faults like the controlling components, which can be monitored by standard fault detection methods.

For benchmarking our technique will be applied to a typical arithmetic circuit which is the Kalman Filter.

3.3 Self-Managed Resource Consumption Optimization by Near-Future User Behavior Prediction at Application Level

Muhammad Imran Asad Gul
(asad.gul@informatik.uni-oldenburg.de,iag@informatik.uni-kiel.de)
Supervisor: W. Hasselbring, O. Theel

Resources can be used in an optimal way if it is known how much they are needed. In our case it is energy consumed by a central processing unit (CPU) which is the major energy consumer in enterprise computation systems. Modern CPUSs are equipped with power saving modes called C-states i.e. C0, C1, C2 and C3. C0 is the active state where C1, C2 and C3 are the idle state. These are managed by operating system power management (OSPM). Depending on the situation, OSPM selects which C-state is be achieved. Higher the number of the C-state, higher is the entering and exit latency. Furthermore, performance issues prevents deep C-state to be achieved.

We present an approach to control CPU cycles based on active prediction of CPU usage pattern at application level. By prediction we mean, precise prediction of CPU activity for the next couple of micro seconds. Our framework consists of monitoring, time prediction model and optimizer. Monitoring continuously monitors the user bahavior and the application behavior and creates user behavior profile and resource usage profile. User behavior profile concerns with how users interact with the application and resource usage profile presents which components are being used while entertaining a request by the user or users. Based on these profiles, we have a time prediction model for the CPU(s) for the next couple of micro seconds. Time prediction activity is a continuous operation which defines which computational unit is needed as an active and for how long (in case of more than one CPU). The optimizer is an adaptation operation which controls the underlying computational units according to the time prediction activity i.e. only those computational units which are needed remain active, other can go to sleep and reactivate as and when required. As time prediction activity is purely based on user behavior profile and resource bahavior profile which are continuous learning operations. Certainty can be guaranteed in the CPU usage pattern.

3.4 Gesellschaftsrechtliche Anforderungen an Risikomanagementsysteme

Christian Jakob (christian.jakob@uni-oldenburg.de)
Supervisor: Prof. Dr. Jürgen Taeger

Unter dem Eindruck bedeutsamer Unternehmenskrisen schuf der Gesetzgeber Vorschriften, die eine frühzeitige Erkennung existenzbedrohender Gefährdungen sicherstellen sollten. Die Diskussion um Risikomanagementpflichten kristallisiert sich insbesondere an der Rechtsform der Aktiengesellschaft, für die mit dem KonTraG die Risikomanagementpflicht des Vorstands eine teilweise Hervorhebung durch § 91 Abs. 2 AktG erfahren hat. Gesetzliche Vorschriften mit Bezug zum Risikomanagement finden sich aber auch an anderen Stellen, etwa im Aufsichtsrecht für den Finanz- und Versicherungssektor. Welches Pflichtenportfolio sich ergibt, hängt maßgeblich von der Rechtsform der unternehmenstragenden juristischen Person ab. Besondere Relevanz erhält das Risikomanagement als Leitungspflicht in Kapitalgesellschaften, wo es sich aus der Sorgfaltspflicht der Geschäftsleitungsorgane ergibt, aufgrund derer sie insbesondere zur Schadensabwehr und Vorteilsmehrung gegenüber der Gesellschaft verpflichtet sind.

Es bestehen unterschiedliche Auffassungen darüber, wie ein betriebliches Risikomanagementsystem konkret gesetzeskonform auszugestalten ist. In der gesellschaftsrechtlichen Literatur wird vertreten, eine Organisationspflicht besteht, bei deren Erfüllung der Vorstand auf einen weiten Ermessensspielraum zurückgreifen könne. Der gesetzlich gebotene Ausbaugrad des Risikomanagements, ist hiernach ausgehend von der Komplexität des Unternehmens zu ermitteln. Vertreter der betriebswirtschaftlichen Literatur nehmen hierzu eine Gegenposition ein. Sie gehen davon aus dass die Einrichtung eines umfassenden Risikomanagementsystems verlangt wird.

Die große Bedeutung der IKT im Unternehmen strahlt auf das Risikomanagementthema aus. Existenzbedrohungen des Unternehmens, die ihren Ursprung in der IKT-Infrastruktur oder in anwandten Softwarelösungen finden, müssen angemessen abgebildet werden. Gleichzeitig bietet sich der Einsatz von Informationstechnik auch als Instrument zur Kontrolle und Steuerung von Risiken aller Art an. Der IKT-Einsatz bei dem Betrieb von Risikomanagementsystemen kann erhebliche Effizienzsteigerungen mit sich bringen. Mögliche Strukturvorteile ergeben sich bei der Unterstützung der Risikoinventur, der plausiblen Auswahl und Analyse von Risikoindikatoren, der Simulation und Klassifizierung von Schadensmodellen sowie der hierauf aufbauenden aktiven Risikosteuerung. Die technische Umsetzung verlangt grundlegend nach der Definition technischer Anforderungsspezifikationen, die nur abhängig vo nden gesetzlichen Anforderungen an ein Risikomanagementsystem formuliert werden können.

3.5 Dependable Data Management in Dynamic Distributed Systems

Kinga Kiss Iakab (kinga.kiss-iakab@informatik.uni-oldenburg.de)
Supervisor: Prof. Dr.-Ing. Oliver Theel, Prof. Dr. Eike Best

Storing and managing data via replication is a well-known technique for improving availability in distributed systems. In this context, a distributed system consists of several interconnected processes, each managing a copy of the data object. In dynamic distributed systems, processes are allowed to fail and recover, but also to deliberately leave and join the system. We assume, that the dynamic distributed system is characterized by a uniform availability of processes and a churn rate, i.e., the percentage of processes that leave and join the system per time unit.

Two operations provided by a replica control algorithm are defined on a data object: writing new data on a subset of replicas and reading data from one up-to-date replica. Read and write operations can be executed using quorums, which are subsets of the set of processes in the system. Replica control algorithms can be classified into two subsets: pessimistic and probabilistic strategies, relying on the theoretical concepts of strict quorum systems and probabilistic quorum systems, respectively. If we take a query-answering distributed system as an example, pessimistic strategies would return consistent data with a consistency probability of 1 or no data at all, if correct data is not available. On the other hand, probabilistic strategies would always return probabilistically consistent data associated with a consistency probability of 1 or less. In the case of pessimistic strategies, there are already solutions for changing replication strategies at runtime to cope with the dynamics of the distributed system. These dynamically adapting pessimistic strategies have the benefit of ensuring consistent read and write operations, but with the downside of high costs, in terms of number of messages, in highly dynamic systems. Our novel approach consists of analyzing and dynamically adapting probabilistic strategies for reducing the costs and improving the operation availability by relaxing the consistency of data.

After analyzing the data consistency vs. operation availabilities trade-off for some probabilistic strategies [1], the objective is to exploit the trade-off analysis when developing dynamically adapting probabilistic strategies, and to compare our approach to dynamically adapting pessimistic strategies.

Bibliography

[1] K. Kiss Iakab, C. Storm, O. Theel *Consistency-driven Probabilistic Quorum System Construction for Improving Operation Availability.* Springer, 2010.

3.6 Diagrammatic Specification and Verification of Mobile Real-Time Systems

Sven Linker (sven.linker@informatik.uni-oldenburg.de)
Supervisor: Prof. Dr. Ernst-Rüdiger Olderog

The spread of mobile real-time systems, i.e. systems which have to react within certain time bounds and to respect spatial constraints and relations, is increasing. Such systems would include e.g. cars organizing themselves automatically as platoons as well as robots and airplane controllers. To ensure correct behaviour in safety-critical environments, formal specification and verification techniques have to be developed. Even though logics able to express temporal and spatial properties have been presented [AB02], these formalisms have not yet been well-studied. Furthermore, specifications and proofs in these logics tend to be hard to read.

Illustration with diagrams is an often used technique to enhance the understanding of mathematical ideas and proofs. Furthermore, diagrams are used for engineering purposes, i.e. for specification. However, such diagrams lack a formal semantics needed to verify correct system behaviour. Recent research [AB96] has shown that diagrams are not only heuristic tools, but that formal reasoning with diagrams instead of formulae is indeed possible.

The aim of our work is to narrow the gap between engineers and scientists, where the former use diagrams and formulae to specify mobile real-time systems, while the latter are in need of formal techniques to verify the correct behaviour of these systems. Hence we develop a diagrammatic language able to represent the crucial aspects of mobile real-time systems. This language is given both a formal syntax and semantics to allow for correct diagrammatic reasoning. The semantics is defined in Shape Calculus [Sch06], a spatio-temporal logic suited for the verification of mobile real-time systems.

Bibliography

[AB02] Marco Aiello and Johan van Benthem. *A Modal Walk Through Space*. Journal of Applied Non-Classical Logics, 2002.

[AB96] Gerard Allwein and Jon Barwise. *Logical Reasoning with Diagrams*. Oxford University Press, 1996.

[Sch06] Andreas Schäfer. *A Calculus for Shapes in Time and Space*. Theoretical Aspects of Computing - ICTAC, 2004.

3.7 Unmasking Fault Tolerance: Masking vs. Non-masking Fault-tolerant Systems

Nils Müllner (nils.muellner@informatik.uni-oldenburg.de)
Supervisor: Prof. Dr.-Ing. Oliver Theel

Fault tolerance is not only important where robustness and reliability of an application or tool is of priority concern, but also for the comfort of services of lesser importance like soft realtime applications where availability is an issue. Hence, tremendous efforts are spent in order to design and build fault tolerant applications and devices that are competitive on the open market. Nevertheless, as devices and requirements become more and more complex, it becomes even more challenging to retain a certain threshold of fault tolerance. On the other hand, new objectives arise such as low energy consumption or service coverage, that even contradict to spending resources on fault tolerance. There are four sorts of fault tolerant systems defined throughout literature: intolerant, fail-safe, non-masking and masking fault tolerant systems. An intolerant system works without any guarantees or assertions about its correctness, i.e., if it works at all. A failsafe system is equipped with detectors that trace faults and bring the system to a halt to prevent it from working off its specifications, i.e., safety properties are never violated but liveness properties might be. On the contrary, non-masking fault tolerant systems employ correctors to maintain an operational status and support repair mechanisms while they do not necessarily comply with their program specification during a repair phase. Masking fault tolerant systems finally comprise of both, detectors and correctors, to delay responses if faults have been detected until they have been corrected. Notably, correctors do not necessarily explicitly require detectors. This work focuses on the gap between non-masking and masking fault tolerant systems. While recent literature focused on the compositional design to be able to add new fault tolerance features to comply with changing requirements, i.e., the extremes of intolerant, fail-safe, non-masking and masking, this thesis takes a steeper look on developing a (probabilistically) masking fault tolerant system and how it can be accomplished in a reasonable inexpansive manner.

3.8 End-User-Requirement-Driven Design of Wireless Sensor Networks

Felix Jonathan Oppermann (oppermann@informatik.uni-oldenburg.de)
Supervisor: Prof. Dr.-Ing. Oliver Theel, Prof. Dr. Wilhelm Hasselbring, Prof. Dr.-Ing. Wolfgang Nebel

As of today, advantages in microelectronics permit to equip individual sensors with limited computing and wireless communication capabilities. An ad-hoc network formed out of such sensor nodes is known as wireless sensor network (WSN). Currently, WSNs penetrate an increasing number of application areas, like habitat monitoring, catastrophe management, and home automation. As WSNs need to operate efficiently in a wide rage of different applications they usually need to be tailored to the mission at hand. This inhomogeneity and the severe constraints, for example in terms of available energy and computation power, render the design and implementation of WSNs difficult and error prone. This is especially true for the envisioned end-users, like biologists or geologists, and thus is a limiting factor for a more widespread adoption of WSN technology outside of the scientific community.

A vast set of solutions to specific problems is already available, but to select the right components for a specific design is far from easy. One needs to make sure that this selection not only fulfills the given requirements, but also allows for efficient operation and works well together in a single application.

In order to make the design process of WSNs easier and more reliable for unacquainted end-users, a methodology is required, to synthesize a possible structure for an applicable WSN based solely on the requirements and constraints of the intended task and the available components. As a first step, the end-user specifies his requirements and important parameters of the mission he intends to solve with the WSN. Based on this mission specification a selection of suitable components implementing the required functionality is obtained. To cope with the resource constrained nature of WSN it is important that this selection is optimized in terms of, for example, energy consumption, memory requirements, and code size. The ultimate goal is to offer the end-user a tool to automatically generate a mission specific selection, out of the available hardware and software components, based on this task definition at "the push of a button."

Such a framework eases the decisions to be made during the design of a sensor network and allows end-users with little WSN specific knowledge to successfully create a mission specific WSN. Still, this approach does not sacrifice the necessary flexibility of tailoring the WSN to the specifics of the intended application. We hope that this work is a step towards a more widespread use of WSN technology.

3.9 Correctness of Graph Programs Relative to Hyperedge Replacement Conditions

Hendrik Radke (hendrik.radke@informatik.uni-oldenburg.de)
Supervisor: Prof. Dr. Annegret Habel, Prof. Dr. Ernst-Ruediger Olderog

As software systems grow more complex, there is a growing need for design concepts that allow an intuitive overview of a system. Visual design concepts try to reach this aim by relying on the visual capabilities of human beings. The desired system properties can be formulated in a graphical way to ease their understanding. One approach to visual modeling are graph transformation systems. The states of a regarded real-world system are modeled through graphs, and changes to the system state are described by graph transformation rules. The rules can be combined to form more complex graph programs as described in (Habel/Plump 2001). First-order properties of the system can be expressed through graph conditions in the sense of (Habel/Pennemann 2009). There are automated tools to check, semi-decidable, whether a graph satisfies a given condition (see Pennemann 2009).

However, the expressive power of graph conditions is not sufficient to formulate certain properties that many real-world problems impose. For instance, it is not possible to express the property "there is a path from node 1 to node 2" with graph conditions. Therefore, an extension to graph conditions that can capture such non-local properties is desired.

We propose an extension of graph conditions by variables in the form of hyperedges. These hyperedges can be replaced by certain graphs, as specified by a hyperedge replacement system.

We call the extension of graph conditions with hyperedges and subconditions *Hyperedge Replacement (HR^+) graph conditions* and show that these conditions are more expressible than monadic second-order (MSO) formulas over graphs. Also, these conditions can be used together with graph programs to build graph specifications in form of a Hoare triple (precondition, program, postcondition). This triple can be transformed into a single condition $pre \Rightarrow wp(program, post)$, where wp transforms the postcondition and the program into a weakest precondition, following the approach of (Dijkstra, 1976). The combined condition can be checked for validity with a theorem prover, enhanced with hyperedge replacement.

In order to show the practical merit of the results, several case studies are performed. We use HR^+ conditions and graph programs to verify a protocol of car platooning (Hsu et al. 1991), where a number of cars on a lane form linear topology networks that can be merged, split or change lanes. Furthermore, HR^+ conditions are applied to the problem of C++ template instantiation type checking. As a third example, we use HR^+ conditions to express and check OCL constraints for UML diagrams.

3.10 Compliance in Kapitalgesellschaften

Bernd Schmidt (bernd.schmidt@uni-oldenburg.de)
Supervisor: Prof. Dr. Jürgen Taeger

Compliance bedeutet die Herstellung von Konformität mit rechtlichen Anforderungen. Als elementarer Bestandteil guter Corporate Governance, wird Compliance zunehmende Aufmerksamkeit in der juristischen, aber auch in der wirtschaftswissenschaftlichen Diskussion gewidmet. Eine ausdrückliche gesetzliche Regelung zur Compliance-Pflicht findet sich in § 33 Abs. 1 Nr. 1 WpHG sowie in den Informationsvorschriften des Deutschen Corporate Governance Kodex. Auch darüber hinaus kommt der Beachtung rechtlicher Vorgaben im Rahmen unternehmerischer Tätigkeit eine herausgehobene Bedeutung zu. Bei schlechter Compliance drohen Schäden für das Unternehmen durch die Verhängung von Bußgeldern, aber auch durch Rufschäden, Gewinnabschöpfungen im Rahmen von Strafverfahren, börsenrechtliche Maßnahmen und Auftragssperren. Hierdurch können Unternehmen kurz-, mittel- und langfristig geschädigt werden.

Ziel der Arbeit ist es, die rechtlichen Implikationen aus der Bedeutung von Compliance herauszuarbeiten. Hierfür wird zunächst die Pflichtenlage der Organe von Kapitalgesellschaften betrachtet. Ein Schwerpunkt der Darstellung soll auf dem Geschäftsführungsorgan liegen. Die Compliance-Diskussion ist für die Gesellschaftsform der Aktiengesellschaft am weitesten fortgeschritten, so dass auch für diese Arbeit von ihr ausgegangen und hierfür erarbeitete Grundsätze in einem folgenden Schritt auf die GmbH und die eG übertragen werden sollen. Ziel ist es, so ein vollständiges Bild rechtlicher Anforderungen an die Herstellung von Compliance in Kapitalgesellschaften zu erhalten. Hierbei sollen insbesondere die Sorgfaltspflichten von Geschäftsleitungsorganen, die Risikomanagementpflicht gem. § 91 Abs. 2 AktG im Hinblick auf die Herstellung von Compliance sowie allgemeine Legalitäts- und Aufsichtspflichten gem. § 130 OWiG berücksichtigt werden.

Ausgehend von dem so herausgearbeiteten Pflichtenstandard soll der zweite Schwerpunkt auf der Darstellung von Maßnahmen zur Erfüllung von Compliance-Pflichten liegen. Hierbei soll die Delegation von Verantwortlichkeiten im Geschäftsführungsorgan und in nachgeordneten Unternehmensteilen betrachtet werden. Angesprochen sind hiermit das Compliance-Ressort im Geschäftsführungsorgan, der Compliance-Officer und die Arbeitnehmer. Bestandteil der Darstellung soll auch die Kontrolle und Steuerung von Compliance-Risiken sowie die disziplinarische Durchsetzung gegenüber den Unternehmensangehörigen sein. Ziel dieses zweiten Teils soll es insbesondere sein, rechtliche Grenzen bei der Erfüllung von Compliance-Pflichten aufzuzeigen.

3.11 Online Capacity Management for Increased Resource Efficiency of Software Systems

André van Hoorn (van.Hoorn@Informatik.Uni-Oldenburg.DE)
Supervisor: Prof. Dr. Wilhelm Hasselbring

Resource efficiency is an increasingly important internal quality attribute of software systems. While the related attribute performance is mainly concerned with metrics quantifying timing behavior and resource usage characteristics, resource efficiency is a measure of a system's resource usage economy. Many software systems are exposed to varying workload conditions significantly influencing its timing behavior. However, the capacity of those systems is typically managed in a static and pessimistic way, causing temporarily underutilized resources, e.g., application servers, during medium or low workload periods.

SLAstic, the self-adaptive approach for online capacity management developed in this work, aims to increase the resource efficiency of distributed component-based software systems employing architectural runtime reconfiguration. A software system is equipped with reconfiguration capabilities that allow to control a system's performance and efficiency properties at runtime in an elastic way, e.g., by migrating and (de-)replicating software components, and (de-)allocating server nodes. Architectural models play an important role in the approach since they are used to specify the system assembly, deployment, instrumentation, reconfiguration capabilities, performance properties etc. At runtime, these models are continuously updated and used for online quality-of-service evaluation, e.g., workload forecasting and performance prediction, in order to determine required adaptations and to select appropriate reconfiguration plans. A prototype implementation of our adaptation framework [1] is used to quantitatively evaluate the approach by simulation and lab experiments, based on an industrial case study system [2].

Bibliography

[1] A. van Hoorn, M. Rohr, A. Gul, and W. Hasselbring. *An adaptation framework enabling resource-efficient operation of software systems*. Proc. ICSE WUP '09, pp. 41–44. ACM, 2009.

[2] A. van Hoorn, M. Rohr, W. Hasselbring, J. Waller, J. Ehlers, S. Frey, and D. Kieselhorst. *Continuous monitoring of software services: Design and application of the Kieker framework*. TR-0921, CS, Univ Kiel, 2009.

4 GRK 1194: Self-organizing Sensor-Actuator-Networks

Prof. Dr.-Ing. Uwe D. Hanebeck
Email: Uwe.Hanebeck@kit.edu
Karlsruhe Institute of Technology (KIT)
Internet: `http://www.grk1194.uni-karlsruhe.de/`

Sensor-actuator-networks consisting of a large number of miniaturized and autonomous nodes offer novel possibilities to collaboratively observe and control distributed phenomena. The individual sensor-actuator-nodes are densely deployed either inside the phenomenon or very close to it. Hence, the network provides a good spatial resolution that can be adapted depending on the dynamics of the phenomenon to be observed. Besides the observation task, the individual nodes are able to interact with the phenomenon by means of integrated actuators. Because of 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 of the nodes, impose severe performance limits with respect to measurement accuracy, communication bandwidth, and processing power. In addition, the topology of the network is continuously changing due to the intrinsic mobility of the nodes, possible node failures, and selective switch-off for energy preservation. For these reasons, new challenges result for communication in sensor-actuator-networks so that classical mechanisms cannot be directly employed. Similar facts hold for information processing in these networks as only limited storage and computing resources are available. Furthermore, each sensor-actuator-node can only access partial information about the considered phenomenon, which together with the limited communication capacity, necessitates a decentralized form of information processing. In summary, sensor-actuator-networks differ from classical communication networks, where typically a large bandwidth and a fixed topology can be assumed. Information processing is also different from classical paradigms, which assume centralized processing, high measurement accuracy, and sufficient storage and computing resources. As a result, sensor-actuator-networks require new architectural models and new methods of operation.

4.1 Privacy and Reciprocity in Sensor-Networks

Martin Heine (Martin.Heine@kit.edu)
Supervisor: Prof. Dr. Klemens Böhm

In the past few years the concepts and the technology for managing, processing and communicating data in sensor-networks have advanced greatly. This is also true for query processing in sensor-networks. Spatio-temporal predicates improve the ease of formualting queries to find objects (e. g. cars, mobile phones, etc.) in a given area. Since a lot of objects can be directly assigned to a person, privacy is gainig importance also in relation to sensor-networks.

Most of the work about privacy is done regarding privacy of users against service providers. A common use-case for sensor networks are location based services, that allow the sharing of location-dependent data between users. In such a case the privacy issue is between the users themselves. In this scenario it is important to find a way to ensure the users' privacy, while facilitating the exchange of data, since this is the very purpose of the system.

In order to consider the users' privacy preferences when processing queries, it is necessary that each user specifies a privacy policy. Conventional privacy policies allow the user for each part of the data he wants to share to specify a group of users that is allowed to access the data. To use this kind of policies effectivly, knowledge about the users and groups is necessary. In large system with lots of users, this can be difficult to achieve and it can also be a hindrance when users are looking for other users willing to share similar data.

We introduce a special kind of privacy policy called reciprocal privacy policy. These policies allow users to formulate policies like "I share my data with everyone of my colleagues, who shares his data with me.". Such policies ensure that data is shared reciprocally and no one-sided observation is taking place. The exact condition, i. e. which kind of data the other users have to share to fullfill the policy, can be further specified. In contrast to conventional policies reciprocal policies do not always reveal data, but only if their condition is fullfilled.

We are looking into ways to formulate reciprocal privacy policies and to check their fullfillment. Another focus of our work is the processing of queries with consideration of reciprocal privacy policies, especially range and top-N queries. While range queries can be evaluated relativly straightforward, effective processing of top-N queries requires algorithms that integrate the computation of the visibility of data into the conventional query evaluation. Regarding the top-N query processing we propose different algorithms using tree-structured spatiotemporal indices as well as histograms.

4.2 Managing Energy Efficient and Reliable Communication in Wireless Sensor Networks

Jens Horneber (horneber@kit.edu)
Supervisor: Prof. Martina Zitterbart

Wireless Sensor Networks consist of a plenty of resource constrained sensor nodes with heterogeneous hardware. Future sensor networks could use the provided hardware diversity not just for one, but for multiple application.

To switch between different applications a common layer with reusable software elements is needed. Therefore a service oriented approach is usually chosen. Every service can be utilized by different applications on top to provide functionality with increased complexity. Each application might have different requirements for communication like reliability, energy-efficiency, measuring precision, latency, security and privacy constraints or a specialized topology. Using services can help to break down application requirements for communication in sensor networks to service requirements. Hence, a self-organizing service management infrastructure has to be researched to realize unattended sensor networks and satisfy application requirements. Instead of sending management messages in concurrency to applications, an opportunistic approach for management messages promises less additional overhead. For example opportunistic communication profits by using aggregation and gossiping or sending messages to multiple potential receivers instead of specific ones. Obviously service requirements are often conflictive, so self-organizing management has to find a trade-off in many cases. Accordingly, new adaptable communication protocols are needed to support such dynamic in-network reconfiguration.

To evaluate and optimize network management and related communication protocols concerning energy efficiency a dedicated sensor network testbed is in development [1, 2]. The primary goal is the research of energy efficient and robust management mechanisms to build a suited management framework for wireless sensor networks.

Bibliography

[1] A. Hergenröder, J. Horneber and J. Wilke. *SANDbed: A WSAN Testbed for Network Management and Energy Monitoring*. Proceedings of 8. GI/ITG KuVS Fachgespräch "Drahtlose Sensornetze", 2009.

[2] A. Hergenröder, J. Horneber, D. Meier, P. Armbruster and M. Zitterbart. *Demo Abstract: Distributed Energy Measurements in Wireless Sensor Networks*. SenSys '09: Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, 2009.

4.3 Nature-Inspired Wireless Sensor-Actuator-Networks

Chih-Ming Hsieh (chih-ming.hsieh@kit.edu)
Supervisor: Prof. Dr.-Ing. Jörg Henkel

In computer systems, people usually have to deal with optimization problems depending on the constraints they encounter. This is especially true in designing a wireless sensor-actuator-network (WSAN) due to its unique characteristic. Numerous energy limited sensor nodes composing a large-scale ad-hoc network, which interacts with the environment in a unattended manner, leads us to the issues such as network lifetime, tolerance to internal and external changes, etc. An optimal way to solve all these issues is however very challenging. Fortunately, many of the natural systems which have evolved throughout the ages are inspiring and indicate an organic way of solutions, such as ant colony optimization, etc. [1, 2]

To adapt to the changes while fulfilling the requirements, a system should have autonomous (Self-X) properties while providing controlling mechanisms to human operators. Therefore in this project, a nature-inspired WSAN system equipped with capabilities of interactions between agents, sensor nodes, and human operators is being developed. At micro level the agents make decisions based on the local information to accomplish their task and a human operator can regulate the system by setting the goal at macro level via control interfaces. An environment which accommodates the agents takes care of the communication between nodes and provisions the vicinity information which can help the agents to make proper decisions. Moreover, an evolutionary mechanism will be developed to improve the efficiency of the system by means of agile adjustment over time. As a result, the system can complete the goal given by human operators while extending its lifetime.

Evaluation of the proposed system will be performed on both simulation and testbed deployment. The system will be implemented using TinyOS and testbed deployment is planned. Accordingly, a simulation environment corresponding to the testbed will be designed as well. This simulation environment, which contains an emulation of the sensor node, can provide a cycle accurate estimation and speed up the development of the system.

Bibliography

[1] M. Dorigo and T. Stützle. *Ant Colony Optimization*. The MIT Press, 2004.

[2] P. Boonma and J. Suzuki. *BiSNET: A Biologically-Inspired Middleware Architecture for Self-Managing Wireless Sensor Networks*. Elsevier, 2007.

4.4 Decentral Task Processing by Cooperation and Interaction: Models Assisting in Local Algorithm Design

Alexander Kettler (alexander.kettler@kit.edu)
Supervisor: Prof. Dr.-Ing. H. Wörn

In order to implement a large network of sensor nodes, the nodes have to form a decentralised system, with each node being autonomous up to a certain degree and acting on the basis of locally available information. Therefore, methods of self-organisation and emergence shall be used, that guarantee decentralisation, robustness and scalability. For the investigations, robotic swarms are considered, which are large groups of small robots with very restricted abilities. The development of the control algorithm that has to be executed locally on each robot in order to achieve a certain global behaviour of the robotic swarm has proven to be difficult, since the classical reductionistic approach is only partially applicable to this kind of problem. Due to this, the resulting behaviour of the robotic swarm often contradicts the initial imagination of the program developer because the effects of the many interactions between the robots cannot be easily anticipated.

One approach of dealing with this problem is to support the algorithm developer with models that allow predictions of the global behaviour long before the implementation on the real hardware or running extensive simulations. This has already been discussed in the literature several times. The model approach pursued here is distinguished from other approaches by the explicit representation of space and velocity of the robots and is therefore well suited for scenarios where the amount of deterministic motion in certain directions is no longer negligible compared to the random collision avoidance moves due to other robots.

The basic model of the robot positions is based on the Boltzmann-Equation that describes the macroscopic space-temporal evolution of an ensemble of particles as a macroscopic density function in phase-space. This physical model was extended to a generic model of communicating robot groups. A formal derivation for generic parameters like the propability of a collision avoidance move has been developed. The model approach has a variety of applications, however, the adaptation to a specific control algorithm is a demanding modeling step.

The proposed model will be validated against several swarm robotic scenarios, both in simulation and with real hardware. Therefore, a new swarm robot has been developed, that especially fullfills the demands of the scenarios under investigation and allows to execute experiments in an easy and automated fashion.

4.5 Scheduling and Ressource Allocation in Decentralized Sensor-actor-networks

Daniel Lyons (lyons@kit.edu)
Supervisor: Prof. Dr.-Ing. Uwe D. Hanebeck

In spacious sensor-networks consisting of a large number of typically small and cheap sensor-nodes, these nodes in general do not have extensive energy resources or processing capabilities. Thus, when designing algorithms and methods for sensor-networks, one is faced with strict energy constraints.

In order to collect as much information about a phenomenon observed by the network as possible, it is desireable that every sensor takes measurements at every time instance. When acting under energy contraints this is not applicable though. On the one hand, turning on the measurement device of a sensor consumes energy resources. On the other hand, once a sensor makes a measurement it will be required that the outcome of the measurement is transmitted to a fusion center or to other sensor-nodes for information fusion. Hence, another factor of energy consumption is turning on the wireless communication device for transmission of data obtained by the measuring process or receiving measurement data from other sensors. Due to the energy constraints not every sensor in the network should be measuring at every time instance. Subproject I.4 is concerned with finding a balance between the information gained through sensor measurements and the overall lifetime of the network under consideration of communication constraints in a decentralized way. The decentralized scheduling approach promises a further reduction of communication overhead and an improved scalability in comparison to centralized sensor scheduling.

In order to deal with stochastic uncertainties, a suitable method for sensor scheduling is a model-based concept. By predicting the impact of future measurements on the current state estimate and assessing the expected informational benefit, sensor scheduling dynamically selects optimal sensor configurations for measurements. The centralized model predictive scheduling algorithm considers all possible sensor configurations in an exhaustive tree search and evaluates their information gain. For decentralized sensor scheduling not all sensor configurations may be known and hence optimality of the found solution cannot be guaranteed. It is therefore necessary to give tight upper bounds on the loss of performance of the decentralized scheduling approach.

4.6 Decentralized Reconstruction of Continuous Distributed Phenomena based on Discrete Measurements

Benjamin Noack (noack@kit.edu)
Supervisor: Prof. Dr.-Ing. Uwe D. Hanebeck

One of the most important fields of application for sensor-actuator-networks is the reconstruction of continuous distributed phenomena based on discrete measurements, which are obtained from a large number of miniaturized sensor nodes. Since the measurements are in general noise-corrupted, recursive Bayesian estimators are used to obtain valuable estimates for the state of the considered phenomenon. But, employing a multitude of spatially distributed sensor systems makes it difficult to apply centralized estimation techniques. Although they can usually guarantee good estimation results, they will lead to a high communication load and, of course, require a central node for the information processing. Therefore, the deployment of decentralized state estimators is aspired, so that information can be processed locally on single sensor nodes instead of aggregating information in the data sink. In the first dissertation in subproject I1, an efficient concept for simultaneously reconstructing and identifying linear phenomena has been presented that can account for unknown correlations. The study and derivation of reconstruction and parameter identification techniques for nonlinear phenonema under unknown stochastic dependencies is the focus of this dissertation. At this, the combination of stochastic and set-theoretic estimation concepts enables us to consider imprecise or even unknown model parameters and to analyse robustness and sensitivity. As a first step, [NKH09] provides a method for simultaneously treating stochastic and systematic uncertainties.

Bibliography

[NKH09] Benjamin Noack, Vesa Klumpp, Uwe D. Hanebeck. *State Estimation with Sets of Densities considering Stochastic and Systematic Errors*. Proceedings of the 12th International Conference on Information Fusion (Fusion 2009), Seattle, 2009.

4.7 V2X Simulation Environment for Comprehensive Design Space Exploration Verification and Test

Christoph Roth (christoph.roth@kit.edu)
Supervisor: Prof. Dr.-Ing. Juergen Becker

Actual research in the area of Vehicle-to-X mainly focusses on protocol and application design as well as software stack development. Implementation and vehicle integration is neglected. Therefore within previous work (subproject H2.1) a V2X communication architecture concept has been developed [2]. For functional verification of the V2X system, up to now a traffic simulator is used which allows for simple functionality and performance tests by stimulating the V2X system using network packets [3]. Subsequently within this work (subproject H2.2) a scalable simulation environment that enables comprehensive design space exploration verification and test of the V2X system architecture is developed [1]. The main goal of this work is to provide the foundation for the test of hardware models as well as real hardware with the help of existing communication-, channel- and mobility-models and for the verification of functionality and performance of V2X hardware, well-matched to the considered application scenario.

The simulation environment comprises three domains namely Network Domain, Traffic Domain and System Domain which are established by the flexible interconnection of three types of simulators: traffic simulators, network simulators and system simulators. By integrating SystemC [6] simulators of V2X systems into an environmental simulation, an observation of their behavior and interaction is possible, which is the basis for a comprehensive design space exploration. Real-time execution facilitates the connection of real V2X hardware for advanced verification. A variety of different test cases based on different environmental scenarios are imaginable which span over several orders of magnitude and therefore demand for high computational power. To cope with the computational complexity, concepts for scalability of the overall simulation are developed within this work. These are based for example on multi-resolution modelling [5] or distributed simulation. Therefore in the context of simulation development, the advantages and disadvantages of the techniques have to be evaluated. It is investigated when to apply which technique and how. A first step is the identification of existing performance bottlenecks. Based on this, the mentioned techniques are developed and implemented. To test their applicability different application scenarios have to be set-up which allow for performance evaluation. The development of the techniques and their evaluation is done in respect of the opportunity for real-time execution.

In general, the interconnection of the simulators will be established by the so called High Level Architecture (HLA) [4] which is a standard specifying the functionality and the interfaces that are necessary for distributed and coupled discrete event-based simulations. Each simulator instance (also called Federate) is connected to a Runtime Infrastructure (RTI) via Ambassadors. The HLA then manages data distribution and timing synchronization.

When interconnecting different types of simulators, one main problem in the context of simulation application is the configuration of the overall simulation prior to the execution of the application scenario. A lot of different configuration-files are needed for each simulator and often have to be edited manually. Especially in the case of greatly varying test cases and the application of multi-resolution modelling, test case generation can become tremendously complex since one must decide which resolution to use in which domain without loosing too much accuracy. Because of that, automatic generation of application scenarios by combining appropriate simulators should be supported by a software toolflow. First investigations into this direction and inital proof of concept implementations are also planned.

Bibliography

[1] Roth, Sander, Huebner, Becker. *Car-to-X Simulation Environment for comprehensive Design Space Exploration Verification and Test.* SAE World Congress 2010, 2010.

[2] Sander, Glas, Roth, Becker, Mueller-Glaser. *Design of a Vehicle-to-Vehicle Communication System on Reconfigurable Hardware.* IEEE (FPT 2009), 2009.

