

Gesellschaft für Informatik (GI)

publishes this series in order to make available to a broad public recent findings in informatics (i.e. computer science and information systems), to document conferences that are organized in cooperation with GI and to publish the annual GI Award dissertation.

Broken down into the fields of

- Seminar
- Proceedings
- Dissertations
- Thematics

current topics are dealt with from the fields of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure the high level of the contributions.

The volumes are published in German or English

Information: <http://www.gi-ev.de/service/publikationen/lni/>

ISSN 1614-3213
ISBN 978-3-88579-442-4



Gesellschaft für Informatik (Hrsg.): Informatiktag 2009

GI-Edition

Lecture Notes in Informatics



Gesellschaft für Informatik (Hrsg.)

Informatiktag 2009

Fachwissenschaftlicher Informatik-Kongress

27. und 28. März 2009

B-IT Bonn-Aachen International Center for
Information Technology, Bonn



Gesellschaft für Informatik (Hrsg.)

Informatiktage 2009

Fachwissenschaftlicher Informatik-Kongress

27. und 28. März 2009

**B-IT Bonn-Aachen International Center for Information
Technology in Bonn**

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Seminars

Series of the Gesellschaft für Informatik (GI)

Volume S-8

ISSN 1614-3213

ISBN 978-3-88579-442-4

Volume Editor

Gesellschaft für Informatik e.V.

Ahrstraße 45

53175 Bonn

E-Mail: gs@gi-ev.de

Redaktion: Ludger Porada

E-Mail: ludger.porada@gi-ev.de

Series Editorial Board

Heinrich C. Mayr, Universität Klagenfurt, Austria (Chairman, mayr@ifit.uni-klu.ac.at)

Hinrich Bonin, Leuphana-Universität Lüneburg, Germany

Dieter Fellner, Technische Universität Darmstadt, Germany

Ulrich Flegel, SAP Research, Germany

Ulrich Frank, Universität Duisburg-Essen, Germany

Johann-Christoph Freytag, Humboldt-Universität Berlin, Germany

Thomas Roth-Berghofer, DFKI

Michael Goedicke, Universität Duisburg-Essen

Ralf Hofestädt, Universität Bielefeld

Michael Koch, Universität der Bundeswehr, München, Germany

Axel Lehmann, Universität der Bundeswehr München, Germany

Ernst W. Mayr, Technische Universität München, Germany

Sigrid Schubert, Universität Siegen, Germany

Martin Warnke, Leuphana-Universität Lüneburg, Germany

Dissertations

Dorothea Wagner, Universität Karlsruhe, Germany

Seminars

Reinhard Wilhelm, Universität des Saarlandes, Germany

Thematics

Andreas Oberweis, Universität Karlsruhe (TH)

© Gesellschaft für Informatik, Bonn 2009

printed by Köllen Druck+Verlag GmbH, Bonn

Wissenschaftliche Tagungsleitung

Prof. Dr. Gottfried Vossen, Universität Münster

Programmkomitee

Prof. Dr. Herbert Göttler, Universität Mainz

Prof. Dr. Katharina Morik, Universität Dortmund

Prof. Dr. Otto Spaniol, RWTH Aachen

Prof. Dr. Alfred Zimmermann, Fachhochschule Reutlingen

Workshops - Themen und Partner

SOA Architekturmuster

Capgemini sd&m AG

Auf der Suche nach Bedienkonzepten der Zukunft

Comma-Soft AG

Web 2.0 - Technologien

iteratec Gesellschaft für iterative Softwaretechnologien mbH

Planen / Optimieren / Steuern

IVU Traffic Technologies AG

Entwicklung von Spielen mit XNA

Microsoft Deutschland GmbH

Schnittstellendesign - Sprachen, Formate, Protokolle

PPI Financial Systems

Übersicht über das Rahmenprogramm und Partner der Informatiktage 2009

Neben den Workshops wurde den Teilnehmerinnen und Teilnehmern ein attraktives Rahmenprogramm geboten. Wir bedanken uns bei unseren Partnern, ohne deren Unterstützung die Informatiktage nicht hätten stattfinden können.

Rahmenprogramm:

Wettbewerbspräsentationen des informatiCup 2008

Absolventen-Workshop

Leitung: Prof. Dr. Alfred Zimmermann, Fachhochschule Reutlingen

Hauptvorträge:

Prof. Dr. Maic Masuch, Universität Duisburg-Essen:

Trends und Perspektiven der Entwicklung von Computerspielen

Prof. Dr. Peter Pepper, TU Berlin:

Von Unterhemden und Oberhemden. Oder: Die Suche nach dem nächsten Top-Modell

Postersessions

Workshops in drei Sessions

Moderation der Veranstaltung:

Prof. Dr. Gottfried Vossen, Universität Münster

Partner und Sponsoren der Informatiktage:

Partner:

B-IT Bonn-Aachen Center

Sponsoren:

Capgemini sd&m AG

Comma-Soft AG

iteratec Gesellschaft für iterative Softwaretechnologien mbH

IVU Traffic Technologies AG

PPI Financial Systems

sd&m AG

Medienpartner:

COMPUTER ZEITUNG

JOBSUMA GmbH

Staufenbiel Media GmbH

Vorwort zum Tagungsband 2009

Im Studienjahr 2007/2008 haben sich gut 30.000 Studienanfänger/innen für das Fach Informatik entschieden. Dies sind rund 4% mehr als im Vorjahr. Dennoch ist die Gesellschaft für Informatik (GI) der Ansicht, dass diese Steigerung angesichts des Bedarfs, der nach wie vor hohen Abbrecherzahlen und der nach wie vor guten Berufsaussichten in der Informatik unbefriedigend ist. Die Zahl der Studienanfänger/innen in der Informatik zu erhöhen, ist daher ein volkswirtschaftlich wichtiges Ziel, von dem die Wirtschaftskraft unseres Landes nachhaltig abhängt. Zur Erhöhung der Studierendenzahlen in Informatik gibt es viele Initiativen, wie die Weiterentwicklung der Curricula, ein Streiten für eine bessere finanzielle und personelle Ausstattung an den Hochschulen und nicht zuletzt ein Arbeiten am Bild der Informatik, das uns nach wie vor nicht attraktiv genug erscheint. Ich freue mich deshalb sehr, dass zu den Informatiktagen jedes Jahr wieder viele interessierte und engagierte Informatikerinnen und Informatiker kommen, die sich im Teamwork praktischen Aufgabenstellungen widmen.

Die Informatik hat sich zum Innovationsmotor in fast allen Branchen entwickelt. Gerade in den etablierten Branchen wie dem Automobil- und Maschinenbau bringt die Verschmelzung klassischer Ingenieurskunst mit der Informationstechnologie die wesentlichen Innovationsschübe und die notwendige Wertschöpfung. Ohne die Informatik ist mittlerweile der größte Teil der Neuerungen an den Produkten der deutschen Schlüsselbranchen gar nicht mehr denkbar. Um weiterhin wettbewerbsfähig zu bleiben, bedarf es allerdings der verstärkten Grundlagenforschung und einer fundierten Ausbildung an den Hochschulen.

Mitte letzten Jahres lag nach einer Statistik des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) in Nürnberg die Arbeitslosenquote aller Berufe bei 10,0 %, die der Berufskennziffer „Datenverarbeitungsfachleute“ nur bei 5,2 %, wobei diese Zahl seit fünf Jahren kontinuierlich abnimmt. Nach wie vor bietet der Arbeitsmarkt gut ausgebildeten Informatikfachleuten also viel versprechende Chancen auf ein interessantes Arbeitsleben.

Ich wünsche Ihnen, den Teilnehmerinnen und Teilnehmern der Informatiktage, dass Sie untereinander und in Gesprächen mit den hier vertretenen Unternehmen einen fruchtbaren Austausch haben, dass Sie auch weiterhin mit Begeisterung Ihr Fach Informatik studieren und im Anschluss an das Studium einen vielversprechenden Arbeitsplatz finden. Den Unternehmen danke ich für ihre Bereitschaft, die Informatiktage zu fördern, dem B-IT für seine Gastfreundschaft und den Vertrauensdozent/inn/en und dem Programmkomitee für die inhaltliche Arbeit.



Prof. Dr. Stefan Jähnichen (GI-Präsident)

Inhaltsverzeichnis

Künstliche Intelligenz

Berlea A., Döhring M., Reuschling N.C., Störl U. Extracting Semantic Relations of High Textual Complexity	15
Borth D., Ulges A., Schulze C., Breuel T.M. Video Copy Detection providing Localized Matches	19
Boschmann A., Kohring A., Reinhardt W. Finden von Experten in Projekten durch Profiling und Verhaltensbewertung	23
Ebert S. Dirichlet Process Mixture Models for Object Category Recognition	27
Holz D. Effiziente 2D-Navigation für Mobile Service Roboter	31
Kutz M. Traffic Simulation for Virtual Environments	35
Schreyer M. Intelligent Printing Technique Recognition and Photocopy Detection for Forensic Document Examination	39
Meier C., Schäfer J. MetaXsA: Metadatenextraktion von externen Informationsquellen unter der Berücksichtigung vorhandener Taxonomien	43
Waldvogel B. Machine Learning Verfahren zur Klassifikation von kurzen Textstücken	47

Softwaretechnik

Blanken M.B.

Komponentenbasierte Software-Entwicklung mit WebObjects 51

Diehl H.

Enterprise Content Management im SAP NetWeaver Portal 55

Finke S.

Workflow-Spezifikation für einen Knowledge Server - Konzeption und prototypische Umsetzung XML-basierter Server-Workflows 59

Förster M.

Verifikation von Datenabhängigkeiten in Ableitungscode 63

Harder A.

Suche nach prozessadäquaten Webservices mit Hilfe eines BPEL Repositories 69

Herrling D., Schindler M.

Rettungsassistenzsystem 75

Kleinschmager S.

A Controlled Experiment for Measuring the Impact of Aspect-Oriented Programming on Software Development Time 79

Kurpick T.

Werkzeuggestützte Prüfungen dynamischer Prozesse 83

Löser M., Leucker M., Nowotka D., Rischen J.

Äquivalenzklassen mit Zeit und Zuständen 87

Nordemann F.

Aggregationsverfahren für Software-Komponenten, 91

Roth S., Dörsam M., Gründling S., Langholz T., Steinbrecht A.

Integrating a Legacy Terminal Application into an SOA 95

Rothmann A.

Entwicklung eines neuartigen Optimierungsverfahrens für die automatische Generierung von Testfällen 99

Mensch-Computer-Interaktion

Abrazhevich V.

Erstellung und Anwendung eines Usability-Kriterienkatalogs für
Online-Bewerbungssysteme an deutschen Hochschulen 103

Bothe S. u.a. (Projekt BesserWissA, Universität Bremen)

AuthorSupportTool 107

Eichel S., Rackow B.

Headtracking für zwei Personen mit der Wiimote 111

Fries S., Lenz K.

Anwendungen in virtueller Realität 115

Lattermann M., Nauerz A., Kriha W.

Implicit Social Network Construction and Expert User Determination for
Context-AwareWeb Portal Environments 119

Mikus A.-C., Nuhn S., Schafmeister C.

PerLE: Persönliche Lernumgebung zur Anzeige und Suche von Wissen 123

Radomski S.

A Framework for Rapid Prototyping of Augmented Reality Interfaces 127

Reck C.

Entwicklung eines Konzepts zur modellbasierten Fehlerdiagnose 131

Scherfgen D.

Integration physischer und mentaler Stressfaktoren in eine
VR-basierte Simulationsumgebung 135

Wagner M.

Probabilistic User Models for the Verification of Human-Computer Interaction 139

Datenbanken und Informationssysteme

Buschjost O.

Entwicklung eines flexiblen Plugin-Systems für das Text-Mining im
Kontext der Technologieplanung 143

Chowaw-Liebmann O.

Indoor Navigation Based on Sectorized Positioning 147

Färber I.

Mining orthogonaler Konzepte in hochdimensionalen Datenbanken 153

Heyde R., Lutter S.

Navigating aggregated Textual Information from the Web with Schema Discovery 157

Hudlet V., Tritschler M.

Towards Providing Complete Knowledge in Constraint-based Database Caching 161

Klein J., Wonneberger H.

Dynamische Konfiguration von Cache Groups beim Constraint-basierten Datenbank-Caching 167

Muthmann K.

Grouping of near duplicates in user generated forum content 171

Neumann S.

Indizierung und Kompression von Schachendspieldatenbanken 175

Stülpnagel J., Terwelp C.

Automatische Adaption von Informationsdiensten für mobile Kommunikationsgeräte 179

Technische und Theoretische Informatik**Cheng D.**

Asynchronous Distributed Indexing of Large Media Collections 183

Egners A.

Versteckte Dienste in öffentlichen Umgebungen 187

Kühne S.

Ein modellgetriebener Architektur-Ansatz zur Umsetzung von verteilten Anwendungen durch ein plattformunabhängiges Modell 191

Meffert F.

Implementierung und Bewertung eines Lokalisierungsdienstes basierend auf aktiver RFID-Technologie 195

Ritsche S.

Parallele Motivsuche in großen Netzwerken 199

Textor A.

Monitoring unternehmenskritischer Anwendungen unter Verwendung modellbasierter Performance Constraints 203

Wirtz H.

A Topology-Aware Overlay for Wireless Mesh Networks 207

Wischweh J.	
Aktivitätsorientierte Kontextadaption in mobilen Anwendungen	211

Graphische Datenverarbeitung

Bergmann K.	
Interfaces based on a hand-held projection screen	215
Campen M.	
A Framework for Geometry Processing based on Hybrid Surface Representations	219
Forkert N. D.	
Visualisier Analyse und dynamische 3D-Visualisierung des Blutflusses von zerebralen Gefäßstrukturen unter Verwendung von 3D- und 4D-Magnetresonanzangiographie-Bildfolgen	223
Jonas S.	
Ein OpenSource Framework für iterative Bildsegmentierung mit Grah Cuts	227
Misiak D.	
Automatische Detektion primärer Neuronen in Fluoreszenzmikroskopbildern	231
Tamm G.	
Deferred Shading	235

Wirtschaftsinformatik

Magnus S.	
Ein Rahmenwerk für Genetische Algorithmen zur Lösung erweiterter Vehicle Routing Problems (VRPSPDMUTW+)	239
Nicolai J.	
Towards extending the Tracker Workflow Capabilities of SourceForge Enterprise Edition	243
Solich S.	
Performanzanalyse von Web Services für das Netzwerkmanagement mit SNMP	247

Informatik und Gesellschaft/Informatik und Ausbildung

Döbbelin S.

Arbeitsumgebungen in der Softwareentwicklung im Team - Themen- und
Expertennetzwerke und deren zeitabhängige Veränderungen 251

Hunke M.

PocketLogo – Ein Logo-Interpreter für Handhelds 255

Maicher D., Scholz D.

Community-Verhalten im Micro-Blogging mit Twitter 259

Meissner G.

Ein Lernspiel zu grundlegenden Problemen in verteilten Systemen 265

Sicherheit – Schutz und Zuverlässigkeit

Frisch W.

Machbarkeitsstudie mit Analyse und Bewertung zur Durchführung von
Onlinedurchsuchungen in der Praxis 269

Sandjaja D.

Identifizierung von Webseiten bei verschlüsselten Verbindungen 273

Extended Abstract: Extracting Semantic Relations of High Textual Complexity

M. Döhring^{1,2}, N. Reuschling^{1,2}, U. Störl¹, A. Berlea²

doehring@stud.fbi.h-da.de, reuschling@stud.fbi.h-da.de,
u.stoerl@fbi.h-da.de, alexandru.berlea@sap.com

¹ Hochschule Darmstadt, ² SAP Research

1 MOTIVATION AND SCOPE

The ability to automatically gain information from unstructured data is an essential task for the purpose of exploiting the enormous potential of information available on the Web. In particular, the promise of the *Semantic Web*, more specifically of being able to automatically capture the meaning of Web content, requires the ability to extract semantic relations from unstructured text. Semantic Relation Extraction builds upon technologies from Natural Language Processing (NLP). While NLP techniques have greatly matured throughout the last decade, improvements are still needed in order to be able to use it fully automatically. For example, despite significant efforts invested in relation extraction in research areas such as Ontology Learning and Ontology Population [BCM05], a qualitatively acceptable procedure for the automatic extraction of ontologies is still a long way off [Zho07]. A main disruptive factor is the yet insufficient addressing of relations with high textual complexity in terms of nested or interchanged syntactic structures as well as the presence of multiple named entities in a sentence. With the biomedical domain in focus, [JVD07] points out the general relevance of sentence complexity for data misinterpretation. [KKT03], applying most prevalent NLP techniques for extracting relations in the biomedical domain, state that 40% of their false negatives are due to the complexity of sentences. Motivated by current research presented in [Doe08], we manually validated similar figures also for the task of recognizing company cooperations in business texts detected by OpenCalais [Reu], a publicly available system for named entity recognition and relation extraction. Below, we give an authentic example for such a sentence with a high degree of complexity.

To provide another authorized Web outlet for its assets, *Viacom* signed a licensing agreement last month with *YouTube* rival *Joost* to distribute video content from *Comedy Central* and *MTV* over the Internet.

What are the main obstacles when automatically extracting relations from the above sentence? Currently, one limitation of prevalent approaches for relation extraction arises from

the expressiveness of the patterns that they are able to recognize. Most approaches use *shallow* pattern matching (in some way similar to regular expressions), that is they are partly oblivious of the grammatical structure of text sentences. In the above example, they could wrongly detect a collaboration example between *Viacom* and *YouTube*, negatively influencing precision as well as recall. The shallow patterns are intrinsically dependent on the order of words in sentences and fail to capture a relation if the relative order of the words changes, even though the expressed relation remains the same. Uninvolved named entities even complicate the detection of a relation or the correct assignment entities directly involved in the relation. We therefore require approaches which abstract away from the representation of a sentence as a sequence of tokens to representations that are more stable with respect to the relative order of words as long as the relation stated among the words or entities of interest remains the same.

One component for the realization of this idea is the exploitation of grammatical dependencies, i.e. linguistic structures that go beyond the flat sentence level, for relation extraction. Scientific work in this area has increased recently, although most of them either rely on a trained model or only address narrow types of relations (such as is-a or part-of). We show that it is possible to achieve improvements in accuracy for relations of high textual complexity also with a defined unsupervised procedure.

2 OUTLINE OF SOLUTION APPROACH

In order to cope with the challenge of extracting high textual complexity semantic relations, we elaborated a generalizable hybrid procedure consisting of shallow pattern matching over word resp. part-of-speech (POS) tokens and the exploitation of grammatical dependencies on the sentence level. With the help of a fixed set of linguistic rules on the tree level, a parse tree with hierarchical sentence chunks and POS-tagged words can be mapped to a dependency tree with each word assigned to its governor [DMM06]. For instance, it indicates the dependency of adjectives on the nouns they modify or the dependency of nouns on their corresponding verbs. In Figure 1, we use a combined representation of the parse tree together with the dependency tree. The dotted arcs are labeled with the dependency type and concurrently indicate the direction of a dependency by pointing on the governing word. For example, the noun '*butterfly*' depends on the verb '*to catch*' via a direct object relation. Basically, our approach relies on the shortest path hypothesis of [BM05], which states that the information needed to determine a semantic relation between two entities in a text lies on the shortest path among them following their grammatical dependency relations. However, we discovered the need to extend this basic assumption to properly work with a current state-of-the-art sentence parser from [KM03], since it sometimes fails to recognize the correct grammatical dependencies. We therefore defined a set of pre- and postprocessing rules in addition to our matching algorithm handling special sentence constellations as for example conjunctions or the above mentioned multiple entity occurrences.

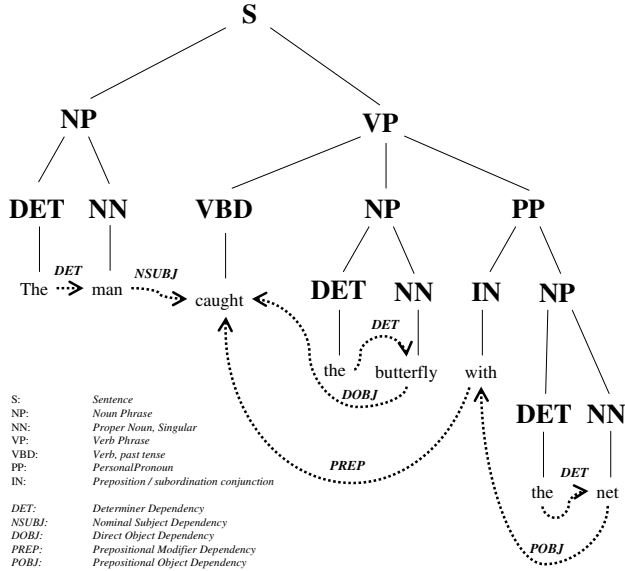


Abbildung 1: Example of a deep parse tree with dependency annotations.

The main advantage of the approach is its ability to overcome the interfering impact of „intermediate“ sentence elements between two candidate entities for a semantic relation in a very long or complex sentence. In our initial example sentence, the direct object grammatical dependency of *Joost* to *Viacom* helps detecting the correct semantic relation.

3 RESULTS AND FUTURE WORK

To be able to asses the usefulness of the proposed procedure especially for our particular application domain of high sentence text complexity, we defined a novel measure for determining the *complexity* of a semantic relation on the textual level. This metric was used for the practical application of our approach on a large¹ up-to-date dataset consisting of business news articles. Our evaluation was performed in direct comparison to OpenCalais with regard to the detection of collaboration relations between companies in the corpus. We chose the top-150 sentences according to the complexity metric and manually evaluated the correctness of the presence or absence of discovered semantic relations. According to the chi-square test value with Yates correction for two categories of data (correctly and wrongly detected relations in the sentences), our presented procedure significantly outperforms the results of OpenCalais at a confidence level of 99,9%. We can conclude that our presented procedure bears potential for improving relation extraction accuracy especially

¹ 550.000 candidate sentences

for relations of high textual complexity.

Some issues remain subject to further research. Testing against a manually built-up gold standard (consisting of about 200 randomly chosen sentences from the corpus), both systems yielded relatively low overall precision and recall values (about 35%), partly due to 'hidden' semantic relations which can only be discovered by a human reader with the use of context and/or world knowledge. A second issue is that in sentences of low complexity according to our metric, OpenCalais seems to yield better results. Consequently, it has to be examined how establish procedures for semantic relation extraction in *simple* text corpora can be combined with our approach. Furthermore, a part of the rules and procedures employed by our system were only introduced due to errors and inaccuracies of the employed sentence parser and thus do not follow any original linguistic motivation. That means, the improvement of sentence deep parsing performance or the employment of a different parsing technique may change circumstances and conditions for our system and therefore might require its adjustment. Additionally, the exploitation of automatic pattern learning algorithms has not been addressed yet.

Literatur

- [BCM05] P. Buitelaar, P. Cimiano und B. Magnini. *Ontology Learning from Text: An Overview*. In P. Buitelaar, P. Cimiano und B. Magnini, Hrsg., *Ontology Learning from Text: Methods, Evaluation and Applications*, Jgg. 123 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2005.
- [BM05] Razvan C. Bunescu und Raymond J. Mooney. A shortest path dependency kernel for relation extraction. In *HLT '05: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Seiten 724–731, Morristown, NJ, USA, 2005. Association for Computational Linguistics.
- [DMM06] M.-C. De Marneffe, B. Maccartney und C. D. Manning. Generating typed dependency parses from phrase structure parses. In *In LREC 2006*, Seiten 449–454, 2006.
- [Doe08] Markus Doehring. EXTRACONN: Extraction and Analysis of Company Networks from News. In Rainer Ruggaber, Hrsg., *Proceedings of the 1st Internet of Services Doctoral Symposium 2008 at International Conference on Interoperability of Enterprise Systems and Applications (I-ESA'08)*, Berlin, Germany, March 25th 2008.
- [JVD07] H. Jose, T. Vadivukarasi und J. Devakumar. Extraction of protein interaction data: a comparative analysis of methods in use. *EURASIP J. Bioinformatics Syst. Biol.*, 7(4):1–9, 2007.
- [KKT03] A. Koike, Y. Kobayashi und T. Takagi. Kinase pathway database: An integrated protein-kinase and NLP-based protein-interaction resource. *Genome Research*, 13:1231–1243, 2003.
- [KM03] D. Klein und C. D. Manning. Accurate unlexicalized parsing. In *ACL-41*, Seiten 423–430, Morristown, NJ, USA, 2003. Association for Computational Linguistics.
- [Reu] Reuters. OpenCalais - Calais WebService. <http://www.opencalais.com>.
- [Zho07] Lina Zhou. Ontology learning: state of the art and open issues. *Inf. Technol. and Management*, 8(3):241–252, 2007.

Video Copy Detection providing Localized Matches

Damian Borth, Adrian Ulges, Christian Schulze, Thomas M. Breuel
d borth@cs.uni-kl.de

German Research Center for Artificial Intelligence (DFKI)
and University of Kaiserslautern

Abstract:

With the availability of large scale online video platforms like YouTube, copyright infringement becomes a severe problem, such that the demand for robust copy detection systems is growing. Such system must find multiple occurrence of copyright protected material within video clips that are created, modified, remixed and uploaded by the user. A particular challenge is to find the exact position of a copy in a – potentially huge – reference database. For this purpose, this paper presents a Content Based Copy Detection system that both detects copies in query videos against a reference database and gives an exact alignment between them. For finding and aligning a matching shot, a fast search for candidates is conducted, and as a second step an exact alignment is found using a dynamic programming minimization of the well-known edit distance from text retrieval. The introduced approach was evaluated on the public available MUSCLE-VCD-2007 [LTJB07] data corpus and showed competitive alignment results compared to the ACM CVPR 2007 evaluation.

GI-Topic: KI-BV (artificial intelligence - image understanding)

1 Introduction

Large scale online video platforms like YouTube rises new critical issues for content owner and platform provider when it come to the illegal distribution of copyright protected video content by its users. Furthermore the nature of such platforms encourage the reediting, modification and mixing of self-created video content with already available video content not created by the user. Often the reused video snippets consist of copyright protected video material that is used without the knowledge of the content owner and therefore should be identified and deleted by the platform provider. Content Based Copy Detection (CBCD), as an alternative to a watermarking technique, solves this issue.

As stated by [LTCJ⁺07], it is important to not confuse similarity driven search in Content Based Video Retrieval (CBVR) with the task of finding a copy in CBCD. A copy is not only an identical or near duplicate video sequence, it is rather a transformed video sequence resulting in a distortion of its original appearance. In Fig. 1 possible video transformations are shown.

2 Our Approach

A common setup of a CBCD systems holds a database of known copyright protected videos (reference database) and checks questioned video (query videos) against it [DLÁ⁺07]. In case a query video contains a reference video the CBCD system should match both

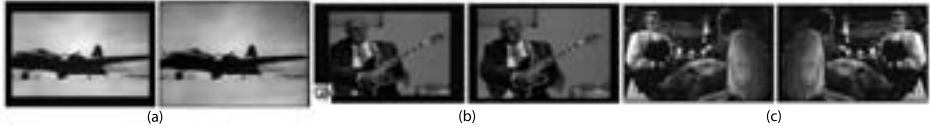


Figure 1: Common video transformations: a) cropping, b) logo insertion and blurring, c) flipping

video sequence regardless of the applied transformations. Our method is utilizing an edit distance [Lev66] based similarity measurement on extracted video signatures to detect and localize copies within query videos. As CBCD systems consist of common components [GB08] [DGJ⁺08], only the differences to those are explained as following:

Preprocessing & Feature Extraction The preprocessing step performs an adaptive keyframe extraction [BUSB08], which additionally segments the processed video into shots. Having these structural information of the video we extract color histograms from the keyframes and MPEG-7 Color Layout Descriptors (CLD) [MOVY01] from the shots. This process is performed on the entire reference database and for each query video.

Candidate Computation & Similarity Check Due to efficiency reasons we first compute k shot candidates using Nearest Neighbor Matching on keyframe level with color histograms features. Then we perform the computational more expensive edit distance calculation on the shot candidates using CLD features, mainly due to their robustness and efficiency in CBVR [KY01]. In Fig. 2 the resulting trellis images of similarity check between a query and a reference shot is shown. The left image visualize the direction selection during the edit distance calculation. A distinct localization of a copy would appear as a diagonal line indicating an identical sequence given by the lowest cost for its frame-by-frame comparison. It can be seen that the left image is too noisy to make a distinct localization possible. An additional Least Square Fitting of the blue trellis path in the right image improves the localization precision dramatically. In combination with the edit distance value and the trellis information we are able to make a global decision if a positive match was detected and a local decision where the match occurred.



Figure 2: The left image visualize the direction selection during the edit distance calculation. The right image displays the result of the Least Square Fitting of the blue trellis path. The green bar below the images indicates the true copy position, the red bar at the top displays the localization result.

Match Marking & Post Processing In this processing step the decision must be made if a query video contains sequences of a reference video. For the global decision we threshold the calculated edit distance of the similarity check against a predefined value where low costs indicate high similarity. In case of a positive match, adjacent query shots with their edit distance result are checked against the threshold and potentially merged together. This leads to a localization based on the entire query video and on reference video level.

3 Experiments & Results

For experimental evaluation of the copy detection system we used the MUSCLE-VCD-2007 data corpus [LTJB07]. The dataset consist of a reference database with 100 hours of video materials and a set of 3 queries with total length of 45 min. This query set belongs to the **ST2** task of the **CVPR 2007 Copy Detection Live Benchmark**, which deals with localized CBCD. The structure of the provided queries can be seen in Fig. 3, where each query consist of randomly selected video sequences form the reference database mixed with non-reference material and modified by one or several of the following transformations: cropping, fade cuts, insertion of logos, moving texts, change of contrast, gamma. Sequences belonging to the reference database must be detected and localized in the query and the reference video with their start and end time code.

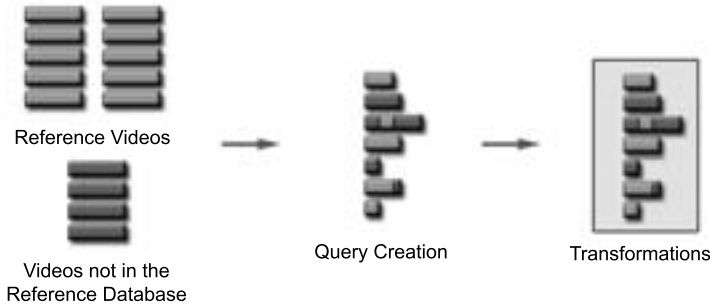


Figure 3: Query structure, mixing random video sequences from a reference and non-reference database together and applying random transformation to each of the query

The used performance criteria for the CBCD evaluation were defines as the following two separate measurements:

$$QualitySegment = \frac{N_{correct} - FalseAlarm}{N_{segments}}$$

QualitySegment is computed from the percentage of mismatched segments in all queries, where *Ncorrect* is the number of correct matches, *FalseAlarm* is the number of false positive matches and *Nsegments* represents the number of all detected segments.

$$QualityFrame = 1 - \left(\frac{N_{mis}}{N_{frames}} \right)$$

QualityFrame represents the percentage of mismatches frames in all queries including non-detected segments, imprecision within correctly detected segments and false positive matches. Here *Nmis* stands for the amount of mismatches frames and *Nframes* for the total number of query frames.

Our system obtained a value of **0.79** for *QualitySegment*, compared with 0.86 as best and 0.33 as worst result of the 2007 evaluation and **0.35** for *QualityFrame*, compared with 0.76 as best and 0.17 as worst result of the 2007 evaluation. An analysis of the *QualityFrame* results identified false negative, like depicted in Fig. 5, as major source of frames mismatches (700 sec.). Frame mismatches for correctly detected copy sequences were in the range of 120 sec. out of a total query length of 45 min. containing 21 min. of video from the reference database.

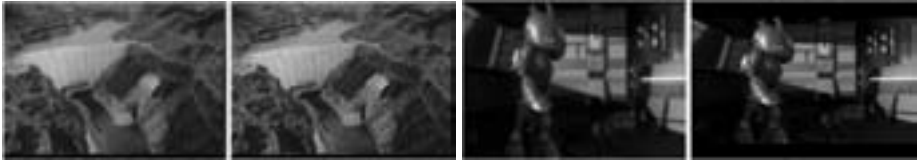


Figure 4: Two examples of correctly detected query sequence which were modified by blur, change of color (left) and crop, blur, zoom (right) transformations

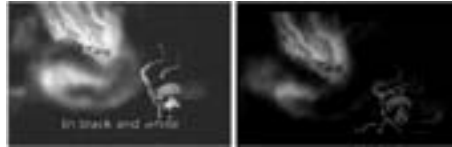


Figure 5: An example of a not detected query sequence modified by gamma transformation, a vertical shift and insertion of subtitles. The lack of color information additionally challenged the detection system due to its focus on color features.

References

- [BUSB08] Damian Borth, Adrian Ulges, Christian Schulze, and Thomas M. Breuel. Keyframe Extraction for Video Tagging and Summarization. In *Informatiktage 2008*, pages 45–48, 2008.
- [DGJ⁺08] M. Douze, A. Gaidon, H. Jegou, M. Marszałek, C. Schmid, and F. Grenoble. INRIA-LEAR’s Video Copy Detection System. In *TRECVID-Workshop*, 2008.
- [DLÁ⁺07] K. Dadason, H. Lejsek, F. Ásmundsson, B. Jónsson, and L. Amsaleg. Videntifier: identifying pirated videos in real-time. In *Proceedings of the 15th international conference on Multimedia*, pages 471–472. ACM Press New York, NY, USA, 2007.
- [GB08] N. Gengembre and S.A. Berrani. The Orange Labs Real Time Video Copy Detection System-TrecVid 2008 Results. In *TRECVID-Workshop*, 2008.
- [KY01] E. Kasutani and A. Yamada. The MPEG-7 color layout descriptor: a compact image featuredescription for high-speed image/video segment retrieval. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 1, 2001.
- [Lev66] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10:707–710, 1966.
- [LTCJ⁺07] J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford. Video copy detection: a comparative study. In *Proceedings of the 6th ACM international conference on Image and video retrieval*, pages 371–378. ACM Press New York, NY, USA, 2007.
- [LTJB07] J. Law-To, A. Joly, and N. Boujemaa. Muscle-VCD-2007: a live benchmark for video copy detection, 2007. <http://www-rocq.inria.fr/imedia/civr-bench/>.
- [MOVY01] B.S. Manjunath, J.R. Ohm, V.V. Vasudevan, and A. Yamada. Color and Texture Descriptors. *IEEE Trans. on Circuits Syst. for Video Techn.*, 11(6), 2001.

RaMBo: Finden von Experten in Projekten durch Profiling und Verhaltensbewertung

Andreas Kohring

Universität Paderborn
Institut für Informatik
ako@upb.de

Alexander Boschmann

Universität Paderborn
Institut für Informatik
sirius@upb.de

Wolfgang Reinhardt

Universität Paderborn
Institut für Informatik
wolle@upb.de

Abstract: Die Metadaten eines Dokumentes können nicht nur zum Indizieren und somit zum Auffinden eines Dokumentes dienen. Sie können auch Aussage darüber geben, mit welchen Themen sich ein Benutzer beschäftigt hat. Vor diesem Hintergrund entstand im Rahmen der Projektgruppe MoKEx IV an der Universität Paderborn eine Softwarekomponente, die es ermöglicht, Metadaten eines Dokuments dafür zu nutzen, eine Aussage über den Kenntnisstand oder die Kompetenz einer Person zu einem bestimmten Thema zu treffen.

Art der Arbeit: Projektgruppe

Betreuer: Wolfgang Reinhardt, Johannes Magenheim

1 Problemstellung

Ein häufiges Problem in Unternehmen ist es, zu spezialisierten Fragestellungen qualifizierte Ansprechpartner zu finden. Besonders bei Unternehmen mit einer großen Zahl an Mitarbeitern ist es für einen einzelnen Mitarbeiter schwierig, herauszufinden, wer der kompetenteste Ansprechpartner zu einem bestimmten Thema ist. Um diese Suche zu vereinfachen, wurde mit RaMBo (Rating Module and Behavior Observation) eine Softwarekomponente entwickelt, die erzeugte und abgerufene Informationsobjekte bewertet.

2 RaMBo - Inhalts- und Verhaltensbewertung

Durch die zentrale und automatische Verwaltung aller Dokumente eines jeden Mitarbeiters durch das Projekt MoKEx IV und die automatisierte Extraktion von Metadaten [MS08] sowie deren semantischer Analyse [SMR08] steht eine Vielzahl an Metadaten zu jedem Dokument eines Benutzers zur Verfügung. Die Metadaten werden im LOM-Format [LOM02] gespeichert, wobei der Standard geringfügig erweitert wurde. Zu den extrahierten Metadaten gehören im Allgemeinen die wichtigsten Schlüsselworte, alle mitwirkenden Autoren

und die Einordnung in den Unternehmenskontext mithilfe von Taxonomien. Für Dokumenttypen wie E-Mails werden zusätzlich Informationen über Anhänge und alle Empfänger der E-Mail einschließlich ihrer Zuordnung zu den Feldern CC, BCC und TO erfasst. Über die Analyse von LOM-Objekten hinaus bietet RaMBo auch die Möglichkeit, Suchanfragen innerhalb des Unternehmenskontextes zu erfassen. Dies geschieht über eine eigene Schnittstelle, die es ermöglicht, Suchanfragen zu analysieren. Dies wird durch eine standardisierte Suchsyntax ermöglicht.

Mithilfe der gesammelten Daten ist es möglich, verschiedene Arten von Auswertungen durchzuführen. Es kann z.B. aufgezeigt werden, welcher Mitarbeiter des Unternehmens als Experte zu einem Themengebiet betrachtet werden kann. Darüber hinaus kann auch detaillierter erfragt werden, wer als Experte zu bestimmten Schlüsselworten gilt.

Durch die Informationen über Autoren und E-Mail Empfänger ist es weiter möglich, ein virtuelles Personennetz auf Basis des Wissensstandes der einzelnen Mitarbeiter zu erzeugen. Beispielhaft ist dies in Abbildung 1 dargestellt: Sie zeigt neben dem aktuellen Benutzer in der Mitte drei weitere Mitarbeiter mit ihrem Punktestand zu einem bestimmten Themengebiet. Die Stärke der sie verbindenden Kanten ist proportional zur Häufigkeit der gemeinsam bearbeiteten Dokumente. Wenn der Benutzer mithilfe von RaMBo einen Experten zu diesem Thema sucht, stellt er fest, dass er mit Michael bereits an einigen Dokumenten zusammengearbeitet hat. Dieser hat zum gleichen Thema viele gemeinsame Dokumente mit Tom, der einen hohen Punktestand zu dem gesuchten Thema hat. Diesen kann er jetzt direkt oder durch Michael um Rat fragen.

Ebenso kann ein Netz von Personen erzeugt werden, das zeigt, wie stark verschiedene Personen über ihr Wissen zusammenhängen. So ist es möglich, zu visualisieren, über welche Themengebiete Mitarbeiter zueinander in Beziehung stehen. Eine weitere Visualisierungsmöglichkeit bietet ein Netz von Schlüsselworten, aus dem ihre gegenseitigen Zusammenhänge abgelesen werden können.

Um solche Informationen zu erzeugen, wurde eine Reihe von Bewertungsrelationen erstellt. Sie dienen dazu, Beziehungen zwischen den Benutzern, den Schlüsselwörtern sowie den Taxonomien und deren Quellen herzustellen. Ebenfalls wird die Häufigkeit des paarweisen Auftretens von Schlüsselwörtern wie auch Taxonomien erfasst. Durch das Erfassen von Paaren von Schlüsselworten, Taxonomien und Benutzern ist es möglich, Aussagen über verwandte Begriffe, Taxonomien und Personen zu treffen. Dies wird durch die Transitivität¹ gewährleistet, siehe Bewertungsrelationen 4 bis 7. Die Relationen enthalten z.T. auch die Quelle in der die Paarung auftraten. Im Folgenden sind die Relationen aufgeführt:

1. Benutzer \times Schlüsselwort \times Bewertung
2. Benutzer \times Taxonomie \times Bewertung
3. Benutzer \times Quelle \times Bewertung
4. Benutzer A \times Benutzer B \times Schlüsselwort \times Quelle \times Bewertung
5. Benutzer A \times Benutzer B \times Taxonomie \times Quelle \times Bewertung
6. Schlüsselwort A \times Schlüsselwort B \times Zähler
7. Taxonomie A \times Taxonomie B \times Zähler
8. Taxonomie \times Schlüsselwort \times Zähler

¹Treten A und B, wie B und C paarweise auf, ist auch eine Beziehung zwischen A und C wahrscheinlich.

3 Bewertungsmetrik

Grundlage für die Bewertungsmetrik ist die Annahme, dass verschiedene Dokumenttypen und verschiedene Aktionen unterschiedlich stark zur Bewertung der Kompetenz der Benutzer beitragen. So gelten zum Beispiel Informationen, die aus einem offenen Wiki stammen oft nicht so gesichert und korrekt, wie die aus einer wissenschaftlichen Arbeit. Ebenso wenig ist es gesichert, dass jemand, der eine Korrektur an einem Dokument vornimmt ebenso qualifiziert ist, wie der Verfasser. Auch das Lesen eines Dokumentes bedeutet nicht, dass der Leser danach so qualifiziert ist wie ein Verfasser. Daher verfügt RaMBo über eine Metrik, die es ermöglicht, genau dieses Problem abzubilden. Es wird zwischen verschiedenen Dokumenttypen wie Office-Dokumenten, Wiki-Artikeln und E-Mails unterschieden. Darüber hinaus wird die Aktion erfasst, die mit dem Dokument durchgeführt wurde. Vier Aktionen werden dabei grundsätzlich unterschieden: suchen, lesen, editieren oder erstellen, für die das System in seiner derzeitigen Form 1, 10, 75 oder 250 Punkte berechnet. Diese Punktzahl wird mit einem Wert zwischen 0 und 1 multipliziert, der die prozentuale Gewichtung einer Aktion auf einem Dokument darstellt und für jedes Paar aus (Dokument \times Aktion) festgelegt ist. Diese Gewichtung erlaubt es beispielsweise, festzulegen, dass das Editieren eines Wiki-Artikels mehr Punkte ergibt, als das Lesen aber weniger, als das Verfassen. Die Punktwerte für die Dokumenttypen sowie die Wertungsfaktoren für die Aktionen lassen sich beliebig festlegen, Tabelle 1 zeigt eine mögliche Belegung, wie sie in der Projektgruppe MoKEx IV verwendet wird. Des Weiteren ist RaMBo in der Lage, die bereits gespeicherten Bewertungen jederzeit unter Berücksichtigung der eingestellten Parameter zu aktualisieren, da nicht nur Ergebnisse der Bestimmung einer Bewertung gespeichert werden, sondern alle nötigen Daten zum Errechnen der Bewertung. Somit kann im laufenden Betrieb eine Feinjustierung der Wertungsfaktoren vorgenommen werden, ohne dass Bewertungen verloren gehen.

	suchen	lesen	editieren	erstellen
Dokument	1	1	1	1
Wiki	0,8	0,8	0,8	0,8
Suche	0,2	0	0	0
E-Mail	0,4	0	0	0,4
E-Mail To	0	0,4	0	0,4
E-Mail Cc	0	0,2	0	0,2
E-Mail Bcc	0	0,2	0	0,2

Tabelle 1: Wertungsfaktoren

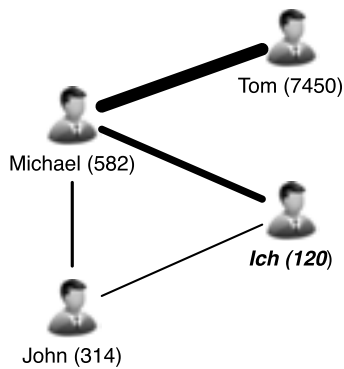


Abbildung 1: Personennetz

4 Zusammenfassung und Ausblick

Im Rahmen der Projektgruppe MoKEx IV war es bislang noch nicht möglich, hinreichende Studien vorzunehmen, um eine praxisnahe Belegung der Wertungsfaktoren zu bestimmen. Es wird daher angestrebt, zum Ende des Projektes diese Werte mit einem Industriepartner zu verifizieren. Des Weiteren soll RaMBo zu einem späteren Zeitpunkt völlig eigenständig ohne die Projekteigenen Kernkomponenten einsetzbar sein und zur Analyse beliebiger LOM-Daten eingesetzt werden können. Dazu zählt auch die Unterstützung verschiedener Ausgabeformate für die Bewertungsergebnisse. Speziell für die erzeugten Netze sollen verschiedene Standardformate wie GraphML unterstützt werden.

Literatur

- [HHM⁺06] Knut Hinkelmann, Kai Holzweißig, Johannes Magenheim, Fabian Probst und Wolfgang Reinhardt. Linking Communities of Practice with Learning Communities in Computer Science Education. In *Education for the 21st Century — Impact of ICT and Digital Resources*, IFIP Computer Science, Seiten 83–92. Springer Berlin / Heidelberg, 2006.
- [MS08] Christian Mletzko und Benedikt Schmidt. Automatische Extraktion von Metadaten aus verschiedenen Dateiformaten. In Gesellschaft für Informatik, Hrsg., *Informatiktag 2008. Fachwissenschaftlicher Informatik-Kongress, 14. und 15. März 2008, B-IT Bonn-Aachen International Center for Information Technology in Bonn*, Jgg. S-6 of *Lecture Notes in Informatics*, Seiten 181–184, 2008.
- [IMS99] University of Stuttgart, Institute for Romance Linguistics und the Institute for Computer Science. Textcorpora und Erschliessungswerkzeuge. <http://www.ims.uni-stuttgart.de/projekte/tc/>, 1999.
- [LOM02] Learning Technology Standards Committee of the IEEE. Draft Standard for Learning Object Metadata. http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf, 2002.
- [SMR08] Benedikt Schmidt, Christian Mletzko und Wolfgang Reinhardt. MetaXsA - Modulares Framework zur semantischen Dokumentanalyse für unstrukturierte Textdokumente. In Gesellschaft für Informatik, Hrsg., *Informatiktag 2008. Fachwissenschaftlicher Informatik-Kongress, 14. und 15. März 2008, B-IT Bonn-Aachen International Center for Information Technology in Bonn*, Jgg. S-6 of *Lecture Notes in Informatics*, Seiten 197–200, 2008.

Dirichlet Process Mixture Models for Object Category Recognition

Sandra Ebert

Advisers: Michael Stark, Mario Fritz and Prof. Bernt Schiele

Technische Universität Darmstadt

Department of Computer Science

Multimodal Interactive Systems

{ebert,stark,fritz,schiele}@mis.tu-darmstadt.de

Abstract: Probabilistic models are widely used in machine learning and information retrieval. Model selection is very important in these models, because of the possibility of under- or overfitting of the observed data. In connection with parametric approaches, this problem can be solved by testing different models and choosing the best one based on a performance measure. Dirichlet processes on the other hand follow a different route: when used as a non-parametric prior in a Bayesian framework, Dirichlet Processes allow the model itself to adapt to the complexity of the observed data. However, the behavior and performance of the models depend strongly on the choice of the hyperparameters. For the present work, we investigated Dirichlet processes in the context of three different types of model: Dirichlet Process Mixture Models, Hierarchical Dirichlet Process Mixture Models, and Nested Chinese Restaurant Processes. The primary goal was to study the properties of the stochastic processes in a comprehensive way for the different models, to determine the optimal hyperparameters, and to compare the nonparametric approaches to parametric ones such as k -means clustering and Latent Dirichlet Allocation. For the Nested Chinese Restaurant Processes, we developed two different heuristic approaches supporting the performance of these models.

1 Introduction

Probabilistic models are used in machine learning to explain observed data, to classify new data or as a tool for Information Extraction. Model selection is essential in these models: if the model is too general or too specific, it results in under- or overfitting, respectively. Parametric Models solve these problems by learning and testing different models with different complexity. Without prior knowledge about the data, this usually results in a large running time overhead. The Dirichlet Process (Ferguson [Fer73]) offers a flexible approach to adapt the complexity of the model to the data.

However, this flexibility of the dirichlet processes requires a good estimation of the hyperparameters, since the process behavior strongly depends on them [MBJ04]. The following issues in the use of these stochastic processes therefore have to be addressed: how can the optimal hyperparameters be found efficiently and quickly? How do the results compare with the parametric approaches? Is there a trustable measure to rate the development of

the models and thus guide us quickly to the optimal solution? How do the results depend on the choice of the hyperparameters?

To answer these questions, we have implemented and investigated the Dirichlet Process in three different contexts in order to obtain insight into the functioning and properties of these processes for all of the models at hand. The performance of the models was evaluated using an unsupervised image re-ranking task as described in [FFPZ05] and the results were compared to those in [FS08].

In the next three sections, the following models are briefly outlined: Dirichlet Process Mixture Models, Hierarchical Dirichlet Process Mixture Models and Nested Chinese Restaurant Processes. The conclusion provides an overview about the structure of my diploma thesis.

2 Dirichlet Process Mixture Models

The Dirichlet Processes have been most commonly used in the Dirichlet Process Mixture Models. These models consist of exactly one Dirichlet Process. Therefore, they are easy to implement and to understand. The Dirichlet Process Mixture Models are often used for Clustering, with different numerous applications, such as Genome Research [XJR07], relational learning [XTY⁺05], Data Mining [IM05] or Natural Language Processing [LWF07]. Depending on the application, we use different distributions, such as the normal distribution in Gaussian Process Mixture Models. In our work, we focused on multinomial distribution. Afterwards, our results were compared with results from parametric k -means Clustering with different k .

3 Hierarchical Dirichlet Process Mixture Models

The Hierarchical Dirichlet Process Models belong to the so-called Topic Models. The data are transformed into a lower dimensional representation using topics (Figure 1). These topics are typical distributions of gradients from the whole image dataset. Each document is then represented as a distribution of topics.



Figure 1: Different topics for face images, such as for a male person, a female person or a background

The current parametric base approach is the Latent Dirichlet Allocation [BNJ03]. Parametric Topic Models have been successfully used for scene classification [QMO⁺05], in Information Retrieval [FFPZ05] and for object category recognition [LJ06], [FS08]. However, the parametric approach is limited by the predefined number of topics. The Hierarchical Dirichlet Process Mixture Models from Teh [TJBB04], [TJBB06] were developed as a counterpart to the parametric Latent Dirichlet Allocation. There are several, hierarchically arranged Dirichlet Processes in this model: one global Dirichlet Process for the whole topic set and one local Dirichlet Process for the document specific topic set for each document.

4 Nested Chinese Restaurant Processes

The Nested Chinese Restaurant Processes were developed by Blei [BGJT04]. They detect the optimal number of topics and sort them in a tree, so that the most general topic is in the root and the most specific topics are in the leaves. Each document is assigned to one path in this hierarchy and consists of the topics in this path. The Dirichlet Process is used to find the optimal tree, since the degree of branching is arbitrary. As a result of these models, we obtain a relationship between topics.

Up to now, the model has been investigated by Sivic [SRZ⁺08] using the Scale-invariant feature transformation and by Bart [BPPW08] in a modified hierarchical Latent Dirichlet Allocation. Because of the adaptable hierarchy, these models are also well suited to quickly add new data.

5 Outline

My diploma thesis is structured as follows. An overview of current development and uses of the Dirichlet Processes and the three models can be found in *Related Works*. In the chapter *Dirichlet Processes*, the theoretical foundations, the functioning and the properties of the Dirichlet Processes, and further fundamentals like Gibbs Sampling are explained. The chapters *Dirichlet Process Mixture Models*, *Hierarchical Dirichlet Process Mixture Models* and *Nested Chinese Restaurant Processes* describe the various models. Each of these chapters contains a brief introduction, details on the implementation, observations concerning convergence and stability of the models, experiments and results, and a final conclusion. In the chapter *Heuristics*, two different heuristics are introduced. These were specially developed by us for the Nested Chinese Restaurant Processes and improve both the accuracy and the running time needs of the model. In the last chapter, *Conclusion*, we give a final summary of our results, an interpretation of our observations and an outlook on different approaches for future work. The appendix contains a glossary with the key words.

References

- [BGJT04] David M. Blei, Thomas L. Griffiths, Michael I. Jordan, and Joshua B. Tenenbaum. Hierarchical topic models and the nested Chinese restaurant process. In *Advances in Neural Information Processing Systems*, page 2003. MIT Press, 2004.
- [BNJ03] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, 2003.
- [BPPW08] Evgeniy Bart, Ian Porteous, Pietro Perona, and Max Welling. Unsupervised Learning of Visual Taxonomies. In *CVPR*, 2008.
- [Fer73] T.S. Ferguson. A Bayesian analysis of some nonparametric problems. *Annals of Statistics*, 1:209–230, 1973.
- [FFPZ05] R. Fergus, L. Fei/Fei, P. Perona, and A. Zisserman. Learning object categories from google’s image search. In *ICCV*, 2005.
- [FS08] Mario Fritz and Bernt Schiele. Decomposition, Discovery and Detection of Visual Categories Using Topic Models. In *CVPR*, 2008.
- [IM05] Hal Daumé III and Daniel Marcu. A Bayesian Model for Supervised Clustering with the Dirichlet Process Prior. *J. Mach. Learn. Res.*, 6:1551–1577, 2005.
- [LJ06] D. Larlus and F. Jurie. Latent mixture vocabularies for object categorization. In *BMVC*, 2006.
- [LWF07] L.J.Li, G. Wang, and L.F. Fei. OPTIMOL: automatic online picture collection via incremental Model learning. In *IEEE Computer Vision and Pattern Recognition or CVPR*, pages 1–8, 2007.
- [MBJ04] Jon D. Mcauliffe, David M. Blei, and Michael I. Jordan. Nonparametric empirical Bayes for the Dirichlet process mixture model. Technical report, UC Berkeley, 2004.
- [QMO⁺05] P. Quelhas, F. Monay, J.-M. Odobez, D. Gatica-Perez, T. Tuytelaars, and L.V. Gool. Modeling scenes with local descriptors and latent aspects. In *ICCV*, 2005.
- [SRZ⁺08] J. Sivic, B. C. Russell, A. Zisserman, W. T. Freeman, and A. A. Efros. Unsupervised Discovery of Visual Object Class Hierarchies. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [TJBB04] Z.W. Teh, M.I. Jordan, M.J. Beal, and D.M. Blei. Sharing clusters among related groups: Hierarchical dirichlet processes. In *NIPS*, 2004.
- [TJBB06] Yee W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet Processes. *Journal of the American Statistical Association*, 101(476):1566–1581, 2006.
- [XJR07] E. Xing, M. Jordan, and R. Roded. Bayesian haplotype inference via the Dirichlet process. *Journal of Computational Biology*, 14:267–284, 2007.
- [XTY⁺05] Zhao Xu, Volker Tresp, Kai Yu, Shipeng Yu, and Hans-Peter Kriegel. Dirichlet enhanced relational learning. In *ICML*, 2005.

Effiziente 2D-Navigation für Mobile Service Roboter

Dirk Holz

Bonn-Rhein-Sieg University of Applied Sciences

`dirk.holz@inf.fh-bonn-rhein-sieg.de`

Abstract: Das Ziel dieser Arbeit ist das Design und die Implementierung eines vollständigen Systems zur robusten Navigation mobiler Roboter in häuslichen Umgebungen. Adressierte Probleme sind die Modellierung von Umgebungen, die Planung von Pfaden sowie die Steuerung eines mobilen Roboters. Das resultierende System wurde erfolgreich in der Robocup@Home-Liga eingesetzt.

1 Motivation und Problemstellung

Serviceroboter, die z.B. im Haushalt assistieren, Besucher durch Museen führen oder in der Alten- und Krankenpflege eingesetzt werden können, bergen ein enormes Potential, den Alltag des Menschen erheblich zu vereinfachen [SHW04]. Um in der Welt, in der wir leben, sicher und effizient agieren zu können, benötigen autonome Roboter Fähigkeiten zur Wahrnehmung und Modellierung ihrer Umgebung, sowie, bis zu einem bestimmten Grad, über diese und in ihr stattfindenden Änderungen Schlussfolgerungen anzustellen. Zur effektiven Navigation und Planung von Aktionen benötigen mobile Roboter interne Modelle ihrer Umgebung. Die Fähigkeit diese selbstständig zu erzeugen sowie sich selbst und andere Objekte in ihnen lokalisieren zu können, wird als Voraussetzung für vollständig autonom agierende Roboter gesehen [Wan04]. Das Problem bei der Modellierung einer unbekannten Umgebung mit mobilen Robotern ist, dass der Roboter seine Position kennen muss, um erfasste Umgebungsstrukturen in einer Karte eintragen zu können. Zur Bestimmung seiner Position benötigt er jedoch eine Karte der Umgebung. Dieses Zyklusproblem wird in der Fachliteratur als *Simultaneous Localization and Mapping* (kurz SLAM) bezeichnet.

Der Kontext dieser Arbeit ist die RoboCup@Home-Liga, ein Wettbewerb für mobile Service Roboter. Hier müssen Roboter in häuslichen Umgebungen navigieren, Objekte suchen, erkennen, greifen und transportieren. Die Kommunikation mit dem Menschen ist auf Sprache beschränkt und erfordert robuste Spracherkennung sowie verständliche Sprachsynthese [NDD⁺08].

Der hierfür verwendete Roboter (Abbildung 1) besteht aus einer mobilen Plattform, mit zwei differentiell angetriebenen und zwei nachlaufenden Rädern zur Stabilität, einem 2D-Laserscanner zur berührungslosen Messung von Distanzen zu umliegenden Umgebungsstrukturen, diversen Kamerasystemen und einem Roboterarm zur Manipulation von Objekten.



Abbildung 1: Roboter

Das Ziel dieser Arbeit ist es, basierend auf den Daten des 2D-Laserscanners, speicher-effiziente Umgebungsmodelle zu erstellen und diese zur Lokalisierung des Roboters, zur Planung von Pfaden und zur Steuerung des Roboters selbst zu verwenden. Hauptaugenmerk liegt hierbei auf der Echtzeitfähigkeit und Robustheit der verwendeten Verfahren, da die Roboter zusammen mit Menschen arbeiten und jede Art von Kollision vermieden werden sollte.

2 Konstruktion des Umgebungsmodells – SLAM

Eine Vielzahl an Lösungsansätzen für das SLAM-Problem wurde über die letzten Jahrzehnte vorgestellt. Diese unterscheiden sich vor allem in der Formulierung des Problems, des eingesetzten Lösungsverfahrens und der Repräsentierung des Umgebungsmodells. Unterschieden werden, unter anderem, diverse probabilistische Verfahren, Graph-basierte Verfahren und solche, die SLAM als Problem der Registrierung von Distanzbildern adressieren. Das hier vorgestellte Verfahren basiert auf der Registrierung von zwei- bzw. dreidimensionalen Punktmengen, wie sie von Laserscannern erfasst werden. Der iterative Algorithmus der nächsten Punkte (*Iterative Closest Point – ICP*) [BM92] ist ein Standard-Algorithmus zur Registrierung von Punktmengen. Der ICP-Algorithmus unterscheidet zwei Punktmengen: das *Modell* M und die *Szene* oder Datenmenge D . Die Registrierung von M und D wird als Optimierungsproblem betrachtet: Der ICP verfeinert iterativ eine Transformation $\mathbf{T} = (\mathbf{R}, \mathbf{t})$, die D auf M abbildet, und somit die folgende Fehlerfunktion minimiert.

$$E(\mathbf{T}) = \sum_{i=1}^{|M|} \sum_{j=1}^{|D|} w_{ij} \|\mathbf{m}_i - (\mathbf{R}\mathbf{d}_j + \mathbf{t})\|^2 \quad \text{mit } w_{ij} = \begin{cases} 1, & m_i \text{ und } d_j \text{ korrespondieren} \\ 0, & \text{sonst} \end{cases} \quad (1)$$

Die Korrespondenzen w_{ij} werden durch die Suche nach dem nächsten Nachbarn $\mathbf{m}_i \in M$ für jeden Punkt $\mathbf{d}_j \in D$ ermittelt. Dieser rechenaufwendigste Schritt des Algorithmus wird beschleunigt durch die Verwendung einer approximierten Suche in einem kD -Baum [MA97]. Für die Minimierung von (1), im zwei- und dreidimensionalen Fall, existieren verschiedene Lösungen in geschlossener Form. Die gesuchte homogene Transformationsmatrix \mathbf{T} ist rigid, beinhaltet also nur Rotation (\mathbf{R}) und Translation (\mathbf{t}). Dass der Algorithmus zu einem (lokalen) Minimum konvergiert, wurde bereits in [BM92] bewiesen, allerdings unter der Annahme, dass für jeden Punkt \mathbf{d}_j wirklich ein korrespondierender Punkt in M existiert. Bei der Registrierung von Tiefenbildern im Kontext des SLAM-Problems ist dies nicht immer der Fall und die zu registrierenden Punktmengen überlappen nur partiell. Daraus resultierende falsche Korrespondenzen haben einen negativen Effekt auf die Registrierung, können aber durch Verwerfung bestimmter Punktpaare umgangen werden, z.b. wenn die Punkt-zu-Punkt Distanz einen Schwellwert überschreitet, oder der Modellpunkt \mathbf{m}_i mehrfach in den Punktkorrespondenzen auftritt [RL01].

Für die Registrierung mehrerer Laserscans und die Erzeugung eines Umgebungsmodells wird hier ein inkrementelles Verfahren verwendet. Der erste Laserscan D_1 bildet hierbei die Basis für das Umgebungsmodell M_1 , das als simple Punktwolke repräsentiert wird. Das Koordinatensystem von D_1 , in dessen Ursprung sich der Roboter befindet, bildet das

Koordinatensystem des Modells. Neue Laserscans D_i werden mit M_{i-1} registriert und die enthaltenen Punkte $d_{i,j}$ werden, nach Anwendung der gefundenen Transformation T_i zum Modell M_{i-1} hinzugefügt. Auf diese Weise wird das Umgebungsmodell inkrementell erweitert ($M_i = M_{i-1} \cup \{R_i d_{i,j} + t_i | d_{i,j} \in D_i\}$). Um den Speicheraufwand des Umgebungsmodells zu minimieren und redundantes Speichern von Punkten zu vermeiden, wird eine weitere Punktpaarsuche durchgeführt. Nur solche Punkte in D_i , deren Distanz zum korrespondierenden Punkt in M einen Schwellwert ϵ_d überschreiten, werden hinzugefügt. Zusätzlich wird die Punktdichte in den Scans reduziert, indem Punkte, deren Distanz zueinander unter ϵ_d liegt, durch ihren Schwerpunkt ersetzt werden. Ein, auf diese Weise erzeugtes Umgebungsmodell ist in Abbildung 2 dargestellt. Es ist um eine Liste von vorgegebenen und erlernten Objekten erweitert.

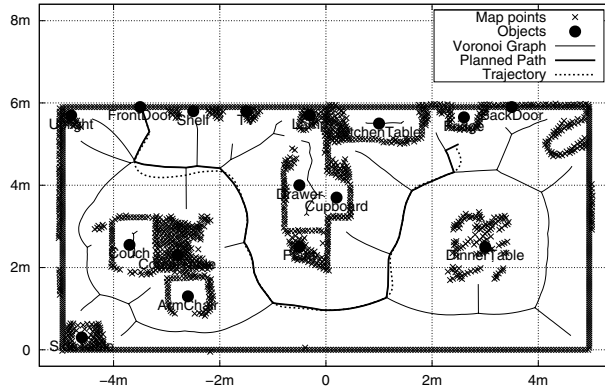


Abbildung 2: Modell einer @Home-Arena mit Objektvektor, geplantem Pfad und Trajektorie.

3 Pfadplanung und Steuerung

Um zu bestimmten Positionen in der Umgebung zu gelangen, muss der Roboter zunächst kollisionsfreie Pfade planen und diesen anschließend folgen. Für die Planung wird der A^* Algorithmus [HNR68] zur Berechnung des kürzesten Pfades zwischen zwei Knoten in einem Graphen verwendet. Um aus einem, wie oben beschriebenen, Punktmodell M einen Graphen für die Pfadplanung zu erzeugen, wird zunächst das Voronoi-Diagramm für die in M enthaltenen Punkte berechnet [For87]. Anschließend werden alle Kanten, die von dem Roboter nicht befahren werden können aus dem Graphen entfernt. Dazu zählen Kanten, deren Abstand zu nächsten Modellpunkten kleiner ist als die halbe breite des Roboters, und Kanten, die außerhalb der konvexen Hülle des Modells und damit in nicht modellierten Bereichen liegen. Die konvexe Hülle wird hier mit *Graham's Scan* Algorithmus [Gra72] berechnet.

Um einen geplanten Pfad abzufahren wird zunächst ein nichtlinearer Regler verwendet, um sich der Zielposition auf einer kollisionsfreien Bahn zu nähern [IC04]. Anschließend wird ein weiterer Regler [Ind99] verwendet, um Zielposition und -orientierung zu erreichen. Zusätzlich werden Verfahren zur Kollisionsvermeidung aus [HLS08] verwendet. Eine daraus resultierende Trajektorie ist in Abbildung 2 dargestellt.