[3] Sander, Glas, Roth, Becker, Mueller-Glaser. *Testing of an FPGA-based C2X-Communication Prototype with a Model Based Traffic Generation.* IEEE (RSP 2009), 2009.

[4] IEEE. *Standard for Modeling and Simulation High Level Architecture.* IEEE Standard 1516, 2000.

[5] Adelantado, Bonet, Siron. *Multiresolution Modeling and Simulation with the High Level Architecture.* 12th European Simulation Symposium, 2000.

[6] IEEE. *Standard SystemC Language Reference Manual.* IEEE Standard 1666, 2005.

4.8 Lifetime Optimization of Wireless Sensor Networks

Dennis Schieferdecker (schieferdecker@kit.edu)
Supervisor: Prof. Dr. rer. nat. Peter Sanders

A *wireless sensor network* (WSN) is a union of small nodes, each equipped with a means of collecting data, processing and communication. Usually, they are only battery-powered. Thus, a considerate use of the available limited energy is important to prolong the lifetime of the network. Optimiziations to yield an energy-efficient sensor network can be applied on many different levels, starting from the selection of power-efficient hardware components and ending with applications that take tight energy constraints into consideration.

The thesis focuses on the algorithmic aspects of the problem of *optimizing the lifetime of wireless sensor networks*. We identify reasonable use-cases of sensor networks and model the respective systems on an abstract level to analyze them from the perspective of *algorithm engineering*. Our primary goal is to provide interesting algorithms for several applications and to give provable good performance guarantees for them. Our studies are focused on two major topics, *scheduling* and *routing*. Both of them also feature many subproblems that are interesting on their own like *detecting topological holes* in the network.

Scheduling. The nodes of a sensor network perform their tasks individually and in concert with the other nodes. But usually, not all of them have to be active all the time. Thus, considering the intended task of the network, it is reasonable to provide a schedule for each sensor node to minimize the energy consumption of the nodes until the are actually needed.

Routing. Communication between sensor nodes is one of the most energy-consuming operations in a sensor network. Thus, it is important to identify and use short and energy-efficient transmission paths for exchanging information between the nodes. But it is also important that individual nodes are not overused and their energy resources depleted prematurely. Furthermore, the arising communication patterns are usually application-dependent and have also to be taken into account.

We focus on algorithms that provide optimal or almost optimal results for static environments and adapt them to work efficiently on network topologies that change with time (i.e. movement or breakdown of nodes). Another focus is providing decentralized variants of our algorithms that work on independent nodes and do not require a single main instance.

4.9 Infrastructureless Localization in WSN

Johannes Schmid (Johannes.Schmid@kit.edu)
Supervisor: Prof. Dr.-Ing. Klaus D. Müller-Glaser

Localization of nodes in wireless sensor networks (WSN) is considered an important research challenge and draws considerable interest for several years now [1], [2], [3]. It rarely makes sense to measure a phenomenon without knowledge of its position [4]. Thus, a lot of WSN applications depend on a correlation of the measured data with the location of the measurement which is often referred to as WSN localization. Within this context, it is the goal of this subproject (H1.2) to explore the possibilities of ad-hoc node localization if little or no information of the environment at all is established previously.

Introduction and motivation One specific instance of this localization problem is the tracking of persons by means of an ad-hoc WSN. A system that provides a robust and scalable way to get to know the location of persons and that allows an instantaneous set-up could be used to help coordinating firefighters that enter a burning building or to track patients that suffer from Alzheimer's disease within the compounds of a retirement home or other application examples. In spite of the number of published research papers, there is still no scalable approach that allows localizing nodes that are attached to (moving) persons. The ad-hoc localization of people in indoor/outdoor environments is a difficult task as systems like GPS are unavailable indoors and state of the art systems like UWB real time location systems (RTLS) are still way too expensive for a mass market adoption. Current WSN localization systems often lack real-world practicability as they only consider certain problem instances and are adapted to special circumstances or environments.

State of the art and basics Localization techniques are composed of two basic steps: the measurement step in which physical parameters are determined, and the actual position estimation step, in which these parameters are used to calculate the current location [3]. Existing measurement approaches include: time of arrival (TOA) or time difference of arrival (TDOA) measurements which estimate a distance from propagation time measurements of one signal or two physically different signals, received signal strength (RSS) measurements which estimate the distance from the electromagnetic field strength, angel of arrival (AOA) measurements with which the angle between emitter and receiver is estimated via a directional antenna or an antenna array and fingerprinting or nearest neighbor approaches that are based on proximity measures [3]. The position estimation is then often performed via tri- or multilateration, tri- or multiangulation in combination with a least squares approach or others. One upcoming field is the use of Baysian filters

like Kalman or Particle Filters for the combination of different informations sources and a probabilistic assessment of position estimation step.

Approach Within the subproject H1.2 it is the goal to evaluate the possibility of setting-up an ad-hoc infrastructure by means of pedestrian dead reckoning and sensor fusion and to localize nodes that are attached to persons within this ad-hoc network. A heterogeneous architecture of nodes with variable sensor configurations is developed and integrated in order to allow the fusion of different information sources. The scope of the project is to develop a prototype implementation of this concept and to experimentally evaluate it in order to understand its behavior. In the next step a system theoretic model will then allow to systematically evaluate the chances and principal limitations of this concept. The conducted approach makes use of different available technologies like the ZigBee communication stack and the upcoming MEMS technology in order to explore the possibilities of an integration of technologies in the field of WSN localization. This integration of technologies will then allow to explore a system concept beyond the current state of the art.

Bibliography

[1] P. Bahl and V. Padmanabhan. *RADAR: An in-building RF-based user location and tracking system.* IEEE infocom, volume 2, pages 775-784. Citeseer, 2000.

[2] A. Boukerche, HAB Oliveira, EF Nakamura, and AAF Loureiro *Localization systems for wireless sensor networks.* IEEE Wireless Communications, 14(6):6-12, 2007.

[3] S. Gezici. *A survey on wireless position estimation.* Wireless Personal Communications, 44(3):263-282, 2008.

[4] I.F. Akyildiz, T. Melodia, and K.R. Chowdhury *A survey on wireless multimedia sensor networks.* Computer Networks, 51(4):921-960, 2007.

[5] L. Hu and D. Evans. *Localization for mobile sensor networks.* Proceedings of the 10th annual international conference on Mobile computing and networking, pages 45-57. ACM New York, NY, USA, 2004.

4.10 Information Dissemination in Urban Wireless Vehicular Networks

Tessa Tielert (tessa.tielert@kit.edu)
Supervisor: Prof. Dr. Hannes Hartenstein

Due to the limited amount and scale of existing real-world implementations of wireless communication networks between vehicles, the evaluation of dedicated protocols and applications is currently based mainly on simulations. Thereby simulation results only allow for conclusions to reality if the underlying modeling was appropriate. An appropriate modeling of wireless vehicular communication is especially challenging for urban environments due to highly complex radio-wave propagation characteristics resulting from a rapidly changing environment. For lack of publicly available appropriate channel models, most information-dissemination protocols for urban scenarios have been evaluated using models intended for open-field communication. For similar reasons, other influences on protocol behavior such as load conditions on the wireless channel, the overhead caused by cryptographic operations or realistic node topologies and movement patterns are often neglected or highly abstracted. Thus, it is not clear if the existing evaluation results of information-dissemination protocols are sufficient to assess the protocols' suitability for real-world deployment.

This work aims at assessing and developing information-dissemination protocols for urban wireless vehicular scenarios under more realistic assumptions and simulation models, taking into account and integrating the aforementioned aspects. Thereby the main scientific challenges faced are balancing the accuracy of models against run-time performance, federating different models for the simulation of ubiquitous networks and the development of protocols for ubiquitous networks taking into account both network conditions and application requirements.

As a sample application of wireless vehicular networks in urban scenarios, the communication between traffic lights and vehicles, enabling vehicles to adapt their speeds according to the traffic lights' green phases, has been evaluated. Furthermore, the impact of state-of-the-art cryptographic protection on a sample information-dissemination protocol has been analyzed. To account for a more realistic modeling of the wireless channel, simulation results obtained from ray-tracing techniques are currently being integrated into the network simulator ns-3. Future work will include a study of mechanisms to control the channel congestion caused by periodic broadcast status messages of vehicles which are a prerequisite of information dissemination. All of these aspects are being integrated into a simulation environment using real-world data of the city of Karlsruhe, in which information-dissemination protocols will be analyzed and developed.

4.11 Algorithmic Aspects of Sensor Networks - Scheduling, Localization and Topology Control

Markus Völker (markus.voelker@kit.edu)
Supervisor: Prof. Dr. Dorothea Wagner

The hardware that is used in wireless sensor networks shrinks constantly, while, at the same time, the size of the networks in terms of the number of involved nodes grows rapidly. This brings along many new and fascinating algorithmic challenges. In particular, there is a special demand for distributed and power efficient algorithms. In this project, several algorithmic issues in the context of sensor networks are considered.

Transmission Scheduling. Concurrent wireless transmissions can interfere with each other and thus result in low throughput or transmission failures. One way to avoid retransmissions and to increase the communication performance is the computation of time-division multiple-access (TDMA) schedules. This project deals with the efficient computation of energy-aware schedules with high throughput. In contrast to most existing work, which is based on simplified graph-theoretic models, we turn special focus on realistic modeling of interference.

Boundary Recognition. Many applications in sensor networks require a certain knowledge about the underlying network topology, especially about holes and boundaries. However, in many scenarios there is no information about node positions available and the nodes are not equipped with GPS receivers or similar tools. For this reason, we examine local algorithms for distributed and location-free boundary recognition. Our methods allow a node to decide autonomously with high precision whether it is on the boundary, based solely on connectivity information of a 2-hop neighborhood.

Indoor Localization. Today, outdoor localization is ubiquitous and used in many navigation systems and GPS devices. In this project, we deal with indoor localization, which is still in the early stages of development. Particularly, we examine localization based on signal strength measurements. We consider scenarios where sensor nodes are distributed within a building, without known positions. Main challenges are the self-localization of the network and the localization of people who move through the building.

Topology Control. Throughput and energy consumption of wireless communication are significantly determined by the quality of used communication links and interferences between concurrent transmissions. Goals of topology control are the avoidance of energy-wasteful links and the reduction of interference within the network. This is achieved by restricting the set of transmission links that is used during communication. In this project, we consider the computation of energy-aware topologies. Again, appropriate modeling of interference is central.

5 GRK 1298: Algorithmic Synthesis of Reactive and Discrete-Continuous Systems (AlgoSyn)

Prof. Dr. Dr. h.c. Wolfgang Thomas
Email: thomas@automata.rwth-aachen.de
RWTH Aachen University
Internet: `http://www.algosyn.rwth-aachen.de/`

The DFG Research Training Group (DFG-Graduiertenkolleg) 1298 "Algorithmic synthesis of reactive and discrete-continuous systems" is an interdisciplinary research project at the RWTH Aachen University led by a group of ten professors representing five different disciplines. Integrating approaches from computer and engineering sciences, the project aims at developing methods for the automatised design of soft- and hardware.

While methods of software validation and verification are by now well established, based on adequate formal models and tested in practical applications, the approach of automatic synthesis of software (and hardware) is as yet only developed in quite rudimentary form. On the other hand, in theoretical computer science as well as in engineering disciplines a rapidly increasing stock of techniques for the development of algorithmic synthesis procedures is emerging, triggered by the demand to decrease development costs by invoking algorithmics in formal modelling frameworks. The approach of program synthesis is applicable in several scenarios, in particular in reactive (multi-agent-)systems with low data complexity and in control systems. Central issues in the area are the establishment 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 aim of the Graduate School is to unify the expertise from Computer science, mathematics, and four engineering disciplines (processor architectures, automatic control, process control engineering, train traffic systems) and to push forward the desired integration of methods. The research is carried out in four subject areas: Algorithmics for agent-based probabilistic and hybrid systems, formal methods of reactive systems and game-theoretic methods, software development and modelling languages, and finally applications and demonstrators.

5.1 Synthesis of Behavioral Controllers for Discrete Event Systems with Augmented Petri Net Models

Kai Bollue (k.bollue@irt.rwth-aachen.de)
Supervisor: Prof. Dr.-Ing. Dirk Abel / Prof. Dr. Dr.h.c. Wolfgang Thomas

Although many approaches have been made to automatic synthesis of controllers for discrete event systems (DES), the standard procedure in practice still consists in doing some intuitive design by hand and refine the controller in an iterative process. The method developed in this work aims to assist an engineer designing a controller by providing a better structured process approach. The focus of most of the work on controller synthesis is on the automatic generation of safety or supervisory controllers, i.e. controllers, which block or force control actions to avoid certain system states previously defined as forbidden ([RamWon89], [Giu92], [HanLue98]), while another (behavioral) controller is still needed for the plant to serve its purpose. The developed algorithm focuses on the synthesis of a behavioral controller from a Petri net model of the uncontrolled plant and goal specifications given by linear constraints on the marking, while still avoiding forbidden states.

On the one hand, structures typically found in plant models can be exploited to efficiently find a valid control algorithm (see [BolAbTho09] for details). On the other hand, some requirements arising from the application in control engineering (like the existence of uncontrollable transitions and timing constraints) make the task more complex than the classical reachability analysis in Petri net theory.

Bibliography

[RamWon89] P. J. G. Ramadge and W. M. Wonham, *The control of discrete event systems*, Proceedings of the IEEE, 1989.

[Giu92] A. Giua, *Petri nets as discrete event models for supervisory control* Ph.D. dissertation, Rensselaer Polytechnic Inst., Troy, NY, USA, 1992.

[HanLue98] H.-M. Hanisch, A. Lueder and J. Thieme, *A modular plant modeling technique and related controller synthesis problems*, IEEE International Conference on Systems, Man and Cybernetics, 1998

[BolAbTho09] K. Bollue, K., D. Abel and W. Thomas, *Synthesis of Behavioral Controllers for Discrete Event Systems with NCES-like Petri Net Models*, Proceedings of the European Control Conference, 2009

5.2 Logics for Quantitative Systems

Diana Fischer (fischer@logic.rwth-aachen.de)
Supervisor: Prof. Dr. Erich Grädel / Dr. Christof Löding

My research topic lies in the field of game theory and its application to logic. I am especially interested in the theory of quantitative systems. A structure is called quantitative if it includes not only qualitative predicates and relations, i.e., a predicate holds or not, but, e.g. the predicates can take values from the field of real numbers. Correspondingly, a logic is called quantitative, if formulae do not evaluate to true or false, but again take values from, e.g, the reals. My aim is to investigate which of the well-known theorems for classical logics still hold in the area of quantitative systems.
The logic that we mainly worked with so far is the μ-calculus which plays a prominent role in verification as it subsumes the temporal logics CTL, LTL, CTL* and PDL that are used in practical applications.
In [FGK09], we defined a quantitative variant of the μ-calculus and a quantitative version of parity games, and showed that an important connection to game theory still holds in the quantitative setting: the model checking problem for the μ-calculus can be translated into the problem to solve a quantitative parity game. As in the classical setting the correspondence goes both ways: the value of a formula in a structure coincides with the value of the associated model checking game, and conversely, the values of quantitative parity games (with a fixed number of priorities) are definable in the quantitative μ-calculus.
As a next step, we have looked at model comparison games. Since we work on modal logics, we have looked at the bisimulation relation. Using its game-theoretic definition via a bisimulation game, we defined a quantitative version and showed that, as in the classical setting, the μ-calculus is invariant under this relation.
Next steps in this project are to look at other logics, e.g. GF, FO, MSO, or also fragments of the μ-calculus, and see for which of these logics we can define natural quantitative variants and which of the theorems from the classical setting can be lifted to the quantitative one.

Bibliography

[FGK09] D. Fischer, E. Grädel, Ł. Kaiser. *Model Checking Games for the Quantitative mu-Calculus.* Theory of Computing Systems, 2009.

5.3 Synthesis of Hardware Simulators for Use in Model Checking

Dominique Gückel (gueckel@informatik.rwth-aachen.de)
Supervisor: Prof. Dr.-Ing. Stefan Kowalewski / Prof. Dr. Rainer Leupers

Today many embedded systems are safety-critical. It is often not sufficient to only *test* the software running on these systems, but to additionally apply formal methods such as model checking. The [MC]SQUARE model checker conducts model checking of microcontroller assembly code. [MC]SQUARE does not require any preprocessing of the code, but can directly load the binary program as it is generated by the compiler. The state space of the program is created by simulating its execution on the real hardware in a simulator. Each instruction of the input program corresponds to a transition in the state space. Furthermore, hardware events such as interrupts can add further transitions. The simulator is a special one that can deal with nondeterminism, abstract from time, and create an over-approximation of the behavior of the real hardware. Nondeterminism arises from the values of components that cannot be modeled (e.g. the environment of the microcontroller, which we must assume can do anything) and also from the abstraction from time (e.g. the device's timers could overflow at any moment). To reduce the risk of state explosion due to instantiation of nondeterministic values, [MC]SQUARE features several abstraction techniques, some of which are based on static analyses of the input program. At the moment, [MC]SQUARE supports Atmel ATmega16 and ATmega128, Infineon XC167, Intel 8051, and Renesas R8C/23 microcontrollers. Additionally, it supports Programmable Logic Controllers (PLCs, German: SPS) and abstract state machines.

Our experience from implementing the simulators for the platforms mentioned above shows that implementing a simulator manually takes between six months and one year. In order to reduce this effort, we have created an architecture description language called SGDL and a synthesis system, which can automatically synthesize the simulator from an SGDL description [GSBK10]. The SGDL description is easier to create and considerably shorter than a full simulator. First results indicate that implementing simulators using SGDL can be achieved with less than a month of total effort.

Bibliography

[GSBK10] D. Gückel, B. Schlich, J. Brauer, and S. Kowalewski. *Synthesizing Simulators for Model Checking Microcontroller Binary Code*. In Proceedings of the 13th IEEE International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS 2010), 2010, to appear.

5.4 Reachability Analysis of Hybrid Systems

Paul Hänsch (haensch@embedded.rwth-aachen.de)
Supervisor: Prof. Dr.-Ing. Stefan Kowalewski / Prof. Dr.-Ing. Dirk Abel

Hybrid systems combine discrete evolution with continuous dynamics. These systems occur whenever a physical process is controlled by, or in interaction with some discrete switching system. An example can be a thermostat which is connected to a heater; the system can be a room, the continuous state is the room temperature and the discrete state is either *on* or *off* as for the heater being turned on or off. But there are also way more complex systems, e.g., tanks in a chemical plant connected with each other in some chemical process which may be not trivial to schedule and includes many continous variables, like temperature, pressure, filling height of the tanks and probably more. Discrete states can represent valves being opened or closed, heaters being turned on or off and so on. But there are also completely natural phenomena that behave in a hybrid way. A bouncing ball for example, is moving due to gravity (and air drag) until it hits the ground. Then a discrete event occurs, inverting the direction of motion of the ball (and reduction of speed due to deformation energy).

In our research group we consider (amongst others) hybrid systems with linear time-invariant dynamics of the form $\dot{x} = Ax + u$. Such systems are widely used in control theory. Just to mention an example, that is also being investigated at our chair, think of a car platoon driving on a highway, where the first car is driven by a human and the consecutive cars are controlled automatically such that they follow the first one in a distance as short as possible but avoiding collisions. Indeed, some of the controllers that are evaluated for this purpose can be modeled accurately with systems of the form $\dot{x} = Ax + u$.

A state of a hybrid system is given by a pair (q, x) of a discrete state $q \in Q$, where q is a discrete state and $x \in \mathbb{R}^n$ is the continuous state of the system. We say the system is n-dimensional. Furthermore, each discrete state is provided with an ordinary differential equation (ODE) of the form $\dot{x} = Ax + u$ with constant matrix $A \in \mathbb{R}^{n \times n}$. Here, u denotes an uncertain input and is bounded by, e.g., $\|u\| \leq \mu$ for some $\mu \in \mathbb{R}$. The ODE defines a vector field that specifies the evolution of the continuous system parameters. I.e. being at state x, the system evolves with the rate $Ax + u$. One can think of u as, e.g., a disturbance to the system. For the car platoon such an input can be acceleration of the first car (that is driven by a human).

In this way, such a system with input u is nondeterministic. As a consequence, simulation tools are not able to give any valuable information about the possible trajectories, starting in a given initial state x_0, since they require a deterministic differential equation. But we investigate an even more general

question; we want to compute the states that are reachable from an initial set of states. In general this set is not computable, so one has to make restrictions. We want an *overapproximation* of the reachable set. This is particularly interesting for the verification of hybrid systems, as for instance for the car platoon. Alternatively we want to verify that a set of given *unsafe* states is not reachble. A further restriction is to fix a time horizon T and to consider evaluations with in time $[0, T]$ not worrying about the evolution of the system beyond this time horizon.

The most common approach for this problem is to discretize the time into $N = T/r$ intervals and to overapproximate the states reachable within time $[ir, (i + 1)r]$ from the initial set. The overapproximations are based on two parameters, one that accounts for the error introduced by discretizing the time and one for the input. One question is whether these parameters can still be reduced, making the overapproximations even more accurate.

Whenever one has to represent subsets of the $\mathbb{R}^n$, one has to decide for one or more geometric set representations. This is the same for reachability analysis of hybrid systems. The very first question is how do we specify the initial set? The easiest way is to give it as an interval hull $[\underline{a}_1, \overline{a}_1] \times \ldots \times [\underline{a}_n, \overline{a}_n]$, i.e., as an axis aligned hyper cube. But representing the reachable states as hypercubes is too restrictive and would yield too coarse overapproximations. More flexible, but still too restrictive are hypercubes. Most flexible in this field are polytopes, i.e., sets of points bounded by hyperplanes. Any convex set can be approximated arbitrarily close by a polytope. But they have some disadvantages in terms of computation time. Other data structures used are ellipsoids, parallelotopes, zonotopes. The latter two are subclasses of polytopes. But they offer more efficient implementations for crucial operations, such as Minkowski sum. In our research we want to further address the problems in the use of zonotopes and to improve efficiency and accuracy of the methods that are known today. One thing to improve is the reduction order problem of zonotopes. During the computation, the representation size of the zonotopes grows larger and larger, thus making the computations more and more time consuming. To tackle this problem, different order reduction techniques already exist, but we assume that these can be improved. Another problem is that intersections of zonotopes with other structures are not zonotopes and in general do not allow good zonotope approximatoins. Intersections are especially needed when the discrete state of the hybrid system changes. In three dimensions the intersection of a zonotope and a hyperplane can be a triangle, and any zonotopes overapproximating a triangle has at least twice the are of the triangle and hence is not very accurate. We want to extend the zonotope approach to allow for more accurate approximations. Further research should be conducted in variable step sizes for the time discretization.

5.5 The Complexity of Winning Strategies in Infinite Games

Michael Holtmann (holtmann@automata.rwth-aachen.de)
Supervisor: Prof. Dr. Dr.h.c. Wolfgang Thomas / Dr. Christof Löding

An infinite game is played by two players on a finite directed graph $G = (V, E)$, and $\varphi \subseteq V^\omega$ is the set of all infinite paths through G which are winning for Player 0. All other paths are winning for Player 1. We say that a player has a winning strategy if he can make appropriate moves such that the resulting play is always winning for him, no matter what the opponent does.

Infinite two-player games constitute a powerful tool for synthesis and verification of reactive systems. The essential part of solving an infinite game is to compute winning strategies for the players from their respective winning regions. Whereas for some winning conditions memoryless strategies suffice others require (usually an exponential amount of) memory. The known algorithmic approaches do not guarantee any particular property of the computed strategies, in particular the used memory may become unnecessarily large. The main goal of this thesis is to get a better understanding of the complexity of a winning strategy, both in terms of the time needed to compute it and especially the space needed for implementation [HL07, GH]. This also includes the designing of algorithms to compute memory optimal winning strategies.

Also, we examine the above memory optimization problem from the opposite point of view. This means we do not consider information about the history of a play, but the opportunity for a player to get a lookahead on the moves of the opponent, i.e. to obtain a finite amount of information about the future of the play [HKT10]. This captures situations in distributed systems, e.g. when buffers are present in communication or when signal transmission between components is deferred. The main result of this part of the thesis is that for a regular game one can decide whether there exists a finite lookahead into the future which helps a player to win the game.

Bibliography

[HL07] Michael Holtmann and Christof Löding. *Memory Reduction for Strategies in Infinite Games*. Springer, 2007.

[HKT10] Michael Holtmann, Łukasz Kaiser and Wolfgang Thomas. *Degrees of Lookahead in Regular Infinite Games*. Springer, 2010.

[GH] Marcus Gelderie and Michael Holtmann. *Memory Reduction for Strategies in Muller and Streett Games*. Unpublished.

5.6 Decision Problems over the Domain of the Real Numbers

Ulrich Loup (loup@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Erika Ábrahám / Prof. Dr. Erich Grädel

Decision problems over the domain of the real numbers form a wide field of research. Beside a theoretical analysis of the border between decidable and undecidable problems in this area, a main part of this thesis is concerned with decision procedures for the real algebra, which is the set of formulas composed of quantified BOOLEan combinations of polynomial constraints over the real numbers. Deciding satisfiability of real-algebraic formulas is of major practical relevance since these formulas are capable of describing, for example, linear and non-linear real constraint satisfaction problems, linear and non-linear real optimization problems as well as problems emerging in geometry like determining the intersection of geometric shapes. Although real algebra is decidable [Ta1948], even current decision procedures for real algebra [DoStWe1997] are often overstrained by solving real problems adequately because of their high complexity, which is often manifested in the level of non-linearity or the immense problem size.

We aim for designing an efficient decision procedure for real algebra by taking advantage of the vast progress in propositional satisfiability checking (SAT): we closely combine a modern SAT-solver, which solves the Boolean structure of the input formulas efficiently, with a real-algebraic solver, which checks satisfiability of conjunctions of real-algebraic constraints. This is generally referred to as satisfiability modulo theories (SMT) [GaHaNi2004]. Currently existing SMT-solvers for real algebra do either not solve full algebraic constraints, or work with non-exact interval arithmetic, or do not propagate much information useful for the solving procedure between the two solvers. We want to fill exactly this gap and develop an SMT-solver for full real algebra essentially using real-algebraic solvers, which are deeply integrated into the SAT-solving procedure.

Bibliography

[Ta1948] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, 1948.

[DoStWe1997] A. Dolzmann, T. Sturm, and V. Weispfenning. *Real Quantifier Elimination in Practice*. 1997.

[GaHaNi2004] H. Ganzinger, G. Hagen, R. Nieuwenhuis, A. Oliveras, and C. Tinelli. *DPLL(T): Fast Decision Procedures*. CAV, R. Alur and D. Peled, eds., Springer, 2004, pp. 175-188.

5.7 Time-Inhomogeneous Markov Chains and Piecewise Deterministic Markov Processes

Alexandru Mereacre (mereacre@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Ir. Joost-Pieter Katoen / Prof. Dr. Berthold Vöcking

1)Model Checking HML On Piecewise-Constant Inhomogeneous Markov Chains This work presents a stochastic variant of Hennessy-Milner logic that is interpreted over (state-labeled) inhomogeneous continuous- time Markov chains (ICTMCs), i.e., Markov chains in which transition rates are functions over time t. For piecewise constant rate functions, the model-checking problem is shown to be reducible to finding the ze- ros of an exponential polynomial. Using Sturm sequences and Newton's method, we obtain an approximative model-checking algorithm which is linear in the size of the ICTMC, logarithmic in the number of bits precision, and exponential in the nesting depth of the formula.

2) LTL Model Checking of Time-Inhomogeneous Markov Chains We investigate the problem of verifying linear-time properties against inhomogeneous continuous-timeMarkov chains (ICTMCs).A fundamental question we address is how to compute reachability probabilities. We consider two variants: time-bounded and unbounded reachability. It turns out that both can be characterized as the least solution of a system of integral equations. We show that for the time-bounded case, the obtained integral equations can be transformed into a system of ordinary differential equations; for the time-unbounded case, we identify two sufficient conditions, namely the eventually periodic assumption and the eventually uniform assumption, under which the problem can be reduced to solving a time-bounded reachability problem for the ICTMCs and a reachability problem for a DTMC. These results provide the basis for a model checking algorithm for LTL. Under the eventually stable assumption, we show how to compute the probability of a set of ICTMC paths which satisfy a given LTL formula. By an automata-based approach, we reduce this problem to the previous established results for reachability problems.

5.8 Aspects of Higher-Order Pushdown Systems - Infinite Regular Games and Strategies

Michaela Slaats (slaats@automata.rwth-aachen.de)
Supervisor: Prof. Dr. Dr.h.c. Wolfgang Thomas / Prof. Dr. Jürgen Giesl

The classical set-up of automatic synthesis of reactive programs is best described by the model of infinite two-person game. The two players are on the one hand the (possible hostile) environment and on the other the controller. The game arena is a finite transition graph where each state belongs to one of the two players. The game is played by moving a token through the arena along the edges, where the player owning the current vertex chooses an edge to a next vertex. An infinite play is won by the "controller" if it satisfies a "regular winning condition". It is known that these games for regular winning conditions can be "solved": One can compute whether the "controller" has a winning strategy starting from a given initial vertex, and one can compute such a winning strategy in the format of a finite input-output automaton. This solvability result (going back to Büchi and Landweber in 1969) is the starting point of the algorithmic theory of infinite games.

We extend this algorithmic result to infinite game arenas. Usually they arise by attaching some infinite store to a given finite arena. It is known (by work of Walukiewicz) how to do this if the store is a pushdown stack. In our work we investigate the structure of "nested stacks", namely higher-order pushdown stacks. A level-1 stack is a standard stack, a level-2 stack is a stack of stacks, and so on. This kind of storage is needed, for example in the implementation of higher-order recursive programs. Even in this quite general framework, it is possible to provide an automatic synthesis of winning strategies and hence of suitable controllers. The main problem in solving games over arenas that involve higher-order pushdown stacks (or counters) is to develop an appropriate format of controller (again with the same storage structure) that has enough power to implement possible winning strategies. We have developed a synthesis method for such controllers based on generalized concepts of "regular stack language" [CS08, HST09a], and we analyze the applicability in several domains, including scheduling problems.

Bibliography

[CS08] A. Carayol and M. Slaats. *Positional Strategies for Higher-Order Pushdown Parity Games*. MFCS, LNCS 5162, 2008.

[HST09a] P. Hänsch, M. Slaats and W. Thomas. *Parametrized Regular Infinite Games and hopd-Strategies*. FCT, LNCS 5699, 2009.

5.9 The Network Design Problem for Railway Infrastructure

Jacob Spönemann (spoenemann@via.rwth-aachen.de)
Supervisor: Prof. Dr. Ekkehard Wendler / Prof. Dr. Berthold Vöcking

The thesis shows how to generate a cost-optimal railway infrastructure by stating and solving the network design problem as a linear optimization problem.

Railway infrastructure is represented by a network consisting of nodes and arcs. The nodes represent stations, the arcs lines connecting the stations. An input instance of the network design problem for railway infrastructure contains two parts. A network without any arcs, i.e. just the stations which has to be connected in a certain way, and a traffic demand which relates each pair of nodes (A,B) to an amount of trains of different types - distinguished by some parameters like length or highspeed of the trains - which has to be routed from A to B in a given time horizon.

A solution of the problem contains on the one hand a new designed network and on the other hand a routing which determines which train has to be routed via which arc to fulfill the traffic demand and capacity constraints.

Such new designed network answers two questions: What is the topology of the network, i.e. which stations are connected to each other and how does each connection respectively line look like (e.g. single track, double track, single track with one overtaking station etc.)?

The observed kind of routing problem can be stated and solved as *multi-commodity flow problem* where every type of train corresponds to a certain *commodity* and the traffic demand mentioned above is modeled by triples of demand, source node and sink node. In general, to find a routing the network has to be given in advance. For the network design problem the network is unknown. But since the nodes are known a *complete* network can be found. This complete network contains one arc for each possible stage of extension of the corresponding line. Finding a routing for a given traffic demand in such complete network is then equal to design the network since the routing *chooses* the arcs needed and so designs the desired network.

In a mathematical formalization the problem can easily written down as an optimization problem. It turns out that the capacity constraint is non-linear and so LP-solvers are not applicable. To overcome this difficulty a transformation is done that allows to reformulate the problem as a mixed integer program which can be solved by standard LP-solvers. Because the mentioned transformation yields to large amounts of variables *column generation* is used to get shorter runtimes.

5.10 Symbolic and Timed Testing

Sabrina von Styp (von-styp@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Ir. Joost-Pieter Katoen / Prof. Dr.-Ing. Ulrich Epple

Formal methods have been applied successfully in the field of conformance testing over the last decade. A well-known representative is the *ioco framework*, where the correct behaviour of an implementation under test is specified using a formal model (a labelled transition system (*LTS*) with input and output actions), and which allows for the automatic derivation of test-suites from this specification. Test-cases cast their verdicts following a formal correctness criterion: the implementation relation *ioco*.

The *ioco* relation expresses that an implementation may only produce outputs if those outputs are also produced by the specification.

The *ioco* theory has been extended in several directions. Recently, an approach to deal with data input and output to and from the implementation, respectively, has been developed. Data is treated symbolically, which avoids the usual problems of state-space explosion and infinite branching that occur when data is represented explicitly in an ordinary *LTS*. Specification and implementation are modelled here as Symbolic Transition Systems (*STS*), which are *LTS* extended with a notion of data and data-dependent control flow based on first order logic. In the context of this theory, the implementation relation *sioco* has been developed, which is defined solely within the FO-Logic framework on *STS* level.

Also several approaches to timed *ioco*-based testing have been developed, mostly on the basis of timed automata. Different notions of conformance have been defined on the basis of timed *LTS*(*TLTS*).

We take first steps towards a testing theory which combines time and data. In this context we have define a conformance relation *stioco* where time and data aspects are both treated symbolically. We introduce *Symbolic Timed Automata* (*STA*) as a formalism to express this combination. *STA* are a straightforward combination of *STS* and *TA*, but which allow data inputs to influence the real-time behaviour. The semantics of *STA* is given in terms of *TLTS*. *stioco* is an extension of *sioco* for *STA* and can be shown to coincide with *tioco* on the semantical, *TLTS* level. Future research will address the definition of test-cases and their automatic derivation.

5.11 Synthesis and Probabilistic Analysis of Online Algorithms for Online Buffering Problems

Melanie Winkler (winkler@cs.rwth-aachen.de)
Supervisor: Prof. Dr. Berthold Vöcking / Prof. Dr. Ir. Joost-Pieter Katoen

In Online Buffering problems, we are dealing with the management of a storage for a commodity with varying prices and demands. We assume that time proceeds in discrete time steps $t \in \{1 \dots t\}$. In every time step an decision maker is given a price c^t (per unit) and a demand d^t of the commodity. The decision maker then has to decide how many units are purchased in time step t. Units which are not used in the step can be saved in the storage of bounded size B for later usage. We consider this problem in an online setting, i.e., the decision how many units the decision maker wants to purchase in time step t can only be made based on c^k, d^k, s^k (filling status of the storage), $k \in \{1, \dots, t\}$, but not on information about future time steps.

There are several applications for those algorithms, for example a canny car driver who has to decide when to refill the tank of his car at the gasoline station. Another application is the management of the buffer in a hybrid car. In a hybrid car there are two engines, one combustion engine and the other operated with electrical energy. The decision maker has to decide when to refill the buffer with electrical energy.

In [1] we have studied this problem for a price sequence modeled by a random walk. We have shown that the algorithm achieving the lowest expected cost only fills the buffer if the price is at the lower boundary. Otherwise it uses units from the storage if possible. Unfortunately, a price sequence modeled by a random walk is not realistic in most applications.