4 Abschließende Bemerkungen

Das resultierende System wurde 2008 erfolgreich in der RoboCup@Home-Liga eingesetzt. Der Roboter belegte den 2. Platz bei den RoboCup German Open in Hannover und wurde Vize-Weltmeister bei den RoboCup Weltmeisterschaften in China. Videos, die den Einsatz der hier beschriebenen Verfahren zeigen, sind verfügbar unter:

<http://www.b-it-bots.de/media>.

Mögliche Erweiterungen des Systems mit einem 3D-Laserscanner zur räumlichen Modellierung der Umgebung und zur 3D-Kollisionsvermeidung wurden in [HLS08] vorgestellt.

Literatur

- [BM92] P. J. Besl und N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [For87] S. J. Fortune. A Sweepline Algorithm for Voronoi Diagrams. *Algorithmica*, Seiten 153–174, 1987.
- [Gra72] R. L. Graham. An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set. *Information Processing Letters*, 1(4):132–133, 1972.
- [HLS08] D. Holz, C. Lörken und H. Surmann. Continuous 3D Sensing for Navigation and SLAM in Cluttered and Dynamic Environments. In *Proceedings of the International Conference on Information Fusion (FUSION)*, 2008.
- [HNR68] P. E. Hart, N. J. Nilsson und B. Raphael. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.
- [IC04] G. Indiveri und M. L. Corradini. Switching linear path following for bounded curvature car-like vehicles. In *Proceedings of the IFAC Symposium on Intelligent Autonomous Vehicles (IAV04)*, 2004.
- [Ind99] G. Indiveri. Kinematic Time-invariant Control of a 2D Nonholonomic Vehicle. In *Proceedings of the 38th Conference on Decision and Control (CDC'99)*, 1999.
- [MA97] D. Mount und S. Arya. ANN: A library for approximate nearest neighbor searching. In *Proceedings of the 2nd Annual Fall Workshop on Computational Geometry*, 1997.
- [NDD⁺08] D. Nardi, J.-D. Dessimoz, P. F. Dominey, L. Iocchi, P. E. Rybski, J. Savage, S. Schiffer, T. Wisspeintner, T. van der Zant und A. Yazdani. RoboCup@Home – Rules and Regulations. Available online, 2008.
- [RL01] S. Rusinkiewicz und M. Levoy. Efficient Variants of the ICP Algorithm. In *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling*, 2001.
- [SHW04] R. D. Schraft, M. Hägele und K. Wegener, Hrsg. *Service - Roboter - Visionen*. Hanser Fachbuchverlag, 2004.
- [Wan04] C.-C. Wang. Simultaneous Localization, Mapping and Moving Object Tracking. PhD Thesis CMU-RI-TR-04-23, Carnegie Mellon University, Pittsburgh, PA, 2004.

Traffic Simulation for Virtual Environments

Michael Kutz

Bonn-Rhein-Sieg University of Applied Sciences
michael.kutz@hochschule-bonn-rhein-sieg.de

1 Introduction

Today's computer graphics and animation techniques are able to produce nearly photo realistic virtual environments which are used for computer games, virtual cities or simulations. While these environments look very realistic, there is a lack of realism in the behavior of their inhabitants, like cars or pedestrians.

This contribution presents the Virtual Environment Traffic Simulator (VETS) – a general approach for the simulation of urban population in virtual three-dimensional environments (VE's) to improve the experience of a real living world to the user.

1.1 State of the art

There are many systems to simulate traffic for analytic use cases (i.e. [YK96, ER01, Hid05]). While these systems are bound to be very good in the prediction of traffic situations, they are not ready to present *realistic looking* traffic to a user of a VE. Furthermore, these systems are focused on motorized traffic and do not consider pedestrians or bicycles.

Better fitting approaches seem to exist in the area of computer gaming. Games that allow the player to discover a surrounding urban world on her/his own need to present some kind of population to the user. While this seems to be exactly what is needed to improve the user's experience of VE's in general, playing these games reveals compromises being made between performance and realism, resulting in behavior irritating to the user.

1.2 Requirements

The overall aim of VETS is to create the illusion of real urban traffic in a VE. To achieve this, several types of urban actors like pedestrians, cars, bicycles etc. must be simulated.

Another very important feature of such a system is interactivity between the virtual urban actors and the user and among themselves. For instance, if the user (controlling a pedes-

trian) decides to cross the street carelessly, the cars should perform any maneuver to avoid the collision with him.

In reality no urban actor shows exactly the same behavior than another. There are differences in the preference of speed or distance to a leading vehicle depending on the personality of the actor. These differences are an important part of urban traffic and should be considered in VETS.

2 Concept

In VETS a need- and model-based agent system (see [RN03]) is used, inspired by several microscopic traffic simulation systems ([YK96, ER01, Hid05]).

2.1 Composition of the agent

The VETS-agent consists of three components: logical, mechanical and human. The logical component is the core of the agent and contains all logic the agent needs for rational acting in urban traffic: handling of traffic rules, knowledge of physical laws, a data structure representing the urban path network and the ability to navigate on it. Since all agents share the same environment, this component is common to all agents.

All basic mechanical features of an agent like weight, dimensions or maximum torque are included in a separate component. The component provides methods to calculate more advanced mechanical parameters like maximum acceleration or deceleration depending on the current speed and steering angle of the agent. These parameters are calculated using an interface to the physics engine NVIDIA PhysX¹. This allows for example to calculate the acceleration of a vehicle depending on the used torque, its weight and the ground's friction.

The human component contains a personality description based on the NEO five factor model (see [MC87]): A vector of five values rating the agent's neuroticism, extraversion, openness to experience, conscientiousness and agreeableness. From these values several behavior modification factors are calculated. For example the speed modification factor defines the preferred speed as a quotient of the maximum speed.

The correlation between the five factors and the behavior modifiers can be determined by mapping the behavior of a test person in an existing simulator to her or his personality values ascertained by a questionnaire. After taking more of such samples a mapping between the behavior of the test persons and their personality can be derived.

¹see <http://www.nvidia.com/physx>

2.2 Update layer model

The update process of the agents' position is divided into four layers of accuracy to reduce computational complexity economically.

On layer 3 the position of the agent is just estimated using the average speed for the current part of the agent's route and ignoring the influence other agents have to the agent's progress. Only the static structure of the traffic network is needed to perform this update. A layer 2 update includes a more accurate calculation of the average speed is performed regarding the speed of the currently leading agent. While on layer 3 and 2 the agents' speed is just estimated, on layer 1 a calculation regarding physical parameters is performed using the mechanical component and its interface to the physics engine. This way the agents' speed is influenced by other agents and the geometry of the current track. Finally layer 0 adds a some random modification to speed and steering to simulate a human error making the movement of the agents look more realistic.

The calculation of the desired speed and the steering angle – only performed on levels 1 and 0 – is inspired by [Rey99].

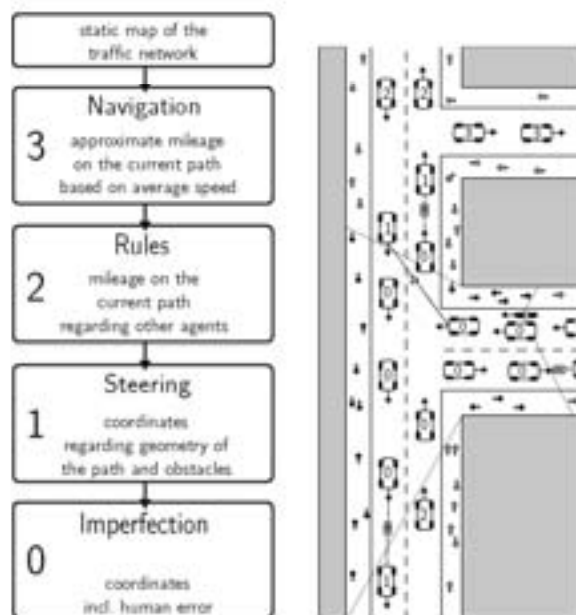


Figure 1: The layer model for agent updates in VETS

As shown in Figure 1 the layer on which an agent's position is updated, depends on its visibility to the user. Only agents directly visible to the user get a full accuracy update and human error. Agents whose position is needed to evaluate traffic rules of level 0 agents receive an update on level 1. Since an update on level 1 takes agents into account which are moving on the same path these agents are updated more accurate – on level 2 – then the rest which is considered for level 3 only.

3 Results & future work

Currently there is only a prototypic implementation of VETS which is not yet ready for evaluation. When this is solved VETS will be integrated into the FIVIS bicycle simulator (see [SBB⁺07] or <http://www.fivis.eu>) and a user study will be performed.

The current concept of character depended behavior makes the assumption that there is only a static personality influencing the behavior. In reality even very calm characters have aggressive moments from time to time which is considered by a system taking moods and emotions into account like described in [EKMT04] or [EM08].

References

- [EKMT04] Arjan Egges, Sumedha Kshirsagar, and Nadia Magnenat-Thalmann. Generic Personality and Emotion Simulation for Conversational Agents. *Computer Animation and Virtual Worlds*, 15(1):1–13, 2004.
- [EM08] Mirjam Palosaari Eladhari and Michael Mateas. Semi-Autonomous Avatars in World of Minds. 2008.
- [ER01] Patrick A. M. Ehlert and Leon J. M. Rothkrantz. Microscopic traffic simulation with reactive driving agents. In *ITSC 2001 2001 IEEE Intelligent Transportation Systems Proceedings (Cat No 01TH8585) ITSC-01*, page 860, 2001.
- [Hid05] Peter Hidas. Modelling vehicle interactions in microscopic simulation of merging and weaving. *Transportation Research Part C Emerging Technologies*, 13(1):37–62, 2005.
- [MC87] Robert R. McCrea and Paul T. Costa. Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, 52(1):81–90, 1987.
- [Rey99] Craig W. Reynolds. Steering Behaviors For Autonomous Characters. *Game Developers Conference*, pages 763–782, 1999.
- [RN03] Stuart Russel and Peter Norvig. *Artificial Intelligence - A Modern Approach*. Pearson Hall International, 2 edition, 2003.
- [SBB⁺07] Oliver Schulzyk, Jens Bongartz, Tobias Bildhauer, Ulrich Hartmann, Rainer Herpers, Bernd Goebel, and Dietmar Reinert. A bicycle simulator based on a motion platform in a virtual reality environment FIVIS project. In *Proceeding in Physics 114: Advances in Medical Engineering*, pages 323–328, Berlin, New York, 2007. Springer-Verlag.
- [YK96] QI Yang and Haris N. Koutsopoulos. A Microscopic Traffic Simulator for evaluation of dynamic traffic management systems. *Transportation Research Part C Emerging Technologies*, 4(3):113–129, 1996.

Intelligent Printing Technique Recognition and Photocopy Detection for Forensic Document Examination

Marco Schreyer, Christian Schulze, Armin Stahl and Wolfgang Effelsberg
mschreye@rumms.uni-mannheim.de
University of Mannheim and
German Research Center for Artificial Intelligence (DFKI)

Abstract: The detection of fraudulent documents is becoming an important issue in large scale office automation. In this context the recognition of a documents creation process is of valuable information in the examination of a questioned documents authenticity. We propose a novel method of printing technique recognition and photocopy detection based on statistical supervised machine learning. This is achieved by the classification of statistical document features that are obtained by an analysis of a given documents spatial and frequency domain. For the purpose of testing our methods we prepared a representative database of documents. Our results showed that (1) most of the documents are classified correctly (2) even at low scan resolutions.

GI-Topic: KI-BV (artificial intelligence - image understanding)

1 Introduction

The progress of digital printing and imaging technologies had a tremendous impact on the way we generate, publish and store information nowadays. This technological progress, as more and more applicable, is not only used for legitimate purposes but also for illegal activities. Especially in the case of banks, insurance companies and tax authorities, processing several thousand documents that are issued by a high number of invoicing parties each day. In such scenarios fully automatic document analysis systems are urgently needed to examine fraud and money laundry [FS07]. Important insights in a forensic document examination can be obtained by answering the questions: How was the suspected document at hand created? Is it an original or a photocopy? An opportunity to address these questions is offered by the appliance of techniques derived from digital image processing and statistical pattern recognition. Only a small number of publications concerning the classification of documents according to their creation process exists nowadays [LMB06, MAC⁺05, Tch04]. To the best of our knowledge none of them is concerned with the detection of photocopied documents.

We will show that photocopied, laser and inkjet printed documents can be distinguished even at low scan resolutions. This is achieved by the examination of the unique characteristics that correspond exclusively to each document creation technique.

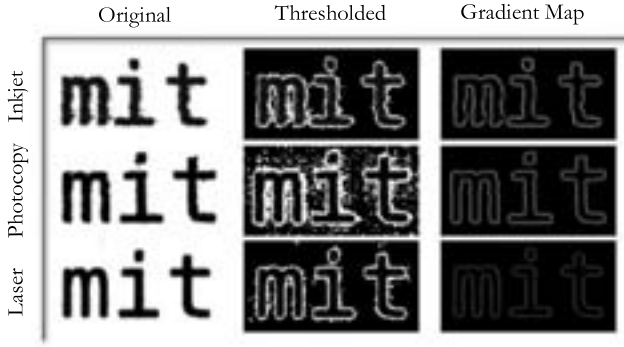


Figure 1: Representative document image extracts of an inkjet printed, laser printed and photocopied document scanned with a resolution of $2400dpi$.

2 Printing Processes and their Characteristics

Printing in a more general perspective can be seen as a “complex reproduction process” in which printing ink is applied to a printing substrate in order to transmit information in a repeatable form [Kip01]. The examination of high resolution scan images of photocopied, laser and inkjet printed documents reveals unique characteristics corresponding to each of these document creation techniques. These characteristics or document image degradations are the main source of evidence in document forensics and can be understood as a documents *“every sort of less-than-ideal properties”* or *“the departure of an ideal version”* [Bia00]. Observing the exemplary document image extracts presented in Figure 1, four discriminative characteristics can be identified: (1) image noise and artifacts, (2) character edge roughness, (3) character edge contrast and (4) uniformity of printed character area.

3 Statistical Feature Extraction

To distinguish different printing techniques and photocopies according to their unique document image characteristics as illustrated in Figure 1 we performed four different types of document analyzes.

Noise Analysis We obtain a documents image noise by applying mean, median and Gaussian filtering techniques as described in [GSW07]. As statistical features the mean and standard deviation are obtained from the noise image. Furthermore, the correlation and mean squared error was calculated from the original and denoised document image.

Gradient Analysis To capture statistical information about fine image intensity variations that correspond to character edges and noisy image regions we apply different gradient filters as proposed by [Tch04]. Subsequently, the gradient histogram is calculated for each document image. As statistical features we obtain the mean and standard deviation is obtained for different histogram intervals.

DCT Frequency Analysis High frequencies are denoted by large graylevel alterations within a small image area and low frequencies are denoted by large areas of nearly constant graylevel values. Therefore, we utilized the concept of frequency using "Discrete Cosine Transformation" (DCT) for the purpose of document classification. This is done by calculating DCT coefficients mean and standard deviation as statistical features at certain frequency subbands as described in [SSSB09].

Multiresolution Wavelet Analysis To achieve independence of the transformations base functions we also utilized Multiresolution Wavelet Analysis [GW07]. This was done using the Haar and Daubechies Wavelets as well as Coiflets. At different scales of wavelet decomposition we obtain the mean and standard deviation as statistical features for document classification.

4 Experimental Setup and Results

Since none of the reviewed document image databases¹ is currently providing a document annotation of the printing technique used in its creation the necessity emerged to create a new document image database annotated with the needed ground truth information. In German speaking countries a document called the "Grauert" letter, implementing the DIN-ISO 10561 standard, is used for the test of printing devices. The "Grünert" letter, which is derived from this document, yields the same results in printer tests. Because of its high similarity in layout and content to regular written business letters, we used the 'Grünert' letter as template for ground truth database creation. The created document database, consists of 49 different laser printouts, 14 different inkjet printouts as well as 46 photocopied documents. The variety of printing and photocopier device manufacturers covers all major brands typically present in (home) office environments, for example Hewlett-PackardTM, EpsonTM, CanonTM and RicohTM. To comply with the "No Free Lunch Theorem" [Wol01], the subsequent feature classification is performed on the basis of two supervised machine learning techniques: a (1) Multilayer Perceptron (MLP) and a (2) Support Vector Machine (SVM).

The experimental results illustrated in Figure 2 confirm our initial hypothesis that printing technique recognition and photocopy detection can be achieved by the analysis of discriminative information obtained from scanned document images. Utilizing both classification techniques remarkable classification results are achieved for documents scanned with resolutions of 400dpi and 800dpi. While the highest classification results 92.92% (99.08%) is achieved utilizing DCT Frequency Analysis and SVM classification at 400dpi (800dpi).

¹UW English Document Image Database I - III, Medical Article Records System (MARS), MediaTeam Oulu Document Database and Google 1000 Books Project

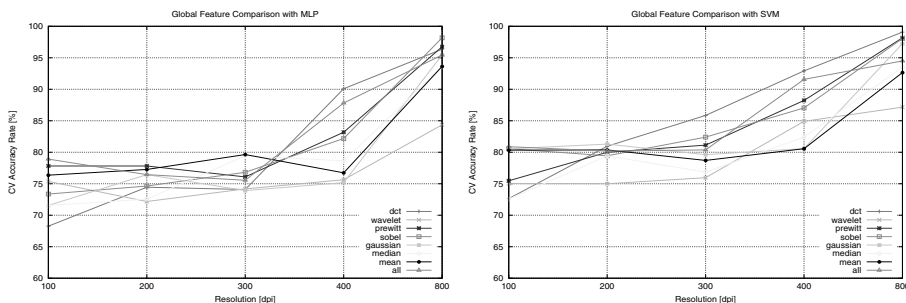


Figure 2: Feature evaluation accuracy results obtained by (left) multilayer perceptron and (right) support vector classification performing 10-fold stratified sampled cross validation.

References

- [Bia00] Henry S. Biard. The State of the Art in Document Image Degradation Modeling. In *Proceedings of the 4th IAPR Workshop on Document Analysis Systems*, pages 1–13, 2000.
- [FS07] Katrin Franke and Sargur N. Shrihari. Computational Forensics: Towards Hybrid-Intelligent Crime Investigation. In *Third International Symposium on Information Assurance and Security*, pages 383–386, 2007.
- [GSW07] Hongmei Gou, Ashwin Swaminathan, and Min Wu. Robust Scanner Identification based on Noise features. In *Proceedings of the SPIE International Conference on Security, Steganography and Watermarking of Multimedia Contents IX*, 2007.
- [GW07] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall International, 3 edition, 2007.
- [Kip01] Helmut Kipphan. *Handbook of Print Media*. Springer, 1 edition, 2001.
- [LMB06] Christoph H. Lampert, Lin Mei, and Thomas M. Breuel. Printing Technique Classification for Document Counterfeit Detection. In *Computational Intelligence and Security (CIS) 2006, Guangzhou, China*, 2006.
- [MAC⁺05] A.K. Mikkilineni, G.N. Ali, P.-J. Chiang, G.T.-C. Chiu, J.P. Allebach, and E.J. Delp. Printer identification based on graylevel co-occurrence features for security and forensic applications. In *Proceedings of the SPIE International Conference on Security, Steganography and Watermarking of Multimedia Contents VII*, volume 5681, pages 430–440, San Jose, CA, 2005.
- [SSSB09] Christian Schulze, Marco Schreyer, Armin Stahl, and Thomas M. Breuel. DCT Coefficient Analysis for Printing Technique and Document Copy Detection. In *Advances in Digital Forensics V (to appear)*. Springer, Boston, 2009.
- [Tch04] Jack Tchan. The development of an image analysis system that can detect fraudulent alterations made to printed images. In *Proceedings of the Photo-Optical Instrumentation Engineers (SPIE) Conference*, volume 5310, pages 151–159, 2004.
- [Wol01] David H. Wolpert. The supervised learning no-free-lunch Theorems. In *In Proc. 6th Online World Conference on Soft Computing in Industrial Applications*, pages 25–42. Springer-Verlag, 2001.

MetaXsA: Metadatenextraktion von externen Informationsquellen unter der Berücksichtigung vorhandener Taxonomien

Christian Meier

Universität Paderborn
Institut für Informatik
chrmeier@uni-paderborn.de

Jens Schäfer

Universität Paderborn
Institut für Informatik
jens21@uni-paderborn.de

Abstract: Dokumentenmanagement ist im Unternehmensbereich, wo mit einem großen Aufkommen von Dokumenten zu arbeiten ist, ein äußerst wichtiger und hilfreicher Bestandteil. Die in diesem Paper beschriebene Komponente MetaXsA ist Teil eines solchen Systems, welches im Rahmen einer Projektgruppe entstand, um automatisiert Metadaten aus Dokumenten zu extrahieren, und die in der Fortführung dieses Projekts erweitert wurde. Diese Erweiterungen beziehen sich auf die Einbeziehung zusätzlicher Informationsquellen wie E-Mails und Wiki-Artikel, die Unterstützung nichttextueller Dateien, Taxonomiezuordnungen, Gewichtungen von Schlüsselwörtern, sowie die Klassifizierung von E-Mails im Kontext eines Unternehmens.

Art der Arbeit: Projektgruppe

Betreuer: Wolfgang Reinhardt, Johannes Magenheim

1 Problemstellung

Im Rahmen der Projektserie „MoKEx“ (Mobile Knowledge Experience) entsteht ein System für Wissensmanagement nach dem Konzept des SPI („Single Point of Information“) [HHM⁺06]. Eine der Komponenten, die in diesem Zusammenhang entstanden sind, ist die Software „MetaXsA“, welche eine Extraktion von vorliegenden Metadaten [MS08] sowie eine semantische Analyse des Inhalts [SMR08] von unstrukturierten Textdokumenten durchführt. Die hierbei gewonnenen Metadaten und Semantiken werden von MetaXsA im standardisierten LOM-Format [LOM02] erstellt und dienen dazu, semantisch präzisere Treffer bei Suchanfragen zu erhalten, als dies bei einer Suche in einem hierarchischen Dateisystem der Fall wäre.

Im Rahmen der aktuellen Projektgruppe wurde diese Softwarekomponente um die Zuordnung von bei der semantischen Analyse gefundenen Schlüsselwörtern zu Taxonomien erweitert. Desweiteren wurde die Gewichtung der gefundenen Schlüsselwörter anhand der Dokumentstruktur sowie die Einbeziehung neuer Informationsquellen wie E-Mails und Wiki-Artikel in die Software integriert. Mit der Gewichtung von Schlüsselwörtern soll eine Verbesserung der Aussagekraft von gefundenen Schlüsselwörtern erreicht werden.

Die Taxonomiezuordnung soll dazu dienen, analysierte Dokumente in Kategorien einzuordnen, um bei späteren Suchanfragen ähnliche Dokumente finden zu können. Zusätzlich werden in einer separaten Komponente (siehe RaMBo [KBR09]) Relationen zwischen Schlüsselwörtern und Taxonomien gespeichert und ausgewertet.

2 MetaXsA - Metadatenextraktion und semantische Analyse

Der grundlegende Ablauf bei der Verarbeitung eines Dokuments durch MetaXsA besteht aus zwei aufeinanderfolgenden Schritten (vgl. Abbildung 1). Zunächst findet eine *Metadatenextraktion* statt (vgl. [MS08]), bei dem in Dokumenten bereits statisch vorliegende Metadaten durch den Einsatz von Metadatenextraktionswerkzeugen ausgelesen werden. Anschließend erfolgt die *semantische Analyse*, deren Ziel es ist, unter Verwendung des GATE-Framework [CMBT07], einer Software zur Verarbeitung natürlichsprachlicher Texte, weitere Metadaten aus dem textuellen Inhalt des Dokuments zu erhalten (vgl. [SMR08]).

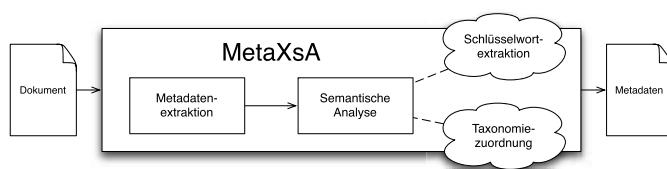


Abbildung 1: MetaXsA Ablauf

2.1 Einbindung externer Informationsquellen

Da der Inhalt einer E-Mail oder eines Wiki-Artikels aus dem eigentlichen Text, sowie evtl. vorhandener Anhänge bzw. eingebetteter Medien bestehen kann, wurde MetaXsA um eine Verarbeitung von Verbunddokumenten erweitert. Während MetaXsA bisher zu einem Dokument genau einen Satz von Metadaten lieferte, so liefert MetaXsA in der neuen Version eine Menge von Metadaten, darunter die Metadaten aus dem Hauptdokument, sowie Metadaten evtl. vorhandener Anhänge bzw. eingebetteter Medien.

Damit externe Informationsquellen durch MetaXsA analysiert werden können, müssen diese zuvor von einer Komponente angefordert und in einem einheitlichen Format zu MetaXsA transportiert werden. Für E-Mails werden deren Header, Text, sowie evtl. vorhandene Anhänge erwartet, bei Wiki-Artikeln sind dies der HTML-Text, zusätzliche Metadaten wie Autor und Änderungsdatum des Artikels, sowie eingebettete Medien.

Die Metadatenextraktion wurde erweitert, so dass die Header-Informationen aus E-Mails ausgelesen werden können um beispielsweise Informationen über Absender und Empfänger zu erhalten. In Kombination mit der Anbindung an ein Benutzerverzeichnis eines Unternehmens können aus den extrahierten E-Mail-Adressen die zugehörigen Personen aufgefunden und zugeordnet werden.

Die semantische Analyse wurde für die neuen Anforderungen an mehreren Stellen erweitert. Zum einen kann eine semantische Analyse nun auch für nichttextuelle Dokumente wie z.B. Bilder durchgeführt werden, um daraus Schlüsselwörter zu extrahieren, wenn zusätzlich zu dem Objekt ein beschreibender Text übergeben wird.

2.2 Gewichtung von Schlüsselwörtern anhand der Dokumentstruktur

Das Modul zur Extraktion von Schlüsselwörtern wurde um eine Funktionalität zur Gewichtung erweitert. Das Gewichtungsschema ist konfigurierbar und kann in MetaXsA für jedes unterstützte Dokumentenformat hinterlegt werden. So kann konfiguriert werden, dass Schlüsselwörter, welche im Betreff einer E-Mail auftauchen, mit einem anderen Faktor gewichtet werden, als Schlüsselwörter aus dem Haupttext. Ferner kann auch die ursprüngliche Dokumentenstruktur bei der Gewichtung berücksichtigt werden. So kann konfiguriert werden, diejenigen Schlüsselwörter, welche in bestimmten HTML-Tags vorkommen, höher zu gewichten als andere. Auch Schlüsselwörter aus Überschriften in Wiki-Artikeln können so mit einer Gewichtung versehen werden. Allerdings ist dabei zu berücksichtigen, dass verschiedene Wiki-Systeme unterschiedliche HTML-Tags für die Formatierung verwenden. Da MetaXsA Wiki-Artikel in einem einheitlichen Format erwartet¹, kann auf eine Wiki-spezifische Implementierung verzichtet werden.

2.3 Zuordnung zu Taxonomien

Das Modul zur Taxonomiezuordnung wurde neu implementiert. Für dieses Modul können verschiedene Taxonomien hinterlegt werden, darunter eine Haupttaxonomie, sowie Taxonomien für verschiedene Projekte. Für die Taxonomiezuordnung wird eine Liste mit Schlüsselwörtern aus dem Dokument erwartet, welche im Modul der Schlüsselwortextraktion gewonnen wird. Anschließend werden die gefundenen Schlüsselwörter in den Taxonomien aufgesucht. Die Obergrenze für die Anzahl der Taxonomiepfade, welche einem Dokument zugeordnet werden können ist konfigurierbar, es bietet sich an, diese möglichst gering zu halten. Sollten mehr Taxonomiepfade gefunden werden, so wird versucht, die gefundenen Pfade zu vereinen, indem nach gleichen Vorgängern gesucht und die Taxonomieknoten entsprechend zusammengefasst werden.

In der gegenwärtigen Implementierung werden verschiedene Taxonomien verwendet, um private Kommunikation von geschäftlicher zu unterscheiden. Bei geschäftlicher Kommunikation findet auf Grundlage von projektspezifischen Taxonomien eine automatische Zuordnung zu Projekten statt. Für diese Zuordnung findet ein Vergleich statt, bei dem überprüft wird, welche dieser Taxonomien die größte Anzahl der Schlüsselwörter enthält. Außerdem muss in der jeweiligen Taxonomie ein Minimum an Schlüsselwörtern gefunden werden, welches prozentual gemessen an der gesamten Anzahl von gefundenen Schlüssel-

¹Die Überführung in dieses einheitliche Format ist die Aufgabe einer Komponente, welche neue Artikel in einem Wiki erkennt und dem System zur Verfügung stellt.

wörtern in der konfigurierbaren Metrik festgelegt ist. Bei dem Vergleich ist zu unterscheiden, ob die maximale Treffermenge eindeutig ist oder nicht. Ist das Maximum eindeutig, so wird das Dokument der Projekttaxonomie mit den meisten Treffern zugeordnet. Ist dies nicht der Fall, d.h. es wurden zu verschiedenen Taxonomien identisch viele Vorkommen gefunden, so kann keine Projektzuordnung vorgenommen werden.

3 Zusammenfassung und Ausblick

MetaXsA wurde erweitert, um zusätzlich zu den bisher unterstützten Dokumentenformaten auch externe Informationsquellen wie E-Mails und Wiki-Artikel einzubinden. Außerdem wurde die Möglichkeit geschaffen, Dokumente Taxonomien zuzuordnen, sowie Schlüsselwörter aus Texten der Formate HTML oder XML zu gewichten.

Durch den modularen Aufbau ist eine Erweiterung durch neue semantische Module auch in Zukunft gewährleistet, denkbar ist die Unterstützung von Dokumenten aus weiteren Informationsquellen wie z.B. RSS-Feeds oder Online-Kalender.

Literatur

- [CMBT07] Hamish Cunningham, Diana Maynard, Kalina Bontcheva und Valentin Tablan. Developing Language Processing Components with GATE Version 4. <http://gate.ac.uk/sale/tao/tao.pdf>, 2007.
- [HHM⁺06] Knut Hinkelmann, Kai Holzweißig, Johannes Magenheimer, Fabian Probst und Wolfgang Reinhardt. Linking Communities of Practice with Learning Communities in Computer Science Education. In *Proceedings of the 19th IFIP World Computer Congress (WCC2006)*, Santiago de Chile, August 20–25 2006.
- [KBR09] Andreas Kohring, Alexander Boschmann und Wolfgang Reinhardt. RaMBo: Finden von Experten in Projekten durch Profiling und Verhaltensbewertung. In *eingereicht bei den Informatiktagen 2009*, 2009.
- [MS08] Christian Mletzko und Benedikt Schmidt. Automatische Extraktion von Metadaten aus verschiedenen Dateiformaten. In Gesellschaft für Informatik, Hrsg., *Informatiktage 2008. Fachwissenschaftlicher Informatik-Kongress, 14. und 15. März 2008, B-IT Bonn-Aachen International Center for Information Technology in Bonn*, Jgg. S-6 of *Lecture Notes in Informatics*, Seiten 181–184, 2008.
- [LOM02] Learning Technology Standards Committee Learning Standards Committee of the IEEE. Draft Standard for Learning Object Metadata. http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf, 2002.
- [SMR08] Benedikt Schmidt, Christian Mletzko und Wolfgang Reinhardt. MetaXsA - Modulares Framework zur semantischen Dokumentanalyse für unstrukturierte Textdokumente. In Gesellschaft für Informatik, Hrsg., *Informatiktage 2008. Fachwissenschaftlicher Informatik-Kongress, 14. und 15. März 2008, B-IT Bonn-Aachen International Center for Information Technology in Bonn*, Jgg. S-6 of *Lecture Notes in Informatics*, Seiten 197–200, 2008.

Machine Learning Verfahren zur Klassifikation von kurzen Textstücken

Benedikt Waldvogel
Hochschule Furtwangen – Fakultät Informatik
mail@bwaldvogel.de

Abstract: Ein häufiger Anwendungsfall von automatischen Klassifizierern ist das Filtern von unerwünschten Nachrichten (Spam). In der Diplomarbeit wurde ein unternehmensspezifischer Anwendungsfall untersucht, um mit modernen Machine Learning Verfahren *kurze* Textstücke automatisch zu klassifizieren. Ein Teil der Ergebnisse wird mit einem ähnlichen Anwendungsfall (Klassifikation von Emails anhand des Betreffs) vorgestellt.

1 Einführung

Die Klassifikation von Textstücken wurde in den letzten Jahren intensiv untersucht. Ein häufiger Anwendungsfall ist das Filtern von unerwünschten Email Nachrichten (Spam), bei dem sich vor allem statistische Klassifikationsverfahren etabliert haben [CHS07]. Bevor ein statistischer Klassifizierer eingesetzt werden kann, muss er zunächst trainiert werden. Dazu werden bereits bekannte Nachrichten verwendet, die ein Mensch zuvor als erwünscht (Ham) bzw. unerwünscht (Spam) eingeordnet hat.

Bei der Klassifikation von Emails steht in der Regel eine Wissensbasis mit tausenden Nachrichten und jeweils einigen Dutzend Wörtern zur Verfügung. Ein schwierigeres Problem ist die Klassifikation von sehr kurzen Texten, wobei jeweils nur einige Schlagworte oder ein (unvollständiger) Satz zur Verfügung steht. Im Rahmen der Diplomarbeit wurde ein unternehmensspezifischer Anwendungsfall zur Klassifikation von kurzen Textstücken untersucht.

Viele Lernverfahren können beschleunigt werden, indem lediglich aus den Änderungen in der Wissensbasis seit dem letzten Training gelernt wird (inkrementelles Lernen). Da die untersuchten Daten jedoch eine hohe Volatilität aufweisen, musste inkrementelles Lernen ausgeschlossen werden. Bereits eine Berechnung aller Änderungen würde zu viel Zeit in Anspruch nehmen. Stattdessen wurden ausschließlich Verfahren betrachtet, die in angemessener Zeit mit der *gesamten* Wissensbasis (mehrere hunderttausend Dokumente) trainiert werden können.

Für dieses Papier wurde stellvertretend das Filtern von Emails anhand der Betreffs untersucht, da Betreffzeilen typischerweise auch weniger als ein Dutzend Worte enthalten [CHS07] und bereits große Korpora öffentlich zur Verfügung stehen.

2 Klassifikationsverfahren: Support Vector Machines

Ein verbreitetes Verfahren um zwei-Klassen Probleme zu lösen, sind Support Vector Machines (SVMs). Sie zählen im Bereich der Textklassifikation zu den Verfahren mit den höchsten Genauigkeiten [SW07, CHS07, Joa98], was bei einem Vergleich mit anderen Verfahren im Rahmen der Diplomarbeit bestätigt werden konnte.

Um ein statistisches Klassifikationsverfahren zu trainieren, werden zunächst Merkmale aus jedem Textdokument extrahiert. Wie die Merkmale gewonnen werden, ist stark vom Anwendungsfall abhängig, wobei von Textdokumenten häufig jedes Wort als Merkmal verwendet wird.

Bei großen Korpora (d.h. vielen Dokumenten) strebt die Dimensionalität der Vektoren gegen die Anzahl der Wörter, im gesamten Wortschatz der zu Grunde liegenden Sprache.

Zur Veranschaulichung kann man sich jedes Dokument als Punkt im Hyperraum vorstellen. Eine Support Vector Machine versucht dazu eine Hyperebene zu finden, sodass sich möglichst alle Punkte einer Klasse oberhalb, und alle anderen Punkte unterhalb der Ebene befinden (siehe Abbildung 1).

Eine gute *Generalisierung* wird erreicht, indem eine Maximierung des Abstands zwischen Punkten und Ebene angestrebt wird. Das daraus entstehende Optimierungsproblem [Wikb] kann bei mehreren (hundert-)tausend Textdokumenten allerdings sehr komplex werden.

Für die Diplomarbeit wurden Support Vector Machine Implementierungen untersucht, die das Problem in kurzer Zeit lösen, sodass auch aus einem sehr großen Korpus gelernt werden kann. Als besonders leistungsfähige Implementierungen haben sich dabei Liblinear¹ [FCH⁺08] und Stochastic Gradient Descent (SGD) SVMs [Bor] herausgestellt, die vor allem zur linearen Klassifikation² geeignet sind.

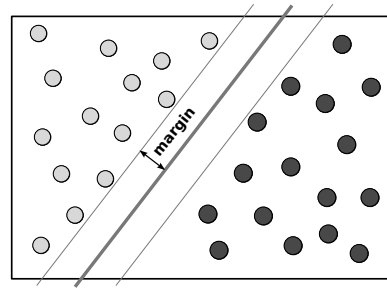


Abbildung 1: Geometrische Veranschaulichung einer Hyperebene im zweidimensionalen Merkmalsraum

¹Liblinear wurde im Rahmen der Arbeit in Java übersetzt. <http://bwaldvogel.de/liblinear-java/>

²bei linearen SVMs werden keine sogenannten Kernelfunktionen verwendet, um die Dimensionalität zu erweitern

3 Orthogonal Sparse Bigrams

Bei einer naiven Merkmalsgewinnung wird die Reihenfolge der Merkmale nicht berücksichtigt, sodass die Beziehung zwischen den Worten verloren geht. Beispielsweise werden “buildings in New York” und “New buildings in York” auf den gleichen Merkmalsvektor abgebildet, obwohl sich die Bedeutung stark unterscheidet.

Mit Orthogonal Sparse Bigrams (OSB) wurde bereits 2004 ein Verfahren zur Merkmalskombination vorgestellt, um Wort-Bigramme effizient auf Merkmalsvektoren abzubilden. Es wird dabei jeweils ein Fenster über die letzten n Worte im Text verschoben, und alle linear unabhängigen Bigramme als Merkmal aufgenommen (siehe [SACY04]). Aus dem Satzteil “buildings in New York” könnten beispielsweise die folgenden Merkmale extrahiert werden:

`buildings*in`, `buildings**new`, `in*new`, `new*york` sowie `in**york`.

Bei den Untersuchungen im Rahmen der Arbeit hat sich gezeigt, dass es bei sehr kurzen Textstücken unerlässlich ist, zusätzlich alle *Monogramme* aufzunehmen.

4 Ergebnisse

Um die Klassifikationsgenauigkeit zu beurteilen, wurde eine 10-fache Kreuzvalidierung [Wika] durchgeführt, wobei als Korpus die 75419 frei verfügbaren Emails des TREC 2007 Corpus [TRE07] verwendet wurden. Aus den 25220 Ham sowie 50199 Spam Emails wurden lediglich die *Betreffzeilen* für die Merkmalsextraktion verwendet.

Die Ergebnisse zeigen, dass mit Liblinear und Stochastic Gradient Descent Support Vector Machines (SvmSgd) in sehr kurzer Trainingszeit eine hohe Klassifikationsgenauigkeit erzielt werden kann. Mit Orthogonal Sparse Bigrams konnte die Fehlerrate außerdem um knapp 20 % gesenkt werden (siehe Tabelle 1).

	Fehlerrate	Trainingszeit
Liblinear	1.29 %	1.17 s
Liblinear + OSB-3 ¹	1.16 %	5.67 s
Liblinear + OSB-4	1.20 %	7.26 s
SvmSgd ²	1.32 %	0.042 s
SvmSgd ² + OSB-3	1.07 %	0.077 s
SvmSgd ² + OSB-4	1.12 %	0.083 s

Tabelle 1: Vergleich von Liblinear sowie SvmSgd in Kombination mit OSB

¹Orthogonal Sparse Bigrams mit Fenstergröße 3

²Mit dem Parameter $\lambda = 10^{-5}$

Literatur

- [Bor] A. Bordes. Stochastic Gradient Descent Support Vector Machine. <http://leon.bottou.org/projects/sgd>, [Online; accessed 07-January-2009].
- [CHS07] G.V. Cormack, J.M.G. Hidalgo und E.P. Sánz. Spam filtering for short messages. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, Seiten 313–320. ACM New York, NY, USA, 2007.
- [FCH⁺08] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang und Chih-Jen Lin. LIBLINEAR: A Library for Large Linear Classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008.
- [Joa98] T. Joachims. *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*. Springer, 1998.
- [SACY04] C. Siefkes, F. Assis, S. Chhabra und W.S. Yeraunis. Combining Winnow and Orthogonal Sparse Bigrams for Incremental Spam Filtering. *LECTURE NOTES IN COMPUTER SCIENCE*, Seiten 410–421, 2004.
- [SW07] D. Sculley und G.M. Wachman. Relaxed online SVMs for spam filtering. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, Seiten 415–422. ACM Press New York, NY, USA, 2007.
- [TRE07] TREC Public Spam Corpus, 2007. <http://plg.uwaterloo.ca/~gvcormac/treccorpus07/>, [Online; accessed 07-January-2009].
- [Wika] Cross-validation — Wikipedia, The Free Encyclopedia. <http://en.wikipedia.org/w/index.php?title=Cross-validation&oldid=251150112>, [Online; accessed 07-January-2009].
- [Wikb] Support vector machine — Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=Support_vector_machine&oldid=258753699, [Online; accessed 07-January-2009].

Komponentenbasierte Software-Entwicklung mit WebObjects

Dipl.-Inf.(FH) Malte B. Blanken

Fachhochschule Osnabrück
Fakultät Ingenieurwissenschaften und Informatik
m.blanken@fh-osnabrueck.de

Art der Arbeit: Überarbeiteter Auszug aus einer Diplomarbeit
Betreuer der Arbeit: Prof. Dr. Frank M. Thiesing

1 Einleitung

Die Software-Entwicklungs-Branche sieht sich permanent mit dem Problem konfrontiert, in immer kürzerer Zeit maximal fehlerfreie Software produzieren zu müssen. Dabei soll diese oftmals auch sehr gut wartbar, leicht austausch- und erweiterbar sein. Als Reaktion auf diese Situation wurden zunehmend Möglichkeiten der Wiederverwendung von Software gesucht. Denn wenn bei der Entwicklung eines neuen (Software-) Produktes bereits vorhandene, getestete Bestandteile wiederverwendet werden können, dann ist eine kürzere Entwicklungszeit und ein fehlerfreieres Ergebnis zu erwarten. So kamen in den 60er Jahren die ersten objektorientierten Programmiersprachen auf, welche allerdings noch nicht zum gewünschten Grad der Wiederverwendung führten (vgl. [So01], S. 310). Daraufhin wurde zunehmend der Ansatz einer komponentenweisen Software-Entwicklung verfolgt. Sowohl von kommerziellen Software-Unternehmen als auch im Rahmen von Forschungs- und Open Source-Projekten wurde mittlerweile eine große Anzahl komponentenbasierter Technologien entwickelt – eine von diesen ist WebObjects¹. Eine vergleichbare aber zur Zeit deutlich bekanntere Technologie ist die Java Platform Enterprise Edition², kurz Java EE. Im Folgenden soll untersucht werden, in wiefern sich die beiden Technologien ähneln, worin sie sich unterscheiden und wo ihre relativen Schwächen und Stärken liegen.

2 WebObjects versus Java EE

Bevor die Unterschiede der zwei Technologien betrachtet werden, sollen hier zunächst einige Gemeinsamkeiten aufgezeigt werden:

- Programmiersprache: Java

¹Eine Einführung zu WebObjects bietet beispielsweise die Web-Seiten [Appl], aber auch [OSWi].

²Ein Überblick über Java EE bietet [Sun].

- XML-basiertes Deployment (mit Apache Ant)
- umfassende Unterstützung von Web Services
- ORM
- Erstellung von Web-Anwendungen und Java-Client-Anwendungen
- kostenlos (unter Berücksichtigung der jeweilige Lizenzierungen)
- kostenlose Entwicklungswerkzeuge³

Beide Technologien ermöglichen die Erstellung von Web- und Client-Server-Anwendungen. Die logischen Schichten der verschiedenen WebObjects- bzw. Java EE-Anwendungen zeigt Abbildung 1.

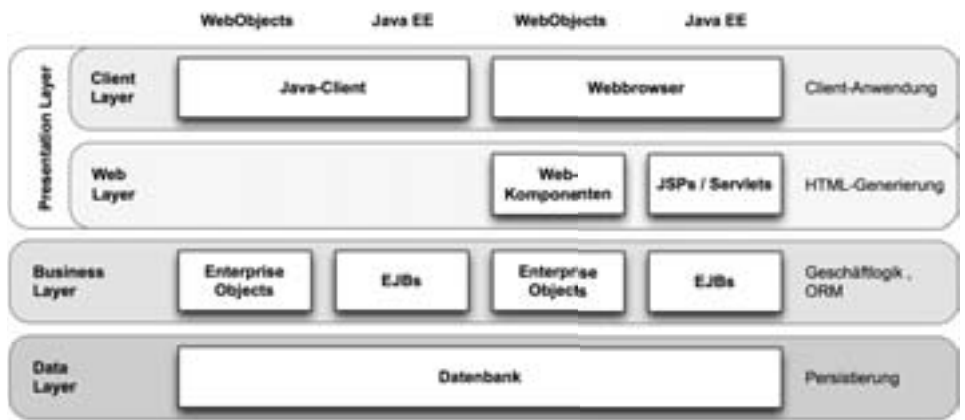


Abbildung 1: Logische Schichten von WebObjects- bzw Java EE- Anwendungen

Im Folgenden sollen nun Web-Komponenten mit JSPs und anschließend Enterprise Objects mit EJBs verglichen werden.

2.1 Web-Komponenten versus JSPs

Der Vergleich von Web-Komponenten mit JSPs liegt nahe, da beide den gleichen Zweck haben: Sie sollen auf HTTP-Anfragen reagieren, indem sie eine HTML-Seite generieren, welche an den aufrufenden Client zurückgesendet wird. Bei der Generierung des HTML-Codes werden bereits fertige HTML-Abschnitte mit dynamisch generiertem Inhalt kombiniert. Es können aber unter anderem die folgenden vier prinzipiellen Unterschiede ausgemacht werden:

Web-Komponente (WebObjects)	JSP (Java EE)
Eine eigene .java-Datei pro Web-Komponente.	Beliebig viele JavaBeans Components mit

³Die IDE für die Entwicklung von WebObjects-Software ist die Eclipse IDE for Java Developers mit dem Plug-in WOLips. Weitere Informationen bieten die Community-Wikis [OSWi].

	beliebig vielen JSPs.
Standard HTML plus ein zusätzliches Tag.	Eine Vielzahl von speziellen Tags und anderen Erweiterungen (Expressions, Skripting, JSTL, Custom Tags, ...)
Komponentenschnittstellenbeschreibung in komponenteneigener .api-Datei.	(Attribute Directives in einer tag-Datei.)
Klare Trennung von Logik (Java-Code) und Präsentation (HTML) mittels .wod-Datei.	Vermengung von Logik und Präsentation im JSP-Code.

Ein nicht zu unterschätzender Vorteil für WebObjects ist, dass sämtliche Angaben, die der Verknüpfung von .html-, .wod- und .java-Dateien einer Web-Komponente dienen, von der Entwicklungsumgebung geprüft werden: Wird im HTML-Code einem <webobject>-Tag ein Name gegeben, der nicht in der .wod-Datei aufgelöst wird, so wird das als Fehler markiert. Wird in der .wod-Datei einem Binding eine Methode zugewiesen, die nicht in der Java-Klasse implementiert ist, so wird dies ebenfalls als Fehler markiert. Die vergleichbaren Zuordnungen bei der Arbeit mit JSP werden mittels Zeichenketten vorgenommen, welche erst zur Laufzeit (bei der Generierung des Servlets) aufgelöst werden. JSP-Entwicklungsumgebungen wie NetBeans bieten zwar eine Validierungsfunktion, mit welcher Fehler an solchen Stellen registriert werden, die Validierung muss allerdings explizit aufgerufen werden und dauert relativ lange.

2.2 Enterprise Objects versus EJBs

Enterprise Objects und EJBs (Entity Beans) dienen in erster Linie dazu Inhalte relationaler Datenbanken als Objekte zu repräsentieren. Neben dem Zugriff auf die Daten und Möglichkeiten diese zu manipulieren, können Enterprise Objects und EJBs aber auch beliebige weitere Methoden (für die Arbeit mit diesen Daten) zur Verfügung stellen.

In der Version 2 der Java Enterprise Edition (J2EE) war der Unterschied zwischen diesen EJBs und Enterprise Objects recht groß – insbesondere bezüglich der Komplexität. Vieles hat sich mit Java EE 5 bzw. der EJB-Version 3 und der Einführung der Java Persistence API geändert. Unter anderem unterstützt nun auch der Java EE-Ansatz Vererbung von Entitäten. Besonders hervorzuheben ist sicherlich, dass Entity Beans, genau wie Enterprise Objects, nunmehr POJOs sind und nicht mehr ausschließlich innerhalb eines EJB-Containers verwendet werden können. Insgesamt scheint der Unterschied zwischen Enterprise Objects und Entity Beans immer kleiner zu werden, wie auch schon [HM04] (S. 274) feststellten.

3 WebObjects mit Java EE

Ohne Zweifel haben beide Technologien ihre Stärken und Schwächen, weshalb es umso erfreulicher ist, dass sie auf verschiedene Weisen kombiniert werden können.

Da sowohl WebObjects als auch Java EE die Möglichkeit bieten, Dienste in Form von Web Services anzubieten und zu konsumieren, besteht hierin eine wohldefinierte Schnittstelle für die Interaktion zwischen diesen beiden Plattformen. Spannender ist jedoch die Integration von WebObjects in Java EE.

JSPs und die aus ihnen generierten Servlets benötigen einen entsprechenden Container als Laufzeitumgebung. Anders verhält es sich mit WebObjects-Web-Anwendungen, denn diese laufen je als eigenständiger Prozess in einer Java-VM. Mit relativ geringem Aufwand ist es dennoch möglich WebObjects-Anwendungen in einem Servlet-Container zu nutzen. Üblicherweise läuft die Kommunikation mit einer WebObjects-Web-Anwendung über einen sogenannten HTTP-Adaptor. Soll sie aber in einem Servlet-Container verwendet werden, so läuft die Kommunikation dort über ein Servlet, welches sich für die Anwendung ganz ähnlich verhält, wie sonst der HTTP-Adaptor (vgl. [HM04], S. 275).

4 Fazit

Was den geringen Bekanntheitsgrad von WebObjects angeht, so kann dieser wohl nur mit mangelndem Marketing und einem über lange Jahre sehr hohen Preis (\$50.000 bis zum Jahre 2000) erklärt werden. Denn Funktionsumfang, Geschwindigkeit, Skalierbarkeit und auch die Unterstützung durch die Entwicklungswerkzeuge lassen etwas anderes erwarten. Zweifelsohne gibt es Bereiche in denen andere Technologien für Software-Komponenten (z.B. CORBA/CCM, .NET) deutlich zu bevorzugen sind. Auf seinem Gebiet, also Web- und Enterprise-Anwendungen, sucht WebObjects, insbesondere was die Einfachheit der Software-Entwicklung angeht, jedoch seinesgleichen. Auch wenn Java EE an manchen Stellen mehr Möglichkeiten bietet als WebObjects, so ist das sicherlich bei einer großen Anzahl von Software-Entwicklungen kein ausreichender Grund dafür, auf die deutlichen Vorteile von WebObjects zu verzichten.

Literaturverzeichnis

- [Appl] Apple Inc. (Hrsg.): Introduction to WebObjects Overview, 2007: http://developer.apple.com/documentation/WebObjects/WebObjects_Overview/Introduction/chapter_1_section_1.html, am 14.01.2009
- [OSWi] Community-Wikis zu den Themen WebObjects, WOLips u.a.: <http://wiki.objectstyle.org>, am 14.01.2009
- [HM04] Hill, Charles; Mallais, Sacha: Practical WebObjects, Apress, Berkeley, 2004
- [McCr] McCraw, Steven Mark; u.a.: EOF-Overview: <http://wiki.objectstyle.org/confluence/display/WO/EOF-Overview>, am 14.01.2009
- [So01] Sommerville, Ian: Software Engineering, Addison Wesley, Harlow u.a., 2001
- [Sun] Sun (Hrsg.): Java EE at a Glance: <http://java.sun.com/javaee/>, am 14.01.2009

Enterprise Content Management im SAP NetWeaver Portal

CMS-Integration und Multi-Site-Management

am Beispiel des Hessenportals

Harald Diehl

Wilhelm Büchner Hochschule
Fachbereich Informatik
h.diehl@saarbase.de

Art der Arbeit: Diplomarbeit

Betreuer/in der Arbeit: Prof. Dr. Christian Wagenknecht

Abstract:

Ziel dieser Diplomarbeit war eine verbesserte CMS-Integration und die Realisierung des Multi-Site-Managements im Hessenportal.

Die Arbeit hat gezeigt, dass mit dem Multi-Site-Management eine portalbasierte Mandanten- und Auftrittsverwaltung mit einer klaren Trennung der fachlichen und technischen Aufgaben realisiert werden kann. Ebenso ist eine stärkere Integration des CMS und der Fachanwendungen durch die Erstellung und Verwaltung einer Metanavigation möglich. Darüber hinaus erlaubt die Fähigkeit für Portalseiten unterschiedliche Quellen zusammenzufassen, eine stärkere Personalisierung für den Anwender.

Der Einsatz des Multi-Site-Managements im Hessenportal wird zu einer signifikanten Reduzierung der Betriebskosten führen, da eine deutlich vereinfachte und kostengünstigere Auftrittserstellung und -verwaltung durch fachlich geschulte Redakteure möglich wird. Die anwenderfreundlich gestaltete Benutzeroberfläche führt zu einer höheren Akzeptanz des Hessenportals bei den Ressorts und Dienststellen.

In der nächsten Phase der Geschäftsprozessintegration kann sich das Multi-Site-Management zu einem zentralen Dienst in einer Serviceorientierten Architektur weiterentwickeln.

1 Enterprise Portale und Enterprise Content Management

„Öffentliche Verwaltungen benötigen heute Portale, die auf Rollen und Arbeitsweisen der einzelnen Benutzer zugeschnitten sind und verwaltungsspezifische Anforderungen erfüllen.“¹

1.1 Enterprise Portale

Bei der Entwicklung von Portalen lassen sich Phasen unterscheiden. Die Differenzierung der Phasen erfolgt dahingehend, was mit den Portalen beabsichtigt wird, an wen sie sich richten, welche Inhalte und Funktionen sie bieten und welche Technologien zum Einsatz kommen. Die einzelnen Phasen kennzeichnen den Reifegrad des Portals. Dieser steigt, je mehr ein Portal in der Lage ist, Geschäftsprozesse vollständig abzubilden und zu unterstützen.²

Erst mit Prozessportalen lassen sich die Möglichkeiten der Verwaltungs- und Dienstleistungsrationalisierung vollständig ausschöpfen. Was sich aus Sicht der Verwaltungen als Rationalisierungsgewinn darstellt, bedeutet für die Zielgruppen von Portalen - die Bürger, Unternehmen und Mitarbeiter - eine erhebliche Steigerung der Dienstleistungsqualität.

1.2 Enterprise Content Management und Multi-Site-Management

Ein Enterprise Content Management System (ECMS) dient im Kontext dieser Arbeit der Erstellung, Verwaltung und Publikation von Informationen in einem Portalumfeld. Dabei kann es sich um Webinhalte, digitale Dokumente, Bilder oder multimediale Daten handeln. Ein ECMS stellt damit in allen Entwicklungsphasen eines Enterprise Portals eine essentielle Komponente dar. Die Verwaltung der Inhalte - von der Erstellung bis zur Archivierung - wird als Content Lifecycle bezeichnet.

Portale verfügen selten über eine leistungsfähige CMS-Komponente. Daher ist häufig die Anbindung eines externen CMS erforderlich. Weiterhin strukturieren Portale Webinhalte nicht nach Auftritten (Sites), sondern präsentieren Inhalte rollenbasiert für den jeweiligen Anwender. Wird ein Portal allerdings im Internet eingesetzt, ist gerade diese Auftrennung nach Auftritten erforderlich, um die komplexe Organisationsstruktur eines Unternehmens oder eines Bundeslandes fachlich und technisch abbilden zu können.

Erforderlich ist eine nahtlose Integration der CMS Navigation in die SAP Portal Navigation. Dabei werden redaktionell erstellte Seiten mit den Portalseiten kombiniert und somit eine übergreifende Navigationsstruktur (Metanavigation), die sowohl Inhalte als auch Anwendungen miteinander verbindet, gestaltet.

¹ [SAP04] SAP, 2004, S. 4

² [GRO05] Großmann/Koschek, 2005, S. 28

Die Verwaltung vieler Webauftritte (Sites) für mehrere Mandanten - Ressorts und Dienststellen - in einer einheitlichen Portallandschaft wird als Multi-Site-Management bezeichnet. Damit lassen sich z.B. themenspezifische oder mehrsprachige Auftritte innerhalb des gleichen Mandanten oder auch mandantenübergreifend abbilden.

2 Semantische Modellierung der Anwendungsarchitektur

2.1 Use Cases und Use-Case-Model

Use Cases definieren die von einem Benutzer erwarteten Leistungen einer Anwendung. Ein Use Case ist eine über mehrere Schritte verlaufende Interaktion zwischen einem oder mehreren Akteuren und der Anwendung, um eine konkrete fachliche Aufgabe zu lösen.

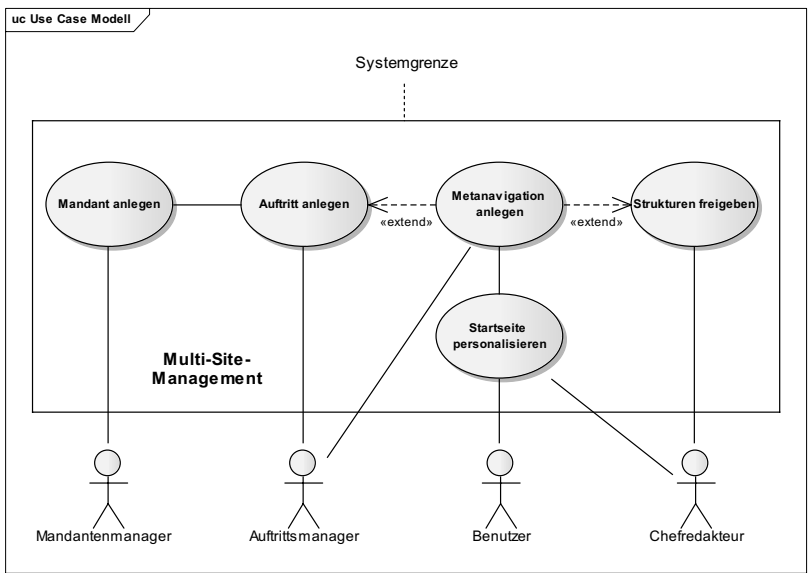


Abbildung 1: Use-Case-Model (Teilsicht „Objekte anlegen und freigeben“)

2.2 Entity-Relationship-Modell

Im folgenden ERM werden einige fachliche Objekte definiert, die im Hessenportal entweder im Portalumfeld oder in ihrer Bedeutung nicht durchgängig vorhanden sind. Insbesondere sind dies Mandanten, Auftritte und die Metanavigation.

Mandanten sind für einen oder mehrere Auftritte verantwortlich, Auftritte wiederum können von einem oder mehreren Mandanten verantwortet werden. Auftritte verfügen über mehrere Navigationspunkte oder -strukturen, welche ebenso von mehreren Auftritten wieder verwendet werden können. Diese Navigationsstrukturen setzen sich aus einer oder mehreren freigegebenen Navigationsstruktur(en) zusammen. Die freigegebenen Strukturen können mehrfach verwendet werden.

Die Navigationsstruktur schließt keine oder eine personalisierbare Startseite ein. Diese Startseite gehört genau zu einem Navigationspunkt.

Im nachfolgenden ER-Diagramm sind diese Objekte (Entitäten) und ihre Beziehungen (Relationship) zueinander aufgeführt.

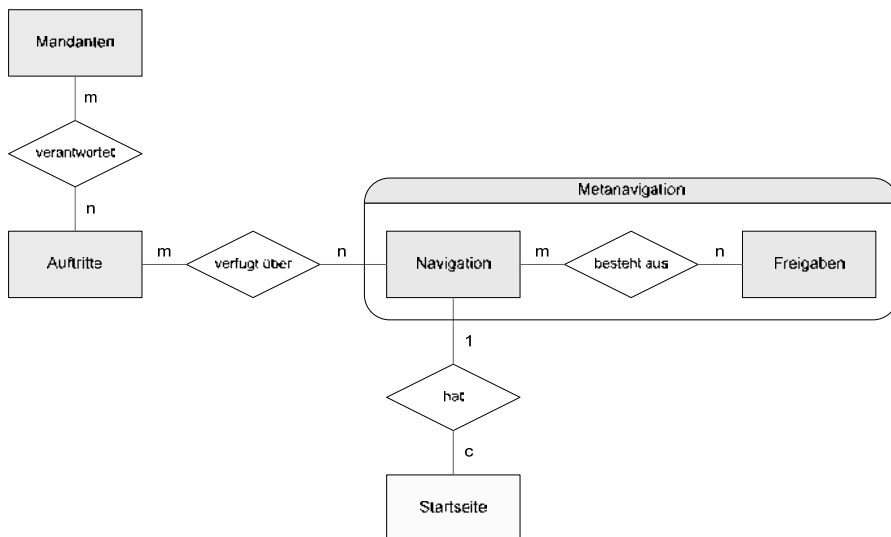


Abbildung 2: Entity-Relationship-Diagramm Multi-Site-Management

Das Multi-Site-Management erfordert die Anpassung von bestehenden und die Entwicklung von zwei neuen Software-Komponenten. Der *MSM Designer* dient der Verwaltung der Mandanten, der Auftritte und der Metanavigation und stellt eine entsprechende GUI zur Verfügung.

Konzeption, Design und Implementierung eines exemplarischen Prototyps sind in der Diplomarbeit ausführlich beschrieben.

Literaturverzeichnis

- [SAP04] SAP: SAP® Enterprise Portal für die öffentliche Verwaltung. SAP for Public Sector 2004. In: <http://www.sap.com/germany/media/50061926.pdf>, zugegriffen am 28.04.2008.
- [GRO05] Großmann, Martina; Koschek, Holger: Unternehmensportale - Grundlagen, Architekturen, Technologien. Springer Verlag 2005.

Workflow-Spezifikation für einen Knowledge Server - Konzeption und prototypische Umsetzung XML-basierter Server-Workflows

Stefan Finke
Universität Paderborn
teffi@upb.de

Abstract: In einem bestehenden Single Point of Information System der Fachgruppe DDI an der Universität Paderborn fungiert ein Knowledge Server als zentrale Instanz. Der Server dient als Schnittstelle zwischen den angebundenen, unterschiedlichen Systemen und ist für die Steuerung, Analyse, Datenhaltung und die Handhabung verantwortlich. Die Steuerung der angeschlossenen Systeme soll flexibel und anpassbar gestaltet werden, um die Bedienung für den Administrator einfacher zu gestalten. Das vorliegende Paper beschäftigt sich mit der Steuerung des Systems über so genannte Workflows und einer Workflow-Engine, die im Server realisiert ist und für die Ausführung der Workflows zuständig ist. Die Workflows werden mittels XML definiert und können in einem grafischen Werkzeug von einem Administrator erstellt und bearbeitet werden.

Art der Arbeit: Diplomarbeit
Betreuer: Wolfgang Reinhardt

1 Problembereich

Im Rahmen der an der Universität Paderborn stattfindenden Projektgruppenserie MoKEx [HHM⁺06] wurde das Single Point of Information-Konzept [HMR⁺07] entwickelt. Bei diesem Konzept werden unterschiedliche Datenhaltungssysteme zusammengefasst und der Zugriff über eine einzige Schnittstelle abgewickelt. Dieser Zugriff stellt sich für den Anwender transparent dar und erleichtert ihm die Handhabung von Informationen.

Das Prinzip des SPI wurde innerhalb der Projektgruppe in verschiedenen Servern umgesetzt. In dieser Arbeit wird die aktuellste Entwicklung betrachtet. Dabei handelt es sich um die Umsetzung eines Knowledge Servers (KNS), der entsprechend dem Konzept als Schnittstelle zwischen den unterschiedlichen angebundenen Systemen fungiert und für die Steuerung, Analyse, Datenhaltung und Handhabung der Informationen verantwortlich ist [Win08].

In der aktuellen Implementierung werden Arbeitsabläufe im System über fest im Java-Quellcode bereitgestellte Workflows realisiert. Diese Lösung ist sehr statisch und das Anbinden einer neuen Ressource im bestehenden Server-System würde eine Veränderung der Workflows und der Workflow-Engine im Server-Quellcode bedingen.