To achieve results which are more generally applicable, we now study online buffering with online learning strategies. In this scenario there are given N experts, every expert gives a strategy to manage the buffer. Depending on the price sequence an expert is chosen in an online setting. The objective is to find a sequence of experts which performs nearly as good as the best expert in hindsight. In general there are no assumptions on the quality of the experts. It is therefore not possible to chive online algorithms which are nearly as good as the optimal offline strategy by combining the experts. Online learning algorithms have widely been studied, for an overview see for example [2].

When using online learning algorithms for online buffering problems an expert would define how many units are bought at a time step. But this might lead to complications when combining expert advice, since the experts in general assume different filling status. So we have to find a way to adapt

the filling status of the storage (s.t. the online strategy is able to fulfill every request) without increasing the cost of the online algorithm to much.

Another open question to achieve more applicable results is how to model the strategy of an expert. Until now we have assumed that the experts are given, but to achieve results which are good in general, we have to develop suitable experts.

For online buffering problem we want to get experts, which decide how many units to buy depending on the current price, demand and filling status of the storage. These decisions can be achieved from assumptions the expert make about the input sequence. We assume that no expert knows the input sequence in advance, but there might be several possible assumptions about the input sequence which can be modeled in different experts. The model has to be defined in a way that the strategy of an expert can be calculated from the model. We want to study how these assumptions can be modeled inside an expert.

One approach to achieve this is to describe the price model by a Markov Chain. If an expert assumes, for example, that a price c_1 is often followed by price c_2 or c_3, then the chain would have a high transition probability from c_1 to c_2 and c_3. Similar we can construct a model for the demand or we can combine both inside one Markov Chain. The optimal number of units bought in a time step can then be calculated from these models, when considering the current filling status of the storage and the cost associated with buying a certain number of units.

Bibliography

[1] Matthias Englert, Berthold Vöcking and Melanie Winkler. *Economical Caching with Stochastic Prices*. Proceedings of the 5th International Symposium, SAGA, 2009.

[2] Dean P. Foster and Rakesh Vohra. *Regret in the On-Line Decision Problem*. Games and Economic Behavior, 1999.

6 GRK 1324: Model-Based Development of Self-Organizing Information-Systems for Disaster Management (METRIK)

Prof. Dr. Joachim Fischer
Email: fischer@informatik.hu-berlin.de
Humboldt-Universität zu Berlin
Internet: `http://www.gk-metrik.de`

Each year, natural disasters cause tremendous devastation and many victims worldwide. The graduate program METRIK attempts to improve the quality of the complex processes involved in disaster management. This is being done in terms of specific issues by developing IT-based methods and technologies in close cooperation with geo-scientists. METRIK is an interdisciplinary cooperation of the Humboldt-Universität zu Berlin, GFZ Potsdam, Fraunhofer FIRST and ZIB Berlin. METRIK focuses on the provision and use of geo-specific services through dynamic, highly flexible and self-organizing information systems. Specific questions concern network routing, replication of distributed databases, ubiquitous monitoring in service networks, automated software component test generation end execution as well as dynamic load balancing on restricted devices. In addition to functional aspects, non-functional aspects such as latency, reliability and availability are taken into account. Within the framework of the European project SAFER, METRIK technology has been applied to the prototype development of a new earthquake early warning system. This warning system has been successfully installed in Istanbul (autumn 2008) and also in italy during aftershocks of the L'Aquila event (April 2009) to monitor damaged buildings. Such benefit had been impossible without the high degree of interdisciplinarity in research and development between computer science at Humboldt-Universität zu Berlin and seismology at GeoForschungsinstitut Potsdam. The combination of seismic signal analysis methods with novel net infrastructures, communication and alarming protocols and specific hardware equipment forms the heart of that early warning system. The main METRIK characteristics are model-driven approaches to the development of described network structures as well as models in combination with powerful analytical methods for controlling complex workflows in disaster management systems on suitable IT structures. Through the use of these methods, it is expected that actors get new room to maneuver in the event of a disaster. Thus, actions can be launched much faster and safer.

6.1 Formal Verification of Self-Organizing Systems

Jan Calta (calta@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Holger Schlingloff, Prof. Dr. Miroslaw Malek

Self-organization is often used as one possible approach to assure adaptivity of a system to changing conditions or environment. A self-organizing system consists of uniform members with limited capabilities. However, if the members establish some kind of interaction (via physical connection, communication, etc.) and a specific structure is established, an advanced functionality emerges in the system. Moreover, this added quality is achieved automatically and locally, without any intervention from outside the system.

An example of a system where self-organization is beneficial is a sensor network for detecting natural disasters. A self-organizing distributed software system can be seen as a homogenous multi-agent system and the members of this system as uniform autonomous agents. An agent in such a system collaborates with his neighbors, without a global view of the whole system and without any central control.

Properties of a software system can be formally investigated by the model checking technique. Model checking uses a rigorous description of the system and the analyzed property. The system is represented by a mathematical model. Then, the question on the model is mapped to a property of the used formalism (e.g., by using a logical expression). Finally, it is proved whether the property holds in the model or not: by exhaustive evaluation in the state space of the system. Thus, a formal verification of the system according to the logical and mathematical rules is performed.

The aim of this thesis is to provide a framework for formal analysis of self-organizing systems.

One part of this framework is a formal for modeling of a real system that exploits the specifics of self-organizing systems, namely their homogeneity, local interaction between the members of the system and the limited size of member's neighborhood. This allows to produce a representation of a real system that is more compact than a conventional model and allows for more informative analysis that can better contribute to the development of a system than traditional model checking technique.

The other part of the framework is a logical language for specification of system properties to analyze. This logic should allow to conveniently and precisely describe the intended adaptive behavior of a self-organizing system in term of relations among system members members and interaction with the environment.

6.2 Self-Aware Adaptive Service Networks with Dependability Guarantees

Andreas Dittrich (dittrich@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Miroslaw Malek

Disasters striking in inhabited areas pose a significant risk to the development and growth of modern societies. The impact of any disaster would be severe. In case a disaster strikes, fast and safe mitigation of damages is important. Information and communication technology (ICT) plays a crucial role in helping reconnaissance and first response teams on disaster sites. Most rescue teams bring their own network equipment to use several IT services. Many of these services (e.g., infrastructure, location, communication) could be shared among teams but most of the time they are not. Coordination of teams is partly done by pen and paper-based methods. A single network for all participating teams with the possibility to *reliably publish, discover and use* services would be of great benefit. Despite the participating teams and course of action being different on every site, described service networks display certain common properties: They arise spontaneously, the number of nodes and their capabilities are subject to high fluctuation, the number and types of services are also fluctuating strongly and there is no global administrative configuration. Because of these properties all network layers involved would need to be configured automatically. Based on the *Internet Protocol* (IP) – the only well-established global networking standard – a number of mechanisms promise to automatically configure service networks. In disaster management scenarios, where various services are critical for operation, mission control could benefit from these mechanisms by getting a live view of all active services and their states. It needs to be investigated if and how they are applicable. Given an ad-hoc, auto-configuring service network, how and to what extent can we guarantee dependability properties such as *availability*, the ability to perform in the presence of faults (*performability*) and ultimately the ability to sustain certain levels of availability or performability (*survivability*) for critical services at run-time? The goal of this dissertation is to provide a comprehensive dependability evaluation for such heterogenous and dynamic service networks. A run-time dependability cycle is being embedded into the network. In this cycle, the network is constantly *monitored*. A distributed service discovery layer provides network-wide service presence monitoring. This will be extended to provide monitoring for availability and performability assessment. Based on monitoring data, dependability properties are *evaluated at run-time*. The survivability of critical services can be estimated by calculating the expected availability or performability with a given fault model. If necessary, *adaptation* measures are triggered which in turn can cause the monitoring to be reconfigured. Even if no adaptation is possible, run-time awareness of critical states is already a huge benefit. This cycle is the base of a *self-aware adaptive service network*.

6.3 Self-Organized Data Replication in Wireless Multihop Ad-Hoc Networks

Joanna Geibig (geibig@informatik.hu-berlin.det)
Supervisor: Prof. Dr. Alexander Reinefeld

My dissertation is motivated by the use of wireless networks (WN) as a support in the Disaster Management (DM). Typical scenarios are collecting environmental data for deducing scientific disaster models (e.g., earthquake models) and communication in the response actions after the disaster. However, wireless networks in comparison to the wired environments are less efficient and unreliable. Wireless links can change their quality and may disappear because of the changes in the traffic or environment. Unattended and exposed to the hostile surroundings nodes can break-down. Additionally, in the disaster scenario, the whole group of nodes may fail (e.g., in a fire).

The objective is to design a replication protocol for achieving a satisfactory availability of data. Replication must be also efficient and self-organizing.

Availability of data provides that data is reachable at the access time. This means that data must survive the disaster and be accessible to the querying node. Data is accessible to the querying node (sink) only if there is a multihop path between this node and a node holding data. Availability of data can be improved by increasing the number of replicas and choosing a good placement strategy. When properly replicated, data will be available to requesting processes even after destruction of some (possibly big) group of nodes and eventual network partitioning. It is desired, that replicas belong to different network partitions, if such emerge.

Efficiency means economizing limited system resources, like bandwidth, energy and storage capacity. To save bandwidth and energy, the amount of send data and number of data retransmissions (number of hops which data traverses) shall be minimized. In order to save storage capacity, replication system should create fewer replicas.

Self-organization supplies a maintenance free adjustment to the changes in the network topology (e.g., after a damage) and load (amount of generated data and number of queries). For example, the protocol shall react to the information about network partitioning by changing its replication policy.

The most important concern when designing the desired protocols for the wireless environment are the *limited resources*. Sending of data is the most resource consuming task. To reduce it, protocols must rely on the *limited knowledge* about the system.

To allow for the efficient replica placement and search the replicas should be load balanced in the network and their locations must be known. A geometric relation for determining replica placement will be used. Relation R assigns to each geometric location a set of points, uniformly spread on the

plane. A data item with the given location will be replicated near the points calculated by the relation R which also belong to the area covered by the network.

To give the nodes required awareness of the network area and their relative position in it, and to inform the nodes about the local resource usage, gossiped based protocols will be used. The gossiping messages will have a constant size.

Moreover, the protocols will exploit natural properties of the underlying wireless network, like broadcast and multihop communication properties. Critical network points (like articulation points and bridges) frequent in the realistic topologies will be economized.

Inshort, the distributed *replication protocol* and associated *retrieval mechanism* will be based on a combination of a structured correlation between data's origin location and the placement of its replicas and an unstructured, gossiped based information retrieval for

- an efficient allocation of replicas to concrete nodes and
- determining the location of actually created replicas (with a high probability).

Resulting self organized replication and retrieval protocols for unreliable, wireless, multihop, ad hoc networks will automatically replicate the data specified by the application in the way best suited to the actual network topology and conditions. The number of replicas will depend on the actual network size and shape. The data retrieval costs will be limited for the nodes in the vicinity of the original data location (High Availability Zone - a system parameter). For other nodes retrieval costs will grow approximately linearly with the distance. In case of network partitioning, each part of the network continues its life autonomously, and allows both for data replication as retrieval. The protocols try to economize critical network elements, like articulation points, in order to further increase efficiency (avoiding hot-spots) and to extend network life time for the battery powered scenarios. The solution will be independent from used MAC and routing protocols. Another important outcome of my dissertation is new knowledge about efficiency, robustness and scalability of gossiping protocols over wireless networks with different characteristics (uniform, realistic, different densities and degree sequences). Examples of protocols designed during this research are efficient and robust gossiping protocols suited for unreliable, wireless multihop ad hoc networks (Wireless Multihop Networks, Wireless Ad-Hoc Networks and Wireless Sensor Networks), like gossiping for calculating local aggregates and a distributed, scalable topology information dissemination protocol.

6.4 Testing Dynamic Component Systems

Siamak Haschemi (haschemi@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Joachim Fischer

Modern software systems are built-up by a collection of reusable components and can evolve (due to new requirements and identified bugs) by replacing one or several components. Some of these systems, called *Dynamic Component Systems* (DCSs), cannot be stopped in order to replace components. Instead, components have to be replaced while the system is running. Thus, components need to be prepared for *dynamic availability* of required functionality provided by other components. The ability of a components to be tolerant to dynamic availability of required functionality is the central problem I want to focus on in my thesis. The goals of my thesis are:

- to develop of a systematic test approach for testing the ability of a component to handle dynamic availability of required functionality; testing should be possible at development time
- integrating the aspects of the approach in a metamodel-based description with the long-term goal of platform and tool independence

My approach is to apply the technique *model-based testing* (MBT) to DCSs. MBT defines a process to generate automatically test cases from formal models, also called *test models*. These models contain the expected behavior of the system under test. The typically infinite number of possible test cases is reduced by using *test selection criteria*. In the phase of *test execution*, the generated test cases are applied to the SUT and expected and actual behavior is compared to find bugs. Applying MBT to DCSs raises following scientific questions, which need to be answered in my thesis:

- **test models:** How can we consider the dynamic availability of components in the creation of test models?
- **test selection criteria:** Which test selection criteria are suitable to find bugs in DCSs related to dynamic availability of required functionality?
- **test case generation:** Which test generation tool supports (or can be extended to support) DCSs, the considered type of test models, and the required test selection criteria?
- **test execution:** What requirements has the test execution on the component runtime environment?

6.5 A Model-based Approach for Transparent Experimentation Workflows

Frank Kühnlenz (kuehnlenz@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Joachim Fischer and Prof. Dr. Jens-Peter Redlich

Experimentation in my work means performing experiments on system models, executed in special computing environments (e.g. a simulation framework or a sensor net) to get model observations. The experimentation process comprises the ordered phases planning, execution and evaluation of experiments, forming a subset of scientific workflows that I call *experimentation workflows*. Often these experimentation workflows are incompletely and not formally described making them intransparent. This means, the experiment results are hardly comprehensible because their provenance is not clear: The experiment can not be repeated due to missing knowledge.

Scientific workflow management systems support experimentation workflows, but they are not designed for them. In addition, using a general and complex scientific workflow management system needs expert knowledge about its handling. Hence, experimenters often do not use any tool support for experimentation. Re-using and exchanging experimentation workflow models across different projects, communities and domains is a common need for experimenters, but problematic due to different management systems.

It is promising to apply the paradigm of MDA also to the domain of experimentation: Evaluating this approach is my research thesis. In particular, my method is identifying and formalizing general experimentation concepts independent from a concrete experimentation domain (e.g. located in geography or in physics) and not limited to a special experimentation workflow management system. Expectingly this results in making experimentation more transparent by ensuring reproducibility in a comprehensible experimentation workflow.

I formalize the general experimentation concepts as a *Domain-specific Language* (DSL) in a meta-model, named *Experiment Management DSL* (EM-DSL). The EM-DSL as general language for formalizing experimentation workflows provides tailored means of expressions in the experimenter's vocabulary. The experimenter models the workflow and typical characteristics of his experiments using the EM-DSL. For example, this comprises defining the system model-specific data structures and the description of how to actually execute this system model in different execution environments. Based on the EM-DSL and all related language tools (e.g. editor, persistence layer using a RDBS), a prototype of an *Experiment Management System* (EMS) will be implemented, realizing the operational semantics of the EM-DSL.

6.6 Wireless Routing Protocols for Earthquake Early Warning Systems

Jens Nachtigall (nachtiga@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Jens-Peter Redlich

The total cost of an Earthquake Early Warning System (EEWS) can be substantially decreased by using Wireless Mesh Networks (WMNs). The Early Warning System triggers on the small-amplitude, but fast P-wave in order to shutdown critical infrastructures before the destructive, but slow S-waves arrive only a few seconds later. To trap false positives the nodes need to interact and repetitively communicate its status which each other. Due to the short warning time, an EEWS demands low-latency communications of high robustness from a WMN and all actions taken by a costumer of an alarm must automated.

The use case of Earthquake Early Warning is unique and requires a special approach [Nachtigall2009]. Even though P-waves are non-destructive, they can cause slight shaking of a few centimeters for strong ($M_w > 6$) and nearby (epicentral distance < 40 km) earthquakes at the very moment, when the Early Warning System would be activated. The sudden small-amplitude P-wave shaking can have an immense impact on the performance of wireless communications.

The swift link quality fluctuations due to multi-path effects and shadowing are very difficult to handle with today's pro- or reactive unicast routing approaches which infer the current quality of links from past measurements. If an earthquake occurs, a wireless link considered of good quality might suddenly become bad as the P-wave arrives, rendering a previously selected routing path unusable. A totally different communications approach like anypath or opportunistic routing is be worth considering for the future. Here, radio links of erratic qualities, which would normally be avoided by current routing protocols, can be used on a trial and error basis. While the concept of opportunistic routing might be beneficial for WMNs in general due to its broadcast medium with lossy links, it should so even more for the application of an EEWS, where link qualities are extremely fluctuating.

Bibliography

[Nachtigall2009] Nachtigall, J., A. Zubow, R. Sombrutzki, and M. Picozzi. *The Challenges of using Wireless Mesh Networks for Earthquake Early Warning Systems.* MESH 2009.

6.7 Analysis of Behavioral Exchangeability of Services

Jarungjit Parnjai (parnjai@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Wolfgang Reisig

In general, a distributed system in disaster management involves an interaction among multiple work units such as a fire brigade, a red cross agent and a landscape engineer department. These work units interact locally and asynchronously without centralized control, according to a global process definition that has been developed before disastrous events occur. Such collaboration of a distributed system is inevitably subject to changes over time due to new requirements or unpredictable circumstances.

Service-oriented computing provides the support for system evolution; for example, by considering a system as a collaboration of interacting services and allowing an exchange of one service by the other one. Nevertheless, it is crucial that an exchange of one service must neither disrupt an interaction with others nor produce an adverse effect to the entire system.

As one single organization may not own the entire collaboration of the services in a distributed system, the verification of the entire system is often not a feasible alternative. This raises the need for an analysis method that guarantees the correct-by-design exchange of a service. The analysis method should ensure that a new version of a service interact properly with other services in the system; thus, maintaining the proper interaction with all service partners of the original service.

The goal of this dissertation is to develop a formal analysis method that provides supports for a service designer, e.g. a domain expert, to systematically analyze the behavioral exchangeability of services at the design time. This method shall allow a service designer to do the followings : to decide if a new version of a service is correct-by-design, to fix an incorrect-by-design version of a service with a minimal cost, as well as to synthesize a a service that is guaranteed to be correct-by-design.

Bibliography

[1] Parnjai, J., Stahl, C., Wolf, K.: *A finite representation of all substitutable services and its applications.* In Proceedings of the 1st Central-European Workshop on Services and their Composition, ZEUS 2009, Stuttgart, Germany, March 2009.

6.8 Security in Self-Organizing Systems

Henryk Plötz (ploetz@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Jens-Peter Redlich

Self-organizing wireless mesh networks in the scope of METRIK's Earthquake Early Warning System (EEWS) will need to be augmented with security functions to become truly useful and dependable. An EEWS that can be subverted by any passer-by who can then inject false alarms will be of no use. An EEWS that can be subverted by a slightly more sophisticated attacker who can pick up and and take control over nodes will only be slightly more useful and might become even more problematic with time as (ill-founded) confidence in the system rises.

A cryptographically secured system has two main aspects:

- The cryptographic primitives that are being used, and
- the key and trust management that is being used to initialize these primitives and ensure the correct operation of the protocol.

While there are a number of standard cryptographic primitives available which can be applied to a broad set of circumstances, the key and trust management still needs to be specifically designed for each application.

Popular key and trust management schemes generally are not self-organizing, hence of limited or no use in the METRIK context. In contrast, a naïve approach to self-organized key management would offer no trust guarantees and therefore no security. Any attempt at a solution should also be able to cope with the challenges of self-configuration and network partition.

Future work is to evaluate the applicability of several key distribution and distributed key management schemes in the METRIK setting, focussing on the impact of the proposed routing and data distribution schemes.

6.9 Analysis and Construction of Choremetic Diagrams

Andreas Reimer (areimer@gfz-potsdam.de)
Supervisor: Prof. Dr. Doris Dransch

Choremetic maps, as introduced by French cartographer R. Brunet in the 1980ies, are a class of diagrams that allow to depict complex situations in a highly synthesized and generalized manner. In essence they are a tool for the structural and iconic representation of complex geospatial situations. They consist of terms and graphics that largely prescind from actual objects and precise cartographic symbols and have enjoyed several documented successes as effective communication tool. Until now, choremetic diagrams have been created manually.

In the case of automatizing the construction of choremetic diagrams, two major research questions need to be answered: How can choremetic diagrams be described formally? Which process steps are needed to derive choremetic diagrams from geodata?

The research consists of two parts: In the first part existing choremetic diagrams found in printed sources such as books, magazines and agency documents are analysed in terms of their attributes. The analysis' results are used to create a formal description of choremetic diagrams. In the second part of the research, a process chain for the construction of choremetic diagrams from geodata is developed on the basis of the formal description derived in part one.

The results are combined into a formal description of choremetic diagrams. This formal description is enriched by concrete measures for specific attributes in form of generalization constraints and a choreme database in which existing mappings of thematic content to choreme usage are stored and classified after scale level structure. The results of the analysis will be used to develop a prototypical process chain for the automated construction of choremetic diagrams from geodata. This is approached as a cartographic generalization problem. Based on the knowledge gained during the analysis, apropriate existing methods are chosen from the generalization domain and modified and parametrized accordingly. One goal is to minimize human intervention in this process chain. The first validation of the process chain will be the recreation of existing diagrams.

6.10 Model-driven Engineering for Environmental Modeling with Cellular Automata

Falko Theisselmann (theissel@informatik.hu-berlin.de)
Supervisor: Prof. Dr. Doris Dransch, Prof. Dr. Joachim Fischer

Disaster management relies on information about the causes, the development, and the mitigation of disasters. Environmental simulation models often generate, store and communicate this information. Therefore simulation model descriptions should be *transparent* and *reusable*.

However, model specifications are usually dependent on single implementation technologies (e.g. modeling tools and programming languages) with substantial limitations with respect to transparency and reusability.

With domain specific modeling languages (DSLs) the issue of transparency has been addressed in many cases already, but the usual realizations of DSLs lack reusability, since these DSLs are bound to monolithic tools. This issue is particularly relevant in multi-disciplinary modeling, where different parts of a complex model might be specified using different coupled DSLs, where existing models might be reused.

In software engineering, the model-driven approach to modeling with DSLs with object-oriented meta-models promises to alleviate such reusability and transparency issues. In this work, this model-driven approach is applied to environmental modeling in order to evaluate it's feasibility in the field of environmental modeling. For this, a DSL for environmental modeling with Cellular Automata (ECA) is defined and realized prototypically.

>From a language development perspective, ECA is a representative example for a number modeling formalisms in environmental science. Cellular Automata are widely used to model disaster management related processes, such as fire spread, seismicity, or the evacuation of buildings. However, in addition to reuse and transparency issues, there is no adequate formalization of this formalism as it is used for environmental modeling.

This work provides a formalization of Environmental Cellular Automata (ECA) by means of the ECA-DSL.

The coupling of DSLs is based on the fact that the coupled languages share common concepts. Many modeling formalisms used for environmental modeling implement common concepts from the domains of simulation modeling and geospatial data processing. In this work, a set of basic language concepts has been identified that can be used for coupling DSLs for environmental modeling and respective models.

In order to evaluate feasibility of the ECA-DSL and the concepts for language coupling, coupled languages (ECA, agent-based modeling, experiment description) have been realized prototypically and applied to a number of published models.

6.11 Verifying Data-Dependent Services

Christoph Wagner (cwagner{@}informatik.hu-berlin.de)
Supervisor: Prof. Dr. Wolfgang Reisig

Distributed systems can be designed according to the paradigm of Service Oriented Computing (SOC). A distributed system consists of several services that communicate via their respective interfaces. The interaction of the services may be correct or may be not correct. Whether the interaction of the services is correct, depends on the behaviour of each single service. The baviour of a service in turn depends in mainly on the data that is contained in the messages which are interchanged by the services. Goal of this work is to model and verify whether the services can interact correctly.

By considering data in the model explicitly, the state space of a service can become very large or even infinitely large. Therefore, the state space of the service is not manageable any more with computational methods alone. Nevertheless, there are regularities in the structure of the state space which can be exploited. Generally, there are several data values which imply the same behaviour of the service.

The goal of this work is to develop data structures and algorithms that allow to efficiently decide whether services can interact correctly with each other. The approach of this work is based on the concept of a so called "Operating Guideline" [LMW2007]. The Operating Guideline of a service contains information for a user of a service on how to interact correctly with that service. Such an Operating Guideline shall be developed specifically for a class of services that process data.

Bibliography

[LMW2007] Niels Lohmann, Peter Massuthe, and Karsten Wolf. *Operating Guidelines for Finite-State Services*. Springer-Verlag, 2007.

6.12 Metamodel-based Technologies for the Development of Optical Nanostructures

Arif Wider (wider@informatik.hu-berlin.de)
Supervisor: Prof. Dr. sc. Joachim Fischer

Introduction

Optical nanostructures are structures that are smaller than the wavelength of light. Because of that they can affect the propagation of electromagnetic waves. Of particular interest are *photonic crystals*, which are designed to affect the motion of photons in a similar way that semiconductor crystals affect the motion of electrons. The long-term goal of the research in this area is the ability to produce photonic components whose features are similar to those of electronic components while having significant advantages, e.g., low heat generation at high frequencies.

Model-driven development (MDD) is a methodology in software engineering that focuses on the creation and processing of *models*. Models are formalized descriptions aiming at a high level of abstraction. Often, models are expressed in *domain-specific (modeling) languages* (DSLs). A DSL is a language that is specially tailored to the needs of a certain problem domain. The goal of a DSL is to be particularly expressive in its domain, i.e., common concepts of the domain can be expressed concisely. These concepts can be described by a *metamodel* of that domain. Such a metamodel-based approach allows for leveraging existing metamodel-based technologies in order to automatically generate editing, transformation or debugging tools.

The overall goal of my work is to show that metamodel-based technologies can be applied to nanostructure development. Therefore, the development of a *Computing Infrastructure for Model-driven development of Optical Nanostructures* (CIMON) serves as a comprehensive example scenario.

Problem statement

In the tooling currently used by physicists to describe experiments in nanostructure development, the specification of an experiment is divided into different aspects. A model-driven infrastructure for nanostructure development has to reflect this by integrating different DSLs, e.g., one for describing the geometrical structure and one for specifying parameters for the simulation. Using different DSLs to describe a system (or in this case an experiment) is called domain-specific *multimodeling*. If the models described in these DSLs overlap, e.g., by having references to each other, consistency between these models has to be ensured. This is called *model synchronization*. Naive implemented model synchronizations that go beyond simple one-to-one mappings

are hard to maintain and to reason about. *Bidirectional transformations languages* aim for a better maintainability and expressiveness of model synchronizations. But existing solutions for metamodel-based applications are either complex or restricted to bijective transformations which is not feasible for most multimodeling scenarios.

Approach

Defining symmetric bidirectional transformations for a relation that is neither surjective nor injective is very hard. This gets more manageable, if the relation is at least surjective, i.e., if one of the models to synchronize is an abstraction of the other. There is a formalism to define such asymmetric bidirectional transformations that is called *lenses*.

A lens comprises a forward and backward transformation. Because information only gets lost in the forward (abstracting) transformation, the backward transformation can use the original complete model to restore the lost information to create an updated complete model. The operational semantics of lenses are defined by term rewriting rules and are implemented using functional programming in order to ensure a one-to-one implementation of the formal specification. Because of that, it is possible to prove certain properties of lenses or to infer their properties by a type system.

My approach is to apply the concepts of lenses to metamodel-based model synchronizations. I propose a *lens-based multimodeling architecture* that incorporates a complete model of the system. Lenses are used to synchronize the models described in different DSLs with that complete model. As a consequence changes in the model of one DSL are propagated to the models of other DSLs representing the same information.

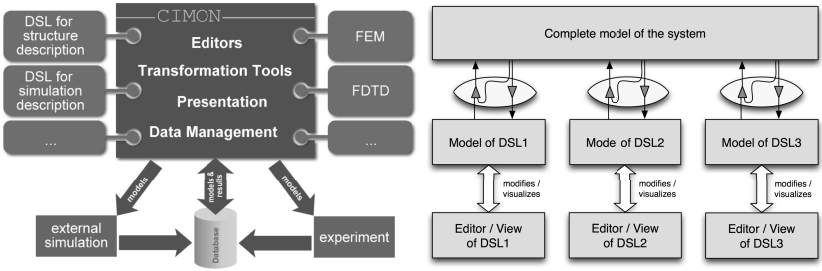

Figure 6.1: A model-driven infrastructure for nanostructure development and the underlying lens-based multimodeling architecture

7 GRK 1362: Cooperative, Adaptive and Responsive Monitoring in Mixed-Mode Environments

Prof. Dr rer. nat. Oskar von Stryk
Email: stryk@sim.tu-darmstadt.de
Technische Universität Darmstadt
Internet: `www.gkmm.tu-darmstadt.de`

The GRK adresses a variety of fundamental, scientific and technological challenges behind networked autonomous entities which accomplish a common task through actively monitoring the environment with a variety of stationary and mobile sensors, that acquire, manage and disseminate data, and finally utilize the data to react and possibly influence the environment. The problem space is defined along several dimensions including (i) perception abilities of entities like sensor complexity and heterogeneity, ranging from sensors producing simple readings, such as temperature or position, to complex sensors producing images or video, and from environments with many homogeneous sensors to environments with multiple types of interacting sensors, (ii) motion and behavior abilities of entities like sensor mobility, ranging from stationary sensors with known location to highly mobile sensors mounted on autonomous ground or aerial vehicles, (iii) communication abilities of ad hoc wireless networks subject to different constraints on bandwith, timeliness of data transmission, node mobility, dependability, and quality of service, and (iv) computing and middleware layers of heterogeneous networked entities. Mixed mode enviroments of networked autonomous entities are characterized by the heterogeneity of resources, capabilities, connectivity and requirements of the system and its components. The resulting, challenging research topics are structured along four main areas: A) Sensing and Monitoring, B) Foundations of Communication, C) Computing and Middleware, D) Cooperative Planning and Control. They require a close, interdisciplinary approach by the participating scientists of the departments of computer science, electrical engineering and information technology, and mechanical engineering. Another special strength of the GRK is the evaluation of methodologies developed in areas A-D in joint experimental setups for monitoring and coordination across networked heterogeneous entities including scenarios with human assistance for task fulfillment.

7.1 People Detection in Challenging Real-World Scenes

Mykhaylo Andriluka (andriluka@cs.tu-darmstadt.de)
Supervisor: Prof. Dr. Bernt Schiele

Localization of human victims is essential part of operation in any search and rescue scenario. In our work we develop computer vision algorithms which enable autonomous robots to detect victims in difficult cluttered scenes, estimate their pose and track their position over time. We envision that our methods will enable not only search of the victims but also assessment of their state using pose estimation and motion cues.

The focus of my work is non-rigid object detection and articulated pose estimation. Recently we have proposed a general model that extends the pictorial structures paradigm and combines discriminative modeling of object appearance with learned generative kinematic tree priors. Due to the large variability in potential human poses we avoid the use of strong priors on the spatial configurations of parts, and represent relative rather than absolute part configurations. Moreover, we employ densely computed local features and boosting to learn powerful discriminative part detectors, which allow to discard significant portions of possible object configurations. Through exact inference based on the learned kinematic tree model we then find the most probable configuration of body parts. We have shown that this hybrid model is equally suitable for both detection and pose estimation tasks, outperforming the state of the art on several recently proposed datasets.

Our model allows us to find people in images and estimate image positions of the key body parts. In our current research we focus on employing these 2D estimates in order to estimate poses of people in 3D and track them over time. We have also applied our people detection approach to recognition of human victims in images acquired from autonomous aerial vehicle. In the letter project we have actively cooperated with several other members of our research training group, who have provided their assistance during image acquisition and analysis of sensor measurements.

7.2 State Estimation in Distributed Systems

Dominik Haumann (dhaumann@rtr.tu-darmstadt.de)
Supervisor: Prof. Dr.-Ing. Jürgen Adamy

One goal in multi-agent systems is to solve common tasks in a distributed manner. The tasks can either be basic tasks such as synchronization or formation, or complex tasks such as rescue missions or surveillance. An important aspect to successfully solve these tasks is the appropriate coordination of the group. To this end, it is inevitable that each member of the group is aware of its state. A state can include robot positions and velocities, obstacles in the environment or arbitrary information that describe the underlying system. Each robot may have a different dynamics and a different kind of state. Hence, the team of robots is heterogeneous in the states and the dynamics.

Due to uncertainties in the sensors, each robot needs to apply filter algorithms to get the best estimate or probability distribution of its state. To increase the accuracy of the estimates, the idea is to incorporate information of the states of other robots in the group. In current researches, this distributed estimation process commonly assumes a synchronous system, i.e., robots can communicate at any time without delays. However, real world scenarios include communication delays and communication loss. Hence, distributed systems are inherently time-variant and highly asynchronous.

A well-known approach to represent distributed systems are graphical models such as Dynamic Bayesian Networks (DBNs). Such graphical representations visualize dependencies in a graph along time. Each node describes a state or measurement and each link between nodes describes a dependency. Information flow between nodes of different robots imply communication. For instance, if nodes representing the state of different robots are interconnected, the system dynamics of those robots are coupled, i.e., the state of one robot may influence states of other robots. This cannot be handled with state-of-the-art inference techniques anymore, if time shifts and different time scales between different robots are taken into account.

The objective of this thesis is to analyze where to take state vectors from other robots into account to maximize the estimation quality. This can either be directly in the system dynamics, leading to a thight coupling of the robots, or in the system outputs. Further, a goal is to synthesize filters especially suited for distributed, asynchronous systems considering the aforementioned constraints.

7.3 Device-Agent Based Middleware Approach for Mixed Mode Environments

Herzog Arthur (aherzog@dvs.tu-darmstadt.de)
Supervisor: Prof. Alejandro Buchmann

Mixed Mode Environments (MME) refer to networks composed of very different kinds of devices, which are distributed among various physical environments and communicate with each other using different communication technologies. The single nodes in the network can be sensors, actuators, robots, Unmanned Vehicles (UV), Human Interface Devices, Mission Control Stations, etc. All those devices have their specific capabilities and constraints. Many of these devices are manufactured by different companies and use different software and operating systems, if any. Those devices can communicate with each other by wire, radio, light, sound or other methods. For each of this communication methods many different communication technologies exist, which use different protocols, frequencies, encoding, etc.

The goal of my research is to enable interoperability among different nodes without the need of adjustments each time new hardware is introduced. The agent based approach offers an abstraction for the different hardware: it sees all the different nodes in the network as independent entities, we call them device-agents [A3ME]. These device-agents know their capabilities and constraints. Depending on its capabilities a device-agent offers services to other agents and can also have tasks it tries to fulfill eventually using services of other agents. The different device types, capabilities, services, etc. can be described using a predefined efficiently encodable classification [A3ME Classification]. The complexity of agents running for example on a small sensor node and on an UV can vary considerably.

Bibliography

[A3ME] Arthur Herzog, Daniel Jacobi, Alejandro Buchmann. *A3ME - An Agent-Based Middleware Approach for Mixed Mode Environments*. UBICOMM 2008, IEEE Computer Society Press, 2008.

[A3ME Classification] Arthur Herzog, Alejandro Buchmann. *Predefined Classification for Mixed Mode Environments*. Technical Report, No. 1878, Technische Universität Darmstadt, Germany, 2009.

7.4 Scopes - A Framework for Multi-Purpose Wireless Sensor Networks

Daniel Jacobi (jacobi@dvs.tu-darmstadt.de)
Supervisor: Prof. Alejandro Buchmann

An application scenario for *Scopes* is the following: After an incident at a chemical plant the equipment is partly destroyed, other parts are still functional. The statically deployed infrastructure to monitor safety and control all the systems is also damaged, but partly working. This infrastructure is deployed as wireless sensor network, so that after an incident it stays functional in not affected areas, at the incident site only single sensors may work and they may not be at the location where they were installed in the first place. So getting the remaining infrastructure back in a working state to be used by emergency response teams, like e.g. firefighters or robots, to get an overview of the situation in a short amount of time and identify areas with leaked chemicals not save for human personnel, is a first priority. For this purpose the robots are able to deploy additional sensor nodes to enhance the radio coverage, connect disconnected parts of the sensor network or to allow monitoring of critical areas.

The goal of the Scopes Framework is to provide a platform for wireless multi-purpose sensor networks. Today's wireless sensor networks are developed for one specific scenario or one specific task. But scenarios foremost from logistics or emergency response suggest, that sensor networks should be capable of multiple tasks. Additionally, as the network is deployed priorities for tasks can change or even new applications could emerge. Different users have to be able to send queries to the network or include new sensor nodes. To access only the specific parts of the sensor network, that are needed for a special task, it is important to be able to restrict the use of nodes via a set of properties.

The term 'multi-purpose' can be found in the framework on different levels. At first multiple applications should be supported in parallel. Next, the framework has to be flexible enough to be efficient in varying scenarios and being deployed on different sensor node platforms. With Scopes being used in these critical areas, there are also demands regarding security, like data security or access control, to name a few.

Bibliography

[1] Daniel Jacobi, Pablo E. Guerrero, Ilia Petrov, Alejandro Buchmann. *Structuring Sensornetworks with Scopes*. In: Adjunct Proceedings of 3rd European Conference on Smart Sensing and Context (EuroSSC08), 2008.

7.5 Multi-Modal Sensor Monitoring for Context-Aware Assistance

Matthias Kropff (Matthias.Kropff@kom.tu-darmstadt.de)
Supervisor: Prof. Dr.-Ing. Ralf Steinmetz

Providing environmental information to the user can be of great interest for several applications. In our rescue scenario we aim to monitor the operator in a mission control center to provide interruption management support by automatically controlling the available communication devices and adopt the information representation.
It has been shown that spontaneous interruptions of cognitive demanding user tasks, like in air traffic control, can have strong negative effects on task completion time and error rate. The analysis of the monitoring scenario showed a strong necessity for (a) semantic sensor description to enable different sensor search and retrieval strategies, agent-based pre-processing and inference, integration of a priori unknown sensor types, as well as context dependent information representation. Further, it was identified that (b) the human operator would very much benefit from a context-aware assistant agent, which adapts the information representation according to his working context.
Therefore the available communication middleware concepts have been analysed and extended by a lightweight semantic sensor description. Besides the semantic type identification, several Quality of Service parameters, like sensing accuracy or energy consumption were integrated to facilitate further optimization techniques for efficient communication. The resulting middleware demonstrates the integration of different sensor platforms, ranging from PCs, embedded systems, mobile phones and low-power wireless sensor nodes. This allows accessing multi-modal sensor information in the operator's proximity. The integration of information from video-, audio-, accelerometer- and psychophysiological sensors, as well as activity related information from PC-usage or RFID-based events facilitates to infer the working context of the operator. Based on the operators self-report, several Maschine Learning algorithms were analysed to infer automatically his level of interruptibility.

7.6 Synchronization of Multi-agent Systems

Kim Listmann (klistman@rtr.tu-darmstadt.de)
Supervisor: Prof. Dr.-Ing. Jürgen Adamy

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 a complex manifold and tools from geometric control theory can be used to derive suitable controllers for the formation.

A further step will be to evolve a motion planning algorithm for this formation of heterogenous robots. In particular the navigation function method offers a general framework to respect the constraints of the different vehicles and inherently design feedback controllers based on the mechanical nature of the systems. Merging the methods of formation control and navigation functions could result in a multi-robot system being able to respect the specific constraints of the different vehicles and to lead the group as a whole to the target points. Further examinations including the effects of communication constraints then have to be done before an application in real systems is possible. Generally, communication plays a crucial role in cooperative control, so that applicable controllers need to be robust with respect to link failures, packet loss and variable time delays when communicating.

Combining the different approaches will lead to a reactive formation control where obstacles/collisions will be avoided and exploration schemes may be enhanced, respecting the communication capabilities of the team.

7.7 Large Scale Replication and Data Consistency

Matthias Majuntke (majuntke@cs.tu-darmstadt.de)
Supervisor: Prof. Neeraj Suri

My current research can be divided into two directions:

Efficient Large Scale Data Replication & Access [1]
As high performance data storage and retrieval has developed into large scale data centers, their growing size also implies dealing with growing number of system and transaction failures. While the fault tolerant distributed systems community has developed multiple replication based sophisticated strategies to maintain data consistency in the presence of failures, the very nuance of "replication overhead" also limits their scalability. My research interests target the development of replication protocols that utilize contextual partitioning of data groups to achieve large scale efficient replication for data centers. With my research I aim to develop fundamental concepts in distributed state machine replication as well as its experimental validation via actual data center deployment studies.

Fork-Consistent Distributed Storage [2]
This reasearch area involves the problem of emulating a shared read/write memory using a storage server prone to Byzantine failures. Although cryptography can be used to ensure confidentiality and integrity of the data, nothing can prevent a malicious server from returning obsolete data. Fork-linearizability guarantees that if a malicious server hides an update of some client from another client, then these two clients will never see each others' updates again. Fork-linearizability is arguably the strongest consistency property attainable in the presence of a malicious server. Recent work has shown that there is no fork-linearizable shared memory emulation that supports wait-free operations. On the positive side, it has been shown that lock-based emulations exist. In ongoing work [2] I have developed for the first time lock-free emulations of fork-linearizable shared memory. Lock-based protocols are fragile because they are blocking if clients may crash.

Bibliography

[1] Dan Dobre, Matthias Majuntke, Marco Serafini, Neeraj Suri *HP: Hybrid Paxos for WANs*. EDCC, 2010. To appear.

[2] Matthias Majuntke, Dan Dobre, Marco Serafini, Neeraj Suri *Abortable Fork-Linearizable Storage*. OPODIS, 2009.

7.8 Real-time Obstacle and Terrain Mapping for Unmanned Aerial Vehicles

Johannes Meyer (meyer@fsr.tu-darmstadt.de)
Supervisor: Prof. Dr.-Ing. Uwe Klingauf

The research interest in unmanned vehicles has grown significantly during the last years. Especially small aerial vehicles (UAVs) can serve many tasks including surveillance, search and rescue, and autonomous exploration and mapping and tend to be deployed in obstacle-rich environments like urban or industrial sites. Even when not maneuvering totally autonomously the operator's awareness about the height above terrain and objects in the vicinity is crucial for the safe conduct of the flight. During autonomous operation obstacle information is used for reactive avoidance maneuvers and as input for strategic path planning. Compared to their ground counterparts, aerial vehicles make high demands on real-time availability of obstacle information due to their faster dynamics. As there are no external aids in most settings, information must be gathered and processed onboard taking into account the limited payload and energy capacity.

In the thesis several methods for autonomous obstacle and terrain mapping will be evaluated with respect to their feasibility for UAVs using simulations and real-flight experiments. Scanning laser range finders (LIDAR) emerged to a state-of-the-art technology for robotic mapping from ground vehicles and for the creation of digital elevation models using airplanes. However, until now only 2D scanners are available within the given constraints, which suffer from the limitation to a single scan plane at each time step. In contrast image-based obstacle detection by reconstructing 3D information from motion or using a stereo camera system can cover the whole field of view at once and is available at lower weight and cost. A combination of both techniques benefits from the accurate distance information on the one side and good sensor coverage on the other.

Once a suitable representation of terrain and obstacle data is available it is integrated into the control system and visualized at the ground control station as overlay to video images in order to maximize the operator's awareness of potential threats. If multiple vehicles are used for the mission, their individual map information can be shared in the network and thus lead to an increased efficiency and precision of the whole system.

For the experimental evaluation, an autonomous ground vehicle and a quad-rotor aircraft have been successfully developed. The vehicles navigate by using several sensors like gyroscopes, accelerometers, a barometric pressure sensor and GPS and can be additionally equipped with a camera, a LIDAR device or a thermal camera depending on the application.

7.9 General Concepts for the Support of Autonomous Teams of Robots by a Human Operator

Karen Petersen (petersen@sim.tu-darmstadt.de)
Supervisor: Prof. Dr. Oskar von Stryk

In the last years, much research has been done on autonomous robot behavior, world modeling, and cooperating teams. Although many results are promising, very often the resulting behavior is suboptimal, and to a human operator possibilities to improve the behavior seem obvious. However, these improvements are often impossible to be achieved by the robot team autonomously for two reasons:

1. Uncertainties or errors in the robots' sensor readings influence the world model, leading to wrong assumptions.

2. The robots' internal models are not as sophisticated as a human's model of the world. With a human's expert knowledge, many decisions of the robot team do not appear reasonable.

This research deals with the question of how a human operator can efficiently support an autonomous team of robots, in order to allow the operator to insert his or her expert knowledge into the system.

There are basically two kinds of interactions between the operator and the robots: interactions that are initiated by the operator, and interactions that are initiated by the robots. This work focuses on operator initiated interactions with the whole robot team. On this level, two interaction modes are identified: implicit influence of the robots' behavior by adapting the common world model, and explicit influence on the team's behavior by adapting the current task allocation.

The presence of an operator allows a lot of implications on the robot team design. The operator station can be used as central knowledge store for the common world model, the mission definition, and task allocation. This allows to support robust distributed planning algorithms by an optimal centralized planner to receive better approximations to the optimal solution. The combination of both allows to profit by the merits of both methods, while eliminating the drawbacks. Simultaneous use of different planners and the demand to allow the operator to adapt the task allocation requires to apply a hierarchical set of rules about which planner is allowed to allocate which tasks and robots, to avoid conflicts.

This concept can be applied to several multi-robot scenarios like cooperative exploration, search and rescue, and also robot soccer.

7.10 Design, Actuation, and Control for Biologically Inspired Four-Legged and Two-Legged Robots

Katayon Radkhah (radkhah@sim.tu-darmstadt.de)
Supervisor: Prof. Dr. Oskar von Stryk

Versatile and energy-efficient legged robots are object of interest in many areas. If compared with biological systems that routinely exhibit dynamic behaviors in complex environments, however, our legged robots still severely lack diversity in locomotion. Bridging the gap between artificial and natural systems requires not only better sensorimotor and learning capabilities but also a corresponding motion apparatus with variable elasticity. Within two parallel projects, various of these challenges are investigated. Work related to a four-legged robot has resulted in a simulation model that is moving as dynamically as an animal and capable of various gaits. Besides, a new concept for the adaptation of the leg stiffness by using extended antagonistic series elastic actuation modules has been developed. This promising concept is currently being compared to well known actuator designs for active and passive compliance. More details concerning the design and model of the robot can be found in [1]. Furthermore, a 3D simulation model for an elastic, biologically inspired bipedal robot has been developed. Focus lies on the realization of walking and running within the same kinematic leg design. Previously, multi-modal locomotion had been shown both in simulation and experiment for the Jena Walker II which represents a novel, elastic and biologically inspired, three-segmented robot that is attached at the trunk to a lateral guide. The modular robot system consists of rigid segments, actuated hip joint modules and adjustable elastic strings, some of them spanning several joints like knee and ankle joints. Currently, the designed novel bipedal model that is capable of stand-alone movements without any lateral guidance is being constructed and produced. In this context, the essential task of a-priori motor selection based on simulation data for elastic bipedal robots has represented an important component of the development process of the robot. For further information we refer to [2].

Bibliography

[1] K. Radkhah et al. *Design Considerations for a Biologically Inspired Compliant Four-Legged Robot*. IEEE RoBio, 2009.

[2] K. Radkhah et al. *Towards Human-Like Bipedal Locomotion with Three-Segmented Elastic Legs: Efficient Optimization of Walking Motion*. ISR, 2010.

7.11 Trajectory Planning and Task Allocation of Cooperating Vehicles: Discrete-Continuous Modeling and Optimization

Christian Reinl (reinl@sim.tu-darmstadt.de)
Supervisor: Prof. Dr. Oskar von Stryk

At its core, optimal task allocation and trajectory planning for cooperating vehicles is characterized by complex problems. A wide range of tasks, whose fulfillment significantly depends on physical motion dynamics, leads to a tight coupling of discrete structure and continuous dynamics in systems analysis as well as in optimal control.

In practical applications, either specific heuristics are used or the non-linear motion dynamics is considered under oversimplifying assumptions. Usually, existing approaches can therefore only limitedly be generalized or transferred to other vehicle classes and tasks.

In a continuously growing area of new applications for cooperative autonomous multi-vehicle systems, the development of broadly applicable methods is of particular significance. Our work presents a consistent concept for modeling, approximation and optimization that is based on the theory of hybrid dynamical systems, on non-linear mixed-integer optimal control and on model-predictive methods from control theory.

Using hierarchical hybrid automata allows for modeling the tight discrete-continuous coupling. By applying an appropriate transformation, the model is made accessible for mathematical optimization. In particular, linear approximations are used for reasons of numerical efficiency and global optimality of the resulting mixed-integer linear optimization problem.

Solving these discrete-continuous optimization problems allows to compute approximate solutions for various problems of cooperative behavior and can – for special cases – already be applied within real-time feedback control architectures in task allocation. For representative benchmark scenarios as well as for new problems – like maintaining wireless communication among vehicles – numerical results are presented, that are demonstrating the competitiveness of the proposed concepts and are sounding out their limits.

The presented methods allow for estimations in systems design and for reference solutions in development of heuristic controller mechanisms – focusing on the central problem of cooperating vehicles, considering physical locomotion dynamics and the characterizing discrete-continuous coupling of states.

7.12 Quality of Information in Wireless Sensor networks

Vinay Sachidananda (vinay@cs.tu-darmstadt.de)
Supervisor: Prof. Neeraj Suri

The core functionality of Wireless Sensor Networks (WSNs) is to deliver the required information considering the evolvable user requirements and operating conditions in the pre-incident, incident and post-incident phases. Usually, WSNs create information from raw data using in network processing and deliver this information to the user [2]. We refer to quality as the degree or grade of excellence, and Quality of Information (QoI) is the quality experienced/perceived by the user concerning the received information.
In the current literature QoI is usually considered for a specific WSN functionality (e.g., raw data collection, in-network data processing, data transport, etc) but isolated from other functionalities. Hereby, the main design objective is to create or deliver the information with the highest possible QoI regardless the varied and evolvable user requirements. However, this usually leads to over provisioning, which means inefficient use of limited and valuable resources such as energy and bandwidth. It is not efficient to use the "best" solutions [1] (processing techniques, protocols etc.) from source to the sink to deliver information with the required QoI. It is not always necessary to increase the quality, but sometimes to decrease it to save resources such as energy, bandwidth, and increase timeliness of information delivery without under-performing the required quality indicators/metrics such as accuracy. To achieve this solution, we can either change the entire deployed WSNs which is not always feasible so we rule this possibility out. We can work on the user evolvable requirements so that it's necessary to tune the different blocks mentioned above.

Bibliography

[1] Zahedi, S.Srivastava, M.B.Bisdikian. *A computational framework for quality of information analysis for detection-oriented sensor networks*, MILCOM 2008.

[2] Hossain M.A, Atrey P.K, El Saddik. *Modeling Quality of Information in Multi-sensor Surveillance Systems*, Data Engineering Workshop 2007.

7.13 Conditional Random Fields for Detection of Visual Object Classes

Paul Schnitzspan (schnitzspan@mis.tu-darmstadt.de)
Supervisor: Prof. Dr. Bernt Schiele

Object detection is one of the key tasks in computer vision and an essential part of unmanned vehicles. On the way to scene interpretation object detection and segmentation of given scenes seem to be fundamental. One promising model stores global statistics of objects of a given class and a classifier is carried forward to unseen images to infer hypothesis of present object instances [Dalal and Triggs]. In the presence of occlusion these global object descriptors often fail to reliably infer the correct hypotheses. Contrary, local or part based object detection was originally designed to handle flexible objects like humans and animals but they also achieved promising results for occlusion scenarios. Modelling the object as a composition of different parts showed to be benefitial since the presence of one object part determines the location of other parts and therefore the location of the entire instance.

Conditional Random Fields (CRFs) [Lafferty et al.] gained increasing attention as they can be designed to model these dependencies directly. Especially segmenting a given image is well covered by CRFs as fixed image regions can be linked to pairwise connected nodes of the model. In my work, I am focussing on CRFs to increase the discriminative nature of these models and thereby achieve better detection results. Furthermore, I am aiming at providing a reliable scene interpretation in order to be able to detect victims with unmanned vehicles. In this setting, I am facing the challenges of multiscale, multiview and heavily occluded object instances. Especially detecting victims after any desaster is very challenging since only parts of victims might be visible and humans could appear in any arbitrary pose.

Bibliography

[Dalal and Triggs] Dalal and Triggs. *Histograms of Oriented Gradients for Human Detection*. CVPR, 2005.

[Lafferty et al.] Lafferty and McCallum and Pereira. *Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data*. ICML, 2001.

7.14 Fault Characterization and Diagnosis in Distributed Systems

Marco Serafini (marco@cs.tu-darmstadt.de)
Supervisor: Prof. Neeraj Suri

My PhD research focuses on the diagnosis of distributed embedded systems. Systems in Mixed-Mode-Environments (MME) are characterized by a high level of interoperation between different distributed entities, which results in a higher degree of interdependencies. In this context, it becomes fundamental for each distributed entity to obtain a view of the operational status of the system, in terms of existing usable physical resources and associated functionalities. Fault diagnosis and resource location need to be done online, i.e., during system operations. They are thus inherently distributed. This in turn raises the problem of obtaining a distributed agreement among all participants in order to enable global reconfiguration actions.

The determination of faulty MME entities in a distributed manner must take into consideration that embedded and wireless networks are particularly subject to transient faults. The reasons are multiple. Since the communication medium is broadcast-based, shared and unreliable, message loss can induce nodes of the distributed system to incorrectly deem other nodes as faulty. Furthermore, the small geometry and low energy consumption of many embedded devices result in a higher rate of soft errors. In both cases, resources appear as faulty but are in fact correct. On the other hand, nodes may have latent faults that only display periodically rather than as permanent failures (crashes). In this case it is important that these nodes are isolated from the system.

Transient faults require system designers to make a tradeoff between over reaction, which unnecessarily makes correct nodes unavailable and may trigger frequent reconfigurations, and under reaction, which increases the latency between fault and reaction and can determine application outages. To this end, I have been working on a diagnostic protocol framework. From a distributed algorithm viewpoint, one wants to build location protocols that are fault tolerant and provably correct. In fact, resource management becomes a safety-critical functionality in disaster scenarios. From a stochastic viewpoint, it is necessary to define probabilistic analysis that allow to fix a correct tradeoff of the filtering given a time to failure and time to recovery of the different nodes.

Subsequent work of my thesis has focused on topics of coordination in distributed environments in different contexts, including loosely connected network characterized by partitions and asynchrony.

7.15 Distributed and Predictive Control of UAV Swarms for Measurement of Scalar Fields in Lower Airspace

Armin Strobel (strobel@fsr.tu-darmstadt.de)
Supervisor: Prof. Dr.-Ing. Uwe Klingauf

In this work I will show a solution to find and assign clouds of gas in the lower atmosphere with a swarm of UAVs. This application has requirements on the control of the UAV swarm. The swarm has to fly without any previous knowledge in the lower airspace where many obstacles and thus less possibilities to avoid collisions are. An decentralized control will support robustness against loss of single UAVs and will provide a flexible solution.

To solve this, I am using a combination of Nonlinear Model Predictive Control (NMPC), potential field and genetic algorithms. This combination provides an integer control system which fulfills all requirements mentioned before. First Simulation results show the advantages of this solution for this kind of application [1-10].

Bibliography

[1] Jude Allred et al. *SensorFlock: An Airborn Wireless Sensor Network of Micro-Air Vehicles*. 2007.

[2] A. Tsourdos, R. Zbikowski, B. A. White *Cooperative Control Strategies for Swarm of Unmanned Aerial Vehicles under Motion Uncertainty*. 2006.

[3] Arthur Richards, John Bellingham *Co-ordination and Control of Multiple UAVsCo-ordination and Control of Multiple UAVs*. 2002.

[4] Bumsoo Kim, Paul Hubbard, Dan Necsulescu *Swarming Unmanned Aerial Vehicles: Concept Development and Experimentation, A State of the Art Review on Flight and Mission Control*. 2007.

[5] David H. Shim, H. Jin Kim, Shankar Sastry *Decentralized Nonlinear Model Predictive Control of Multiple Flying Robots in Dynamic Environments*. 2003.

[6] William B. Dunbar, Richard M. Murray *Distributed Receding Horizont Control with Application to Multi-Vehicle Formation Stabilization*. 2004.

7.16 Towards Enablement of Mobile Orientation-aware Services

Farid Zaid (Farid.Zaid@kom.tu-darmstadt.de)
Supervisor: Prof. Dr.-Ing. Ralf Steinmetz

Orientation-aware services [Layar] are special kind of Location-based Services that in addition to *location* leverage the notion of direction. This is facilitated by integrating orientation sensors like compasses and accelerometers in the mobile device. However, such services face some challenges like sensor imperfections (both static and dynamic) and the need to efficiently model the surrounding 3D objects in a user-centric way (in contrast to 2D maps which offer a map-centric view) [LVis].

To tackle the aforementioned problems, the *iView* reference architecture is proposed. Basically, iView supports the mobile spatial interaction process by splitting the task between the iView Server and the iView Clien. As a response to a query associated with position ρ, the iView server returns an $iView(\rho)$ model, which holds both geometry description of the *scene* around position ρ together with points of interest registered in that scene. Here, an object geometry is simplified to its bounding box, which confines the maximum dimensions of the object and describes its location and orientation. Such simplified model reduces the need for complex processing on the client side, and enables user-centric computing of visible objects with minimal need to communicate with the server, thus saving both power and bandwidth consumption and allowing real-time interaction with the surrounding.

The iView client executes local spatial queries on the geometry included in the obtained iView model, and using the sensory data (position, tilt and compass bearing). A query aims at identifying the most probable object along the line of sight. This is achieved by both computing the effective area of the bounding boxes as they appear in the user's field-of-view, and by modelling the sensor error bounds. Probabilistic querying help associate query answers with confidence levels, which can convey to the user the quality of the answer and prompt her to give feedback to the query.

Bibliography

[Layar] *Layar: Browse the world.* http://layar.com/. Last accessed March 2, 2010

[LVis] R. Simon: *Mobilizing the Geospatial Web - A Framework and Conceptual Model for Spatially-aware Mobile Web Applications.* PhD Thesis. TU Wien, 2008

8 GRK 1387/1: dIEM oSiRiS - Integrative Development of Modeling and Simulation Methods for Regenerative Systems

Prof. Dr. rer. nat. habil. Adelinde M. Uhrmacher
Email: lin@informatik.uni-rostock.de
University of Rostock, `www.diemosiris.de`

The Research Training Group dIEM oSiRiS aims at developing modeling and simulation methods for regenerative systems. Regenerative systems are able to overcome significant perturbations and maintain autonomously their functionality in dynamic and uncertain environments. Thus, they are becoming of increasing interest in designing e.g. robust technical systems or effective medical therapies. In the Research Training Group dIEM oSiRiS we focus on cell biological systems, to be more concrete on the role of the Wnt-signaling pathways in the differentiation of neural progenitor cells. Signaling pathways are reaction networks, wich allow cells to respond to internal or external stimuli in terms of changes in gene expression. They are highly complex, since they influence multiple processes at different levels, from the molecule, over molecule complexes up to structural changes at the level of cells and cell ensemble. Thus, as an application field, signaling pathways are likely to propel the development of new modeling, simulation, analysis, and visualization methods. A salient feature of signaling pathways is that they operate on different spatial and temporal scales. This is reflected in the past and ongoing work in the research training school. Among variants of the π-calculus formalism addressing space and multi-level descriptions, annotations that support a retrieval and reuse of models, a portfolio of Gillespie related spatial simulation algorithms, means for automatically configuring simulators, and the integration of visualization methods, that make use of innovative layouts and linked and coordinated views, target the challenges in analyzing such cell-biological processes. The work in the Research Training Group is highly interdisciplinary. Scientists from Computer Science, Medicine, and Biology join forces to reveal new insights in the area of cell biological systems, to help establishing modeling and simulation as experimental methods in Biology and Medicine and to develop new innovative modeling and simulation methods for the analysis of regenerative systems. Methodological developments are driven by diverse fields of computer science which allows to address the various challenges of modeling and simulating regenerative systems in a comprehensive manner.

8.1 Brownian Dynamics Simulations for Cell Biological Applications

Fiete Haack (`fiete.haack@uni-rostock.de`)
Supervisor: Prof. Adelinde M. Uhrmacher, Prof. Ronald Redmer

The last years have seen an increase of interest in micro-scale approaches to cell biological and molecular simulations, as some biological processes like diffusional association or molecular crowding cannot be adequately treated at higher levels of abstraction. Among those methods Brownian Dynamics (BD) plays a central role, as it is particularly well suited to describe a wide range of cell biological processes like protein-protein association, molecular crowding and protein translocation. Incorporated in a structure-based simulation approach, BD can be used for the estimation of reaction or association rates [SGW07].
However, despite recent advances in the development of BD algorithms [KT08], they remain computationally expansive, which prohibits the simulation of longer periods in time and large scale systems.
The goal of the project is to develop and combine suitable algorithms to support an effective and efficient simulation of BD. Thereby the role of BD in Computational Biology and its potential application for cell biological simulations shall be explored.
Based on that, we aim for a simulation-based estimation of diffusion- and association rates of Nucleoredoxin (NRX) and Dishevelled (Dvl) interaction. The redox-sensitive association of NRX with Dvl could provide a new regulatory mechanism within the Wnt signalling pathway [FM06]. A further field of application will be the simulation of multi-particle systems by BD in order to study the effect of molecular crowding on the diffusion of macromolecules.

Bibliography

[KT08] K. Takahashi. *An Exact Brownian Dynamics Method for Cell Simulations*. In: Computational Methods in Systems Biology 6th International Conference CMSB 2008, Rostock, Germany, October 12-15, 2008.

[SGW07] M. Stein, Razif R. Gabdoulline and Rebecca C. Wade. *Bridging from molecular simulation to biochemical networks*. Curr. Op. in Struct. Biol. 17: 166-172, 2007.

[FM06] Y. Funato, T. Michiue, M. Asashima and H. Miki. *The thioredoxin-related redox-regulating protein nucleoredoxin inhibits Wnt-beta-catenin signalling through Dishevelled*. Nat. Cell Biol. 8:501-508, 2006.

8.2 Distributed Information Retrieval of Computational Biological Models

Ron Henkel (henkel@informatik.uni-rostock.de)
Supervisor: Prof. Andreas Heuer

Applying Information Retrieval (IR) techniques on model retrieval is gaining more importance with the fast growing number of computational biological models (hereafter named bio-models). Several work groups develop and use different formalisms and interchange formats for bio-models and store those models in various model data bases and repositories. Mostly, models are lost to the community because they lack a proper annotation or documentation or because they use an uncommon or new formalism.

Searching for biological models is hardly sufficient by querying raw model code of a certain interchange format. Presumably, it is worthwhile to include additional information about a model: (1) What is known about a model's constituents? (2) What kind of relations exist between the constituents? (3) How does a model behave under certain conditions? These questions can often be answered by annotations. Thus, a thorough annotation of bio-models enhances the semantic description of the modeled system by far and can contribute to sophisticated IR techniques. Efforts for enhanced model annotation, such as the *Minimum Information Requested in the Annotation of Biochemical Models* (MIRIAM) [N+05] approach, are already applied to some modeling formalisms, e. g. the *Systems Biology Markup Language* (SBML).

Overcoming the aforementioned losses will enable models to be shared and reused between communities. To render sharing and reuse of bio-models possible the ability to *find* bio-models (i. e. to efficiently *retrieve* them from a repository) is mandatory. However, the current state-of-the-art is to provide the user with an unranked list of models for his query [KMHK09]; as currently done in BioModels Database or in the CellML Model Repository.

The main question for this Ph.D. work is how the information available about a bio-model can be gathered, processed, analyzed, and stored to efficiently *retrieve* bio-models. Furthermore, a research question is how to *rank* the retrieved models with respect to a user's query and needs.

Bibliography

[N+05] Nicolas Le Novère. *Minimum Information Requested In the Annotation of biochemical Models (MIRIAM).* Nature, 2005.

[KMHK09] Köhn et al. *Towards Enhanced Retrieval of Biological Models Through Annotation-based Ranking.* Springer, 2009.

8.3 Efficient Non-spatial and Spatial Simulation of Biochemical Reaction Networks

Matthias Jeschke (matthias.jeschke@uni-rostock.de)
Supervisor: Prof. Dr. rer. nat. Adelinde M. Uhrmacher

Efficient simulation algorithms should find a good balance between the effort and resources required to produce results and their quality. The search for such methods does not necessarily start with the development of new techniques; there might be a plethora of existing algorithms already suitable for the task, so it could be a better idea to first have a closer look at these.

Evaluating the performance of algorithms, e.g., how fast or accurate they execute, involves thinking in *algorithm setups*, which include the actual implementation and possibly a set of parameters and sub-algorithms. Especially the influence of the latter should not be underestimated. For example, there are numerous variants for event queues and random number generators, so which one should be used for a) a given model and b) a selected algorithm? These considerations are taken into account in the first part of the thesis. General requirements for performing evaluation studies are defined and exemplarily applied to a large-scale design space exploration of non-spatial and spatial stochastic simulation algorithms.

Results from the evaluation could hint to possible *improvements*, which are central to the second part of the project. A special type of algorithm – an approximative variant – is extended to the spatial realm and compared to other methods. Having difficulties with larger models, some if its operations are additionally parallelized to exploit modern multi-core processors.

It eventually might turn out that neither existing nor improved methods are suitable for a task, which requires the *development* of new techniques. Individual-based simulation algorithms, e.g., based on Brownian dynamics, are very accurate, but often require much effort and computational resources to execute models containing a large number of entities. In contrast, population-based methods, while usually being faster, do not allow a detailed modeling of particle interactions. The final part of the dissertation is focused on the development of a hybrid between both simulation types, which allows to model at different levels of detail: focusing on individual particles but still allowing the presence of other species as populations and interactions between both levels.

8.4 Experimental Model Validation

Stefan Leye (stefan.leye@uni-rostock.de)
Supervisor: Prof. Adelinde M. Uhrmacher

Model validation is an essential step in the modeling and simulation workflow. While a plethora of methods exist to face the various aspects of model validation, current studies revealed that a better support is required, as users often lack the required mathematical background choose and configure the right method for the right validation task [1].

The aim of this research project is to develop an architecture that seamlessly integrates these methods in a flexible manner and makes them available to users at the appropriate step of a validation experiment. If possible the right methods should be chosen automatically, according to a given specification.

A first step for the flexible organization of validation methods is the identification of principal tasks in validation experiments. This facilitates a specific application of methods for each task and allows the combination of different methods facing individual tasks in a flexible manner. The identified tasks are the specification of requirements, configuration of the model, model execution, observation, analysis, and evaluation [2]. The specification of requirements is an essential step for the formalization of validation experiments. A format to represent the requirements should be adaptable to various experiment designs and furthermore form an interface to formal experiment descriptions in order to make those repeatable and comparable. The configuration of the model corresponds to the generation of test cases in software testing. Points in the model parameter space are selected, that need to be investigated in order to achieve the desired information about the validity of the model. Different methods of configuration algorithms, like optimization algorithms should be supported. The execution of the model may be biased by the simulation components due to accuracy issues or implementation bugs. Thus, it is important to investigate different simulator components (e.g., RNGs, event queue, etc.) and their impact. Furthermore, to identify interferences between specific model and simulator parameters, a separated configuration of both is required. The outputs of the model execution are retrieved during the observation. The goal of the this task is to collect as much information as necessary to allow a proper analysis of the results but as little information as possible to save memory as well as computation costs during the analysis. The analysis comprises two steps: the analysis of single simulation runs (e.g., calculation of the steady state) and the analysis of replication sets (e.g., mean of the steady states of different replications). During the evaluation the analysis results are used twofold. On the one hand feedback is produced for the configuration in order to identify additional interesting parameter combinations. On the other hand the result of the validation experiment are

presented, e.g., by creating a figure to allow a face validation.

The design of the "Flexible Architecture for Model Validation" (FAMVal) is based on the six tasks of a validation experiment [3]. It exploits the plug-in based structure of the simulation environment James II [4] to allow a flexible integration of methods for the configuration, analysis, and evaluation tasks. Those methods are realized as plug-ins and coupled to FAMVal by interfaces, which allows exchange, extension, and reuse of them. A hierarchical representation of experiment settings works as interface between formal specifications (or user interfaces) and FAMVal. The representation is flexible and tree-based and can be adapted to hold the settings of experiments with arbitrary structures. The execution of the model is done by external simulation tools (e.g., the BaseExperiment of James II), that can be coupled to FAMVal by implementing appropriate communication handlers. Thereby, different simulation algorithms, components, and environments can be used during the validation process. Information about the minimum data to be observed for the analysis, are offered by FAMVal to allow a proper instrumentation of the simulation runs.

In the future an experiment language (based on the representation of experiment configurations) shall be designed that facilitates the specification of validation experiments. Since simulation and analysis can be computationally expensive, a parallelization of FAMVal in every aspect possible shall be aspired. Finally, additional validation methods should be implemented.

Bibliography

[1] L. Felipe Perrone, C. Cicconetti, G. Stea, and B. C. Ward *On The Automation of Computer Network Simulators*. 2nd International ICST Conference on Simulation Tools and Techniques, 2009.

[2] S. Leye, J. Himmelspach, and A. M .Uhrmacher *A discussion on experimental model validation*. Proceedings of the 11th International Conference on Computer Modeling and Simulation, 2009.

[3] S. Leye and A. M. Uhrmacher *A flexible and extensible architecture for experimental model validation*. 3rd International ICST Conference on Simulation Tools and Techniques, 2010.

[4] J. Himmelspach and A. M. Uhrmacher *Plug'n simulate*. Spring Simulation Multiconference, 2007.

8.5 Multi-Level Modelling for Systems Biology

Carsten Maus (carsten.maus@uni-rostock.de)
Supervisor: Prof. Dr. rer. nat. Adelinde M. Uhrmacher

Systems biology is an integrative study and research field focussing on complex interactions in biological systems. The aim of systems biology is to study biological systems, like a whole organism or a single cell, in a holistic rather than a reductionist way. This general approach for understanding complex systems follows the ancient but still valid dictum of Aristotle: "The whole is more than the sum of its parts".

Modelling and simulation are important techniques in systems biology for analysing the highly non-linear complexity of such systems. Thereby, thorough decisions regarding abstractions have to be made, as modelling always means to abstract from the real world. The chosen level of detail naturally depends on the granularity of the available data and knowledge about the system of interest, but also on the objective of the model, i.e. its purpose or intention, and on the computational efforts that are needed for analysing the model. Therefore, different modelling techniques and formalisms supporting various abstraction levels are used in this field of research. While the application of ordinary differential equations (ODEs) for modelling biological systems goes back to the 1950s, in the last years a bunch of sophisticated model description languages have been developed that reflect the diversity of different problems to be encountered by systems biology today. ODEs can efficiently be simulated and they fit well to the metaphor of well-mixed biochemical solutions with high molecule amounts. However, low molecule amounts require stochastic methods and sometimes an individual-based model is desired. There is also an increasing need for modelling spatial aspects. To sum up, different modelling formalisms are suitable for describing models at different abstraction levels which are required due to diverse needs in studying the complex interactions of biological systems.