Um das Anbinden neuer Systeme zu erleichtern und die Steuerung der schon angeschlossenen Systeme flexibel und anpassbar zu gestalten, sollen zukünftig Workflows zum Einsatz kommen, die ohne Eingriff in den Serverkern erstellt werden können. Im Server selbst wird eine Open-Source Workflow-Engine eingesetzt, die für das Ausführen der Workflows zuständig ist.

2 Konzeption und Umsetzung

Zur Umsetzung des neuen KNS-Systems wurde ein Konzept entwickelt, das sich im Grunde in drei Aspekte unterteilen lässt: Die Anbindung neuer Adapter an den Server und die Konfiguration des Servers selbst, die Generierung und Bearbeitung von Workflows mittels eines Editors und die Beschreibung der Workflows.

Die Anbindung neuer Adapter an den Server, die die Dienste zur Verfügung stellen, sowie die Konfiguration des Servers selbst verlaufen generisch, wobei die erforderlichen Konfigurationen in XML-Dateien beschrieben und beim Start des Servers ausgewertet werden. Auch die Adapter verfügen über eine XML-Datei, die deren Funktionalität beinhaltet. Diese wird, ebenfalls beim Startvorgang des Servers, von diesem angefordert und ausgewertet, so dass der Server die entsprechenden Funktionen der Dienste im Editor anbieten kann. Dies führt zu dem zweiten Aspekt des Konzepts. Der Editor zur Generierung und Bearbeitung von Workflows stellt sich auf die aktuell im Server angebotenen Adapter ein und generiert generisch seine Oberfläche. Diese repräsentiert somit nur Service-Schnittstellen für einen Workflow an, wenn die entsprechende Funktion von einem Dienst, der im Server angemeldet ist, angeboten wird.

Der dritte Aspekt betrifft die Beschreibung der Workflows selbst, die mittels XML definiert werden. Dies gestaltet das Workflow-Management flexibler und löst die Bindung der Workflows an den Server-Quellcode. Somit können diese außerhalb des Servers generiert und bearbeitet werden und danach wieder in der Workflow-Engine im Server deployed werden. Die Kommunikation mit den Adapter und dem Editor läuft dabei über eine speziell dafür entwickelte Nachrichten-Bibliothek.

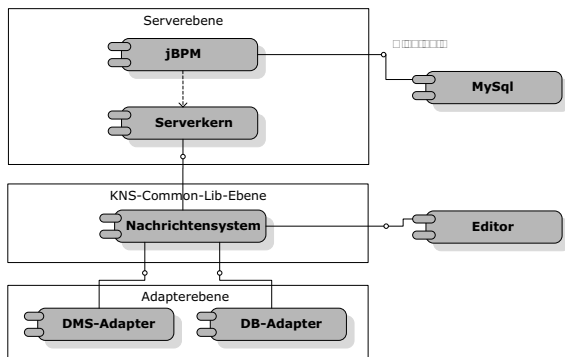


Abbildung 1: Architektur des entworfenen KNS-Systems.

Die Umsetzung dieses Konzepts ist in Abbildung 1 zu sehen. Auf die wichtigsten Elemente wird in den folgenden Abschnitten eingegangen.

2.1 Die Wahl der Workflow-Engine

Im Rahmen der Arbeit wurden mehrere Open Source Workflow-Engines und Workflow-Beschreibungssprachen evaluiert. Darunter fielen unter anderen YAWL [AH05], XPDL [Coa05] und jBPM [Rüc05].

Die Wahl fiel auf die jBPM-Engine mit der eigens dafür entwickelten Beschreibungssprache jPDL von JBoss (www.jboss.org). Die Gründe für diese Entscheidung werden in den folgenden beiden Auflistungen dargestellt. Die erste Liste führt dabei die Vorteile der Engine und die zweite Liste die Vorteile von jPDL auf. Vorteile der jBPM sind:

- Standard Java-Applikation, die als Stand-Alone-Lösung eingesetzt werden kann
- Die Bereitstellung als Java-Bibliothek bietet eine einfache Einbindung in ein bestehendes Java-System
- Einfach Konfiguration über XML-Dateien
- Anbindung einer Persistenzdatenbank über *Hibernate* möglich
- Versionierung von Workflows
- Speicherung von Variablen im Workflow-Kontext

Auswahlkriterien für jPDL waren:

- Zur Beschreibung und Ausführung von Workflows geeignet
- Abbildung der meisten Geschäftsprozesse möglich
- Test getriebene Entwicklung (ohne Persistenzdatenbank, etc.)
- Einfache Einbindung von Java-Quellcode über spezielle Elemente
- Open Source Editor steht als Eclipse-Plugin zur Verfügung

Wie in der zweiten Auflistung zu sehen ist, wird auch der Open Source jPDL-Editor als Auswahlkriterium aufgeführt. Dieser wurde im Rahmen dieser Arbeit um die für den KNS notwendigen Aspekte erweitert und modifiziert. Er wird im folgenden Abschnitt genauer erläutert.

2.2 Der Workflow-Editor

Als Grundlage des entwickelten Editors dient der jPDL-Editor von JBoss. Dieser steht als ein Open Source Eclipse-Plugin zur Verfügung, was eine Modifikation und Erweiterung gestattet.

Der Anwender muss sich beim Startvorgang des Editors zunächst beim Server anmelden.

Nach einem erfolgreichen Login fordert der Editor dann bei dem KNS die für die Generierung notwendigen Daten an. Dazu gehören unter anderen die im Server bereit gestellten Client-Requests und die zur Verfügung stehenden Service-Aufrufe. Aus diesen Daten werden in der Oberfläche entsprechende Elemente (Knoten) generiert, die dann zu einem Workflows zusammen gesetzt werden können. Zwischen diesen Knoten können Abhängigkeiten entstehen. So kann z.B. das Ergebnis eines Service-Aufrufs als Eingabe eines weiteren Dienstes dienen. Diese Abhängigkeiten und die für einen bestimmten Service-Aufruf anzugebenden Daten, für die zusätzliche Eingabeelemente im Editor bereit gestellt werden, verwaltet eine zentrale Instanz im Editor und werden wenn nötig in der XML-Beschreibung des Workflows gespeichert. Der aktuell bearbeitete Workflow kann dann, über eine weitere Ansicht im Editor, in die jBPM, mittels eines Nachrichtenaustauschs, deployed werden. Des Weiteren unterstützt der Editor das Laden schon im KNS vorhandener Workflows. Diese werden in der gleichen Ansicht angezeigt, in der auch das Hochladen von Workflows angeboten wird. Selektiert der Anwender einen Workflow, öffnet sich ein Pop-Up, in dem die grafische Darstellung des Workflows visualisiert wird, was die Auswahl des gesuchten Workflows erleichtert. Diese kann dann in dem Editor weiter bearbeitet werden.



Abbildung 2: Screenshot der Oberfläche des Editors am Beispiel eines einfachen Such-Workflows. Quelle: Eigene Darstellung

Abbildung 2 zeigt einen Screenshot der Oberfläche des erweiterten, modifizierten jPDL-Editors.

3 Zusammenfassung

In dieser Arbeit wurde das Workflow-Management in dem KNS-System der Projektgruppenserien MoKEx der Fachgruppe DDI an der Universität Paderborn, das auf dem Prinzip des SPI beruht, ersetzt. Das Workflow-Management des KNS lag als reine Java-Implementation im Server vor, was es sehr statisch gestaltete und bei jeder Änderung eines Workflows,

einen Eingriff in den Quellcode des Servers erforderte.

Zur Lösung dieses Problems wurde in der vorliegenden Arbeit ein Workflow-Management auf der Basis von XML-Workflow-Beschreibungen entwickelt. Dazu wurden bestehende Workflow-Engines und -Beschreibungssprachen evaluiert. Als Ergebnis dieser Untersuchung wurde die jBPM-Engine und die dazugehörige Sprache jPDL von JBoss als Grundlage für das zu entwickelnde System gewählt. Im Verlauf der Arbeit wurde die XML-Workflow-Beschreibung von jPDL um weitere Elemente ergänzt. Die jBPM-Engine wurde in den Server integriert und hat das bestehende Management-System ersetzt. Zur Verwaltung der Workflow-Definition und der Workflow-Instanzen wurde eine SQL-Datenbank verwendet, die die Objekte persistent speichert. Sie wird über Hibernate von der jBPM angesprochen.

Des Weiteren wurde die Anbindung der Adapter an den Server flexibler gestaltet. Die Service-Schnittstellen des Systems werden über XML-Beschreibungen definiert. Diese werden im Server, nach der Anmeldung der Adapter, vom Server angefordert, der dann die darin definierten Dienste zur Verfügung stellt.

Des Weiteren wurde zur Erstellung der Workflows ein Editor entwickelt. Dieser baut auf dem jPDL-Editor von JBoss auf. Der Editor wurde so erweitert, dass er sich beim Start an dem KNS anmeldet und die KNS-spezifischen Daten abfragt. Aus diesen Informationen generiert der Editor seine Oberfläche.

Der Editor bietet neben der Erstellung noch weitere Funktionen an. Zum Einen können Workflows in die jBPM im Server hochgeladen und zum Anderen aus der jBPM in den Editor geladen werden.

Literatur

- [Coa05] Workflow Management Coalition. Process Definition Interface – XML Process Definition Language. Bericht, Workflow Management Coalition, 2005.
- [HHM⁺06] K. Hinkelmann, K. Holzweißig, J. Magenheim, F. Probst und W. Reinhardt. Linking Communities of Practice with Learning Communities in Computer Science Education In IFIP International Federation for Information Processing. 2006.
- [HMR⁺07] Knut Hinkelmann, Johannes Magenheim, Wolfgang Reinhardt, Tobias Nelkner, Kai Holzweißig und Michael Mlynarski. KnowledgeBus - An Architecture To Support Intelligent And Flexible Knowledge Management. 2007.
- [Rüc05] Bernd Rücker. Vorstellung der Open Source BPM-Engine JBoss jBPM. Bericht, Camunda, 2005.
- [AH05] W. M. P. van der Aalst und A. H. M. ter Hofstede. YAWL: yet another workflow language. *Inf. Syst.*, 30(4):245–275, 2005.
- [Win08] Jens Winter. Konzeption und prototypische Implementierung eines Knowledge-Servers mit Adaptern zur Integration von E-Learning- und Knowledge-Management-Applikationen. Diplomarbeit, Universität Paderborn, 2008.

Verifikation von Datenabhängigkeiten in Ableitungscode

GI Fachbereich Softwaretechnik
Michael Förster, LuFG Informatik 12, RWTH Aachen
Michael.Foerster@rwth-aachen.de

Abstract: Mit Automatischen Differenzierung wird durch einen Compiler ein gegebener Eingabecode in Ableitungscode transformiert. Um einen effizienten Code zu erhalten wird eine Datenabhängigkeitsanalyse des Eingabecodes durchgeführt. Das Ergebnis der Analyse wird dann von einem weiteren Compiler verifiziert.

1 Einleitung

Die Datenabhängigkeiten in einem Programmcode sind von essenzieller Bedeutung für die Bildung des Ableitungscode. Daher ist eine automatische Verifikation von Datenabhängigkeiten für einen gegebenen Programmcode sinnvoll. Die Automatische Differenzierung (AD) [1] von C-Quellcode wird durch den Derivative Code Compiler (DCC) durchgeführt. Der DCC wird am LuFG Informatik 12: Software und Werkzeuge für Computational Engineering [3] der RWTH Aachen entwickelt. Die Verifikation soll der Verified C Compiler (VCC) leisten, welcher uns freundlicherweise vom “European Microsoft Innovation Center” [4] zur Verfügung gestellt wurde.

Es sei ein C-Quellcode mit einer Funktion gegeben, die n Eingabewerte erhält und daraus m Ausgabewerte berechnet. Wir betrachten also ein Programm P als Funktion $y = f(x)$, mit $f : \mathbb{R}^n \mapsto \mathbb{R}^m$. Der DCC transformiert den Code der Funktion f in den Ableitungscode $y' = f'(x) * \dot{x}$, wobei $f'(x)$ die Jacobimatrix und \dot{x} eine Richtung im Ursprungsraum von f ist.

In typischen Anwendungen ist man aber nicht an allen Ableitungen von den Ausgabewerten in Bezug auf alle Eingabewerte interessiert, sondern nur an ganz bestimmten. Die gewünschten Ableitungen geben wir dem DCC vor und dieser führt dann eine Aktivitätsanalyse durch. Diese Analyse ergibt dann für jede Zuweisung im Programm den Status aktiv oder passiv. Nur eine aktive Zuweisung wird vom DCC transformiert, so dass deren Ableitung berechnet wird. Sollte die Analyse aber für eine Zuweisung als Ergebnis passiv erhalten, obwohl diese aktiv ist, würde deren Ableitung nicht berechnet. Daher wäre die Berechnung der Ableitung ab dieser Stelle falsch. Aus diesem Grund soll der VCC uns die Passivität der Zuweisungen im Originalcode P verifizieren.

2 Aktivitätsanalyse

Bei einer Zuweisung $y=e$ ist y eine Programmvariable und e ein Ausdruck, der aus einer Teilmenge von Programmvariablen berechnet wird. Ist x in dieser Teilmenge enthalten schreiben wir dafür $x \in VAR(e)$. Durch die Zuweisung $y=e$ beeinflusst die Variable x die Variable y der linken Seite. Das bezeichnen wir mit $x \prec y$. Beeinflusst eine Variable x die Variable y über eine Kette von Zuweisungen bezeichnen wir das mit $x \prec^+ y$. Eine beliebige Zwischenvariable v heißt *varied*, wenn sie von einer unabhängigen Variable beeinflusst wird, das heißt es gilt:

$$v \text{ varied} : \iff \exists x \in x_I : x \prec^+ v$$

Die Menge x_I ist die Teilmenge der Eingabevariablen, an denen wir interessiert sind. Die Menge der für uns interessanten Ausgabevariablen heißt y_D . Als *useful* bezeichnen wir v , wenn sie mindestens eine der abhängigen Ausgabevariablen aus y_D beeinflusst, d.h. es gilt:

$$v \text{ useful} : \iff \exists y \in y_D : v \prec^+ y$$

Ist eine Programmvariable v sowohl *varied* als auch *useful*, so nennen wir sie *aktiv*. Andernfalls ist sie *passiv* und deren Ableitung sollte weder berechnet noch benutzt werden. Ist v nicht *varied*, dann ist deren Ableitung Null. Ist sie nicht *useful* wird die Ableitung gar nicht benötigt.

Betrachten wir nun eine Originalanweisung $I : a = \phi(B)$, wobei wir mit $B = \{b_1, \dots, b_k\}$ die skalaren Argumente von ϕ bezeichnen. Diese Anweisung generiert eine Zuweisung:

$$\dot{a} = \sum_{i=1}^k \frac{\partial \phi}{\partial b_i} \cdot \dot{b}_i$$

Ist a nicht aktiv, dann ist auch keines der b_i aktiv und die Zuweisung kann wegfallen. Ist irgendein b_i nicht aktiv, dann ist der Summand $\frac{\partial \phi}{\partial b_i} \cdot \dot{b}_i$ gleich Null. Daher können unnötige Operationen eingeschränkt werden und der Ableitungscode wird einfacher und effizienter.

2.1 Beispiel

Listing 1: Originalcode

```

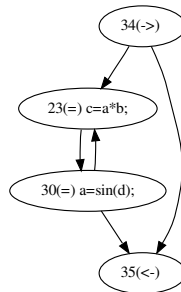
void f( double & a, double & b, double & c, double & d)
//Sad indep a b
//Sad dep c
{
    while(a<b) {
        c=a*b;
        a=sin(d);
    }
}

```

In dem Beispielcode aus Listing 1 sind wir an der Ausgabevariablen c und an den Eingabevariablen a und b interessiert. Das teilen wir dem Compiler durch die Zeilen 2 und 3 mit.

Den Kontrollflussgraph sehen wir in Abbildung 1. Bei diesem überschaubaren Beispiel kann man nachvollziehen wie das Ergebnis der Aktivitätsanalyse aussieht. Die Varied-Analyse läuft vorwärts durch das Programm, d.h. es werden alle möglichen Pfade von Knoten 34 zu Knoten 35 durchlaufen. Da a und b varied sind wird die Zuweisung in Knoten 23 varied. Die Useful-Analyse durchläuft rückwärts alle möglichen Pfade von Knoten 35 nach 34. Zuerst ist nur c useful und nach dem ersten Besuch von 23 dann a und b. Danach wird in Knoten 30 auch d useful. Danach ändert sich bis Knoten 34 nichts mehr. Als Ergebnis erhalten wir also, dass 23 aktiv ist und Knoten 30 passiv, da dieser zwar useful aber nicht varied ist. Somit müssen wir die nur die Zuweisung aus 23 ableiten. Die Ableitung wird in Listing 2 in den Zeilen 13 bis 19 berechnet. Die Zuweisung aus Knoten 30 wird in Zeile 22 einfach übernommen.

Abbildung 1: Kontrollflussgraph von Listing 1



Listing 2: Ableitungscode

```

5  void d0_f(double& a, double& d0_a, double& b, double& d0_b,
    double& c, double& d0_c, double& d, double& d0_d)
    //Sad indep a b
    //Sad dep d0_c
10 {
    double v0_0=0;
    double v0_1=0;
    double v0_2=0;
    double d0_v0_0=0;
    double d0_v0_1=0;
    double d0_v0_2=0;
    while (a<b) {
15         d0_v0_0=d0_a;
            v0_0=a;
            d0_v0_1=d0_b;
            v0_1=b;
            d0_v0_2=v0_1*d0_v0_0+v0_0*d0_v0_1;
            v0_2=v0_0*v0_1;
            d0_c=d0_v0_2;
20         c=v0_2;
            d0_a=0;
            a=sin(d);
    }
}

```

3 Verifikation der Aktivitätsanalyse

Der DCC generiert nicht nur den Ableitungscode sondern auch den Code, den der VCC anschließend verifizieren soll. Für das Beispiel 1 sehen wir diesen in Listing 3. Die Varied-Analyse findet in Zeile 17 bis 32 statt und die Useful-Analyse in Zeile 34 bis 58. Anschließend überprüft der VCC ab Zeile 59 die Behauptungen über die Zuweisungen. Dort wird behauptet, dass die erste Zuweisung aktiv und die zweite passiv ist.

Listing 3: Verifikationscode

```
/* Source code for verification of file d0.test.c */
#include <gcc2.h>
int sin(int);
5 void d0.f(int* a, int* d0.a, int* b, int* d0.b, int* c, int* d0.c, int* d, int* d0.d)
maintains( mutable(a) ) maintains( mutable(d0.a) ) maintains( mutable(b) )
maintains( mutable(d0.b) ) maintains( mutable(c) ) maintains( mutable(d0.c) )
maintains( mutable(d) ) maintains( mutable(d0.d) )
requires( a!=d0.a && a!=b && a!=d0.b && a!=c && a!=d0.c && a!=d && a!=d0.d )
requires( b!=d0.b && b!=c && b!=d0.c && b!=d && b!=d0.d )
10 requires( c!=d0.c && c!=d && c!=d0.d ) requires( d!=d0.d )
writes( a, d0.a, b, d0.b, c, d0.c, d, d0.d )
{
    bool assignment_23_varied = 0; bool assignment_30_varied = 0;
    bool assignment_23_useful = 0; bool assignment_30_useful = 0;
15    bool a_varied = 1; bool b_varied = 1; bool c_varied = 0; bool d_varied = 0;
    bool a_useful = 0; bool b_useful = 0; bool c_useful = 1; bool d_useful = 0;
    // vary analysis
    { // while statement
        // for all x elem LHS(B): bool old_x_varied = x_varied , result_x_varied;
20        bool old_a_varied = a_varied , result_a_varied;
        bool old_c_varied = c_varied , result_c_varied;
        // 1. iteration
        c_varied = a_varied || b_varied; // c=a*b;
        assignment_23_varied = assignment_23_varied || c_varied;
25        a_varied = d_varied; // a=sin(d);
        assignment_30_varied = assignment_30_varied || a_varied;
        // for all x elem LHS(B): result_x_varied = result_x_varied || x_varied;
        result_a_varied = result_a_varied || a_varied;
        result_c_varied = result_c_varied || c_varied;
30        a_varied = old_a_varied || result_a_varied;
        c_varied = old_c_varied || result_c_varied;
    } // end of while statement

    // useful analysis
35    { // while b do B1
        // for all x elem Var(B1): bool old_x_useful = x_useful
        bool old_a_useful = a_useful;
        bool old_b_useful = b_useful;
        bool old_c_useful = c_useful;
40        bool old_d_useful = d_useful;
        // useful analysis for B1
        // 1. iteration
        assignment_30_useful = assignment_30_useful || a_useful;
        if( a_useful ) { a_useful = 0; d_useful = 1; } // a=sin(d);
45        assignment_23_useful = assignment_23_useful || c_useful;
        if( c_useful ) { c_useful = 0; a_useful = 1; b_useful = 1; } // c=a*b;
        // 2. iteration
        assignment_30_useful = assignment_30_useful || a_useful;
        if( a_useful ) { a_useful = 0; d_useful = 1; } // a=sin(d);
50        assignment_23_useful = assignment_23_useful || c_useful;
        if( c_useful ) { c_useful = 0; a_useful = 1; b_useful = 1; } // c=a*b;
        // end of iterations
        // for all x elem Var(B1): x_useful = old_x_useful || x_useful;
55        a_useful = old_a_useful || a_useful;
        b_useful = old_b_useful || b_useful;
        c_useful = old_c_useful || c_useful;
        d_useful = old_d_useful || d_useful;
    } // end of while stmt
60    assert( ( assignment_23_varied && assignment_23_useful ) );
    assert( !( assignment_30_varied && assignment_30_useful ) );
}
```

Literatur

- [1] A. Griewank: Evaluating Derivatives, SIAM 2000.
- [2] Hascoët, Naumann, Pascual: 'To be recorded' analysis in reverse-mode automatic differentiation, Future Generation Computer Systems 21, 1401-1417, Elsevier 2005.
- [3] <http://www.stce.rwth-aachen.de>
Webpage LuFG Informatik 12: Software und Werkzeuge für Computational Engineering
- [4] <http://www.microsoft.com/emic>
European Microsoft Innovation Center

Suche nach prozessadäquaten Webservices mit Hilfe eines BPEL Repositories

Andreas Harder

FH Aachen
Elektrotechnik und Informationstechnik
andreas.harder@gmail.com

Art der Arbeit: Diplomarbeit

Betreuer/in der Arbeit: Prof. Dr.-Ing. Thomas Ritz

Abstract: Das Auffinden geeigneter Dienste ist eine der zentralen Aufgaben in einer SOA. Dieser Artikel zeigt, wie durch Informationen, die aus BPEL-Prozessen eines Unternehmens gewonnen werden, die Ergebnisse einer konventionellen Webservice Suchanfrage verbessert werden können.

1 Einleitung

Die Verwendung einer Service Oriented Architecture (SOA) in Unternehmen erfreut sich wachsender Beliebtheit [Dos05]. Als Implementierung wird dabei meist auf Webservices gesetzt. Dieser Artikel beschäftigt sich mit einer der zentralen Aufgaben in einer auf Diensten basierender Architektur - das Auffinden eines Dienstes zur Erfüllung einer geforderten Funktion. Als Hilfsmittel wird derzeit meist das Universal Description, Discovery and Integration (UDDI)-Verzeichnis eingesetzt. Für einige Zwecke ist das UDDI jedoch ungeeignet. Dazu gehört vor allem das Suchen eines Dienstes von dem lediglich die zu erfüllende Funktionalität bekannt ist. Ursachen dafür finden sich im Aufbau der UDDI Suchfunktionen. Diese sind für Benutzer ausgelegt, die bereits wissen welchen Dienst sie verwenden möchten und nach Details, wie beispielsweise den Endpunkt oder die WSDL-Beschreibung, suchen.

Hier setzt das erstellte Konzept an. Die Verwendung von Informationen, die aus BPEL-Dokumenten gewonnen werden, eröffnet neue Möglichkeiten Suchergebnisse zu verbessern oder neue Suchverfahren einzuführen. Dazu ist meist kein großer zusätzlicher Aufwand notwendig, da viele Unternehmen bereits BPEL zur Komposition von Diensten in ihrer SOA einsetzen.

2 Konzeption einer Suchmaschine für Webservices

2.1 Erstellen einer Datenbank

Als Basis für die Suchmaschine muss zunächst ein Index mit allen relevanten Informationen angelegt werden. Dieser Index umfasst zwei Teile. Ein BPEL-Repository und ein vereinfachtes Abbild des UDDI-Verzeichnisses.

Das UDDI-Verzeichnis enthält umfangreiche Informationen zu den in Unternehmen verwendeten Diensten. Bei der Suche nach diesen Diensten wird jedoch beim UDDI lediglich eine Suche nach den Namen angeboten. Eine Durchsuchung der ebenfalls enthaltenen Beschreibungstexte ist nicht möglich. Um diese Informationen nutzbar zu machen, wird ein vereinfachtes Abbild des UDDI-Verzeichnisses in den Datenbestand übernommen.

Den zweiten Teil des Indexes bildet das BPEL-Repository. Zum einen werden die Dokumentationstexte der BPEL-Prozesse zusätzlich zu den UDDI-Beschreibungstexten als Basis für Suchanfragen genutzt. Zum anderen werden Strukturinformationen zum Implementieren von Rankingverfahren verwendet.

2.2 Einführen eines Rankings

Eine Ergebnismenge von hunderten oder gar tausenden Diensten ist zum Prüfen zu umfangreich und damit sinnlos. Die Einführung eines Rankings bietet hier Abhilfe. Dieses erlaubt das Sortieren der Ergebnisse nach Relevanz, was zu einer schnelleren Auffindung des gesuchten Dienstes führt.

Der Rang oder auch die Relevanz eines Suchergebnisses wird aus mehreren Faktoren, die sich in zwei Kategorien einteilen lassen, errechnet [Lew05].

Anfrageabhängige Faktoren Diese Faktoren ermitteln für die Ergebnisse jeder Suchanfrage zugehörige Relevanzen. Beispielsweise erhält ein Suchergebnis einen höheren Rang, falls es ein häufigeres Auftreten des Suchbegriffes aufweist.

Anfrageunabhängige Faktoren Diese Faktoren bestimmen für jeden Dienst eine generelle Relevanz, die in bestimmten Abständen errechnet und mit in den Datenbestand aufgenommen wird. Beispielsweise kann davon ausgegangen werden, dass ein Dienst der häufig von BPEL-Prozessen aufgerufen wird eine hohe generelle Relevanz besitzt.

Der Gesamtrang eines Suchergebnisses wird durch die Kombination des anfrageabhängigen und anfrageunabhängigen Ranges errechnet. Mit welcher Gewichtung die jeweiligen Faktoren in die Rechnung eingehen kann variiert werden. Möglich wäre beispielsweise die Multiplikation beider Ränge.

2.3 Erstellen einer Such-API

Zuletzt müssen dem Benutzer Funktionen zur Verfügung gestellt werden, um den aufgebauten Datenbestand zu durchsuchen. Da es sich bei der Umgebung um eine SOA handelt, ist es sinnvoll diese Funktionen, ähnlich wie beim UDDI, als Webservices zu implementieren.

2.4 Komponenten

Das bisher geschilderte Konzept beschreibt indirekt fünf Komponenten, die im Zusammenspiel in Abbildung 1 dargestellt werden.

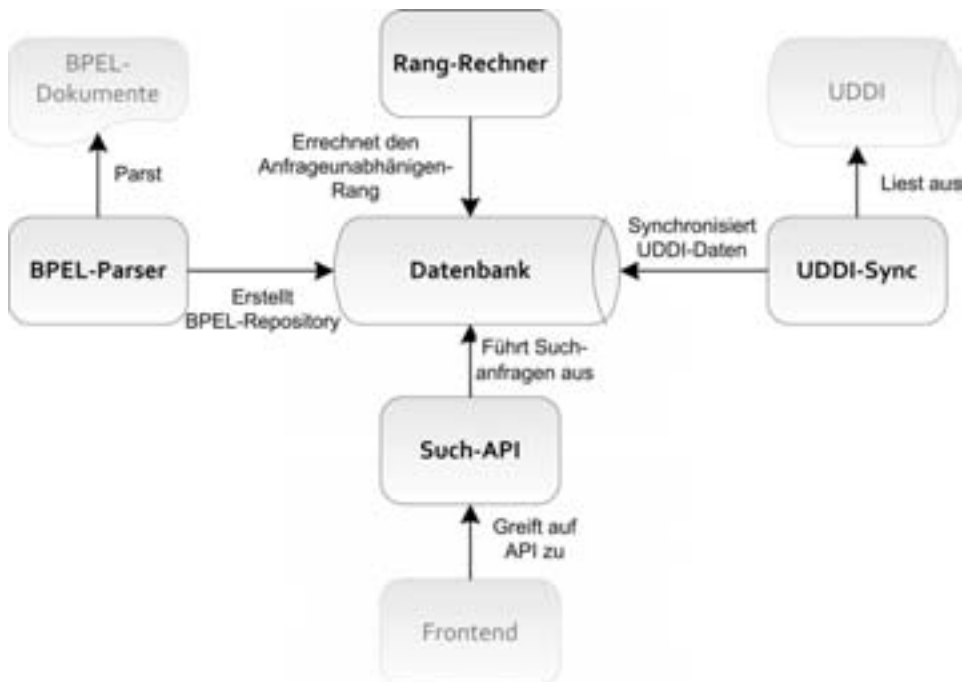


Abbildung 1: Komponenten

3 Fazit / Ausblick

Die konzipierte Suchmaschine kann in vielen Umgebungen ohne großen zusätzlichen Aufwand integriert werden. Einmal eingerichtet bietet sie in vielen Bereichen deutliche Vorteile.

Die Einbeziehung von Beschreibungen und Dokumentationen sowie die Verwendung einer modernen Volltextsuche führt zu deutlich mehr Ergebnissen auf eine Suchanfrage. Die Sortierung der Ergebnisse nach Relevanz bewirkt, dass selbst bei großen Ergebnismengen eine hohe Wahrscheinlichkeit besteht den richtigen Dienst schnell zu finden.

Strukturinformationen, die sich im BPEL-Repository befinden, ermöglichen eine ganze Reihe weiterer, nützlicher Suchfunktionen. Beispielsweise könnte man dem Benutzer, der einen gesuchten Dienst gefunden hat, einen Hinweis auf weitere Dienste geben, die potentiell relevant sein könnten. Das oftmalige Auftreten zweier Dienste im selben BPEL-Dokument könnte auf einen Zusammenhang und damit auf eine solche Relevanz für den Suchenden hinweisen.

Literatur

- [Dos05] Wolfgang Dostal. *Service-orientierte Architekturen mit Web Services: Konzepte - Standards - Praxis*. Elsevier Spektrum Akad. Verl., München, 1. Aufl., Auflage, 2005.
- [Lew05] Dirk Lewandowski. *Web Information Retrieval: Technologien zur Informationssuche im Internet*, Jgg. 7 of *Reihe Informationswissenschaft der DGI*. Dt. Ges. für Informationswiss. und Informationspraxis, Frankfurt am Main, 2005.

Rettungsassistenzsystem

Mirco Schindler, Dirk Herrling

Technische Universität Clausthal
Institut für Informatik – Lehrstuhl für Software Systems Engineering
mirco.schindler@tu-clausthal.de
dirk.herrling@tu-clausthal.de

Art der Arbeit: Studienprojekt
Betreuer/in der Arbeit: Prof. Dr. Andreas Rausch (andreas.rausch@tu-clausthal.de),
Dipl.- Inf. Dirk Niebuhr (dirk.niebuhr@tu-clausthal.de)

Abstract: Ziel dieses Projektes war es, einen Prototyp zu entwickeln, welcher Rettungskräfte in einem Massenanfall von Verletzten unterstützt und die Fähigkeiten der an der TU-Clausthal entwickelten DAiSI-Middleware veranschaulicht. Das Rettungsassistenzsystem wird auch auf der CeBIT 2009 vorgestellt.

1 Einleitung

Ein Massenanfall von Verletzten (MANV) wird als *„Notfall mit einer größeren Anzahl von Verletzten oder Erkrankten sowie anderen Geschädigten oder Betroffenen“*, definiert *„der mit der vorhandenen und einsetzbaren Vorhaltung des Rettungsdienstes aus dem Rettungsdienstbereich versorgt werden kann.“* (Deutsches Institut für Normung e.V., 2002).

Nach einer Alarmierung sind am Einsatzort eine örtliche Einsatzleitung (ÖEL) sowie verschiedenste Rettungseinheiten (Notärzte, Sanitäter, Feuerwehrleute, Soldaten) anwesend. Die Notärzte kategorisieren in mehreren Sichtungsdurchgängen die Verunfallten in Triage-Klassen¹ und werden von der Einsatzleitung koordiniert.

Bei einem MANV ist das Hauptproblem der Einsatzleitung ein Mangel an Informationen über Aufenthaltsort und Tätigkeit der Rettungskräfte und Zustand verschiedener Verletzter. Die Einsatzleitung wird zurzeit ausschließlich per Sprechfunk von den Rettungskräften informiert. Im Folgenden werden wir ein System vorstellen, das dieses Problem löst. Andere Projekte wie zum Beispiel das SoKNOS-Projekt² oder Med-on@ix³, beschäftigen sich mit anderen oder ähnlichen Problemstellungen.

¹ Klassifizierung von Verletzten oder Erkrankten nach Schwere der Verletzung/Erkrankung

² <http://www.soknos.de/>

³ <http://www.medonaix.de>

2 Lösungsansatz

Viele von der Einsatzleitung benötigten Informationen können von verschiedenen Sensoren leichter aufgenommen werden als von Menschen. Darüber hinaus können diese Informationen über digitalen Datenfunk deutlich effektiver übertragen werden als über herkömmlichen Sprechfunk. Deshalb wollen wir Sensoren zur Vitaldatenerfassung mit digitaler Datenübertragung im Rettungsassistenzsystem verwenden.

Das Rettungsassistenzsystem umfasst tragbare Interaktionsgeräte für Notärzte, ein stationäres für die Einsatzleitung, Verletzteneinheiten und Sensoreinheiten. Jede Einheit kann in einer Softwarekomponente abgebildet werden. Es lag nahe ein System zu entwickeln, das Komponenten zur Laufzeit integrieren und entfernen kann. Diese Eigenschaften sind erforderlich, da nach einer Alarmierung die Rettungskräfte zu unterschiedlichen Zeitpunkten eintreffen und Verletzte auch während des Einsatzes abtransportiert werden.

Die Verletzteneinheiten werden einem Verletzten von Notärzten beim Auffinden statt der üblichen Patientenanhängerkarte umgelegt. Diese speichert persönliche Daten, Position und die aktuelle Triage Klasse. In Kombination mit der Sensoreinheit werden Vitaldaten wie Puls, Sauerstoffsättigung und Blutdruck aufgenommen. Sie werden in Echtzeit an die Notärzte und die Einsatzleitung übertragen.

Position und aktuelle Aktivität der Notärzte sind der Einsatzleitung durch die Notarzteinheit bekannt. So kann sie schneller bessere Entscheidungen treffen, zum Beispiel welcher Notarzt einen bestimmten Patienten aufsucht oder welche Verletzten von welcher Sammelstelle in welcher Reihenfolge abtransportiert werden.

3 Umsetzung

Um die Vitaldaten von Verletzten aufnehmen zu können wurde eine Sensoreinheit entwickelt die über Puls-, Sauerstoffsättigungs- und Blutdrucksensoren verfügt. Die Messwerte werden über eine ZigBee⁴ ähnliche Schnittstelle von einem SunSPOT⁵ an die Einsatzleitung übertragen. Auch die Verletzteneinheit ist durch einen SunSPOT realisiert welche ebenfalls mit der Einsatzleitung kommuniziert.

Die Notarzteinheiten kommunizieren über WLAN mit der Einsatzleitung. Über diese Schnittstelle kann die Einsatzleitung die aktuellen Aktivitäten einsehen oder Anweisungen erteilen.

⁴ <http://standards.ieee.org/getieee802/802.15.html>

⁵ <http://www.sunspotworld.com/>

Befindet sich ein Notarzt in der Nähe eines Verletzten werden die Daten, welche auf der Verletzteneinheit gespeichert sind, auf dem Gerät des Notarztes angezeigt. Der Notarzt hat so die Möglichkeit mit der notwendigen Behandlung zu beginnen, ohne vorher Vitaldaten manuell messen zu müssen. Zusätzlich kann er die persönlichen Daten des Verletzten wie Name, Alter und Geschlecht erfassen. Sie werden dann wiederum auf der Verletzteneinheit gespeichert und automatisch an die Einsatzleitung übertragen.

Um die dynamische Integration der verschiedenen Komponenten zu ermöglichen wurde auf die DAiSI-Middleware⁶ (D. Niebuhr, 2007), die an der TU-Clausthal entwickelt wurde, zurückgegriffen. Laufzeittests zur Komponentenintegration erhöhen die Zuverlässigkeit des Gesamtsystems.

4 Fazit

Unser System erhöht den Informationsfluss während des Einsatzes und dient als Ergänzung des Sprechfunks. Der Einsatzleiter hat nun präzisere Informationen über Anzahl, Lage und Triage-Klasse aller bisher gefundenen Verletzten, so dass die beschränkten Ressourcen zielgerichteter eingesetzt werden können. Durch die Aufbereitung der gesammelten und gemessenen Daten steht auch eine genauere Planungsgrundlage für alle logistischen Probleme zur Verfügung.

Ein weiterer Vorteil ist die permanente Überwachung der Vitaldaten. Ohne Rettungssassistentensystem wird eine Änderung des Gesundheitszustandes eines Verletzten nur erkannt, wenn der Verletzte gerade von einem Notarzt behandelt wird. Die Praxis zeigt jedoch, dass dies in einem Katastrophenfall nicht für alle Verletzten gewährleistet werden kann. Das System kann solche Änderungen jedoch automatisch erkennen und so kann schnellstmöglich darauf reagiert werden.

Das Rettungssassistentensystem kann auch bei Rettungsübungen und Schulungen von Rettungskräften eingesetzt werden. Hierfür können vielfältige Kombinationen aus realen und simulierten Hard- und Softwarekomponenten verwendet werden. So können Zustandsänderungen bei Unfalldarstellern erstmals realistisch während einer Übung simuliert werden. Außerdem können die gesammelten Daten für die Übungsnachbereitung genutzt werden.

Das Projekt wurde in Zusammenarbeit mit Siemens, SUN Microsystems und Sinosys umgesetzt. Fachliches Know-how wurde von den Mitarbeitern der Rettungsschule Niedersachsen beigesteuert. Es wird erstmals auf der CeBIT 2009 vorgestellt und soll im Anschluss bei Übungen in der Praxis getestet werden, um Informationen über die Anforderungen und die Handhabung eines solchen Systems zu sammeln.

Weitere Informationen sind unter folgender URL zu erhalten:

www2.in.tu-clausthal.de/~Rettungssassistentensystem

⁶ <http://ftp.tu-clausthal.de/ftp/institute/informatik/sse/web/Publikationen/Dokumente/2007-SSE-NKA+07-DAiSI%20-%20Dynamic%20Adaptive%20System%20Infrastructure.pdf>

Literaturverzeichnis

Deutsches Institut für Normung e.V. (2002). DIN 13050. Beuth-Verlag.

D. Niebuhr, H. Klus, M. Anastasopoulos, J. Koch, O. Weiß, A. Rausch : DAiSI – Dynamic Adaptive System Infrastructure, IESE-Report No. 051.07/E, 2007

A Controlled Experiment for Measuring the Impact of Aspect-Oriented Programming on Software Development Time

Sebastian Kleinschmager

Universität Duisburg-Essen
Institut für Informatik und Wirtschaftsinformatik
sebastian.kleinschmager@stud.uni-due.de

Art der Arbeit: Bachelorarbeit

Betreuer/in der Arbeit: Dr. Stefan Hanenberg

Abstract: Dieses Dokument fasst Thematik, Aufbau und Ergebnisse der genannten Bachelorarbeit kurz und bündig zusammen.

1 Einleitung und Motivation

Die Aspektorientierung ist eine relativ neue Programmier Technik, die die Kapselung bestimmter Querschnittsfunktionen in Programmen zum Ziel hat, welche mit der reinen Objektorientierung nicht oder nur sehr schwierig modular zu erfassen sind. Sie geht auf ein Paper von Kiczales et. al zurück [KLMM97], hat sich seitdem aber weiterentwickelt.¹ Obwohl der Ansatz und die bisherigen Implementierungen der Technik sehr vielversprechend sind, gibt es bisher nur wenige empirische Studien zu dem Thema.² Dieses Dokument beschreibt zusammenfassend den Aufbau und das Ergebnis eines kontrollierten Experiments, welches im Rahmen der Bachelorarbeit durchgeführt wurde. Das Experiment wurde aufgesetzt, um die allgemeine Arbeitsgeschwindigkeit von Entwicklern bei der Nutzung des objektorientierten und des aspektorientierten Ansatzes für eine bestimmte Anzahl an Aufgaben vergleichbar zu machen. Dabei wurde der primäre Fokus darauf gelegt messbar zu machen, ob die Aspektorientierung einen positiven Einfluss auf die Entwicklungszeit im Vergleich zur Objektorientierung bei der gleichen Aufgabe hat und ab welchem Umfang sich ein Aspekt lohnt. Inhalt der Aufgaben war in allen Fällen eine mehr oder weniger große Anzahl an mit redundantem Code zu erweiternden Methoden. Die Fortschritte der Teilnehmer wurden über ein Software-Tool erfasst und in unterschiedlichen Formen ausgewertet.

¹ Eine sehr ausführliche Ausarbeitung zur Aspektorientierung findet sich in [Ha05]

² Viele der Studien dazu sind im Literaturverzeichnis aufgeführt: [BH07] [CC04] [CRGF08] [GBFD07] [LKC04] [LV99] [MKL97] [PP04] und [WBM01]

2 Experimentaufbau

Der Versuchsaufbau war folgender: Die Teilnehmer sollten nach Ausfüllen eines Fragebogens zu ihren Programmierkenntnissen an einem selbstgeschriebenen Programm insgesamt neun unterschiedlich großen Aufgaben durchführen, wobei jede Aufgabe verlangte eine gewisse Anzahl von Methoden mit zusätzlichem Code zu versehen. Diese Aufgaben sollten sowohl objektorientiert als auch aspektorientiert durchgeführt werden. Welcher Teilnehmer mit welcher Technik zuerst begann wurde per Zufall ausgewählt. Im Zuge einer anderen Arbeit wurde speziell für dieses Experiment ein Tool entwickelt [Jo08], welches dafür zuständig war, möglichst viele Interaktionen der Teilnehmer mit der Entwicklungsumgebung aufzunehmen und in eine Datenbank zu schreiben, um sie später in den unterschiedlichsten Formen auszuwerten.

Die genutzten Technologien waren Java, AspectJ³ (die Implementierung des Aspektorientierten Ansatzes für Java) und die Eclipse-Entwicklungsumgebung. Alle Teilnehmer erhielten zuerst eine Mini-Schulung von ca. 1 ½ Stunden in der Programmierung mit AspectJ und bekamen dann die Aufgabenbeschreibungen, die sie in der vorgegebenen Reihenfolge bearbeiten sollten. Die Teilnehmer konnten anhand von Testfällen für jede Aufgabe überprüfen, ob sie mit der Aufgabe fertig waren und durften erst dann zur nächsten Aufgabe übergeben. Die Testfälle waren noch in einer anderen Hinsicht ein wichtiger Aspekt des Experiments, da sie benutzt wurden um später während der Auswertung den Fortschritt der Teilnehmer für jede Aufgabe nachvollziehbar zu machen und sowohl eine Zeitreihe über den Fortschritt erstellen zu können als auch die generell benötigte Zeit bis zur Erfüllung messbar zu machen.⁴

3 Auswertung und Ergebnisse

Nachdem das Experiment vorbei war wurden die Daten in eine auswertbare Form gebracht, um sie statistisch auszuwerten, sie darzustellen und zu interpretieren. Dabei wurde der Fokus einerseits auf den Vergleich der gesamten Entwicklungszeit von Objektorientierung und Aspektorientierung gelegt, um zu schauen für welche Teilnehmer und bei welchen Aufgaben sich ein positiver Effekt durch die Nutzung von Aspektorientierung bemerkbar machte, und wie groß die Differenzen und Unterschiede im negativen Fall sein würden. Andererseits wurde in diesem Kontext versucht, eine Art Break-Even Analyse durchzuführen, indem per linearer Regression idealisierte Fortschrittskurven der objektorientierten Aufgabenbearbeitungen erstellt wurden und den Break-Even Punkt zu berechnen (siehe dazu Abb. 1). Dieser bestand in diesem Fall aus der Anzahl der zu bearbeitenden Methoden (in Abb. 1 als Anzahl der Erfolge, d.h. positive Testfälle gekennzeichnet), ab welcher der Zeitaufwand für Objektorientierung und Aspektorientierung ungefähr gleich waren. Ziel der Analyse in dieser Form, war es herauszufinden, ob es einen generellen Trend zu einem Break-Even Punkt für bestimmte Aufgabengrößen geben würde.

³ <http://www.eclipse.org/aspectj/>

⁴ Eine gute und übersichtliche Einführung in den Aufbau von Experimenten in der Softwaretechnik bietet [Pr01]

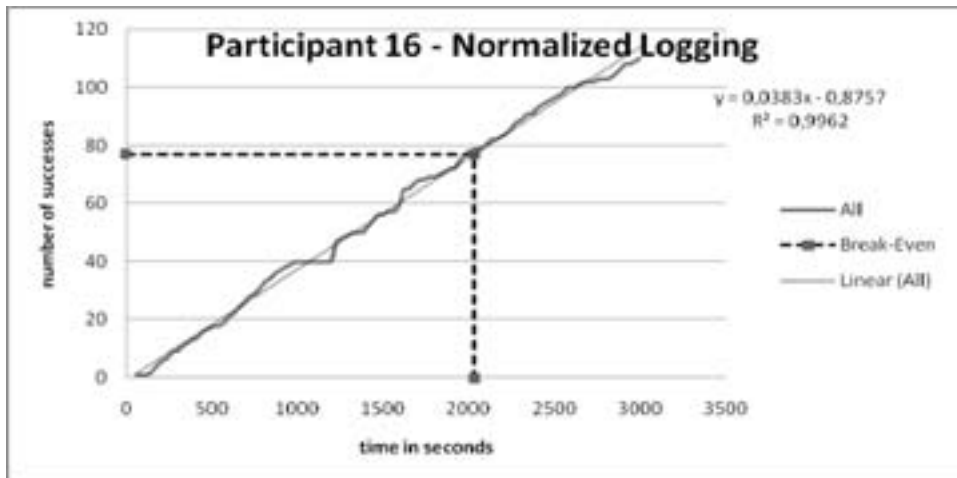


Abbildung 1: Beispielhafte Darstellung einer Break-Even Analyse

Die Daten für die Zeiten die die Teilnehmer für beide Programmieransätze benötigten wurden sowohl über alle Teilnehmer und Aufgaben verglichen und statistisch aufbereitet (beispielsweise durch deskriptive Funktionen wie arithmetisches Mittel und Median), als auch für jeden Teilnehmer einzeln betrachtet.

Für den Vergleich aller Teilnehmer wurden folgende statistische Verfahren angewandt: Kolmogorov-Smirnov- und Shapiro-Wilk-Test zur Prüfung der Hypothese auf Normalverteilung, welche abgelehnt werden musste. Statt eines T-Tests wurde ein Wilcoxon-Test angewandt, der zumindest eine signifikante Abweichung der objektorientierten von den aspektorientierten Ergebnissen nachweisen konnte.

Für die Auswertung der Ergebnisse für jeden einzelnen Teilnehmer wurden die Ergebnisse der aspektorientierten Lösungen und der objektorientierten Lösungen unter anderem in direkte Verhältnisse gesetzt, um sie besser interpretierbar zu machen.

Die Ergebnisse der Datenauswertung waren insgesamt leider nicht eindeutig genug, um eine allgemeine Aussage aus ihnen ableiten zu können. Allerdings konnte ein definitiver Trend festgestellt werden, dass die Aspektorientierung ab einer wachsenden Größe äquivalenten redundanten objektorientierten Codes auf jeden Fall die Entwicklungszeit verkürzen kann, was sich vor allem aus den Ergebnissen der größeren Aufgaben herauskristallisierte. Zusätzlich konnte anhand der Informationen aus den Fragebögen herausgefunden werden, dass Entwickler, die bereits viel mit Aspektorientierung gearbeitet hatten, eine weitaus größere Zeitersparnis einfuhren.

4 Konklusion

Auch wenn die Daten keine eindeutigen Schlussfolgerungen zulassen, war das Experiment ein Schritt in die richtige Richtung den Einfluss der Aspektorientierung auf die Entwicklungszeit und den gesamten Entwicklungsprozess besser zu verstehen. Gerade da es einen allgemeinen Mangel an empirischen Studien in der Softwaretechnik gibt [Ti97], ist es wichtig an einem Punkt anzusetzen um sukzessive das Verständnis einer Thematik sowohl inhaltlich zu verbessern als auch empirische Erkenntnisse dazu zu gewinnen. Dies wurde mit dieser Studie getan und bietet die Möglichkeit der Experimentreplikation, breiterer Auswertung der Daten und einer Weiterentwicklung zukünftiger Experiment zum Thema auf Basis der gesammelten Erfahrungen und Erkenntnisse.

Literaturverzeichnis

- [BH07] Bartsch, Marc and Harrison, Rachel: An exploratory study of the effect of aspect-oriented programming on maintainability. 2007.
- [CC04] Colye, Adrian and Clement, Andrew: Large-scale AOSD for Middleware. 2004.
- [CRGF08] Coelho, Roberta, Rashid, Awais, Garcia, Alessandro, Ferrari, Fabiano Cutigi, Cacho, Nélío, Kulesza, Uirá, von Staa, Arndt, Pereira de Lucena, Carlos José: Assessing the Impact of Aspects on Exception Flows: An Exploratory Study. 2008.
- [GBFD07] Greenwood, Phil, Bartolomei, Thiago, Figueiredo, Eduardo, Dosea, Marcos, Garcia, Alessandro, Cacho, Nelio; Sant’Anna, Cláudio, Soares, Sergio, Borba, Paulo, Kulesza, Uirá: On the Impact of Aspectual Decompositions on Design Stability: An Empirical Study. 2007.
- [Ha05] Hanenberg, Stefan: Design Dimensions of Aspect-Oriented Systems. 2005
- [KLMM97] Kiczales, Gregor, Lamping, John, Mendhekar, Anurag, Maeda, Chris, Lopes, Cristina, Loingtier, Jean-Marc, Irwin, John: Aspect-Oriented Programming. 2007
- [LKC04] Li, Jingyue, Kvale, Axel Anders and Conradi, Reidar: A Case Study on Improving Changeability of COTS-Based Systems using Aspect-Oriented Programming. 2004.
- [LV99] Lippert, Martin and Videira Lopes, Cristina: A Study on Exception Detection and Handling. 1999.
- [MKL97] Mendhekar, Anurag, Kiczales, Gregor and Lamping, John: RG: A case-study for aspect-oriented programming. 1997.
- [PP04] Papapetrou, Odysseas and Papadopoulos, George A: Aspect Oriented Programming for a component-based real life application: A case study. 2004.
- [Pr01] Prechelt, Lutz. Kontrollierte Experimente in der Softwaretechnik: Springer-Verlag. 2001
- [Jo08] Josupeit-Walter, Manuel: Entwicklung eines Eclipse Plugins zur Aufzeichnung der Aktivitäten eines Entwicklers. 2008
- [Ti97] Tichy, Walter: Should computer scientists experiment more? 1997
- [WBM01] Walker, Robert J., Baniassad, Elisa L.A. and Murphy, Gail C.: An Initial Assessment of Aspect-oriented Programming. In Proceedings of the 21st International Conference on Software Engineering, 1999; S. 120-130.

Werkzeuggestützte Prüfungen dynamischer Prozesse

Thomas Kurpick

thomas.kurpick@rwth-aachen.de

Abstract:

Die Modellierung von Prozessen ist nicht trivial, weil sie verschiedenartigen Bedingungen unterliegt. Die informatische Unterstützung der Prozessmodellierung mittels eines Prozessmodelleditors, der (a) verschiedenartige Bedingungen an Prozessmodellen überprüft und ggf. präzise und konstruktive Fehlermeldungen berechnet sowie (b) die Modellierung von Bedingungen selbst erlaubt, ist demnach wertvoll. Es wurden daher Konzepte erarbeitet, mittels derer ein solcher Prozessmodelleditor mit wenig Aufwand entwickelt werden kann.

1 Zusammenfassung der Diplomarbeit

Die Wertschöpfung aller Unternehmen besteht in der Durchführung spezieller Arbeitsprozesse, im Folgenden einfach Prozesse genannt. *Prozessmanagementsysteme (PMS)* können die Verwaltung und Durchführung dieser Prozesse informatisch unterstützen. *Allgemeines fachliches Prozesswissen* über Prozesse sowie die *Ablaufstruktur* bestimmter Typen von Prozessen als auch *Fortschrittsinformationen* konkreter Prozesse werden über formale und überwiegend graphische Modelle zwischen einem PMS und dessen Benutzern kommuniziert. Diese Modelle sind stark *änderungsanfällig*, insbesondere bei dynamischen Prozessen, die sich nur unvollständig vor ihrer Ausführung modellieren lassen. Modelländerungen unterliegen jedoch gewissen *Rahmenbedingungen*, deren Missachtung zu *technischen Problemen* im PMS (z.B. nicht-initialisierte Variablen) oder auch rein *fachlichen Schwierigkeiten* führt (z.B. Auslassung gesetzlich vorgeschriebener Aktivitäten).

Ziel der Diplomarbeit war daher die Konzeption und Realisierung eines *Prozessmodelleditors*, der PMS-Benutzer (z.B. Versicherungssachbearbeiter) bei der Einhaltung der Rahmenbedingungen unterstützt. Der Prozessmodelleditor ist in der Lage, unterschiedlich abstrakte Prozessmodelle graphisch anzuzeigen und bietet Funktionen für spezielle Modelländerungen. Die Modelländerungen werden hierbei auf technische und fachliche Richtigkeit überprüft.

Die Abstraktionsstufen für Prozessmodelle sind hierbei

- *fachliche Modelle*, die grobe Zusammenhänge zwischen Aktivitäten eines bestimmten Anwendungsbereichs wie der Versicherungsbranche widerspiegeln,
- *Prozessdefinitionen*, die Ablaufstrukturen gleichartiger Prozesse, z.B. die Schadensregulierung, beinhalten sowie

- *Prozessinstanzen*, die z.B. Fortschrittsinformationen konkret durchgeführter Prozesse modellieren.

Prozessdefinitionen werden in dieser Arbeit in der Simplified Business Process Execution Language (SimBPEL) ausgedrückt. SimBPEL ist eine *Variante* des Standards WS-BPEL [Jur06]. Ihr fehlen alle technischen Elemente, wie die genaue Spezifizierung der Schnittstellen und der Datentypen. SimBPEL-Modelle sind bewusst zu detailarm, um Prozesse automatisch zu erzeugen und auszuführen. Ihr Hauptaugenmerk liegt auf dem funktionalen, datenfluss- und verhaltensbezogenen Aspekt [JB97] eines Geschäftsprozesses.

Die *Business Process Compliance Language* (BPCL) dient dazu, besagte *fachliche Zusammenhänge* von Geschäftsprozessen auszudrücken. Dabei steht nicht die Ausführung eines Geschäftsprozesses im Vordergrund, sondern das fachliche Wissen über die Geschäftsprozesse eines Unternehmens. Es ist mit BPCL möglich, *generelle Bedingungen* für Geschäftsprozesse zu definieren. Diese Regeln gelten für alle Geschäftsprozesse eines bestimmten Typs. Die Semantik von BPCL wurde in der Diplomarbeit mittels temporärer Logiken (CTL*) mit Bezug auf gängige mathematische Strukturen (Kripkestrukturen) [EMCGP99] formalisiert.

Abb. 1 zeigt den Prozessmodelleditor mit einem geöffneten BPCL-Modell (oben) und ein SimBPEL-Modell (unten) sowie die entsprechenden Fehler- und Warnmeldungen, die auf Verletzungen technischer und fachlicher Bedingungen hinweisen. In der Prozessdefinition wird die Variable `cd` vor dem ersten Zugriff nicht initialisiert und im Prozessmodelleditor als Fehler markiert. Die fachlichen Bedingungen des BPCL-Modells werden auch durch die Prozessdefinition verletzt und im Editor als Warnungen angezeigt. Zum einen ist die Reihenfolge der beiden Aktivitäten `i` und `r` vertauscht und zum anderen fehlt eine Aktivität vom Typ `C` in der Prozessdefinition. Zusätzlich kann der Prozessmodelleditor mit dem PMS IBM WebSphere Process Server gekoppelt werden, um Prozessinstanzen darin laufender Prozesse darstellen und dynamisch ändern zu können. Auch diese Prozessinstanzen können nach dynamischen Änderungen wieder auf technische und fachliche Bedingungen überprüft werden.

Die Realisierung des Prozessmodelleditors geschah *modellgetrieben*. Dazu wurde die abstrakte Syntax der verschiedenen Prozessmodellarten formal mittels diverser *Metamodelle* definiert. Die verwendete Metamodellierungssprache *Ecore* [BBM03] selbst ist dabei jedoch zu ausdruckschwach zur Formulierung komplizierter struktureller Bedingungen in Prozessmodellen. Daher wurden die Metamodelle um Ausdrücke der *Object Constraint Language* (OCL) [WK99] erweitert.

Durch die Verzahnung der verschiedenen Metamodelle von SimBPEL und BPCL können in OCL-Ausdrücken nicht nur festgeschriebene technische Bedingungen formal definiert, sondern auch variable fachliche Bedingungen. Letzteres wird bewerkstelligt, indem sowohl das zu prüfende Prozessmodell als auch das Modell, das die fachlichen Bedingungen enthält, als ein gemeinsames syntaktisches Objekt aufgefasst werden, auf das sich komplexe OCL-Ausdrücke beziehen. Beide Bedingungsarten wurden also in einem *einheitlichen Ansatz* berücksichtigt.

Der Prozessmodelleditor wurde als *Eclipse-Plugin* [Dau07] realisiert. Für die Definition der abstrakten Modellsyntax der Prozess- und Bedingungsmetamodelle wurde das *Eclipse*

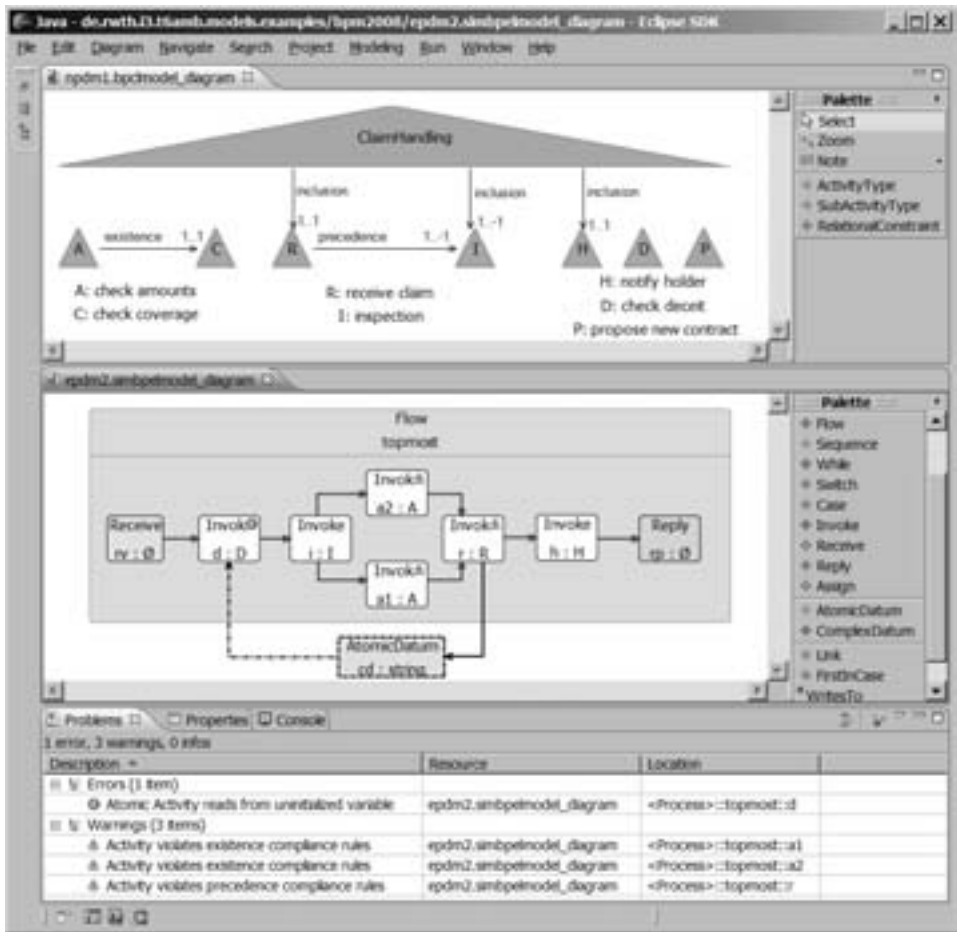


Abbildung 1: Screenshot des Prozessmodelleditors [WKH09b]

Modeling Framework [BBM03] genutzt. Die Definition der konkreten Modellsyntax der Prozesse und Bedingungen wurde mit Hilfe des *Graphical Modeling Framework* [Gro09] erstellt. Da der Ansatz mit vorhandenen Rahmenwerken zur Quellcode-Generierung abgestimmt ist, konnte eine *aufwändige Programmierung* in einer herkömmlichen Programmiersprache *vermieden* werden. Vielmehr können sowohl die Metamodelle der verschiedenen Prozessmodelle als auch die in OCL formulierten Bedingungen zur Berücksichtigung neuer Anforderungen *rasch angepasst* oder *erweitert* werden.

Die Ergebnisse wurden auf verschiedenen Konferenzen vorgestellt und werden demnächst in [WKH09a] und [WKH09b] veröffentlicht.

Literatur

- [BBM03] Frank Budinsky, Stephen A. Brodsky und Ed Merks. *Eclipse Modeling Framework*. Pearson Education, 2003.
- [Dau07] Berthold Daum. *Rich-Client-Entwicklung mit Eclipse 3.3. Anwendungen entwickeln mit Eclipse RCP, SWT, Forms, GEF, BIRT, JPA u.a.m.* dpunkt.verlag, 3., überarb., erw. a.. Auflage, 2007.
- [EMCGP99] Jr. Edmund M. Clarke, Orna Grumberg und Doron A. Peled. *Model checking*. MIT Press, Cambridge, MA, USA, 1999.
- [Gro09] Richard C. Gronback. *Eclipse Modeling Project: A Domain-specific Language Toolkit: A Domain-Specific Language (DSL) Toolkit*. Addison-Wesley Educational Publishers Inc, 1. Auflage, March 2009.
- [JB97] Stefan Jablonski und Markus Böhm. *Workflow-Management*. dpunkt Verlag, 1997.
- [Jur06] Matjaz B. Juric. *Business Process Execution Language for Web Services BPEL and BPEL4WS 2nd Edition*. Packt Publishing, 2006.
- [WK99] Jos Warmer und Anneke Kleppe. *The object constraint language: precise modeling with UML*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.
- [WKH09a] René Wörzberger, Thomas Kurpick und Thomas Heer. Checking Correctness and Compliance of Integrated Process Models. In *Proceedings of the 10th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 2009. to appear.
- [WKH09b] René Wörzberger, Thomas Kurpick und Thomas Heer. On Correctness, Compliance, and Consistency of Process Models. In *Proceedings of the 17th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE)*, 2009. to appear.

Äquivalenzklassen mit Zeit und Zuständen

Martin Leucker*, Moritz Löser†, Dirk Nowotka‡, Joachim Rischen§

Abstract: In diesem Papier wird eine Erweiterung der Klassifikationsbaummethode zur Spezifikation von Echtzeitbedingungen vorgestellt. Ferner wird ein Verfahren beschrieben, welches automatisch für die resultierenden Äquivalenzklassen Testfälle ermittelt, die somit Echtzeitinformationen beinhalten. Die vorgestellte Erweiterung erlaubt die Anwendung der Klassifikationsbaummethode zur Beschreibung von Signalverläufen mit entsprechender automatischer Testfallgenerierung und wurde in Zusammenarbeit mit Bosch Engineering entwickelt, wo sie seit kurzem erfolgreich eingesetzt wird.

1 Überblick

Im Rahmen des äquivalenzklassen-basierten Testens werden Eingabegrößen eines zu entwickelnden Systems in Äquivalenzklassen eingeteilt. Da in der Praxis erwartet werden kann, dass sich ein System bei äquivalenten Eingaben gleich verhält, reduziert sich der Aufwand des Testens eines Systems auf das Testen von Vertretern der einzelnen Äquivalenzklassen.