So far so good, but the demand for high levels of details in selective parts of a model while other parts – due to available data or computational power – have to be kept rather abstract, raises the need for different abstraction levels within the same model. Here, multi-level approaches come into play, e.g. the combination of population-based and individual processes or quantitative and qualitative approaches. Also diverse reaction kinetics could be subject of multi-level methods. The aim of this PhD project is to analyse modelling languages with respect to their ability for multi-level modelling of biological systems and thereby to formulate general requirements for the description of such models. Also part of this project is to extend formalisms and to develop languages facilitating multi-level modelling in systems biology.

8.6 Tutoring System for Modelling and Simulation for Biologists

Géraldine Ruddeck (`Geraldine.Ruddeck.uni-rostock.de`)
Supervisor: Prof. Alke Martens

Getting basic skills in modelling and simulation has become necessary for biologists dealing with systems biology. However, there is a very big step to do from the wet-lab to the dry-lab which requires practical training complementary to the theoretical lectures. For this purpose is a tutoring system the most appropriate: it enables a long-term training directly on the target medium and individual support.

First of all, it is necessary to identify the special needs of biologists while learning computational modelling and simulation. Which specific skills are essential to enable the collaboration between biologists and modelers, and which ones are particularly difficult to understand for the biologists?

Once the teaching content is identified, different technical areas have to be explored in order to develop a such tutoring system. One challenge concerns the checking of the user's entries while modeling in order to compute appropriate feedback at running time. Uncomplete models have to be analysed in this respect, as it is of little use for the learner to get no feedback until finishing the modeling process. The computing of the feedback itself involves to take into account teaching principles, intelligent user modelling and error classification.

The whole tutoring system should be reusable and expandable for other teaching areas or in order to update the teaching content or style. Developing and testing a modular and reusable framework for tutoring systems is also part of the thesis aims.

8.7 Mathematical Modeling of Cell Signalling: Kinetic Properties of the Wnt Pathway

Yvonne Schmitz (schmitz@informatik.uni-rostock.de)
Supervisor: Prof. Olaf Wolkenhauer, Dr. Katja Rateitschak

Processes in biological cells, like signal transduction, metabolism and proliferation, are affected by complex networks of biochemical reactions. Systems Biology aims at identifying relations between structure and function of these networks by mathematical modelling and computer simulations. The mathematical models, which describe temporal changes of protein concentrations, are based on nonlinear ordinary differential equations and on the interactions between the different network proteins. The model properties are investigated by formal and numerical methods, e.g. bifurcation analysis.

The Wnt signaling pathway plays a significant role in the differentiation of stem cells. Dysfunction of the Wnt pathway can yield serious consequences like pathological diseases such as Parkinson's disease and Cancer. The purpose of modelling and simulation in this project is a better understanding of the biology of signalling cascade and the cellular transport mechanisms that crucially influence the distribution of the proteins, as well as its role in neuronal differentiation.

In a Systems Biology approach we developed a mathematical model of the Wnt signaling pathway. Thereby we focused on the main protagonist β-catenin and its antagonists APC, Axin and GSK3, which control its cytoplasmic degradation. We investigated several hypotheses about their particular cellular distributions which led to the design of new experiments.

The best known model on Wnt signaling was developed and investigated by [1]. It has previously been shown that this Wnt model is 'sloppy', meaning that the dynamic behavior of the systems depends sensitively on only a few 'stiff' combinations of model parameters, suggesting that fits to collective properties may be an effective alternative to measuring parameters to extract model predictions [2]. We are using the approach of 'sloppy parameters' to investigate the sensitivity of the dynamical behavior of the system to experimental errors in the measured values and to rather arbitrarily estimated values used to set the parameters of the model.

Bibliography

[1] Lee *et al.*: The roles of APC and Axin derived from experimental and theoretical analysis from the Wnt Pathway. *PLoS Biology*, 2003.

[2] Gutenkunst *et al.*: Universally sloppy parameter sensitivities in systems biology models. *PLoS*, 2007.

8.8 Explorative Graph Visualization

Hans-Jörg Schulz (hjschulz@informatik.uni-rostock.de)
Supervisor: Prof. Heidrun Schumann

The exploration of graphs for analysis purposes has received increasing interest over the last years. What started out as a socio-mathematical tool for network researchers is nowadays applied to networks in many application domains ranging from homeland security to cell biology. Yet, the larger these graphs grow the more analysis needs to integrate explorative, visual methods alongside computationally complex algorithms. This has not only the benefit of utilizing the extremely efficient visual system of the human analyst, but also of bringing the analyst with her background knowledge and common sense back into the previously more and more automated analysis process. This development falls under the only recently coined term *Visual Analytics* [1].

In the context of the GRK dIEM oSiRiS, Visual Analytics of graphs is employed to gain insight into the time-varying structures of cell biological models. This raises several visualization challenges imposed by the models themselves and by the available associated data including biological experiment results from the wet-lab and simulation experiment data from the dry-lab:

- the models may be of different graph types and on different levels of scale – from small-scale network structures of signalling pathways to large-scale hierarchical structures of agent-based multi-level models,
- the inherently uncertain biological context from which the model parameters have been derived must be communicated – e.g., accuracy and precision of the corresponding lab experiment,
- the large quantity and stochasticity of the resulting data for multiple experiment setups and multiple simulation runs must be handled and visually integrated.

An integrated visual analysis of cell biological model structures generating a synergy, an added analytical benefit of the above aspects, is one of the main aims of the combined visualization efforts in the GRK dIEM oSiRiS.

Depending on the type of the model structure (e.g., hierarchy, network, hypergraph), different integrated visualization approaches have been developed in the GRK. For the coupling between model and simulation data, this includes for example a table-based visualization technique for attributed π-Calculus models, which exhibit a time-varying, bipartite graph structure during simulation. Its table-layout allows for an intuitive exploration of structural relations as well as node attributes at the same time [2]. Time-value

plots enhance the exploration and hint at time points of interest by evaluating a model's structural complexity and other parameters over time [3]. This visualization is used for model debugging and refinement.

For the integration of wet-lab data from the biomedical domain with the model, this includes the utilization of the visual analysis framework Caleydo for the combined exploration of pathways and gene expressions. In this specific case, the focus lies on the co-located, cooperative visual analysis by bringing together domain experts from different fields of biology and medicine and visually guiding them along the exploration process [4]. This is yet another current research topic in Visual Analytics and our GRK is a prime example for the necessity of interdisciplinary, cooperative analysis.

This summarizes some of the research on Visual Graph Exploration in the scope of the GRK dIEM oSiRiS. Other research topics within the GRK interlock with this work, as multiple research challenges need to be addressed simultaneously and in close cooperation in order to find a comprehensive solution. One example is the heterogeneity of the data and the visual integration of the multiple available data sources from wet-lab and dry-lab experiments, as well as data from online databases (e.g., pathways) along the lines of the interaction concept developed in [4]. Other research within the GRK more extensively explores time-varying properties of model structures and the integration of visualization into the model validation process.

Bibliography

[1] J. Thomas and K. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Press, 2005.

[2] H.-J. Schulz, M. John, A. Unger, and H. Schumann. *Visual Analysis of Bipartite Biological Networks* in VCBM'08: Proceedings of the Eurographics Workshop on Visual Computing for Biomedicine. pages 135–142, 2008.

[3] M. John, H.-J. Schulz, H. Schumann, A. Uhrmacher, and A. Unger. *Exploring Time-Varying Hypergraphs* in InfoVis'09: Poster Compendium of the IEEE Information Visualization Conference, 2009.

[4] M. Streit, H.-J. Schulz, D. Schmalstieg, and H. Schumann. *Towards Multi-User Multi-Level Interaction* in CoVIS'09: Proceedings of the Workshop on Collaborative Visualization on Interactive Surfaces, 2009.

9 GRK 1424: Multimodal Smart Appliance Ensembles for Mobile Applications(MuSAMA)

Prof. Dr.-Ing. Thomas Kirste
Email: thomas.kirste@uni-rostock.de
Universität Rostock
Internet: `http://www.informatik.uni-rostock.de/musama_homepage.html`

MuSAMA is based on the hypothesis that ubiquitous 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. The resultant concept 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 userŠs need emerges from local interaction of individual appliances. The application scenario for MuSAMA are instrumented rooms that support teams in knowledge exploration and knowledge integration based on distributed display facilities.

9.1 Smart Privacy Management in Ubiquitous Computing Environments

Christian Bünnig (christian.buennig@uni-rostock.de)
Supervisor: Prof. Dr. rer. nat. Clemens H. Cap

Privacy in ubiquitous computing environments mostly is considered as a problem of hiding personal information. While this is an important aspect of privacy on the data level it neglects that people often *want* to show certain information to other entities for the purpose of communication, customization or self-representation. Naturally humans intuitively decide which personal face to show depending on the current situation. The crucial issue is how ubiquitous computing environments with its potentially vast amount of communicated personal information influences intuitive capabilities to practise privacy. The typical approach to control information disclosure is to compile privacy policies describing which information to disclose to which entity in which situation. Unfortunately this contradicts to the way users normally practise privacy – they require to specify preferences in advance in an abstract manner. This works as long as users have a clear idea of situations to come and how to handle information in these situations. It fails when situations and communciated information gets more complex or are not known in advance.

An alternative to precompiled policies is to *learn* a user's privacy preferences. The basic idea is to automate and assist the creation and maintenance of privacy policies by observing users in their disclosure behavior in order to automatically correlate situations and disclosed information. The challenges on learning a users disclosure behavior are (1) input selection (determining disclosure-relevant context information and its levels of abstraction), (2) learner implementation (choosing an appropriate method with regard to accuracy) and (3) output representation (a user's view on and interaction with a learned output model).

This work investigates these three issues. A testbed has been developed which enables us to observe users in their interaction with services in ubiquitous computing environments concerning disclosure of personal information. The testbed consists of a tool for mobile devices which enables users to understand and control the flow of personal information and an environment emulator which provides privacy related interaction with typical services found in smart environments. The testbed is used to observe and learn user disclosure behavior and to evaluate the proposed learning based disclosure assistance system. The overall goal of this work is to provide an ad-hoc disclosure system for practising privacy more dynamically and intuitive than existing approaches.

9.2 Synthesizing Probabilistic Generative Models for Intention Analysis

Christoph Burghardt (Christoph.Burghardt@uni-rostock.de)
Supervisor: Prof. Thomas Kirste

Ubiquitous computing devices are becoming increasingly important in a modern, mobile world. Such devices are used in various situations, e.g. while repairing a machine, while giving a lecture in a smart classroom or during a team meeting. Nevertheless the user should not have to focus on the system, but on the primary task. Thus the idea is to have a system that can assist the user *proactively*.

Bayesian Filtering is the process of – given a sequence of sensor observations – finding out the most probable current state or intention of the user. Current research in ubiquitous computing uses sequential Monte Carlo methods to make predictions about the intention of a user within certain situations. The problems so far are the underlying user-models, which grow exponential with the numbers of possible actions, persons or devices involved. The main problem of this type of probabilistic model is how to generate *efficiently* these probabilistic models. The designer has to calculate all possible human behavior and determine the a-priori probabilities of each action based on common sense.

The focus of this thesis is the automatic generation of such models from a formal description of human behavior. I focus especially on two methods: Task models are widely used in the domain of human computer interaction to model the interaction of a user with a system. They describe the next possible actions, depending on the current context (execution history). Thus the directed acyclic graph of a task model represents the structure of the corresponding Markov model.

Another possible solution for solving this problem is to employ methods of partial order planning to define the structure and transition probabilities of the Markov model. We describe each action with preconditions and effects. These actions form the input of the planning process could be given by the respective participating devices. Just like devices today have a handbook or manual, the possible interactions with the device are expressed in form of preconditions and effects. This is like an electronic version of a manual.

The technique allows model that are automatically rebuilt when the current context changes, as devices or people leave or enter the ensemble. The model is adapted further to the users behaviour at runtime, as it learns the transition probabilities with the help of the Expectation Maximization algorithm. The output of the intention recognition is a declarative description of user goals. It thus forms the foundation for all further programs that try to support the inferred goal in the current ensemble in an optimal manner.

9.3 Distribution Strategies of Data and Functionality in a Spontaneous Linked-up Sensor Network

David Gassmann (david.gassmann@uni-rostock.de)
Supervisor: Prof. Dr. rer. nat. Clemens H. Cap

During the first funding period of the GRK MuSAMA the fellowship holder Henry Ristau extended the communication paradigm of Publish/Subscribe by introducing a new role - the processor. The created concept of Publish/Process/Subscribe (PPS) is not only able to publish and subscribe but also manipulate information. The data flows from the source (via the processors) to the sink by using the routing algorithm Announcement/Subscription/Publication (ASP) developed by H. Ristau.

A general problem of dynamic networks is the integration of new nodes or the discontinuation of existing nodes. Changes to the infrastructure lead to changes in the routing paths of the network. In terms of PSS it is likely that a processor is disappearing with the discontinuation of an existing node. As a result, there is a high potential risk that subscriptions can not be served (temporally). The implication of the mobility and selection of the processor needs to be examined in more detail.

ASP uses the best effort strategy. The delivery of a publications is not guaranteed. This is not adequate for all application scenarios. As a result various levels of fault tolerance up to transaction properties (ACID) should be supported depending on the kind of sent/processed publications as well as the context of publication (location of the sender/receiver, value of the data etc). The extended spectrum of new applications, e.g. payment service, makes the concept of PPS/ASP attractive to companies. A raising number of applications leads to more and more mobile devices participating in the network. As a consequence the mobile devices compete for the limited resources of the network. This makes it necessary to think about the integration, implementation and compliance of constraints in the concepts of PPS and ASP. These constraints include among others hardware restrictions (size of the internal memory, speed of the CPU etc), visibility restrictions (e.g. not every node has the permission to communicate with its neighbours) as well as security restrictions (protect the data from unauthorized reading and spoofing).

The communication paradigm PPS is characterized by the fact of time decoupling. That implies it is impossible to support applications based on a chronology of events. To support such applications there is a demand to overcome the time decoupling.

The opened issues compose the foundation of my research activities within the scope of the GRK MuSAMA.

9.4 Development of a Flexible, Locatable and Safe Radio Platform

Enrico Heinrich (enrico.heinrich@uni-rostock.de)
Supervisor: Prof. Ralf Salomon

Naturally, different appliances use different transmission technologies and protocols. Herein, the wireless technologies, such as Wi-Fi, Bluetooth and ZigBee, are of particular importance. In his preliminary work, Enrico Dressler [Dre08] has proposed a method for the seamless integration of Bluetooth devices into WLAN infrastructures. Moreover, the pertinent literature has presented a similar approach based on a sample application (conference system). All these approaches require an administered infrastructure, which restricts their use to certain localities. On the other hand, the support for ad-hoc meetings in unprepared locations, e.g., a spontaneous meeting in a train, is of increasing relevance. However, a general and robust wireless networking technology for this type of application is not yet available. In addition, the rapid development in the field of wireless communications does indicate a further divergence of technologies and protocols.

In order to support smart appliances ensembles within the context of MuSAMA, the development of a wireless communications platform appears to be of particular importance. This platform offers a simple way to integrate various techniques (such as WLAN and Bluetooth) and new protocols. For the application domain at hand, it would be particularly favorable if the involved devices themselves would also provide the necessary capabilities. Another important aspect is the requirement of the user to protect his data (privacy) [Bue09]. This is particularly important in open scenarios, such as the ones considered by MuSAMA. For example, during a large business meeting, competing partners need to protect their data from each other. Therefore, it is crucial that the devices can not be compromised by software attacks and that the user can control the availability of its information to other participants and / or devices.

Bibliography

[Dre08] Enrico Dressler and Raphael Zender and Ulrike Lucke and Djamshid Tavangarian. *A new Architecture for Heterogeneous Context Based Routing*. Proceedings of the 13th International CSI Computer Conference, Springer, 2008.

[Bue09] Christian Buennig. *Smart Privacy Management in Ubiquitous Computing Environments*. Human Interface and the Management of Information, 2009.

9.5 Interaction in High-resolution Display Environments

Anke Lehmann (anke.lehmann3@uni-rostock.de)
Supervisor: Prof. Dr. Oliver Staadt

Nowadays, large high-resolution displays (HRDs) are used in a wide application area like product engineering, geospatial imaging or scientific visualization. The advantage of scientific visualization on HRDs is the presentation of complex data in a higher level of detail as well as in the context of surrounding information. Common 2D desktop metaphors and traditional input devices (mouse, keyboard) are not well suited in large HRD environments. Consequently existing interaction techniques and interface metaphors have to be extended or adapted for large HRDs. Towards this challenge, various novel virtual reality interfaces and 3D interaction techniques were developed, for example, pen-based techniques, laser-pointers, tracking or gesture-based techniques. The question is how the user benefits to perform general or domain specific task on large high-resolution displays [1]. Especially the integration of and interaction with high-resolution display systems are explored insufficiently in the field of smart ensembles and multimodal appliances. The goal of this thesis is the development of methods and mechanism to integrate efficiently high-resolution interactive display systems in smart appliance ensemble in the context of MuSAMA. The Visual Computing Laboratory at the University of Rostock provides a high-resolution tiled display wall (6 x 4 tiles with a total resolution of 11520 x 4800) controlled by a cluster system, a 3D stereo display, various interaction devices and various camera systems (e.g., infrared, time of flight). In a smart ensemble many available devices may be integrated dynamically into the system and they cooperate with each other spontaneously. The challenge is the cooperation between the available sensor technology of the smart ensemble and the tracking system of the tiled display wall. An important factor is user tracking. The user's current position and task define the display system as an (inter-) active or passive device. A natural user interaction is achieved by tracking and gesture recognition in large display environments. Such interaction techniques are easy to learn and they are similar to the natural behaviour of humans. The system must be capable of detecting the most appropriate tracking mode for the user and the interaction device related to the current context of use.

Bibliography

[1] T. Ni, G. S. Schmidt, O. G. Staadt, M. A. Livingston, R. Ball, R. May. *A Survey of Large High-Resolution Display Technologies, Techniques, and Applications*. In Proceedings of the IEEE Conference on Virtual Reality, 2006, pp. 223-236.

9.6 Intelligent Integration and Dissemination of Services in a Smart Environment

Philipp Lehsten (philipp.lehsten@uni-rostock.de)
Supervisor: Prof. Dr.-Ing. habil. Djamshid Tavangarian

Many applications for teaching and research exist as conventional stand alone software components. In the future, these services shall be accessible as services within a smart environment. By now, there are some components which are already usable in a service-oriented architecture. But the hype on SOA in the last years has led to a multitude of mostly incompatible service technologies. One approach towards SOA interoperability is the use of an abstract service description and request language like the Service Technology-independent Language (STiL) that will be refined during this work.
Additionally, a process-oriented model will be used to identify potential service candidates from different areas and to embed them into the smart environment. Beside the service interface, user interactions and context data shall also be used for intelligent service brokerage.
Another target of the thesis is the evaluation of fusion concepts of existing service islands in heterogeneous, cross-institutional environments as well as the development of principles for their integration into smart environments. Research challenges are the interoperability of service technologies, their user-oriented transparency in usage and task migration in case of resource failures. Furthermore the integration of context information and user intentions into the process-oriented model and the functionality of the services will be focused.
The focused solutions for communivation between services and devices will be evaluated in the graduate school's reference scenario of a "Pervasive University".
One approach deals with the interconnection of mobile internet enabled devices and learning management systems with consideration of context information for service brokerage to enable a mobile and context-aware usage. Next steps intended to create abstract service interfaces to existing student management systems and the usage of location based services in a campus environment. These services form the foundation of the process-oriented model. By knowing the existing services the intentions of the user can be mapped to processes. In the next step a smart environment within the cloud of services can be formed by an intelligent and context-aware broker. Finally the whole chain needs to be tested and evaluated.

9.7 Query Processing in Ad-hoc Environments

Andre Peters (ap@informatik.uni-rostock.de)
Supervisor: Prof. Dr. Andreas Heuer

At MuSAMA data access, query evaluation on mobile devices and retrieval techniques face problems such as resource constraints, power limitations and time delays in information exchange. At this point a solution to manage such problems needs to be devised. Following the idea of data sources accessible via services, I have to concentrate on **indexing support** for service discovery, data replication/synchronization and data management at spontaneous linked environments. The objective is to develop mechanisms that will enable efficient adaptable processing of queries in ubiquitous environments. In general, query processing systems are tailored for specific types of queries on restricted data sets according to a particular semantic. However, answering special queries may require certain tasks. To illustrate this, consider a scenario of information service applications in Smart Labs. Among other features all available devices should be able to inform about their features and their stored data in real time. So standardized interfaces are needed to support such information application. Services are able to fulfill this task. A service interface typically is described by $\{f, X, Y\} : f$ the name or identifier of the service, X the input parameters and Y the output parameters. Additional parameters like *preconditions* and *postconditions* can be given. Using these information, heterogeneous software and hardware resources may be discovered and integrated transparently. An illustration of service data providers can be found in [1].
In order to evaluate the ideas I'm working on BlueS, a framework for information exchange in ad-hoc environments based on data and information services in Bluetooth environments. Here and at my research the focus lies on indexing solution as presented in [2]

Bibliography

[1] V. Cuevas-Vicenttin. *Towards multi-scale query processing*. Data Engineering Workshop. IEEE 24th International Conference on Data Engineering, 2008.

[2] A. Peters, A.Heuer. *BlueS: Indexing Support for Data Exchange in Ubiquitous Environments*. Ubicomm, 2009.

9.8 User-driven Visual Interfaces for Smart Meeting Rooms

Axel Radloff (axel.radloff@uni-rostock.de)
Supervisor: Prof. Dr. Heidrun Schumann

User-centered information visualization is a new challenging task, since different aspects have to be taken into account:

1. Properties of the user (perception, preferences, knowledge)
2. Current intentions and tasks
3. Data of interest
4. Used output devices

Developing a user-centered visual representation requires formalized appropriate models for describing these aspects. In the HCI-community, workflow models describing the role of a user, user models and task models are applied to designing multiple human-computer interfaces. In information visualization, however, only specific aspects are addressed, such as perception-based mapping design (see Healey at al.). The goal of my PhD thesis is to develop a systematic view on this problem and to provide solutions against the background of information representation in smart meeting rooms, especially the perception-based prevention of visual clutter supported by smart environments.

The human visual perception influences the quality and amount of information that can be extracted from a graphical representation. In contrast, visual clutter decreases the recognizability of information. Thus, concepts, such as the preattentive recognizability, have to be combined with the capabilities of the smart room, especially the tracking of a user, and the provided models (user model, workflow description via task models, online evaluation feedback) to reduce the visual clutter problem on multiple output devices.

9.9 Intelligent and Distributed Content Addressable Memory

Henry Ristau (henry.ristau@uni-rostock.de)
Supervisor: Prof. Dr. Cap

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.

9.10 Component-based Modeling and Simulation for Smart Environments

Alexander Steiniger (alexander.steiniger2@uni-rostock.de)
Supervisor: Prof. Dr. rer. nat. Adelinde M. Uhrmacher

Modeling and simulation, the experimentation with models, can be applied in various domains and with different intentions. Therefore, the range of possible applications of modeling and simulation in smart environments is wide. Human behavior models can be used to detect anomalies in the activity of users or to predict users' future actions in order to automate the execution of services. Signal models, representing the output of sensors deployed in a smart environment, can be used to generate test data for simulation-based testing of specific implementations. The modeling of test data can thereby interact with the simulation directly or can be done in a pre-process separately.

The thesis addresses the question, how modeling and simulation could be used in different stages of the development process of smart environments and their software in order to support the designers and developers. The focus is thereby on the development of modeling concepts and a library of model components using the plug-in-based modeling and simulation framework JAMES II (Java-based Multipurpose Environment for Simulation II) with its comprehensive portfolio of modeling and simulation methods. Addressing different levels of abstractions is one major challenge in developing such modeling concepts and component library.

In order to get an overview of the use of modeling in the domain of smart environments, existing modeling approaches were examined in consideration of their purpose and their modeling subject. According to the requirements to a model, different abstractions and simplifications of the real system can be discovered. Furthermore, there are approaches which combine different models to one complex model regarding different aspects, such as the user behavior or the positions of computational devices embedded in the smart environment. Furthermore, we started to examine, when and where simulation is used in the development process of smart environments. As the current research indicates, simulation is usually applied relatively late in the development process. For instance, simulation is used to test the run time behavior, accuracy, robustness and scalability of implementations in terms of simulation-based testing. As a next step we are going to examine the requirements to simulation in general and address the question, if there are further interesting use cases for simulation in the development process of smart environments.

9.11 An Integrated Approach for Task Specification for Smart Environments

Maik Wurdel (maik.wurdel@uni-rostock.de)
Supervisor: Peter Fobrig

Specifying users' tasks for smart environments can be a cumbersome action to be performed to develop systems which are able to anticipate users' intention and needs. On one hand approaches made in the ubiquitous computing research area usually define the task space on a rather low level of abstraction. On the other hand task modeling languages from the human computer interaction (HCI) research community lack expressiveness. However modeling tasks is still an important step in development to gain insights about users' needs using such a complex software system. Task modeling can be a vehicle to specify requirements about the envisioned software system as well a tool to start off with development. HCI task modeling techniques do not offer capabilities to handle complex scenarios occurring in smart environments. Especially the environment of task execution of the user needs to be taken into account to reflect the situational context of the user when performing a task. This is a current research issue.

In this thesis a new task modeling framework is currently under development which takes into account several models reflecting the situational context of task execution. Therefore a set of models were investigated and integrated using a generic approach. In particular the cooperation of actors, the location of objects and actors as well the systems state are taking account. However the framework is designed to able to further integrate other model types such as sensor values.

Additionally the framework has to be embedded in a development approach supporting incremental, iterative development which needs further tool support and integration into software engineering tool environments. Such an environment comprises creation, manipulation, visualization, validation and verification facilities.

The kernel of the resulting framework is based upon a specification language which basically defines in which sequence tasks can be performed according to the system state. Such a language of the potential task execution is particularly from interest when developing a system which should be able to anticipate users' intention, because a model of the potential task execution narrows the set of valid actions within the smart environment. A labeled transition system can be derived from the task specification representing all possible task traces which can be further enriched using probabilistic algorithms.

10 GRK 1480: Programm- und Modell-Analyse (PUMA)

Prof. Dr. Helmut Seidl
Email: seidl@in.tum.de
TU München and Ludwig Maximilian Universität München
Internet: `http://puma.in.tum.de`

The doctorate program (Graduiertenkolleg) PUMA brings together the four fundamental approaches of program and model analysis, namely, type systems, theorem proving, model-checking, and abstract interpretation. Its goal is to develop new analysis techniques through cross-fertilization. The new methods should prototypically be implemented and be evaluated on selected analysis problems of software-intensive systems.

In the last years, four competing approaches for program and model have been developed: verification through theorem proving, model checking, abstract interpretation and via type systems. The Munich universities hosting the program have expert researchers in all of these areas. Our goal is to stimulate cross-fertilization between these approaches resulting in a better understanding of their common basis and their distinctive properties, and leading to better algorithms and tools. Our vision is the Verifying Compiler, i.e., the development of methods and tools that examine not only whether a program or model is syntactically correct, but also whether it behaves according to its specification.

10.1 Counterexample Generation for Higher-order Logic Using Logic Programming

Lukas Bulwahn (bulwahn@in.tum.de)
Supervisor: Prof. Tobias Nipkow

Finding errors in specifications by failed proof attempts can be extremely time-consuming. Therefore, the theorem prover should also try to generate counterexamples to the statements the user is attempting to prove to point out mistakes.

The theorem prover Isabelle [3] already provides a counterexample generator based on random testing, similar to the Haskell tool QuickCheck [2]. This tool, with its unsystematic state space exploration, is mostly restricted to statements with weak preconditions and no quantifie But many statements occurring in practice do involve strong preconditions and quantifiers.

Our new approach explores the state space more systematically by reformulating the preconditions as logic programs that compute values meeting the preconditions. This way, a much smaller number of test cases suffices to find errors in specifications.

We adopt various techniques from logic programming: Formulas in predicate logic with quantifiers are translated to logic programs, which enables the execution of quantifiers when the system detects that the values of the quantified variables can be computed. Furthermore, recursive functions defined by equations are translated to logic programs by flattening, which enables the inversion of functions. The resulting logic program is analyzed by a mode inference [1]. This returns a data flow used to generate a functional program that explores the state space.

The work is driven by and evaluated on the numerous specifications developed in Isabelle/HOL, e.g., semantics of programming languages and specifications of cryptographic protocols.

Bibliography

[1] Stefan Berghofer, Lukas Bulwahn, Florian Haftmann. *Turning inductive into equational specifications*. Springer, 2009.

[2] Koen Claessen, John Hughes. *QuickCheck: a lightweight tool for random testing of Haskell programs*. ACM, 2000.

[3] Makarius Wenzel, Lawrence C. Paulson, Tobias Nipkow. *The Isabelle Framework*. Springer, 2008.

10.2 GAVS: Game Arena Visualization and Synthesis

Chih-Hong Cheng (chengch@in.tum.de)
Supervisor: Prof. Alois Knoll

We present an open-source tool called GAVS (an acronym derived from Game Arena Visualization and Synthesis), for the visualization and synthesis of infinite two-person games over finite directed graphs (see [2] for an overview of infinite games).

GAVS is designed for the following purposes:

- **(Education)** The teaching of infinite games can benefit from the tool. The front-end editor of GAVS supports features for game editing, graphical specification and visualization of strategies. The back-end verification engines, based on different acceptance criteria, are implemented with symbolic techniques (reachability, safety, weak-parity, Büchi, Staiger-Wagner), known heuristic algorithms (parity [3]), or game reduction techniques (Muller, Streett).

- **(Research)** As GAVS is open-source (both the front-end GUI and the back-end engine), researchers can modify the tool freely, and implement novel algorithms on top of the software framework. For example, we also use back-end engines of GAVS alone for other projects.

GAVS is released under the GNU General Public License (v3). The software package (`.jar` executable, test cases, manual, source code) is available at [1].

Bibliography

[1] GAVS project, http://www6.in.tum.de/~chengch/gavs/

[2] Grädel, E., Thomas, W., Wilke, T. (eds.): Automata, Logics, and Infinite Games: A Guide to Current Research. LNCS, vol. 2500, Springer, Heidelberg (2002)

[3] Vöge, J. and Jurdziński, M.: A discrete strategy improvement algorithm for solving parity games. In: 12th International Conference of Computer Aided Verification (CAV'00), LNCS, vol. 1855, pp. 202-215. Springer, Heidelberg (2000)

10.3 Probabilistic Verification Techniques and Probabilistic Program Analysis

Andreas Gaiser (gaiser@model.in.tum.de)
Supervisor: Javier Esparza / Helmut Seidl

10.3.1 Research during the first year in PUMA

My studies during the first year were centered around computing least fixed points of polynomial equation systems over semirings. In the beginning we derived small and rather technical results, e.g. computable conditions which showed the existence of nonnegative fixed points for special classes of polynomial equation systems. Our main project, together with Stefan Kiefer, was the design of efficient algorithms for two computational problems which are useful for studying *Probabilistic Systems of Polynomials*, a special kind of equation systems widely used in various areas of science. This work is described in more detail in the last section of this report. See also [1].

10.3.2 Vision of the thesis

My most recent project and presumably main topic of my thesis is probabilistic static analysis of programs using Abstract Interpretation. Some fundamental approaches were already developed in this field, e.g. in [3, 2, 4]. The authors develop theoretical foundations but hardly ever talk about efficient implementations or concrete instances of analyses.

We investigate possibilities of performing interval analysis on probabilistic programs. Hereby we especially want to compute the expectation value or "frequency of occurrences" of program configurations. We develop an abstract domain for the semantics of probabilistic programs with integer variables using a generalization of domains for interval analysis. We also had first advances in efficiently computing some of the needed operations. Our studies revealed some interesting mathematical problems. We also intend to work out useful and efficient methods for widening and narrowing operators and investigate efficient data structures for operations defined over the abstract domain. Sophisticated generalizations of Binary Decision Diagrams, like Interval Decision Diagrams, seem to be suited for that task.

We plan to combine techniques of the authors named above to develop a rather efficient approach for probabilistic program analysis, which we hope will turn out to be useful in practice. We also collect scenarios in which the information gathered by such analyses seems to be useful, e.g. compiler optimization, automatic parallelization of programs etc. In the long run, we would like to extend our currently still intraprocedural approach to an

interprocedural one and implement the developed algorithms in an analysis tool for simple probabilistic programs.

10.3.3 Summary of the presentation

In the presentation I will talk about the main research topic of my first year in the PUMA graduate school. This is joint work with Javier Esparza and Stefan Kiefer. We study systems of equations of the form

$$X_1 = f_1(X_1, \ldots, X_n), \ldots, X_n = f_n(X_1, \ldots, X_n),$$

where each f_i is a polynomial with nonnegative coefficients that add up to 1. We call them *Probabilistic Systems of Polynomials*. The least nonnegative solution, say μ, of such equation systems is central to problems from various areas, like physics, biology, computational linguistics and probabilistic program verification. We give a simple and strongly polynomial algorithm to decide whether $\mu = (1, \ldots, 1)$ holds. Furthermore, we present an algorithm that computes reliable sequences of lower and upper bounds on μ, converging linearly to μ. Our algorithm has these features despite using inexact arithmetic for efficiency. We implemented prototypes of both algorithms using the computer algebra system Maple and report on case studies, originating from applications in the area of nuclear physics and probabilistic verification, that show the performance of our algorithms.

Bibliography

[1] J. Esparza, A. Gaiser and S. Kiefer. *Computing least fixed points of probabilistic systems of polynomials*. Symposium on Theoretical Aspects of Computer Science (STACS), Nancy, France, 2010.

[2] D. Monniaux. *Analyse de programmes probabilistes par interprétation abstraite*. Thèse de doctorat, Université Paris IX Dauphine, 2001.

[3] G. Ramalingam. *Data flow frequency analysis*. SIGPLAN '96 Conference on Programming Language Design and Implementation, 1996.

[4] Alessandra Di Pierra, Chris Hankin and Herbert Wiklicky. *On probabilistic techniques for data flow analysis*. ENTCS, Vol. 190, Issue 3.

10.4 Quantitative Information Flow Analysis in Isabelle

Johannes Hölzl (hoelzl@in.tum.de)
Supervisor: Tobias Nipkow

An important question in computer security is whether some program leaks any confidential data. We assume the variables in a state of a program P are split into private and public variables. *Non-interference* states that the public output L of the program P is independent of its private input data H. There are various static analysis to ensure that a program obeys the non-interference property.

For some classes of programs this non-interference property is too strong. For example a password checker or the dining cryptographers protocol require to declassify small amounts of their private data. The *quantified information flow* analysis of the leaked data allows to quantify the amount of leaked data in bits. It uses *information theory* to quantify the *mutual information* between the private input H and the public output L. Mu and Clark [Mu] provide an automatic analysis to determine an upper bound of leaked bits for deterministic WHILE-programs. Similar to a model checker the state spaces of the analysed programs have a fixed size.

Coble formalized in his doctoral thesis [Co] information theory in the *interactive theorem prover* HOL4. He proved the anonymity of the dining cryptographers protocol for an arbitrary size in HOL4. He also implemented a proof method to verify the amount of leaked bits for fixed size problems.

My research project is to develop a framework using a Hoare-style logic in Isabelle/HOL to do quantitative information flow analysis. The information theory is formalized similar to Cobles [Co]. For programs of fixed size I want to implement an analysis similar to Mu and Clarks [Mu]. This needs to employ and extend the numerical decision procedures in Isabelle/HOL. A further goal is to prove *probabilistic* programs with finite state spaces of arbitrary size. Such a framework needs to reasons about the distribution of the results, here we look into utilizing an automatic FOL theorem prover using linear arithmetic.

Bibliography

[Co] Aaron R. Coble. *Anonymity, Information, and Machine-Assisted Proof.* PhD thesis, King's College, University of Cambridge, 2009.