Die Klassifikationsbaum-Methode [GG93] ist eine Methodik zur iterativen und strukturierten Definition und Beschreibung einzelner Äquivalenzklassen, die üblicherweise aus einem Anforderungsdokument gewonnen werden. Für die Methodik gibt es mit dem Classification -tree Editor (CTE) [Weg] auch eine adäquate Werkzeugunterstützung.

Für die resultierenden Äquivalenzklassen können dann üblicherweise automatisch geeignete Testfälle generiert und auf dem zu testenden System ausgeführt werden. Der manuelle Aufwand des Testens wird somit auf die Erstellung der einzelnen Äquivalenzklassen reduziert.

Im automotive Umfeld findet die Äquivalenzklassenmethode häufig Einsatz, da in diesem Umfeld umfangreiche Anforderungsdokumente vorliegen, die, wenngleich mit viel Aufwand, zur Definition von geeigneten Äquivalenzklassen herangezogen werden können.

Ein Problem der Äquivalenzklassenmethode stellt die Definition von Echtzeitanforderungen und insbesondere die sich anschließende Generierung geeigneter Testfälle dar. Somit lassen sich auch komplexe Signalverläufe nur unzureichend in geeignete Äquivalenzklassen zerlegen, die zeitliche Aspekte berücksichtigen. In diesem Paper wird eine Erweiterung der Klassifikationsbaummethode zur Definition von Echtzeitbedingungen vorgestellt, die eine automatische Generierung von Testfällen für die resultierenden Äquivalenzklassen erlaubt. Aufbauend auf dieser Methode wurde ein Testprozess definiert, der seit kurzem bei Bosch Engineering eingesetzt wird. Durch die Verwendung des Testprozesses wurde in geeigneten Testprojekten bereits eine wesentliche Kostenersparnis erzielt.

*leucker@in.tum.de

†dermoritz@arcor.de

‡dirk.nowotka@informatik.uni-stuttgart.de

§Joachim.Rischen@de.bosch.com

2 Beitrag

Die vorgestellte Arbeit erweitert die Klassifikationsbaummethode zur Spezifikation von Testfällen um Echtzeitbedingungen unter der Bedingung das in der Testfallgenerierung diejenigen Äquivalenzklassen, die nicht in den Anforderungen genannt werden aber trotzdem gültig sind, in jeder möglichen Kombination getestet werden. Genau diese Kombination stellt herkömmliche Verfahren vor große Probleme.

Bisherige Ansätze zur Erweiterung der Klassifikationsbaummethode um das Zeitverhalten von Tests abzubilden, wie in [Con05] und [CK06] beschrieben, eignen sich nur bedingt für unsere Zwecke. In [Con05] wird ein Testfall in ein Sequenz von Testfällen unterteilt, um die zeitliche Abfolge von Äquivalenzklassen zu beschreiben. Dieser Ansatz hat zum Nachteil, dass er erstens das Zeitverhalten vom Äquivalenzklassenbaum trennt und zweitens die unübersichtliche Darstellung von Sequenzen mit vielen Schritten. Der zweite Nachteil wurde zwar in der Arbeit [CK06] adressiert, aber die prinzipiellen Einwände gegen die Trennung des Zeitverlaufs vom Äquivalenzklassenbaum bleiben bestehen. Diese Trennung heben wir mit der vorgestellten Darstellung auf. Dies wirkt sich außerdem auf die Lösung des Problems der Testfallgenerierung mit erschöpfender Kombination von Äquivalenzklassen, die nicht in den Anforderungen aufgeführt werden aber trotzdem gültig sind, positiv aus. Außerdem werden in unserem Ansatz Ausgabewerte zusammen mit den Eingaben modelliert. Dies erlaubt die automatische Generierung nicht nur von Eingaben für Testfälle sondern gleichzeitig auch die erwartete Ausgabe, so dass die generierten Testfälle automatisch auf dem zu untersuchenden System sowohl abgelaufen und das System hinsichtlich Korrektheit überprüft werden kann. Die vorgestellte Erweiterung des Äquivalenzklassentests wird semantisch (intern) mit Hilfe von Zeitautomaten modelliert. Zeitautomaten bilden einen ausdrucksstarken Formalismus, der sich besonders bei der Erzeugung der Testfälle auszahlt. Weitere Einzelheiten sind in [Lös08] beschrieben. Abbildung 1 dient

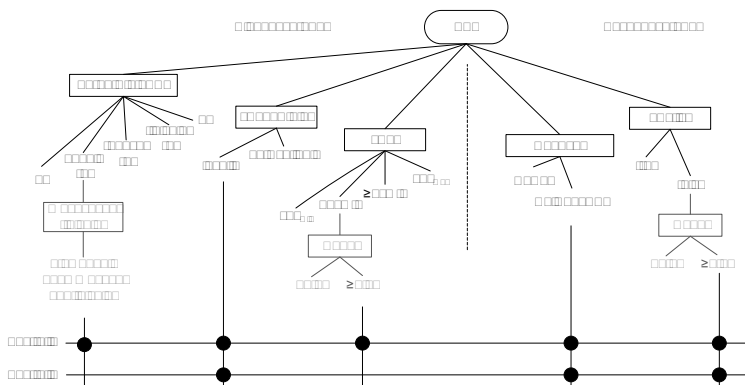


Abbildung 1: Beispiel BBF

als laufendes Beispiel zur Illustration unseres Ansatzes. Die zwei beschriebenen Testfälle stehen für die folgenden Anforderungen: (1) „Wenn TargetStateBBF von enable auf

finishing wechselt und vVeh_Status den Wert invalid besitzt und vVeh für mindestens 0,4s unter 3km/h ist dann setzt BBF ReqType auf den Wert RelReq_BBF und bActive auf den Wert false.“ und (2) „Wenn vVeh_Status den Wert invalid besitzt dann setzt BBF ReqType auf den Wert RelReq_BBF und bActive auf den Wert false.“

Der erste Testfall verlangt, dass die beiden Forderungen „0,1s enable und dann Wechsel auf finishing“ für TargetStateBBF und „unter 3km/h für mindestens 0,4s“ für vVeh im Testfall zum selben Zeitpunkt realisiert werden müssen. Eine Konjunktion von Stellen der entsprechenden Zeitautomaten zusammen mit einem geeigneten Füllwert setzt diese Anforderung um (siehe Abb. 2).

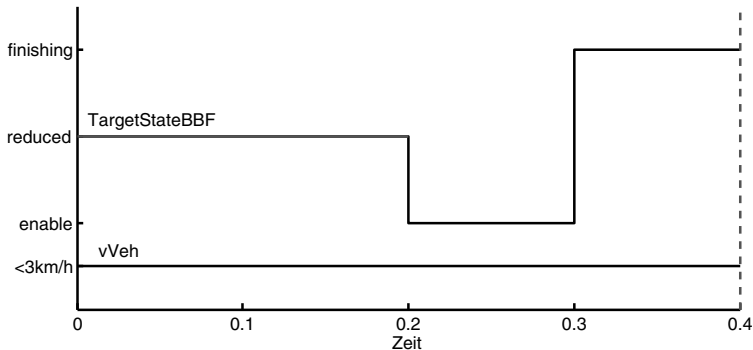


Abbildung 2: Füllwert

Zur Realisierung des zweiten Testfalls ist die sequenzielle Kombination von insgesamt neun Kombinationen (inklusive der Kombination von zeitbehafteten Klassen) notwendig. Im vorgestellten Ansatz entsprechen diese Kombinationen eine Kombination der Zeitautomaten für jede Äquivalenzklasse derart, dass im Fall festgelegter Anforderungen bestimmte Stellen und im Fall offener Anforderungen alle Stellen durchlaufen werden müssen. Abbildung 3 illustriert das Resultat einer solchen Kombination für den zweiten Testfall.

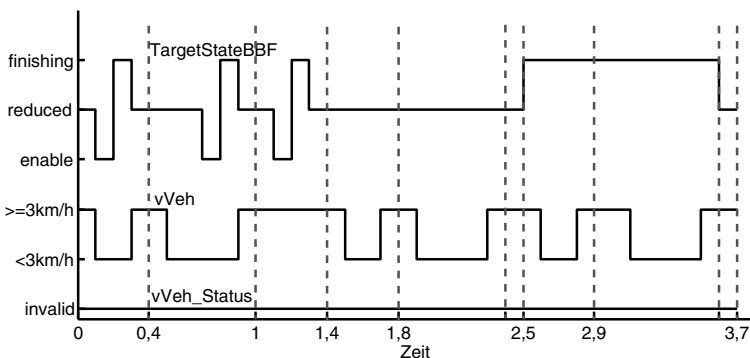


Abbildung 3: Kombination

Gerade diese Möglichkeit der automatischen Testfallgenerierung in einer mit Echtzeitbedingungen erweiterten Klassifikationsbaummethode und einer vollständigen Kombination nicht mit Anforderungen belegten Äquivalenzklassen stellt eine Besonderheit des vorgestellten Ansatzes dar.

3 Evaluation

Das hier vorgestellte Verfahren wird seit Oktober 2008 im Bereich Modultest eingebetteter Systeme bei Bosch Engineering produktiv eingesetzt. Die ersten Ergebnisse sind sehr vielversprechend. Wie erwartet konnte insbesondere für Tests von Echtzeitsystemen sowohl die Testeffizienz als auch die Testeffektivität erhöht werden. Durch die Einführung der Klassifikationsbaummethode wurde zudem die Qualität der Dokumentation verbessert. Damit einhergehend ist eine Verbesserung der Tests bezüglich ihrer Überprüfbarkeit und Anpassbarkeit.

Ein wesentlicher Nachteil der Klassifikationsbaummethode, nämlich der hohe zeitliche Aufwand zur Erstellung der Klassifikationsbäume, bleibt erhalten. Insbesondere die für die Prozessautomatisierung nötige Attributierung der Klassifikationsbäume ist sehr zeitintensiv. Zum einen die hohe Anzahl der Attribute aber auch einige Mängel in der Benutzerschnittstelle des CTE sind die Ursachen dafür. Weiterhin besteht die Problematik, dass im Rahmen der Klassifikationsbaummethode genau eine oder keine Klasse pro Klassifikation gewählt werden muss, so dass ungleich- und oder-Bedingungen nicht direkt darstellbar sind. Um im obigen Sinn vollständig zu Testen muss für jede Möglichkeit solche Bedingungen zu erfüllen, ein Testfall angelegt werden. Als Ausweg könnte zukünftig der in [CK06] vorgestellte Ansatz dienen, in dem Klassen durch verschiedene Symbole mit entsprechender Semantik markiert werden können. Eine Integration der in [CK06] mit der hier vorgestellten Erweiterung um Realzeit-Bedingungen scheint jedoch problemlos möglich und ist Teil zukünftiger Arbeit, die Modifikationen am CTE nach sich ziehen.

Literatur

- [CK06] M. Conrad und A. Krupp. An Extension of the Classification-Tree Method for Embedded Systems for the Description of Events. *Electronic Notes in Theoretical Computer Science*, 164(4):3–11, 2006.
- [Con05] Mirko Conrad. Systematic Testing of Embedded Automotive Software - The Classification-Tree Method for Embedded Systems (CTM/ES). In Ed Brinksmä, Wolfgang Grieskamp und Jan Tretmans, Hrsg., *Perspectives of Model-Based Testing*, number 04371 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2005.
- [GG93] K. Grimm und M. Grochtmann. Classification Trees for Partition Testing. *Software Testing, Verification and Reliability*, 3(2):63–82, 1993.
- [Lös08] Moritz Löser. Äquivalenzklassen-basiertes Testen. Diplomarbeit, Universität Stuttgart, 2008.
- [Weg] Classification Tree Editor, verfügbar unter: <http://www.systematic-testing.com>.

Aggregationsverfahren für Software-Komponenten

Dipl.-Inf. (FH) Frank Nordemann

Fachhochschule Osnabrück
Fakultät Ingenieurwissenschaften und Informatik
info@franknordemann.de

Art der Arbeit: Auszug aus der Diplomarbeit „Untersuchung eines Komponentenmodells für .NET Rich Client Applications“

Betreuer/in der Arbeit: Prof. Dr. Frank M. Thiesing (Fachhochschule Osnabrück), Dr. Rainer Roosmann (ROSEN Technology and Research Center GmbH)

Abstract: Die Aggregation (Vereinigung) von Software-Komponenten ist ein bedeutendes Merkmal der komponentenbasierten Software-Entwicklung. Mit Hilfe der Aggregation von Komponenten können mit der Reduktion von Komplexität und der Geheimhaltung von Informationen wesentliche Ziele der Komponentenorientierung erreicht werden. Zur Realisierung der Komponentenaggregation bieten sich im Umfeld des .NET-Frameworks von Microsoft eine physische sowie eine logische Vorgehensweise an. Die erste Aggregationstechnik beschreibt die physische Zusammenführung der Subkomponenten zu einer neuen Makrokomponente. Bei der logischen Aggregation bleiben dagegen die Subkomponenten vorhanden. Richtlinien stellen in diesem Fall den korrekten Gebrauch des Aggregats sicher. Der nachfolgende Artikel beschreibt die unterschiedlichen Konzepte und führt ihre Vor- und Nachteile auf.

1 Einleitung

Die Software-Technik steht immer komplexeren und umfangreicheren Software-Projekten gegenüber. Die komponentenbasierte Software-Entwicklung hat das Ziel, die Komplexität heutiger Projekte in einem beherrschbaren Rahmen zu halten, die Kosten für die Entwicklung zu reduzieren und den Entwicklungsprozess durch den Einsatz von wiederverwendbaren Komponenten zu beschleunigen [Zw05]. Dabei setzt sie auch auf das Zusammenführen von vorhandenen Software-Komponenten zu neuen Aggregaten. Nachfolgend werden mit der physischen und der logischen Aggregation zwei Verfahren vorgestellt, mit denen vorhandene Software-Komponenten vereinigt werden können.

Den in diesem Artikel vorgestellten Konzepten liegt das .NET-Framework von Microsoft als Softwareentwicklungsplattform zu Grunde. Die Verwendung der Aggregationsverfahren ist jedoch nicht auf das .NET-Framework begrenzt: Sofern die nötigen Werkzeuge auch für andere Plattformen verfügbar sind, ist eine Portierung der Konzepte durchaus vorstellbar.

2 Aggregationskonzepte zur Bildung von Makrokomponenten

Das Zusammenführen von Software-Komponenten zu Aggregaten, sogenannten Makrokomponenten, ist ein mächtiges Werkzeug der komponentenbasierten Software-Entwicklung. Die Bildung von Aggregaten unterstützt die Einhaltung der Prinzipien der Abstraktion und des Information Hidings [Vo05], die wesentlich zur Komplexitätsreduzierung von wiederverwendbaren Komponenten beitragen. Verwendet ein Entwickler eine Makrokomponente, befasst er sich ausschließlich mit den Eigenschaften der Makrokomponente und erhält keine Informationen über die aggregierten Subkomponenten. Makrokomponenten entsprechen daher dem Prinzip einer Black-Box [An04]. Die Realisierung von Makrokomponenten kann mit Hilfe unterschiedlicher Konzepte umgesetzt werden.

2.1 Physische Aggregation

Die erste Möglichkeit zur Bildung von Makrokomponenten beschreibt die physische Aggregation. Bei diesem Verfahren werden die zu aggregierenden Komponenten physisch zu einer Makrokomponente zusammengefügt. Konkret bedeutet dies, dass die .NET-Assemblies der Subkomponenten im Vereinigungsprozess (Merge-Vorgang) zu einer neuen Makrokomponentenassembly zusammengesetzt werden. Abbildung 1 veranschaulicht den Merge-Vorgang an den Assemblies (.dll-Dateien) der Komponenten A, B, und C.

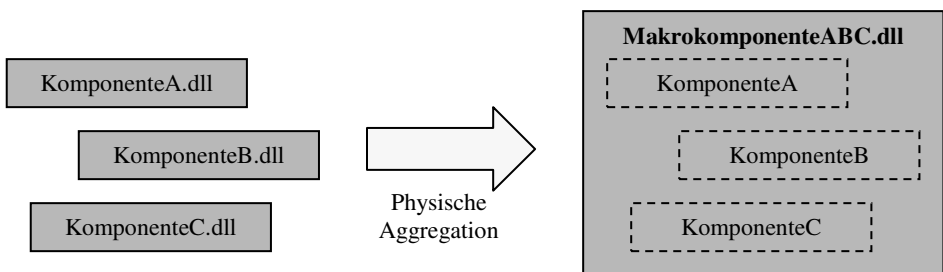


Abbildung 1: Physische Aggregation von Komponenten

Die drei Komponenten werden durch den Merge-Vorgang zu einer neuen Makrokomponente ABC zusammengefügt. Realisiert wird der Merge-Vorgang durch das freie Kommandozeilenprogramm ILMerge [II09] von Microsoft. Ein Software-Entwickler verwendet anschließend nur noch die Makrokomponente ABC. Über die darin gekapselten Subkomponenten A, B, und C benötigt und erhält er keine Informationen (Black-Box-Prinzip [An04]). Um den Gebrauch der Makrokomponente ohne Zugriff auf die enthaltenen Komponenten zu realisieren, verfügt das Aggregat über eine Fassade [Ga95]. Die Fassade wird vor dem Zusammenführungsprozess von einem Komponentenentwickler geschrieben und reicht nur die Funktionen der Subkomponenten nach außen, die von der Makrokomponente zur Verfügung gestellt werden sollen. Wichtig ist in diesem Zusammenhang, dass jegliche Kommunikation mit der Außenwelt über die Fassade geführt wird. Ein direkter Zugriff auf die enthaltenen Subkomponenten verletzt das Architektur-Prinzip des Information-Hidings [Vo05] und ist untersagt.

2.2 Logische Aggregation

Eine Makrokomponente nach logischem Konzept besitzt einen wesentlichen Unterschied zu ihrem physischen Äquivalent: Die zu aggregierenden Subkomponenten werden nicht im physischen Sinn zu einem Aggregat zusammengefügt, sondern verbleiben in ihrer bisherigen Form als .NET-Assemblies. Aus diesem Grund besitzen logische Makrokomponenten Abhängigkeiten zu den Assemblies ihrer Subkomponenten und sind mit diesen auszuliefern. Abbildung 2 veranschaulicht den logischen Aggregationsprozess.

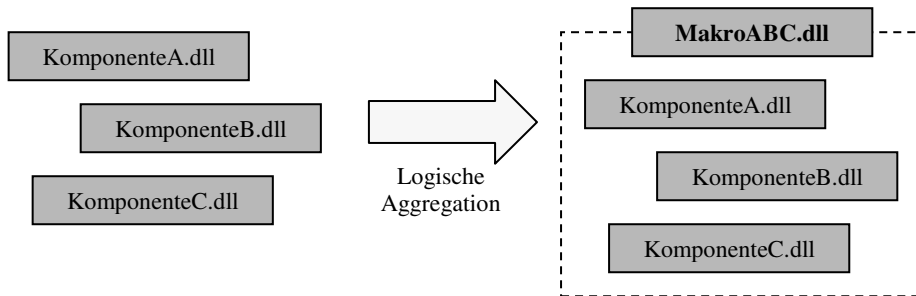


Abbildung 2: Logische Aggregation von Komponenten

Der logische Aggregationsprozess besteht aus der Erstellung der Komponente Makro-ABC. Sie entspricht dem Muster einer Fassade und kapselt den Zugriff auf die Subkomponenten. MakroABC stellt somit die logische Makrokomponente dar. Wie in Abbildung 2 angedeutet, ist sie jedoch nicht ohne die Assemblies der Subkomponenten A, B und C funktionsfähig. Eine Referenzierung der .NET-Assemblies der Subkomponenten sowie ihre direkte Verwendung verletzen die Richtlinien der logischen Aggregation und können mit Hilfe von Programmen wie FxCop [Fx09] festgestellt und unterbunden werden.

3 Fazit

Die Aggregation als ein Mittel der beständigen Komposition von Software-Komponenten unterstützt die komponentenbasierte Software-Entwicklung maßgeblich bei der Reduktion von Komplexität und sollte Bestandteil jedes Komponentenmodells sein. Software-Entwickler können sich bei der Verwendung von Aggregaten (Makrokomponenten) auf die konkret von der Aggregation definierten Anforderungen konzentrieren, interne Details werden nach dem Prinzip einer Black-Box [An04] vor ihnen verborgen.

Um die Aggregation von Komponenten zu ermöglichen, bieten sich grundsätzlich das physische und das logische Aggregationsverfahren an. Beiden Vorgehen gemein ist zunächst die Erstellung einer Fassadenkomponente, die den Zugriff auf die Subkomponenten steuert und die Funktionalität des Aggregates freigibt. Beim physischen Konzept werden daraufhin die Subkomponenten mit der Fassade in eine Makrokomponente zusammengeführt. Die Komponente besteht physisch aus einer .NET-Assembly, die durch die Kapselung der Subkomponenten keine Abhängigkeiten zu weiteren Assemblies besitzt.

Bei der logischen Aggregation erfolgt dagegen keine Zusammenführung der Fassade mit den Subkomponentenassemblies. Bei diesem Konzept regeln Vorschriften die Verwendung der entstehenden Makrokomponente und unterbinden den direkten Zugriff auf die Subkomponenten.

Durch die Kapselung der Subkomponenten in der entstehenden Makrokomponente besteht das Resultat des Merge-Vorgangs bei der physischen Aggregation aus einer einzigen .NET-Assembly, die keine Abhängigkeiten zu weiteren Assemblies besitzt. Dieses Ergebnis stellt eine gute Grundlage zur Schaffung von Black-Box-Komponenten dar. Die logische Aggregation kann derartige Vorgaben nur durch die Definition von Richtlinien und die Verwendung zusätzlicher Werkzeuge zu deren Überprüfung gewährleisten. Das physische Zusammenfügen lässt allerdings auch Nachteile mit beträchtlichen Konsequenzen entstehen, die beim logischen Konzept ausgeschlossen werden können: Aggregieren mehrere Makrokomponenten die gleiche Subkomponente, besteht die Gefahr von Quellcoderedundanzen. Werden diese Makrokomponenten anschließend in einer gemeinsamen Anwendung benutzt, können Komplikationen nicht ausgeschlossen werden. Darüber hinaus tragen Coderedundanzen nicht zu einer soliden Software-Architektur bei. Wird eine Aktualisierung einer Subkomponente vorgenommen, sind alle Aggregate neu zu erstellen und auf Funktionsfähigkeit zu prüfen. Weitere Untersuchungen zur Flexibilität von Software-Komponenten werden in [St08/09] durchgeführt.

Eine Entscheidung für den Gebrauch des physischen oder des logischen Aggregationsverfahrens sollte nach einer gründlichen Prüfung der eigenen Anforderungen getroffen werden. Allgemein kann die Verwendung von physischen Aggregationen nur bei Projekten empfohlen werden, die einen begrenzten Umfang besitzen und deren Komponenten später nicht in anderen Anwendungen wiederverwendet werden sollen. Bei mittleren und größeren Projekten ist das logische Konzept zu bevorzugen, da es durch den Verzicht auf eine physische Kapselung flexibler ist und weniger Komplikationen zu erwarten sind. Alternativ kann die physische Aggregation das logische Vorgehen unterstützen, indem es die Anzahl der Komponenten assemblies zur Auslieferungszeit minimiert, aber nicht Bestandteil des eigentlichen Komponentenmodells wird.

Literaturverzeichnis

- [An04] Andresen, A.: Komponentenbasierte Softwareentwicklung mit MDA, UML 2 und XML, Carl Hanser Verlag, München, 2004.
- [Fx09] FxCop, <http://msdn.microsoft.com/en-us/library/bb429476.aspx>, März 2009.
- [Ga95] Gamma, E. et al.: Design Patterns – Elements of Reusable Object-Oriented Software, Addison Wesley, 1995.
- [Il09] ILMerge, <http://www.microsoft.com/downloads/details.aspx?familyid=22914587-b4ad-4eae-87cf-b14ae6a939b0&displaylang=en>, März 2009.
- [St08/09] „Strategische Flexibilität durch komponentenbasierte Software-Entwicklung“ ist ein AGIP-Forschungsprojekt an der FH Osnabrück in Kooperation mit der RTRC GmbH.
- [Vo05] Vogel, O. et al.: Software-Architektur – Grundlagen - Konzepte - Praxis, Spektrum Akademischer Verlag, Heidelberg, 2005.
- [Zw05] Zwintzsch, O.: Software-Komponenten im Überblick – Einführung, Klassifizierung & Vergleich von JavaBeans, EJB, COM+, .Net, CORBA, UML 2, W3L Verlag, 2005.

Integrating a Legacy Terminal Application into an SOA

Michael Dörsam, Stefan Gründling,
Thomas Langholz, Sascha Roth, Alexander Steinbrecht

Hochschule Darmstadt - University of Applied Science
— Fachbereich Informatik —

Abstract: This paper is an experience report of an SOA integration project in cooperation with Software AG University Relations. An application of separating application and technical infrastructure architecture is demonstrated.

Keywords: SOA Integration, Screen Scraping, ApplinX, EntireX, IBM 3270

1 Introduction

Today's organizations' application landscapes still consist of terminal based legacy systems, running on reliable mainframes. Grown over the past decades, these legacy systems are involved into mission critical parts of organizations' core businesses. Within the establishment of a service-oriented application landscape, these mission critical systems have to be integrated in a loosely coupled way.

The related project's mission is to implement an online portal for a luxury travel agency. This portal uses different backend systems. One of these backend systems is provided by Software AG (SAG) University Relations [SAG09-4]. This backend system simulates a mainframe within a virtual machine (VM).

A legacy application written in Natural [SAG09-3] called *SAG-Tours* is running on this mainframe. *SAG-Tours* is accessible via an IBM 3270 terminal for maintaining and booking of yachts. The project's major goal is the integration of *SAG-Tours* and backend systems for hotel and flight booking to provide a state-of-the-art web based travel portal for luxury travels.

Besides of these backend systems there are added-value services integrated into the portal to give customers general information about the chosen destination e.g. weather and literature recommendations. These added value services are shared from third-party providers on the Internet, e.g. Amazon or Wikipedia.

2 Project Management Issues

Since a full report of the project approach is beyond the scope of this paper, this section focuses on project meeting culture and risk management.

During weekly project meetings the participants have to answer three major questions: *What have you done? What open issues do you have? What will you do next?* Based on this information, we are (re-)assessing the project plan from the begin of the semester.

Furthermore, interface specifications are communicated during project meetings to move the different teams towards relatively autonomous work. Also, the architecture (see sec. 3) enhances the discussion and assures reasonable interfaces.

Plenty of new products are used to create the new portal. That requires an *evaluation phase* of new development tools and integration platforms which takes place on every development cycle. Each evaluation has a *strategic deadline*. When due, the product/technology has to be up and running without any project compromising bugs. Otherwise, the technology has to be switched to an alternative, e.g., Open Source products. This should reduce the risk of a project failure through technology risks.

3 Architecture

The architecture follows the principle of *separation of concerns*. Concretely, logical applications and physical implementation are separated.

The *logical* Application Landscape (AL) Component Architecture [EHH+08, p. 159ff] is shown in Figure 1. It hides the complexity of the actual implementation, and shows which of the different AL components are interacting.

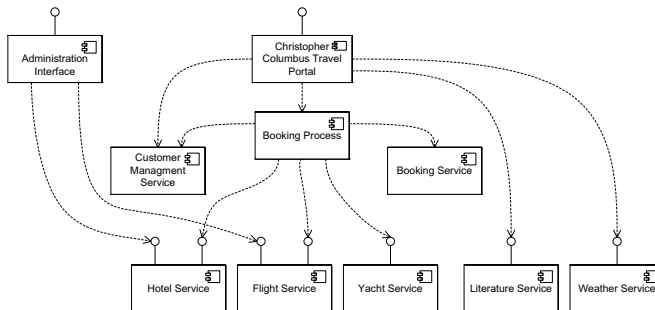


Figure 1: Application Landscape Component Architecture

Figure 2 illustrates the *Technical Infrastructure (TI) Component Architecture* [EHH+08, p. 196ff] as a four-tier architecture. *Client Tier* components provide an interface to end users. E.g. the travel portal is accessible via a web browser. Components of the *Business Process Tier* implement a business processes via orchestration of services within lower

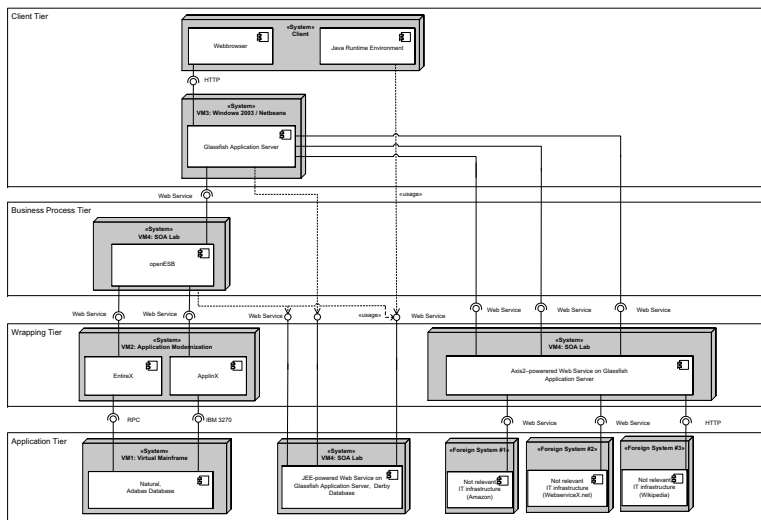


Figure 2: Technical Infrastructure Component Architecture

tiers. The *Wrapping Tier* provides Web Service interfaces. Components within this tier realize an adapter for the Application Tier. Caching algorithms are used to guarantee reliable response times and to enforce better decoupling to upper tiers. The *Application Tier* provides components with any kind of interfaces. These components can be used or wrapped by upper tiers. Also, foreign systems like Amazon or Wikipedia reside in this tier.

An example for the complexity reduction of the AL component architecture is given by the literature service. Literature recommendations are shown to a customer in the portal on the Client Tier. To do that, the portal calls a Web Service on the Wrapping Tier which implements an adapter and caching algorithms to provide proper response times. The data itself resides in a component on the Application Tier provided by Amazon. Another example is the service *booking process* in the Business Process Tier. This service orchestrates different services of the Wrapping and Application Tier.

4 Implementation

Screen scraping is a variant used to integrate applications of legacy systems by accessing its presentation. Because this variant uses the presentation, neither changes to the source code are required nor the legacy system has to go offline during deployment of the new application. In our project a proprietary product called ApplinX [SAG09-1] is used to integrate the IBM 3270 terminal of SAG-Tours.

Besides ApplinX, EntireX [SAG09-2] is used to access SAG-Tours programmed in Natural [SAG09-3] via Remote Procedure Calls (RPCs) to share the method as Web Services. The

portal itself is implemented with Java Server Faces (JSF). A Business Process Execution Language (BPEL) Engine which is part of OpenESB [Col09] is used for a travel booking process.

5 Learnings

At the end of the project we conclude the following learnings:

Discuss interface specifications as early as possible to enable teams to work autonomously. Collaborative work can be enforced due to early interface specification. Combined with loosely coupled components a high degree of autonomous development is supported.

Technical evaluation of project critical development tools and runtime environments is crucial for a successful project. Software for implementing an SOA still needs to mature. A validation of WSDLs against their XML-Schema does not guarantee any compatibility to other vendors. We exchanged the Business Process Engine of SAG because of compatibility issues while integrating Java Enterprise Edition (JEE) Web Services. Within this context, sticking to deadlines is necessary to guarantee the project's success.

In our project, the architecture (see sec. 3) evolved over time. The AL architecture can be used to discuss new features. Combined with a mapping to the TI architecture, developers know what systems are affected by a specific new feature. The four-tier TI architecture structures the application landscape, even for future components.

References

- [Col09] CollabNet: *OpenESB, The Open Source ESB for SOA & Integration*.
<https://open-esb.dev.java.net>, last access: 10.01.2009
- [EHH+08] Gregor Engels, Andreas Hess, Bernhard Humm, Oliver Juwig, Marc Lohmann, Jan-Peter Richter, Markus Voß, Johannes Willkomm: *Quasar Enterprise – Anwendungslandschaften serviceorientiert gestalten*. dpunkt.verlag, Heidelberg, 2008.
- [Mas07] Dieter Masak: *SOA? Serviceorientierung in Business und Software*. Springer Verlag, Heidelberg, 2007.
- [SAG09-1] Software AG: *Applinx*.
<http://www.softwareag.com/de/products/az/applinx.asp>, last access: 10.01.2009
- [SAG09-2] Software AG: *EntireX*.
<http://www.softwareag.com/de/products/az/entirex.asp>, last access: 10.01.2009
- [SAG09-3] Software AG: *Natural*.
<http://www.softwareag.com/de/products/natural/default.asp>,
last access: 10.01.2009
- [SAG09-4] Software AG: *University Relations*.
<http://www.softwareag.com/university>, last access: 10.01.2009

Entwicklung eines neuartigen Optimierungsverfahrens für die automatische Generierung von Testfällen

Andreas Rothmann

Hochschule Offenburg
Angewandte Informatik
andy.rothmann@gmx.de

Art der Arbeit:
Betreuer der Arbeit:

Bachelor-Thesis
Prof. Dr.-Ing. Daniel Fischer

Abstract: Diese Ausarbeitung beschäftigt sich mit der automatischen Testfallerstellung mit Hilfe eines evolutionären Algorithmus. Als Vorbild dienen Algorithmen, welche im Bereich des Pairwise Testing genutzt werden. Eine Einschränkung dieser Algorithmen ist, dass sie nur einfache Zielfunktionen benutzen, welche lediglich das Vorhandensein aller möglichen paarweisen Kombinationen bei der Testskripterstellung gewährleisten. Benutzt man einen evolutionären Algorithmus, besteht die Möglichkeit, die Reihenfolge der Testfälle im Skript zu optimieren und so eine Kosten- und Zeitersparnis zu erreichen. Leistungsmessungen haben ergeben, dass durch diesen Ansatz eine erhebliche Optimierung möglich ist.

1 Einleitung

Der Entwurf effizienter Testfälle stellt immer noch eine der Herausforderungen im Bereich des Testdesigns dar. Vielfältige Methoden des Testdesigns - wie beispielsweise Äquivalenzklassenbildung, Grenzwertanalyse, Test besonderer Werte, Evolutionäres Testen, Zustandsorientiertes Testen, Fuzz Testing sowie Stress Testing – sind heute Stand der Technik ([Lig02], [Hof08]). Während einige dieser Methoden zufallsgesteuert Testdaten generieren, erheben andere Methoden den Anspruch, strukturiert Testfälle zu erzeugen, mit welchen eine möglichst hohe Testüberdeckung (Test Coverage) erreicht wird.

Diese Methoden berücksichtigen allerdings nicht explizit so genannte Doublemode-Fehler, welche im Gegensatz zu Singlemode-Fehlern nur dann in Erscheinung treten, wenn zwei Eingangsparameter oder zwei Systemkomponenten gleich sind, beziehungsweise aus der gleichen Äquivalenzklasse stammen. Beim Pairwise Testing handelt es sich um eine Methodik des Testdesigns, die sicherstellt, dass jede mögliche Kombination zweier Eingangsparameters oder zweier Systemkomponenten mindestens einmal auftritt und dabei die Anzahl der notwendigen Testfälle minimiert wird ([LT98], [Wil00], [GOA04]).

Vielfältige Algorithmen wurden bisher für Pairwise Testing entwickelt. Meist wird dabei bei der Bewertung (Zielfunktion) nur berücksichtigt, dass jede paarweise Kombination nur einmal auftritt. In der Praxis sind allerdings gerade im Bereich des Systemtests weitere wichtige Randbedingungen zu berücksichtigen. Einzelne Systemkomponenten sind mit wesentlich höherem Zeit- oder Kostenaufwand wie andere zu testen, ebenso sollten bei der Abarbeitung der Testfälle möglichst wenige Konfigurationsänderungen aufgrund der notwendigen Umrüstzeiten stattfinden. Unter Berücksichtigung dieser und weiterer Randbedingungen ergibt sich somit ein nicht lineares Optimierungsproblem. Im Rahmen dieser Arbeit wird ein neuartiger evolutionärer Algorithmus beschrieben, welcher das gegebene Problem in endlicher Zeit löst.

Als mögliche Anwendungsgebiete seien die Telekommunikations- und die Automobilbranche genannt. In beiden Bereichen müssen beim Systemtest unterschiedliche Konfigurationen von Systemkomponenten paarweise getestet werden. Beispielsweise müsste man sicherstellen, dass jede Motorvariante mit jeder Variante des Bordcomputers funktioniert.

2 Neuartiger Ansatz mit einem evolutionären Algorithmus

Mittels evolutionärer Algorithmen lassen sich vielfältige nicht lineare Optimierungsprobleme lösen ([ES03]). Im Gegensatz zu anderen klassischen Optimierungsverfahren werden mehrere mögliche Lösungen (Individuen) gleichzeitig betrachtet. Alle Individuen bilden die so genannte Population. Bei evolutionären Algorithmen wird die biologische Evolution nachgestellt. Ausgehend von einer Anfangspopulation werden durch Rekombination neue Individuen generiert, welche die Gene (Eigenschaften) von den Eltern erben. Mittels einer Kostenfunktion wird die Fitness der Individuen bestimmt. Durch Mutation werden einzelne Gene verändert, während durch die Selektion weniger optimale Lösungen verworfen werden. Über mehrere Generationen hinweg entstehen somit immer besser angepasste Individuen. Nach dem Abbruch stellt das beste Individuum die beste gefundene Problemlösung dar.

Einige Teilschritte eines evolutionären Algorithmus sind problemunspezifisch, während andere Teilschritte wiederum problemspezifisch sind. Bei der vorliegenden Problemstellung sind Testfälle so zu generieren (Testskript), dass der Zeit- und Kostenaufwand für die Durchführung der Testfälle minimiert wird. Ebenso ist zu gewährleisten, dass jede paarweise Kombination mindestens einmal auftritt. Ein Individuum stellt somit ein mögliches Testskript dar, wie in Tabelle 1 exemplarisch dargestellt ist. Ein System soll dabei aus den Systemkomponenten A, B und C bestehen. Jede Systemkomponente sei in zwei Varianten vorhanden.

Testfall	Komponente A	Komponente B	Komponente C
1	A1	B1	C1
2	A2	B1	C2
3	A1	B2	C2
4	A2	B2	C1

Tabelle 1: Testskript als Individuum

Problemspezifisch ist auch die Bewertung eines Individuums durch eine Kostenfunktion, die jedem Individuum eine Fitness zuweist. Diese Fitness hat Auswirkungen darauf, welche Individuen bei der Rekombination verwendet und welche bei der Selektion verworfen werden. Bei der Selektion wurde als Selektionsstrategie „Elitismus“ realisiert.

Ein evolutionärer Algorithmus stellt wiederum selbst ein Optimierungsproblem dar, da dieser durch viele Parameter, wie beispielsweise die Größe der Population oder die Realisierung einer Selektionsstrategie, in seinem Konvergenzverhalten beeinflusst wird. Folglich waren umfangreiche Untersuchungen notwendig, um das Konvergenzverhalten zu optimieren.

3 Experimentelle Resultate

In diesem Kapitel sollen exemplarisch zwei Ergebnisse dieser Untersuchungen vorgestellt werden. Abbildung 1 zeigt hierbei die Fitness des besten Individuums in Abhängigkeit von der Generation. Während im linken Diagramm die Kosten (Reziprokwert der Fitness) für die Durchführung eines Systemtests um 20% gesenkt werden konnte, waren es im rechten Diagramm durch Variation der Parameter des evolutionären Algorithmus sogar 30%.

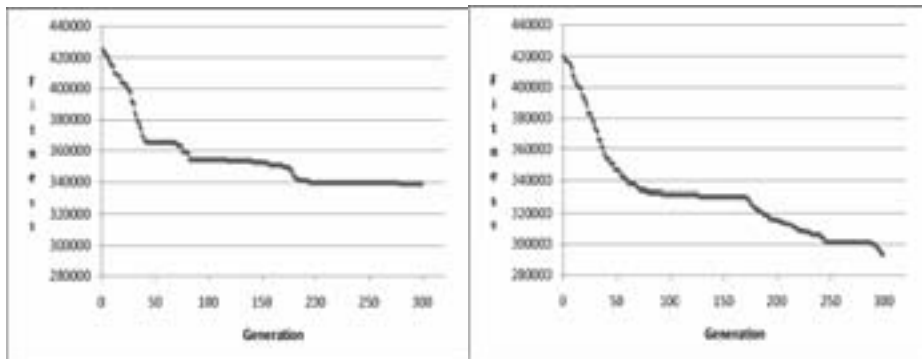


Abbildung 1: Verlauf der Fitness (bestes Individuum)

4. Zusammenfassung

Die Ergebnisse zeigen, dass es durch den evolutionären Algorithmus möglich ist, bis zu 30% an Kosten und Zeit für einen Systemtest einzusparen. Der Algorithmus lässt sich ebenso noch erweitern. Denkbar wäre der Einsatz von weiteren Selektionsstrategien wie beispielsweise „Monte Carlo“ oder „Ranking Verfahren“. Auch die Performance des erstellten Prototypen lässt sich durch den Einsatz von Datenstrukturen verbessern, welche auf Zugriffsgeschwindigkeit optimiert sind.

Literatur

- [GOA04] Grindal, Matts; Offutt, Jeff; Andler, Sten F.: Combination Testing Strategies: A Survey. GMU Technical Report ISE-TR-04-05, July 2004.
- [Hof08] Hoffmann, Dirk: Software-Qualität. Springer, 2008.
- [Lig02] Liggesmeyer, Peter: Software-Qualität – Testen, Analysieren und Verifizieren von Software. Spektrum Verlag, 2002.
- [ES03] Eiben, Agoston E.; Smith, James E.: Introduction to Evolutionary Computing. Springer Verlag, 2003.
- [LT98] Lei, Y.; Tai, K.C.: In-parameter-order: A test generation strategy for pair-wise testing. In Proceedings of the third IEEE High Assurance Systems Engineering Symposium, pp. 254-261, IEEE, November 1998.
- [Wil00] Williams, A.W.: Determination of test configurations for pair-wise interaction coverage. In Proceedings of the 13th International Conference on the Testing of Communicating Systems (TestCom 2000), Ottawa, Canada, pp. 59-74, August 2000.

Erstellung und Anwendung eines Usability-Kriterienkatalogs für Online-Bewerbungssysteme an deutschen Hochschulen

Volha Abrazhevich

Leibniz Universität Hannover
Das Fachgebiet Software Engineering
avolha@gmx.de

Art der Arbeit: Masterarbeit

Betreuer der Arbeit: Prof. Dr. Kurt Schneider, M. Sc. Leif Singer

Abstract: Bislang existiert im Bereich Online-Bewerbungssysteme an deutschen Hochschulen noch keine allgemein anerkannte Vorgehensweise dafür, wie mit Problemen der Benutzbarkeit umgegangen werden soll. Dieses wirkt sich auf die Bewerber aus, da ihre Bewerbung möglicherweise unnötig verkompliziert wird. Im Rahmen der folgenden Masterarbeit soll in Zusammenarbeit mit der Hochschul-Informations-System GmbH ein Usability-Kriterienkatalog für Online-Bewerbungssysteme an deutschen Hochschulen erstellt und auf ein bestehendes Online-Bewerbungssystem angewandt werden.

1 Einführung

Usability steht für die Benutzbarkeit von Softwaresystemen und stellt ein wichtiges Qualitätsmerkmal eines Softwaresystems dar. Um beim Entwickeln eines Softwaresystems seine Benutzbarkeit berücksichtigen und gegebenenfalls bewerten zu können, werden oft Kriterienkataloge verwendet.

Die Online-Bewerbung an deutschen Hochschulen muss für die Bewerber verständlich sein, sie muss verschiedene Fälle, auch viele Sonderfälle beachten. Bislang gibt es auf diesem Gebiet noch keine genaue Vorgehensweise dafür, wie mit Problemen der Benutzbarkeit genau umzugehen ist. Der Bewerbungsprozess wird dadurch unnötig verkompliziert. Eine mögliche Lösung dieses Problems ist die Erstellung eines Usability-Kriterienkatalogs für Online-Bewerbungssysteme an deutschen Hochschulen, womit sich diese Masterarbeit befasst.

2 Erstellung, Validierung und Anwendung eines Usability-Kriterienkatalogs

Die Erstellung des Usability-Kriterienkatalogs umfasste die folgenden vier Schritte: Erstellung einer ersten Version des Katalogs, Validierung, Optimierung und Anwendung.

Erstellung einer ersten Version

Die erste Version des Usability-Kriterienkatalogs wurde auf der Basis von Usability-Vorschriften: Norm DIN 9241, Usability-Richtlinien, Usability-Empfehlungen unter der Berücksichtigung der Spezifik der Domäne Online-Bewerbungssysteme an deutschen Hochschulen erstellt. Der erstellte Kriterienkatalog umfasste 116 Usability-Kriterien, davon sind 52 - Schritte des BITV-Tests und 64 - generiert aus den Usability-Richtlinien und den Anforderungen der Domäne Online-Bewerbungssysteme.

Validierung

In der Validierungsphase wurden die Usability-Kriterien auf ihre Relevanz zur Aufnahme in den Kriterienkatalog durch Tests (mit Testfällen und Sollwerten) sowie durch Usability-Tests (mit Probanden) überprüft. Die 64 zu validierenden Usability-Kriterien wurden hierzu aufgeteilt in: solche, die ohne Usability-Tests (35, technische Aspekte) und diejenigen, die nur oder zusätzlich mit den Usability-Tests geprüft werden müssen (29, psychologische Aspekte). Die Usability-Tests wurden mit 21 Testpersonen, größtenteils Schüler der 13. Klasse sowie einigen Testpersonen mit einem Hochschulabschluss zur Bewerbung für ein Zweitstudium validiert.



Abbildung 1: Usability-Tests mit den Abiturienten

Bei den Usability-Tests haben sich die Testpersonen an drei verschiedenen Online-Bewerbungssystemen beworben: 1) Universität Flensburg (Hersteller: Hochschul-Informations-System GmbH), 2) Universität Hamburg (Hersteller: Firma „Datenlotsen“), 3) Technische Fachhochschule Berlin (von der Hochschule selbst entwickelt). Außerdem wurden verschiedene Rollen der Bewerber berücksichtigt: 1) Standardfall (deutscher Bewerber, Erststudium), 2) besonderer Fall (z.B. ausländisch/höheres Semester/ Härtefall), 3) ‚worst-case‘-Fall (z.B.: ausländischer Bewerber, Zweitstudium). Die bei der Validierung verwendeten Systeme befanden sich im Einsatz an den Hochschulen und somit war eine Logfile-Analyse nicht durchführbar.

Optimierung und Anwendung

Nach der Validierung wurde der Kriterienkatalog in Bezug auf die Ergebnisse der Validierung optimiert und auf das Online-Bewerbungssystem der Hochschul-Software-System GmbH, entwickelt im Rahmen der neuen Softwaregeneration HISinOne, angewandt. Die Anwendung fand im Rahmen der Usability-Tests statt. Die Kriterien wurden wie bei der Validierung entsprechend aufgeteilt. Zusätzlich dazu fand teilweise eine Logfile-Analyse statt. Zurzeit werden die Ergebnisse der Anwendung ausgewertet. Das Ziel ist, eine Usability-Beurteilung des Systems zu erstellen und daraus die Empfehlungen für Systementwickler zu generieren.

3 Ergebnisse

Da im aktuellen Bearbeitungsstatus der Masterarbeit die Auswertungen der Usability-Tests zur Anwendung des optimierten Kriterienkatalogs noch durchgeführt werden, sind erst die Zwischenergebnisse zu berichten. Die Validierung der Kriterien des auf der Theoriebasis erstellten Kriterienkatalogs lieferte folgende Ergebnisse:

a) Es haben sich systemübergreifend viele im Kriterienkatalog angenommene Kriterien bestätigt. z. B.:

- Fehlermeldungen (auch system faults): müssen verständlich beschrieben sein.
- Hilfen für Formate (z.B.: Geburtsdatum, HZB-Note) müssen vorhanden sein.

b) Es haben sich systemübergreifend keine im Kriterienkatalog angenommenen Kriterien nicht bestätigt.

c) Es wurden neue Kriterien entdeckt, die im Katalog noch nicht vorhanden waren, z. B.:

- Es muss eine Übereinstimmung der Information auf der Homepage und in der Online-Bewerbungsanwendung geben.
- Es müssen Auswahlmöglichkeiten für Fächerkombinationen vorhanden sein.

In Bezug auf die Ergebnisse der Validierung wurde der Katalog optimiert und aufgeteilt: zum einen basierend auf der Relevanz der Kriterien: (führt zum) Fehler/ bedeutend/ unbedeutend, zum anderen basierend auf den Bereichen, in welchen die Kriterien erfüllt werden sollen: vor der Bewerbung/ im Inhalt der Bewerbung etc.

Literaturverzeichnis

- Dix, Alan; Finlay, Janet; Abowd, Gregory D.; Beale, Russell: Human-Computer-Interaction. 3. Auflage, Pearson Education Limited, 2004.
- Lauesen, Soren. User Interface Design. A Software Engineering Perspektive. Pearson Education Limited, 2005.
- Mayhew, Deborah, J. The Usability Engineering Cycle. Academic Press, 1999.
- Nielsen, Jakob. Usability Engineering. Academic Press, 1993.
- Beier, Markus; Vittoria von Guzycki. Usability. Nutzerfreundliches Web-Design. Springer, 2002.
- Brau, Henning; Diefenbach, Sarah; Hassenzahl, Marc; Koller, Franz; Peissner, Matthias; Röse, Kerstin. Usability Professionals 2008.
- German Chapter der Usability Professionals Assoc., Stuttgart, 2008.
- Gross, Tom. Mensch& Computer 2008. Oldenbourg Wissenschaftsverlag, 2008.
- Gross, Tom. Mensch& Computer 2007. Oldenbourg Wissenschaftsverlag, 2007.
- Heinsen, Sven; Vogt, Petra. Usability praktisch umsetzen. Hanser, 2003.
- Lucke, Ulrike; Kindsmüller, Martin Christof; Fischer, Stefan; Herczeg, Michael; Seehusen, Silke. Workshop Proceedings der Tagungen Mensch & Computer 2008, DeLFI 2008 und Cognitive Design 2008. Logos, Berlin, 2008.
- Seehusen, Silke; Lucke, Ulrike; Fischer, Stefan. DeLFI 2008: Die 6. e-Learning Fachtagung Informatik der Gesellschaft für Informatik e.V. 07.-10. September 2008, Lübeck. Gesellschaft für Informatik, Bonn, 2008.
- Nielsen, Jakob; Loranger, Hoa. Web Usability. (broschiert) 1. Auflage. Addison-Wesley, München, 2008.
- Nielsen, Jakob. Designing Web Usability. Markt+Technik, 2001.
- Nielsen, Jakob; Loranger, Hoa. Prioritizing Web Usability. Nielsen Norman Group, 2006.
- Krug, Steve. Don't make me think! Web Usability – das intuitive Web. 2. Auflage, Pearson Education Inc, 2006.
- Tidwell, Jenifer. Designing Web Usability. O'Reilly Media, 2006.
- McGovern, Gerry; Norton, Rob; O'Dowd, Catherine. The Web Content Style Guide. Pearson Education Limited, 2002.
- Wirth, Thomas. Missing Links. Über gutes Webdesign. 2., erweiterte Auflage. Hanser, 2004.
- Manhartsberger, Martina; Musil, Sabine. Web Usability. Galileo Design. Das Prinzip des Vertrauens. Galileo Press, 2002.
- Nielsen, Jakob; Tahir, Marie; Homepage Usability. 50 enttarnte Websites. Markt+Technik, 2002.
- Jacobsen, Jens. Website-Konzeption: Erfolgreiche Websites planen und umsetzen (broschiert). 4. erw. u. aktualis. Auflage. Addison-Wesley, München, 2007.
- Neumann, Bettina. Usability im World Wide Web: Kriterien, Techniken und Lösungen (broschiert). 1. Auflage. Vdm Verlag Dr. Müller, 2007.
- Dahm, Markus: Grundlagen der Mensch-Computer-Interaktion. Pearson Studium, München, 2006.
- Weiermann, Thomas. Evaluation der Usability von Websites mittels Kriterienkatalogen: Vom Keevil Usability Index zum Web Usability Index: Gemeinsamkeiten, Unterschiede, Optimierungspotenziale. 1. Auflage (broschiert). GRIN, 2007.
- Tullis, Tom; Albert, Bill. Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics (Interactive Technologies): Collecting, Analyzing, and ... Kaufmann Series in Interactive Technologies). (Taschenbuch). Morgan Kaufmann, 2008.
- Bartel, Torsten. Die Verbesserung der Usability von WebSites. Wiku-Verlag, 2004.
- Chisnell, Dana, und Jeffrey Rubin. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. Wiley & Sons, 2008.
- Sarodnick, Florian, und Henning Brau. Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung. Huber, Bern, 2006.
- Schweibenz, Werner, und Frank Thissen. Qualität im Web: Benutzerfreundliche Webseiten durch Usability-Evaluation. Springer, Berlin, 2002.

AuthorSupportTool

Projekt BesserWissA, Universität Bremen
Steffen Bothe, Ali Ayhan Dardagan, Marcus Ermler, Patrick Fehrmann,
Christoph Jahn, Marshall Nana, Vitali Richter, Benjamin Vollmer
Betreuung: Prof. Dr. Karl-Heinz Rödiger

grp-astproject@informatik.uni-bremen.de

Abstract: Wie können moderne Textverarbeitungssysteme den Schreibprozess einer wissenschaftlichen Arbeit unterstützen? Mit diesem Thema beschäftigt sich das studentische Projekt BesserWissA („Besser Wissenschaftlich Arbeiten“) im Studiengang Informatik an der Universität Bremen. Das AuthorSupportTool ist eine Erweiterung, die die Textverarbeitungs-Software OpenOffice.org 3 Writer um sinnvolle Funktionen für wissenschaftliches Schreiben ergänzt.

1 Einleitung

Beim Verfassen wissenschaftlicher Arbeiten verbringen Autoren¹ neben der eigentlichen Denk- und Formulierungsarbeit üblicherweise viel Zeit mit formalen (z.B. Einhaltung einer einheitlichen Formatierung) und arbeitsorganisatorischen Angelegenheiten (z.B. Koordinierung des kooperativen Arbeitens, Verwaltung von Quellen). Diese Aufgaben wollen wir, das studentische Projekt BesserWissA an der Universität Bremen, Autoren abnehmen oder zumindest vereinfachen. Da es wenig sinnvoll ist, die Grundfunktionalität eines n-ten Textverarbeitungssystems nach zu programmieren, haben wir zunächst vorhandene analysiert. Wir haben uns für OpenOffice.org 3 Writer als Basis für eine Erweiterung entschieden, da dieses quelloffen und plattformunabhängig ist. Der Erweiterung haben wir den Namen AuthorSupportTool gegeben.

2 AuthorSupportTool

Das AuthorSupportTool bietet Eigenschaften, die für die Erstellung von wissenschaftlichen Texten hilfreich sind. So kann sich der Autor auf die eigentliche Arbeit konzentrieren, ohne sich mit technischen Einzelheiten auseinandersetzen zu müssen.

¹Obwohl aus Gründen der Lesbarkeit im Text nur die männliche Form gewählt wurde, beziehen sich die Ausführungen auf Angehörige beider Geschlechter.

2.1 Grundeinstellungen und kooperatives Arbeiten

Zu Beginn hilft ein Assistent dem Verfasser, Einstellungen vorzunehmen, die das Layout des späteren Dokuments bestimmen. Hierzu wählt der Autor die Art seiner Arbeit (Diplomarbeit, Dissertation, Aufsatz, Hausarbeit etc.) aus. Die verschiedenen sogenannten Templates definieren einen einheitlichen Stil für Formatierungen, Typographie und Hervorhebungen von Textelementen. Anschließend entscheidet sich der Benutzer für ein passendes Seitenlayout wie ein- oder beidseitiges Drucken, die Größe des Korrekturrandes und weitere Details. Das Thema der Arbeit und die Daten der Autoren werden ebenfalls im Assistenten eingetragen.

Verschiedene Autoren können unabhängig von Ort und Zeit an einem Dokument arbeiten. Dokumente werden dabei versioniert und mit einem System zum Software Configuration Management (hier Subversion, SVN) verwaltet. Damit werden u.a. Versionskonflikte aufgezeigt und nach Möglichkeit automatisch behoben. Verschiedene Schreib- und Lese-rechte werden mit einem Rollenkonzept verwaltet. Je nach Bedarf gibt es neben Autoren auch Gutachter, Lektoren oder Leser, welche ebenfalls einen geregelten Zugriff auf das Dokument haben. Zugriffsrechte werden von den jeweiligen Autoren vergeben.

2.2 Zitat- und Quellenverwaltung

Ein wichtiges Element wissenschaftlichen Schreibens ist das korrekte Zitieren. Hierfür bietet AuthorSupportTool zwei Werkzeuge: die Zitatsammlung und die Zitatgenerierung. In der Zitatsammlung werden Zitate zusammen mit den notwendigen bibliographischen Angaben verwaltet. Die Sammlung wird als Liste angezeigt; sie kann nach Stichwörtern durchsucht werden. Aus dieser Aufstellung lässt sich jedes Zitat zur Bearbeitung öffnen oder zur Übernahme als Zitat im Text auswählen. Dabei stehen verschiedene Bausteine als Anregung für die Einbettung des Zitats in den Text zur Auswahl.

Ähnliches bietet AuthorSupportTool zur Quellenverwaltung und zur Generierung der Literaturliste: Aus den zitierten Büchern, Artikeln oder Online-Quellen wird eine wissenschaftlich korrekte Literaturliste erstellt, die sich alphabetisch oder nach Auftreten im Text sortieren lässt. Als Alternative zu reinen URL-Angaben können Online-Quellen auch komplett abgespeichert werden. Alle Quellen sind thematisch strukturierbar und werden so verwaltet, dass sie bei Bedarf projektübergreifend (wieder-) verwendet werden können. Während der Arbeit am Dokument werden alle relevanten Elemente (Titelblatt, Seitenzahlen oder verlinkte Verzeichnisse wie Glossar, Inhalts-, Abbildungs- oder Tabellenverzeichnis) aktuell gehalten und laufend ergänzt. Über eine geschachtelte Elementübersicht wird deren Anordnung visualisiert. Mit dieser Übersicht kann man direkt zu einem bestimmten Element im Textkörper springen. Durch Umstellen der einzelnen Elemente innerhalb der Baumstruktur wird das Dokument entsprechend restrukturiert.

2.3 Hilfen zur Strukturierung und Arbeitsorganisation

Um beim Schreiben den sogenannten roten Faden nicht zu verlieren, können Autoren jedem Kapitel Schwerpunkte zuordnen. AuthorSupportTool prüft dann, ob die Schwerpunkte in den entsprechenden Kapiteln vorkommen. Wird ein Schwerpunkt nicht gefunden, erhält der Autor einen Hinweis. Schwerpunkte und deren Beziehungen zueinander werden automatisch visualisiert.

Der Arbeitsfortschritt wird mittels Zeitlinie, Terminplan, Meilensteinen und Fortschrittsanzeigen in Text- und Tabellenform dokumentiert. So lässt sich zu jedem Zeitpunkt erkennen, wie weit die Arbeit fortgeschritten ist. Wort- und Zeichenanzahl der verschiedenen Kapitel werden ebenfalls numerisch und grafisch dargestellt.

Damit die Ergebnisse der Arbeit nicht verloren gehen, bietet AuthorSupportTool Backup-Funktionen, um Datensicherungen in mehreren Revisionen vorzunehmen. Sicherungen erfolgen manuell oder automatisch. Revisionen werden mit einem Datum-Uhrzeit-Stempel versehen und können jederzeit aufgerufen werden.

Nach Abschluss der Arbeit kann mittels der Schwerpunkte und mit den zu Beginn eingegebenen Metadaten die Rohfassung eines Abstracts generiert werden.

3 Fazit

Mit Hilfe von AuthorSupportTool werden arbeitsorganisatorische Aufgaben wie die Zusammenarbeit mehrerer Autoren an einem Dokument vereinfacht. Im Gegensatz zu gängigen Versionsverwaltungssystemen wird kein zusätzliches Programm benötigt, um diese Anforderung umzusetzen. Ebenso ist die Zitat- und Quellenverwaltung integraler Bestandteil des AuthorSupportTool. Damit benötigt der Autor keine externen Hilfsmittel wie Literaturdatenbanken, bibliographische Nachweise oder Werkzeuge, um große Mengen an Material und Quellen zu verwalten. Die Rohfassung des Abstracts verlangt allerdings noch intellektuellen Aufwand seitens des Autors; sie muss in jedem Fall nachbearbeitet werden. AuthorSupportTool ist nach unserer Meinung eine gute Unterstützung beim Schreiben wissenschaftlicher Werke. Die kreativen Anteile dieser Tätigkeit kann es jedoch keinesfalls substituieren.

4 Literatur

- [Eco 1989] Eco, U.
Wie man eine wissenschaftliche Abschlußarbeit schreibt.
Müller Juristischer Verlag, 2. Aufl., Heidelberg 1989
- [Friedrich/Rödiger 1991] Friedrich, J. und K.-H. Rödiger (Hrsg.)
Computergestützte Gruppenarbeit (CSCW).
Teubner Vlg., Stuttgart 1991
- [OpenOffice Extensions 2008] OpenOffice.org Extensions. OpenOffice.org Repository
Plattform für Extensions. 2008.
<http://extensions.services.openoffice.org>
(Aufruf: 22.01.08)
- [Rechenberg 2002] Rechenberg, P.
Technisches Schreiben (nicht nur) für Informatiker.
Hanser Vlg., München 2002
- [Textor 1985] Textor, A.M.
Sag es treffender. Rowohlt Vlg., Reinbek 1985

Headtracking für zwei Personen mit der Wiimote

Bastian Rackow, Simeon Eichel

Hochschule Reutlingen
Informatik

Bastian.Rackow@student.Hochschule-Reutlingen.de

Simeon.Eichel@student.Hochschule-Reutlingen.de

Art der Arbeit: Projekt im Kurs „Visualisierung“ im 2. Mastersemester

Betreuer/in der Arbeit: Prof. Dr. Gabriela Tullius, Prof. Dr. Uwe Kloos

Abstract: Durch Headtracking (dt. Nachverfolgung des Kopfes) kann die Position und die Bewegung von zwei Spielern bestimmt und damit eine perspektivisch richtige Sicht auf die Spielwelt berechnet werden. Die zwei Spielansichten der beiden Spieler werden übereinander projiziert und sollen durch Polarisation getrennt werden, so dass jeder Spieler seine eigene Spielumgebung sieht.

1 Idee

Die Idee dieses Projekts ist eine spielerische Anwendung, die es zwei Spielern ermöglicht zur selben Zeit, am selben Display, nebeneinander stehend, in ihrer eigenen perspektivischen Sicht in der selben Welt interagieren zu können.

Die Grundlage dafür wurde von Lee [JLE08] geschaffen, welcher ein sehr kostengünstiges aber effektives Headtracking mit dem Controller („Wiimote“) der Spielkonsole „Wii“ von Nintendo ermöglicht. Die Kopfbewegungen werden von einem solchen Controller erfasst und in eine perspektivisch korrekte Sicht des Benutzers auf die dargestellte Umgebung umgerechnet.

In Lees Arbeit ist Headtracking nur für eine Person an einem Fernseher (oder Computer) umgesetzt, da es dem Ausgabegerät nicht möglich ist zwei überlagerte Bilder unabhängig von einander darzustellen. Wir erweitern diese Arbeit um die zweite Sicht.

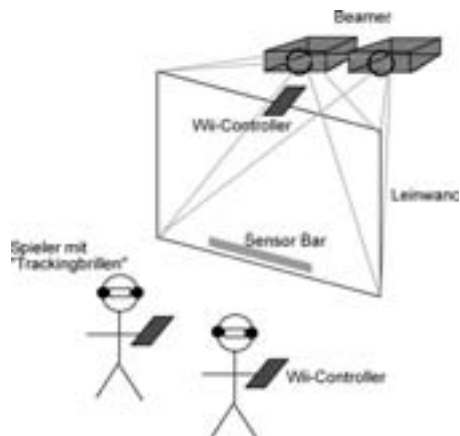


Abbildung 1: Aufbau des Systems

2 Konzept

Um die perspektivische Sicht für zwei Personen zu ermöglichen, müssen auf der Eingabeseite die Kopfpositionen und damit die jeweilige korrekte Sicht auf die Welt beider Spieler getrennt berechnet werden. Die Ausgabe wird dann ebenfalls für beide Personen getrennt, damit jeder nur das für ihn passende Bild sieht (vgl. Abbildung 1). In einer Spielidee soll es den Spielern ermöglicht werden, Zielscheiben zu treffen. Hierfür kommen weitere Wiimotes als Interaktionsgeräte zum Einsatz.