[Mu] Chunyan Mu and David Clark. Quantitative analysis of secure information flow via probabilistic semantics. In *Proc. of the Int. Conf. on Availability, Reliability and Security (ARES)*, pages 49–57, 2009.

10.5 An Automatic Amortized Analysis of the Resource Consumption of Functional Programs

Jan Hoffmann (jan.hoffmann@ifi.lmu.de)
Supervisor: Prof. Martin Hofmann

The determination of the quantitative resource behavior of an algorithm is a classic problem of computer science. It is often referred to as the *analysis of algorithms* and elaborately discussed in many textbooks like *The Art of Computer Programming*. Quantities that are subject to the analyses include execution time and memory usage but also particular properties like the amount of data passed through a channel.

Quantitative analysis of algorithms is a non-trivial problem. Often, one is not only interested in the asymptotic behavior of an algorithm but rather in an exact determination of the resource costs of a concrete implementation. In fact, this aspect is present in *The Art of Computer Programming* where algorithms are implemented in an assembly language for the MIX architecture to determine their exact use of clock cycles and memory cells. Such concrete bounds can be employed in various ways in software development. Most notably, they can be used to determine the hardware requirements of embedded systems and to ensure the safety of real-time systems.

Even for basic programs, a manual analysis of the specific (non-asymptotic) costs can be tedious and error-prone. The problem gets increasingly complex for high-level programming languages since one needs to be aware of the translation of the compiler. As a result, automatic methods for analyzing the resource behavior of programs have been subject to extensive research.

Of course, one can not expect the full automation of a manual analysis that involves creativity and sophisticated mathematics. But in most resource analyses the greater part of the complexity arises from the glut of detail and the program size rather than from conceptual difficulty.

The state of the art in resource analysis research builds on various approaches of program analysis. The field of worst-case execution time (WCET) is mainly focused on the analysis of code with given inputs and deals especially with architectural features like caches and instruction pipelines. Complementary, there are methods to derive bounds on the number of loop iterations and recursive calls. For instance, the COSTA[1] project has made recent progress in the automation of the classic approach of deriving and solving recurrence relations that describe the program behavior. Another approach is to cleverly annotate programs with counters and use automatic invariant discovery between their values using off-the-shelf program analysis tools which are based on abstract interpretation (the SPEED[2] project).

[1] http://costa.ls.fi.upm.es

[2] http://research.microsoft.com/en-us/um/people/sumitg/pubs/speed.html

We developed a system that is an automation of the potential method of amortized analysis which was initially introduced by Sleator and Tarjan [Tar85] to analyze the efficiency of data structures. This approach was pioneered by Hofmann and Jost [Hof03] to infer *linear bounds* on the heap-space consumption of functional programs by integrating it into a type system.

Similarly, we used the potential method to develop a type-based automatic analysis system [Hof10] that computes *polynomial bounds* on the (worst-case) resource behavior of (first-order) functional programs. The analysis works without any program annotations and is fully automatic if a maximal degree of the polynomials is given. Our system is parametric in the resource and can compute bounds for every quantity that can be associated with an atomic step of the operational semantics. This includes clock cycles, heap space, and stack space.

Since the problem of deciding whether a given program admits a polynomial resource bound is undecidable in general there will always be programs with polynomial bounds for which our analysis unsuccessfully terminates. However, an implementation of the system showed that our method can compute time and space bounds for a number of interesting functions such as quick sort, merge sort, insertion sort, longest common subsequence via dynamic programming, and sieve of Eratosthenes.

The computation of the bounds is efficient and takes only a few seconds for the above programs. Our experiments showed that the computed bounds match exactly the measured worst-case behavior of many functions. A prototype implementation along with the examples is available online[3]. It is easy to use, adequately documented, and can be run directly in a web browser.

Compared to other approaches our system seems to better deal with recursion and inductive data structures. It is for example the only one that can automatically analyze functions like quick sort.

Bibliography

[Tar85] R. E. Tarjan. *Amortized Computational Complexity.* SIAM, 1985.

[Hof03] M. Hofmann and S. Jost. *Static Prediction of Heap Space Usage for First-Order Functional Programs.* POPL, 2003.

[Hof10] J. Hoffmann and M. Hofmann. *Amortized Resource Analysis with Polynomial Potential - A Static Inference of Polynomial Bounds for Functional Programs.* ESOP, 2010.

[3] http://raml.tcs.ifi.lmu.de

10.6 Verified Generic Local Fixpoint Algorithms

Aleksandr Karbyshev (aleksandr.karbyshev@in.tum.de)
Supervisor: Helmut Seidl

Generic local engines are core components of compilers and program analysis frameworks. They have first been proposed for the analysis of logic programs [2, 3, 4, 5] and model-checking [6], but recently have also attracted attention in interprocedural analyzers of imperative programs [7, 8].

In many application the information of interest can be specified as a *constraint system* $\mathbf{x} \sqsupseteq f_{\mathbf{x}}, \mathbf{x} \in V$, over some complete lattice D, where V is a set of variables and the right-hand side $f_{\mathbf{x}}$ of each variable $\mathbf{x}$ is given as a function of type $(V \to D) \to D$ implemented in some programming language. A *generic* solver computes a solution of a constraint system S making no assumption on the application domain D. A *local* generic solver, when started with a set $X \subseteq V$ of *interesting* variables, tries to determine the values for the X of a solution of the constraint system by touching as few variables as possible.

Efficient solvers for constraint systems exploit that often right-hand side functions query the current variable assignment only for few variables. A generic solver, however, must consider right-hand sides as *black boxes* which cannot be preprocessed for variable dependencies before-hand. Therefore, efficient generic solvers rely on *self-observation* to detect and record variable dependencies on-the-fly during evaluation of right-hand sides. The local generic solver **TD** by van Hentenryck [2] as well as the solver **RLD**, presented in this work, add a recursive descent into solving variables before reporting their values. Both self-observation through side-effects and the recursive evaluation make these solvers intricate in their operational behavior and therefore their design and implementation are error-prone.

The first issue in proving any generic solver correct is which kind of functions safely may be applied as right-hand sides of constraints. In the companion paper [1] we therefore have presented a semantical property of *purity*. The notion of purity is general enough to allow any function expressed in a pure functional language without recursion, but also allows certain forms of (well-behaved) stateful computation. Purity of a function f allows f to be represented as a *strategy tree*. This means that the evaluation of f on a variable assignment σ can be considered as a sequence of variable lookups followed by local computations and ending in an answer value.

It is w.r.t. this representation that we prove the local generic solver **RLD** correct. Our proof is formalized by means of the interactive theorem prover Coq. Related formal correctness proofs have been provided for variants of Kildall's algorithm for dataflow analysis [9, 10, 11]. This fixpoint algorithm is neither generic nor local. It also exploits variable dependencies which,

however, are explicitly given through the control-flow graph.

Our plans for a future work are to verify optimized versions of fixpoint algorithm **RLD**, to design local algorithms for special kinds of constraint systems as well as application of these algorithms in verified tools for program analysis.

Bibliography

[1] Martin Hofmann and Aleksandr Karbyshev and Helmut Seidl. *What is a pure functional?*. 2010, submitted.

[2] Baudouin Le Charlier and Pascal Van Hentenryck. *A Universal Top-Down Fixpoint Algorithm*. CS-92-25, 1992.

[3] Christian Fecht, *GENA - A Tool for Generating Prolog Analyzers from Specifications*. LNCS 983, 1995.

[4] Christian Fecht and Helmut Seidl. *Propagating Differences: An Efficient New Fixpoint Algorithm for Distributive Constraint Systems*. LNCS 1381, 1998.

[5] Christian Fecht and Helmut Seidl. *A Faster Solver for General Systems of Equations*. Sci. Comput. Program., 1999

[6] Niels Jorgensen. *Finding Fixpoints in Finite Function Spaces Using Neededness Analysis and Chaotic Iteration*. LNCS 864, 1994.

[7] Michael Backes and Peeter Laud. *Computationally sound secrecy proofs by mechanized flow analysis*. ACM Conference on Computer and Communications Security, 2006.

[8] Helmut Seidl and Vesal Vojdani. *Region Analysis for Race Detection*. LNCS 5673, 2009.

[9] Gerwin Klein and Tobias Nipkow. *Verified bytecode verifiers*. Theor. Comput. Sci., 2003.

[10] David Cachera and Thomas P. Jensen and David Pichardie and Vlad Rusu. *Extracting a Data Flow Analyser in Constructive Logic*. LNCS 2986, 2004.

[11] Solange Coupet-Grimal and William Delobel. *A Uniform and Certified Approach for Two Static Analyses*. TYPES, 2004.

10.7 Automatic Error Correction of Java Programs

Christian Kern (kernch@in.tum.de)
Supervisor: Prof. Dr. Javier Esparza

Software debugging is an expensive and time consuming task. It has been estimated that software debugging accounts for 50% to 80% of the total software development and maintenance costs within a software project [1]. A lot of effort has been spent in this area and tools were developed to support the programmer. However, these tools are not sophisticated enough to locate and explain the bug automatically. We propose an approach for the automatic correction of bugs in Java programs, i.e., automatic debugging.
We present an algorithm, that takes as input a Java program showing some unexpected behavior with respect to a complete specification. Our implementation builds up on the JavaPathFinder[4] model checker. The algorithm tries to construct a corrected program, i.e., a program satisfying the specification. We assume the bug can be fixed by mutating the program code at automatically selected program points.
Therefore, heuristics are applied, using static code analysis techniques to extract program locations where a bug could have been unintentionally injected by the developer. Next, replacements for those program locations are automatically generated. Our algorithm uses search strategies to derive correct program candidates, that satisfy the specification for a given set of test inputs. A model checker is used to definitely verify the candidate programs for correctness with respect to the specification.
We show that our approach is able to automatically repair real world defective programs (for example sorting algorithms), extracted from various internet sources.
We are currently creating more sophisticated, heuristics which are able to reason about complex programs that make use of advanced data structures, like for example linked lists. Furthermore, we want to extend our approach to concurrent programs. We will also create a plugin for the Eclipse Framework[5], to facilitate developers debugging their code more efficiently and in an easy way.

Bibliography

[1] JS Collofello and SN Woodfield. Evaluating the effectiveness of reliability-assurance techniques. *The Journal of systems and software*, 9(3):191–195, 1989.

[4] http://javapathfinder.sourceforge.net
[5] http://www.eclipse.org

10.8 Information Flow Analysis in Business Processes

Máté Kovács (kovacs@pst.ifi.lmu.de)
Supervisor: Prof. Dr. Martin Wirsing

During the last three decade much effort has been put into the research on the specification and verification of information flow policies on computer programs. The formal requirement against a secure program is referred to as the noninterference property [1], which intuitively states that values of variables of higher security classes should not interfere with lower priority variables. Denning and Denning showed a method to verify lattice-based information flow policies on structured programs [2], the soundness of which has been proved by Volpano et al. [3]. The current state of the art offers solutions for the compilation time verification, and the run time monitoring of the compliance with information flow policies of programs written in various high level and object-oriented programming languages.

However, today's Web services based distributed information processing technologies pose new challenges. The canonical data representation format used in messages and as values of variables in orchestration languages (e.g. BPEL [4]) is XML, information on business items is maintained hierarchically allowing for the usage of the appropriate abstraction level at design time. On the other hand these languages are used to implement business logic, many times conducting autonomous communication with external partners through open interfaces.

It is a valid expectation of workflow designers to be able to assign information flow policies to different pieces of information describing one specific business item residing in the same document tree. In my work I investigate how information flow can be controlled in Web service based environments, tanking into consideration the structure of tree-like data.

Bibliography

[1] J. A. Goguen and J. Meseguer. *Security Policies and Security Models.* IEEE, 1982.

[2] D. E. Denning. and P. J. Denning. *Certification of Programs for Secure Information Flow* ACM, 1977.

[3] D. M. Volpano, C. E. Irvine and G. Smith. *A Sound Type System for Secure Flow Analysis.* Journal of Computer Security, 1996.

[4] A. Alves, A. Arkin and S. Askary et al. *Web Services Business Process Execution Language Version 2.0* . OASIS, 2007.

10.9 Automata for Program and Model Analysis (working title)

Markus Latte-Freiberger (markus.latte@ifi.lmu.de)
Supervisor: PD. Dr. Martin Lange

An generic extension of the branching-time logic CTL is introduced by refining the temporal "until"- and "release"-operators. For instance, a language may determine the moments along a path that an "until"- property may be fulfilled. We consider several classes of languages, represented by automata on finite words, which lead to a range of logics with different expressive power and complexity. The importance of such logics is motivated by their use in model checking, program synthesis, abstract interpretation, etc.

The whole machinery is intended to be applied to CTL* (and other logics). However, the presented algorithm which decides whether a CTL-formula is satisfiable bases on a tableau for CTL. To apply previous extensions to CTL*, also a tableau for CTL* is useful. Due to the lack of such tableaus, an appropriate one for CTL* is proposed [1] as an intermediate goal.

Bibliography

[1] O. Friedmann, M. Latte, and M. Lange. A decision procedure for CTL* based on tableaux and automata. To appear in Proc. of the 5th Int. Joint Conference on Automated Reasoning, IJCAR'10, Edinburgh, UK, 2010.

10.10 Program Analysis with Horn Clauses

Andreas Reuss (a.reuss@in.tum.de)
Supervisor: Prof. Dr. Helmut Seidl

My research so far led to two main results. In particular I have developed two new *extremely efficient* (meaning asymptotically optimal) algorithms for the computation of two very general graph relations, namely

- dominance (which is used for many purposes in program analysis and optimization, e.g. for control dependence computation)
- the binary relation induced by the CTL operator **AF** which can be used to compute – in linear time – the general version of control dependence [RABHD] for arbitrary graphs where no dominance relation needs to be defined (and therefore cannot be computed)

For the latter, no efficient algorithm was known, whereas for dominance computation previous algorithms were very complicated and therefore not favoured in practice. The new algorithm is purely graph search based and has only little overhead.

My main direction of research however is to investigate in which way logic (program) specifications – in particular Horn clauses – can be used to do or support program analyses. Horn clauses are very suitable for the control flow analysis of programs and protocols. In [NNS] the Spi calculus is analyzed by means of Horn clauses. Further directions of research also include SAT solving.

Bibliography

[RABHD] Venkatesh Prasad Ranganath, Torben Amtoft, Anindya Banerjee, John Hatcliff, Matthew B. Dwyer. *A new foundation for control dependence and slicing for modern program structures.* ACM Trans. Program. Lang. Syst., 2007.

[NNS] Flemming Nielson, Hanne Riis Nielson, Helmut Seidl. *Normalizable Horn Clauses, Strongly Recognizable Relations and Spi.* 9th Static Analysis Symposium (SAS), LNCS 2477, 20-35, 2002.

10.11 A Type System for Amortised Heap-Space Analysis of Object-Oriented Programs.

Dulma Rodriguez (dulma.rodriguez@ifi.lmu.de)
Supervisor: Prof. Martin Hofmann

Space analysis of computer programs is important for many applications, especially in the ever-growing area of embedded systems because of the stringent space requirements in such applications. Moreover, memory usage influences time complexity. In languages with garbage collection like Java, where memory exhaustion will invoke garbage collection, manipulations of a heap indirectly slow down execution.

A variety of approaches to resource analysis have been proposed based in particular on recurrence solving, abstract interpretation and amortised analysis [HJ06]. In the amortised analysis case, very good bounds are produced based on the solution of linear inequalities.

We are currently developing a type system that provides upper bounds on the heap-space requirements of object-oriented programs with explicit deallocation.[6] The system was first described and proved sound by Hofmann and Jost in [HJ06]. We are interested in making it as expressive as possible, to maximise its usefulness. For example, we can code linked lists and appropriate operations, and sorting algorithms like insertion sort and merge sort.

We provided an efficient type checking algorithm [HR09a] and a subtyping algorithm [HR09b] for the system. Our next goal is the type inference, i.e. finding typing annotations automatically by a constraint-based analysis. This will allow analysis to be performed on programs with no type annotations, i.e. code similar to standard Java code. The resulting type annotations can be regarded as a certificate of bounded resource consumption that can then be attached to the program using Proof-Carrying Code (PCC) technology.

Bibliography

[HJ06] M. Hofmann and S. Jost. *Type-Based Amortised Heap-Space Analysis (for an Object-Oriented Language)*. ESOP, 2006.

[HR09a] M. Hofmann and D. Rodriguez. *Efficient Type-Checking for Amortised Heap-Space Analysis*. CSL, 2009.

[HR09b] M. Hofmann and D. Rodriguez. *Membership Checking in Greatest Fixpoints Revisited*. FICS, 2009.

[6] Resource Aware JAva (RAJA), http://raja.tcs.ifi.lmu.de.

11 GRK 1487: Self-organized Mobile Communication Systems for Disaster Scenarios

Andreas Mitschele-Thiel
Email: gs-mobicom@tu-ilmenau.de
Ilmenau University of Technology
Internet: `www.gs-mobicom.de`

Todays mobile communication systems are highly static and inflexible with regard to changes. This prevents the fast implementation of innovative ideas as well as the fast adaptation to changes of the environment, e.g. to changing service demands, changing traffic loads or traffic distributions in mobile environments.

Our research goal is to develop mechanisms for the autonomous, dynamic and distributed coordination (self-organization) of future mobile communication systems. We apply self-organization to the radio system, to protocols and the network as a whole. Our specific application scenarios are disaster networks, requiring an aggressive approach to network operation. The coordinated adaptation of radio, protocols and network aspects is essential to provide communication services in disaster scenarios, e.g. to support the localization of individuals or emergency communication.

Important topics addressed by our research are:

- Decentralized Information Management, focusing on the reliable and robust provision of information in mobile communication networks as the base for self-organized decision making,
- Self-organized Service Recovery, focusing on mechanisms to recover from the failure of services,
- Cognitive Management of Transport Resources, focusing on mechanisms to provide transport resources, comprising wireless as well as wired links, and
- Reconfigurable Radio Interfaces, focusing on fundamental issues and basic requirements for

Common to all these topics is their requirement for an autonomous and distributed coordination.

11.1 Routing Optimization Using Traffic Prediction

Abutaleb-Abdelmodhi Turky (abutaleb-abdelmohdi.turky@tu-ilmenau.de)
Supervisor: Andreas Mitschele-Thiel / Horst-Michael Gross

The efficiency of Traffic Engineering (TE) schemes mainly depends on route optimization. Most routing algorithms use the information of available bandwidth to choose the paths between the source and destination pairs. The provided QoS depends on the accurate measurement of the available BW. Due to the varying nature of the available BW, updating the link state with the current measured BW is not an efficient approach to represent the link utilization. Therefore, new approaches perform an estimation of the link utilization in the future from the actual traffic profile. In the case of disaster scenarios, it is very important to handle link failures. Therefore, one of the vital targets is the enhancement of the performance of routing recovery method using self-organized approaches.

The proposed routing recovery method should improve the network survivability, decrease the cost of the computation complexity and consider the importance of reducing the interference between the requests in the future. In this study, we introduce a new efficient approach for optimizing the routing performance in IP based networks. The approach uses an Artificial Neural Network (ANN) to build an adaptive predictor that predicts future link loads. Combining the predicted link load with the current link load is an effective method in order to optimize the link weights. ANN offers accurate prediction capabilities with different types of network traffic and has the ability to be adaptive. Another research object is introducing a new efficient prediction-based decentralized routing algorithm, which is based on the Ant Colony optimization. In both algorithms, we try to reduce the rejection ratio of requests, maximize the percentage of accepted bandwidth and reroute the requests upon link failure in an optimal way.

The prediction-based decentralized routing algorithm requires a decentralized information management system to handle the interaction between the mobile agents and to acquire the link state information.

11.2 Optimization of Spectrum Handoff in Cognitive Radio Ad-Hoc Networks

Md. Akbar Hossain (mdakbar.hossain@tu-ilmenau.de)
Supervisor: Andreas Mitschele-Thiel

Due to the enormous development in communication technology, demand for the spectrum is expected to grow even more tremendously in the coming years. To cope up with the ongoing demand, one of the most promising concepts to facilitate the flexible usage of the radio environment is Cognitive Radio (CR) technology. Disaster scenario is the most amazing usage case of CR networks, where CR node can use the best available spectrum to establish an ad-hoc network through its cognition capability and reconfigurability. However, with most of the spectrum being already allocated, it is becoming exceedingly hard to find unoccupied bands either to deploy new services or to enhance the existing one due to the lack of frequency bands in radio spectrum. In this case, the most important challenge is to share the licensed spectrum without interfering with transmission of other licensed users. Hence, CR node should immediately vacate the spectrum upon detection of licensed user on the particular frequency band, which known as spectrum handoff or spectrum mobility. In order to manage the spectrum handoff, current researches are mainly focus to stay in licensed spectrum. In addition, there are some other proposals that take in consideration of both licensed and unlicensed frequency band. However, most of the cases they manages the channel state information in static manner which it is very much impractical due to dynamic nature of the spectrum allocation. The idea of my work is to manage the information about the availability of spectrum in such a way that CR users can reduce the spectrum handoff. To achieve that, my focus is to develop an algorithm to discover the spectrum opportunities as fast as possible to incorporate with Dynamic Spectrum Access (DSA). The benefits and drawbacks of such strategies will be compared with more conventional approaches. A mixture of simulation and analysis will be used to assess performance of the algorithm. Game theory and Markov analysis will be particularly important analytical tools for the spectrum selection process among the secondary users.

11.3 Self-Organized Service Placement in Mobile Communication Networks

Christian Brosch (christian.brosch@tu-ilmenau.de)
Supervisor: Prof. Dr.-Ing. habil. Andreas Mitschele-Thiel

In today's mobile communication networks services usually implemented in a static and centralized way. For this reason failures of several network elements in disaster scenarios lead to extensive service breakdowns. But especially in disasters services are very important for an effective work of the actions forces. Here notably important services are telephony or messaging for the distribution of mission orders or navigation services for the observance of open roads.

Services in communication networks need three basic resources: Transmission resources are necessary to receive requests and transmit the results to the service's user. Furthermore these resources are used to synchronize the several entities of the service within the network. Memory resources are required to store temporary or permanent data required by the service, e.g. large databases or multimedia files. Last but not least, computation resources are needed to process the requests. Existing service placement solutions most often take only one or two of these resources into account.

The goal of this work is the development of algorithms for service placement, which are completely self-organized and take all service resources into account. Through the self-organization, high availability and sufficient reliabilty of the services should be ensured. The algorithms to be developed should be as general as possible to cover many services.

This topic has some natural interfaces with other topics of the Graduate School. The placement algorithms surely require some topology and resource information from the resource map (Jin Yang). The decentralized information management (Elizabeth Ribe-Baumann) may be helpful to store some parameters for the service placement algorithms.

Bibliography

[Wittenburg08] Wittenburg, G. and Schiller, J., *A Survey of Current Directions in Service Placement in Mobile Ad-hoc Networks*, Proceedings of the 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications, 548-553.

[Gramoli08] Gramoli, V., Kermarrec A. and Le Merrer, E.*SONDe, a Self-Organizing Object Deployment Algorithm in Large-Scale Dynamic Systems*, EDCC 2008. Seventh European Dependable Computing Conference, (2008) 157-166.

11.4 A Framework for Robust and Efficient Movement-Based Message Forwarding

Markus Brückner (markus.brueckner@tu-ilmenau.de)
Supervisor: Prof. Dr.-Ing. habil. Andreas Mitschele-Thiel

This work will develop a system for message-based, non-interactive mobile communications in a disaster scenario. The basic communicating entitity is a mobile node equipped with any kind of peer-to-peer capable radio transmission system (e.g. WiFi or Bluetooth) and being carried by a person or vehicle. These nodes form a Disruption Tolerant Network (DTN) that can be used to forward any kind of message from a sender to one or more targets. The main focus of the work is on two things: node localization/movement estimation and – based on the movements predicted from different information sources – optimized message forwarding.

To localize a node it is assumed that a small percentage of all nodes are equipped with a reference positioning system like GPS or Galileo. These nodes serve as beacons based on which all other nodes can calculate their positions to a certain precision in a distributed fashion just taking node contacts over time into account. The work also assumes that individual node movement in case of a disaster is not random. It more or less targeted towards a point on the map (e.g. a shelter or hospital) and follows roads and ways that already existed before the disaster. Even flying vehicles do not fly in a random manner but follow a mission that leads them on an efficient way from a source point to a target. Exploiting the non-randomness in the movement pattern was shown to enable prediction of future node positions with a certain accuracy (see [LOC08]). The work intends to improve on the existing results by adapting them to disaster scenarios.

Using the predicted position data we establish the concept of a time- and space-dependent routing graph which allows for the use of existing path finding algorithms to calculate routes from a source to a target. It enables a node to predict the effects of a forwarding descision and thereby improve the descision quality. The rationale behind this is that an improved descision quality results in less communication overhead, lower delay and higher delivery ratios. Existing works like PROPHET [PRO03] or MobySpace [MOB05] try to improve descision quality by taking repeated contact pattern or prediction based on the node movement into account. This work will try to improve on that by trying to predict multiple forwarding steps. We hope that this will enable us to predict routing paths more accurate as well as derive additional information like delivery probability and expected delay.

The central problem of the routing algorithm is the efficient creation and update of the forwarding graph. In a disaster scenario computational power is limited, so there is a need for efficient algorithms to fulfill the task. In

order to create a flexible and robust solution the routing service needs to be able to cope with outdated, incomplete or even plain wrong data. The work will therefore also explore possibilities to recover from route failures.

In order to provide a solution for message-based communications in disaster scenarios that can scale over a wide range of network environment from sparsely connected areas to high-densitiy scenarios we strive for an integration of the different routing approaches being worked on in the graduate school. Ideally the system always selects the forwarding that provides an optimal solution for the current situation. Denser areas should be covered with ad-hoc routing protocols, remaining fixed infrastructure might be included using gateways, while sparsely populated areas stay in the domain of DTN protocols. All these possibilities shall be transparent to the application which only uses a message-based, non-interactive communication interface.

Bibliography

[LOC08] Pei Zhang, Margaret Martonosi. *LOCALE: Collaborative Localization Estimation for Sparse Mobile Sensor Networks*. IEEE Computer Society, 2008.

[PRO03] Anders Lindgren, Avri Doria, Olov Schelén. *Probabilistic Routing in Intermittently Connected Networks*. SIGMOBILE Mobile Computing and Communication Review, 2003.

[MOB05] Jeremie Leguay, Timur Friedman, Vania Conan. *DTN Routing in a Mobility Pattern Space*. Proceedings of the 2005 ACM SIGCOMM workshop on delay-tolerant networking, 2005.

11.5 Cognitive Control of a FPGA-based RF Interface for Cognitive Radio in Disaster Scenarios

Michael Grimm (michael.grimm@tu-ilmenau.de)
Supervisor: Prof. Dr.-Ing. habil. Reiner S. Thomä

Until now, there are predominantly static mobile communication systems in tightly adjacent frequency ranges, which have a low spectrum utilisation in the spectral, temporal or spatial domain. Cognitive Radio (CR) can enhance the spectral utilisation a lot and provides advanced flexibility and services for a secondary use of the radio resources. Therefore, a software-defined RF front-end (SDR) with cognitive and adaptive properties is necessary, which can adapt its parameters without human intervention. In the case of disaster scenarios - the main application of this research program - this reconfigurability is helpful to rebuild partly destroyed links or to establish new ones. The key functions are composed of two parts: the Spectrum Sensing and the intelligent access to the medium. However, this reconfigurability is attendant by some challenges. First, the secondary use has to be done in cooperation with the primary users, i.e. secondary users must not interfere with them. Next, there are some special hardware requirements for the front-end, e.g. a low power consumption, a wideband frequency-agility, a low-noise amplification and reception with a high dynamic range as well as the handling of the resulting deteriorating effects.

The thesis deals with digital signal processing, Dirty RF and sensing issues. The focus of research is on the study of the front-end impairments and their mitigation by signal processing as well as on interference mitigation of primary users, i.e. to suppress strong jammers of licensed users. Opportunities and constraints as well as metrics for the front-end will be investigated. Furthermore, conceptual studies for ad-hoc connections and the extension to a CR network will be done. The main goal is a verification of the algorithms by implementation in real hardware.

Basic features of a single CR node will be demonstrated on the commercial available FPGA-based RF front-end, the Universal Software Radio Peripheral (USRP2). This thesis is an important preliminary work for the development of a first Cognitive Radio demonstrator within the Graduate School and combines various related doctoral studies. There are direct connections to the reconfigurable antennas (Noman Murtaza), Over-the-air testing methods (Alexander Krah), data link layer/MAC protocol (Saleh Hussin) and the methods for Spectrum Sensing and Collaborative Sensing (Rami Nicolas).

11.6 Electromagnetic Wave Field Synthesis for an Over-the-Air Test of Radios in the Graduate School

Alexander Krah (alexander.krah@tu-ilmenau.de)
Supervisor: Prof. Dr.-Ing. Albert Heuberger, Prof. Dr.-Ing. habil. Reiner S. Thomä

After a disaster scenario the availability of radio resources may be strongly constricted and it's beneficial to have a high flexibility of the communication system. The best prerequisite to overcome the situation possesses a cognitive radio system. It can observes its environment and adapt the transmission technology to the extant radio resources. So it can serve as a support system for example. Focused on the environment observation reliability of the measuring data is important. This can be the spectral occupation, the estimation of the signal direction or the fading behavior. To assure good performance of cognitive devices appropriate tests are required. Most realistic is an over-the-air (OTA) test. It is used to verify the functionality of first prototypes especially the influence of the antenna system. The OTA test takes place in an anechoic chamber. A number of probe antennas are arranged around the device under test. Goal is the development of synthesis algorithms for the emulation of a predefined radio channel using a wave field synthesis. In reality each signal path that arrives at the device is described by the following parameters: power, delay, doppler shift, angle of arrival and polarization. All these parameters for each path must be transformed by the synthesis software into suitable impulse responses that will be mapped onto the probe antennas. The big advantage of such an OTA test is the possibility to test the interaction between the antenna system, the RF front end and the predefined radio channel [1] [2]. So it is possible to verify the spatial recognition of a primary user and the adaptation to this (fade-out or support). Furthermore it will be possible to test the sensing of different frequency bands. For this purpose a wave field synthesis for different primary user channels will be performed, e.g. WLAN and GSM.

Bibliography

[1] P. Kyoesti, J. Kolu, J. Nuutinen, M. Falck, P. Maekikyroe. *OTA Testing for Multiantenna Terminals*. COST 2100 TD(08)670, 2008.

[2] J. Krogerus, P. Maekikyroe, P. Vainikainen. *Towards an applicable OTA test method for multi-antenna terminals*. COST 2100 TD(08)671, 2008.

11.7 Monitor and Control System of Distributed Data Exchange in MANETs for Dissaster Scenarious

Pavlo Krasovsky (pavlo.krasovsky@tu-ilmenau.de)
Supervisor: Prof. Dr. rer. nat. habil. Jochen Seitz

Mobile Ad-Hoc Networks (MANETs) are a very special type of wireless networks where a set of mobile devices with wireless network interfaces could form a temporal network, without any additional infrastructure in decentralized manner. Wireless mobile devices as nodes of storage and information transfer can communicate with each other in an ad-hoc mode building self-organizing networks. These networks can be classified according to topological features and operating principles. Each node in the network communicates only with its immediate neighbors. It is important because it allows free flow of information without any moderation or selective rate control. In order to integrate devices into larger networks, self-management mechanisms for distributed data exchange must be implemented. Such a system should provide different services for data retrieval, recovery and healing as well as system topology, especially in disaster scenarios. The visual representation of internal processes in such networks gives an opportunity to evaluate the work of the system and to realize available resources. Monitor and control allow uniting the resources into the organic whole. Combining these entities is an easy and comfortable way to manage changing environment with proper human facility. So, the visualization tool could be a great contribution to resource localization, monitoring and control. The new mobile smartphones appear constantly nowadays, hence the interoperability interface between different mobile platforms for data interchange will play a key role for self-organized mobile networks. The main goal of our proposed system is building a monitor and control system of mobile devices with interoperability and connectivity between different mobile nodes that could represent environmental resources as well as system topology. It allows autonomic decentralization with the ability to control the network from any (authorized) network node. The system provides different services such as network discovery mechanism, network topology, current state and resources of mobile nodes etc. The mobile agents, installed on mobile devices, perform target operations. These agents are very simple units because the element of heterogeneity will be increased in future networks and the interactions among them should not become complex. A unit is an entity that represents a distributed mobile service such as: low battery indicator, update data, data recovery, communication interfaces (WLAN, UMTS, etc.) etc. This module architecture and services could be easily extended by third-party developers that could follow to increase the popularity of the system.

11.8 Self-Organized Routing in Cognitive Radio Ad Hoc Networks

Omar Ahmed (omar-hussein.jan@tu-ilmenau.de)
Supervisor: Prof. Dr.-Ing. habil. Andreas Mitschele-Thiel

Insufficient unlicensed spectrum along with unexploited licensed spectrum leads to utilize this spectrum availability which also called spectrum holes or white, cognitive radio has been investigated as an essential solution to opportunist the spectrum holes. The essential point is to provide routing in the multi hop connections with cognitive radio technology. Currently, Routing in Cognitive Radio Ad Hoc Networks (CRAHN) is the beggist chanllenge. In CRAHN, link availability depend on the PUs activity. This leads to high dynamic link availability in the network. Thus, it is not feasible to apply classical Ad Hoc routing protocol because can simply leads to fluctuate connection. The best solution is integrate the routing mechanism with spectrum availability.The path selection should lies on information pass up through MAC/PHY layers to Network Layer. In addition, route mechanism must flawlessly be completed without interfering PU activiry. We proposed a spectrum-aware on-demand Ad-Hoc routing protocol, which modify traditional on-demand routing to meet the challenges requirement of CRAHN. [1, 2, 3].

Bibliography

[1] Ian F. Akyildiz, Won-Yeol Lee and Kaushik R. Chowdhury. *CRAHNs: Cognitive radio ad hoc*

[2] Hicham Khalife, Satyajeet Ahuja, Naceur Malouch and Marwan Krunz. *Probabilistic Path Selection in Opportunistic Cognitive Radio Networks*.Proc. IEEE GLOBECOM, 2008.

[3] Geng Cheng, Wei Liu, Yunzhao Li and Wenqing Cheng. *Joint On-Demand Routing and Spectrum Assignment in Cognitive Radio Networks*. Proc. IEEE ICC, 2007.

11.9 Robust, Decentralized Information Management

Liz Ribe-Baumann (liz.ribe-baumann@tu-ilmenau.de)
Supervisor: Prof. Kai-Uwe Sattler

During the response and recovery phases after a disaster, various working levels produce and require an immense amount of information. In order to ensure efficient disaster management, this data must be reliably available to a quickly growing group of users, despite possible breaks in network links or node failures due to, for example, power loss or mobile devices that have moved out of broadcast range. Increasingly, large amounts of data are being stored in a distributed manner over wide area networks using distributed hash tables (DHTs), which essentially map data to network nodes and provide routing algorithms for lookups. A DHT's structure provides users with guarantees on data availability while its randomness provides the scalability and robustness that are so important during disaster management. The aim of this work is to develop a highly robust, scalable DHT protocol that integrates node location and additional node information (such as node power availability or reliability) in order to obtain improved data availability and resource allocation, with a central interest placed on the balance between network structure and randomness.