2.1 Perspektivische Sicht für zwei Personen

Vom Tracking der Kopfposition bis zur Ausgabe des jeweiligen Bildes sind mehrere Schritte nötig (vgl. Abbildung 1). Die Kopfposition jedes Spielers wird zunächst von je zwei Infrarot-LEDs an einer Brille getrackt. Das Tracking wird von der Wiimote übernommen, die die LEDs mit der eingebauten Kamera erkennt. Die Positionsdaten der LED-Punkte im Kamerabild werden von der WiimoteLib von Peek [PEEK08] empfangen. Die insgesamt vier LED-Punkte müssen dann dem passenden Spieler zugeordnet werden. Die Zuordnung geschieht hauptsächlich über die x-Position im

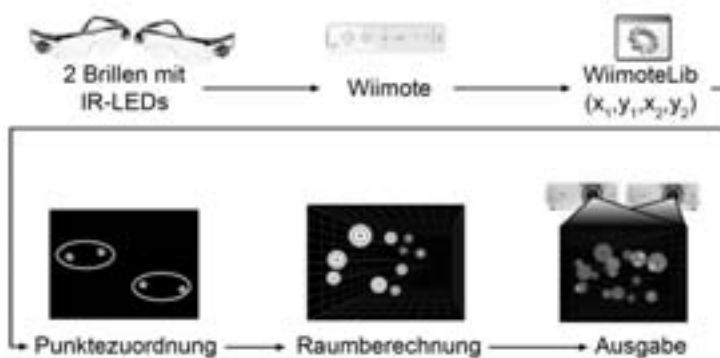


Abbildung 2: Grafische Darstellung des Konzepts

Kamerabild, aber auch über die ID der Punkte, die von der Wiimote vergeben wird. Mit Hilfe von verschiedenen Parameter (Größe des Anzeigefläche, Abstand der LEDs an der Brille, Punktposition im Kamerabild) kann dann die Position des Spielers im Raum, und damit die Position der virtuellen Kamera berechnet werden. Jeder Spieler steuert also die Position seiner virtuellen Kamera mit dem Kopf. Die Ausgabe erfolgt auf zwei Beamern, von denen jeweils einer das Bild eines Spielers darstellt und auf die gemeinsame Leinwand projiziert. Dort überlagern sich die Bilder der beiden Spieler.

2.2 Interaktion mit weiteren Controllern

Zum „Abschießen“ der Zielscheiben hat jeder Spieler eine weitere Wiimote in der Hand. Diese werden wie vom Hersteller ursprünglich vorgesehen verwendet. Die Kameras dieser Controller tracken die Infrarot-LEDs einer Sensorbar von Nintendo, die unter der Leinwand angebracht ist. So kann pro Controller ein Fadenkreuz zum Zielen gesteuert werden.

3 Hardware

Um die angedachte Idee realisieren zu können, bedarf es einiger Hardwarevoraussetzungen. Benötigt werden drei Controller der Spielekonsole „Wii“ von Nintendo (zwei zur Interaktion, einen zum Tracking). Zwei Brillen mit jeweils zwei Infrarot-LEDs und Polarisationsfiltern, sowie zwei Beamer, die ebenfalls mit Pol-Filtern ausgestattet sind.

3.1 Brillen

Mit zwei im Handel erhältlichen Schutzbrillen mit ausrichtbaren, weißen LED-Leuchte sollen die zwei Spieler getrackt werden. Dafür müssen die weißen LEDs durch sehr helle Infrarot-LEDs ausgetauscht werden, damit das Tracking mit Hilfe der Wiimote überhaupt möglich ist. Die LEDs sollten so leuchtstark wie möglich sein und trotzdem einen breiten Abstrahlwinkel besitzen, damit auch leicht seitliche Kopfbewegungen möglich sind.

Damit der Spieler nur seine Ansicht sieht und nicht die des Mitspielers, werden vor die Gläser Polfilter geklebt. Die gleichen Filterausrichtungen sind auch vor den Beamern angebracht, welche das Bild für den Spieler projiziert. Dadurch sieht der Spieler nur das für ihn bestimmte Bild, und nicht auch das des anderen Spielers.

3.2 Beamer

Die Beamer sind wie in 3.1 erwähnt ebenfalls mit Polfiltern bestückt, um jedem Spieler nur das für ihn bestimmte Bild anzuzeigen. Das Bild wird in der so genannten „Rückprojektionstechnik“ (vgl. Abbildung 1) von hinten auf eine Leinwand projiziert. Dies ist sinnvoll, damit sich die Spieler nicht im Strahlengang der Beamer befinden.

3.3 Controller

Wie im Abschnitt „Konzept“ erläutert, werden drei Wii-Controller („Wiimotes“) benötigt. Jeder Spieler hält einen Controller in der Hand um das das Fadenkreuz zu steuern. Der dritte Controller, welche die Kopfbewegungen erfasst, wird oberhalb der Leinwand angebracht und auf die Spieler ausgerichtet. Für die Datenübertragung werden die Controller mit dem PC via Bluetooth verbunden.

4 Ausblick

Durch das Projekt wurde eine Grundlage zur Interaktion von zwei Personen vor einer Leinwand in ihrer eigenen perspektivischen Sicht und eine Umsetzung als Spiel geschaffen. Das Projekt wird gegenwärtig z.B. durch folgende Erweiterungen fortgeführt:

- Erweiterung der Spielidee (z.B.: Spieler muss mit Kopfbewegungen verschiedenen Objekten ausweichen)
- Headtracking mit mehreren Wiimotes für mehr Freiheitsgrade
- Verbesserung der Technik (hellere LEDs, bessere Stromversorgung)
- Darstellung beider Bilder auf einem Monitor und Trennung der Bilder mit dem Anaglyphenverfahren

Literaturverzeichnis

- [JLE08] Johnny Chung Lee: Head Tracking for Desktop VR Displays using the Wii Remote, <http://www.cs.cmu.edu/~johnny/projects/wii/>, Stand: 10.01.2009
- [PEEK08] Brian Peek: Managed Library for Nintendo's Wiimote, <http://www.codeplex.com/WiimoteLib>, Stand: 10.01.2009

Multitouch Anwendungen in virtueller Realität

B.Sc. Simone Fries
sfries@stud.fbi.h-da.de

Prof. Dr. Katja Lenz
k.lenz@fbi.h-da.de

Fachbereich Informatik, Hochschule Darmstadt

Abstract:

Multitouch gilt als neue Art der Mensch-Maschine-Interaktion und bedeutet das Agieren mit mehr als einem Finger auf berührungsempfindlichen Oberflächen. Das vergleichsweise alte Konzept wird derzeit in den Medien immer öfter in Verbindung mit neuen Technologien genannt. Multitouch Systeme gelten als einfacher und natürlicher bezüglich der Möglichkeit simultan zu arbeiten. Dennoch bestehen Grenzen für die Einsatzmöglichkeiten und die Evaluierung geeigneter Interaktionsprinzipien ist notwendig. In der vorliegenden Arbeit wird das Spiel Pong auf Basis verschiedener Interaktionsstudien um Multitouchfunktionalitäten erweitert. Die Realisierung erfolgt an einem Multitouch Table in virtueller Realität als Multiuser Szenario. Anhand dieses Anwendungsbeispiels werden die Aufbereitung und Verarbeitung der nötigen Daten vorgestellt.

1 Einleitung

Die Interaktionserweiterung von handelsüblichen Singletouch-Displays, wie sie beispielsweise in öffentlichen Informationsterminals für die Fahrplanauskunft, zum Einsatz kommen, hält viele neue und bisher wenig erforschte Möglichkeiten bereit. Neben komplexen Interaktionen können mehrere Personen simultan am gleichen Gerät interagieren. Durch Produkte wie Microsofts Surface [Mic08] oder Apples iPhone [App08] wird mit Multimedia-Applikationen die Technologie mehr und mehr marktfähig gemacht. In solchen, oft bunt gestalteten Anwendungen wird der Umgang mit dem Computer spielerisch und intuitiv ermöglicht. Jedoch bietet Multitouch mehr als beispielsweise die Anordnung von einfachen Media-Daten, wie Bildern oder Videos. Komplexe Interaktionen und damit die simultane Manipulation mehrerer Freiheitsgrade¹ kann auf einfache Weise erreicht werden und ist besonders in 3D-Umgebungen von großer Relevanz. Eine real wirkende Interaktion mit den darin befindlichen Objekten wird erreicht, indem mehrere Parameter gleichzeitig verändert werden können. Herkömmliche Eingabemedien, wie die Maus, ermöglichen nur ein simultanes Steuern von zwei Parametern, was einen sequenziellen Ablauf zur Folge hat.

¹Der Freiheitsgrad gibt an, wie viele Parameter eines Objektes sich gleichzeitig manipulieren lassen.

2 Multitouch Technologie

Die Berührungsempfindlichkeit von Oberflächen kann auf verschiedene Weise erreicht werden. Neben kapazitiven Systemen, wie beispielsweise Rekimotos Smartskin [Rek02], hat sich, führend durch die Arbeit von Han [Han05], die optische Sensortechnik der "Frustrated total internal reflection (FTIR)" durchgesetzt. Han's Konzept nutzt dabei das optische Phänomen der Totalreflexion: Licht wird innerhalb eines Mediums solange gehalten, bis dessen Oberfläche verändert (frustrated) wird (siehe Abbildung 1). Für den Einsatz in einem Multitouch System wird das für Menschen nicht sichtbare infrarote Licht verwendet. Die Fingerberührungen treten durch die Unterbrechung der Totalreflexion als helle Punkte im Kamerabild auf und werden anschließend mit Bildverarbeitungsverfahren in Koordinaten umgerechnet. Solche Installationen können hohe Auflösungen erreichen und sind dabei preiswert zu realisieren. Die Technik kann problemlos mit Rückprojektionsoberflächen kombiniert und skalierbar eingesetzt werden.

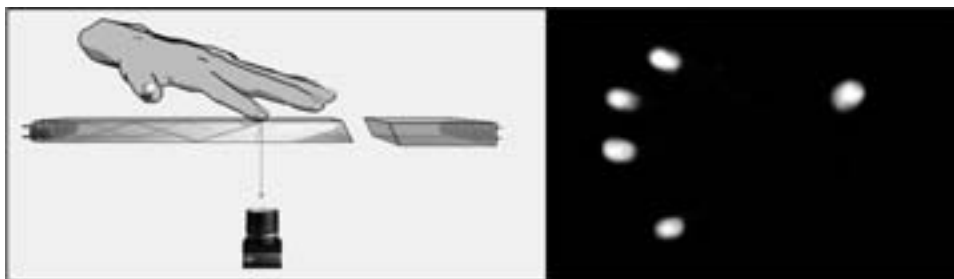


Abbildung 1: FTIR als optische Touchtechnologie

3 Interaktionskonzept

Die neuen, für Nutzer zunächst ungewohnten und über einfache Mausemulation hinausgehenden Interaktionsformen sollen über ein bekanntes Spielprinzip zugänglich gemacht werden. Der Spielverlauf des vereinfachten Tennis-Spiel "Pong", kann ohne aufwendiges Lernen von Regeln verstanden werden, so dass die Aufmerksamkeit auf die Interaktion gerichtet wird. Allerdings sollten verschiedene Aspekte für die Entwicklung geeigneter Interaktionsprinzipien beachtet werden. Dafür wurden in [Fri08] Interaktionsstudien wie unter anderen [FWSB07], [BM86], [KBS94], [Hug08], [Mos07], [SS91], [AZ03] evaluiert und unter Beachtung dieser Ergebnisse und der Größe des Multitouch Tables eine beidhändige Interaktion für die Steuerung des Schlägers gewählt. Die Skalierung des Schlägers wird durch den Fingerabstand zueinander bestimmt. Durch Verschieben zweier Finger wird die Translation, durch Drehen der Finger umeinander die Rotation erreicht (siehe Abbildung 2). Die Interaktionen können simultan ausgeführt werden, so dass mit zwei Fingern beispielsweise drei Freiheitsgrade gleichzeitig, durch Translation und Rotation, beeinflusst werden können. Diese Interaktion erscheint natürlich, da mit realen Objekten in gleicher

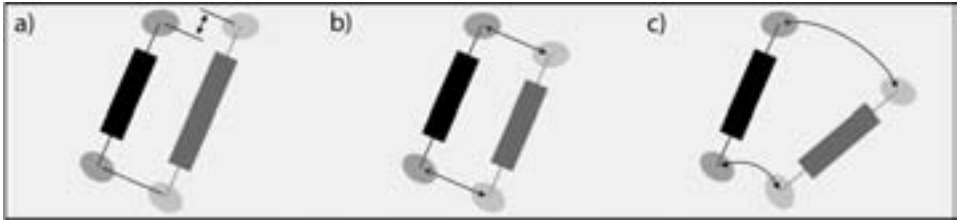


Abbildung 2: Interaktionen mit dem Schläger: a) Skalierung; b) Translation; c) Rotation

Weise interagiert werden kann. Des Weiteren kann der Ball durch eine Geste mit vier Fingern vor dem eigenen Schläger gehalten und beschleunigt werden. Dies erfordert eine schnelle Reaktionsfähigkeit, welche von den Spielern aber einfach durch Hinzunahme von Fingern erreicht wird. Dies zeigt das Potenzial weitere Parameter einfach beeinflussen zu können. Die große Tischoberfläche bietet eine gute Überschaubarkeit und es kann ein real wirkendes Multiplayer-Spiel realisiert werden. Im Gegensatz zur Maus ist die Touch-Eingabe nicht so präzise, worauf die Größe der Schläger angepasst wird. Das inhärente Problem der Verdeckung des dargestellten Bildes an den Berührstellen wird durch Hilfslinien behoben. Diese visualisieren den größeren Interaktionsradius, wodurch der Schläger nicht genau gefasst werden muss. Zusätzliches visuelles Feedback bietet die Simulation der Berührungspunkte auf der Oberfläche. Die zweidimensionale Eingabe wird durch Trackingverfahren des Bildverarbeitungsprogramm InstantVision² als X- und Y-Koordinaten bestimmt. Die Visualisierung der Objekte wird in der Modellierungssprache X3D beschrieben. Zunächst liegen die erfassten Koordinaten im Kamerakoordinatensystem vor. Um diese in der virtuell erstellten Szene nutzen zu können, ist die Umrechnung in Bildschirmkoordinaten notwendig. Dafür wird im Kalibrierungsprozess die affine Transformationsmatrix bestimmt und auf die Berührungen angewendet.

4 Zusammenfassung und Ausblick

Im Rahmen der Bachelorarbeit [Fri08] wurden Multitouchinteraktionen für Pong entwickelt und implementiert. Es stellte sich heraus, dass sich das Spielprinzip für Multitouch- und Multiuserinteraktionen eignet, die über die einfache Mausemulation hinausgehen. Die gleichzeitige Kontrolle und Manipulation mehrerer Funktionen und Freiheitsgrade wird ermöglicht und das Potenzial dieser Technik als Eingabegerät für 3D-Umgebungen veranschaulicht. Wie in der Anwendung gezeigt, können simultan mehrere Personen am Multitouch Table agieren. Der Einsatz als Multiuser System setzt allerdings eine Benutzeridentifikation, also die Zuordnung von Benutzer und Berührung voraus. Erste Ansätze sind in den multitouchfähigen Multiuser Systemen DiomandTouch [DL01] und dem von Dohse et al. [DDSP08] entwickelten System vorgestellt. Allerdings besteht noch Forschungsbedarf für die Entwicklung geeigneter Methoden.

²Vision-Modul des am Fraunhofer IGD entwickelten InstantReality Framework [Fra08]

Literatur

- [App08] Apple Inc. iPhone, 2008. <http://www.apple.com/de/iphone/>, Stand 10.01.2009.
- [AZ03] Pär-Anders Albinsson und Shumin Zhai. High precision touch screen interaction. *SIGCHI conference on Human factors in computing systems*, Seiten 105 – 112, 2003.
- [BM86] William Buxton und Brad A. Myers. A study in two-handed input. *SIGCHI Conference*, 17:321 – 326, 1986.
- [DDSP08] K. C. Dohse, Thomas Dohse, Jeremiah D. Still und Derrick J. Parkhurst. Enhancing Multi-user Interaction with Multi-touch Tabletop Displays Using Hand Tracking. *International Conference on Advances in Computer-Human Interaction*, 1:297–302, 2008.
- [DL01] Paul Dietz und Darren Leigh. DiamondTouch: A multi-user touch technology. *ACM symposium on User interface software and technology*, 14:219 – 226, 2001.
- [Fra08] Fraunhofer IGD. InstantReality, 2008. <http://www.instantreality.org/home/>, Stand 01.10.2008.
- [Fri08] Simone Fries. Multitouch Anwendungen in virtueller Realität: Evaluation und Verarbeitung fingerbasierter Multiuser Interaktionsprinzipien, 2008.
- [FWSB07] Clifton Forlines, Daniel Wigdor, Chia Shen und Ravin Balakrishnan. Direct-Touch vs. Mouse Input for Tabletop Displays. *SIGCHI Conference on Human factors in computing systems*, Seiten 647 – 656, 2007.
- [Han05] Jefferson Y. Han. Multi-Touch Sensing through Frustrated Total Internal Reflection. *International Conference on Computer Graphics and Interactive Techniques*, 2005.
- [Hug08] Tomer Moscovich John F. Hughes. Indirect Mappings of Multi-touch Input Using One and Two Hands. *SIGCHI Conference on Human factors in computing systems*, 26:1275–1284, 2008.
- [KBS94] Paul Kabbash, William Buxton und Abigail Sellen. Two-handed input in a compound task. *SIGCHI Conference on Human factors in computing systems: celebrating interdependence*, Seiten 417 – 423, 1994.
- [Mic08] Microsoft. Surface, 2008. <http://www.microsoft.com/surface/index.html>, Stand 10.01.2009.
- [Mos07] Tomer Moscovich. *Principles and Applications of Multi-touch Interaction*. Dissertation, Brown University, 2007.
- [Rek02] Jun Rekimoto. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. *SIGCHI Conference on Human factors in computing systems: Changing our world, changing ourselves*, Seiten 113 – 120, 2002.
- [SS91] Andrew Sears und Ben Shneiderman. High precision touchscreens: design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies*, 34:593 – 613, 1991.

Implicit Social Network Construction and Expert User Determination for Context-Aware Web Portal Environments

Mareike Lattermann, Andreas Nauerz, Walter Kriha
mareikel@de.ibm.com, andreas.nauerz@de.ibm.com, kriha@hdm-stuttgart.de

Abstract: Information overload and therefore the complex and time-consuming act of finding the designated content has become the biggest problem of today's web portals. This paper [Lat08] deals with ways to overcome this problem by exploiting social networks and the expertise of users to provide support for others. The goal is to speed up the process of getting to the sought-after information by recommending people based on current and long-term behavior as well as knowledge, that are able to point out relevant content and give support in the current context.

1 Introduction

Web Portals were introduced to provide a single point of access to great amounts of content, but a new challenge for portals has evolved: managing the increasing information overflow efficiently and user-oriented. This paper's goal is to leverage social networking and communication tools to provide people looking for information with additional support apart from search engines and navigation. The idea is to go beyond explicit social networking and construct implicit social networks based on similar navigational behavior of users. The underlying assumption is that users consuming the same content or performing similar tasks on the system, might benefit from each other. Furthermore an expertise model is constructed to determine the knowledge domains of users, since these experts can be recommended to users that are less familiar with the system. Finally information about others currently being in the same context is provided to give users the chance to ask for help. Additionally content recommendations using the recommender engine LikeMinds (c.p. chapter 3) will be evaluated.

2 Conceptual Design

Figure 1 visualizes the idea of this paper, which was initially proposed in [NG08]. Each of the components of the figure are described in detail in the following sections.

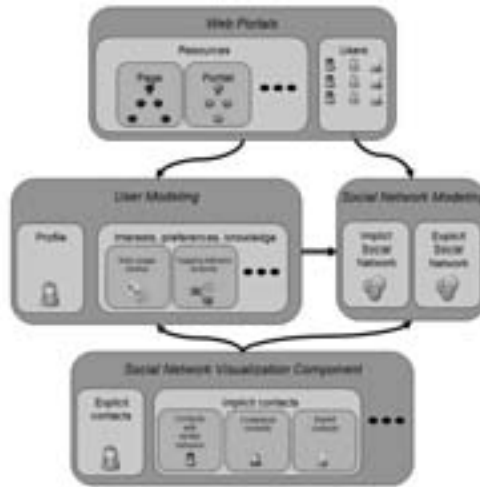


Figure 1: Conceptual Overview of the Idea of Implicit Social Networks

2.1 User Modeling

User models aim to store all information useful for improving communications between users and a system. There are several aspects that can influence a user model, and different information sources to extract data for it e.g. web data mining [Liu07]. Since the focus of this work is to try to enhance the user experience in Portals *implicitly*, data sources that do not require the user to contribute explicit interests of any kind are used, the navigational behavior within the community in particular. Basically the clickstream of each user is stored in a keyword vector [BKN07] where the keywords are represented by the visited resources and the weight by the number of visits. Subsequently information about visited resources, duration and frequency of visits can be extracted, which leads to user-profiles showing the interest and usage intensity and allow to interpret these as e.g. a measure of expertise or to detect similarities concerning users' interests and knowledge.

2.1.1 Expertise Determination

Expertise location is the process of extracting users knowledge and determine in what areas they can be called experts and thereby used to provide support for others. This paper extracts the expertise by interpreting the information stored in the user model. The foundational assumption for this is that expertise is expressed by frequent views and usage over a longer period of time. For example: If a User visits the page "Soccer" and interacts with the "Soccer-Results"-portlet frequently, it can be assumed that he is interested in the topic itself and secondly knows something about soccer and the usage of the page.

2.1.2 Similarity Detection

To be able to find similarities between users based on the introduced keyword model, simple vector-distance calculation can be applied. One way of determining that is to use the so-called "Extended Jaccard Coefficient" or "Cosine Similarity" [GGK⁺05]. The extended Jaccard coefficient or Tanimoto coefficient of two vectors a and b is defined as the angle α between a and b . The normalized version of the equation, the Cosine Similarity, emits a number between 0 and 1: the higher the more similar. Users that have been found similar can now be recommended to each other for collaboration purposes or social interaction depending on the community-focus.

2.2 Social Network Modeling

The idea is to base the connections between users on similar behavior and thereby on similar interests in relation to the content rather than personal vicinity. The network will therefore not necessarily or exclusively be created by people already knowing each other and adding each other to some sort of contact-list but by people-recommendations depending on the currently viewed content or the overall navigational behavior.

3 System and Solution

The prototype that has been developed for this work features a database for storing the user models and an application layer that consists of two components dealing with the people recommendations on the one hand and the content recommendations with LikeMinds on the other. To visualize all the implemented ideas, two portlets, one contacts portlet that has the look and feel of an instant messaging contacts-list and one sample DVD-shop application, have been developed that allow implicit recommending of people and content, based on the users' navigational behavior. The recommendations can either be calculated dependent on the usage within the whole web portal, where resources are represented by pages and portlets, or on the shop, where resources would be purchasable items. Four types of people recommendations are presented to the user:

- Similar behaving users (Recommend movies he is very likely to be interested in)
- Expert users (Provide further information about the movie)
- Contextual contacts (Contact for a discussion)
- Dissimilar behaving users (Help decide what a user might not like)

Additionally, the shop portlet features content recommendations created via the LikeMinds recommendation engine, which allows three ways of generating recommendations [JG00]. The subengines used for this paper are the Clickstream-engine, which stores the navigation behavior of a user and recommends content based on similarly behaving users, and

the Item-Affinity-engine, which gathers information about associations between items or resources and recommends the related ones. Summarized, this solution's goal is to provide further information and support in finding what a user is really looking for in combination with building up a social network that is helpful in every possible situation.

4 Conclusion and Future Work

The developed application represents one way of scaling down the problem of information overflow in web-portals. The social networking approach combined with people recommendation boosts the idea of community and collaboration within portals and is able to speed up the search for content. Users can rely on other peoples experiences and thereby not only get to the designated information quicker but also fall back on trustworthy users. This work demonstrates how much can be gained from the people-focused approach of evaluating information. The goal is not to find a one-and-only answer to the information-flood problem though, but to explore different approaches, try to extract the best features and combine these qualities to a framework of several solutions that are able to support the user in various ways. Room for improvement can be found in the limitation of information sources to just the clickstream, which will be addressed in the future. More and especially different types of data sources need to be explored and added to make the user model more accurate, e.g. the analysis of a users contributions could be a next step, which also allows the systems calculations to remain undiscovered.

IBM and WebSphere are trademarks of International Business Machines Corporation in the United States, other countries or both. Other company, product and service names may be trademarks or service marks of others.

References

- [BKN07] Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl, editors. *The Adaptive Web: Methods and Strategies of Web Personalization*. Lecture Notes in Computer Science. Springer, Berlin, June 2007.
- [GGK⁺05] Osvaldo Gervasi, Marina L. Gavrilova, Vipin Kumar, Antonio Laganà, Heow Pueh Lee, Youngsong Mun, David Taniar, and Chih Jeng Kenneth Tan, editors. *Computational Science and Its Applications - ICCSA 2005, International Conference, Singapore, May 9-12, 2005, Proceedings, Part II*, volume 3481 of *Lecture Notes in Computer Science*. Springer, 2005.
- [JG00] Mark Endrei Maria Miccolis Erly Serrano John Ganci, Michael Adams. *WebSphere Personalization Solutions Guide*, 2000.
- [Lat08] Mareike Lattermann. *Implicit Social Network Construction and Expert User Determination for Context-Aware Web Portal Environments*. ACM, Stuttgart, Germany, 2008.
- [Liu07] Bing Liu. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications)*. Springer, January 2007.
- [NG08] Andreas Nauerz and Georg Groh. Implicit Social Network Construction and Expert User Determination in Web Portals. In *Proceedings of the AAAI Spring Symposium on Social Information*, Stanford University, California, USA, March 2008.

PerLE: Persönliche Lernumgebung zur Anzeige und Suche von Wissen

Christian Schafmeister

Universität Paderborn
Institut für Informatik
schafm@upb.de

Sebastian Nuhn

Universität Paderborn
Institut für Informatik
spider@upb.de

Ann-Cathrin Mikus

Universität Paderborn
Institut für Informatik
acb@zitmail.upb.de

Abstract: Die Clientsoftware PerLE ist ein Benutzerinterface um auf ein Single-Point-of-Information-System zuzugreifen, welches im Rahmen der Projektgruppe MokexIV eingesetzt wird. Neben der grundlegenden Funktionalität zur Dokumentenrecherche wurde der Schwerpunkt auf die Visualisierung von zusätzlichen Informationen gelegt. Startet der Benutzer eine Suchanfrage, werden neben den Suchergebnissen auch Experten und Verwandte Begriffe in Form von Listen und Netzen angezeigt; jeweils im Kontext zum aktuellen Suchwort. Um eine möglichst große Interoperabilität zu gewährleisten, wurde PerLE in Actionscript entwickelt und ist somit auf einer großzahl von Plattformen verfügbar.

Art der Arbeit: Projektgruppe

Betreuer: Wolfgang Reinhardt, Johannes Magenheim

1 Problemstellung

Auf dem Markt existieren viele Softwarelösungen, die dem Wissens- und Informationsmanagement dienen. Der Fokus der genannten Software liegt in dem Bestreben, dem Anwender die Bearbeitung von Arbeitsprozessen zu erleichtern. Eine Schwachstelle dieser Systeme liegt im Bereich der Gewinnung und Archivierung von Wissensobjekten. Ein Ansatz Wissensobjekte zu verarbeiten stellt das Konzept des *Single Point of Information* (SPI) [HHM⁺06] dar. Das einheitliche Erzeugen und Speichern, sowie die Suche von Informationsobjekten, steht bei diesem Konzept im Vordergrund. Die existierenden Softwarelösungen zur SPI-Umsetzung basieren darauf, den Benutzer beim Organisieren von Informationen und Aufgaben zu unterstützen. Im Rahmen der Projektgruppe MoKEx wird ein Gesamtkonzept entwickelt, welches auf dem des SPI-Konzept basiert. Bei der Weiterentwicklung wird das Konzept um neue Funktionalitäten erweitert. Neben der Speicherung und Suche von Wissensobjekten werden zusätzlich die Zusammenhänge zwischen diesen erfasst und visualisiert. Im Folgenden wird das Benutzerinterface PerLE¹ vorgestellt.

¹Personal Learning Enviroment

2 PerLE - Benutzerschnittstelle

Die Oberfläche der Clientsoftware PerLE besteht aus statischen und dynamischen Elementen, welche vom Benutzer frei konfiguriert werden können. Dabei kann der Benutzer wählen, welche Informationselemente er in einer vergrößerten Darstellung anzeigen haben möchte.

Die Elemente beinhalten beispielsweise die Suchergebnisse sowie Listen und Graphen, welche im Folgenden vorgestellt werden. Umfangreiche Suchfunktionen ermöglichen es dem Benutzer Informationen am SPI zu lokalisieren. Dabei erhält der Benutzer die Möglichkeit, die Suchtreffer hinsichtlich verschiedener Kriterien, wie zum Beispiel den Typ des Informationsobjektes, einzuschränken. Suchanfragen werden protokolliert und an eine Ratingkomponente [KBR09] übergeben, um das Wissen des jeweiligen Benutzers, über ein bestimmtes Thema, im System zu erfassen.



Ergänzend zu den Suchtreffern, wird basierend auf einem Taxonomiebaum, ein Netz mit verwandten Begriffen angezeigt. Begriffe zwischen denen ein kürzerer Pfad im Taxonomiebaum liegt, werden räumlich entsprechend näher abgebildet. Der Benutzer hat die Möglichkeit über jeden Knoten/Eintrag des Begriffsnetz/-liste eine gezielte Suche nach Informationsobjekten oder Experten zu starten. Die Top-Experten, zum jeweils verwandten Begriff, werden direkt beim Eintrag in der Liste dargestellt.

Zusätzlich zum Begriffsnetz/-liste, wird ein Expertennetz/-liste im Kontext zur aktuellen Suchanfrage visualisiert. Dabei basieren die Kanten im Expertengraph auf deren Kommunikation untereinander. Experten, welche sich verstärkt über ein Thema austauschen, werden visuell entsprechend mit einer größeren Kantenstärke im Graphen dargestellt als andere. Der Benutzer kann über die Knoten im Netz oder über die entsprechenden Einträge in der Expertenliste, direkt Kontakt mit dem jeweiligen Experten aufnehmen, so wird der Wissensaustausch und gegenseitige Hilfestellung durch das System unterstützt. Die benötigten Informationen werden über die Rating-Komponente angefordert.

Die Arbeitsschritte des Benutzers werden während der Session erfasst und in einer Chronik dargestellt, welche hinsichtlich des Umfangs, des zu erfassenden Zeitraums und der Inhalte konfigurierbar ist.

Die Software ist als Stateless-Client² konzipiert. Somit kann der Benutzer von jedem Arbeitsplatz aus die Anwendung aufrufen und startet mit seiner gewohnten Arbeitsumgebung. Zu den gespeicherten Daten gehören Konfigurationseinstellungen der Informationsquellen, Konfigurationen der angezeigten Fenster sowie Einstellungen für die Listen/Netze und Templates, durch die das Aussehen beeinflusst wird.

²Alle Daten werden serverseitig gespeichert und verwaltet.

3 PerLE - eingesetzte Technologien

Die Clientsoftware PerLE wird mit Hilfe von Adobe Flex³ realisiert. Diese Technologie stellt ein Entwicklungsframework zur Erstellung von *Rich Internet Applications (RIA)* zur Verfügung. Eine RIA kann als Schnittmenge aus Desktop, Internet und mobilen Technologien interpretiert werden [Wid03].

Der Adobe Flex Code der Anwendung (bestehend aus ActionScript3 und MXML-Code) wird durch den Compiler in ein Flash Objekt übersetzt, welches in jedem Internetbrowser mit installiertem Flash-Plugin⁴ genutzt werden kann. Das Plugin ist mittlerweile für viele Plattformen verfügbar und laut einer Statistik⁵ auf 99% aller internetfähigen PCs installiert. So ist ein einfaches Deployment sowie Updates der Software möglich, da keine lokale Installation der Clientsoftware vorgenommen werden muss.

Da die serverseitigen Komponenten Java Technologien nutzen und mit Hilfe des *Java Messaging Systems (JMS)* kommunizieren, benötigt die Clientsoftware einen Adapter, um Informationen austauschen zu können. Als Adapter wird hier der BlazeDS Server [Bla08] eingesetzt. Es handelt sich dabei um eine Java-basierte Remoting- und Messaging-Technologie. Adobe Flex Clients können über ein *Remote-Object* Methoden aufrufen um Daten zu übermitteln. Um in unserem Szenario mit dem

KNS [Win08] kommunizieren zu können, besitzt der BlazeDS Server eine Anbindung an den JMS Kanal des KNS, über den er mit Hilfe von Java Messages kommuniziert. Mit Hilfe des BlazeDS wird somit indirekt ein Kommunikationskanal zwischen Clientsoftware und Server hergestellt. Die Abbildung 1 veranschaulicht dieses Verfahren grafisch.

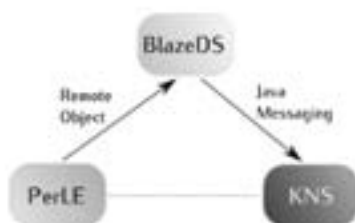


Abbildung 1: Überblick Kommunikation

4 PerLE - Ausblick

Adobe Flex besitzt bedingt durch die Einbettung in eine Browserumgebung gewisse Einschränkungen. Ein Zugriff auf das lokale Dateisystem, bzw. deren Überwachung ist nur bedingt realisierbar. Deshalb wird eine zusätzliche Version auf Basis der Adobe-Air-Technologie angestrebt. Adobe Air erweitert dabei die bekannten *Rich Internet Applications (RIA)* zu *Rich Desktop Applications (RDA)* [Wid03]. Der Mehraufwand der Erweiterung ist sehr gering. So ist es möglich, einen Features zu realisieren, die durch die Einschränkung der Flex-Technologie nicht möglich sind. Diese umfassen u.a. die Überwachung eines Upload Ordners, um neue Informationsobjekte in das System zu integrieren bzw. vorhandene Objekte zu modifizieren.

Die Anbindung zusätzlicher Informationsquellen wie RSS-Feeds, Wikis, Kalender und Onlinedienste wie Twitter und Flickr ist ebenfalls denkbar.

³<http://www.adobe.com/products/flex>

⁴http://www.adobe.com/products/player_census/flashplayer

⁵Stand: Sep. 2008

Literatur

- [Bla08] BlazeDS Developer Guide. http://livedocs.adobe.com/blazeds/1/blazeds_devguide/blazeds_devguide.pdf, 2008.
- [HHM⁺06] Knut Hinkelmann, Kai Holzweißig, Johannes Magenheim, Fabian Probst und Wolfgang Reinhardt. Linking Communities of Practice with Learning Communities in Computer Science Education. In *Proceedings of the 19th IFIP World Computer Congress (WCC2006)*, Santiago de Chile, August 20–25 2006.
- [KBR09] Andreas Kohring, Alexander Boschmann und Wolfgang Reinhardt. RaMBo: Finden von Experten in Projekten durch Profiling und Verhaltensbewertung. In *Eingereicht bei Informatiktag 2009*, 2009.
- [Wid03] Simon Widjaja. *Rich Internet Applications mit Adobe Flex 3*. Hanser Fachbuch, 2008-07-03.
- [Win08] Jens Winter. Konzeption und prototypische Implementierung eines Knowledge Servers mit Adaptionen zur Integration von eLearning- und Knowledge-Management-Applikationen. http://ddi.uni-paderborn.de/fileadmin/Informatik/AG-DDI/Arbeiten/Diplomarbeiten/2008/2008_Winter_web.pdf, 2008.

A Framework for Rapid Prototyping of Augmented Reality Interfaces

Stefan Radomski

sradomski@mintwerk.de

Abstract: This paper presents the jsArf, a framework for rapid development of augmented reality applications through scripting within a customized JavaScript runtime environment. The problem is to find intuitive *user interface paradigms* to interact with computer-generated, three-dimensional worlds. By describing these worlds and interaction techniques in an evaluated scripting language, more rapid iterations, incorporating user feedback become possible.

1 Motivation

In an AR system, computer-generated objects are continuously superimposed on a user's perception of reality, supplementing objects in the real world with spatially referenced information [ABBF01]. It is a new approach to *Human Computer Interaction* and expected to be part of an eventual successor to traditional interaction techniques [Dam97].

Most conventional frameworks to support the exploration of the AR design space provide a full-featured set of abstractions in a compiled programming language to implement high-performance and distributed AR systems [FHSS02] [BBKMSW03] [BFHLO04]. For the evaluation of different user interface paradigms, however, a simplified development process with rapid iterations is more suitable.

In this paper, an approach to implement AR systems as scripts within a JavaScript (JS) runtime is presented, reducing the time needed to iterate the process steps of implementation and evaluation, thus further accelerating the exploration of the AR design space.

2 Approach

An AR framework has to offer the means to describe the appearance of information as visual objects, to *register* them with objects from the real world, and to provide a way for a developer to specify the behavior of these objects when acted upon by a user.

Three-dimensional virtual worlds containing visual objects are commonly represented as a scene graph on top of the platform-specific functionality. The framework presented in this paper offers only three components to assemble scene graphs. Nodes containing

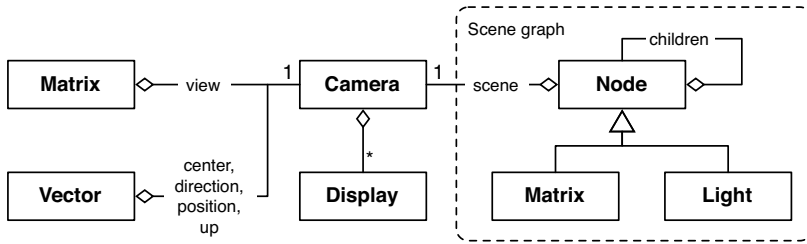


Figure 1: Object model of the jsArf.

descriptions of visual objects, matrices to rotate, scale and translate the nodes, and light to illuminate the scenery.

The position and orientation of objects in the real world is represented as matrices within the scene graph. This *pose* information is derived from the recognition of patterns printed on cardboard, *fiducials*, attached to the real world objects.

User input is discretized, and passed as events to appropriate JS handlers for processing. This requires special types of events and handlers for the various input devices, but preserves the generality needed to support a wide array of input devices.

To acquire projections of the scene graph on a two-dimensional plane, virtual cameras can be positioned in the scene. The cameras are assigned to a display, while specifying the extent and position of the camera's projection plane. The projections are processed on a more powerful stationary computer, and the images to be superimposed are transmitted on a wireless network connection to the display devices.

3 Implementation

The programming language chosen to implement the jsArf is C++ as it provides a wide range of libraries, is available on virtually all computer platforms, and is fast enough to fulfill the real-time aspect associated with interactive computer systems.

SpiderMonkey, the JavaScript implementation from the Firefox web-browser was used to provide the runtime environment for the scripts that modify the scene graph in response to events. For network transparency, the gSOAP Web Service implementation was chosen. The scene graph implementation, used to elevate the concepts on top of the system libraries for three-dimensional data is the OpenSceneGraph. Registration of virtual objects with their real-world counterparts is achieved by using the corresponding functionality provided by the ARToolkit. Furthermore, libraries to provide concurrency and reference counting were employed.

Extending the framework to support unforeseen applications is possible by selectively exposing additional abstractions from the underlying libraries as classes in the JS runtime environment.

4 Application

To verify the approach, a mock-up application for nautical navigation was developed using the framework. A fiducial attached to a PDA and a webcam, mounted on a tripod above a paper map act as an approximation to a desk dedicated to navigation [PS07]. The PDA is spatially registered and used as a *window into the information space* [Fitz93] to navigate the virtual world containing dynamic features of the displayed terrain. The user can place waypoints in the virtual world to denote coordinates for an eventual, external GPS system. After modeling the static setup of the application in C++, it took less than 150 lines of JS code to describe the actual application logic.

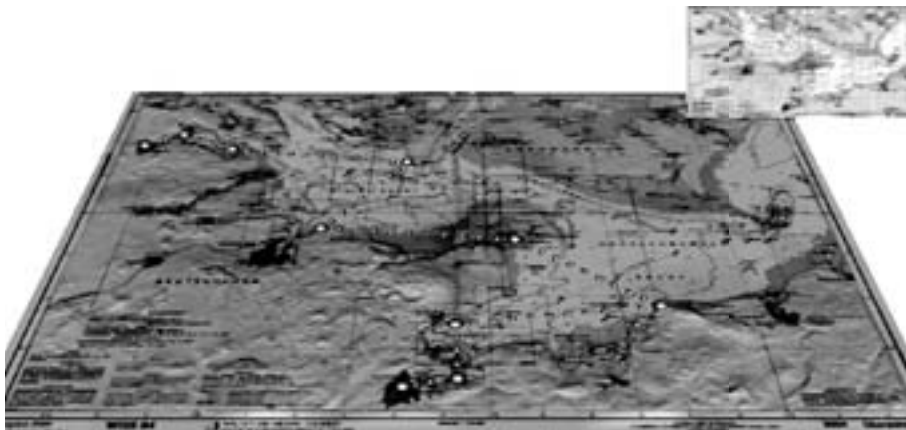


Figure 2: Baltic sea with height information and waypoints.

5 Conclusion

Modeling user interactions with AR systems as events, modifying a scene graph in a JS runtime environment, is a feasible approach to balance the forces of expressiveness, verbosity and rapid prototyping. The performance penalty involved with evaluating the JS scripts at runtime proved to be negligible when processing low-frequent events, such as discretized user input.

The functionality of remote displays is not yet implemented. The projections of the virtual worlds are instead rendered on a local screen. It has been shown, however, that wireless transmission of the images to be superimposed is a feasible solution [PW03].

There remains an untapped resource of *syntactic sugar* available within the JS language. Additional language features, such as *operator overloading*, will become available with the emergence of the JavaScript 2.0 standard, further minimizing the verbosity involved with modeling prototypical AR applications

References

- [ABBF01] R. Azuma, Y. Baillet, R. Behringer, S. Feiner (2001): *Recent advances in augmented reality*, in: Computer Graphics and Applications (2001) vol. 21 (6) pp. 34 - 47
- [BBKMSW03] A. MacWilliams, C. Sandor, M. Wagner, M. Bauer, G. Klinker, B. Bruegge (2003): *Herdin Sheep: Live System Development for Distributed Augmented Reality*, in: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality (2003) pp. 123
- [BFHLO04] J. Ohlenburg, I. Herbst, I. Lindt, T. Frhlich, W. Broll (2004): *The MORGAN framework: enabling dynamic multi-user AR and VR projects*, in: Proceedings of the ACM symposium on Virtual reality software and technology (2004) pp. 166 - 169
- [Dam97] A. van Dam (1997): *Post-WIMP user interfaces*, in: Communications of the ACM (1997) vol. 40 (2) pp. 63 - 67
- [FHSS02] D. Schmalstieg, A. Fuhrmann, G. Hesina, Z. Szalavari (2002): *The Studierstube augmented reality project*, in: Teleoperators and Virtual Environments (2002) vol. 11 (1) pp. 33 - 54
- [Fitz93] G. Fitzmaurice (1993): *Situated information spaces and spatially aware palmtop computers* in: Communications of the ACM (1993) vol. 36 (7) pp. 39 - 49
- [PS07] V. Paelke, M. Sester (2007): preprint: *Design Exploration of Augmented Paper Maps*. Hanover
- [PW03] W. Pasman, C. Woodward 2003: *Implementation of an augmented reality system on a PDA*, in: Second IEEE and ACM International Symposium on Mixed and Augmented Reality, Proceedings (2003) pp. 276 - 277

Entwicklung eines Konzepts zur modellbasierten Fehlerdiagnose

Christina Reck

Universität Paderborn
Institut für Informatik
AG Mensch-Computer-Interaktion und Softwaretechnologie
creck@upb.de

Art der Arbeit: Diplomarbeit

Betreuer der Arbeit: Prof. Dr. Szwillus

Abstract: In komplexen System ist oftmals auch die Suche nach Fehlerursachen komplex. Mit einem Konzept zur modellbasierten Fehlerdiagnose sollen die Abhängigkeiten in einer Maschine ausreichend beschreibbar sein, so dass den Bedienern eine tatsächliche Hilfe im Produktionsalltag geboten wird.

1 Projektentstehung

Die vorliegende Diplomarbeit „Entwicklung eines Konzepts zur modellbasierten Fehlerdiagnose“ entstand aus einer Kooperation mit der Firma MediSeal, die Medikamentenverpackungsmaschinen herstellt und weltweit vertreibt. In einer vorangegangenen Studienarbeit mit dem Titel „Usability im Betrieb – Untersuchung des Nutzungskontextes zur Bedienung von Verpackungsmaschinen am Beispiel MediSeal“ [REC07] wurden zuvor die Nutzungsprobleme im Umgang mit den Maschinen untersucht. Dabei ist deutlich geworden, dass eine Vielzahl von Faktoren zu Fehlbedienungen führt, was den Produktionsalltag erheblich stört. Ein besonderes Problem dabei ist die geringe Akzeptanz der aktuell existierenden Maschinenhilfe, die nur sehr selten bis gar nicht verwendet wird. Die Anwender erwarten von dort keine relevante Hilfestellung, da Probleme oftmals nicht erfasst oder nicht ausreichend behandelt werden. Dies ist durch Beobachtungen und Befragungen der Bediener offensichtlich geworden. Es zeigte sich, dass keine hinreichende Unterstützung zur Ursachenforschung geboten wird. In der Studienarbeit wurden bereits Vorschläge zur Verbesserung des Hilfesystems gemacht, die nun in der Diplomarbeit aufgegriffen wurden.

2 Zielsetzung

2.1 Fehlermodell

Ziel ist es, die integrierte Maschinenhilfe so zu verändern, dass sie einen effektiven Beitrag zur Suche nach Ursachen darstellt und dadurch für die Anwender zu einem nützlichen Hilfsmittel wird, das von diesen auch akzeptiert wird. Dabei soll systematisch eine modellbasierte Arbeits- und Denkweise umgesetzt werden. Die vielen in einer solchen Maschine bestehenden Abhängigkeiten sollen in einem Modell festgehalten und dann als Grundlage für das Geben von Hilfe dienen. Da der Ort der Fehlerentstehung vielfach nicht dem Punkt entspricht, an dem der Fehler sichtbar wird, ist es wichtig, die Propagation von Fehlern nachvollziehen zu können, um im geeigneten Maß den Bedienern im Störfall diese als Hilfe anzubieten. Der modulare Aufbau der Verpackungsmaschinen soll dabei explizit im Modell aufgegriffen werden.

2.2 Anforderungen an die Hilfe

Um die Aufgabenstellung und Problematik der Arbeit vertiefend zu beschreiben, sollen im Folgenden kurz die Anforderungen an das modellbasierte Hilfesystem skizziert werden.

Wiederverwendbarkeit. Die Hilfe und somit auch das Modell müssen gleichermaßen auf bereits existierenden Maschinen als Aufrüstung zu verwenden sein, wie auch bei Neuentwicklungen durch die Firma MediSeal einsetzbar sein. Die Industrie unterliegt einem stetigen Wandel und aus diesem Grund werden die Verpackungsmaschinen fortdauernd weiterentwickelt. Daher ist die Funktionalität der Wiederverwendbarkeit notwendig. Auf diesem Hintergrund ist es dann auch erforderlich, dass das Modell jederzeit veränderbar und erweiterbar ist.

Erlernbarkeit. Zudem muss das Erstellen bzw. Anpassen des Modells für die Mitarbeiter von MediSeal leicht zu realisieren sein. Hierfür ist es sinnvoll einen Editor zu planen, mit dem solche Arbeiten komfortabel umzusetzen sind und der über eine benutzerfreundliche Bedienung verfügt. Die detaillierte Ausführung und Programmierung wurde in einer parallel laufenden Studienarbeit durchgeführt.

Wechselwirkungen. Während der vorangegangenen Studienarbeit „Usability im Betrieb“ wurde festgestellt, dass an den Maschinen gänzlich unterschiedliche Benutzertypen arbeiten – dies reicht von einfachen Anfängern bis zu erfahrenen Mechanikern. Das Hilfesystem muss diesem Problem durch entsprechende Flexibilität begegnen.

Granularität bzw. Vollständigkeit. Das Modell wird von den Mitarbeitern von MediSeal erstellt. Allerdings muss das Konzept gewährleisten, dass jede Art von Fehler zu modellieren ist. Darüber hinaus sollte das Modell in geeigneter Granularität betrachtet werden können, so dass der Mitarbeiter die Übersicht über das Modell behält, er aber ebenfalls, wenn es notwendig ist, Störungen sehr detailreich darstellen kann.

3 Konzept und Umsetzung

Auf dem Hintergrund der beschriebenen Anforderungen wurde ein Konzept zur Realisierung des Hilfesystems entwickelt. Dieses beinhaltet eine Reihe wesentlicher Elemente, die kurz vorgestellt werden.

Hierarchischer Aufbau. Die Maschinen sind modular und hierarchisch in sogenannte Module, Bauteile und Parameter strukturiert. Um die Übersichtlichkeit zu steigern und um den Mitarbeitern der Firma MediSeal ein komfortables Arbeiten mit dem Modell zu ermöglichen, wird dieser hierarchische Aufbau im Modell nachvollzogen.

Verknüpfungen zwischen den Elementen. Damit kenntlich gemacht werden kann, dass zwei Elemente in Abhängigkeit zueinander stehen und dass sich auf diese Weise Fehler fortpflanzen, müssen Verknüpfungen eingerichtet werden können. Eine solche Verbindung repräsentiert daher immer eine Möglichkeit der Fehlerpropagation; ihr werden Ursachen- und Abhilfebeschreibungen zugeordnet. Eine solche Verknüpfung muss keineswegs wechselseitig bestehen, daher muss die „Leserichtung“ definiert sein.

Ausblenden nicht vorhandener Elemente. Es existieren für Kunden unterschiedliche Ausfertigungen der Maschine, so dass nicht alle Elemente der Maschine aktiviert sein müssen. Jedoch dürfen ausgeschaltete Bauteile nicht als Ursache präsentiert werden. Dies muss über einen Abgleich der aktuellen Maschinenkonfiguration gesichert werden.

Propagation. An Hand des Modells werden Ursachenketten ermittelt. Sehr oft sind nicht nur einzelne Fehler im Spiel, da Fehler sich durch Seiteneffekte auf andere Elemente auswirken können. Es entsteht eine Art „Fehlernetz“, worin die unterliegende Ursache für den Benutzer eventuell noch schwierig auszumachen ist. Durch die Propagation im Modell sollen solche Ketten und deren Ausgangspunkte ermittelt werden.

Reihenfolge für den Benutzer. Für eine Störung können gegebenenfalls mehrere Ursachen existieren und diese werden im Modell auch aufgespürt. Um die Nachhaltigkeit der Hilfe zu gewährleisten, ist es erforderlich, dass in der Praxis gerade die Ursachenketten angeboten werden, die auch zu einer wirklichen Lösung des Problems führen. Dazu sind Kriterien notwendig, mit denen Ursachenketten bzw. Abhängigkeiten bezüglich gewertet werden. Dazu werden die Wahrscheinlichkeit einer Ursache, die Schwere der Behebungsmaßnahmen und zum Teil (insbesondere für erfahrene Mechaniker) auch die unwahrscheinlichen Ursachen erfasst. Die Reihenfolge wird durch den aktuellen Bediener parametrisierbar gestaltet.

4 Ergebnisse

Die oben dargestellten Ideen wurden zu einem direkt umsetzbaren Implementationskonzept im Rahmen der vorliegenden Diplomarbeit „Entwicklung eines Konzepts zur modellbasierten Fehlerdiagnose“ ausgearbeitet. Es entstand so ein leistungsfähiges Modell zur Auffindung und Behandlung von Fehlern in komplexen Systemen. Durch die Abstützung auf ein geeignetes Maschinenmodell aus dem Hilfetexte und Ursachenbeschreibungen abgeleitet werden, muss der Entwickler der Hilfe nicht jede einzelne globale Fehlersituation beschreiben, was fehleranfällig und mühselig ist, und leicht in Gefahr gerät, unvollständig zu sein. Außerdem können die Fehlermodellanteile des einzelnen Moduls jedesmal wiederverwendet werden, wenn in einer Maschine dieses Modul eingebaut wird. Die Bewertung von Abhängigkeiten bietet zudem die Möglichkeit, Ursachen gezielt an die Vorkenntnisse der Bediener anzupassen.

Im Rahmen der Diplomarbeit und der parallel entstandenen Studienarbeit zur Implementation [SCH08] wurden exemplarisch einige Module und Abhängigkeiten modelliert. Dies hat sich als durchaus realistisch durchführbar erwiesen – der modellbasierte Ansatz hat sich in dieser Hinsicht als erfolgreich anwendbar gezeigt. Die Modellierung einer kompletten Maschine und Tests mit diesem Modell stehen noch aus – sind aber in Zusammenarbeit mit MediSeal angedacht. In einem nächsten Schritt sollen dann die bisher nur konzeptionell ausgearbeiteten Ansätze zur Implementation der tatsächlichen Hilfe umgesetzt werden.

Zwar wurde das Vorhaben auf die Produkte der Firma MediSeal zugeschnitten, jedoch ist eine Übertragung in gänzlich andere Bereiche ebenfalls denkbar. So können etwa Abhängigkeiten von Dialogschritten in Dialogsystemen modelliert und die flexible Behandlung von Eingabefehlern über dieses Verfahren realisiert werden. Eine weitere Studienarbeit hat sich auf andere Weise dem Problem genähert [NAS07]. Das Konzept soll daher verallgemeinert und für andere Anwendungsbereiche einsetzbar gemacht werden.

Literaturverzeichnis

[MAJ08] MediSeal (Alois Jandausch), Christian Schafmeister, Christina Reck: Diskussion über die Darstellung von Fehlerverknüpfungen im Modell, 2008, Persönliche Kommunikation

[NAS07] Mike Störmer und Steven Nahimi, Entwicklung einer intuitiven Webschnittstelle unter der Verwendung von AJAX am Beispiel der Berechnung von Kfz – Versicherungstarifen, 2007

[REC07] Christina Reck, Usability im Betrieb – Untersuchungen des Nutzungskontextes der Bedienung von Verpackungsmaschinen am Beispiel MediSeal, 2007

[SCH08] Christian Schafmeister, Konzeption und Implementation eines Editors zur Fehlermodellierung am Beispiel der MediSeal-Verpackungsmaschinen, Studienarbeit, Universität Paderborn, 2008

Integration physischer und mentaler Stressfaktoren in eine VR-basierte Simulationsumgebung

David Scherfgen

Hochschule Bonn-Rhein-Sieg
david.scherfgen@hochschule-bonn-rhein-sieg.de

Abstract: Im Rahmen des Projekts „FIVIS“ (Fahrradsimulation in der immersiven Visualisierungsumgebung „Immersion Square“) wird an der Hochschule Bonn-Rhein-Sieg ein immersiver Fahrradsimulator entwickelt. Neben der Verkehrserziehung soll dieser auch für Forschungsprojekte im Bereich der Arbeitssicherheit eingesetzt werden. Dazu müssen physischer und mentaler Stress getrennt voneinander skalierbar erzeugt werden können. Diese Bachelor-Arbeit beschreibt die Entwicklung und Evaluation der Simulationssoftware sowie eines Szenarios zur Erzeugung von Stress.

1 Einleitung

FIVIS ist ein Kooperationsprojekt der Hochschule Bonn-Rhein-Sieg in Sankt Augustin und des RheinAhrCampus in Remagen. Ziel ist die Entwicklung eines immersiven Fahrradsimulators. Geplante Anwendungen umfassen:

- **Stressforschung:** Hier liegt der Schwerpunkt dieser Arbeit. Fahrradfahren erzeugt sowohl physischen als auch mentalen Stress, Letzteres insbesondere im Straßenverkehr. Dies macht einen Fahrradsimulator attraktiv für die Erprobung von Messmethoden für kombinierte physische und mentale Belastungen, die auch bei arbeitsicherheitsrelevanten Tätigkeiten auftreten [Ste08].
- **Verkehrserziehung für Kinder:** Korrektes Verhalten im Straßenverkehr kann eingeübt werden, insbesondere in gefährlichen Situationen.
- **Training für Radsportler:** Mit der immersiven Visualisierung kann das Indoor-Training abwechslungsreicher und möglicherweise effektiver gestaltet werden.

Der Immersion Square ist ein immersives Dreiwand-Projektionssystem [HHHH05]. Für den Fahrradsimulator wurde eine kleineres, ähnliches System namens „FIVISquare“ entwickelt. Der FIVISquare besteht aus drei Projektionsleinwänden der Größe $1,4\text{ m} \times 1\text{ m}$, die ihr Bild per Rückprojektion über Spiegel von drei SXGA-Projektoren empfangen (Abbildung 1). Mit Hilfe einer Matrox TripleHead2Go-Einheit werden die drei Projektoren zu einem großen virtuellen Display kombiniert (bis zu 3840×1024 Pixel).

Während der Fahrt sitzt der Benutzer auf einem gewöhnlichen Mountainbike, das mit Sensoren ausgestattet ist, um die Interaktion mit der Simulation zu ermöglichen. Im Einzelnen sind dies ein Potentiometer-basierter Lenksensor zum Messen des Lenkwinkels und ein elektro-optischer Sensor zur Erfassung der Drehgeschwindigkeit des Hinterrads. Am Hinterrad ist eine Magnetbremse angebracht, um Tretwiderstand zu erzeugen. Das Fahrrad ist auf einem hydraulischen Hexapod mit sechs Freiheitsgraden montiert, was eine realistische Simulation von Anstiegen oder Kurvenfahrten erlaubt [SBB⁺07].



Abbildung 1: Fahrrad auf Hexapod im FIVISquare

Im Rahmen dieser Arbeit sollte eine erweiterbare Simulationssoftware entworfen und implementiert werden, die es dem Benutzer erlaubt, mit Hilfe des Fahrrads realitätsnah durch eine virtuelle Welt zu navigieren. Zudem sollte ein konkretes Szenario zur Erzeugung physischer und mentaler Belastungen entwickelt werden.

2 Konzept

Der Simulator basiert auf einem mehrschichtigen Simulationsmodell. Für die Erzeugung von Stress wurden verschiedene Methoden entwickelt. Die Intensitäten der einzelnen Stressfaktoren lassen sich unabhängig voneinander kontrollieren.

2.1 Simulationsmodell

Zur Simulation der Objekte (inklusive Fahrrad) und der Umgebung wurde ein abstraktes Modell entwickelt. Objekte in der Simulation werden dabei auf mehreren aufeinander aufbauenden Schichten abgebildet:

- **Physikalische/visuelle Schicht:** Auf dieser Ebene erfolgt die physikalische Simulation und die Visualisierung einzelner Objektkomponenten, wie etwa Rahmen, Lenker, Vorderrad und Hinterrad bei einem Fahrrad. Die Physiks simulation geschieht

mittels Festkörperdynamik. Komponenten können physikalisch durch Gelenke verbunden werden, die einen oder mehrere Freiheitsgrade blockieren. Auf diese Weise kann auch die Dynamik eines Fahrrads simuliert werden [FSR90].

- **Logische Schicht:** Physikalische/visuelle Komponenten werden zu logischen Objekten kombiniert. Durch Manipulation der Komponenten wird dem logischen Objekt erlaubt zu agieren (wie das Beschleunigen eines Autos durch Ausüben eines Drehmoments auf die Räder).
- **Verhaltensschicht:** „Controller“ abstrahieren das Verhalten von Objekten. Sie steuern die ihnen zugewiesenen Objekte, indem sie die von ihnen angebotenen Aktionen nutzen. Zum Steuern des virtuellen Fahrradobjekts durch den Benutzer existiert ein Controller, der die Daten des Lenk- und des Drehgeschwindigkeitssensors ausliest und auf das virtuelle Fahrrad überträgt.
- **Semantische Schicht:** In dieser Schicht reagiert die Simulation auf stattfindende Ereignisse, deren Bedeutung vom konkreten Szenario abhängt.

2.2 Erzeugung von Stress

Physischer Stress lässt sich durch die Natur des Fahrradfahrens vergleichsweise einfach erzeugen. Vom Benutzer wird verlangt, in begrenzter Zeit eine vorgegebene Strecke in der nachmodellierten Innenstadt von Siegburg abzufahren. Um die Belastung zu skalieren, können sowohl der Umsetzungsfaktor zwischen der Tretgeschwindigkeit und der Geschwindigkeit des virtuellen Fahrrads als auch die verfügbare Zeit verändert werden.

Zur Erzeugung von mentalem/emotionalem Stress wurde die Szene um entgegenkommende Fahrzeuge und zufällig auftauchende Hindernisse erweitert. Der Benutzer soll Kollisionen vermeiden. Die Frequenz und Geschwindigkeit der Fahrzeuge und Hindernisse können angepasst werden.

3 Implementierung

Der Simulator wurde größtenteils in C++ implementiert. Zudem wurde eine Schnittstelle zur Python-Skriptsprache integriert, worüber die Erweiterung des Simulators um neue Szenarien oder Objekte möglich ist. Die Visualisierung (Abbildung 2) wurde mit der Open Source-Grafik-Engine „Ogre3D“ realisiert, während die Physiksimulation auf der kostenlos verfügbaren Physik-Engine „PhysX“ von NVIDIA basiert. Bei der Darstellung der 3D-Grafik wird die Geometrie der angewinkelten Leinwände berücksichtigt, um eine für den Benutzer korrekte Perspektive zu erzeugen. Die Beschreibung der zu simulierenden Szenarien erfolgt mittels XML.



Abbildung 2: Screenshot des Simulators mit dem Siegburg-Modell

4 Ergebnisse und Fazit

Zur Evaluation wurden mehrere Tests durchgeführt. Im Rahmen der Verkehrserziehungsanwendung absolvierten einige Grundschüler im Simulator einen Rundkurs durch die Siegburger Innenstadt. Es wurde festgestellt, dass ihr Verhalten im Simulator vergleichbar mit dem in der Realität ist.

Mit dem Szenario zur Stresserzeugung konnten die physische und mentale Belastung eines Probanden, dessen physiologische Daten mit Hilfe verschiedener Sensoren aufgezeichnet wurden, durch Veränderung der Simulationsparameter beeinflusst werden.

Insgesamt hat sich der entwickelte Simulator als gut geeignet für die geplanten Einsatzzwecke herausgestellt. Im weiteren Verlauf des Projekts sollen eine realistische Straßenverkehrssimulation integriert [KH08] und die Fahrphysik des virtuellen Fahrrads weiter verbessert werden.

Literatur

- [FSR90] G. Franke, W. Suhr und F. Rieß. An Advanced Model of Bicycle Dynamics. 11(2):116–121, März 1990.
- [HHHH05] R. Herpers, F. Hetmann, A. Hau und W. Heiden. The Immersion Square – A Mobile Platform for Immersive Visualizations. In *Aktuelle Methoden der Laser- und Medizinphysik: 2. Remagener Physiktage*, Seiten 54–59, 2005.
- [KH08] M. Kutz und R. Herpers. Urban Traffic Simulation for Games. In *Future Play 2008*, Seiten 181–184, 2008.
- [SBB⁺07] O. Schulzyk, J. Bongartz, T. Bildhauer, U. Hartmann, B. Göbel, R. Herpers und D. Reintert. A Bicycle Simulator Based on a Motion Platform in a Virtual Reality Environment – FIVIS project. *Proceeding in Physics 114: Advances in Medical Engineering*, Seiten 323–328, 2007.
- [Ste08] H. Steiner. Monitoring of Multi Causal Strain at Mobile Workplaces by Using Diverse Physiological Data. Bericht, Hochschule Bonn-Rhein-Sieg, 2008.

Probabilistic User Models for the Verification of Human-Computer Interaction

Markus Wagner

wagnermar@uni-koblenz.de

Abstract: We present a method that allows the formalization, analysis, and verification of probabilistic human-computer interaction (HCI) with the support of model checking tools. Based on a method for the formalization of HCI under security aspects, our method allows us to model probabilistic user behavior as well as to offer support for the computation of the interaction's cost. It enables us to verify quantitative aspects of the interaction and to answer questions like “what is the probability that the user will send confidential information to unauthorized individuals” and “will the user (on average) be able to send the email within a given amount of time?”

1 Introduction

Interaction between computer and user is a two-way communication process, where the user enters commands and the system responds to the input. This is referred to as interactive control [ASO⁺07]. Unfortunately, this control does not always work perfectly and errors in HCI occur. An error can be regarded as a deviation of the achieved result from the intended result. A huge number of errors in HCI can be traced back to user interfaces that are insufficiently secured against these errors, irrespective of whether the user interacted incorrectly with the computer, or someone attacked the system. The consideration of user errors and their overall impact on the system should form an important part of an analysis of a system's usability, safety, and security.