Typically, DHTs fail to incorporate any information about either nodes' locations or resource availability, both of which have heightened importance in disaster scenarios where nodes are often mobile with limited power and bandwidth availability. Node location is ignored twofold in standard DHTs: Firstly, messages are routed on roundabout physical paths - a message originating in Berlin and headed for Paris may hop from continent to continent before returning to its destination - detrimental behavior for a heavily loaded network with nodes of limited resources. Secondly, data is stored in a DHT independent of the actual physical locations at which that data is most needed, so that network failures, partitioning, or overloaded links can lead to inaccessible data. Meanwhile, the lack of attention to resource availability can lead to unnecessary node failures and loss of connectivity: Since routing in DHTs is oblivious to the resource availability of the nodes through which a message passes, nodes with restricted resource availability are burdened the same as nodes with unrestricted resource availability. Thus, steps to increase the robustness and data availability of a DHT for disaster scenarios include: keeping communication as local as possible; placing data physically close to those nodes which most need that data; and conserving resources by sparing nodes with restricted resources from unnecessary activity.

Unfortunately, the integration of location information often comes at the cost of robustness and scalability by reducing the randomness that was intentionally integrated into the design of DHTs: the random selection of

identifiers ensures scalability with the load distributed uniformly throughout the network and increases robustness and peer autonomy while random links - such as those used in Symphony [MBR03] - reduce the expected routing length in the overlay network. The balance of structure and randomness, although not fully understood, appears to play an important role in a network's routing capabilities, as demonstrated by Kleinberg on a set of small-world graphs [Kle00].

In this work, the integration of location and resource awareness will be considered in conjunction with effects of randomness and structure using three main approaches:

- **Location and resource aware network overlay design:** Combining a small-world network in the physical space with a location/resource-optimized DHT could yield benefits of both approaches - the high robustness and short paths of small-world networks along with the scalability and effective routing of DHTs - and even more through heightened robustness and data availability due to location and resource awareness.

- **Location aware replication:** By placing data replication physically near to where they are needed and in a manner that anticipates the dynamic changes in the network, data availability would likely increase should nodes fail or partitioning occur.

- **Gossip protocols for network maintenance** Gossiping geared at network maintenance, resource allocation, and location-aware replication support is integral to the success of a highly dynamic network.

Solutions in these three areas, as well as the development of measures with which to compare the location awareness, resource conservation, scalability, robustness, and routing efficiency of various distributed data management systems, will hopefully lead to a better understanding of how the integration of node information influences fundamental network characteristics.

Bibliography

[MBR03] G.S.Manku, M.Bawa, and P.Raghavan. *Symphony: Distributed hashing in a small world.* In Proceedings of the 4th USENIX Symposium on Internet Technologies and Systems, pages 127-140. 2003.

[Kle00] J. Kleinberg. *The small-world phenomenon: An algorithm perspective.* In STOC'00: Proceedings of the thirty-second annual symposium on theory of computing, pages 163–170. 2000.

11.10 Opportunistic Spectrum Access in Cognitive Radio Ad Hoc Networks

Saleh Hussin (saleh.hussin@tu-ilmenau.de)
Supervisor: Andreas Mitschele-Thiel / Reiner Thoma

There are more than 70% of the licensed spectrum bands unused. So, cognitive radios (CRs) have emerged as the solution to the problem of spectrum scarcity. In disaster scenario many of base stations will be no longer operational, may be they are just lacking power or an operational backhaul. The operational station will have a lot of communication through it, and then these stations will be congested. However, there will be a lot of licensed spectrum band unused. If the wireless devices have CR ability, it will be easy to communicate without infrastructure network. So, it can use all spectrum bands without or with less congestion.

Most related problems of CR associated with physical layer and Media Access Control (MAC) layer [1]. CR MAC protocols deal with access to the available channel in licensed band and select the best channel. There are a lot of problems such as: the multichannel hidden terminal problem, heterogeneity of channels, using Common control channel (CCC), etc [1]. Current approaches try to solve these problems by different access mechanism, but a new problem is generated which is synchronization between CR users [2]. My work will avoid these problems by designing a new protocol that using random access mechanism, this protocol is based on Carrier sense multiple access with collision avoidance (CSMA-CA), it do not need time synchronization. Also, it will avoid the CCC problem.

My access mechanism relies on: information about available free channels that will provide by spectrum sensing, information about the channel history that will provide by the information management system, after choosing the access channel, it required to adapt the antenna power to transmit data. For multi hop routing, it will provide information about the common communication channels between hopes.

Bibliography

[1] I. F. Akyildiz, W. Y. Lee and K. R. Chowdhury. *CRAHNs: Cognitive Radio Ad Hoc Networks*. Ad Hoc Networks Journal (Elsevier), 2009.

[2] C. Cormio and K. R. Chowdhury. *A survey on MAC protocols for cognitive radio networks*. Ad Hoc Networks Journal (Elsevier), 2009.

11.11 Parameters for Distributed MIMO Systems

Dominik Schulz (dominik.schulz@tu-ilmenau.de)
Supervisor: Prof. Dr. Martin Haardt

Current wireless communication systems rely on a well deployed network of sites in order to serve every user appropriately. The proposed doctoral thesis focuses on a large area scenario where a lot of those sites are damaged as a result of a disaster. Only a few intact base stations are left. Hence it is unlikely being served by one of them. In order to make use of the remaining base stations and to restore a certain amount of coverage it is proposed to make use of virtual antenna arrays consisting of several relay nodes.

Today's mobile devices are mainly not capable of exploiting MIMO gains. Thus virtual antenna arrays may be utilized to transmit over long distances and in rich scattering environments. The proposed thesis focuses on non-line-of-sight transmissions where virtual antenna arrays benefit from multipath propagation. For the upper layers this increases performance concerning throughput and bit error rate.

Within our work we consider a transmit terminal, a receive terminal, and several relay nodes each being equipped with on antenna. Distributed communication schemes for such configurations are mostly divided into three phases: Within the first phase the transmit terminal transmits the data symbols to a group of relay nodes. This group is called transmit cluster. During the second phase the transmit cluster will send the symbols to the receive cluster. The receive cluster is a group of relay nodes too. Within the final phase the receive cluster will forward the data to the receive terminal.

All of those three phases give rise to several problems like distributed transmission, distributed channel estimation and distributed decoding and encoding. As we especially concern with disaster scenarios we consider robustness against node failures as a vital point that has to be investigated.

This doctoral thesis is being done as part of the *International Graduate School on Mobile Communications* which is funded by the *Deutsche Forschungsgemeinschaft* (DFG).

11.12 Self-organized Mission Planning and Control for UAV Communication Platforms

Tobias Simon (tobias.simon@tu-ilmenau.de)
Supervisor: Prof. Dr. Andreas Mitschele-Thiel

In major disaster scenarios like earthquakes or floods, huge parts of the existing terrestrial network infrastructure are functionally degraded or destroyed. Since search and rescue management relies on a working communication infrastructure, a fast and robust network reconstruction is essential. Adding mobile network resources in form of UAVs to the damaged network is a promising approach for repairing and augmenting such a "disaster network".

Since central planning algorithms for mission generating and assignment, as well as for control are less adaptive, robust and scalable, other forms of fleet management need to be taken into account. Self-organization techniques and similar meta-heuristics offer valuable properties and may be used for the mission planning, the mission assignment and the control algorithms of the UAV swarm. Since the network reconstruction by UAV placement requires information about the existing network infrastructure, special flight formations need to be evolved.

Another aspect of this work is related to cognitive robotics, meaning how an UAV is able to avoid obstacles and other "no-flight-areas". Behavioristic robot mission generating, coupled with on-line simulation will be used to respect resource constraints like the limited battery lifetime. Avoiding complex communication and relying on rather simple inter-UAV-communication like gossiping is also one goal towards a robust multi-UAV mission planning and control system.

For evaluation, a system simulator needs to be established. Once having a reliable simulated system, the proposed algorithms will be implemented in the demonstrator platform, which is based on a "quadrocopter".

11.13 Self-Organized Network Optimization via Placement of Additional Nodes

Mikhail Tarasov (mikhail.tarasov@tu-ilmenau.de)
Supervisor: Univ.-Prof. Dr. rer. nat. habil. Jochen Seitz

In our time, mobile communication technologies play an important role practically in all spheres of life and technology. The development of communication facilities provides functionality for a multitude of branches: from the entertainment and communications industry to the management of complex transport systems. Sometimes, health and even the lives of people depend on the correct functioning of a communication system, such as for example the mobile telephony network. Reliable work of a communication infrastructure provides chances to rescue many people in zones of natural disasters, such as flooding, hurricanes, tsunamis and earthquakes.

In extreme situations, the most vulnerable part of communication systems consists of elements of the terrestrial infrastructure, such as base stations and wireless access points. The optimal decision in this situation is the creation of hybrid networks consisting of both infrastructure and ad-hoc networks, manageable with a minimum number of nodes in the terrestrial infrastructure. But for such a network, at least one gateway is necessary to connect to an external network (for example, the Internet). Thus, mobile access points based on quadrocopters, for example, can provide the functionality of these gateway nodes. In this case, quadrocopters flying over the network coverage zone provide the substitute for terrestrial base stations.

Such a structural organization of the network leads to the occurrence of a variety of problems, for which it is necessary to find effective solutions. The first of all, the network structure is previously unknown and changes quickly. Because of that, it is necessary to use dynamic routing. Then, the restrictions on the power consumption of mobile nodes and base stations must be considered.

Within the scope of this work, the criteria to decide where to place additional nodes must be defined with respect to the limited resources in the network. Thereafter, the optimal position for additional nodes, for example the quadrocopters, must be determined in a 3 dimensional space with obstacles. The proper communication must be established between the new nodes and the existent network. At the computed position, further optimization in the local environment might be necessary with respect to the current network situation, requested quality of service (QoS) and priorities of ongoing communication associations.

To reach these goals, a set of distributed algorithms must be developed, implemented and tested on the target quadrocopter hardware platform.

11.14 Car-to-X Communications

Tatsiana Ulazniuk (tatsiana.ulazniuk@tu-ilmenau.de)
Supervisor: Prof. Reiner S. Thomä

The challenging topic of self-organization in communication networks is addressed to the area of vehicular ad-hoc networks. Car-to-car or car-to-x communication poses many challenges on the wireless transmission and network protocols in the last years. Vehicles are equipped with communication devices and aim to be nodes in an ad-hoc communication network. Such vehicle terminals send, receive and/or relay an information to other network terminals, and can be, therefore, very useful in disaster situations and provide self-organization functionality of the communication network. Evaluation of the system performance and cooperation efficiency depends on radio channel of each communication link. The goal of this work is the research on so-called large scale statistics of radio environment.

WINNER model is a geometry-based stochastic channel modeling approach which will be considered in this work. It enables link level as well as system level simulations. Large scale parameters like shadow fading, delay and angular spreads are considered as an average over a distance of some number of wavelengths and play an important role by radio system engineering. In addition, possibility to predict channel fading behavior using measurements of past and present behavior could be an advantage by system level design. Large scale statistics as well as correlations between channels for car-to-x communications are not sufficiently considered in WINNER model. Moreover, other channel modeling approaches and research in area of vehicular ad-hoc networks mostly consider small scale parameters. Therefore, an extension of WINNER model for car-to-car communications and derivation of large scale parameters and correlation properties will be solved by means of measurement campaign and further processing of measured data as well as computer simulations.

The topic of car-to-car communications is strongly related to the topics of Graduate School which deal with distributed MIMO systems, since the goal of the research is to provide relation between large scale parameters and performance of distributed MIMO systems in car-to-x scenario. These are "A novel approach to network coding and control of distributed and heterogeneous MIMO" (Bilal Zafar), "Virtual antenna arrays in large disaster areas" (Dominik Schulz). The topic "Self-organization in future radio access networks" (Muhammad Naseer ul Islam) considers LTE standard and requires the influence of radio channel on system performance. Here, cooperation in sense of channel modeling is also provided.

11.15 Self-organization Data Harvesting and Fusion for Building a Resource Map

Jin Yang (jin.yang@tu-ilmenau.de)
Supervisor: Prof.Dr.-Ing.habil.Kai-Uwe Sattler

There will be some resources not available in a mobile communication system after disaster occurs. The damages caused by the disaster will reduce the availability of resources and lead to dramatic changes to the network topology, either forcing part of the devices function abnormally or causing some devices quit entirely from the system. Lacking knowledge about current status of available resources after a disaster becomes a main obstacle for rescuers or system administrators to react effectively and efficiently. Building a resource map using self-organization techniques, especially data harvesting and data fusion, can increase the chances of making correct decisions to recover the system. Structure[1]and structure-free techniques [2] techniques are used to aggregate data in sensor networks.These techniques are well suited for applications such as fusion with statistic index (average, maximum, minimum) on single data (temperature, power index, humidity). However, several special requirements result from the generation of "radio resource map" in disaster scenarios which can not be easily met using these aggregation protocols. First, the fusion process depends on the interplay of four variables among nearby nodes according to radio propagations, which increases the difficulty of the fusing process. Second, the data harvesting strategy should be robust enough to suffer the disaster scenario itself, which necessitates a self-organizing design. This work explores the potential of gossip-based data harvesting and data fusion framework. The problem considered is a multi-source single-sink problem, where the data collection and fusion are tightly coupled with query processing in wireless communications systems. This work belongs to information management and has close relation with the other topics in the Graduate School of MobiCom.

Bibliography

[1] W.Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Micorsensor Networks", in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, vol.2, Jan. 2000.

[2] Kai-Wei Fan, Sha Liu and Prasun Sinha, "Stucture-free Data Aggregation in Sensor Networks", in *IEEE Transactions on Mobile Computing*, Vol.6, August 2007, pp.929-942.

11.16 Control of Distributed Heterogenous MIMO

Bilal Zafar (bilal.zafar@tu-ilmenau.de)
Supervisor: Martin Haardt

MIMO (Multiple Input Multiple Output) systems are the most promising option for system designers aspiring to provide very high data rates to users while maintaining a high quality of service (QOS) as well as keeping the system to low levels of complexity and required power. The reason is that MIMO systems can provide Spatial multiplexing gain as well as Diversity gain. However, MIMO requires multiple antennas at the transmitter and the receiver which is difficult to implement in practice. Hence we have to look at other options that can display the MIMO benefits while being practically implementable with relative ease. Co-operative MIMO is the most innovative option in this regard. The main idea of co-operative MIMO is the recruitment of nearby idle nodes to assist transmitting and receiving data thus transmitting nodes use idle nodes as relays to reduce multi-path fading effect in wireless channels. The relays can be used in either "Amplify and forward" or "Decode and forward" or "Coded Cooperation" configuration. Hence we can create a "Virtual MIMO" system in which we can achieve MIMO gains even with one antenna per-node, e.g. in open-spectrum meshed/ad-hoc networks, sensor networks, etc. We will investigate a cluster-based distributed MIMO system. The scheme is especially suitable for disaster scenarios since multi-hop MIMO has the ability to communicate over long ranges even in the absence of some infrastructure components and its distributed nature renders it suitable for self-organization as well. We will focus mainly on "Cooperative MIMO system with STBC and code combining". Two major challenges await us in the design of such a system:

- Clustering mechanism to support self organization in Disaster scenarios
- MIMO communication strategies

12 International Graduate School: Dynamic Intelligent Systems

Prof. Dr. Wilhelm Schäfer
Email: wilhelm@upb.de
Universität Paderborn
Internet: `http://pace.upb.de`

The International Graduate School Dynamic Intelligent Systems at the University of Paderborn is one of seven NRW Graduate Schools which are initiated and supported by the Ministry of Innovation, Science, Research and Technology of the federal state North-Rhine-Westphalia, Germany. Since Oct. 09, the Graduate School is funded by the program "NRW-Forschungs-schulen" of the state of North-Rhine-Westphalia. The Graduate School is run by the faculties of Electrical Engineering, Computer Science and Mathematics (EIM), Business Administration and Economics and Mechanical Engineering.

The overall goal of the School is to provide a stimulating environment for research in the rapidly developing field of software, systems and network engineering. The field deals with the construction of so-called embedded or mechatronic systems. Such systems are already in daily use and the social and economic welfare of the modern information society increasingly depends on them. Examples are simple systems like electronic teller machines and complex life critical systems like a modern aircraft or a railway system. Such systems only function correctly if the often complex interplay between mechanical, electrical and software components is carefully defined and works well. Complexity is further increased when the systems are connected in networks. Challenges for future research in the area are particularly due to the fact that these systems must increasingly become dynamically reconfigurable, such that their components (software and hardware) can be replaced or upgraded during runtime. Such changes can only be contemplated if their success is assured in advance. Not only must the new component work properly, but it also must operate in its new environment without causing any bad side effects. In addition, the evolution of such systems and the implant of components into different, evolving environments require that components become more and more intelligent, i. e., that they adapt automatically to their new or changing environment. Such changing, adaptive systems we call dynamic intelligent systems. They are the focus of the School's research program.

12.1 Developing Cognitive Functions in Self-Optimizing Systems with Solution Patterns

Roman Dumitrescu (Roman.Dumitrescu@hni.uni-paderborn.de)
Supervisor: Prof. Dr.-Ing. Juergen Gausemeier

The increasing penetration of mechanical engineering by information technology enables considerable benefits aptly expressed by the term mechatronics, which means the close interaction of mechanics, electronics, control engineering and software engineering. The progressive integration of information technology will enable mechatronic systems with partial intelligence. We refer to such systems as self-optimizing systems. S.o. systems have the ability to react autonomously and flexibly on changing operation conditions. Therefore such systems have to perform system functions such as "to percept the current situation" or "to determine objectives". Those functions come with the territory of cognitive systems and are known as cognitive functions. Cognition can be characterized as the ability that enables not only autonomous and adapting, but also more reliable, effective and viable systems regarding their purpose. My PhD work addresses the development of s.o. systems, which is an interdisciplinary task. Additionally to mechanical, electrical, control and software engineers experts from mathematical optimization and artificial intelligence are involved. Thus, the domain-independent storage and retrieval of domain-specific expertise matter enormously. This approach focuses on the reuse of once proven solutions in form of patterns. Such patterns comprise not only the principle solution for developing certain system functions, but also the constraints for their use. We distinguish two kinds of solution patterns: working principles, which are based on physical effects and active pattern, which relay on methods of information processing. Whereas working principles are well establishes, only abstract compilations for active pattern exist. Therefore I developed a new kind of solution pattern, the active pattern for self-optimization (APSO), which describes methods from the field of artificial intelligence and supports the development of cognitive functions in s.o. systems. A few aspects of APSO are introduced briefly: The aspect "functions" covers those system functions that can be realized by implementing a specific APSO. The functions are arranged in a hierarchy to support the classical design methodology approaches which start bottom up. The aspect "methods" details the combination of methods realizing the specified functions. The aspect "structure" describes which elements are needed to implement the methods and their interrelations. The aspect "behavior" specifies the performance of single elements or even of groups of elements including their states and activities during operation mode. For the validation of the new solution patterns, several APSO were documented and a software support was developed.

12.2 A Framework for Change Management of Business Process Models

Christian Gerth (gerth@uni-paderborn.de)
Supervisor: Prof. Dr. Gregor Engels

In model-driven software development approaches, business process models are used at different levels in the development process. For instance, in Service-Oriented Architectures (SOA) high-level business process models become input for the development of IT systems and in running IT systems executable process models describe choreographies of Web Services. A key driver behind this development is the necessity for a closer alignment of business and IT requirements, in order to reduce the reaction times in software development on frequent changes in competitive markets.

Similar to other software artifacts, business process models underlie constant change, i.e., they are created and refined by different business modelers and software architects in distributed environments. This results in different versions reflecting the different views of the involved stakeholders. At some point in time, different versions of a process model have to be compared and merged with each other to obtain an integrated version. This task is called change management of business process models.

Change Management of business process models can be compared to widely used concurrent versioning systems (CVS) for textual documents. However, in contrast to textual documents, that are compared syntactically line by line, the comparison and merging of business process models must consider the graph-like structure of process models and can be improved by considering the semantics of process models.

In my thesis, I develop a framework for change management of business process models that solves the following problems: First, in typical business process modeling scenarios, no change log is available that records changes applied to different model versions. The reasons for this are the use of different modeling tools by different stakeholders and the distributed environment of large software development projects. As a consequence, different versions of a process model need to be compared to identify differences between the versions before they can be merged. For each difference detected, appropriate change operations have to be derived which together can be considered as a reconstructed change log. These change operations must be close to the intended meaning of the change to be understandable by a business user. I approach this problem by decomposing process models into so called single-entry-single-exit (SESE) fragments, which are non empty subgraphs of a process model with a single entry and a single exit edge. These fragments improve the comparison of different process model versions and help to group differences into intuitively understandable compound change operations [1].

Second, to enable a high degree of automation within integration of different process model versions, it is important to understand dependencies and conflicts of changes. Informally, if two changes are dependent, then the second one requires the application of the first one. If two changes are in conflict, then only one of the two can be applied. As a consequence, an approach for computing dependent and conflicting compound changes is required. To address this issue, I capture our set of compound operations in terms of model transformations and apply a critical pair analysis to identify dependent and conflicting transformations [2].

The third problem that I approach in my thesis arises from the fact that change management is a modeling language-dependent problem, i.e., a solution for a particular modeling language cannot be reused easily for another language, due to different syntax and semantics of the languages. To solve this problem, I investigate how a change management solution for a specific business process modeling language can be generalized in terms of a framework for change management of process models [3]. The framework contains an intermediate representation for process models that serves as a common denominator for different process models. Based on the intermediate representation, differences, dependencies, and conflicts are computed. This way, the framework can be instantiated for change management of different process modeling languages.

There are some issues that are currently not covered by my approach to change management of business process models which have to be addressed in future work. For instance, the framework does not support change management across modeling language boundaries, yet. In those scenarios, a (partial) mapping between the meta models of different modeling languages is required.

Bibliography

[1] J. M. Küster, C. Gerth, A. Förster, and G. Engels. *Detecting and Resolving Process Model Differences in the Absence of a Change Log*. In M. Dumas and M. Reichert, *BPM'08*, volume 5240 of *LNCS*, pages 244–260. Springer, 2008.

[2] J. M. Küster, C. Gerth, and G. Engels. *Dependent and Conflicting Change Operations of Process Models*. In R. Paige, A. Hartman, and A. Rensink, *ECMDA-FA'09*, volume 5562 of *LNCS*, pages 158–173. Springer, 2009.

[3] C. Gerth, J. M. Küster, and G. Engels. *Language-Independent Change Management of Process Models*. In A. Schürr and B. Selic, *MODELS'09*, volume 5795 of *LNCS*, pages 152–166. Springer, 2009.

12.3 Knowledge-Based Agent Societies

Thomas Kemmerich (kemmerich@upb.de)
Supervisor: Prof. Dr. Hans Kleine Büning

In Multiagent Systems multiple autonomous entities, known as *agents*, interact in an environment to solve a given problem in a decentralized way. Such systems raise different kinds of questions, including areas like coordination, cooperation, learning, trust, social interactions, and so forth.

Our work deals with coordination issues in Multiagent Systems composed of simple agents with limited abilities and resources. In particular we are interested in the following research questions:

1. How can external storage media that are located in an environment support coordination?

2. How can agents learn to solve a coordination problem?

3. How to coordinate the assembly of autonomous agents around a given location?

To investigate these questions, we consider the following problem that deals with the repeated partitioning of agents to targets in a two-dimensional Euclidean world. The considered environments may contain obstacles and special regions that influence movement, visibility, as well as communication of agents. The agents operate at different positions in the environment by executing position-related jobs. At the same time, all agents determine a target and move towards their selection, i.e. they calculate a partitioning to the target set in a decentralized way. Afterwards, the agents redistribute in the environment choosing new positions if the job at their previous position is done. This leads to dynamically changing configurations that the agents have to solve one after the other.

Each partitioning is rated according to three contrary objectives. The first objective demands a uniform partitioning, i.e. the number of agents assigned to a target should be equal (± 1) for all targets. The second objective asks for minimizing the total distance sum between agents and their selected target. The last objective is to minimize costs produced by the agents according to a cost model. In a sense, the latter objective demands efficient algorithms running on simple agents with as limited abilities as possible. The cost objective is especially interesting if different approaches should be compared.

The contradictions of these objectives become clear when we consider one objective at a time disregarding the other two. For instance, concentrating on the distance objective means that all agents should select their nearest target as this minimizes the distance sum to an optimum. If, however, a

single target is closer to all agents than any other target then all agents would select it which obviously leads to bad distribution results.

In the end, the overall goal of the agents is to coordinate their decisions to maximize the average partitioning quality.

Based on local information and knowledge that agents gain while operating, the idea is to construct learning agent societies that – as a whole – benefit from knowledge of its individuals. Since agents move around, break down, or are replaced by new agents, individual knowledge can get lost or might become useless, for instance if knowledge is position-related. Enabling an entire society to profit from individual knowledge thus means to share knowledge and to make it available for relevant society members at the right location and time. One approach of our work is to store knowledge and information on external media that are located in the environment. We investigate the role of these media as external coordination objects. Dealing with knowledge, questions on how it has to be structured and how it can be used in agent internal reasoning processes arise, too.

Because of the high number of agents in the system a large state and action space has to be considered when learning algorithms are applied. Due to many interacting agents the learning environment becomes non-stationary. Learning to solve the problem thus is hard and convergence to optimal solutions can not always be guaranteed, especially since no central controller is available. Hence, we investigate novel machine learning techniques that involve external media to get insights in the role external resources can play in the learning process.

Besides these questions, another practical part of our work deals with movement issues that occur when many agents move towards a common target position. How could they coordinate based on local information and limited sensor ranges to find compact arrangements and what are the impacts of obstacles?

Answers and mechanisms developed to solve the presented iterative partitioning problem may also be applied to other domains. Examples include load balancing in large scale networks where users have to be partitioned to servers according to some objectives or the more general (iterative) graph partitioning problem.

12.4 Maintaining Connectivity of Autonomous Agents Using Mobile Ad-hoc Robotic Network

Emi Mathews (emi@hni.upb.de)
Supervisor: Prof. Dr. rer. nat. Franz Josef Rammig

Scenarios such as multi-robot exploration of an unknown terrain, urban search and rescue etc. usually have several exploring autonomous agents and one or more base station(s). For collaborated tasks in such situations require communication between the base station and agents. These scenarios lack infrastructure for network connectivity. A stable, high bandwidth communication link that surpasses the obstacles is also necessary. We propose the formation of a Mobile Ad-hoc Robotic Network (MARNET), with the help of cheap mobile router nodes, to connect the agents and the base station(s). A Worker-Router strategy is designed for the router nodes, so that apart from being a router, it assists the agents in their mission and provides services too. The agents move autonomously as well as drop the router robots at particular locations. The router robots move in a limited fashion within their comfort zone and try to relocate in case of link failures. We propose a solution, that could maximize the coverage area, maintaining the connectivity of agents and base station(s) and minimize the number of router robots. The node densities of Hexagonal grids and 2D r-strip structures are found to be close to the optimal density [1] for the area coverage problem. We have developed a localized algorithm for the movement of router nodes that could form such grids and thus obtain the near-optimal Coverage-Connectivity. We have also designed a new routing algorithm for the MARNET, which is a hybrid of Geographic and Topology-Based routing. This algorithm could cope with the inefficiency of geographic routing due to localization errors by forming hierarchal layers using the topology information. The cluster heads, formed from multi-hop clustering, keeps information about their one-hop neighbors only. The empirical study shows that the overlay graph of the cluster heads is mostly planar and hence the Geographic routing algorithms with planar void handling could be used in these graphs.

Bibliography

[1] R. Iyengar, K. Kar, S. Mukherjee *Low-coordination topologies for redundancy in sensor networks*. Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing, May 2005.

13 Research College: Service-Oriented Systems Engineering

Prof. Dr. Andreas Polze
Email: andreas.polze@hpi.uni-potsdam.de
Hasso Plattner Institute at University of Potsdam
Internet: `http://kolleg.hpi.uni-potsdam.de`

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.

Commonly used technologies, such as J2EE and .NET, form de facto standards for the realization of complex distributed systems. Evolution of component systems has lead to web services and service-based architectures. This has been manifested in a multitude of industry standards. All these achievements lead to the paradigm of Service-Oriented Systems Engineering which 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 28 PhD students and one postdoc.

In context of the research college, the HPI works on the following topics:

- Service-Oriented Geovisualization Systems
- Modeling and Verification of Self-Adaptive Service-Oriented Systems
- Tools and Methods for Software Engineering in Service-Oriented Systems
- Security Engineering of Service-Based IT Systems
- Formal Semantics of Service-Oriented Systems
- Evolutionary Transition of Enterprise Applications to Service-Orientation
- Operating System Abstractions for Service-Oriented Computing
- Services Specification, Composition, and Enactment
- Human Computer Interaction for Service-Oriented Systems

13.1 Information Integration in Service-oriented Computing

Mohammed AbuJarour (mohammed.abujarour@hpi.uni-potsdam.de)
Supervisor: Felix Naumann

Information integration has been the typical approach to data-driven applications in several domains, such as enterprise applications or news portals. Applying information integration techniques in Service-oriented Computing (SOC) is not straightforward, because of the lack of adequate description artifacts about the considered services.

We propose a novel approach and platform to alleviate this problem and investigate the benefits of information integration in Service-oriented Computing.

The increasing number of Web Services and the easiness of creating Web Services from existing software applications have been increasing the complexity of SOC systems. This increasing complexity has been making service discovery and selection, evaluating service quality, and providing fundamental and value-added features more and more challenging. The main reason for these limitations is the lack of enough information and descriptions associated with these services.

The key challenges, which drive our research include:

- Increasing complexity of Web Services and business needs.
- Dynamic SOA and business environments.
- Different techniques and notions used to describe and provide services.
- Inadequate information for service discovery and selection.

In our research, we introduce a new information integration environment for SOC applications, where we integrate information about services from different sources to provide the required features and value-added features. Our research statement is summarized in this question: How to enrich, integrate, and manage service descriptions efficiently and what are the consequences of enriching service descriptions in SOC?

All parties involved in a SOC application are taken into consideration and information from these parities is gathered. This information includes *data* from service providers, e.g., WSDL, *metadata*, e.g., category, *community annotations and consumers feedback, invocations metadata*, and *usage history*. All these different types of information are then used to create a unified service description for each Web Service using non-traditional information integration techniques.

13.2 Context-aware Reputation Framework

Rehab Alnemr (rehab.alnemr@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Christoph Meinel

Open systems such as service oriented architectures (SOA), internet of services(IoS), and cloud environment demand a continuous attention to trust relationships between their participants. One way to ensure trust is to use reputation concepts. Existing work on reputation systems focuses on improving the calculation of reputation values, preventing malicious actions, and the deployment into the business world where reputation is mostly represented in a singular value form. The main goal of this research is to model a correct representation for reputation, reflecting its real-world notion. The final goal is to facilitate reputation portability between different platforms. The proposed framework present a new reputation form that holds detailed information about an entity's reputation; *Reputation Object*.

A Reputation Object (RO) holds a summary of the behavior of an entity in several contexts in the form of *list of contexts* and their corresponding *List of rated values*. For example, a seller in an e-market place will have an object with contexts *"product-quality", "delivery", "customer-services"* with a rating for each context. A *context* can be *rating-criteria*, *quality-attribute*, or any *concept/ontology*. The final form is stored in an RDF format to be easily transfered and interpreted between different systems. This way several goals are achieved:

- entity's reputation is more meaningful because it is associated with the context it was earned in. Each organization can have different set of ratings' contexts but it will still be easy to understand the meaning of the reputation. Also, one can easily extend these criteria.
- automizing reputation transfer
- helping customers place requests according to their customized needs, in e-Markets.

The presented framework is described in [Article 1] with some use cases for using the reputation object (i.e. service selection, reputation distribution in agent-based systems, and breach management for service level agreements).

Bibliography

[Article 1] Rehab Alnemr, Justus Bros, & Christoph Meinel. *Constructing a Context-aware Service-Oriented Reputation Model using Attention Allocation Points*. Proceedings of the IEEE International Conference on Service Computing, 2009.

13.3 Context-oriented Service Computing

Malte Appeltauer (malte.appeltauer@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Robert Hirschfeld

The Software Architecture Group, led by Prof. Dr. Robert Hirschfeld, is concerned with fundamental elements and structures of software. Methods and tools are developed for improving the comprehension and design of large complex systems.

Software systems often require context information for user-specific computation. However, state-of-the-art approaches do not represent context-dependent behavior with first class entities, which hinders separation of concerns and software evolution. *Context-oriented programming* (COP) explicitly represents context-specific concerns and their dynamic composition. COP allows for the definition of *layers* that can crosscut object-oriented decomposition and encapsulate behavioral variations. Depending on the execution context, layers are composed into a system at run-time.

In the course of this thesis, the Java language extension *ContextJ* has been developed. The COP languages implemented so far, including ContextJ, support selective activation and deactivation of layer compositions, expressing programmatically when the application enters and leaves certain contexts. It is, however, not enough to regard context as being entirely under programmer control; instead, context can impose itself on the running application "from the outside". Based on this observation, we distinguish control-flow specific from event-specific contexts. Two key properties serve to characterize them: (1) Event-based context can overlap several control flows, unlike control-flow specific context that is confined to a single control flow. (2) Event-based context entry and exit often cannot be localized at specific points in the control flow. Instead, context entry depends on asynchronous events independent from the main control flow. Moreover, a certain context is often valid until another event changes the composition. To meet these requirements, we develop *JCop* based on our experiences with ContextJ. For a better separation of layer composition from application logic, JCop provides declarative composition statements.

In service-based systems, context awareness gains increasing relevance for mobile, personalized applications. Therefore, we conduct several case studies in which we apply JCop to the development of service-based applications. We analyze the expressiveness of JCop's language abstractions for the representation of static and dynamic variability of service interfaces and implementations.

13.4 Modeling and Verification of Self-Adaptive Service-Oriented Systems

Basil Becker (Basil.Becker@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Holger Giese

Self-Adaptive systems behave different from classical software systems. Typically self-adaptive systems are used in situations where the system's environment of operation is changing in an unpredictable manner. In order to be able to react on environmental changes self-adaptive systems often rely on one or more feedback loops as they were well known in control theory. Service-oriented system are adaptive systems or even self-adaptive systems – depending on the definition.

For the modelling of a self-adaptive system the above observations results in some serious issues. The degree of uncertainty is much higher than it is for a classical software system. Hence, it has to be figured out in which way the modelling techniques which software engineers currently are used to can be used to model self-adaptive systems. Especially the system's dynamics have to be expressible within the models. I propose to use attributed and typed graph transformation systems to model self-adaptive systems. Beside the modelling the verification of self-adaptive systems is a major part of my thesis. The problems concerning the modelling of self-adaptive systems hold for their verification, too. It should be obvious that in situations one not even knows the system's goals it is very hard to verify that the systems fulfills all these goals. This results in the following problems that have to be solved: Beside the models also the verification techniques have to be available at runtime, verification techniques must be enhanced to meet runtime restrictions and they have to be adopted to infinite state systems. During the design phase of the system only some properties could be verified. So is it possible to verify for each rule that the rule fulfills certain safety properties. At runtime verification becomes necessary because changing goals may result in changed rules that have not been verified before. Following online verification is one aspect of the verification of self-adaptive systems. Due to the missing possibility of offline verification the verification technique itself has to meet several requirements concerning time consumption.

Most of verification techniques known today are only able to handle systems with an finite state space. Self-adaptive systems' state space often is infinite. So either one has to combine existing verification techniques, systems specifications and modelling techniques to elegantly come around this or verification techniques that are able to handle infinite state spaces have to be used or developed. For graph transformation systems as well as timed graph transformation systems we have developed such techniques.

13.5 Programming Models for Multicore Computers

Frank Feinbube (frank.feinbube@hpi.uni-potsdam.de)
Supervisor: Andreas Polze

A recent trend in microprocessor design is the increase of the number of processing cores on a single processor chip, where the resulting products are often called multicore or manycore processors. This trend originates from the desire to utilize the increased number of transistors which can be accommodated on a single chip, following the prediction of Moore's law. Other strategies for utilizing more transistors, such as pipelined and superscalar execution, have mostly been exhausted, leaving the integration of many computing cores as the major strategy to provide an ongoing increase of computing power. This trend can be seen both in the system central processors as well as in graphics processors.