Several techniques available can be used to verify the functional correctness of a system. Examples of such techniques are theorem proving, simulation/testing and model checking. We focus on the last technique, namely formal verification by means of model checking [CGP99]. The goal of our technique is to proof properties of the interaction, or more specifically, to formally prove that all possible ways of HCI in a given scenario conform to the requirements. The method of [BB06] allows the formalization, analysis, and verification of secure user interfaces and serves as the basis of our work. It is based on GOMS [JK96], which is a well-established user modeling method. In [BB06], the formal GOMS models are augmented with formal models of the application and the user's assumptions, finally allowing formal reasoning about the security of user interfaces.

2 Erroneous Interaction

The user of an application is not always aware of all possible consequences of his/her interaction potential. While in most cases only “time” is lost due to incorrect interaction, money is lost as well, or even humans injured. To distinguish between the different effects

of actions the term of *action costs* is introduced. This term is used to take quantities like “amount of time”, “amount of money” or “amount of resources” into consideration. Later, this is used to reason about the effects of minor errors, such as mistyping characters when writing an email, and major errors, such as sending confidential data to unauthorized individuals. In order to undo errors, additional actions have to be carried out, introducing *error costs*. In other words, human errors add extra costs to the total sum of the sequence of actions. Based on the total interaction cost and on the idea that the user has an amount of resources he is willing to invest during the interaction, the sequences can be classified into three categories: (1) optimal sequences with minimal total costs, (2) acceptable sequences with total costs that do not exceed the budget, and (3) unacceptable sequences with total costs that exceed the user’s budget.

The designers of systems often consider human frailty and try to reduce errors and usability problems. A popular approach is to stick to informal lists of design rules [Shn98, ASO⁺07, Lev97]. In general, most rules are obtained reflectively from experience, meaning that they are based on a combination of expert knowledge and empirical approaches to refine the rules. Furthermore, a number of techniques is known to help in making predictions about the reliability of a system when humans interact with it [SG83, Kir94]. They consist mainly of databases containing probabilities of different kinds of human errors. The error data comes from many sources, e.g. nuclear power plants, simulator studies, laboratory experiments etc. A method is provided for combining the data in order to produce estimations of the erroneous executions of tasks.

3 Formal Analysis of Erroneous Actions

In the following, we present our approach to the formalization, analysis, and verification of probabilistic HCI. Following [BB06], we assume that a user can only interact correctly with a system if he always has a correct assumption about the configuration of the application, including the internal state and relevant data. This means in terms of Linear Temporal Logic $\mathbf{G}((a_0 \leftrightarrow c_0) \wedge (a_1 \leftrightarrow c_1) \wedge \dots \wedge (a_n \leftrightarrow c_n))$, where a_0, \dots, a_n represent the critical properties of the application, and c_0, \dots, c_n represent the user’s assumption about critical properties. If a user fulfills this requirement, he is not error-prone. If he does not, the following scenarios are possible: (1) parts of the user’s assumptions are wrong, and (2) parts of the user’s assumptions are missing (over-abstraction).

In [BB06], *Input Output Labeled Transition Systems* (IOLTS) are used to model the components of the interaction. An IOLTS is a tuple $L = (S, \Sigma, s_0, \rightarrow)$ where S is a set of *states*, $s_0 \in S$ is an *initial state*, $\rightarrow \subseteq S \times \Sigma \times S$ is a *transition relation*, and a set of *labels* $\Sigma = \Sigma? \cup \Sigma! \cup \Sigma I$. We call $\Sigma?$ the *input alphabet*, $\Sigma!$ the *output alphabet*, and ΣI the *internal alphabet*. For example, the software’s state is represented by its internal state and data. Whenever the user executes an action, the corresponding state transition is performed. The corresponding edges in the automaton are annotated with a label denoting the action.

In our approach, three components are needed in order to apply automated reasoning to HCI: (1) a formal GOMS model and its corresponding IOLTS, (2) a component representing the user’s assumptions of the software’s state and its corresponding IOLTS, and (3) a component representing the application itself and its corresponding IOLTS. The method

of [BB06] uses these three components to produce a complete model of the interaction, which can be used as input for a model checker.

In our expansion, we continue to model the GOMS description using an IOLTS. The application’s model is extended by numerical values, representing the costs for the execution of each single step. We use a *probabilistic IOLTS with costs* to model the user’s assumptions. Thus, not only the probabilistic behavior can be modeled, but also “personal” arising costs. As sources for the statistical data, we suggest databases that are mentioned in the previous section. However, this is not mandatory and the analyst can come up with his/her own values. For the component representing the application itself, an *IOLTS with costs* is used. Now, the mutual composition of these three components provides the expanded model of the HCI (see Figure 1), incorporating probabilistic user behavior as well as the interaction’s costs.

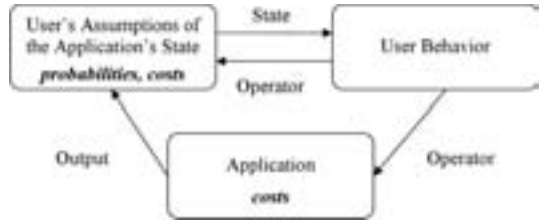


Figure 1: Extended model of the interaction.

The probabilistic model that is the output of our method can be used to prove quantitative aspects of the interaction. In general, probabilistic model checking is an automatic formal verification technique for the analysis of systems that exhibit stochastic behavior [HKNP06]. In our setting, properties of the HCI are expressed in Probabilistic Computation Tree Logic (PCTL) [HJ94] and verified using the probabilistic model checker PRISM [KNP02].

For demonstration purposes, let us think of a scenario where a user interacts with an email program. She intends to write a confidential email and send it to *Alice*. However, with a certain probability, she can choose *Bob* as well, which will result in high costs. For example, the following situations are possible: (1) the user accidentally selects *Bob* as the addressee, notices her mistake, and corrects her mistake, (2) she selects *Bob* again, but does not notice her mistake, and sends the email to *Bob*, and (3) she selects *Bob*, notices it, but fails to change the addressee, and sends the email to *Bob*. Once the model of the interaction is constructed using our method, properties of the interaction can be formulated in PCTL and then checked by PRISM. For example, if the property $P_{\geq 0.95}[\text{true } U \text{ sentTo}(\text{Alice})]$ holds, it is formally proven that the email is sent to *Alice* in at least 95% of the the interactions. The property, $P_{\leq 0.20}[\text{chosen}(\text{Bob}) U \text{ sentTo}(\text{Bob})]$ can be used to express that “with a probability of at most 0.20, the selection of *Bob* as the addressee is not corrected”. A way to lower the probability of sending the email to *Bob* would be to introduce additional dialogue boxes that would require the user to confirm her selection. If the user is modeled to react “reasonably” (versus “inattentively”) to a confirmation, it can be proven that the probability to send the email to *Bob* is lowered. Finally, with PRISM’s support of cost analysis, the average total cost for sending an email can be computed by the property $R = ? [F \text{ email_sent}]$. Due to the possibility of erroneous interaction, the average total cost here will be higher than in a scenario where errors do not result in additional cost.

References

- [ASO⁺07] Larry W. Avery, Thomas F. Sanquist, Peter A. O'Mara, Anthony P. Shepard, and Daniel T. Donohoo. U.S. Army weapon systems human-computer interface style guide. Version 2, April 30 2007.
- [BB06] Bernhard Beckert and Gerd Beuster. A Method for Formalizing, Analyzing, and Verifying Secure User Interfaces. In *ICFEM*, volume 4260 of *Lecture Notes in Computer Science*, pages 55–73. Springer, 2006.
- [CGP99] Edmund M. Clarke, Orna Grumberg, and Doron A. Peled. *Model Checking*. The MIT Press, Cambridge, Massachusetts, 1999.
- [HJ94] Hans Hansson and Bengt Jonsson. A Logic for Reasoning about Time and Reliability. *Formal Aspects of Computing*, 6(5):512–535, 1994.
- [HKNP06] Andrew Hinton, Marta Z. Kwiatkowska, Gethin Norman, and David Parker. PRISM: A Tool for Automatic Verification of Probabilistic Systems. In *TACAS*, volume 3920 of *Lecture Notes in Computer Science*, pages 441–444. Springer, 2006.
- [JK96] Bonnie E. John and David E. Kieras. The GOMS family of user interface analysis techniques: comparison and contrast. *ACM Transactions on Computer-Human Interaction*, 3(4):320–351, December 1996.
- [Kir94] Barry Kirwan. *A Guide to Practical Human Reliability Assessment*. Taylor and Francis, London, UK, 1994.
- [KNP02] Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM: Probabilistic Symbolic Model Checker. *Lecture Notes in Computer Science*, 2324:200–204, 2002.
- [Lev97] Nancy G. Leveson. Analyzing Software Specifications for Mode Confusion Potential. In *Proc. of the Workshop on Human Error and System Development*. 1997.
- [SG83] Alan Swain and H. E. Guttman. *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*. Sandia National Laboratories, 1983.
- [Shn98] Ben Shneiderman. *Designing the User Interface*. Addison Wesley Longman, third edition, 1998.

Entwicklung eines flexiblen Plugin-Systems für das Text-Mining im Kontext der Technologieplanung

Oliver Buschjost

HEINZ NIXDORF INSTITUT
Universität Paderborn
oliver.buschjost@hni.upb.de

Abstract: In der Technologieplanung ist es notwendig über die aktuellen Entwicklung neuer Technologien informiert zu sein und diese rechtzeitig zu antizipieren. Die regelmäßige Überwachung der Veränderungen ist ein aufwendiger, manueller Prozess. Die am Heinz Nixdorf Institut entwickelte Innovations-Datenbank unterstützt die strategische Produkt- und Technologieplanung mit Hilfe eines strukturierten Prozesses. Zurzeit bietet die Innovations-Datenbank jedoch keine automatisierten Verfahren zur Recherche nach Technologie-Informationen. Das im Rahmen einer Diplomarbeit entwickelte Plugin-System setzt hier an. Es bietet ein Verfahren zur Automatisierung großer Teile des Rechercheprozesses und versorgt die Mitarbeiter mit direkt nutzbaren digitalen Informationen. Das Ergebnis ist ein integriertes Recherchewerkzeug, das die Anwender in ihrer Arbeit unterstützt.

Art der Arbeit: Diplomarbeit

Betreuer: Wolfgang Reinhardt, Stephan Ihmels

1 Problemstellung

Unternehmen werden heute durch den Wandel der Märkte und die dynamische Entwicklung von Technologien vor neue Herausforderungen gestellt. In dieser Situation besteht – vor allem auch aufgrund kürzer werdender Produktzyklen – die Gefahr, Erfolgspotentiale für neue Geschäfte zu übersehen. Eine Ursache ist die zu starke Fokussierung auf die Anforderungen von Kunden für die Produktentwicklung. Daraus resultierend folgen die Unternehmen dem Markt und gestalten ihn nicht aktiv. Für Produktinnovationen ist es nötig, zu Beginn des Produktentstehungsprozesses bereits Erfolgspotentiale zu erkennen [BI07, WV07].

Zur Unterstützung der Unternehmen im strukturierten Produktentwicklungsprozess wurde am Heinz Nixdorf Institut – im Kontext der strategischen Produktions- und Technologieplanung – die Innovations-Datenbank (IDB) entwickelt und in verschiedenen Unternehmen erfolgreich eingeführt [BIH08]. Mit Hilfe der Innovations-Datenbank wird den Unternehmen ein Instrumentarium zur Verfügung gestellt, bestehende und neu aufkommende Technologien zu verfolgen und die aus aktueller Forschung resultierenden Chancen bewerten zu können [BI07].

Als Ausgangspunkt für eine realistische Bewertung kann nur eine fundierte und stets aktuelle Datenbasis dienen. Die regelmäßige Aktualisierung und Pflege von Technologie-Informationen erzeugt bei den verantwortlichen Mitarbeitern einen hohen Aufwand. Dieser regelmäßig anfallende Zeitaufwand soll durch ein zielgerichtetes Information-Retrieval System reduziert werden. Hier ergibt sich die Möglichkeit, den Mitarbeitern vorgefilterte Informationen aus einer breitgefächerten Auswahl elektronischer Datenquellen an die Hand zu geben, um den Personalaufwand zu verringern.

Aufgrund der Vielzahl an möglichen Datenquellen und der ständigen Veränderung der dieser unterliegen, ist es notwendig ein flexibel anpassbares System zu entwickeln, das einfach an weitere Datenquellen angepasst werden kann. Neben den Datenquellen selbst – in diesem Fall u.a. allgemeine Suchmaschinen, Archive von Konferenzen, Patentdatenbanken, soziale Netzwerke sowie Zeitungen und Zeitschriften – können auch die gelieferten Resultate in einer Vielzahl an zu berücksichtigenden Formaten – Office-Dokumente, (evtl. gescannte) PDF-Dateien, HTML, etc. – vorliegen.

2 Konzept

Für die Ablaufsteuerung wird die Windows Workflow Foundation von Microsoft eingesetzt. Innerhalb des hierdurch gesteuerten Prozesses (Abbildung 1) werden alle gefundenen Plugins¹ in das System geladen und ausgeführt. Da die Recherche und Analyse der Daten sehr zeitaufwendig ist, wurden diese Schritte in einen separaten Server ausgelagert. Ein Ablauf der Client-Server-Kommunikation ist in Abbildung 2 dargestellt.

Das System soll den Anwendern zu den gewünschten Technologien relevante Informationen liefern, ohne dass weitere Eingaben notwendig sind. Dazu werden bereits in der Innovations-Datenbank enthaltene Daten an den Workflow übermittelt und während der Suchkontext-Erweiterung mit weiteren Informationen angereichert.

Die Auswahlkriterien für die anzubindenden Dienste und Programme umfassen:

- Lizenzbestimmungen - Die verwendeten Datenquellen müssen frei verfügbar oder lizenziert sein und den automatisierten Abruf der Inhalte erlauben. Dadurch entfallen einige interessante Datenquellen (Patentdatenbanken, Fachmagazinen und Konferenzen).
- Günstige Realisierung - Die verwendeten Werkzeuge und Datenquellen sollten unter einer liberalen Open-Source-Lizenz stehen oder günstig lizensierbar sein.
- Relevanz der Informationen für die Technologieplanung
- Vorhandensein einer API zum direkten Zugriff

¹Die Plugins sind als .NET Assemblies realisiert.



Abbildung 1: Workflow

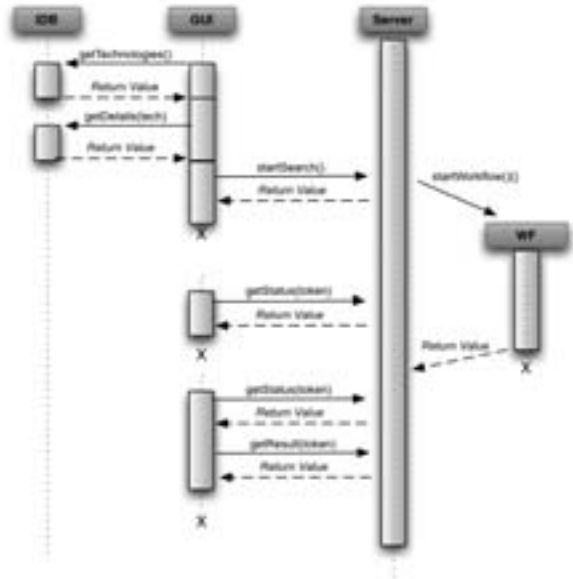


Abbildung 2: Ablauf der Client-Server Kommunikation

Unterstützt sind aktuell folgende Typen von Plugins:

Datenquellen - Diese Plugins² generieren zu den übergebenen Daten passende Suchanfragen für die konkrete Datenquelle und liefern die gefundenen Resultate der Originalquelle.

Suchkontext-Erweiterung - Diese Plugins liefern – basierend auf einer Ontologie des zu analysierenden Themenumfeldes und mit Hilfe von Synonymdatenbanken und social-tagging – weitere passende Suchwörter bzw. Kombinationen zur Abfrage der Datenquellen.

Datendownload - Diese Plugins laden die gefundenen Dokumente von den verschiedenen Quellen in das lokale System.

Text-Extraktion - Es werden Texte aus den Dokumenten extrahiert und zur späteren Weiterverarbeitung (z.B. Entfernen von Markup) aufbereitet.

Sprachbestimmung - Bestimmung der Sprache der extrahierten Texte. Realisiert durch eine statistische Analyse auf Basis von n-Grammen [Kra00, CT94].

Metadaten-Extraktion - Extraktion der Metadaten³.

²Aktuell existieren Plugins für die Suchmaschinen Google, Yahoo und eine firmenintern; für den sozialen Bookmarkdienst del.icio.us sowie Google-News.

³Zur Nutzung bietet sich das MetaXsA System [MS08, SMR08] an.

Text-Processing - Verschiedene Schnittstellen, um die Texte zum Clustern vorzubereiten. Umgesetzt sind u.a. das Aufsplitten in einzelne Wörter, das Filtern von Stoppwörtern, die Erzeugung von BagOfWords⁴ und Stemming⁵, Interessante Möglichkeiten ergeben sich hier bei Anbindung von Part-of-Speech-Tagging und Named-Entity-Recognition Verfahren.

Bestimmung der Termgewichte - Die Termgewichte werden aus der Häufigkeit des Vorkommens von Wörtern innerhalb eines Dokumentes bestimmt und dienen zur Berechnung der inversen Dokumentfrequenz, welche die Relevanz einzelner Wörter im Dokument – im Vergleich zu einem Referenzkorpus – angibt.

Clustering - Plugins zur Identifikation von Clustern in den Daten. Es ist die Übergabe von hierarchisch geclusterten Daten ebenso möglich, wie von partitionierend geclusterten⁶ und es können zusätzlicher Daten (z.B. die Relevanz eines Clusters und ein Titel) übermittelt werden. Die Plugins zur Clustering haben Zugriff auf alle extrahierten Daten und können auch Metadaten einbeziehen.

3 Zusammenfassung und Ausblick

Das System erlaubt dem Anwender, die notwendigen Recherchen dank maschineller Unterstützung effizienter durchzuführen. Der modulare Aufbau des entwickelten Plugin- und Workflow-Konzeptes erlaubt es, mit geringem Aufwand weitere Funktionen und Datenquellen einzubinden und Verfahren zur Analyse und Clustering gegen andere auszutauschen. Damit ist es auf die Praxisanforderungen vorbereitet.

Literatur

- [BI07] V. Brink und S. Ihmels. Strategische Produkt- und Technologieplanung mit der Innovations-Datenbank. *Konstruktion*, 11/12:65–69, 2007.
- [BIH08] V. Brink, S. Ihmels und J. Haug. Einführung eines Instrumentariums zur strategischen Produkt- und Technologieplanung in einem Unternehmen der elektrischen Antriebstechnik. In *Tagungsband des 4. Symposium für Vorausschau und Technologieplanung*, 30. Oktober 2008, Berlin. Heinz Nixdorf Institut, 2008.
- [CT94] William B. Cavnar und John M. Trenkle. N-Gram-Based Text Categorization. In *Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval*, Seiten 161–175, Las Vegas, US, 1994.
- [Kra00] Simon Kranig. Evaluation of Language Identification Methods. Diplomarbeit, Universität Tübingen, 2000.

⁴Eine Datenstruktur, welche alle Wörter eines Textes und die Häufigkeit des Auftretens beinhaltet. Diese wird als Input für die Clustering genutzt.

⁵Reduzierung von verschiedenen morphologischen Varianten auf Ihren Wortstamm

⁶Aktuell wird k-Means zur Clustering verwendet [Nou04].

- [MS08] Christian Mletzko und Benedikt Schmidt. Automatische Extraktion von Metadaten aus verschiedenen Dateiformaten. In Gesellschaft für Informatik, Hrsg., *Informatiktage 2008. Fachwissenschaftlicher Informatik-Kongress, 14. und 15. März 2008, B-IT Bonn-Aachen International Center for Information Technology in Bonn*, Jgg. S-6 of *Lecture Notes in Informatics*, Seiten 181–184, 2008.
- [Nou04] Nebil Nouri. K-Means-Clustering für semistrukturierte Daten. Diplomarbeit, Universität Dortmund, FB Informatik, 2004.
- [SMR08] Benedikt Schmidt, Christian Mletzko und Wolfgang Reinhardt. MetaXsA - Modulares Framework zur semantischen Dokumentanalyse für unstrukturierte Textdokumente. In Gesellschaft für Informatik, Hrsg., *Informatiktage 2008. Fachwissenschaftlicher Informatik-Kongress, 14. und 15. März 2008, B-IT Bonn-Aachen International Center for Information Technology in Bonn*, Jgg. S-6 of *Lecture Notes in Informatics*, Seiten 197–200, 2008.
- [WV07] E. Westkämper und A. Verl. *Tagungsband von Werkzeuge des Technologiemanagements: Portfolio, SWOT, Roadmapping & Co. Fraunhofer IPA Workshop F 152. 11. Oktober 2007, Stuttgart*, 2007.

Indoor Navigation Based on Sectorized Positioning

Ory Chowaw-Liebman

ory.chowaw@rwth-aachen.de

Abstract: Navigation aids have usually concentrated on the great outdoors, whether driving on highways or, more recently, walking through cities. These systems use GPS for position information. Indoor navigation cannot rely on GPS. The following presents a indoor system currently being developed, based on a novel technique called whispering.

1 Introduction

Many Travelers today use technological aids to find to their destination. The path is given by a set of way points, and connected by the means of transportation. Commonly, it is the travelers responsibility for finding his or her way at a location to the next conveyance, a challenging task in large train stations or airports.

Electronic navigation systems rely on GPS to obtain position information. Unfortunately the radio waves used by the GPS system do not propagate through roofs (and even dense foliage). In order to continue using existing navigation solutions, trilateration, and more recently multilateration, are used to compute precise positions from Mobile network base stations or WLAN access points. Being computationally intensive, these methods will most likely take too long on current mobile devices to provide a satisfactory user experience.

A technique named Whispering [KK08] which trades accuracy for speed was developed at the Department of Computer Science, Informatik 4. of the RWTH Aachen. This method is based on wireless LAN, and associates the mobile device with the closest access point. This is done by controlling the transmitting power WLAN interface: The device is set to minimum power, then queries access points for management information. The closest access point is defined as the one answering to the request with the minimum amount of transmission power used.

Being based on indoor infrastructure reception should not be a problem, assuming existing infrastructure. But WLAN for Internet access is more and more provided as a service to guests. The whispering method is also robust against signal multi-path propagation and power oscillations or automated adaption of access points.

2 Geometry Representation

The whispering technique defines a mapping from a point in space to the closest element in a set of special points, the access points. This type of mapping, known as Voronoi Diagrams, is an important concept in computational geometry [AK, Aur91]. Efficient algorithms to compute the voronoi diagram for a set of points exist [AK, ?]. Thus, a natural, though idealized, sectorization of a floor plan can be generated automatically:

1. Calculate Voronoi diagram of access points
2. Clip sectors to floor plan

These sectors, formally known as voronoi regions, form the basic element of navigation. Except where clipped to corners, they are convex polygons (again [AK]). Finding paths in such a representation is a simple application of the A* algorithm [RN03, Dal03].

A path to some destination is represented as a series of sectors. The user guidance can be broken down into two components:

1. Guide the user to the next sector,
2. Guide the user to the target in the destination sector,

The actual map used for navigation consists of the floor plan, the sectorization and all points of interest like check-ins, gates, platforms, cafes, restrooms and so on. The sectors can further be grouped together in order represent form long hallways and large rooms, traversal of which may need less or more information content.

A complete way may include multiple consecutive e.g. in an airport after check in a bookstore is to be visited, some coffee is needed, the gate must be reached in time, the duty free store and maybe a rest-room (constrained to have a baby station) before boarding the plane.

So a complete way consists of a number of tasks, possibly ordered, which may have time constraints. From every task, a path to the next one is needed. While a given location may imply a number of tasks, users can add to the base minimum at any time.

3 User Interaction

Presentation of path information to users with mobile devices is an active area of research. Use of Visual landmarks is considered on the one hand, e.g. [EPK05, TL, LAS⁺], natural language instructions on the other [Rei, KB, PE, TL]. The fact that pedestrians are less constrained than cars in their movement is considered in [GF08]

Landmarks can be anything, as long as it is easily seen. Stores and other corporate entities (preferably with a nice iconic logo) are useful. In cities, Statues, fountains, buildings, the cologne cathedral and the tour d'Eiffel, and other elements of the surroundings are used

which are less interesting for indoors applications. Staircases, elevators, emergency exits and phone booths (while they exist) are exemplary of architectural features which make for good indoor landmarks. Landmarks can be displayed on the map by drawing icons and logos at the appropriate places.

Landmarks are regularly used by humans in natural language instructions. Human languages offer a plethora of ways to express absolute directions (straight ahead, turn left, in and out), relative locations (By the stairs, next to the record store) and description (colors, shapes, labels).

4 Services

A number of services must be provided by a complete indoor system. First and foremost is map acquisition. Context based information like a flight's check in or platform changes, and emergency information like fire alerts also need to be taken into account to provide a user with necessary information.

The classic approaches are either completely pre-loading all necessary data (e.g. maps of countries) which is not flexible, or a mobile Internet connection, GPRS or UMTS which are expensive.

The whispering technique on the other hand assumes the existence of a WLAN, which can be used to access a local server. This network can be used to provide local information to the mobile devices without the need for a mobile phone network data connection, and avoids the use of a global server.

A simple directory service can be offered by the servers in order to browse for specific targets. In various cases information beyond navigation becomes necessary for the user to make decisions. For example, while the closest rest-room is usually sufficient, choice of hotels depends on available rooms and pricing and food may be desired to be pork-free or vegetarian. Further restrictions may be applied by disabled user.

5 Conclusion

A system for indoor navigation based on the Whispering technique was presented. It is designed to operate on a rough approximation of the current position. The system is being developed in the course of a diploma thesis at the Department for Computer Science I4.

References

- [AK] Franz Aurenhammer and Rolf Klein. Voronoi Diagrams.
- [Aur91] Franz Aurenhammer. Voronoi Diagrams - A Survey of a Fundamental Geometric Data

- Structure. *ACM Computing Surveys*, 23(3), 1991.
- [BDH95] C. Bradford Barber, David P. Dobkin, and Hanu Huhdanpa. The Quickhull Algorithm for Convex Hulls. *ACM Transactions on Mathematical Software*, 1995.
 - [Dal03] Daniel Sánchez-Crespo Dalmau. *Core Techniques and Algorithms in Game Programming*. New Riders Publishing, 2003.
 - [EPK05] Birgit Elias, Volker Pelke, and Sascha Kuhnt. Concepts for the Cartographic Visualization of Landmarks. 2005.
 - [GF08] Christian Gaisbauer and Andrew U. Frank. Wayfinding Model For Pedestrian Navigation. 2008.
 - [KB] Christian Kray and Anselm Blocher. Modeling the Basic Meanings of Path Relations.
 - [KK08] Karl-Heinz Krempels and Martin Krebs. Improving directory-less wlan positioning by device whispering. 2008.
 - [LAS⁺] E. Lazkano, A. Astigarraga, B. Sierra, J. M. Martínez-Otzeta, and I Rañó. Natural landmark based navigation.
 - [PE] Bert Paepen and Jan Engelen. Using a Walk Ontology for Capturing Language Independent Navigation Instructions.
 - [Rei] Ehud Reiter. Building Natural Language Generation Systems.
 - [RN03] Stuart Russel and Peter Norvig. *Artificial Intelligence: A Modern Approach, 2nd Ed.* Pearson Education, 2003.
 - [TL] Barbara Tversky and Paul U. Lee. Pictorial and Verbal Tools for Conveying Routes.

Mining orthogonaler Konzepte in hochdimensionalen Datenbanken

Ines Färber

ines.farber@rwth-aachen.de

RWTH Aachen University

Lehrstuhl für Informatik 9, Prof. Dr. rer. nat. Thomas Seidl

Art der Arbeit: Diplomarbeit

Betreuer der Arbeit: Dipl.-Inform. Emmanuel Müller, Dipl.-Inform. Stephan Günnemann

Abstract: Ziel meiner Diplomarbeit ist es, einen Clusteringalgorithmus zu entwickeln, welcher für hochdimensionale Daten möglichst relevante Gruppierungen liefert. Für die Relevanz eines Clusters ist dabei nicht nur der Cluster selbst maßgeblich, sondern auch seine Beziehung zu anderen erkannten Clustern. Relevant sollen Cluster genau dann sein, wenn sie neue Informationen enthalten, also entweder neue Objekte abdecken, oder für bereits abgedeckte Objekte neue orthogonale Konzepte aufdecken.

1 Einleitung

Clustering ist ein spezielles Data Mining Verfahren, welches Gruppierungen in den Daten ermittelt. In einem erfolgreichen Clustering weisen Datenobjekte innerhalb eines Clusters untereinander eine hohe Ähnlichkeit auf, Datenobjekte aus unterschiedlichen Clustern sind dagegen möglichst unähnlich zueinander.

Auf vielen Datensätzen lässt sich das traditionelle Clustering, z.B. aufgrund zu hoher Dimensionalität nicht mehr anwenden, da keine Cluster im hochdimensionalen Raum zu erwarten sind. In realen Anwendungen sind für Datenobjekte oft mehrere Interpretationen denkbar, sodass verschiedene Teilräume unterschiedliche Strukturen der Daten offenlegen können. Das Beispiel in Abbildung 1 zeigt, dass die Attribute Sport und Obstkonsum das Konzept Gesundheitsbewusstsein aufdecken, wohingegen Rock- und Klassikkonzertbesuche das Konzept Musikvorlieben repräsentieren. Kern dieses Ansatzes ist es gezielt verschiedenartige Strukturen, bzw. orthogonale Konzepte der Daten zu extrahieren bei gleichzeitiger Minimierung der Redundanz und Maximierung des Informationsgehaltes.

Einen Lösungsansatz bieten bereits bekannte Subspace Clustering Verfahren, welche Cluster in allen möglichen Teilräumen suchen [AGGR98]. Durch die im hochdimensionalen sehr große Ergebnismenge, bei der häufig die Clusteranzahl die Datenbankgröße übersteigt, wird allerdings die Datenanalyse erheblich erschwert. Eine Alternative stellen Projected Clustering Verfahren dar, welche in allen Teilräumen nach einer Partitionierung der Daten suchen [AWY⁺99]. Die Partitionierung erzwingt eine Zuordnung jedes Objektes zu genau einer Gruppe, wodurch die Erkennung verschiedener Konzepte verhindert wird.

Ziel meiner Arbeit ist es die Vorteile beider Ansätze zu kombinieren, sodass bei der Erkennung orthogonaler Konzepte ein enormes Anwachsen der Subspace Cluster Ergebnismenge, durch Einschränkung auf relevante Cluster, vermieden wird.

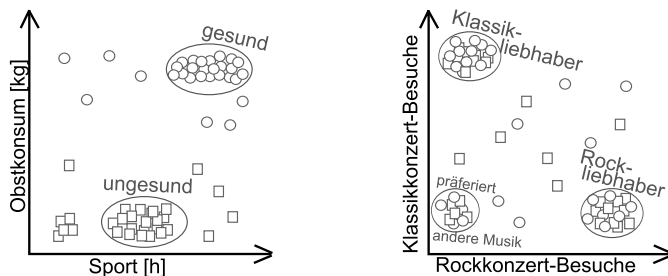


Abbildung 1: verschiedene Sichten auf die gleichen Datenobjekte (Durchschnittswerte eines Jahres).

2 Multi-View Clustering

Zur Erkennung orthogonaler Konzepte existieren bereits Ansätze, welche jedoch einige Nachteile mit sich bringen. Ein Lösungsansatz des Multi-View Clusterings [YC07] verwendet den Clusteringalgorithmus K-Means [Mac67], bei dem allerdings durch Vorgabe der Clusteranzahl ein großes Vorwissen vorausgesetzt wird und zusätzlich lediglich sphärische Cluster erkannt werden. Durch Einsatz der globalen PCA-Dimensionsreduktion [DHZS02] vor einem Orthogonalisierungsverfahren zur Generierung unterschiedlicher Konzepte leidet zudem die Interpretierbarkeit der Ergebnisse, da einem Cluster keine eindeutige Attribut-Teilmenge mehr zuzuordnen ist.

Die Schwächen dieses Ansatzes motivieren die Entwicklung eines eigenen Verfahrens zur Erkennung global optimierter, orthogonaler Konzepte. So können die bekannten Nachteile des K-Means Verfahrens durch Einsatz eines Dichte-basierten Clusteransatzes [MEX96, KK04] vermieden werden. Die wesentliche Neuerung meines Ansatzes findet sich allerdings in der deutlichen Verbesserung der Interpretierbarkeit der Ergebnisse durch Clustering in Originalteilräumen, aber vor allem durch eine globale Betrachtung der Clusterinteressantheit.

2.1 Globales Interessantheits-Modell

Viele Clusterdefinitionen, wie z.B. in CLIQUE [AGGR98], erfüllen ein Monotonie-Kriterium, sodass ein Cluster in einem Raum S auch in den Projektionen aller Teilräume von S wieder einen Cluster bildet. Zwar werden durch Projektionen in Teilräume weitere Punkte abgedeckt, der Informationsgewinn ist in Anbetracht der exponentiell vielen neuen Cluster jedoch zweifelhaft. Einen Ansatz zur Vermeidung lokaler Redundanz durch Cluster in Teilräumen anderer Cluster liefert das Subspace Clustering Verfahren

INSCY [AKMS08]. In dieser Arbeit soll eine globale Redundanzanalyse auf Cluster aller Teilräume ausgeweitet werden, die zu einem gegebenen Teilraum keine ausreichende Orthogonalität sicherstellen. Um einen derartigen Redundanz-minimierenden Clusteralgorithmus zu entwickeln, muss zunächst ein Interessantheitsmaß definiert werden, welches die Clusterinteressantheit auch in Relation zu anderen Clustern bewertet. Das Interessantheitsmaß beinhaltet zwei Komponenten, zum einen die lokale Clusterinteressantheit, auf deren Grundlage der Informationsgehalt des Clusterings maximiert wird und zum anderen die globale Clusterinteressantheit, welche zur Vermeidung von Redundanz dient. Die lokale Clusterinteressantheit bewertet Kriterien, wie Größe, Dichte und Dimensionalität des Clusters und definiert damit, analog zu bekannten Subspace Cluster Verfahren, die Güte eines Clusters unabhängig von anderen Clustern. Die globale Clusterinteressantheit bewertet den Cluster im Kontext aller Cluster der Ergebnismenge und betrachtet so den Informationsgewinn durch diesen Cluster.

Definition 1 (globale Interessantheit):

Für einen Cluster $C_i = (O_i, S_i)$: $interest_{global}(C_i) = \frac{|O_i|}{\left| \bigcup_{C_j \in Conceptgroup(C_i)} [O_j \cap O_i] \right|}$
wobei: $Conceptgroup(C_i) = \{C_j \in results \mid |S_j \cap S_i| \geq \beta * |S_i|\}$

In der Menge $Conceptgroup(C_i)$ befinden sich alle Cluster C_j , deren Teilraum S_j sich zu einem gewissen Prozentsatz β mit dem Teilraum S_i des betrachteten Clusters C_i deckt. Bei Kenntnis aller der in $Conceptgroup(C_i)$ enthaltenen Cluster wird die globale Interessantheit $interest_{global}(C_i)$ des Clusters C_i durch das Verhältnis der Objektanzahl O_i mit der Anzahl der davon bereits durch Cluster in $Conceptgroup(C_i)$ entdeckten Objekte bestimmt. Auf diese Weise werden Cluster, die bereits vorhandene Cluster in Unterräumen erneut finden, aus der Ergebnismenge ausgeschlossen. Zu Ergebnisklustern außerhalb der $Conceptgroup(C_i)$ -Menge liefert der Cluster C_i allein durch seine neue Sicht einen Informationsgewinn.

2.2 Algorithmische Herausforderungen

Es kann gezeigt werden, dass die Suche nach der im Modell definierten Ergebnismenge aus der Menge aller Cluster NP-schwer ist, sodass ein effizienter Algorithmus nur für Approximationen dieser Zielvorgabe realisiert werden kann.

Neben diesem Optimierungsproblem, bleibt zusätzlich das Problem der exponentiell vielen Teilräume, welchem sich auch die Subspace Cluster Verfahren gegenübersehen, bestehen. In allen Teilräumen können sich potentielle Cluster befinden, sodass diese ohne geeignetes Ausschlusskriterium ausnahmslos zu untersuchen wären.

Für eine Optimierung des Laufzeitverhaltens, wird für die algorithmische Lösung ein Greedy-Ansatz gewählt, der nach heuristischen Kriterien, wie z.B. Entropie-basierter Raumanalyse [CFZ99] und Orthogonalitäts-basierter Raumanalyse, nur die Teilräume untersucht, die mit hoher Wahrscheinlichkeit interessante Cluster, gemäß der Definition liefern. Meine Arbeit ermöglicht somit eine effiziente Berechnung orthogonaler Konzepte basierend auf einem global optimierenden Redundanzmodell.

Literatur

- [AGGR98] Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopulos und Prabhakar Raghavan. Automatic subspace clustering of high dimensional data for data mining applications. In *SIGMOD '98: Proc. of the 1998 ACM SIGMOD international conference on Management of data*, Seiten 94–105, New York, NY, USA, 1998. ACM.
- [AKMS08] I. Assent, R. Krieger, E. Müller und T. Seidl. INSCY: Indexing Subspace Clusters with In-Process-Removal of Redundancy. In *Proc. IEEE International Conference on Data Mining (ICDM 2008), Pisa, Italy*, 2008.
- [AWY⁺99] Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu, Cecilia Procopiuc und Jong Soo Park. Fast algorithms for projected clustering. In *SIGMOD '99: Proc. of the 1999 ACM SIGMOD international conference on Management of data*, Seiten 61–72, New York, NY, USA, 1999. ACM.
- [CFZ99] Chun-Hung Cheng, Ada Waichee Fu und Yi Zhang. Entropy-based subspace clustering for mining numerical data. In *KDD '99: Proc. of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, Seiten 84–93, New York, NY, USA, 1999. ACM.
- [DHZS02] Chris Ding, Xiaofeng He, Hongyuan Zha und Horst D. Simon. Adaptive dimension reduction for clustering high dimensional data. In *ICDM '02: Proc. of the 2002 IEEE International Conference on Data Mining (ICDM'02)*, Seite 147, Washington, DC, USA, 2002. IEEE Computer Society.
- [KK04] P. Kröger K. Kailing, H.-P. Kriegel. Density-connected subspace clustering for high-dimensional data. In *Proc. 4th SIAM Intl. Conf. on Data Mining*, Seiten 246–257, 2004.
- [Mac67] J. B. MacQueen. Some methods of classification and analysis of multivariate observations. In *Proc. of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Seiten 281–297, 1967.
- [MEX96] J. Sander M. Ester, H.-P. Kriegel und X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining*, Seiten 226–231, 1996.
- [YC07] Jennifer G. Dy Ying Cui, Xiaoli Z. Fern. Non-redundant Multi-view Clustering via Orthogonalization. In *ICDM '07: Proc. of the 2007 Seventh IEEE International Conference on Data Mining*, Seiten 133–142. IEEE Computer Society, 2007.

goolap.info: Navigating aggregated Textual Information from the Web with Schema Discovery

Ralf Heyde, Steffen Lutter

Technische Universität Berlin
Department of Software Engineering and Theoretical Computer Science
Database Systems and Information Management Group

blnkoo@cs.tu-berlin.de
slutter@cs.tu-berlin.de

Supervisors: Alexander Löser and Volker Markl

Abstract: One of today’s biggest problems in the web is finding information without browsing and reading dozens of documents. The “GOOLAP” project solves this problem by recognizing two different search intentions aka “show all extracted facts about entity x” and “list all entities with attribute y”. We describe our prototype including fact extraction and fact aggregation and fact navigation.

1 Introduction

The World Wide Web is today one of the largest databases. Lots of information is provided as unstructured text in e.g. news feeds. Most of the content is user generated content and hardly to aggregate or control. Today’s search engines are able to find keywords, but the user still has to review the data sources to find the information he searches for. The main idea of the “GOOLAP” project is to recognize the search intention of a user and to provide the information in real time from an aggregation of web data sources of different kinds. We focus on simple intentions like “show me all information about the entity ‘Barack Obama’ ” or list members of the class “president”. For solving these queries our GOOLAP.info prototype addresses three main problems: Extract entities and relations from a large corpus of a news-data¹, recognize the search intentions from the user and display aggregated extracted data in a simple, fast and intuitive fashion.

¹ 11000 news documents from <http://en.wikinews.org> covering the years between 2004 and 2008

2 System Architecture and Analysis Process

The GOOLAP architecture (Figure 1) consists of three main parts: The extraction, the data management and the web frontend.

1. **The extraction part** focuses on the analysis of the text. Its base extraction component extracts information like Part-Of-Speech (e.g. verb, noun), named entities (such as persons, places, organizations) based on Stanford NER/POS² and LingPipe. Following [4], relation extraction, such as "works for" or "located at" is realized via grammatical patterns e.g., noun groups followed by verbs and articles. The extraction part is implemented via UIMA³, a framework for the analysis of unstructured information and is enriched by the opencalais.org extraction framework.
2. **The data management, data analysis and data aggregation part** is responsible for the aggregation and schema generation. In text, often one sentence only incorporates a binary relationship. This part aggregates all binary relationships for a single entity. E.g., all attributes for a person are aggregated. For aggregating extracted data we rely on object identifiers from the service "opencalais.org". In spirit to [1,2,3] we provide two types of additional lineage information for each extracted fact: First, each fact is related to the text document and the position in the text document. Hence a user can browse through facts and documents and prove the correctness and source of the fact by examining the correlated document. Second, we aggregate identical facts and their distribution over all facts of the same person. E.g., the fact that the person "George W. Bush" was a president is more interesting for the authors in our corpus than the fact that he was a governor.
3. **The web frontend is the human computer interface:** It receives the user queries and displays the generated schemas. Often a user is unfamiliar with the analysis schema provided by an extractor or with possible extracted instances. Therefore we use auto completion and dynamically visualize all relationships for a particular entity.

² <http://nlp.stanford.edu/software/>

³ <http://incubator.apache.org/uima/>

Search and analysis process. A user poses his search query in the search box. In the current version the result schema is based on the extracted facts of the text corpus and generated in an offline processing step. We use Lucene⁴ to find documents containing the given keywords and to provide document identifiers. Based on the search query at he set of available documents is projected to documents matching the query. For these documents the extraction framework provides a set of extracted binary relationships. In a next step we aggregate all relationships for one entity and for one entity type and visualize extracted relationships and documents. Please note that while we here use a static corpus, GOOLAP can operate on data sources of different kinds, e.g. Google News, CNN or Yahoo News.

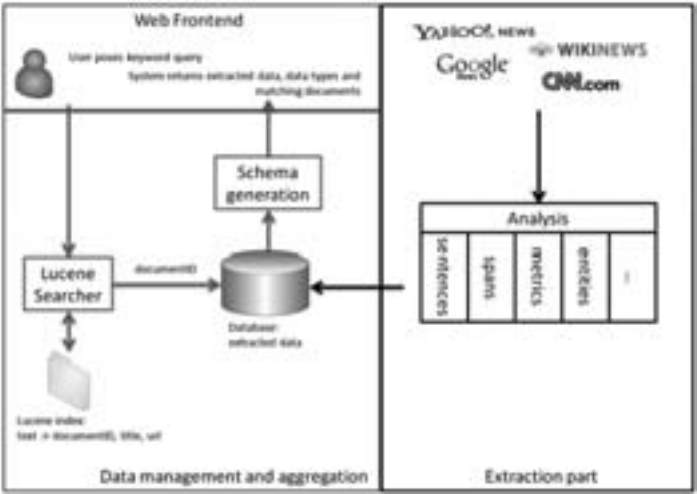


Figure 1: Architecture and Process

3 Demonstration

The result page in Figure 2 shows information about the entity “George W. Bush” and source documents found in the wiki-news corpora, which captures user-generated news between 2004 and 2008. A search starts by either typing in a named entity (e.g., “George W. Bush”) or a class of entities (e.g., president). GOOLAP recognizes these two intentions, finds and lists information which was intended by the search query and shows the correlating documents for each fact. To gather additional lineage information one might click on a link “show more” next to a fact to highlight correlating documents. To navigate the schema of extracted data, in a second iteration the user may click on a shown fact (e.g., that Laura Bush is the wife of George W. Bush). A result the entity “Laura Bush” becomes the focus of the goolap.info application and all information about this entity is shown. Thus, goolap.info allows the user to explore and navigate in a large and often unfamiliar schema, which has been created by the extraction framework.

⁴ <http://lucene.apache.org>



Figure 2: The search query "George W. Bush" returns aggregated information for this person. On the left side aggregated data and on the right side documents referring to the facts is shown.

4 Future Work

The current prototype is limited to the fixed data set of Wiki News contributions between 2004 and 2008, which have been crawled, extracted and aggregated a-priori. In our future work, the result will be generated ad-hoc by extracting and aggregating a large sample of news information e.g., top-k results from a web search engine. However, cost intensive tasks, such as crawling, extraction and aggregation, need to be parallelized. Therefore our near-time vision [5] is to leverage massive parallel cloud computing e.g., distributing these tasks on the Hadoop platform. We are currently investigating declarative query languages on top of Hadoop, such as JAQL⁵, which are capable to translate a declarative data flow specification into a functional map-reduce program. Furthermore we are currently working on incorporating machine learning techniques and large corpora for structured data, such as Freebase⁶, into our relationship extraction framework.

Acknowledgements: GOOLAP.info is a project of ten students at the Technische Universität Berlin: Martin Bach, Konstantin Domoratsky, Ralf Heyde, Marc Hofer, Robert Kubiak, Steffen Lutter, Teng Qiu, Michael Saecker, Martin Schröder and Mareike Strüwing. The project took place under supervision of Alexander Löser and Volker Markl at the Database Systems and Information Management Group of the TU-Berlin.

⁵ <http://jaql.org/>

⁶ <http://www.freebase.com>

5 References

- [1] Tao Cheng, Xifeng Yan, Kevin Chen-Chuan Chang: EntityRank: Searching Entities Directly and Holistically. VLDB 2007: 387-398
- [2] Gjergji Kasneci, Fabian M. Suchanek, Georgiana Ifrim, Maya Ramanath, Gerhard Weikum: NAGA: Searching and Ranking Knowledge. ICDE 2008: 953-962
- [3] Juan Manuel Pérez, Rafael Berlanga Llavori, María José Aramburu Cabo, Torben Bach Pedersen: R-Cubes: OLAP Cubes Contextualized with Documents. ICDE 2007: 1477-1478
- [4] Oren Etzioni, Michele Banko, Stephen Soderland, Daniel S. Weld: Open information extraction from the web. Commun. ACM 51(12): 68-74 (2008)
- [5] Alexander Löser, Fabian Hueske, Volker Markl: Situational Business Intelligence. BIRTE workshop collocated with VLDB 2008.

Towards Providing Complete Knowledge in Constraint-based Database Caching

Martin Tritschler, Volker Hudlet

Databases and Information Systems Group
Dept. of Computer Science, University of Kaiserslautern
P.O.Box 3049, 67653 Kaiserslautern, Germany
{m_trit, v_hudlet}@informatik.uni-kl.de

Form of work: study project

Advisor: Prof. Dr. Theo Härder

GI department: Database and Information Systems

Abstract: Database caching accelerates access to frequently requested data of a remote database (backend) by keeping subsets of records close to applications. Thereby, the constraint-based approach primarily uses constraints to determine what predicate extensions are complete and which queries can be answered. Equipping the backend with so-called *complete knowledge* about the caches' configurations enables it to make sufficient predictions about the contents stored in the caches. Furthermore, complete knowledge obviates the unnecessarily high communication costs caused by the currently supported simpler *structural knowledge* during update propagation. In addition, with complete knowledge it becomes possible to prepare the loading of contents on backend side. This paper describes the basic requirements and proposes an index structure that enables complete knowledge on the backend.

1 Motivation

Database caching accelerates access to data of a remote database system by providing frequently requested data close to applications. Using constraint-based database caching (CbDBC), a cache stores sets of records which represent predicate extensions whose consistency is controlled and preserved by constraints. This enables a dynamic adaptation of the cache contents to the evolving query workload, a behavior that cannot be assured by approaches based on full-table caching or materialized views [LGZ04]. In CbDBC one or more caches are connected to the primary copy which is managed by the backend. It performs all database updates and propagates the needed changes to the cache instances. Correct and efficient change propagation is an essential prerequisite for implementing a global transaction control.

Without further assistance the backend cannot predict the subset of caches affected by a certain change operation. Therefore, it is unable to avoid propagation to all caches, which in turn provokes an unnecessary communication overhead. In order to achieve a

target-oriented propagation the backend needs to maintain information about the state of its connected caches. We will show in section 3 that by providing the backend with *complete knowledge* (metadata about the cache which allows to reconstruct its exact contents) it is enabled to perform a precise preselection of change propagations, which leads to a reduced communication load. However, to keep this metadata consistent with the caches' loading and unloading behavior, it has to be accessed and modified very frequently. Therefore, we will propose an index structure in section 4 which enables efficient look-up and maintenance of this metadata. The following section will explain the required basics of CbDBC.

2 Principles of Constraint-based Database Caching

A CbDBC system caches records of predicate extensions from backend tables in related, equally structured cache tables. A backend table T has a cache table T_C assigned, which only contains a logical subset of its data. Furthermore, the cache manages two kinds of constraints: *filling columns* (FC) and *referential cache constraints* (RCC). Both constraints are defined using the concept of *value completeness*: The value v in a column $T_C.a$ is *value complete* if and only if all records of the related backend table T containing v are present in the cache table T_C [HB07]. The set of all cache tables and constraints is called *cache group*.

An RCC $S_C.s \rightarrow T_C.t$ defines a constraint in a cache group between two cache table columns, where the column $S_C.s$ is called source column and the column $T_C.t$ is the target column. An RCC is satisfied if and only if all values v in the source column are value complete in the target column. So a value v appearing in the source column implies value completeness of v in the target column [Bra08].

Filling columns control the process of loading values into the cache. Given a value v , which is a member of the set of loadable values K , a query referencing this value in a filling column $T_C.f$ triggers the loading of v to make it value complete in $T_C.f$. If $T_C.f$ has outgoing RCCs, the loading of records needs to be recursively continued to preserve all affected constraints [BHM06, KBM09].

3 Advantage of complete knowledge

The ACCache project has implemented a middleware-based CbDBC prototype [BHM06]. For consistency control a primary-copy strategy is used in which all writing operations take place in the backend. A component called *Write Set Manager* captures all changes which occur in the backend database within a transaction and maintains information about them in a data structure called *Write Set (WS)*, which is used for the pinpointed propagation to the affected cache instances [THK09].

Since a cache only stores a subset of the backend data, propagation of all changes to all

caches (and consequently to unaffected caches) would cause substantial, but unnecessary communications. If the backend just knows the definition of each cache group (*structural knowledge*), it can only perform a coarse-grained selection of records. For both update and delete operations, false positives can occur whenever a certain cache contains an affected table, but no affected records for it. To enable a more fine-grained and accurate preselection, the backend needs additional knowledge about the records stored in each cache. To provide efficient look-up of these extended metadata (*complete knowledge*), the backend maintains an index structure which can deliver for every modified record an indication of the caches storing it. Thus, for a given WS, all affected caches can be detected.

Moreover, the additional task of handling the loading and unloading behavior of caches [KBM09] can benefit from this index structure. The index structure can be used to determine the set of records which have to be loaded or unloaded. Given a certain RCC and its source column, it is easy to identify the records that have to be present in the target column due to value completeness. Therefore, RCC metadata has to be included in the index structure, which will be shown in the remainder of this paper.

4 Implementation of complete knowledge

As seen in the previous section, complete knowledge is used to identify the exact set of caches which are affected by the data manipulation operations contained in a WS. This task is supported by a hash-based index structure, which allows fast and direct access to the requested metadata. In order to avoid duplication of records, the implementation only indexes row identifiers (RowIDs) provided by the underlying DBMS. For each table, a separate index is maintained.

Figure 1a gives an example where, for the given table name CUSTOMER and the given RowID 3 (both are provided by the WS), all caches containing this record are identified. In this example, cache 1 and cache 2 are affected and so the change information will only be propagated to these two caches instead of all. The determination of the record set to

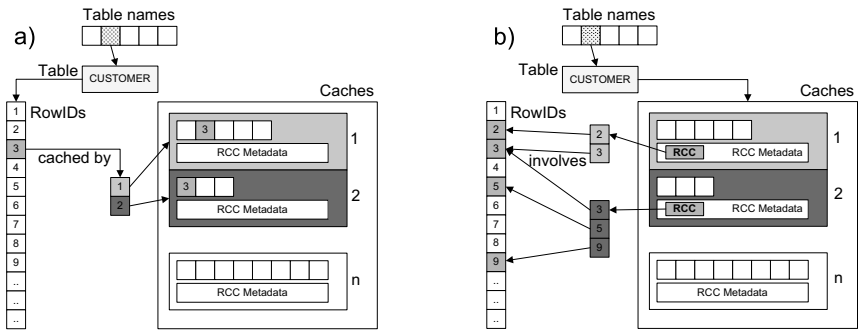


Figure 1: Sample situations of index use: a) identification of caches affected; b) determination of an RCC-based record set

be loaded in order to satisfy the RCC requirements—the second task mentioned in the previous section—uses the same index structure in a reverse way. Figure 1b illustrates this scenario: For the given table name CUSTOMER and cache 1, all RowIDs, which are implied by an RCC, are found.

A particular challenge is to keep this structure consistent with the cache contents. Whenever a cache instance unloads a subset of records, this information needs to be propagated to the backend. In this case of unloading, the adjustment of the index can be handled in a lazy way, because a temporary inconsistency only results in an unnecessary communication overhead as mentioned before. In the case of loading, however, this adjustment has to be applied immediately, because the backend must ensure that all changes needed in a specific cache instance are delivered in time to guarantee correct cache-evaluated query results.

5 Conclusion

As we have shown, complete knowledge avoids unnecessary communication. The presented index structure allows the identification of the exact set of affected caches for each change recorded in a WS. By using this index, the change propagation can be performed in an efficient way, which is an important basis for implementing a global transaction control mechanism. Additionally, the index structure helps to select the exact subset of records that needs to be loaded into the cache at the backend side. A mechanism called *prepared loading* [KBM09] can be implemented more efficiently.

It still has to be explored whether the additional administrative overhead caused by the index maintenance is justified by the saved communication overhead. Therefore, a performance analysis evaluating the benefits of different knowledge levels at the backend side will be subject of future research.

References

- [BHM06] Andreas Bühmann, Theo Härder, and Christian Merker. A Middleware-Based Approach to Database Caching. In *Proc. ADBIS*, LNCS 4152, pages 182–199, Thessaloniki, 2006.
- [Bra08] Susanne Braun. Implementierung und Analyse von Synchronisationsverfahren für das CbDBC. Master’s thesis, Technische Universität Kaiserslautern, 2008.
- [HB07] Theo Härder and Andreas Bühmann. Value Complete, Column Complete, Predicate Complete – Magic Words Driving the Design of Cache Groups. *VLDB Journal*, 17(2):805–826, 2007.
- [KBM09] Joachim Klein, Susanne Braun, and Gustavo Machado. Selektives Laden und Entladen von Prädikatsextensionen beim Constraint-basierten Datenbank-Caching. In *Proc. BTW*, 2009.
- [LGZ04] Per-Åke Larson, Jonathan Goldstein, and Jingren Zhou. MTCache: Transparent Mid-Tier Database Caching in SQL Server. In *Proc. ICDE*, pages 177–189. IEEE, 2004.
- [THK09] Martin Tritschler, Volker Hudlet, and Joachim Klein. Aspekte der Konsistenzsicherung beim Constraint-basierten Datenbank-Caching. In *Proc. BTW, Stud.-Prog.*, 2009.

Dynamische Konfiguration von Cache Groups beim Constraint-basierten Datenbank-Caching

Harald Wonneberger, Joachim Klein

AG Datenbanken und Informationssysteme, Fachbereich Informatik
Technische Universität Kaiserslautern
Postfach 3049, 67653 Kaiserslautern
{h_wonne, jklein}@informatik.uni-kl.de

Art der Arbeit: Studienarbeit

Betreuer: Prof. Dr. Theo Härder

GI-Fachbereich: Datenbanken und Informationssysteme

Abstract: Beim Datenbank-Caching werden Teilmengen relationaler Daten eines zentralen Datenbankservers in der Nähe von Anwendungen vorgehalten. Das Constraint-basierte Datenbank-Caching (CbDBC) hält dabei in den Cache-Instanzen Satzmenge von Prädikaten vor, deren Vollständigkeit mithilfe von Constraints kontrolliert und eingehalten wird. Hierbei spielt die Konfiguration einer Cache-Instanz eine wichtige Rolle für die Leistung des Gesamtsystems. In dieser Arbeit wird gezeigt, wie die bisher statische und manuell vorzunehmende Konfiguration der Cache-Instanzen dynamisch während der Laufzeit angepasst werden kann.

1 Motivation

Das Datenbank-Caching versucht den Zugriff auf meist weit entfernte Daten zu beschleunigen, wobei häufig zugegriffene Datenelemente in der Nähe von Anwendungen bereitgestellt werden. Während Konkurrenzansätze auf der Verwendung materialisierter Sichten [LGZ04] basieren, hält das Constraint-basierte Datenbank-Caching (CbDBC) Satzmenge für einzelne Prädikate vor, deren Vollständigkeit mithilfe von Constraints kontrolliert und eingehalten wird. Somit können Cache-Inhalte dynamisch an die auftretende Anfragelast angepasst werden [BHM06], was durch das Vorhalten kompletter Tabelleninhalte (Full-Table-Caching [Ora08]) nicht gewährleistet werden kann.

Während Cache-Inhalte bereits dynamisch geladen und entladen werden, muss die Cache-Konfiguration manuell vorgenommen werden. Eine dynamische Anpassung der Konfiguration würde auf sich ändernde Anfrageschemata ohne manuellen Eingriff reagieren und dadurch versuchen, die Leistung des Gesamtsystems zu verbessern. Ziel ist es daher, das Gesamtsystem zu beobachten, um so gezielt Entscheidungen über eine Anpassung der Cache-Konfiguration treffen zu können. Da im CbDBC eine der wichtigsten Anpassungen das Hinzufügen und Entfernen der dort verwendeten referentiellen Cache Constraints

(RCCs) ist, beschränken wir uns in dieser Arbeit darauf, die Überlegungen hierzu darzustellen. Die theoretischen Grundlagen dafür werden im nächsten Kapitel kurz beschrieben.

2 Grundlagen

Beim CbDBC werden die Konfigurationen innerhalb einer sog. *Cache Group* abgelegt. Diese beschreibt die Menge einzuhaltender Constraints sowie die Cache-Tabellen. Die Definition einer Cache-Tabelle T_C entspricht dabei der ihr zugeordneten Backend-Tabelle T bis auf Fremdschlüsseldefinitionen, die nicht übernommen werden. Bisher werden zwei verschiedene Constraint-Typen unterschieden: der *referentielle Cache Constraint (RCC)* und die *Füllspalte (filling column, FC)*. Über einen RCC (wie z. B. $T_C.b \rightarrow S_C.d$, vgl. Abb. 1) wird entschieden, welche Sätze in der jeweiligen Zieltabelle (hier S_C) vollständig vorzuhalten sind. Mit Hilfe von FCs wird entschieden, wann Sätze in den Cache zu laden sind. Dies ist immer dann der Fall, wenn z. B. ein bisher nicht vorhandener Wert v in einer FC $T_C.a$ durch eine Anfrage explizit referenziert wird. Ist v Element einer der Füllspalte zugeordneten Kandidatenmenge, werden die Sätze $\sigma_{a=v}T$ *vollständig* in den Cache geladen. Ein entsprechender Ladevorgang setzt sich rekursiv über ausgehende RCCs fort. Hierbei garantieren diese, dass alle geladenen RCC-Quellspaltenwerte *wertvollständig* in der RCC-Zielspalte vorgehalten werden (vgl. Abb. 1) [HB07]. Durch die Vollständigkeit der Prädikatextensionen kann so entschieden werden, welche Anfragen durch den Cache beantwortbar sind.

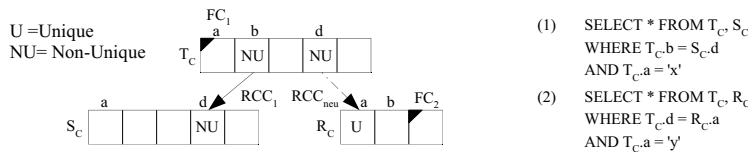


Abbildung 1: Schematische Darstellung einer Cache-Group-Konfiguration

3 Hinzufügen und Entfernen von RCCs

Hinzufügen von RCCs. In einem ersten Beispiel soll auf einer Cache-Instanz mit der bestehenden Cache Group aus Abb. 1 die Anfrage (1) ausgeführt werden. Dies ist ohne Probleme möglich, wenn der Wert x durch die FC $T_C.a$ vorgehalten wird. Der vorzunehmende Gleichverbund $T_C.b = S_C.d$ wird dabei durch RCC_1 ermöglicht. Wird durch mehrfaches Anfordern der Anfrage (2) aus Abb. 1 ein Schwellwert überschritten, so könnte man entschieden, einen neuen RCC (hier: RCC_{neu}) zur Cache Group hinzuzufügen, um so die Anfragebeantwortung zu beschleunigen. RCC_{neu} erweitert die Menge der Constraints und erzwingt so die Wertvollständigkeit in $R_C.a$ für die in $T_C.d$ enthaltenen RCC-Quellspaltenwerte. Diese muss jedoch zunächst durch das Nachladen fehlender Sätze her-

gestellt werden, bevor RCC_{neu} für die Anfragebeantwortung sichtbar wird. Der Ladeprozess startet dabei nicht wie bisher ausgehend von einer FC, sondern wird in einer RCC-Zielspalte für die fehlenden Quellspaltenwerte ausgelöst. Um initial zu entscheiden, welche RCC-Quellspaltenwerte wertvollständig nachzuladen sind, kann evtl. vorhandenes Wissen genutzt werden. Die nachfolgenden Überlegungen lassen auch auf den Fall übertragen, dass durch das Einfügen eines neuen RCCs Zyklen entstehen, wobei die Struktur der sog. *atomaren Zonen* entsprechend anzupassen ist [HB07]:

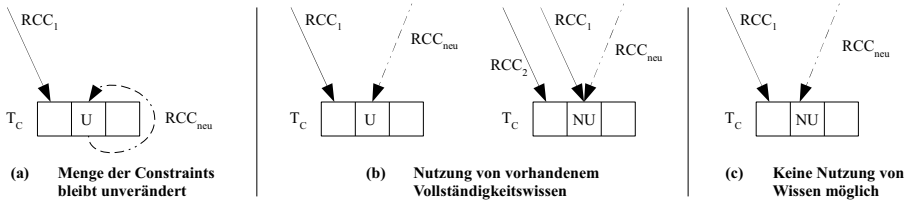


Abbildung 2: Nutzung von bestehendem Vollständigkeitswissen zur Initiierung des Nachladevorgangs

- Die Einfügung eines redundanten RCC [BH07] (wie RCC_{neu} in Abb. 2a) wird ausgeschlossen, da hierbei kein neues Vollständigkeitswissen entsteht. Somit muss auch kein Nachladevorgang ausgelöst werden, um Wertvollständigkeit zu garantieren.
- Ist die Zielspalte eines neu hinzugefügten RCC als Unique (U) markiert (RCC_{neu} in Abb. 2a), sind bereits alle vorhandenen Werte implizit wertvollständig, sodass nur noch nicht vorhandene RCC-Quellspaltenwerte nachzuladen sind. Bei Non-Unique-Spalten (NU) (vgl. Abb. 2b) lässt sich die Wertvollständigkeit nur über bereits vorhandene RCCs, die auf dieselbe Spalte verweisen, ableiten. RCC-Quellspaltenwerte, die über entsprechende RCCs bereits geladen sind, müssen somit nicht nochmals geladen werden.
- Ist die Zielspalte eines hinzugefügten RCC eine NU-Spalte und kein anderer RCC verweist ebenfalls auf dessen Zielspalte (wie bei RCC_{neu} in Abb. 2c), so kann kein bestehendes Vollständigkeitswissen genutzt werden, da nicht auf die Wertvollständigkeit einiger vorhandener Werte geschlossen werden kann. Daher müssen sämtliche Werte der Quellspalte des neu eingefügten RCC wertvollständig nachgeladen werden.

Entfernen von RCCs. Unterschreitet die Häufigkeit der Anfrage (2) aus Abb. 1 einen zuvor zu definierenden Schwellwert, so kann der RCC_{neu} wieder aus der Cache Group entfernt werden. Dadurch würden sämtliche Sätze, die durch RCC_{neu} in R_C wertvollständig geladen wurden, ihre Existenzberechtigung verlieren und müssten entladen werden, sofern kein anderer RCC ihre Existenz fordert. Größere Änderungen ergeben sich, wenn RCC_1 entfernt wird, da dann die Cache-Tabelle S_C von keinem weiteren RCC mehr referenziert wird und daher komplett entfernt werden müsste (vgl. Abb. 1). Dieser Vorgang setzt sich rekursiv durch die Cache Group fort, wodurch weitere Sätze bzw. Cache-Tabellen weggelassen können.

Sichtbarkeit. Ohne zusätzliche Maßnahmen muss ein neu hinzugefügter RCC solange invalidiert bleiben, bis alle benötigten Sätze nachgeladen sind; erst dann kann er zur Anfragebearbeitung freigegeben werden. Beim Löschen von RCCs entstehen keine Inkonsistenzen bzgl. der Anfragebearbeitung, sodass eine entsprechende Löschung sofort wirksam werden kann (vgl. Abb. 3a). Eine weiterführende Maßnahme besteht darin, nur einzelne Werte eines RCC zu invalidieren. Dies würde es ermöglichen, bereits validierte Werte als *Einstiegspunkt* zu verwenden (vgl. [HB07]). Weiterhin könnten Equi-Joins durchgeführt werden, solange ein solcher Join keine invalidierten Werte miteinander verbindet (vgl. Abb. 3b). Eine entsprechende Evaluation des Joins ist jedoch ohne spezielle Maßnahmen schwierig [Bra08].

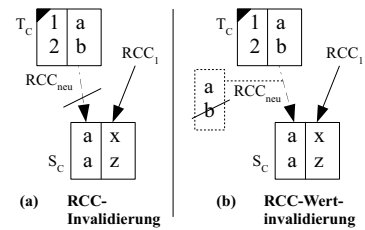


Abbildung 3: RCC Invalidierung und RCC-Wertinvalidierung

4 Fazit

In dieser Arbeit wurde dargestellt, wie eine dynamische Anpassung von Cache Groups grundsätzlich ermöglicht werden kann, wodurch sich weitere Vorteile zur Optimierung der Anfragebeantwortung bieten. Neben RCCs müssen jedoch noch detaillierte Analysen für die Anpassung von Füllspalten, Werte der Kandidatenlisten sowie Cache-Tabellen durchgeführt werden. Eine darüber hinaus wichtige Frage betrifft die Auswirkungen, die sich auf andere Komponenten des CbDBC-Systems wie z. B. der Transaktionsverwaltung ergeben. Die Freigabe von Anpassungen der Cache Group für die Anfragebeantwortung (z. B. Freischalten eines RCCs) darf die bei der Transaktionsverwaltung vorgesehene Snapshot-Isolation nicht verletzen [Bra08]. Völlig offen ist jedoch noch, anhand welcher statistisch zu ermittelnder Schwellwerte, Entscheidungen über Anpassungen an einer Cache Group getroffen werden können.

Literatur

- [BH07] Andreas Bühmann und Theo Härder. Making the Most of Cache Groups. In Kotagiri Ramamohanarao, P. Radha Krishna, Mukesh K. Mohania und Ekawit Nantajeewarawat, Hrsg., *DASFAA*, Lecture Notes in Computer Science, Seiten 349–360. Springer, 2007.
- [BHM06] Andreas Bühmann, Theo Härder und Christian Merker. A Middleware-Based Approach to Database Caching. In *Proc. ADBIS*, LNCS 4152, Seiten 182–199, Thessaloniki, 2006.
- [Bra08] Susanne Braun. Implementierung und Analyse von Synchronisationsverfahren für das CbDBC. Diplomarbeit, Technische Universität Kaiserslautern, 2008.
- [HB07] Theo Härder und Andreas Bühmann. Value Complete, Column Complete, Predicate Complete. *VLDB Journal*, 17(2):805–826, 2007.
- [LGZ04] Per-Åke Larson, Jonathan Goldstein und Jingren Zhou. MTCache: Transparent Mid-Tier Database Caching in SQL Server. In *Proc. ICDE*, Seiten 177–189. IEEE, 2004.
- [Ora08] Oracle Corporation. Internet Application Server Documentation Library, 2008.

Grouping of near duplicates in user generated forum content

Klemens Muthmann

klemens.muthmann@gmx.de

Supervisors:

Dr. Alexander Löser (TU Berlin)

Wojciech Barczynski (SAP Research)

Dr. Thomas Springer (TU Dresden)

Handed in as Diploma thesis at TU Dresden, chair of Computernetworks

1 INTRODUCTION

More than 10 gigabytes of ‘user-generated content’ is created in the World Wide Web daily [RT07]. One particular example are web-based question-answering forums. Search engines in such forums miss the ability to group threads with near-duplicate content in their results. E.g., we examined forum content for a software developer network forum where more than one million users contribute each day more than 5000 postings on approximately 60 moderated sub-channels. On a snippet of 219,000 postings we measured that on average a user will wait nearly one day until the first response (not necessarily a correct one) is posted as a follow-up. We expect that a grouping algorithm will help users to spot an existing answer in the search results immediately. Furthermore, answering users will be able to focus on unanswered threads and to create answers with high quality by searching existing threads. To solve that problem, we develop a new similarity model and grouping algorithm. We tested them on a sample of our data set. Our evaluation results show that we are able to group threads with nearly 75% precision.

2 METHODOLOGY and PROCESS

Common methods to detect duplicates on machine-generated content are near-duplication detection approaches, such as [Bro00, Cha02]. These methods are developed for automatically replicated web pages or web pages describing the same entity (e.g., a product) but presented by different agents (e.g., vendors). However, unlike replicated web-pages, forum threads with similar question or answer intentions often differ significantly in the

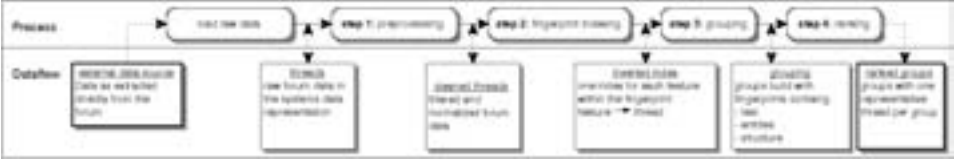


Figure 1: Process for identifying near-duplicate threads

Feature type	s
Content-based	$s_{t,e}(A, B) = \begin{cases} \frac{ fn(A) \cap fn(B) }{ fn(B) } & fn(B) \neq 0 \\ 0 & \text{otherwise} \end{cases}$
Structure-based	$s_l(A, B) = \begin{cases} 1 & fn(A) \cap fn(B) \geq 1 \\ 0 & \text{otherwise} \end{cases}$

Table 1: Similarity equations for feature types

syntax and language. In fact, in our initial experiments we could observe that these existing metrics build for ‘machine-generated duplicate web pages’ result in only a decent precision of ca. 50% and are not applicable to detect near-duplicates in human-generated forum content.

What are common reasons for near-duplicates in forum threads? Near-duplicate posts occur as symptoms of missing overview over forum content. In an initial analysis we collected reasons for the existence of duplicate posts:

- **Aggressive posting users:** These users submit the same question to several threads to increase the visibility of their question to potential expert members.
- **Thread closed to early:** If a thread owner sets a thread to ‘answered’ and suddenly notices that the answer was not as helpful as he thought, he creates a new thread with the same question.
- **‘Copy and Paste user’:** Users copying existing posts results in threads containing syntactically similar posts that often even share the same user.

Grouping and Analysis Process: Figure 1 shows the grouping process for finding near-duplicates in a community forum. It is intended as a batch process, grouping new threads and storing the groups to a database. The process consists of the following steps:

Step 0. Preliminary data loading. Raw data (H) is loaded from an external source. It incorporates unstructured forum content and additional structured meta information, e.g., semantic information, links, information about the author, etc.

Step 1. Pre-Processing cleans raw content to prepare fingerprint generation.

Step 2. Fingerprint generation produces a fingerprint for each thread. Fingerprints are created from thread content and thread structure: We incorporate domain-independent text

Experiment	Threshold	Found groups	Near duplicates	Precision	Group size		
					2	3 - 5	5+
Baseline	100% similarity	18	36	100%	18	0	0
External link	2 shared links	5	10	20%	5	0	0
Semantics	50% similarity	142	501	28,14%	91	38	13
Semantics	80% similarity	15	31	80,65%	14	1	0
Text	75% similarity	139	272	48,82%	96	29	14
Text	80% similarity	88	186	57,80%	76	9	3
All	34% similarity	126	194	74,33%	118	8	0

Table 2: Summary of experimental results

features, such as n-grams and text snippets; domain-independent external link, such as anchor texts to external help pages; forum-internal links, which are frequently used to refer a querying user to an existing postings and domain-dependent ‘semantic’ features; such as extracted products, error messages etc.

Step 3. Thread grouping identifies similar threads and combines them into a group. The similarity of two fingerprints (threads or groups) is calculated using equation 1.

$$sim(A, B) = \frac{\sum_{i=1}^n (w_i * s_i(A, B))}{\sum_{i=1}^n (w_i)} \quad (1)$$

This equation takes two fingerprints A and B and calculates their similarity as a weighted sum of the similarities of every feature type. This sum is normalized to get a value between 0 and 1. The function s calculates the similarity of two fingerprints concerning one feature type. We distinguish between two similarity methods, shown in table 1. The function fn returns the features of a fingerprinted object. The equation for text and entities is based on its Jaccard similarity [Jac01]. Two fingerprints are considered similar if they have an overlap larger than some threshold T .

Step 4. Ranking and choosing group representant. From each group we currently identify the newest thread, to represent the entire group in the search results.

3 EXPERIMENTAL RESULTS

Data Set and Evaluation Methodology. Our experiments are based on a set of more than 2000 threads. To compare different grouping approaches we follow the evaluation methodology introduced by [Hen06]. We compared the experiments based on precision and the duplicate threads identified for the entire corpus. Calculating precision required human evaluation by an expert, who labeled the pairs as correct or incorrect.

Discussion and Lessons learned. We compare a set of seven experiments. Table 2 shows the results. As base line an exact overlap between initial thread postings is calculated. Five experiments with different thresholds measure the individual contribution of content-based features, such as external links, text based features and extracted entities. Finally the

experiment ‘All’ measures the contribution of all features. We summarize our observation as follows:

- **Existing matching strategies are not sufficient.** Experiments Text 75 and 80 show the inability of common metrics for web-pages based on content features to group threads with high precision.
- **Structure-based metrics help.** The ‘all’ experiment clearly shows that adding information about the link structure within a forum or between threads of the same channel drastically increases precision while the amount of found groups nearly remains at the same level.
- **Length of the thread is important.** In our corpus more than 80% of the threads could be answered with four or less posts. We assume, that these threads have a clear intention in the question, while longer threads are often disturbed by contributions not directly fitting the topic of the thread.

Conclusion and Further Work. To the best of our knowledge, no other work solved this problem so far. In this work we identified technical factors and human factors, as reasons for duplicate posts. For spotting near-duplicate threads we used a fingerprint matching approach. We developed a feature model incorporating several feature types and developed methods to compute and compare fingerprints from these features. Our experimental study has shown that only a combination of all feature types results in high precision. In our future work we will for example apply machine learning techniques to classify postings e.g., as thank you posting, referral posting, intermediate question etc.

References

- [Bro00] Andrei Z. Broder. Identifying and Filtering Near-Duplicate Documents. In *COM 00: Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, pages 1–10, London, UK, 2000. Springer-Verlag.
- [BYRN99] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, May 1999.
- [Cha02] Moses S. Charikar. Similarity estimation techniques from rounding algorithms. In *STOC 02: Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*, pages 380–388, New York, NY, USA, 2002. ACM.
- [Hen06] Monika Henzinger. Finding near-duplicate web pages: a large-scale evaluation of algorithms. In *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 284–291, New York, NY, USA, 2006. ACM.
- [Jac01] Paul Jaccard. Distribution de la flore alpine dans le bassin des Dranses et dans quelques rgions voisines. *Bulletin del la Socit Vaudoise des Sciences Naturelles*, 37:241–272, 1901.
- [RT07] Raghu Ramakrishnan and Andrew Tomkins. Toward a PeopleWeb. *IEEE Computer*, 40(8):63–72, 2007.

Indizierung und Kompression von Schachendspieldatenbanken

Severin Neumann
severin.neumann@altmuehlnet.de

Universität Passau
Fakultät für Informatik und Mathematik

Betreuer: Prof. Martin Kreuzer (martin.kreuzer@uni-passau.de)

Art der Arbeit: Bachelorarbeit

Themengebiet: Datenbanken und Informationssysteme

Abstract: In dieser Arbeit wird ein neues Format für Schachendspieldatenbanken vorgestellt. Bei der Entwicklung wurde das Zusammenspiel der Indizierung und Kompression untersucht und optimiert. Als Kompressionsverfahren wurden Datacomp, GZip, BZip2, LZO und LZMA verwendet und verglichen.

1 Einleitung

Eine Schachendspieldatenbank speichert für jede Stellung mit fester Figurenzahl, welcher Spieler bei beidseitig fehlerfreiem Ziehen gewinnt (vgl. [Tay01, A.2]). Durch eine Indizierungsfunktion für die Stellungen können die Ergebnisse in einem *Array* linear angeordnet werden. Damit ist ein schneller wahlfreier Zugriff gegeben. Die Größe eines Eintrags ist hierbei ein Bit, wenn gespeichert wird ob der Spieler am Zug gewinnt, oder ein Byte, wenn zusätzlich angegeben wird, in wie vielen Zügen das Matt erreicht ist.

Theoretisch kann mit den Schachendspieldatenbanken perfekt gespielt werden. Jedoch setzt in der Praxis der Speicher Grenzen, da mit jeder weiteren Figur auf dem Schachbrett die Größe der Datenbanken exponentiell wächst. Aktuell gibt es Stellungsinformationen für alle Endspiele mit zwei bis sechs verbleibenden Figuren. Sie umfassen 1.153 Gigabyte. Dabei machen 99% der Datenmenge die Dateien mit sechs Figuren aus. Schon jetzt sind die Datenbanken durch Optimierungen des Indizierungsverfahrens und durch Kompression verkleinert. Dennoch gibt es weitere Optionen sie zu reduzieren. In dieser Arbeit wird gezeigt, dass die verwendeten Funktionen für Indizierung und Kompression voneinander abhängig sind und man durch Änderungen am Indizierungsverfahren die Komprimierbarkeit der Datenbanken deutlich verbessern kann. Es wird zunächst ein neues Indizierungsverfahren beschrieben und mit dem von Eugene Nalimov verglichen. Um die Größen der beiden Datenbankformate in komprimierter Form zu vergleichen, werden die Kompressoren von Datacomp [Kad92], GZip, BZip2, LZO und LZMA auf sie angewendet.

2 Indizierung von Schachendspieldatenbanken

Die einfachste Art alle Stellungen mit n verbleibenden Figuren zu indizieren ist folgende Funktion, wobei x_i die Zeile und y_i die Reihe der i . Figur ist:

$$\sum_{i=0}^{n-1} ((x_i + y_i * 8) * 64^i) \quad (1)$$

Eine Datenbank, die mit dieser Methode erstellt wurde, hat 64^n Einträge. Das heißt, wenn $n := 5$ gilt und jeder Wert in ein Byte gespeichert wird, ist die Endspieldatenbank ein Gigabyte groß.

Es gibt für diese Funktion verschiedene Optimierungen, die illegale oder redundante Stellungen entfernen (vgl. [NHH00]). Sie sind in Tabelle 1 zusammengefasst. Aus ihr ist ebenso ersichtlich, welche Verbesserungen Nalimov und das neue Verfahren verwenden.

Verbesserung	Neumann	Nalimov
Symmetrie	ja	ja
Figuren des gleichen Typs	ja	ja
Mehrfach belegte Felder	nein	teilweise
Illegale Stellungen am Ende	nein	ja
Bauern auf Reihen 2 bis 7	ja	ja
Gezieltes Entfernen von Stellungen	nein	ja

Tabelle 1: Verbesserungen für die Indizierung (siehe hierzu auch Tabelle 2 in [NHH00, S. 149])

Für die neue Methode wurden nur die Optimierungen gewählt, deren Verbesserungseffekt auch nach der Kompression noch deutlich ist. Zum Beispiel ist es nicht nötig alle illegalen Stellungen am Ende der Datei zu entfernen, da sie durch Kompression auf wenige Bytes zusammengefasst werden. Die Auswahl hat zudem den Vorteil, dass das neue Verfahren einen Index um 50% schneller als das von Nalimov berechnet.

Desweiteren hat das hier vorgestellte Format noch folgende Eigenschaften:

1. Es ist immer Weiß am Zug. Den Wert für Schwarz erhält man durch eine Spiegelung.
2. Bei einem Remis oder einer Niederlage wird keine weitere Information angegeben.
3. Für illegale Stellungen wird der Wert für Remis, wenn Weiß überlegen ist, der Wert für eine Niederlage im anderen Fall gespeichert.
4. Der Wert einer Stellung ist in einem Byte gespeichert. Der größte Wert ist 254. Sollten mehr Züge benötigt werden, wird das in einem separaten Bereich gespeichert.

Besonders die Punkte 2 und 3 tragen dazu bei, dass die Datenbanken besser komprimierbar sind. Beide erhöhen die Wahrscheinlichkeit, dass Einträge mit gleichem Wert häufiger vorkommen und sich zusammenfassen lassen. So lassen sich vor allem Datenbanken, die fast nur aus dem Wert für Verloren oder Remis bestehen, hervorragend verkleinern.

3 Kompression von Schachendspieldatenbanken

Für die Schachendspieldatenbanken von Nalimov wird das Kompressionsverfahren Datacomp verwendet. Das Verfahren ist ein Lempel-Ziv-Derivat mit einem Huffman-Kodierer. Die besondere Eigenschaft von Datacomp ist, dass es die Daten in zwei bis 64 Kilobyte große Blöcke unterteilt und über diese einen Index anlegt. Damit ist es beim Lesen eines Eintrags aus einer Datei nicht nötig alle Daten zu dekomprimieren. Soll ein anderes Verfahren verwendet werden, muss es mit ihm auch möglich sein blockweise zu kodieren. Das ist jedoch mit einer einfachen Funktion für jedes Kompressionsverfahren realisierbar. Datacomp wurde mit den Kompressionsverfahren GZip, BZip2, LZMA und LZO verglichen. Die Verfahren wurden gewählt, da sie frei verfügbar sind, eine Bibliothek für die Verwendung existiert und sie vergleichsweise schnell sind. Bei der Untersuchung spielte nicht nur die Dateigröße eine Rolle, sondern auch die Zeit welche nötig ist, um aus einer Datenbank zu lesen. Für den Vergleich der Kompressionsverfahren wurden 11 Endspieldatenbanken in unterschiedlichen Formaten und Blockgrößen komprimiert. Anschließend wurden aus einer der Datenbanken 1 Million Einträge ausgelesen. Die Daten waren dabei nicht zusammenhängend und jeder Wert wurde einzeln abgerufen. Da nur die Lesezeit gemessen wurde und die Dauer der Indexberechnung nicht berücksichtigt wurde, war hierbei irrelevant, welches Format die Datenbank hat. Tabelle 2 fasst die Ergebnisse zusammen.

	8 KB			64 KB		
	Nalimov	Neumann	μs	Nalimov	Neumann	μs
Datacomp	136	90	18,9	131	83	129,6
GZip	136	101	21,4	128	85	113,3
BZip2	138	102	158,5	126	79	1.152,6
LZO	171	131	17,0	153	106	129,9
LZMA	126	88	73,8	100	63	427,0

Tabelle 2: Größen der komprimierten Dateien für beide Formate in MB und die Zugriffsdauer in μs .

4 Ergebnis

Der Test zeigt, dass sich die Datenbanken im neuen Format um bis zu 40% besser komprimieren lassen. Damit ist das neue Verfahren nicht nur schneller in der Indexberechnung, sondern die erzeugten Datenbanken sind auch deutlich kleiner als in Nalimovs Format. Bei dem Vergleich der Kompressionsverfahren scheiden BZip2 und LZO aufgrund der hohen Zugriffsdauer beziehungsweise der schlechten Kompression aus. Für Datacomp, GZip und LZMA muss je nach Anwendungsfall entschieden werden. Ist es wichtig in kurzer Zeit viele Einträge zu lesen, so ist der Einsatz von Datacomp oder GZip mit kleinen Blöcken die beste Entscheidung. Ist Zeit jedoch nicht maßgebend und möchte man eine gute Kompression erreichen, so ist LZMA mit großen Blöcken zu wählen. Besonders im Hinblick auf Datenbanken mit sieben Figuren könnte der Wechsel hierzu sinnvoll sein.

Literatur

- [Kad92] Andrew Kadatch. Efficient algorithms for image compression. Master's thesis, Novosibirsk State University, 1992.
- [NHH00] Eugene Nalimov, Guy McC. Haworth, and Ernst A. Heinz. Space-Efficient Indexing of Endgame Tables for Chess. *ICGA Journal*, 23(3):148–162, September 2000.
- [Tay01] Aaron Tay. A guide to Endgame Tablebase.
<http://horizonchess.com/FAQ/Winboard/egt.html>,
April 2001.

Automatische Adaption von Informationsdiensten für mobile Kommunikationsgeräte

Christoph Terwelp

Janno von Stülpnagel

christoph.terwelp@rwth-aachen.de

janno.stuelpnagel@rwth-aachen.de

Abstract: Mobile Kommunikationsgeräte bieten heutzutage mehr als nur elementare Telefonfunktionen an, sondern unterstützen auch ein weites Spektrum an Informationsdiensten. Die große Vielfalt hiervon führt zu Problemen im Bereich der Softwareentwicklung, durch die bedingte Anpassungsanforderungen an die sehr unterschiedlichen Gerätetypen und die Gestaltung der Benutzerschnittstellen und Benutzerinteraktionen. Die stets wachsende Anzahl an Informationsdiensten erschwert die Auswahl der relevanten Informationen durch einen Anwender. Diese Voraussetzungen führen zu weiteren Anforderungen für die Entwicklung von Anwendungen für mobile Kommunikationsgeräte, welche die Benutzerpräferenzen bezüglich der Informationsauswahl, der Informationsdarbietung und der Anwendungssteuerung abdecken.

Die identifizierten Dimensionen für die Spezifikation dieser Anforderungen werden hier vorgestellt und drei Ansätze für die automatisierte Integration von Informationsdiensten beschrieben.

1 Problembeschreibung

Die Vielfalt von mobilen Kommunikationsgeräten und Informationsdiensten erfordert seitens des Benutzers einen aufwendigen Konfigurationsprozess, um gewünschte Informationen auf seinem Mobilgerät in der gewünschten Darbietung zu erhalten. Um diesen Prozess zu Verkürzen werden die Präferenzen des Benutzers für die Informationsauswahl und die Informationsdarbietung mit den technischen Eigenschaften seines Mobilgerätes erfasst und anschließend eine automatisierte Adaption der Informationsdienste vorgenommen. Die sechs identifizierten Dimensionen werden in Abbildung ?? dargestellt und im Folgenden weiter beschrieben.

Präferenzen für die Benutzerschnittstelle beschreiben was der Benutzer von der Anwendungsschnittstelle erwartet. Zum Beispiel, ob Informationen in einer bestimmten Art angezeigt oder ob sie vorgelesen werden, ob die Eingabe mit einem Stift auf dem Display, mit einer Tastatur oder sprachlich erfolgt. Die Spezifikation dieser Informationen wird von dem MPEG-21 Part 7 Standard [?][?] vorgenommen.

Präferenzen für Interessengebiete beschreiben die Informationen an denen der Benutzer interessiert ist. Zum Beispiel, ob er sich für Sport, Politik oder Finanzen interessiert. Des Weiteren ist es hier vorstellbar, dass Reiseinformationen von Interessen sind, wie beispielsweise die Verspätung von Zügen. Die Spezifikation dieser

Präferenzen kann ebenfalls mit Hilfe des MPEG-21 Part 7 Standards [?][?] vorgenommen werden.

Kontext des Benutzers beschreibt im Wesentlichen die aktuelle Zeit und den Aufenthaltsort des Benutzers. Hierdurch kann jedoch ein erweiterter funktionaler Kontext abgeleitet werden, wie beispielsweise: Zug, Bus, Bahnhof, Flughafen, Kino, etc. Der Benutzerkontext kann noch um zusätzliche Mobilitätsinformationen erweitert werden, beispielsweise Reisemittel, Alternativen, etc. Ergänzend zum MPEG-21 Part 7 Standard [?][?] kann der Kontext des Benutzers mit Hilfe von WURFL (Wireless Universal Resource File) [?] und UAProf (User Agent Profile) [?] spezifiziert werden.

Geräteeigenschaften beschreiben die verfügbare Hard- und Software des Kommunikationsgeräts. Wesentlich hierbei sind die Ein- und Ausgabemöglichkeiten des Geräts, wie beispielsweise die Bildschirmauflösung, die Farbtiefe, Unterstützung von sprachlicher Ein- und Ausgabe, aber auch die zusätzliche Funktionalität, welche durch Software bereitgestellt wird, wie beispielsweise Schrifterkennung oder Erkennung von sprachlichen Dialekten.

Beschreibung der Dienstschnittstelle beschreibt die Kommunikationsprozedur mit dem Dienstanbieter, insbesondere die verwendeten Datentypen und Kommunikationsmethoden. Zum Beispiel, ob der Dienst über einen Web Service oder als Web-Seite (HTML) angeboten wird (Web-Schnittstelle oder Web Service [?]). Bei Web-Schnittstellen werden die Informationen in Form von HTML-Seiten angeboten [?]. Für Web Services ist keine Benutzerschnittstelle definiert, da ihre Verwendung durch andere Dienste oder Anwendungen vorausgesetzt wurde.

Beschreibung des Dienstkontexts beschreibt die Informationen, welche über einen Dienst angeboten werden und auf die Präferenzen für die Interessensgebiete des Benutzers abgebildet werden. Beispielsweise, eine Reiseinformationsdienst, ein Hotel-Reservierungsdienst, etc.

Durch die Beschreibung der identifizierten Dimensionen kann die Benutzerschnittstelle einer Anwendung für ein Mobilgerät automatisiert an die spezifizierten Benutzerpräferenzen adaptiert werden, wobei die technischen Möglichkeiten seines mobilen Kommunikationsgeräts verwendet werden.

2 Ansätze für die automatische Adaption von Informationsdiensten

Die Adaption von Informationsdiensten an die Fähigkeiten von mobilen Kommunikationsgeräten und die Präferenzen von Benutzern kann in Abhängigkeit des Verarbeitungs-orts entsprechend einem der folgenden Ansätze oder einer Kombination davon zugeordnet werden.

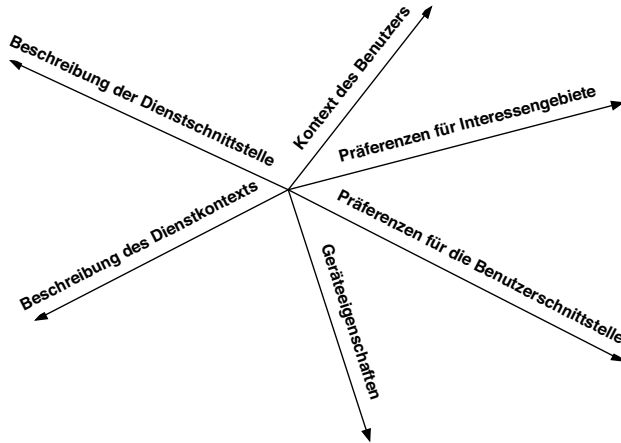


Abbildung 1: Dimensionen für die automatische Adaption von Informationsdiensten

2.1 Adaption auf dem mobilen Kommunikationsgerät

Die Anpassung der Benutzerschnittstelle und die Filterung der Informationen findet auf dem mobilen Kommunikationsgerät statt. Vorteil dieses Ansatzes ist, dass der Benutzer keine Informationen von sich an Dritte weitergibt und der Verwendungszweck der Daten verborgen bleibt. Nachteil ist dagegen die hohe Datenmengen, die an das mobile Endgerät übertragen werden müssen und von denen oft nur ein kleiner Ausschnitt verwendet wird, sowie der hohe Energieverbrauch, der durch die entstehende Datenkommunikation und die erforderliche Datenverarbeitung für die Adaption entsteht.

2.2 Adaption auf einem Proxy

Die Anpassung der Benutzerschnittstelle und die Filterung der Informationen findet auf einer Komponente (Proxy) zwischen dem mobilen Kommunikationsgerät und dem Dienstanbieter statt. Hierfür werden die Eigenschaften des mobilen Kommunikationsgeräts und die Präferenzen des Benutzers an den Proxy weitergegeben und die Adaption der Informationsdienste wird von dem Proxy durchgeführt. Der Vorteil des Ansatzes liegt in der Entlastung des mobilen Kommunikationsgerätes bezüglich Datenaufkommen und Datenverarbeitung. Nachteil ist die Preisgabe der Benutzerpräferenzen sowie ein höherer Änderungsaufwand von diesen.

2.3 Adaption durch einen Dienst

In diesem Ansatz findet die Adaption der Benutzerschnittstellen direkt im Dienst statt. Dies unterscheidet sich vom Proxy im wesentlichen dadurch, dass es sich um einen einzigen Informationsdienst handelt, der die automatisierte Adaption unterstützt. Vorteil ist auch hier die Entlastung des mobilen Kommunikationsgerätes bezüglich des Datenaufkommens und der Datenverarbeitung. Nachteil ist wie auch beim Proxy die Preisgabe der Benutzerpräferenzen. An den Dienstanbieter werden in diesem Ansatz jedoch die zusätzlichen Anforderungen gestellt, dass er eine Datenbank mit Geräteeigenschaften und Gerätefähigkeiten pflegen muss. Man kann hier zwischen *single authoring* und *multiple authoring* [?] unterscheiden. Bei *single authoring* liegt die Information nur in einem Format vor und wird dann konvertiert, während bei *multiple authoring* die Informationen in schon in verschiedenen Formaten vorliegen, von denen das passende Format verwendet wird.

Literatur

- [All] Open Mobile Alliance. WAG UAProf. <http://www.openmobilealliance.org/tech/affiliates/wap/wap-248-uaprof-20011020-a.pdf> [Accessed Januar 26, 2009].
- [Dat08] Refsnes Data. HTML Forms and Input, 2008. \verb, http://www.w3schools.com/html/html_forms.asp, [Accessed December 17, 2008].
- [HM] Rotan Hanrahan und Roland Merrick. Authoring Techniques for Device Independence. <http://www.w3.org/TR/di-atdi/> [Accessed Januar 26, 2009].
- [Pasa] Luca Passani. Introducing WALL: a Library to Multiserve Applications on the Wireless Web. <http://wurfl.sourceforge.net/java/tutorial.php> [Accessed December 17, 2008].
- [Pasb] Luca Passani. Welcome to the WURFL the Wireless Universal Resource File. <http://wurfl.sourceforge.net/index.php> [Accessed Januar 26, 2009].
- [RV02] Erhard Rahm und Gottfried Vossen, Hrsg. *Web & Datenbanken - Konzepte, Architekturen, Anwendungen*, Kapitel 10 Web Services, Seiten 293–331. Dpunkt-Verlag, Heidelberg, 2002. Beitrag von Markus Keidl, Alfons Kemper, Stefan Seltzsaam und Konrad Stocker.
- [VT03] Anthony Vetro und Christian Timmerer. Text of ISO/IEC 21000-7 FCD â Part 7: Digital Item Adaptation, 2003. \verb, <http://www.chiariglione.org/mpeg/working-documents.htm#MPEG-21>, [Accessed December 4, 2008].
- [VT05] Anthony Vetro und Christian Timmerer. Digital Item Adaptation: Overview of Standardization and Research Activities. *IEEE TRANSACTIONS ON MULTIMEDIA*, 7(3):418–426, June 2005.

Asynchronous Distributed Indexing of Large Media Collections

Diana Cheng

Rheinisch-Westfälische Technische Hochschule Aachen
Lehrstuhl für Informatik 4
diana.cheng@rwth-aachen.de

Type of Work: Master Thesis

Supervisor: Prof. Dr. Dr. h.c. Otto Spaniol

Abstract: Large sets of media assets available from existing archives or from real-time feeds must be processed, i.e. stored, analyzed, and indexed in reasonable time to allow for different searches to be performed on them. Given this sheer amount of media and the computational cost of some of the processing tasks, we seek to develop a scalable and performing architecture that is flexible and extensible, and that allows later integration of new media processing tasks. In this Master Thesis, we design, implement, and evaluate an Asynchronous Distributed Media Indexing System, and compare it with other architectural approaches in terms of the throughput, scalability, and response time that it shows when executing the tasks required for analyzing and indexing large media collections. This Master Thesis is done in cooperation with the Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) at Sankt Augustin, Germany.

1 Multimedia Analysis and Indexing

Multimedia represents the convergence of different content forms such as audio, text, images, video, etc. Its usage and application lies in various areas including advertisement, entertainment, education, business, and scientific research amongst others. Due to the increases in data storage and network bandwidth observed in the last years, these industries now own large media content archives. Multimedia documents, though rich in content, generally lack structured and descriptive metadata that would allow indexing, searching, and cross-linking (e.g. clustering) [1]. They represent a body of knowledge that can only be exploited if the appropriate means for their discovery are provided to users through, for example, multimedia search engines that facilitate the retrieval of documents from these large media content archives.

When textual annotations on media documents are not present, several content-based methods can be used to process them and extract the necessary metadata [1] in order to automate their high-level description, which would otherwise need to be done manually [2]. Once collected, the media assets' metadata should also be indexed, thus allowing for

search queries to quickly and accurately locate and retrieve the relevant data. Content processing techniques depend on the type of media being analyzed. Examples of these media analyzing techniques are:

- Text: Topic extraction, document categorization, summarization, etc.
- Video: Temporal segmentation, key frame extraction, object recognition, etc.
- Audio: Audio segmentation, speech detection, speaker's gender detection, etc.

Given the vast amount of media available, and the fact that some of the aforementioned processing tasks have proven to be computationally expensive, an efficient multimedia indexing system needs to be designed to perform the needed functions in a reasonable time, and has to be based on an architecture that is flexible and extensible, and that allows for later integration of new media processing tasks.

2 Architecture of a Multimedia Indexing System

Different architectural approaches can be chosen in order to realize a system that executes a media analyzing and indexing workflow. However, special care should be taken when designing the system's architectures, in order to ensure scalability and a reasonable overall throughput when dealing with massive-scale data volumes.

One of these approaches is the one based on a Centralized System Architecture. The advantage of such a solution is that it is easy to realize, so no special measures need to be additionally taken in order to deal with network communications. However, a key drawback of this approach is that it usually requires extremely costly central servers and its resources regarding CPU and memory, become insufficient in a short time. Centralized architectures cannot scale infinitely and, at a given point, they will not be able to meet the processing requirements of increasingly large media collections.

Distributed Systems on the other hand, offer numerous advantages including the ability to scale, i.e. the capability to keep functioning and working efficiently while growing in orders of magnitude [3]. Distributed systems, by computing things in parallel, enable larger problems to be solved without resorting to larger computers [4]. Achieving higher throughput and having the ability to continue operating correctly even when components fail are other valuable benefits.

There are several ways of architecting and implementing a distributed system. One of them is the use of Application Clustering. By distributing processing tasks to different nodes in a cluster, higher processing capacity, high availability, and load balancing can be achieved. Unfortunately, not only are cluster implementations operating system-dependant (UNIX-based), but also most of the clustering software is built around C-based messaging APIs, a fact that restricts the use of other programming languages [5].

Another approach to distributed computing is the use of Remote Procedure Call (RPC) technologies, which enable call and execution of routines on another computer, without

having to code the details for the remote interaction. However, even when these technologies provide a higher level of abstraction for the developer, the implementation of a distributed system is far from being transparent, since it is still necessary to explicitly deploy code at every node, and desired features such as load balancing need to be implemented by the developer. Moreover, most of these technologies do not allow interfacing of components developed with different programming languages.

2.1 The Need for Integration

For a number of years, the field of Multimedia Information Retrieval has been working on solving the problem of multimedia content analyzing and indexing, where both several existing components (legacy) and new techniques from current research results, often need to be integrated into the final solution. Hence, a multimedia indexing system not only faces the challenge of a providing scalability and performance, but also should allow for interoperability with legacy code and integration of new processes.

The latter is a challenge faced by several large organizations that need to integrate and link their legacy applications (often built and running on different platforms) with newly developed software, in order to share data or business rules across departments.

2.2 Service Oriented Architecture and Enterprise Service Bus

Service-Oriented Architecture (SOA) is an architectural style for designing software systems addressing the aforementioned needs of integration, data sharing, and interoperability across technologically-heterogeneous systems [6]. The typical realization of SOA uses Web service standards, with the SOAP-based XML Web service stack gaining broad support in industry due to its standardization and use of open standards [7].

SOA-based systems typically require a coordinating mechanism to control business process workflows and specify business rules. An Enterprise Service Bus (ESB) is an architectural style and middleware software that provides this functionality, decoupling clients and services via asynchronous event-based messaging. This messaging mechanism is the communication backbone between service providers and consumers and has a set of benefits including a higher level of reliability (message delivery is retried until it succeeds). Moreover, in an asynchronous model, a service client does not need to wait for the return of control from a service provider, which can avoid bottlenecks and greatly improve the overall throughput of a system.

That said, although both XML Web services and ESB address the need of application integration, they still face some challenges in terms of performance and scalability [8]. The standards that comprise them, such as XML, HTTP, together with the asynchronous event-based messaging, can entail some drawbacks when compared to the performance of other types of distributed applications, such as cluster-based and RPC-based distributed systems.

2.3 Expected scientific/technological contribution

Therefore, the purpose of this study is designing a loosely coupled, asynchronous, distributed SOA media indexing system, making use of Web services and Enterprise Service Bus technologies, and to compare it with different application distribution approaches regarding throughput, scalability, and response time when executing the tasks required for analyzing and indexing large media collections. We also aim at finding the necessary technical mechanisms to alleviate the performance issues and bottlenecks of Web services and ESB, and to provide a set of guidelines for the construction of a performing and scalable SOA-based system.

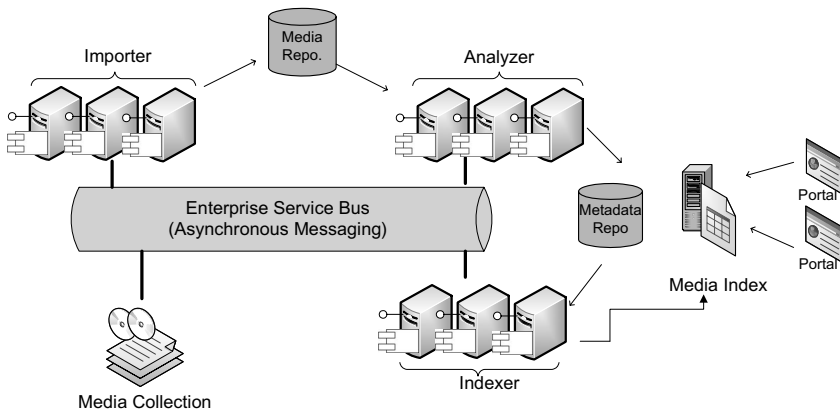


Figure 1: An asynchronous distributed Multimedia Indexing System with Web Services and ESB.

References

- [1] J. V. Thong, S. Blackwell, C. Weikart, and H. A. Mandviwala, "Multimedia content analysis and indexing: Evaluation of a distributed and scalable architecture," Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, INTERNATIONAL, August 2003.
- [2] N. Dimitrova, "Multimedia content analysis: The next wave," in CIVR, pp. 9–18, 2003.
- [3] R. Steinmetz and K. Wehrle, Peer-to-Peer Systems and Applications (Lecture Notes in Computer Science). Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.
- [4] J. Farley, Java Distributed Computing. Sebastopol, CA, USA: O'Reilly & Associates, Inc., 1998.
- [5] T. Sterling, E. Lusk, and W. Gropp, eds., Beowulf Cluster Computing with Linux. Cambridge, MA, USA: MIT Press, 2003.
- [6] T. Erl, SOA Principles of Service Design (The Prentice Hall Service-Oriented Computing Series from Thomas Erl). Upper Saddle River, NJ, USA: Prentice Hall PTR, 2007.
- [7] E. Newcomer and G. Lomow, Understanding SOA with Web Services (Independent Technology Guides). Addison-Wesley Professional, 2004.
- [8] K. Ueno and M. Tatsubori, "Early capacity testing of an enterprise service bus," in ICWS '06: Proceedings of the IEEE International Conference on Web Services, (Washington, DC, USA), pp. 709–716, IEEE Computer Society, 2006.

Versteckte Dienste in öffentlichen Umgebungen

Andre Egners

RWTH-Aachen - Lehrstuhl für Informatik IV

andre.egners@rwth-aachen.de

Abstract: Durch Anonymisierungsnetzwerke ist es möglich, anonym Dienste anzufragen und zu nutzen. Der somit gebotene Schutz bietet sich also für den Nutzer des Dienstes und schließt den Dienstanbieter nicht mit ein. Bestehende Verfahren die es ermöglichen, anonym Dienste *anzubieten* sind sehr komplex und wenig anwenderfreundlich. Im Folgenden wird ein Verfahren zum Betreiben von versteckten Diensten mit einfachen Mitteln und basierend auf offenen Standards vorgestellt.

1 Einführung

Die von Tor[RP04] bereitgestellte Möglichkeit, versteckte Dienste zu betreiben wird heute als aktueller Stand der Technik angesehen. Dienste anonym anzubieten ermöglicht es, dass der Einsatzort, hier die IP-Adresse des Dienstes, nicht öffentlich bekannt ist und bietet Schutz gegen Denial of Service-Angriffe. Bei den angebotenen Diensten kann es sich sowohl um kontroverse Dienste handeln, aber auch um solche mit sensiblen Inhalten oder um Dienste, die aufgrund von unrechtmäßiger Zensur unerreichbar sind. Die sogenannten *Hidden-Services*¹ bieten den Vorteil, nicht einfach gesperrt oder gestoppt werden zu können. Bei einem versteckten Dienst kann es sich beispielsweise um eine unabhängige Nachrichtenseite handeln die der Zensur unterliegt. Durch den verbesserten Schutz gegen Abschaltung oder Sperrung werden die Freiheitsrechte der Betreiber gestärkt. Die Komplexität des Hidden-Service Designs birgt Implikationen, die sich direkt auf die vom Nutzer wahrgenommene Qualität auswirken. Der hier vorgestellte Ansatz unterscheidet sich durch die Benutzung offener Standards und geringere Komplexität.

2 Tor Hidden-Services

Das Hidden-Service Protokoll von Tor erlaubt es, anonym einen Dienst anzubieten. Da hierbei die IP-Adresse des Dienstes vor den Clients verborgen wird, ist der Dienstanbieter nicht nur anonym, sondern kann auch gegen Denial Of Service-Angriffe geschützt werden. Das Design zeichnet sich durch Zugriffsschutz, Robustheit, Verleumdungsschutz und Anwendungstransparenz aus.

Im Folgenden wird der Protokollablauf[RL] beschrieben der nötig ist, um einen versteck-

¹dt. Versteckter Dienst

ten Dienst zu veröffentlichen und eine Anfrage entgegenzunehmen. Alle in Abbildung 1 dargestellten Verbindungen sind anonymisiert². Zuerst erstellt der Dienstanbieter ein

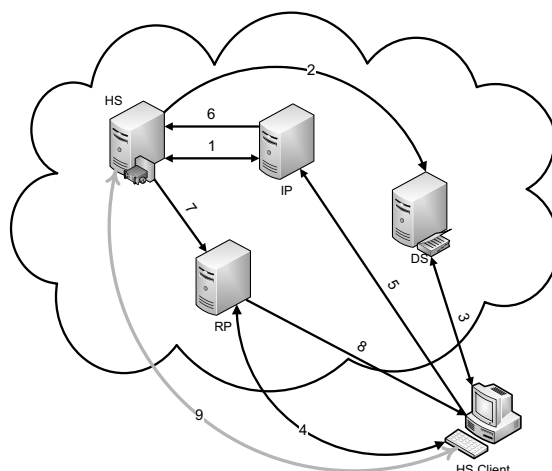


Abbildung 1: Tor-Hidden Service mit einem Introduction-Point

Schlüsselpaar mit dem er eindeutig identifizierbar ist. Außerdem wählt er zufällig einige Onion-Routern aus, die als *Introduction-Points* (IP) fungieren sollen (1). Der Dienstanbieter kontaktiert dann den *Verzeichnisserver* (DS³) und übergibt diesem den Hidden-Service Deskriptor (*öffentlicher Schlüssel* und *Introduction-Points*)(2). Möchte ein Nutzer mit dem versteckten Dienst kommunizieren, benötigt er den *Hash-Wert* des öffentlichen Schlüssels des Dienstes. Mittels dieses Hashwertes kann der Nutzer weitere Details über den versteckten Dienst von den Verzeichnisservern erfahren (3). Der Dienstnutzer erstellt nun einen Tor-Circuit zu einem zufällig ausgewählten Onion-Router. Dieser Tor-Knoten soll für ihn als *Rendezvous-Point* (RP) fungieren (4). Jetzt baut der Dienstnutzer einen weiteren Tor-Circuit zu einem der *Introduction-Points* des versteckten Dienstes auf (5). Der *Introduction-Point* leitet nun eine, vom Dienstnutzer für den versteckten Dienste verschlüsselte, Nachricht an diesen weiter. Die Nachricht enthält die Adresse des vom Dienstnutzer bestimmten *Rendezvous-Point*, einen *rendezvous-cookie* sowie den ersten Teil des Diffie-Hellman Handshakes (6). Falls der Dienstanbieter mit dem Nutzer kommunizieren möchte, erstellt er einen Tor-Circuit zum *Rendezvous-Point* und schickt den *rendezvous-cookie*, den zweiten Teil des Diffie-Hellman Handshakes und den Hashwert des gemeinsamen Sitzungsschlüssels an den *Rendezvous-Point* (7). Der *Rendezvous-Point* verbindet nun beide Streams miteinander (8). Dienstanbieter und Dienstnutzer können miteinander kommunizieren, ohne dass der *Rendezvous-Point* deren Identitäten kennt (9).

²Es handelt sich um Tor-Circuits die aus mindestens 2 Onion Routern bestehen

³directory service

2.1 Probleme

Die meisten Probleme des aktuellen Designs ergeben sich aus dessen Komplexität. In verschiedenen Studien wurden sowohl die Eigenschaften, als auch mögliche Verbesserungen des Designs untersucht. Neben der statistischen Aufbereitung angestellter Messungen von QoS Parametern[KW08], wie beispielsweise *Verbindungsaufbaudauer* und *Antwortzeit* wurden außerdem im Hinblick auf verbesserten Schutz vor DDoS-Angriffen versucht, das bestehende Design zu vereinfachen[LS06, LS07]. Diskutierte Designänderungen[Rog, KC] bringen weitere Verbesserungen, wobei viele von diesen nicht mehr ohne weiteres integriert werden können.

3 Neuer Ansatz

Der hier vorgestellte Ansatz basiert auf einem vom Lehrstuhl für Informatik 4 der RWTH Aachen entwickelten Anonymisierungsverfahren. Die Verwendung von offenen Standards und die geringe Komplexität stehen hierbei im Vordergrund. Die Kommunikation ist ähnlich wie in Tor über Proxyserver aufgebaut und verschlüsselt Nachrichten nach dem Onion-Routing[GRS96] Schema. Verbindungen werden mit dem HTTP Protokoll über SSL-fähige Proxyserver aufgebaut und tunneln Anfragen von Benutzern zum Ziel.

Um nun in diesem Kontext versteckte Dienste zu ermöglichen, wird die HTTP[RJJ⁺99] Implementierung eines Proxyservers um einen Befehl erweitert, der es erlaubt, einen weiteren Serverprozess zu starten. Dieser neue Prozess ist über einen anonymisierten Tunnel mit dem versteckten Dienst verbunden, nimmt Anfragen entgegen und leitet diese weiter. Im Gegensatz zu dem Design der versteckten Dienste in Tor kommunizieren Clients direkt über einen Kontaktpunkt mit dem versteckten Dienst. Rendezvous- und Introduction-Point werden hierbei zusammengelegt und senken somit die Komplexität des Protokolls wie auch dessen Reaktionszeit.

4 Zusammenfassung

Neben der geringeren Komplexität stellt dieser Ansatz vor allem eine schnellere Reaktionszeit des versteckten Dienstes in Aussicht. Diese vom Benutzer tatsächlich wahrgenommene Antwortzeit wurde als eine der wichtigsten Eigenschaften für die Akzeptanz und den Erfolg von versteckten Diensten identifiziert. Die Auswirkungen des einfachen Designs wie auch die Verwendung von standardisierten Protokollen müssen gründlich untersucht werden. Vor allem die Sicherheitsimplikationen müssen analysiert werden.

Literatur

- [GRS96] David M. Goldschlag, Michael G. Reed und Paul F. Syverson. Hiding Routing Information. In R. Anderson, Hrsg., *Proceedings of Information Hiding: First International Workshop*, Seiten 137–150. Springer-Verlag, LNCS 1174, May 1996.
- [KC] Karsten Loesing und Christian Wilms. Discussion of Possible Design Changes to Improve Tor Hidden Service Performance.
- [KW08] Christian Wilms Karsten Loesing, Werner Sandmann und Guido Wirtz. Performance Measurements and Statistics of Tor Hidden Services. In *Proceedings of the 2008 International Symposium on Applications and the Internet (SAINT)*, Turku, Finland, July 2008. IEEE CS Press.
- [LS06] Lasse Øverlier und Paul Syverson. Valet Services: Improving Hidden Servers with a Personal Touch. In George Danezis und Philippe Golle, Hrsg., *Proceedings of the Sixth Workshop on Privacy Enhancing Technologies (PET 2006)*, Seiten 223–244, Cambridge, UK, June 2006. Springer.
- [LS07] Lasse Øverlier und Paul Syverson. Improving Efficiency and Simplicity of Tor circuit establishment and hidden services. In Nikita Borisov und Philippe Golle, Hrsg., *Proceedings of the Seventh Workshop on Privacy Enhancing Technologies (PET 2007)*, Ottawa, Canada, June 2007. Springer.
- [RJJ⁺99] Robert Fielding, Jim Gettys, Jeff Mogul, Henrik Frystyk, Larry Masinter, Paul Leach und Tim Berners-Lee. Hypertext Transfer Protocol: HTTP/1.1. Internet Engineering Task Force: RFC 2616, June 1999.
- [RL] Roger Dingledine und Karten Loesing. Tor Rendezvous Specification. <https://tor-svn.freehaven.net/svn/tor/trunk/doc/spec/rend-spec.txt>.
- [Rog] Roger Dingledine, Karsten Loesing, Steven J. Murdoch und Christian Wilms. Design and Evaluation of Tor Hidden Service Performance Improvements.
- [RP04] Nick Mathewson Roger Dingledine und Paul Syverson. Tor: The Second-Generation Onion Router. In *Proceedings of the 13th USENIX Security Symposium*, 2004.

Ein modellgetriebener Architektur-Ansatz zur Umsetzung von verteilten Anwendungen durch ein plattformunabhängiges Modell

Stephan Kühne

Fachhochschule Brandenburg
Fachbereich Informatik und Medien
kuehnestephan@gmx.de

Art der Arbeit: Masterthesis

Betreuerin der Arbeit: Prof. Dr. rer. nat. Gabriele Schmidt

Abstract: Modernen Internetanwendungen stellen verteilte Systeme dar. Die zunehmende Nutzung der Möglichkeiten der Klientplattformen erfordert eine koordinierte Entwicklung der beiden Systemkomponenten Klient- und Serveranwendung. Beide Komponenten unterscheiden sich dabei meist durch die Verteilung einzelner Methoden. In dieser Arbeit wurde ein UML-Profil entwickelt, welches den Modellelementen die Möglichkeit gibt, diese Informationen für die verschiedenen Komponenten zu markieren.

1 Einleitung

Moderne verteilte Anwendungen nutzen immer stärker klientseitige Möglichkeiten aus. Die Entwicklung geht weg von einfachen HTML-Seiten, hin zu komplexen Internetanwendungen, die auf pluginbasierten Technologien¹ oder Javascript aufbauen. In diesen Anwendungen beruhen die beiden Komponenten (Klient- und Serveranwendung), meist auf einem gemeinsamen Entitymodell². Jede Änderung dieses Entitymodells muss in beiden UML-Modellen der Komponenten eingepflegt werden. Diese Arbeit ist besonders bei umfangreicheren Multiplattform-Systemen³, bei denen mehrere verschiedene Klientanwendungen auf einen Serverdienst zugreifen, aufwendig.

Bei diesen Anwendungen werden, in Abhängigkeit von ihren Architekturtypen, bestimmte Funktionen in der Klient- bzw. der Serverkomponente realisiert.

¹ Gemeint sind beispielsweise Anwendungen, die den Flashplayer als Browserplugin benötigen

² Mit Entitymodellen sind hier die UML-Modelle gemeint, welche die Daten und Beziehungen der Businessobjekte (auch Entitäts) beschreiben

³ Aus einem UML-Modell werden Komponenten für verschiedene Plattformen / Programmiersprachen erstellt.

Dementsprechend unterscheiden sich die Entitymodelle für beide Komponenten durch die Implementierung der Methoden, welche die Aufgabenstellung ausführen. Als Beispiel wird hier die Berechnung eines Betrages in einer dementsprechenden Software angeführt. Je nach Klientarchitektur wird diese bei einer einfachen HTML-Seite durch den Server, in einem modernen Smart-Klient⁴ dagegen direkt in der Klientanwendung umgesetzt. Aufgrund der Verteilung der Funktionen muss aber bei der Entwicklung eines solchen Gesamtsystems für jede Systemkomponente ein eigenes UML-Modell genutzt und auf einem aktuellen Stand gehalten werden.

Die modellgetriebene Entwicklung bietet für diese Arbeit viele Vorteile an. In ihrem Ablauf werden aus einem plattformunabhängigen Modell mehrere plattformspezifische Modelle oder direkt die benötigten Plattformquelltexte generiert. Es existiert aber keine Lösung, um die komponentenabhängige Verteilung von Funktionen in einem gemeinsamen UML-Modell zu berücksichtigen.

2 UML-Profil zur Verteilung

Mit Hilfe des in der Arbeit entwickelten UML-Profiles ist es möglich, die Merkmale der verteilten Funktionen in einem gemeinsamen plattformunabhängigen Modell zu hinterlegen. Aus diesem Modell können anschliessend die unterschiedlichen Komponentenmodelle generiert werden. Das UML-Profil besteht aus dem Stereotyp mit mehreren TaggedValues, wobei jeder TaggedValue ein Komponentenmodell repräsentiert. Der Stereotyp kann mit verschiedenen Modellelementen (Class Interface, u.a.) benutzt werden, wobei er aber hauptsächlich für die Methoden der Entityklassen gedacht ist. Der Datentyp der TaggedValues ist Boolean. Ein negativer Wert bedeutet für das markierte Modellelement, dass es in der zugeordneten Komponentenversion nicht enthalten ist. Da der Standardwert der TaggedValues 'true' ist, müssen nur die Elemente markiert werden, die in einer bestimmten Version nicht vorkommen. Die Anzahl der Versionen ist bei diesem Stereotyp auf die Anzahl der TaggedValues beschränkt.

In der Masterarbeit wird noch ein alternativer Stereotyp (Distributed2) vorgestellt, mit dem eine uneingeschränkte Versionsanzahl markiert werden kann. Dieser nutzt ein Array von Node⁵-Elementen als Datentyp des TaggedValues. Die Node-Elemente stammen aus dem UML-Verteilungsdiagramm und stellen jeweils eine Komponentenversion dar. Das hat den Vorteil, dass die Anzahl und Beziehungen der einzelnen Versionen im UML-Projektmodell vermerkt werden.

⁴ Klientanwendungen, die u.a. eigene Businesslogik implementieren und u.U. einen eigenen Datenbestand besitzen.

⁵ Node Elemente stellen in den Verteilungsdiagrammen die Zielkomponenten (Server, Klientversion1, Klientversion2, etc.) dar.



Abbildung 1: Die Stereotypen zum Verteilen von Modellelementen

Nach dem Markieren werden die Modelle in die einzelnen Komponentenmodelle transformiert. Aus einer markierten Klasse können so unterschiedliche Versionsklassen entstehen.

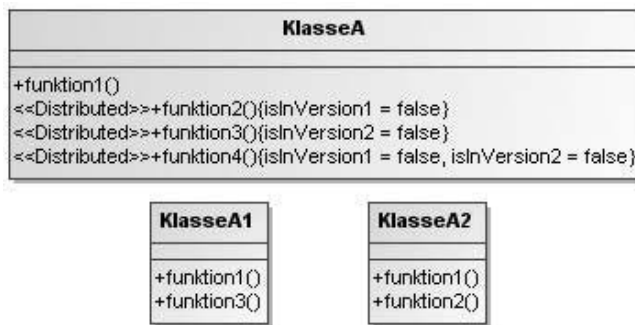


Abbildung 2: Die Ursprungsklasse und die daraus generierten Klassen für die Version 1 und Version 2

3 Zusammenfassung

Das UML-Profil ermöglicht es, die Informationen über die verteilten Modellelemente (z.B. Funktionen) in einem Modell zu erfassen. Dies ermöglicht eine zusammenhängende Entwicklung der verschiedenen Komponenten.

Die Praxistauglichkeit des Profils wurde mit Hilfe des AndroMDA⁶-Frameworks erfolgreich getestet. Dazu wurde der Quelltext von vier Komponenten erfolgreich aus einem gemeinsamen UML-Modell generiert.

Neben dem eigenen wurden auch bestehende UML-Profile (UML-Profile for Components [Gi06]; UML-Profile for Distribution [Si04] u.a.) auf Nutzbarkeit untersucht. Allerdings wurde kein Profil mit den benötigten Eigenschaften gefunden.

⁶ www.andromda.org

Literaturverzeichnis

- [Gi06] Hyun Gi Min, Soo Dong Kim, A Technique to Represent and Generate Components in MDA/PIM for Automation*, Konferenzbericht, 2006, Department of Computer Science, Soongsil University, Korea, FASE 2006 , LNCS, 3922, Seiten: 293 – 307, Springer-Verlag Berlin Heidelberg,
- [Si04] Raul Silaghi, Frédéric Fondement, Alfred Strohmeier, Towards an MDA-Oriented UML Profile for Distribution, Konferenzbericht, 2004, , Proceedings of the 8th IEEE International Enterprise Distributed Object Computing Conference, , , Seiten: 227-239, IEEE Computer Society, <http://lgl.epfl.ch/members/silaghi/publications.html>

Implementierung und Bewertung eines Lokalisierungsdienstes basierend auf aktiver RFID-Technologie

Frank Meffert

Fachhochschule Wiesbaden
Fachbereich Design Informatik Medien
Studiengang Allgemeine Informatik
`fmeff002@student.informatik.fh-wiesbaden.de`

Art der Arbeit: Bachelor-Thesis
Referent: Prof. Dr. Reinhold Kröger
Korreferent: Prof. Dr. Martin Gergeleit

1 Motivation

Heutzutage besitzen Informationen über den Aufenthaltsort oder die Bewegung eines Objektes einen hohen Stellenwert. Das Problem, Position oder Bewegung von Objekten zu ermitteln, ist die Hauptaufgabe sogenannter Lokalisierungsdienste. Sie stellen diese Informationen höheren Diensten zur Verfügung (Location-Based Services).

In vielen Bereichen der Industrie setzt man derzeit auf RFID-Technologie (Radio Frequency Identification) zur kontaktlosen Identifikation von Objekten über Funk ([Fin06]). Zu einem solchen RFID-System gehören Transponder, die zur Identifikation an Gegenständen angebracht werden und Daten über sie dezentral speichern. Diese Daten können mit RFID-Readern ausgelesen und zur Weiterverarbeitung an Rechensysteme übertragen werden. Bei aktiver RFID-Technik erfolgt die Energieversorgung eines Transponders mit einer eigenen Energiequelle, wie zum Beispiel einer Batterie. Die Entwicklung spezieller Dünnschichtbatterien ermöglicht eine Identifikation von Objekten auch über Distanzen von bis zu 100 Metern und lässt den Einsatz dieser Technik im Bereich der Lokalisierung, z.B. auf einem Werksgelände oder in einer Halle, möglich erscheinen ([ZZY06]).

Um aus rohen Sensordaten Lokationsinformationen, wertet ein Lokalisierungsdienst Messdaten verschiedener Sensoren aus, wie z.B. RSSI (Radio Signal Strength Indicator), d.h. die Signalstärke zwischen Sender und Empfänger. Dabei kommen verschiedene Methoden zum Einsatz, insbesondere auch statistische Filtermethoden ([HB04], [YZN07], [WKP06]). Diese Filter schätzen einen Systemzustand eines dynamischen Systems, hier also u.a. die Position eines Objektes, durch Wahrscheinlichkeitsfunktionen, die auf den Sensordaten basieren. Einer dieser Filter ist der Partikel-Filter ([RAG04], [Sch05]), der Systemzustän-

de durch eine große gewichtete Menge von Partikeln beschreibt. Ein Vorteil von Partikel-Filtern gegenüber anderen Filtern ist, dass er keine Annahme bzgl. der Linearität des Prozessmodells und normalverteilter Störgrößen macht.

2 Ziel der Arbeit

Diese Bachelor–Thesis entstand im Labor für Verteilte Systeme (DOPSY) an der Fachhochschule Wiesbaden in Zusammenarbeit mit IdentPro, einem Spezialunternehmen für Identifikation und Authentifikation auf Basis von RFID–Technik in St. Augustin bei Bonn. Ziel der Arbeit war es, einen Prototypen eines Lokalisierungsdienstes basierend auf RFID–Technologie unter Benutzung von Partikel–Filtern zu entwickeln und zu bewerten. Da der Einsatz von Partikel–Filtern in eingebetteten Systemen in der Praxis aufgrund ihres hohen Rechenaufwands bisher kaum möglich war, sollten Tests die Einsetzbarkeit solcher Filter auf heutigen Systemen im industriellen RFID–Bereich prüfen.

3 Architektur und Implementierung

Im Rahmen dieser Arbeit wurde eine geschichtete Architektur für einen Lokalisierungsdienst entworfen (vgl. Abb. 1), die es erlaubt, verschiedene Sensorsysteme zur Positions-schätzung anzubinden. Neben der genannten RFID–Technik können weitere Sensorsys-teme angebunden werden, deren Messdaten gemeinsam fusioniert werden. Ferner ist es möglich, mehrere Ortungsverfahren zur Verbesserung der Genauigkeit zu kombinieren. Der Lokalisierungsdienst selbst führt eine Liste über alle lokalisierten Objekte und deren geschätzte Zustandsvektoren. Diese Daten werden über ein plattformunabhängiges Client–Interface zur Verfügung gestellt und können entsprechend von höheren Diensten genutzt werden. Die Parameter des Dienstes können über eine austauschbare Konfiguration ver-ändert werden. Das vom Lokalisierungsdienst erfasste Terrain wird durch ein Geometrie–Modell beschrieben.



Abbildung 1: Architektur des Lokalisierungsdienstes

Zentraler Bestandteil der Implementierung war die Umsetzung eines Partikel-Filters als Lokalisierungsverfahren. Hier wurde zunächst ein Beobachtungsmodell implementiert, in dem geregelt wird, wie neue Sensordaten die Gewichte der Partikel verändern und somit zu einer Schätzung beitragen. Dazu wurde für die gegebene aktive RFID-Technologie eine Kennlinie als optimale Schätzfunktion für den Zusammenhang zwischen Signalstärke und Entfernung aufgestellt. Hierbei hat sich bestätigt, dass die Signalstärke etwa quadratisch mit der Distanz abnimmt.

4 Bewertung und Ausblick

Zur Bewertung wurde ein Testfeld in Form eines gleichseitigen Dreiecks mit drei RFID-Sensoren in den Ecken im Freien vermessen. Dafür wurde die Position eines RFID-Tags innerhalb des Dreiecks variiert, und die wahrgenommene Signalstärke (RSSI) des Tags aus Sicht der drei Sensoren zusammen mit der Position gespeichert. Die gemessenen Daten wurden dann als Eingabe für den Partikel-Filter verwendet. Dabei wurden die geschätzten Positionen mit den tatsächlichen verglichen. Die Auswertung der Messdaten ergab, dass im Freien in mehr als 75% aller Fälle eine Genauigkeit von 1–4 m erzielt werden kann (vgl. Abb. 2). Eine genauere Analyse des Partikel-Filters offenbarte allerdings, dass der Algorithmus sich nicht immer wie erwartet verhält, falls zu wenige Sensoren zum Einsatz kommen.