Todays commercial off-the-shelf computer systems are multicore computing systems as a combination of CPU, graphic processor (GPU) and custom devices. For some years graphic cards were not only used to render pictures to screens, but also for numerical processing. In these applications, shader languages or vendor specific languages like *AMD Brook+*, *AMD Cal* or *Nvidia Cg* were applied. Todays frameworks like *Nvidia CUDA* and *AMD Stream Computing SDK* are based on the *C* programming language with few extensions and have a general purpose nature. The next step will be the application of the emerging *OpenCL* programming framework. It allows to write programs that use either the CPU or GPU as the underlying processing device.

In comparison with CPU cores, graphic cards are capable to execute hundreds up to thousands compute units in parallel. To benefit from these GPU computing resources, applications have to be parallelized and adapted to the target architecture. Our current research focusses on the issues that one encounters when trying to applicate graphic cards for general purpose programming (GPGPU). The current programming models still have a lot of restrictions and one has to know the underlying hardware in detail to get acceptable performance [Feinbube2010].

In addition we are interested in design patterns for multicore programming, especially the ones that are used in GPGPU-programs.

Bibliography

[Feinbube2010] Frank Feinbube. *On Programming Models for Multi-Core Computers* in Proceedings of the 4th Ph.D. Retreat of the HPI Research School on Service-oriented Systems Engineering. University of Potsdam, 2010.

13.6 Mega-modeling the Development of Service-Oriented Enterprise Systems

Regina Hebig (regina.hebig@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Holger Giese

For the development of software various approaches can be applied. In the simplest cases a single programming language is used and after a compilation step the developed software can be applied. In more complex cases the developer has to deal with different building steps and representations of a system. For example, in the Model-Driven Engineering (MDE) approach the system is first given in an abstract representation, which is refined with more and more information during each development step.

Mega-models are special models that can be used to illustrate the relationship between other models, language specifications, tools and meta-models. Thus, mega-models can be used to reason about and document different approaches for developing software. The system that is developed might be represented parallel in multiple languages, which might also change several times during the development. Thus, not only a programming language, but the whole correlation of artifacts that can be captured with the mega-models has strong influence on productivity, efficiency and maintainability.

Especially the development of service oriented architectures (SOA), where the program modules are implicitly coupled via language independent interfaces, leads to complex and diverse mega-models. This diversity and complexity makes it difficult to ensure that development is as productive and efficient as possible. Consequently, empirical studies are required to identify best practices and potential for the improvement of the development of service oriented systems.

It is a goal to define a mega-modeling methodology that allows documenting and planning approaches for the development of service oriented architectures. Further, it is a goal to identify best practices for the development based on mega-models. As a result it might become possible to compare mega-models of different approaches to develop a system for specified circumstances and state, which potentials can better be used with which approach.

13.7 Services for Real-Time Computing

Uwe Hentschel (uwe.hentschel@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Andreas Polze

Service oriented computing became more and more important in the last years. One Example is the increase in popularity of web services for business applications. The principles of service orientation are also interesting for real-time applications. Examples are medical remote monitoring and health care, synchronization and control of different measurement devices within complex measurement and test sequences as well as the cooperation of different control units inside of a car.

One point of interest are the communication networks. In some systems the challenges are the nondeterministic behavior of common network protocols, like TCP (Transmission Control Protocol) and IP (Internet Protocol), as well as the influence of additional network components. In order to analyze the influence of network components a round-trip-delay measurement was made in an IP based network environment [1]. In three different network scenarios the round-trip-delay for UDP (User Datagram Protocol) datagrams was captured as a series of measurements each with a million individual measurements using a simple client-server application and Microsoft Windows as operating system.

The Fontane project [2], a telemedicine project, aims to improve the medical attendance of patients with cardiovascular diseases in the region Nordbrandenburg, Germany. The medical reasons to select this region are the low number of ambulant working cardiologists and the high death rate due to cardiovascular diseases. In the context of this project several case studies are supported as well as new medical devices and a self-adaptive prioritizing middleware are developed. The data capture part and communication part are of special interest. The vital signs of a patient are to be captured and analyzed as well as stored and transmitted to the telemedicine center in normal and emergency cases. One challenge here is the use of a mobile communication network and the resultant limited bandwidth.

Bibliography

[1] Uwe Hentschel. *Messung der Übertragungszeiten in IP-basierten Netzwerken unter Windows*, September 2009.

[2] Friedrich Köhler. *Fontane: Gesundheitsregion Nordbrandenburg: Antragsskizze für die Realisierungsphase im BMBF-Wettbewerb "Gesundheitsregionen der Zukunft – Fortschritt durch Forschung und Innovation"*. BMBF-GRDZ-46-105 "FONTANE-Projekt", April 2009.

13.8 Service-Oriented Visualization of 3D Geovirtual Environments based on Standards and Images

Dieter Hildebrandt (dieter.hildebrandt@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Jürgen Döllner

Visual representations of geospatial information proved to be valuable means to facilitate thinking, understanding, and knowledge construction about human and physical environments, at geographic scales of measurement. Massive amounts of distributed and heterogeneous geospatial information and geospatial computing functionality are increasingly available as distributed resources that can be accessed through the Internet. This increased availability has created the demand and opportunity to build distributed systems that leverage these resources for visualizing and interacting with geospatial information. For the implementation of such distributed systems, the application of the architectural concept *service-oriented architecture* (*SOA*) and standardization proposals by the *Open Geospatial Consortium* (*OGC*) are commonly proposed.

This research focuses on the higher-level research question of how to apply service-oriented principles in the geospatial domain in an effective and value-adding way. More precisely, this research focuses on the question of how to provide, interact with and style 3D geovirtual environments (3DGeoVE) in a service-oriented infrastructure, based on standards and 2D images a the primary representation form.

I propose an approach that allows a human user to explore, interact with, and define the visual appearance of a remote 3DGeoVE through the Internet and by using lightweight clients such as web-based clients and Smartphones. The use of service-orientation and standards facilitates designing a distributed system that is interoperable and can easily be adapted to changing requirements. The image-based provisioning of visual representations allows for providing high-quality visual representations by dedicated services and consuming them on lightweight clients without having to download massive amounts of data. The image-based provisioning of visual representations of 3DGeoVE is implemented by a Web View Service (WVS, OGC discussion paper). Means for interaction and exploration of the 3DGeoVE are provided by lightweight clients that request images as visual representations of the 3DGeoVE from the WVS. The clients employ image-based and point-based modeling and rendering techniques for the local reconstruction of the remote 3DGeoVE from novel viewpoints. Styling of the visual representations of the 3DGeoVE is facilitated by employing the concepts pre-render styling (implemented within the WVS) and post-render styling (implemented in the WVS or provided as a dedicated service).

13.9 Quantitative Modeling and Analysis with FMC-QE

Stephan Kluth (stephan.kluth@hpi.uni-potsdam.de)
Supervisor: Prof. Dr.-Ing. Werner Zorn (em.)

The modeling and evaluation calculus FMC-QE, the Fundamental Modeling Concepts for Quantitative Evaluation, extends the Fundamental Modeling Concepts (FMC) for performance modeling and prediction. In this new methodology, the hierarchical service requests are in the main focus, because they are the origin of every service provisioning process. Similar to physics, a service requests is a tuple of value and unit, which enables hierarchical service request transformations and therefore the hierarchical modeling. The reducing of the model complexity by decomposing the system in different hierarchical views and the distinction between operational and control states reduces the computational complexity by orders of magnitude.

According to FMC, the system performance parameters are described in three arbitrarily fine-grained hierarchical bipartite diagrams. From the structures and parameters of the performance model, a hierarchical set of equations is derived. The calculation of these equations is done on the assumption of stationary processes and is based on fundamental laws of the performance analysis: Little's Law and the Forced Traffic Flow Law.

While the initial ideas of FMC-QE were founded by the supervisor of the PhD student, the planned thesis will further describe and validate the methodology. Furthermore, it extends the methodology for the handling of closed networks through the integration of the summation method in an iterative computation, as well as the handling of multiclass and semaphore synchronization scenarios with the goal of precision and reduction of computational complexity.

Through the background in FMC and FMC-QE, the hypothesis: 'If a system can be represented by a server model without any inter-server control flows, the underlying problem is of type PFQN (Product Form Queueing Network), otherwise of type NPFQN (Non Product Form Queueing Network).'[WZ07] was raised. In the thesis, this hypothesis will also be analyzed and validated with the example of BCMP networks.

Bibliography

[WZ07] Werner Zorn. *FMC-QE - A New Approach in Quantitative Modeling.* In: H.R. Arabnia: International Conference on Modeling, Simulation and Visualization Methods (MSV 2007), CSREA Press, 2007.

13.10 Towards a Service Landscape for a Project Management Dashboard

Thomas Kowark (thomas.kowark@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Hasso Plattner

The continuing trend towards distributed collaboration and the corresponding indirect, digital communication imposes new challenges on project management. Counterproductive developments within the team dynamics are much harder to recognize since continuous on-site observation by project managers is rarely feasible. Accordingly, they need to rely increasingly on indicators provided by digital communication artifacts to guide project teams.

The focus of our work is the creation of a service landscape that enables such process introspection for software development processes by collecting and analyzing digital collaboration artifacts. We are using the d.store [1], a service-oriented platform for capturing team collaboration spaces with resource-based information networks, as the basis for our research. Even though it was not specifically designed to capture common software development process artifacts, such as source-code management system revisions or bug tracker items, its adherence to communication and data representation standards facilitated their simple integration.

In order to test the platform extensions in a realistic scenario, we have created a testbed that allows us to reenact large-scale software development processes within an educational environment. The approximately 80 students that annually participate in this exercise are required to work on a single task (e.g., developing an enterprise resource planing system) in a collaborative effort. The chosen team setup and process model require a substantial amount of inter and intra-team communication and, thereby, allow us to reason about characteristic, problem-indicating patterns within the collected networks of digital collaboration.

Additionally, the feedback that we get while using the tools and examining the data helps us to identify potential extensions to the service landscape. Amongst others, services that ensure privacy policies in a traceable manner, visualize the data in dashboard-like overviews, allow simple querying using natural language patterns, and share gathered insights with other project managers have been identified as viable additions and will be subjects of future research.

Bibliography

[1] M. Uflacker and A. Zeier. *d.store: Capturing Team Information Spaces with Resource-based Information Networks.* IADIS International Conference WWW/Internet, Freiburg, Germany, 2008.

13.11 Data Placement Strategies for Cloud Platforms

Johannes Lorey (johannes.lorey@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Felix Naumann

As a rather new concept in Information Technology, the term Cloud Computing is used ambiguously and usually comprises a number of different service offerings, including Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). Nevertheless, some general traits are common in all these variations: on-demand resource allocation, no up-front payment and dynamic scalability.

Thus, migrating data from local storage and hosting it on a Cloud infrastructure should yield several benefits, such as reducing hardware costs while maintaining or increasing scalability. Different commercial providers already offer dedicated data storage services in the Cloud, usually referred to as Database as a Service (DaaS). However, there has been little academic or industrial research on how Cloud storage can be regarded as an additional tier for data placement in the context of Information Lifecycle Management (ILM). In my research, I identified different functional and non-functional issues that must be considered before migrating data to the Cloud.

First of all, accessing resources remotely typically results in higher latency and lower throughput rates, which could pose limitations for time-critical applications or might render them unusable. Moreover, depending on the nature of the data and application, not all Cloud infrastructures might be suited for hosting due to their design and lack of data operators. For example, financial operations rely heavily on the concept of transactions. Here, whenever data is modified, this change is considered atomic, i.e. modifications are either applied to all of the data in question or to none of it. Therefore, the database as a whole is always in a consistent state. In a Cloud environment, on the other hand, data is generally distributed among different physical machines, making it harder to synchronize transactions and offer an always-consistent view on the data. The dilemma of choosing between consistency and availability of data in distributed environments has been formalized in the CAP theorem and poses a major research topic.

In addition to these rather technical problems, storing data on a publicly available infrastructure may also result in security concerns depending on the sensitivity of the information and corresponding policies within an institution. Also, uptime considerations and other non-functional aspects of a Cloud infrastructure specified in the service level agreements (SLA) are essential factors when investigating data migration options. In summary, the goal of my research is to develop a Cloud-based framework that incorporates strategies for automated information placement, processing and retrieval depending on the nature of the data and the applications that use it.

13.12 Model-driven Generation of Security Policies for Service-oriented Architectures

Michael Menzel (Michael.Menzel@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Christoph Meinel

Service-oriented Architectures support the provision, discovery, and usage of services in different application contexts. The Web Service technologies (such as WSDL, UDDI, and SOAP) provide the means to describe, locate, and invoke a service based on XML. The independent nature of the services, with respect to operating systems and system architectures, facilitate a composition of different services.

However, the seamless and straightforward integration of cross-organisational services conflicts with the need to secure and control access to these services. Security-related requirements are stated in security policies to negotiate these requirements with service consumers. Since such policies are hard to understand and even harder to codify, I foster a model-driven approach in which security intentions are annotated in service orchestrations and are translated to deployable security policies.

My solution consist of three layers: Security Requirements, expressed at the modelling layer are translated to a platform independent model. This model constitutes the foundation to generate security configurations. To integrate security intentions in SOA system models, the integration schema introduced by SecureUML is used. My security design language SecureSOA can be mapped and merged with design modelling languages such as FMC or BPMN to create a security aware modelling dialect. For instance, such a dialect can be used to visualize services and to annotate these services with authentication intentions.

The transformation of these intentions is challenging, since different strategies might exist to enforce a security intention. To use expertise knowledge to facilitate the generation of these policies, I foster the usage of design patterns that encapsulate this knowledge and guide the transformation process.

In conclusion, I intend to integrate security engineering and service-oriented systems engineering by enabling the specification of security requirements at the modelling layer. Security intentions are translated to security policies using predefined security patterns that encapsulate security expert knowledge.

13.13 Resource Management in BPMN

Andreas Meyer (andreas.meyer@hpi.uni-potsdam.de)
Supervisor: Mathias Weske

In this thesis, we present an approach to extend BPMN 2.0 for achieving resource awareness. Today, BPMN is the most used standard for graphical modeling and process visualization and moves forward into the direction of process execution as well. Currently, BPEL is the standard chosen for execution but lacks visualization opportunities, which BPMN in turn provides. Additionally, a mapping between BPMN and BPEL is not straight forward as for instance Dijkman and Weidlich proofed. Therefore, BPEL will not be the end of development and evolution of process execution languages and BPMN might take over and become the most used standard for process modeling, visualization, and execution.

The current version of the BPMN 2.0 specification improves resource awareness but using the reference patterns by the Workflow Patterns Alliance, the coverage is still below 25%, which provides further improvement opportunities to get BPMN completely resource aware. Modeling wise, several important aspects as allocation of resources to tasks and restricting constraints for these allocations have been added with the latest specification, but execution wise, resources are not considered at all in the execution semantics. The current opportunities are mostly expression-based but lack a nice presentation. Furthermore, collaboration of resources is not considered yet.

Transition from BPEL to BPMN takes time and therefore, we provide a mapping from BPMN to BPEL4People, a resource extension to BPEL, comprising BPMN's current resource support state. The contribution of this mapping is a blueprint what to do with a current BPEL4People engine, the current BPMN specification, and the wish or need to model in BPMN including workarounds to provide new functionalities coming from BPMN in BPEL using known constructs as a first step. Afterwards, we will propose additions to BPMN covering the modeling and visualization aspects, which will extend the previous mapping and, if needed, BPEL4People as well. Consequently, this mapping provides a round trip starting from modeling in BPMN over executing in BPEL including monitoring to modifying in BPMN again to improve the model based on the made observations.

The contribution of this thesis is a mapping to combine the visualization of BPMN and the execution capabilities from BPEL covering basic resource management as first step towards BPMN as a resource aware and extensive process modeling, visualization, and execution language covering more than 90% of the workflow resource patterns and additional use cases, for instance in the area of resource collaboration.

13.14 Web Systems Design with Process-based Self Adaptive Behavior

Emilian Pascalau (emilian.pascalau@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Mathias Weske

Future Internet finds itself in a continuous transformation. Transformation that brings together services, service aggregation, knowledge representation, semantic relationships between concepts, reasoning techniques, processes, a paradigm shift in how people perceive, interact and use the Web. All these are the mold from which a new generation of intelligent applications must be build and more over constantly improved based on how users interact with the systems. This research addresses these problems in the context of mashups that run in the browser and involves defining a process based architecture capable of learning and adapting itself based on users behavior.

Mashups are a new genre of Web applications that fosters innovation and creativity in a generative environment, which Zittrain denotes as the Future Internet, and are probably "one of the more powerful capabilities coming out of the Web 2.0 wave". There are different approaches that deal with how things happen in such complex environment, i.e. semantic technologies that understand the meaning of information and facilitate the accessibility of the content, business rules which in the last 10 years were employed to declaratively describe policies, business processes and practices of an enterprise.

In addition actual technologies such as Asynchronous JavaScript and XML (AJAX), and the new browser improvements allow the development of Rich Internet Applications (RIAs), applications that typically run in a web browser, and do not require software installation. Several Web 2.0 applications use heavily AJAX in order to provide desktop-like behavior to the user. The number of RIAs is increasing because of the broad bandwidth of today's Internet connections, as well as the availability of powerful and cheap personal computers, also browsers provide richer functionalities, and extension points.

As process mining techniques have underlined in most of the cases the initial processes that have been embedded into systems do not correspond with *real* processes. As van der Aalst states "people tend to think in terms of highly simplified processes and their views on these processes often contain an initial bias. Therefore, it is vital to have an objective understanding of reality. Moreover, it is often not sufficient to understand things at an aggregate level." The architecture of next generation web systems and not only must tackle these type of problems by taking advantage of hybrid techniques that include semantic approaches, knowledge representation, rules and processes. Such systems must reason, learn and improve/adapt their behavior based on users' behavior.

13.15 Dynamic Service Analysis

Michael Perscheid (michael.perscheid@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Robert Hirschfeld

The Software Architecture group, led by Prof. Dr. Robert Hirschfeld, develops new methods, concepts, and tools for improving the comprehension and design of large complex systems.

One of the most important aspects in software development is to understand programs, with a special focus on those internal details that constitute system functionality. Current development environments offer a wide range of perspectives to support developers in understanding source code, but the task of program comprehension is still time-consuming and laborious. This is in part due to most tools' limiting their scopes to static views on source code without considering program behavior.

With such additional behavioral information available, new perspectives for comprehension could be offered, ranging from internal communication between objects to orchestration at the architectural level. Although the widespread creative use of debuggers for understanding behavior by developers indicates a need for dedicated behavioral views. There are only few development tools addressing this requirement. Most prominently, high memory consumption and performance penalties render current approaches impractical in software development.

We suggest a practical, lightweight, and incremental approach to dynamic analysis based on test cases as entry points into system behavior. By observing and enriching concrete examples of program executions, we investigate three perspectives on software systems with the main focus being on improving behavior comprehensibility.

First, we design and implement *dynamic analysis tools* that are seamlessly integrated into current development environments. Using these, developers can introspect the behavior of the system under test. Based on a light-weight call graph representation used for navigation, developers can state their points of interest and all further information will be computed on demand by re-executing tests. We distribute the overhead of dynamic analysis over multiple runs so that there is no need to collect all data at once, whereby memory consumption and performance impacts are kept low. Thus, we have a practical approach for behavioral views that will be evaluated regarding its practicability and improvements for program comprehension in the near future.

Next, our concept of *dynamic software architectures* merges static and behavioral (dynamic) system architecture. We investigate new system perspectives that are intended to support software engineering tasks. Tool support will be provided that guides developers to potential locations for,

for example, caching, redundancies, or unused code. Moreover, new tools will compare dynamic paths through the system and investigate anomalies, leveraging more suitable failure detection or hints at design disharmonies.

Finally, *post-traceability of requirements* is considered a critical component in software understanding, as it allows developers to comprehend the system from the user's point of view. However, existing traceability approaches often comprise tedious processes with a small degree of automation. We propose a new requirement traceability approach that automatically combines feature localization techniques with acceptance tests. Our concept is not limited to specific system properties, so that we can propagate requirement traceability knowledge across service and system boundaries. In future work, we will automate the manual step that deals with the connection between acceptance tests and requirements, enrich development tools with traceability information, and address some typical feature localization problems of service-oriented systems.

To summarize, dynamic service analysis offers new and deeper understanding of how services and objects belong together. Moreover, new starting points are given for improving reusability, robustness, and clarity of services and their implementations.

13.16 Abstraction of Process Specifications

Artem Polyvyanyy (Artem.Polyvyanyy@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Mathias Weske

Abstraction is the result of the generalization or elimination of properties in an entity or a phenomenon in order to reduce it to a set of essential characteristics. Information loss is the fundamental property of abstraction and is its intended outcome.

Process specifications are special types of entities that describe behavior observed or intended to be implemented in reality, which result from creative engineering practices. Software engineers and process analysts constantly deal with problems of designing, analyzing, and improving process specifications, e.g., source code, service compositions, or process models. Every process specification is a compromise between two points: On the one hand engineers strive to operate with less modeling constructs which conceal irrelevant details, while on the other hand the details are required to achieve the desired level of customization for envisioned process scenarios.

In our research, we develop mechanisms to vary the level of details, i.e., the abstraction level, of process specifications. The challenge lies in identifying what are the units of process logic suitable for abstraction and afterwards performing the abstraction. Once abstraction artifacts are identified, they can be eliminated or replaced by concepts of higher abstraction levels which conceal, but also represent, abstracted detailed process behavior. Finally, individual abstractions must be controlled in order to achieve an abstraction goal—a process specification that suits the needs of a use case.

A methodology for abstraction of process specifications is a combination of the techniques for discovering, performing, and controlling individual abstractions. In [1], we discuss different abstraction use cases and propose an abstraction slider as a mechanism for abstraction control. In [2], we propose a mechanism for the discovery of abstraction candidates, i.e., units of process logical suitable for abstraction. Afterwards, abstraction candidates get generalized or eliminated in a process specification.

Bibliography

[1] Polyvyanyy, A., Smirnov, S., Weske, M. *Process model abstraction: A slider approach.* In: EDOC., Munich, Germany, IEEE Computer Society (September 2008) 325–331

[2] Polyvyanyy, A., Smirnov, S., Weske, M. *The triconnected abstraction of process models.* In: BPM. Volume 5701 of LNCS., Ulm, Germany, Springer Verlag (September 2009) 229–244

13.17 Operating System Support for Monitoring Parallel Software Systems

Alexander Schmidt (alexander.schmidt@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Andreas Polze

Alexander Schmidt investigates operating system (OS) related aspects in service-oriented computing infrastructures at the Operating Systems and Middleware Group headed by Prof. Dr. Andreas Polze.

Service-oriented computing recently received quite significant attention in both the industry and research. As previous research has shown, monitoring is a key challenge in service-oriented computing. Monitoring is necessary at various stages of the software development life cycle. For example, a developer may use monitoring for tuning a service's performance or for debugging purposes. A service consumer on the other hand may use monitoring for checking service-level agreement compliance.

In recent years, there has been an ongoing trend towards many/multi-core architectures. This drives application developers to parallelize their applications in order to increase the performance and to benefit from the existence of multiple compute nodes. However, with parallel programs becoming main stream, monitoring those must be re-considered regarding an application's consistency constraints.

The task of monitoring an application comprises two phases: (1) collecting application specific information during its runtime and (2) analyzing this information with respect to some criteria. Monitoring is also a key part at various stages of the software development life cycle. Within this thesis, we concentrate on the first part.

If the gathered information consists of complex data types for which an access require more than a machine word to read, *e.g.*, when extracting a whole object, it is crucial that this read access is properly synchronized with the application. Otherwise the resulting data may be invalid within the state space of the application. Consequently, any analysis is doomed to fail based on this data.

To address this issue, we propose KStruct, a monitoring framework that encorporated in an operating system kernel, *e.g.*, the Windows Research Kernel, that facilitates consistent access to any application data with respect to the consistency properties of the application. KStruct therefore provides a domain specific language to make the application's locking model explicit through annotations, which are used to generate an access driver. If annotations are missing or incomplete, KStruct can perform a data flow analysis to infer the locking model itself. The access driver can be used either by a tracing or sampling monitor to gather consistent data.

13.18 Reuse of Legacy Software Components via Decompilation

Jan-Arne Sobania (jan-arne.sobania@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Andreas Polze

In the past, companies have relied on custom-developed software to support and automate their internal processes for a long time. Although standard software exists and is used for almost all aspects of business processes today, there often was a considerable effort of manual adaption required before rollout. With the advent of service-oriented architectures, companies now face the problem of how to interface them with their existing systems. Due to cost and time constraints, *redevelopment* is usually not an option; on the other hand, *porting* legacy components to new systems is not an easy task either, depending on the size and state of the program's source code or even whether original development systems are still available. Further complications arise if source code has been lost or is considerably out-of-sync with deployed executables.

The thesis discusses *decompilation* – the process of reversing the operation of a regular compiler – as a general tool for software reuse. As deployed executables represent a known-working state of a system, they can also serve as a foundation for porting. In addition, decompilation could be used for a variety of other tasks as well, including verification of seldomly-used code paths in regular compilers, safety analysis of programs for constrained environments, as well as security research and understanding of malware.

Basic decompilation techniques have been known since the 1960s [1] and implemented in a number of specialized decompilers. These decompilers have traditionally been written to support only one particular processor type, although remarkable parts of their core analysis algorithm are independent of the processor type or even the processor family. Few researchers have worked on *retargetable* decompilers in the past; i.e., decompilers supporting processors from different families with non-identical feature sets. However, these approaches generally failed to deliver a single core analysis that worked as well as a processor-specific one while still supporting changing the target architecture.

Our work focuses on the architecture and algorithms for a general, retargetable decompiler suitable both for classic as well as contemporary processor architectures.

Bibliography

[1] C. Cifuentes. *Reverse Compilation Techniques*. PhD thesis, School of Computing Science, Queensland University of Technology, 1994.

13.19 Reliable Digital Identities for SOA and the Web

Ivonne Thomas (Ivonne.Thomas@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Christoph Meinel

Open identity management models provide a way of sharing identity information across several trust domains in a controlled manner. The basic idea is to have multiple places to manage a user's identity data (so called identity providers) and to exchange identity attributes between entities holding identity information (the identity providers) and those consuming it (the relying parties). Open protocols and standards exists to ensure the interoperable exchange of identity attributes as security tokens between identity providers and relying parties (cf. e.g. OASIS Identity Metasystem Interoperability specification 1.0). However, the willingness to trust on such information is low as current approaches mainly consider the attribute value itself, but hardly how this value was collected or whether any verification process took place. In my research, I follow up the thesis that in order to enable service providers to rely on information from a foreign source, an identity management for the Internet should be able to deal with attributes with a strong verification besides attributes without any verification which are managed by the users themselves. Moreover, it should allow a relying party (such as a service) to assess the value of received identity information in terms of correctness and integrity. In previous work, we argued that this assessment should be done on the granularity level of the identity data – meaning, that the decision to trust should not only be made between the issuing and the relying party on a all-comprising level, but for each identity attribute, which is exchanged, separately. To give an example, we could consider a university which is trusted to make right assertions about whether a user is a student, but not about whether this user pays its telephone bills. Therefore, exchanging identity information over the Internet between unknown parties requires information in addition to the attribute value itself to make right assertion about the credibility of an identity attribute. This meta identity information is all information additionally to the attribute value itself which enables a relying party to decide whether it trusts the received value with regard to an intended transaction. As a proof of concept, we implemented an identity provider to manage reliable digital identities for SOA and the web. In the centre of the identity provider is the management of digital identities and associated attributes for Internet users. As protocols to exchange identity information, OpenID and InformationCard are supported. Current work includes the integration of identity information from different sources, in particular federated identity providers and their assessment in terms of trustworthiness for the attributes they can issue.

13.20 Visualization of Complex Software Systems

Jonas Trümper (jonas.truemper@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Jürgen Döllner

Large software systems, in particular service-oriented software systems, typically (a) consist of millions lines of code, (b) are maintained over a long period of time and (c) are developed by a large, diverse team. This poses an enormous challenge to developers in several dimensions. For example, (1) the knowledge about the whole system is typically distributed. That is, a single developer is no more able to memorize the complete system structure with all its details. More precisely, each developer working on the system typically has detailed knowledge about one or a few parts of the system and is roughly aware of the big picture. In addition to that, there are "orphaned" parts of the system's implementation that are not known in detail anymore by any developer. (2) The dependencies between system components may not be explicitly documented or visible. Dynamic binding due to polymorphism in object-oriented software systems complicates this even further as the actual dependencies are only visible at runtime. (3) Documentations and actual system implementations often exhibit significant differences in practice. Hence, the main reliable information sources are represented by the actual implementation artifacts, e.g., source codes and binaries.

Traditional development tools such as IDEs and debuggers only provide partial support for developers to cope with the challenges inherent to large software systems. Especially parallel execution poses a huge challenge for developers as it raises the system's runtime complexity by orders of magnitude. For example, synchronization has to be handled and each thread of execution has its own local stack and state. Thus, e.g., locating performance bugs in such parallel systems is a hard task.

With increasing market share of multi-core processors and multithreading in modern software systems, the aforementioned aspects need to be considered and appropriate tool support is important for high-quality software products. This work leads to visualization techniques and tools for developers of parallel software systems.

13.21 Models and Infrastructure Support for Self-Adaptive Service-Oriented Software Systems

Thomas Vogel (thomas.vogel@hpi.uni-potsdam.de)
Supervisor: Prof. Dr. Holger Giese

Runtime adaptability or even self-adaptability is often a crucial requirement for complex software systems that operate in a highly dynamic world [1, 2]. To achieve self-adaptation, a *managed system* and its environment are monitored and analyzed, and if changes are required, adaptations are planned and executed on the system. These activities are performed by *autonomic managers* that are usually located externally to a managed system.

The concept of *Service-Orientation* is helpful for realizing self-adaptation, since service-oriented systems are based on modular architectures that are compositions of loosely-coupled services and each service is specified by a self-describing contract. Therefore, service-based architectures are a promising abstraction mechanism for self-adaptive systems [2]. However, service-oriented systems are usually complex, large-scaled, and highly distributed, which requires specific solutions for self-adaptation. Instead of centralized solutions, only distributed solutions that decentralize highly adaptation activities might be feasible. However, the sum of all decentralized activities have to be globally consistent and have to fulfill global systems goals.

Within the thesis the usage of models as foundation for adaptation activities performed by autonomic managers is investigated. Models provide a rich semantic base for runtime adaptation by covering the system architecture with its environment, constraints, goals, and concerns of interest, like performance or reliability. Therefore, the *Model-Driven Engineering* (MDE) discipline is helpful though it is focused on using models at the design, implementation and deployment stages of software development. However, models and techniques from MDE can also be used at runtime for self-adaptation, which requires applicable and efficient online solutions, e.g., to maintain models at runtime or to validate whether a model reflects a managed system properly or not. Consequently, an infrastructure as a runtime environment that provides support for model maintenance and adaptation activities is required. Without infrastructure support, the idea and the realization of self-adaptive systems would be in most cases too laborious. To sum up, my research focuses on concepts for a generic model-driven infrastructure for distributed software systems employing the service-oriented computing paradigm, which enables distributed adaptations with the help of models. For proof of the concepts, the infrastructure will be prototyped.

First results of my work cover the usage of architectural runtime models for monitoring and adapting software systems [3, 4, 5]. Instead of employing one complex and platform-specific model at a low level of abstraction as a basis

for monitoring and adaptation, several models are employed simultaneously. Each model focuses on a specific concern of interest, like performance to address self-optimization, and abstracts from the platform of a managed system. Thus, these models are less complex and at a higher level of abstraction, and they provide appropriate views for autonomic managers. This eases adaptation activities performed by managers. Moreover, platform-independent models leverage the reusability of managers across different managed systems. Multiple runtime models are maintained automatically using model transformation techniques that synchronize online and incrementally models specified by different metamodels. Moreover, the development of maintaining models at runtime was considerably eased due to these techniques. The current implementation targets systems based on *Enterprise Java Beans 3.0* technology, while the models and MDE techniques are based on the *Eclipse Modeling Framework*. Next steps of my work include distributed and concurrent adaptations based on different runtime models.

Bibliography

[1] B.H.C. Cheng, R. de Lemos, H. Giese, P. Inverardi, J. Magee, et al. *Software Engineering for Self-Adaptive Systems: A Research Road Map*. In Software Engineering for Self-Adaptive Systems, Vol. 5525 of LNCS, pp. 1–26. Springer, 2009.

[2] E. Di Nitto, C. Ghezzi, A. Metzger, M. Papazoglou, and K. Pohl. *A journey to highly dynamic, self-adaptive service-based applications*. In Automated Software Engineering, Vol. 15(3-4), pp. 313–341, Springer, 2008.

[3] T. Vogel, S. Neumann, S. Hildebrandt, H. Giese, and B. Becker. *Model-Driven Architectural Monitoring and Adaptation for Autonomic Systems*. In Proc. of the 6th Intl. Conference on Autonomic Computing and Communications (ICAC), pp. 67–68. ACM, 2009.

[4] T. Vogel, S. Neumann, S. Hildebrandt, H. Giese, and B. Becker. *Incremental Model Synchronization for Efficient Run-Time Monitoring*. In Models in Software Engineering, Workshops and Symposia at MODELS 2009, Reports and Revised Selected Papers, Vol. 6002 of LNCS, pp. 124–139. Springer, 2010.

[5] T. Vogel and H. Giese. *Adaptation and Abstract Runtime Models*. In Proc. of the 5th ICSE Workshop on Software Engineering for Adaptive and Self-Managing Systems (SEAMS), ACM, 2010.

13.22 Dynamically Adaptive Data Quality Web Services

Tobias Vogel (tobias.vogel@hpi.uni-potsdam.de)
Supervisor: All Professors of HPI, primarily Prof. Dr. Felix Naumann

Data quality plays an important role for entrepreneurial success. However, many companies do not recognize the importance of data quality in their ERP or CRM systems, as recent studies show[1,2]. Many different technical measures can be employed to increase data quality, e.g., data normalization, duplicate detection, data fusion, etc. In my research, I concentrate on the detection of duplicates, i.e., multiple representations of same real world objects. Traditional duplicate detection employs well-established algorithms and heuristics, which - in short - search through a database and estimate the similarity of pairs of tuples based on data type, value, and additional information to identify these pairs as possible duplicates.

However, sometimes the amount of available information is restricted: the schema might not be up-to-date, the field-mapping is unclear, privacy issues prevent full access to all the data, etc. Due to their nature, Web Service implementations of data cleansing methods (Data Quality Web Services) share many characteristics with these restrictions and are used as foundation for evaluating duplicate detection algorithms.

Thus, the research question is how appropriate results can still be achieved under these given conditions. In other words: which information is essential for a duplicate detection process and which information therefore has to be inferred from the data.

Successful duplicate detection within unpredictable data and structure requires some effort before the actual detection of duplicates can take place.

- The fields of a record might not be clearly distinguishable. Therefore, the separator has to be found and entities have to be recognized.
- Records might have different schemas. Meta information, e.g., thesauri or ontologies, help resolve synonym relations between terms.
- There is a large variety of metrics to compare two entities. However, no metric works best on all types of data. That is why knowledge about the actual type (the "semantics") of data is beneficial for the selection and tuning of the respective algorithms.

[1] http://www.pbinsight.com/about/newsroom/press-releases/detail/3390_data-quality-study-reveals-businesses-still-face-significant-challenges/
[2] http://www.gartner.com/it/page.jsp?id=589207

Supported by:

Deutsche
Forschungsgemeinschaft

DFG

verlag mainz

ISBN 3861301466