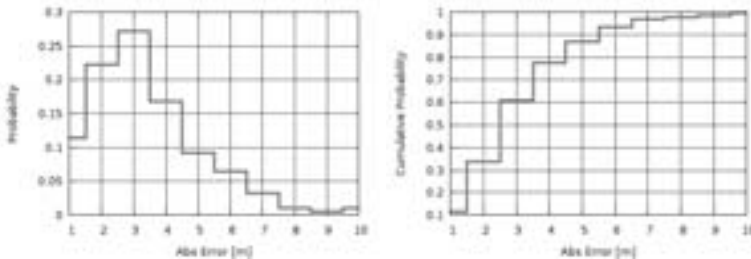


Abbildung 2: Fehlerwahrscheinlichkeit des Partikel-Filters (links), kumulativ (rechts)

In aktuellen Arbeiten wird der Lokalisierungsdienst um eine dynamische Objektverfolgung (Tracking) in Form eines Bewegungsmodells im Partikel-Filter erweitert. Weiterhin wird ermöglicht, an bekannten Positionen Transponder als Referenzpunkte zu platzieren. Mit einem Kunden von IdentPro soll ein prototypischer Einsatz evaluiert werden. In der Fachhochschule Wiesbaden ist eine Weiterentwicklung im Rahmen einer AAL-Umgebung (Ambient Assisted Living) in Vorbereitung.

Literatur

- [Fin06] Klaus Finkenzeller. *RFID Handbuch*. Carl Hanser Verlag, 2006.
- [HB04] Jeffrey Hightower und Gaetano Borriello. *Particle Filters for Location Estimation in Ubiquitous Computing: A Case Study*. 2004.
- [RAG04] Branko Ristic, Sanjeev Arulampalam und Neil Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House Publishers, 2004.
- [Sch05] Gregor Schwarzenberg. Objektverfolgung mit Partikel-Filtern. Diplomarbeit, Institut für interaktive Echtzeitsysteme, Fakultät für Informatik, Universität Karlsruhe (TH), 2005.
- [WKP06] Widyawan, Martin Klepal und Dirk Pesch. Influence of Predicted and Measured Fingerprint on the Accuracy of RSSI-based Indoor Location Systems. Bericht, Centre for Adaptive Wireless Systems Cork Institute of Technology, Ireland, 2006.
- [YZN07] Wilson M. Yeung, JunYang Zhou und Joseph K. Ng. Enhanced Fingerprint-Based Location Estimation System in Wireless LAN Environment. In *EUC Workshops*, Seiten 273–284, 2007.
- [ZZY06] Junhui Zhao, Yuqiang Zhang und Mengjie Ye. Research on the Received Signal Strength Indication Location Algorithm for RFID System. In *Communications and Information Technologies, 2006. ISCIT '06. International Symposium on*, Seiten 881–885, 2006.

Parallele Motivsuche in großen Netzwerken

Sven Ritsche

Martin-Luther-Universität Halle-Wittenberg
Institut für Informatik
ritsche@informatik.uni-halle.de

Abstract: Für verschiedene Anwendungen, die über Netzwerke modellierbar sind, besteht ein großes Interesse, kleine wiederkehrende Strukturen, sogenannte Netzwerk-Motive, aufzufinden. Ziel der Arbeit war die Entwicklung paralleler Algorithmen zur Suche nach Motiven in großen Netzwerken. Dabei stand die effiziente Ausnutzung der Speicherhierarchien der benutzten Systeme mit gemeinsamem Speicher im Vordergrund.

1 Einleitung

Als Motiv eines Netzwerks, im Folgenden auch nur Motiv genannt, bezeichnet man kleine induzierte Subgraphen, deren Auftreten sich innerhalb eines zu untersuchenden Netzwerks häufen können. Das vermehrte Auftreten eines bestimmten Motivs deutet dabei auf eine strukturelle Eigenschaft hin, die für die Analyse großer Netzwerke von Bedeutung sein kann. Zum Beispiel können komplexe biologische und chemische Vorgänge innerhalb von Organismen als sehr große Netzwerke dargestellt werden, und häufig auftretende Strukturen können besonders effiziente biologische bzw. chemische Funktionen abbilden, da sie durch evolutionäre Prozesse gebildet wurden. Verwendung findet die Motivsuche aber auch in der Technischen Informatik, um bestehende Schaltungen hinsichtlich Platz oder Zeitverbrauch zu optimieren, indem sehr oft auftretende Teilstrukturen optimiert werden.

Ein effizienter sequentieller Algorithmus zur Suche nach einem gegebenen Motiv lag mit der Arbeit [GK07] vor. Insbesondere kann dieser Algorithmus zur Lösung von Problemen aus der Klasse der Subgraph-Isomorphie-Probleme benutzt werden. Diese Klasse ist bekannterweise NP -vollständig und verdeutlicht somit die Komplexität des Problems. Einfache Beispiele für Motive in Netzwerken sind in der Abbildung 1 zu sehen.

Dieser Algorithmus nutzt die vorhandenen Automorphismen des gegebenen Motivs, um den Berechnungsaufwand zu reduzieren bzw. doppelte Zählung zu vermeiden. Mit der zunehmenden Verbreitung von *symmetric multi processor*-Systemen und Mehrkernarchitekturen steht die Frage nach einer effizienten Ausnutzung der Ressourcen. Da zur Berechnung der Signifikanz eines Motivs eine Vielzahl an strukturähnlichen Netzwerken untersucht werden muss, beschränken sich die meisten Vorschläge zur Parallelisierung auf die nebenläufige Berechnung der Motive für die verschiedenen Netzwerke. Eine parallele Suche auf einem Netzwerk kann vorteilhaft die Struktur bzw. Speicherdarstellung dieses

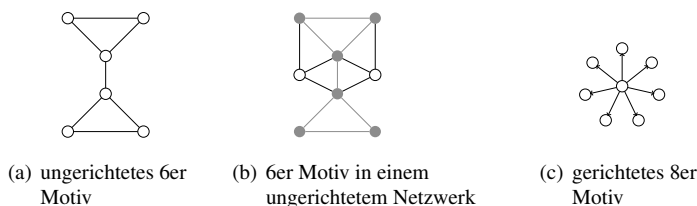


Abbildung 1: Beispiel für ein Motiv und dessen Auftreten in einem Netzwerk

Netzwerks ausnutzen. Insbesondere für *symmetric multi processor*-Systeme, die auf dem *shared memory model* [RR07] basieren, bietet es sich an, deren Architektur auszunutzen. Ziel der Arbeit war es verschiedene Parallelisierungen des Algorithmus zu entwickeln, sowie die Suche auf gerichtete Netzwerke zu erweitern.

2 Algorithmus nach [GK07]

Der als Grundlage gewählte sequentielle Algorithmus aus [GK07] ist so konzipiert, dass er für ein gegebenes Motiv alle isomorphen Motive innerhalb des zu durchsuchenden Netzwerks berechnet. Zur Reduzierung des Berechnungsaufwands werden die Symmetrien des gegebenen Motivs ausgenutzt. Dies geschieht durch Festlegung einer partiellen Ordnung auf den Knoten des Motivs, so dass die Kardinalität der Menge aller Motive, die isomorph zu dem gegebenen Motiv sind und dieser Ordnung genügen, eins beträgt.

Die Suche nach dem Motiv kann nun durch schrittweise rekursive Erweiterung eines jeden Knotens des Netzwerks durchgeführt werden. Im Detail erweitert man eine gegebene Menge von Knoten um einen ihrer Nachbarn und überprüft, ob nicht ausgeschlossen werden kann, dass diese erweiterte Menge ein isomorphes Bild des zu suchenden Motivs ist. Zur Reduzierung des Berechnungsaufwands können durch die gegebene partielle Ordnung zusätzliche Abbruchkriterien erzeugt werden, die redundante Aufzählungen der isomorphen Bilder verhindert. Außerdem können Knoten, für die bereits alle Motive berechnet wurden, aus der Menge der Knoten zur Erweiterung entfernt werden, da es ansonsten zu redundanten Aufzählungen kommt.

3 Parallelisierungsansätze

Es wurden drei Ansätze zur Parallelisierung erarbeitet. Ein trivialer Ansatz ist die nebenläufige Suche nach dem Motiv für jeden Knoten des Netzwerks (die Berechnungen der Erweiterungen für zwei Startknoten sind unabhängig voneinander). Ein theoretischer Speedup in Höhe der Anzahl der verwendeten Prozessoren stellt sich jedoch nur ein, wenn die einzelnen Motivsuchen für verschiedene Startknoten entsprechend verteilt werden. Der

Berechnungsaufwand für die einzelnen Aufgaben kann dabei stark variieren, da das zu suchende Motiv sowie die Struktur des zu durchsuchenden Graphen großen Einfluss auf die Berechnung besitzen.

Ein weiterer Ansatz ist eine 2-Phasen-Suche. Hierbei wird in einer ersten Phase das Netzwerk in mehrere kleinere Netzwerke partitioniert und für jedes dieser Teilnetzwerke eine Motivsuche gestartet, die jedoch nur Motive berechnet, die vollständig in diesem Teilnetzwerk liegen. In einer zweiten Phase wird nach all den Motiven gesucht, die sich über mehrere Teilnetzwerke erstrecken. Hauptsächliches Ziel dieser Methode ist es aber, die Speicherhierarchien heutiger Systeme auszunutzen, denn sofern die Teilnetzwerke klein genug sind, um in den Cache zu passen, könnte die Motivsuche wesentlich schneller sein. Dies ist bedingt durch die verschiedenen Zugriffszeiten auf Hauptspeicher bzw. Cache. Für die Suche nach Motiven, die sich über mehrere Partitionen erstrecken, muss jedoch eine kantenbasierte Motivsuche genutzt werden. Diese beginnt den Erweiterungsschritt mit einer zweielementigen Knotenmenge und erweitert diese nur um Knoten, die über keine partitionsübergreifende Kante verfügen, die "kleiner" als die Startkante ist.

Eine weitere Möglichkeit liegt darin, das Netzwerk auch wieder zu partitionieren. Jedoch soll nicht nur innerhalb der Partitionen gesucht werden, sondern es sollen auch Motive gefunden werden, die über die Partitions Grenzen hinweg existieren. Ziel dieser Methode ist auch wieder die Ausnutzung der Speicherhierarchien heutiger Systeme, indem man nur auf kleinen Teilnetzwerken sucht. Jedoch hofft man, nur sehr selten auf Knoten außerhalb der Partition zugreifen zu müssen, denn sofern dies nicht allzu oft passiert, kann ein Großteil der Motivsuche wieder innerhalb der Partition stattfinden. Vorteil gegenüber der vorherigen Methode ist, dass keinerlei Nachberechnungen durchzuführen sind.

4 Ergebnisse

Die erarbeiteten Ansätze wurden mit Threads (Boost-Thread-Bibliothek [BTB]) implementiert und auf einem Sun Fire X4600 M2 Server getestet. Dieser besitzt 8 Dual-Core-Prozessoren mit jeweils einer Taktfrequenz von 2.6 GHz [ASF]. Für die Implementierung der trivialen Variante zeigte sich zwar ein Geschwindigkeitsgewinn, der jedoch hinter den theoretischen Erwartungen zurück blieb. Mit hoher Wahrscheinlichkeit ist dies zurückzuführen auf die Architektur des verwendeten Testsystems und dessen Speichermanagement.

Für die 2-Phasen-Suche konnte gezeigt werden, dass die Partitionierung eines Netzwerks und die Suche innerhalb der Partitionen die Laufzeit verringert und man somit einen vorteilhaften Einfluss des Caches annehmen kann. Insbesondere im Vergleich zur trivialen Variante zeigte sich ein deutlicher Vorteil. Die Laufzeiten können sich unterschiedlich auf die zwei Phasen verteilen. Bei einem Netzwerk, das eine geringe Anzahl an partitionsübergreifenden Kanten besitzt, beschränkt sich der Großteil des Berechnungsaufwands auf die erste Phase. Es entstehen jedoch für die zweite Phase immer gewisse Grundkosten, da über alle Kanten des Netzwerks iteriert werden muss. Die dritte Variante zeigte ein noch besseres Laufzeitverhalten. Zurückzuführen war dies auf die Vermeidung der Kosten

Threads \ Partitionen	10	20	40	80	160	320
	2-Phasen-Suche					
seq.	802.45s	786.48s	776.95s	767.83s	764.08s	769.19s
2	430.33s	415.20s	413.00s	412.83s	409.62s	407.89s
4	228.57s	217.05s	223.19s	216.43s	214.80s	213.38s
8	114.35s	109.89s	109.73s	109.93s	110.03s	108.58s
Threads \ Partitionen	1-Phasen-Suche					
seq.	693.94s	691.71s	689.64s	687.39s	691.35s	699.28s
2	378.64s	366.08s	389.72s	375.86s	377.01s	369.98s
4	221.85s	195.58s	194.46s	197.20s	193.32s	201.07s
8	155.03s	117.57s	104.35s	106.92s	106.40s	104.31s

Tabelle 1: Laufzeit für Motivsuche (Abbildung 1(c)) in einem Zufallsnetzwerk mit 1000000 Knoten und 4000000 Kanten

für eine zweite Phase und der positiven Auswirkungen der Partitionierung.

Beispielhaft sei hier auf experimentelle Werte in Tabelle 1 verwiesen, welche die durchschnittlichen Laufzeiten für die Suche auf einem gerichteten partitionierten Netzwerk enthält. Das Netzwerk wurde mittels eines Zufallsgenerators erzeugt. Für die Motivsuche wurde das Motiv aus Abbildung 1(c) verwendet. Mittels dieser Werte wird deutlich, dass die Suche auf einem partitionierten Netzwerk von Vorteil sein kann, da die Laufzeiten für die Berechnungen mit steigender Anzahl an Partitionen rückläufig ist und einen Minimalwert erreicht. Insbesondere wirkt sich die 1-Phasen-Suche bei Verwendung von mehreren Prozessoren vorteilhaft aus. Jedoch bestehen mehrere Größen, die die Laufzeiten beeinflussen können. Grundsätzlich sind dies natürlich die Struktur des zu durchsuchenden Netzwerks, aber auch die Struktur des Motivs sowie die Architektur des Testsystems. Ein weiterer starker Faktor ist auch die Anzahl gefundener Bilder bzw. die mittlere Abbruchtiefe, um nur einige Beispiele zu nennen.

Für weitere Untersuchungen bietet es sich an, nacheinander nach mehreren Motiven zu suchen und somit nur einmal die Partitionierung berechnen zu müssen. Interessant ist es auch, die Algorithmen unter Berücksichtigung verschiedener Netzwerkstrukturen zu optimieren.

Literatur

- [ASF] SUN FIRE X4600 M2 SERVER ARCHITECTURE .
<http://www.sun.com/servers/x64/x4600/arch-wp.pdf>.
- [BTB] Boost Thread Bibliothek Dokumentation .
http://www.boost.org/doc/libs/1_35_0/doc/html/thread.html.
- [GK07] Joshua A. Grochow und Manolis Kellis. Network Motif Discovery Using Subgraph Enumeration and Symmetry-Breaking. In *RECOMB*, Seiten 92–106, 2007.
- [RR07] T. Rauber und G. Runger. *Parallele und Verteilte Programmierung*. Springer, 2007.

Monitoring unternehmenskritischer Anwendungen unter Verwendung modellbasierter Performance Constraints

Andreas Textor

Fachhochschule Wiesbaden
Fachbereich Design Informatik Medien
Studiengang Allgemeine Informatik

Art der Arbeit: Bachelor-Thesis

Betreuer der Arbeit: Prof. Dr. Reinhold Kröger, Dipl.-Inform. (FH) M.Sc. Marcus Thoss

1 Zielsetzung

Unternehmenskritische Anwendungen, also Anwendungen, deren reibungsloser Betrieb wichtig für den wirtschaftlichen Erfolg eines Unternehmens ist, werden zunehmend umfangreicher und komplexer. Für die Entwicklung und Wartung derartiger Systeme sind strukturierte Methoden des Software-Engineering notwendig. Anwendungsmodelle werden aus Anforderungen konstruiert und dienen als Grundlage der Entwicklung der Software. Für die Modellierung hat sich insbesondere die Unified Modeling Language (UML) der Object Management Group (OMG, [Objb]) etabliert. Arbeitsabläufe (Workflows) von Business-Prozessen werden ebenfalls durch Modelle konkretisiert, und Teile von Workflows, die in Form einer Anwendung umgesetzt werden, können durch UML-Aktivitätsdiagramme formuliert werden. Zur Sicherstellung der Performance einer kritischen Anwendung muss sie laufend überwacht werden; dies bezeichnet man als Monitoring. Für das Monitoring müssen Daten aus den verschiedenen Komponenten der Anwendung gesammelt und aufbereitet werden, um mögliche Leistungsengpässe zu finden und darauf reagieren zu können.

Ziel dieser Arbeit ist es, im Workflow-Modell einer Anwendung Regeln für die Auswertung von Monitoring-Daten der Anwendung formal definieren zu können. Die Regeln, sogenannte Constraints, erlauben es, Aussagen über die Soll-Ausführungsreihenfolge und -dauer von Abschnitten des Workflow-Modells der zu überwachenden Anwendung machen zu können. Monitoring-Daten, die die Anwendung zur Laufzeit liefert, sollen über spezielle Mechanismen bestimmten Entitäten im zugehörigen Workflow-Modell eindeutig zugeordnet werden können. Dies erlaubt es, die definierten Regeln anhand des Modells und der zugeordneten Monitoring-Daten zur Laufzeit automatisch zu überprüfen. Auf diese Weise ist es möglich, die zur Entwicklung und Weiterentwicklung einer Anwendung entworfenen Modelle auch für ihr Monitoring zu verwenden. Für das in dieser Arbeit entwickelte modellbasierte Performance Monitoring werden Performance Constraints auf

der Ebene von Workflow-Modellen, die in Form von UML-Aktivitätsdiagrammen vorliegen, mittels der Object Constraint Language (OCL, [Obj06]) festgelegt, einer deklarativen Sprache zur Formulierung von Ausdrücken und Bedingungen in objektorientierten Modellen, die Teil der UML ist. Da OCL momentan keine Möglichkeit bietet, Aussagen über zeitliche Abläufe zu treffen, wird untersucht, wie Constraints über Ausführungsreihenfolge und -dauer von Workflow-Abschnitten festgelegt werden können.

2 Einbettung der Arbeit

Diese Arbeit entstand als externe Arbeit im Rahmen des Forschungsprojektes „PerManEntA“¹ des Labors für Verteilte Systeme der Fachhochschule Wiesbaden beim Projektpartner Lufthansa Systems Passenger Services (LHS), wo die entwickelten Methoden auf die LHS-Check-In-Anwendung „PICTURES“ angewendet wurden. Ziel des PerManEntA-Projekts ist die Entwicklung von Methoden und Werkzeugen für das Performance Management von komplexen verteilten Anwendungen auf einer hohen Abstraktionsebene. Im Rahmen des Projektes entstanden Methoden für die modellbasierte Instrumentierung einer Anwendung, bei denen Stellen des Quellcodes einer Anwendung zur Instrumentierung ausgezeichnet werden können, indem ihnen sogenannte Instrumentierungspunkte zugeordnet werden, ähnlich dem Anbringen eines Breakpoints in einem Debugger. Im Gegensatz zu einer manuellen Instrumentierung, bei der entsprechende Code-Fragmente in den Quellcode eingefügt werden, wird die Instrumentierung nun innerhalb eines Anwendungsmodells eingebettet. Instrumentierungspunkte können im Instrumentierungsmodell in logischen Gruppen, sogenannten Patterns, zusammengefasst werden. Die entsprechenden Code-Fragmente für die Instrumentierung können anhand des Modells automatisch generiert werden. Dieser generierte Code erzeugt, je nach Konfiguration, zur Laufzeit Log- oder Messausgaben mittels der Frameworks Log4J oder ARM (Application Response Measurement, [Ope03]).

Außerdem entstand in dem Projekt ein für diese Arbeit genutztes Workflow-Meta-Modell, das als eine Abstraktion von Programmcode oder Programmabschnitten auf einer hohen Ebene verstanden werden kann. Dieses einfache Workflow-Modell dient als kleinster gemeinsamer Nenner von Arbeitsabläufen in Geschäftsanwendungen und ermöglicht die Konvertierung oder Abbildung unterschiedlicher etablierter Workflow-Notationen wie UML oder BPMN [Obj06] in ein gemeinsames Format und erlaubt außerdem die Weiterverwendung bestehender Modeling-Tools und UML-Anwendungsmodelle. Es gibt fünf grundlegende Bausteine, aus denen jedes Workflow-Modell, so wie hier definiert, zusammengesetzt ist: Workflow (Top-Level-Element), Action (stellt einen einfachen Befehl oder Programmabschnitt dar und hat genau einen Eingangs- und einen Endpunkt), Branch (Eine Verzweigung mit einem Eingangs- und beliebig vielen Endpunkten), Merge (Gegenteil zu Branch mit beliebig vielen Eingangspunkten und einem Endpunkt), und Parallel. Instanzen dieser Bausteine werden kombiniert und bilden ein konkretes Anwendungsmodell. Um das Workflow-Modell mit seinem zugehörigen Code-Abschnitt zu koppeln, wird

¹Performance Management of Enterprise Critical Applications (gefördert durch die BMBF unter Förderkennzeichen 1706X07), [Fac08]

eine Zuordnung von Code-Stellen zu Workflow-Punkten benötigt. Zu diesem Zweck existiert eine Abbildung des Workflow-Modells auf das Instrumentierungsmodell, bei der ein Workflow-Punkt einem Instrumentierungspunkt zugeordnet ist.

3 Ergebnisse der Arbeit

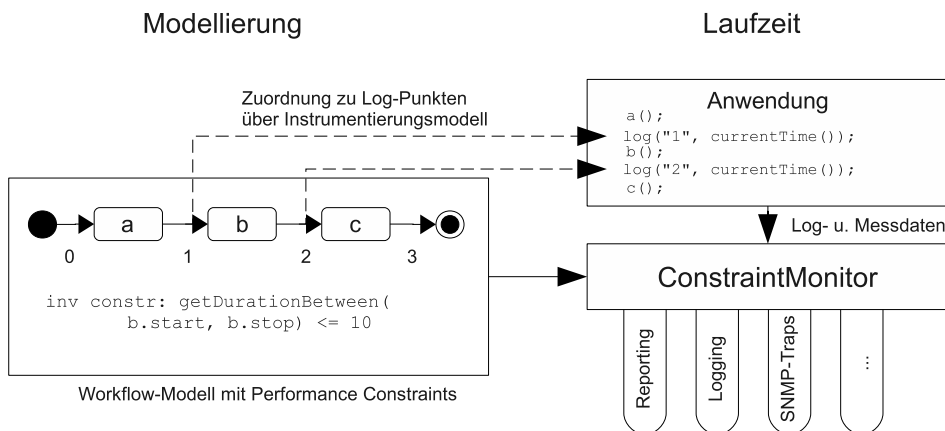


Abbildung 1: Einsatz des ConstraintMonitors

Abbildung 1 zeigt die Architektur des entwickelten Systems: Für das Monitoring der zu überwachenden Anwendung wird zur Laufzeit der sogenannte ConstraintMonitor verwendet. Der Strom von Performance-Daten (also sowohl Log- als auch Messdaten), die von dem mittels der modellbasierten Instrumentierung in der Anwendung angebrachten Code erzeugt werden, werden vom ConstraintMonitor gelesen und dem Instrumentierungsmodell zugeordnet, das der Codestelle oder Komponente entspricht, von der die Daten stammen. Dies kann über unterschiedliche Wege geschehen: Ein Log-Punkt, der über die Methode der modellbasierten Instrumentierung in der Anwendung angebracht wurde, verfügt über eine eindeutige ID; Log-Punkte, die unabhängig davon in der Anwendung vorhanden sind, können über den voll qualifizierten Klassennamen und die Zeilennummer eindeutig identifiziert werden. Eine ARM-Messung kann anhand ihrer Transaction-ID einem Pattern im Instrumentierungsmodell zugeordnet werden.

Anschließend werden diese Daten vereinheitlicht: Zeitstempel von Log-Meldungen und Zeitdauern von ARM-Messungen werden in ein gemeinsames Format gebracht, das es erlaubt, Constraints auszuwerten, unabhängig davon, in welcher ursprünglichen Form diese Daten vorlagen. Hierzu wird zunächst die Verbindung von Instrumentierungspunkten zu Punkten im zur Anwendung passenden Workflow-Modell hergestellt. Die Zuordnung wird zum Modellierungszeitpunkt vorgenommen und ist als Teil des Workflow-Modells für den ConstraintMonitor verfügbar. Weil in Constraints Aussagen über die Dauer von

Abschnitten der Anwendung getroffen werden, müssen die gemessenen Log-Punkte mit Zeitstempeln in Zeitdauern umgewandelt werden. Dabei gehören immer diejenigen Log-Punkte zusammen zu einem Abschnitt, die auf den Eingangs- bzw. Endpunkt einer Action im Workflow-Modell abgebildet werden; die Zeitdifferenz zwischen den beiden Zeitstempeln gibt also die Zeitdauer dieser Action an.

Auf diese Weise baut der ConstraintMonitor intern eine Datenbasis der Zeitdauern der durchlaufenen Abschnitte der Anwendung auf, die dazu verwendet wird, die im Workflow-Modell festgelegten Performance Constraints zu überprüfen. Für das Parsen und Auswerten der OCL-Constraints kommt im ConstraintMonitor die OCL-Komponente des EMF (Eclipse Modeling Framework, [emf]) zum Einsatz.

Der ConstraintMonitor verfügt über ein Modulsystem, das unterschiedliche Reaktionen auf die so entstandenen Events erlaubt. Performance-Daten können vom ConstraintMonitor sowohl online ausgewertet werden, also während die zu überwachende Anwendung läuft, oder offline, indem zuvor aufgezeichnete Log-Files und Messwerte als Eingabestrom verwendet werden. Für die Offline-Auswertung kann mittels eines Reporting-Moduls ein Bericht erstellt werden, der aufzeigt, welche Constraints wie oft verletzt wurden; Leistungsengpässe in bestimmten Anwendungskomponenten können so auch im Nachhinein gefunden und zusammengefasst werden. Wird der ConstraintMonitor für die Online-Auswertung verwendet, kann ein Modul eingesetzt werden, das bei Constraint-Verletzungen SNMP-Traps an eine übliche Management-Konsole verschickt, die Informationen über den untersuchten Workflow und das verletzte Constraint enthalten. Außerdem haben Module die Möglichkeit, die Auswertungsumgebung der OCL-Constraints zu beeinflussen, Werte von OCL-Variablen können so zur Laufzeit verändert werden. Dies erlaubt die dynamische Anpassung von Constraint-Grenzwerten und kann beispielsweise verwendet werden, um automatisch den Detaillierungsgrad von Ausgaben zu erhöhen oder zu reduzieren. Eine Anwendung für das Selfmanagement der überwachten Anwendung ist ebenfalls denkbar und wird zur Zeit untersucht.

Literatur

- [emf] Eclipse Modeling Framework. <http://www.eclipse.org/modeling/emf/>.
- [Fac08] Fachhochschule Wiesbaden - Labor für verteilte Systeme. *Performance Management of Enterprise Critical Applications (PerManEntA)*, 2008. <http://www.vv.informatik.fh-wiesbaden.de/projekte/permanenta.html>.
- [Obja] Object Management Group. *Business Process Modelling Notation (BPMN) Specification, Version 1.1*. <http://www.omg.org/spec/BPMN/1.1/PDF/>.
- [Objb] Object Management Group. *Unified Modeling Language (UML) Specification, Version 2.1.2*. http://www.omg.org/technology/documents/modeling_spec_catalog.htm#UML.
- [Obj06] Object Management Group. *Object Constraint Language Specification, Version 2.0*, 2006. <http://www.omg.org/technology/documents/formal/ocl.htm>.
- [Ope03] The Open Group. *ARM 4.0 Java Language Binding Technical Standard 4.0*, October 2003. <http://www.opengroup.org/arm/uploads/40/3945/C037.pdf>.

A Topology-Aware Overlay for Wireless Mesh Networks

Hanno Wirtz, hanno.wirtz@rwth-aachen.de

Distributed Systems Group, RWTH Aachen

Abstract: In a Wireless Mesh Network (WMN), routers communicate over successive radio links. As no central managing entity is present, the overall structure of the network is unknown. Thus, addressing of routers and data items stored at routers is a challenge. In this work, we first introduce a spring-based approach to assigning virtual geographic coordinates to routers to efficiently route through the WMN. Furthermore, we construct an index structure on top of these coordinates that allows for unique addressing and distributed management of data items. As communication in the WMN is expensive, both the assignment of coordinates and the construction of the index structure take the underlying network topology into account to avoid loops and indirections in the routing and data retrieval process.

1 Introduction

The growing capabilities of wireless network technologies have lead to pervasive deployment of wireless routers. Besides providing access to a wired network, communication between two routers is also possible. The set of connections between wireless routers then forms a network structure that is independent from the wired infrastructure, called a *wireless mesh network* (WMN). Advantages of WMNs are simple deployment, as the installation of new routers suffices to expand the network, and the ability of single routers to automatically configure themselves with regard to the surrounding network.

Wireless mesh routers, so called mesh nodes, are able to forward data over a certain number of links or *hops* in the mesh. Forwarding data requires a means to address the next hop, each node however only knows about its immediate neighbors. While it is possible to discover routes to each node in the network by managing tables of complete routes at every node, this approach is exhaustive in terms of communication and memory. Especially in dynamic networks, routing tables need to be updated constantly. To establish efficient communication between non-neighboring nodes, nodes need to have identifiers that are meaningful in the whole network. The assignment of these identifiers needs to be unique, to furthermore ensure progress when routing towards an identifier, the assignment should be topologically meaningful, this means that physically close nodes in the network should have similar identifiers.

Another challenge in a WMN is the discovery of data items or services. Consistently handling data items or services accessible by all nodes in the network requires a means of storing and retrieving each item by a unique identifier. To this end, addressing data items requires a mechanism that assigns each identifier to a node in the network that handles the

underlying item. A structure that provides such a mechanism is a Distributed Hash Table (DHT). In a DHT, responsibility for the space of identifiers is distributed among nodes in a structured manner that allows for an efficient look-up of items. In the DHT, nodes that share adjacent parts of the identifier space will forward requests based on the current distribution of responsibility to route the request to the responsible node. We introduce an approach to construct a DHT on top of a WMN, a so called *overlay network*, with the goal of preserving the underlying topology in the topology of the overlay to account for the fact that communication in the WMN is expensive.

2 Problem Analysis

The scope of our work consists of two conceptual different problems. First, we need a unique identifier for each node to address a node in the routing process. Using this identifier structure, we implement a greedy routing scheme. Second, to facilitate data lookup and retrieval, a means of addressing data consistently in the network is needed. There are existing approaches to both problems, these however fail to address the characteristics of a WMN. This is because they do neither account for the dynamic structure of the network nor the expensive and unreliable communication between nodes.

Rao et al. [RRP⁺03] present an approach to assigning *virtual coordinates* to nodes by finding nodes on the perimeter of the network. Subsequently, each node assigns a position to itself based on its distance in hops to each of these perimeter nodes. Distributing the distance information requires flooding the network with distance messages. The main disadvantage of this approach is the fact that a node joining the network on the outside of the current perimeter either renders obsolete the distance and position estimates of inner nodes or fails to estimate a correct position for itself when aligning itself inside the perimeter.

There are several classical design approaches to constructing a DHT such as CAN[RFH⁺01] and Chord[SMK⁺01]. In these designs, nodes join the DHT by randomly choosing an identifier from the identifier space. This way, neighbor-relations between nodes do not take the physical location of these nodes into account. Forwarding a request between neighbors in the DHT may thus require traversing large parts of the WMN.

3 Our Approach

The main focus of our approach is the construction of both the node identifiers and the overlay with regard to the underlying topology. The motivation for this approach is to guarantee progress when routing towards a node or an overlay identifier without indirections. As stated in the previous section, this approach is divided in *i)* the assignment of an identifier to each node and *ii)* the construction of the DHT using this identifier distribution.

i) To assign an identifier to each node, we employ an approach using virtual coordinates

presented by Moore et al. [NM05], modified to better meet our requirements. To derive a virtual coordinate, a node only needs the virtual coordinates of its direct and two-hop neighbors, which is received by periodic broadcasts without any additional overhead. The calculation itself is based on a relaxation-algorithm with the participating forces being *attraction* between one-hop neighbors and *repulsion* between two-hop neighbors. Using only the distance to all known neighbors, a node calculates an *energy level* for a given coordinate. Each node tries to find a coordinate that minimizes this energy level. This way, each node calculates a topologically oriented unique identifier using only local communication. This approach satisfies the requirement of similar coordinates at physically close nodes while allowing for joining and departing nodes.

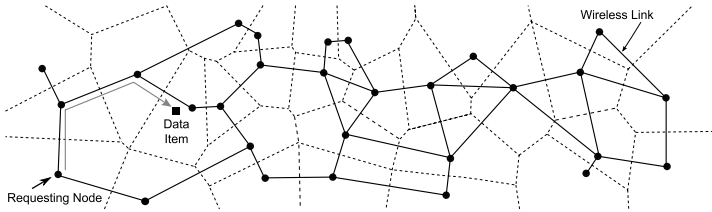


Figure 1: Voronoi fragmentation

ii) As stated in section 3, we construct an overlay on top of these coordinates to facilitate a data-centric addressing mechanism. We choose a *2D Voronoi fragmentation* as shown in figure 1 to calculate a *virtual region* of responsibility for each node. This region contains all points in the identifier space that lie closer to this node than to any other node. In terms of maintenance, the calculation solely depends on the coordinates of the surrounding nodes in the overlay and thus flexibly adapts to the topology information of the underlying coordinates.

A node joins the overlay in the usual DHT manner by first finding the node that is currently responsible for its identifier. To calculate its own virtual region, the joining node will gather the coordinates of the current neighbors of its proxy as the adjacent nodes of the proxy's region will likely be the neighbors of the joining node. After calculating its own region, the joining node announces its presence to its own neighbors in the overlay, which in turn will adapt their virtual region to the new topological situation.

As the overlay construction solely depends on the virtual coordinates of nodes, it is easily adjustable in case a node leaves or joins the network. Furthermore, by incorporating the topological information of virtual coordinates, directional routing in the overlay directly corresponds to routing progress in the physical network.

4 Evaluation

To evaluate our approach, two measurements are significant. First, as we try to preserve the network topology in the overlay, the set of overlay neighbors should contain a high

fraction of physical one- or two-hop neighbors. Second, the distribution of responsibility in the overlay should result in virtual regions of equal size.

Measurements show that for 20% – 30% of all nodes the fraction of local neighbors in virtual neighbors is 100% and for 45% – 60% of nodes this fraction still is between 80% and 100%. Regarding these results, the main objective of our approach is fulfilled at a high percentage of nodes. For up to 10% of the nodes, the fraction of local neighbors is small due to the fact that they lie on the borders of the coordinate distribution and their next neighbors coordinate-wise are farther away.

Comparing the size of virtual regions yields the result that 60% – 80% of all nodes has a similar virtual region size. Again, around 10% of all nodes have larger virtual regions due to the fact that these nodes have coordinates on the boundary of the identifier space. As the identifier space needs to be completely fragmented, these nodes are responsible for virtual regions that are adjacent to regions managed by nodes farther through the network.

5 Conclusion and Outlook

We have presented an approach to scalably and flexibly assign topologically meaningful identifiers to mesh nodes to facilitate efficient routing between nodes in the network. To establish consistent data storage and retrieval, we implement an overlay on top of these identifiers that is topology-oriented as well, ensuring efficient data retrieval.

As the overlay structure is based on the virtual coordinates of nodes, a more equal distribution of these would improve the quality of both the virtual neighbor distribution and the virtual region size. This aspect of our approach lies in the scope of our future work.

References

- [NM05] Gregory K. Egan Nick Moore, Ahmet Sekercioglu. Virtual Localization for Mesh Network Routing. In *Proc. Network and Communication Systems (NCS 2005)*, 2005.
- [RFH⁺01] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, and Scott Shenker. A Scalable Content-Addressable Network. In *Proceedings of the ACM SIGCOMM '01 Conference*, San Diego, California, August 2001.
- [RRP⁺03] Ananth Rao, Sylvia Ratnasamy, Christos Papadimitriou, Scott Shenker, and Ion Stoica. Geographic routing without location information. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 96–108, New York, NY, USA, 2003. ACM.
- [SMK⁺01] Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan. Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. In *Proceedings of the ACM SIGCOMM '01 Conference*, San Diego, California, August 2001.

Aktivitätsorientierte Kontextadaption in mobilen Anwendungen

Jan D.S. Wischweh

Arbeitsbereich Verteilte Systeme und Informationssysteme, Universität Hamburg
mail@wischweh.de

Abstract: Mobile Geräte wie z.B. Smartphones durchdringen immer weiter den Alltag. Dabei bringen die geringe Größe und der enorm steigende Funktionsumfang der Geräte erhebliche, bisher kaum gelöste Nutzungsschwierigkeiten mit sich. Die Kenntnis des aktuellen Nutzungskontextes kann helfen, die Interaktion mit mobilen Geräten zu erleichtern. Aktivitätskontext (Ziele und Tätigkeiten des Nutzers) ist dabei ein Teilbereich von Kontext, der spezifische Potentiale zur Verbesserung von Interaktion hat, jedoch bisher kaum genutzt und untersucht wurde. Diese Potentiale und Ansätze zu ihrer Realisierung werden im Folgenden geschildert.

1 Einleitung

Das Marktforschungsunternehmen Gartner schätzt, dass im Jahr 2008 weltweit 1.28 Milliarden Mobiltelefone verkauft werden [Gara]. Dabei wächst der Anteil von Smartphones, Telefonen deren Funktionalität weit über bloßes Telefonieren hinausgeht, stetig [Garb]. Den explosionsartig wachsenden Möglichkeiten der Geräte stehen prinzipielle Schwierigkeiten bei deren Bedienung gegenüber: Bedingt durch die Mobilität der Geräte steht nur wenig Platz für Ein- und Ausgabemöglichkeiten zur Verfügung [LK06]. Eine Strategie um diese Probleme zu reduzieren, ist das Ausnutzen des jeweiligen Kontextes, in dem Interaktion mit einem mobilen Gerät erfolgt. Im Bereich der Ausgabe ist es einer Anwendung beispielsweise möglich Schriftgröße, Lautstärke, Helligkeit und Einstellungen zur Privatsphäre an die gegebene Situation anzupassen. Auf Eingabeseite kann eine Anpassung der angebotenen Auswahlmöglichkeiten und der Eingabemethode an die aktuelle Situation erfolgen. Denkbar ist beispielsweise der Einsatz von zur Umgebung passenden Audiofiltern, um die Zuverlässigkeit der Spracherkennung zu optimieren [Sch00].

Kontext lässt sich grob in fünf Bereiche gliedern. Der *persönliche Kontext* beschreibt Eigenschaften des Nutzers, z.B. Stimmung, Wissen, Fähigkeiten und Behinderungen. Der *Umweltkontext* erfasst die unmittelbare Umgebung des Nutzers, wie z.B. vorhandene Gegenstände, verfügbare technische Dienste, anwesende Personen. Regeln und Rollen des Nutzers welche seine Handlungen beeinflussen, bilden den *Sozialen Kontext*. Ein weiterer Bereich ist der *zeitlich-räumliche Kontext* und schließlich der *Aufgabenkontext*, der sich aus den Aktivitäten und Zielen des Nutzers ergibt [KPC06, S.12f]. Im Bereich der Mobiltelefone konzentrierte sich die Forschung zur Nutzung von Kontext bisher vor allem auf die Aspekte der physikalischen Umgebung (Ort [DRD⁺00], Umweltbedingungen [Sch00]

und verfügbare technische Ressourcen [PSGP01]) und zum Teil auf die aktuelle soziale Situation des Nutzers [ORT05]. Aspekte wie Informationen über den Nutzer, seine Aktivitäten und zeitliche Zusammenhänge wurden bisher weniger genutzt [KA04].

2 Aktivitätsorientierte Kontextadaption in mobilen Anwendungen

Dabei verspricht die Kenntnis von aktuellen, vergangenen oder zukünftigen Aktivitäten des Nutzers spezifische Möglichkeiten, die Interaktionsabläufe zwischen Nutzern und mobilen Geräten zu verbessern:

- Durch Aktivitäten werden Beziehungen zwischen Personen, Orten und Objekten hergestellt. Diese Beziehungen können analysiert und genutzt werden, um Vorhersagen für die Relevanz bestimmter Personen, Orte oder Objekte in einer gegebenen Situation zu treffen.
- Die Tätigkeiten des Nutzer bestimmen, wo gerade sein Aufmerksamkeitsfokus liegt und wie hoch seine aktuelle kognitive Last ist. Diese Kenntnisse können als Grundlage dienen, um dynamisch Elemente der Benutzerschnittstelle an die aktuellen Bedürfnisse des Nutzers anzupassen.
- Kenntnisse über übliche Tätigkeitsmuster eines Nutzers lassen sich zur Vorhersage wahrscheinlicher zukünftiger Aktivitäten nutzen. Umgekehrt können Abweichungen von üblichen Tätigkeitsmustern Hinweise auf einen veränderten Nutzungskontext sein.

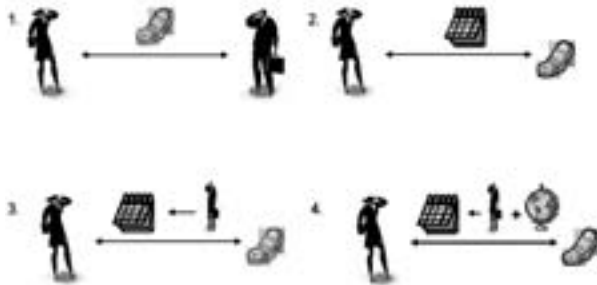


Abbildung 1: Beispielszenario: Aus vergangenen Aktivitäten wird auf zukünftig relevante Personen und Orte geschlossen.

Angenommen Alice telefoniert mit Bob (siehe Abbildung 1 Punkt 1). Anschließend öffnet Alice ihren digitalen Terminplaner (2.). In dieser Situation ist es höchst wahrscheinlich, dass Alice einen Termin "Treffen mit Bob" in ihren Kalender eintragen will (3.). Wahrscheinliche Orte für das Treffen sind Orte, die mit Bob in Verbindung stehen, wie z.B.

Bobs Haus oder sein Arbeitsplatz. Auch Orte an denen Alice und Bob sich häufig getroffen haben oder die Alice gerne für Treffen nutzt, könnten der Ort des Termins sein (4.). Ein mögliches Szenario zur Nutzung von Aktivitätskontext sind somit kontextadaptive Auswahllisten, die anstatt einer statischen Sortierung Personen und Orte in Reihenfolge ihrer angenommenen aktuellen Relevanz anbieten.

3 Ansatz

Im vorherigen Abschnitt wurde der Nutzen von aktivitätsorientiertem Kontext verdeutlicht. Um zu einer Ausnutzung von Aufgabenkontext zu gelangen, wird folgendermaßen vorgegangen:

Ausgehend von der Fragestellung “Wie lässt sich der Aufgabenkontext (Tätigkeiten und Ziele des Nutzers) zur Laufzeit erfassen und zur Verbesserung der Interaktion mit mobilen Geräten nutzen?” müssen die Teilfragen “Wie können Aktivitäten erfasst und modelliert werden?”, “Welche Schlüsse können aus diesen Aktivitäten gezogen werden?”, “Wie lassen sich diese Informationen sinnvoll nutzen?” erörtert werden. Zu diesem Zweck ist die Entwicklung eines generischen aktivitätsorientierten Kontextmodells erforderlich. Ein generisches Kontextmodell propagiert eine einheitliche Art und Weise, wie Kontext aufgefasst und dargestellt wird. Es ist Grundlage für die gemeinsame Nutzung von Kontextinformationen durch unterschiedliche Applikationen und ermöglicht die Unterstützung der Abfrage, Nutzung und des Austausches dieser Informationen mittels einer Middleware. Eine solche Middleware implementiert das entwickelte Kontextmodell und stellt Anwendungen eine Schnittstelle zur Nutzung des Aktivitätskontextes bereit und bietet Möglichkeiten zum Erfassen der auftretenden Nutzeraktivität.

Schon während der Vorarbeit sind dabei einige spezifische Eigenschaften von Aktivitätskontext deutlich geworden, die sich in solch einer Architektur zur Nutzung von Aktivitätskontext widerspiegeln müssen: Anders als in den meisten bisherigen kontextabhängigen System können Anwendungen Aktivitätskontext nicht nur nutzen, sondern der Anwender stellt diesen durch seine Aktivitäten auch selbst her. Somit ist ein Informationsfluss in beide Richtungen, also vom Kontextsystem zu kontextabhängiger Anwendung und umgekehrt, notwendig. Bei Aktivitäten handelt es sich um zeitlich ablaufende Sequenzen von Aktionen. Mehrere Aktivitäten können parallel ablaufen und die gleiche Art von Aktivität kann unterschiedliche Verlaufsformen annehmen. Der möglichst einfache Umgang mit zeitlichen Eigenschaften, Parallelität und alternativen Sequenzen sollte somit von Modell und Architektur berücksichtigt werden.

Um die in Abschnitt 2 vorgestellten kontextadaptiven Auswahllisten zu realisieren muss, als ein Beispiel für Schlussfolgerungsmöglichkeiten aus dem Aktivitätskontext, eine Heuristik entwickelt werden, welche die Relevanz von einzelnen Personen und Orten zu einem bestimmtem Zeitpunkt abschätzt und in Echtzeit lauffähig auf mobilen Geräten ist.

4 Zusammenfassung

Die Möglichkeiten zur Erleichterung von Interaktion mit mobilen Geräten, die sich aus Analyse und Nutzung des Aktivitätskontextes zur Laufzeit ergeben, wurden bisher nahezu kaum ausgenutzt. Zur Unterstützung der Entwicklung von kontextadaptiven Anwendungen in diesem Bereich sollte eine Middleware Aktivitätskontext modellieren und verwalten. Eine prototypische Implementierung wird die prinzipielle Umsetzbarkeit von Kontextmodell, Middleware und Beispielszenario aufzeigen.

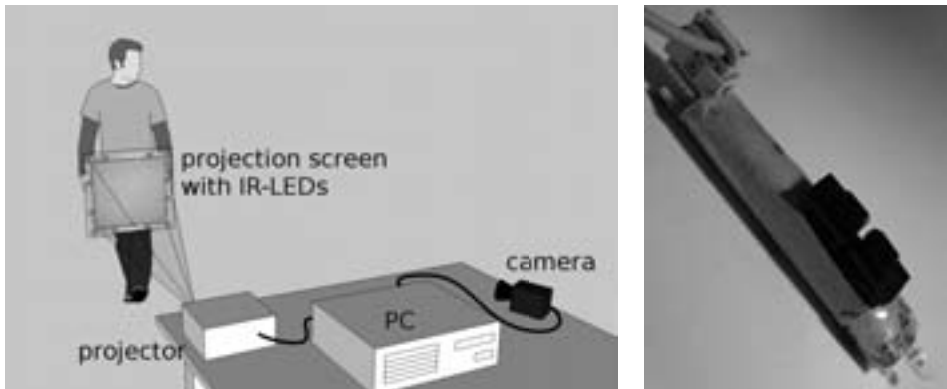
Literatur

- [DRD⁺00] Alan Dix, Tom Rodden, Nigel Davies, Jonathan Trevor, Adrian Friday und Kevin Palfreyman. Exploiting space and location as a design framework for interactive mobile systems. *ACM Trans. Comput.-Hum. Interact.*, 7(3):285–321, 2000.
- [Gara] Pressemitteilung Gartner. Gartner Says Worldwide Mobile Phone Sales Increased 12 Percent in Second Quarter of 2008. <http://www.gartner.com/it/page.jsp?id=747414>, Stand: 30.10.2008.
- [Garb] Pressemitteilung Gartner. Gartner Says Worldwide Smartphone Sales Grew 16 Per Cent in Second Quarter of 2008. <http://www.gartner.com/it/page.jsp?id=754112>, Stand: 30.10.2008.
- [KA04] Manasawee Kaenampornpan und Bath Ba Ay. Modelling context: an Activity Theory approach. In *Ambient Intelligence: Second European Symposium, EUSAI 2004*, Seiten 367–374. Springer, 2004.
- [KPC06] Anders Kofod-Petersen und Jörg Cassens. Using activity theory to model context awareness. *Modeling and Retrieval of Context: Second International Workshop, MRC 2005, Revised Selected Papers. Volume 3946 of LNCS (LNAI)*, Seiten 1—17, 2006.
- [LK06] Ulrich Leiner und Sonja Krzonkalla. Mobile Endgeräte: Einfache Alleskönner - ein Widerspruch? *i-com - Zeitschrift für interaktive und kooperative Medien*, 1(1):41–47, 2006.
- [ORT05] Antti Oulasvirta, Mika Raento und Sauli Tiitta. ContextContacts: re-designing Smartphone's contact book to support mobile awareness and collaboration. In *IEEE Photon. Technol. Lett.*, Seiten 167–174. Portal, 2005.
- [PSGP01] Thai-Lai Pham, Georg Schneider, Stuart Goose und Arturo Pizano. Composite Device Computing Environment: A Framework for Situated Interaction Using Small Screen Devices. *Personal Ubiquitous Comput.*, 5(1):25–28, 2001.
- [Sch00] Albrecht Schmidt. Implicit human computer interaction through context. *Personal and Ubiquitous Computing*, 4(2):191–199, Juni 2000.
- [SLP04] Thomas Strang und Claudia Linnhoff-Popien. A Context Modeling Survey. In: *Workshop on Advanced Context Modelling, Reasoning and Management, UbiComp 2004 - The Sixth International Conference on Ubiquitous Computing, Nottingham/England, 2004*.

User Interfaces based on a hand-held projection screen

Kristian Bergmann

1 Übersicht



(a) Das System besteht aus einer frei beweglichen, ultraleichten Leinwand und handelsüblicher Kamera und Beamer.

(b) Der Stift wird als Eingabegerät direkt auf der Leinwand benutzt.

Abbildung 1: (a) Systemskizze und (b) Foto des Stifts

Im Rahmen meiner Diplomarbeit “User Interfaces based on a hand-held projection screen” wurden Hardware und ein Softwaresystem entwickelt, um virtuelle 3D-Umgebungen auf intuitive Art und Weise erkunden zu können. Eine wichtige Fragestellung im Gebiet der Mensch-Computer-Interaktion ist, wie man durch virtuelle 3D-Räume und -Daten navigieren kann ([HPGK94, Han97, Sub00] fassen einige Aspekte dessen zusammen). Die Idee eines Anzeigegerätes, das seine Lage im Raum kennt, geht auf Fitzmaurice et al. [FZC93] zurück. Inspiriert durch die Arbeiten von Lee et al. [LHSD05, LHD07] entstand die Idee ein System zu entwickeln, das im Gegensatz zum Boom Chameleon [TFK⁺02] leicht genug ist, um in der Hand gehalten zu werden. Ein weiteres Ziel war es, anders als bei ähnlichen Systemen wie z.B. dem Personal Interaction Panel [SG97], kein kompliziertes oder teures Hardwaresystem zu benötigen.

Die im Rahmen der Diplomarbeit entwickelte Hardware (Abb. 1a) besteht aus einem etwa 40x30cm großen, mit Projektionsfolie bespannten, Rahmen (der Leinwand) und einem Stift mit zwei Knöpfen (Abb. 1b). Beides zusammen wiegt ca. 140g und wird vom Nutzer in den Händen gehalten. Auf dem Rahmen der Leinwand und in der Stiftspitze befinden sich Infrarot-Leuchtdioden (IR-LED), welche vom Softwaresystem mit einer feststehen-

den Kamera aufgenommen werden. Aus den Bilddaten kann die Pose (Position und Orientierung) der Leinwand im Raum und die Position des Stifts auf der Projektionsfolie bestimmt werden. Diese Informationen werden genutzt, um das von einem feststehenden Beamer projizierte Bild vor der Projektion so zu verzerren, dass es passgenau und unverzerrt auf der beweglichen Leinwand erscheint. Außerdem können die Leinwandpose und die Position des Stifts als Eingabegerät mit insgesamt 8 Freiheitsgraden genutzt werden.

Zusätzlich zur Entwicklung des beschriebenen Systems, wurde es als Eingabegerät in Fibermesh [NISA07], ein skizzenbasiertes Programm zur Erstellung frei definierbarer 3D-Flächen, integriert. Dazu wurden verschiedene Methoden implementiert, mittels derer der Nutzer, durch die Leinwandpose, die Sicht auf das modellierte Objekt einstellen kann. Diese wurden dann in einer informellen Nutzerstudie bewertet und mit einer üblichen Maussteuerung verglichen.

2 Software

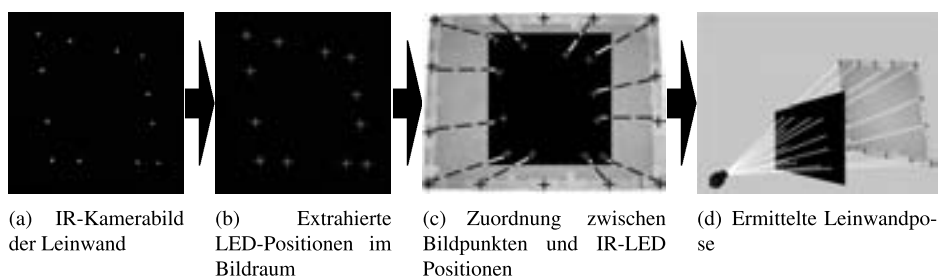


Abbildung 2: Ausgehend vom Kamerabild wird in mehreren Stufen (a)-(d) die Lage der Leinwand im Raum bestimmt. Anschließend wird das projizierte Bild so verzerrt, dass es, nach der Projektion, auf der Leinwand passgenau und unverzerrt erscheint.

Das Softwaresystem zur Bestimmung der Leinwandpose besteht aus drei Komponenten. Die erste Komponente (Abb. 2, a→b) basiert auf einer rechenzeitoptimierten Adaptation des in [Hor86] beschriebenen Algorithmus zur “blob-extraction”. Dieser findet helle Leuchtpunkte im Bild, die das Abbild der, durch einen infrarot-pass Filter aufgenommenen, IR-LED sind. Dazu werden zusammenhängende Mengen heller Pixel im Kamerabild bestimmt und deren Schwerpunkte ermittelt.

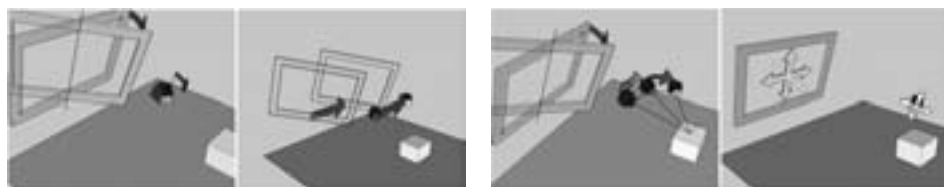
Der nächste Schritt (Abb. 2, b→c), zur Ermittlung der Leinwandpose, ist die Bestimmung einer Paarung zwischen den im Bild gefundenen Punkten und den bekannten Positionen der IR-LED auf der Leinwand. Hierzu wird der Umstand genutzt, dass die LED auf dem Rahmen der Leinwand in Geraden angebracht sind und dort auf jeder Seite andere Längenverhältnisse zwischen den einzelnen LED bestehen. Diese Längenverhältnisse lassen sich im Bild wieder finden. Anhand der so miteinander identifizierten Bild- und Rahmengeraden können auch die Leuchtpunkte mit den LED identifiziert werden.

Letztendlich wird die Pose der Leinwandfläche, anhand der Bildraumpositionen der Leucht-

punkte und der bekannten Leinwandgeometrie, bestimmt (Abb. 2, c→d). Dazu müssen zuerst die Bildpunkte mit von der Kamera ausgehenden Strahlen im Raum in Verbindung gebracht werden, wozu die intrinsischen Parameter der Kamera bekannt sein müssen. Den Vorgang zur Ermittlung derselben nennt man Kamerakalibrierung. Hierzu benutzt mein Softwaresystem eine für Videodaten angepasste Variante des von Zhang vorgestellten Algorithmus [Zha99]. Nach erfolgreicher Kalibrierung der Kamera liefert der Algorithmus von Zhang zusätzlich für jedes analysierte Bild die Pose der Kamera relativ zur Leinwand (auch extrinsische Kameraparameter genannt), aus der die Leinwandpose abgeleitet wird.

Zusätzlich dazu kann das System interaktiv die Pose (relativ zur Kamera) und die intrinsischen Parameter des Beamers ermitteln. Mit diesen Informationen wird dann das Bild, vor der Projektion durch den Beamer, so verzerrt, dass es genau auf der Leinwand erscheint.

3 Interaktion mittels des Bildschirms



(a) Absolut: Bewegungen des Rahmens werden direkt in Kamerabewegungen umgesetzt.

(b) Relativ: Neigen des Rahmens führt zu fortlaufender Rotation der Kamera ums Objekt. Verschiebungen werden tangential zur Rotation ausgeführt.

Abbildung 3: (a) Absolute Kamerasteuerung und (b) Relative Kamerasteuerung

Bei der Interaktion mittels der Leinwandpose (der große, rechteckige Rahmen in Abb. 3) galt das Hauptaugenmerk im Rahmen meiner Diplomarbeit der Steuerung der (virtuellen) Kamera (das kleine, bewegte Objekt in Abb. 3), durch die ein einzelnes Objekt (in Abb. 3 weiß) im virtuellen 3D-Raum betrachtet wird.

Die absolute Kamerasteuerung (Abb. 3a) überträgt direkt jede Bewegung der Leinwand in eine gleichartige Bewegung der kontrollierten Kamera. Hierbei ergibt sich das Problem, dass die IR-LED immer von der (realen) Kamera aus sichtbar sein müssen, was Rotationen der Leinwand auf etwas unter 90° in jede Richtung einschränkt. Dadurch ist die Sicht auf das Objekt auf eine Objektseite eingeschränkt. Indem man Nutzern erlaubt das Objekt auf Wunsch “festzuhalten”, kann es trotz dieser Einschränkung von allen Seiten betrachtet werden.

Eine wesentlich überzeugendere Variante die Einschränkung aufgrund der optischen Natur des Systems zu umgehen, ist die relative Kamerasteuerung (Abb. 3b). Dabei rotiert die (virtuelle) Kamera solange um das betrachtete Objekt wie der Benutzer die Leinwand aus ihrer neutralen Lage neigt. Geschwindigkeit und Richtung der Kamerarotation ergeben sich dabei aus dem Ankippwinkel und der Kipprichtung der Leinwand.

Varianten und Kombinationen dieser beiden Grundmechanismen wurden implementiert und in einer informellen Nutzerstudie gegenübergestellt. Die in der Studie gewonnenen Daten deuten darauf hin, dass eine Kombination der beiden Kamerasteuerungsarten ähnlich gut angenommen wird, wie eine etablierte Maussteuerung der Sicht.

Mögliche, und für einige Anwendungen nötige, technische Verbesserungen am vorgestellten System beinhalten die Erhöhung der Stiftpositionsauflösung und die Verringerung der bestehenden Orientierungseinschränkungen. Außerdem soll in einer Nutzerstudie die Effizienz des Systems für Kamerasteuerungsanwendungen sowie andere 3D-Interaktionen mit Computern erfasst werden.

Literatur

- [FZC93] George W. Fitzmaurice, Shumin Zhai und Mark H. Chignell. Virtual reality for palmtop computers. *ACM Trans. Inf. Syst.*, 11(3):197–218, 1993.
- [Han97] C. Hand. A Survey of 3D Interaction Techniques. In *Computer Graphics Forum*, Jgg. 16, Seiten 269–281, 1997.
- [Hor86] Berthold K. Horn. *Robot Vision*. McGraw-Hill Higher Education, 1986.
- [HPGK94] Ken Hinckley, Randy Pausch, John C. Goble und Neal F. Kassell. A survey of design issues in spatial input. In *UIST '94: Proceedings of the 7th annual ACM symposium on User interface software and technology*, Seiten 213–222, New York, NY, USA, 1994. ACM.
- [LHD07] Johnny Lee, Scott Hudson und Pau Dietz. Hybrid infrared and visible light projection for location tracking. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, Seiten 57–60, New York, NY, USA, 2007. ACM.
- [LHSD05] Johnny C. Lee, Scott E. Hudson, Jay W. Summet und Paul H. Dietz. Moveable interactive projected displays using projector based tracking. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, Seiten 63–72, New York, NY, USA, 2005. ACM.
- [NISA07] Andrew Nealen, Takeo Igarashi, Olga Sorkine und Marc Alexa. FiberMesh: designing freeform surfaces with 3D curves. *ACM Trans. Graph.*, 26(3):41, 2007.
- [SG97] Zsolt Szalavári und Michael Gervautz. The personal interaction panel - a two-handed interface for augmented reality. In *Computer Graphics Forum*, Seiten 335–346, 1997.
- [Sub00] Sriram Subramanian. Survey and classification of spatial object manipulation techniques. In *Proceedings of OZCHI 2000, Sydney, Australia*, Seiten 330–337, 2000.
- [TFK⁺02] Michael Tsang, George W. Fitzmaurice, Gordon Kurtenbach, Azam Khan und Bill Buxton. Boom chameleon: simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. In *UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology*, Seiten 111–120, New York, NY, USA, 2002. ACM.
- [Zha99] Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1:666–673 vol.1, 1999.

A Framework for Geometry Processing based on Hybrid Surface Representations

Marcel Campen

Marcel.Campen@rwth-aachen.de

Abstract: We present a framework that allows for the composition of custom-tailored data structures for hybrid representation of geometry and supports the development of associated geometry processing methods. Besides others, a novel hybrid approach for the evaluation of Boolean expressions on polygon meshes is elaborated in this context.

1 Introduction

When dealing with three-dimensional geometric objects in computational applications, there are some fundamentally different approaches for representation of these objects. In widespread use are the explicit or parametric form that, for instance, represents an object by a polygonal mesh and the volumetric form that, e.g., represents it by inside/outside, surface/non-surface or distance information on a number of discrete positions in space.

For various processing operations that can be applied to geometric objects there exist algorithms for both types of representation. While those algorithms that operate directly on an explicit representation usually yield very accurate results, their design is often quite cumbersome due to the necessary handling of many case distinctions and dealing with numerical inaccuracies and resulting inconsistencies. Hence those algorithms often either lack robustness and sometimes fail to generate correct output, or are highly complex [Hof01]. Those algorithms that operate on a volumetric geometry representation, on the other hand, show quite different advantages and disadvantages. Their conceptual design normally is fairly comprehensible and robustness issues rarely have to be taken into account. But due to the necessary use of a finite spatial resolution in the volumetric representation, the resulting output often contains aliasing artifacts and lacks sharp and thin geometric features, hence differs from the exact result. Considering the case of the input being a polygon mesh, we hence have the choice to either directly apply an algorithm of the first kind and live with its advantages and disadvantages, or first convert it into a volumetric representation and apply an algorithm of the latter kind, bearing other pros and cons.

This apparently necessary compromise between accuracy and robustness or complexity, that arises when polygon meshes are dealt with, may be circumvented by using methods that rely on hybrid representations of geometry. These try to combine the accuracy of the direct explicit methods with the robustness and simplicity of the volumetric methods while leaving the disadvantages of either category mainly behind.

2 Framework of Hybrid Representations

Fundamental to the aforementioned hybrid methods is an appropriate data structure that – as opposed to a simple polygon mesh or voxel grid – carries two kinds of information, explicit and volumetric, thus embeds both representations in a hybrid way. Of course, different algorithms can impose completely different requirements on their underlying data structure. Hence it is not straightforward to design *the* universal hybrid representation that fits every need. Such a fully versatile, thus overloaded structure would probably be quite ponderous, leading to bad performance in many applications. By combining the elements of both approaches in specific ways, dependent on the particular purpose, certainly more efficiency can be achieved. Our framework enables the user to easily compose such custom-tailored hybrid data-structures out of pre-defined components as required.

Notwithstanding the mentioned differences, hybrid geometry representations as well as geometry processing algorithms that operate on them still have particular commonalities. Explicit information like polygonal meshes as well as volumetric information like distance fields need to be managed, interlinked, and modified. Our framework allows these tasks common to many hybrid algorithms to be performed through a high-level interface from the user's point of view. The low-level details are hidden in order to allow for a fast and convenient development and implementation of hybrid geometry processing algorithms using the framework. It provides additional functionalities and tools that allow for generation, modification, processing and conversion of the hybrid information. Hence a variety of hybrid algorithms shall be able to be developed faster, safer, and more conveniently than by creating specialized implementations from scratch. Due to the abovementioned particularities and the wide range of variations among the hybrid representations and tools that are required by different algorithms, we designed the framework as flexible and as adaptable as possible, while still maintaining performance and narrowing the overhead that is usually introduced by generalization.

The main properties of the framework can be outlined as follows:

Flexibility Using a layered architecture in conjunction with techniques of generic programming, flexibility and genericity can be achieved without runtime overhead. By employing a traits mechanism and a dynamic property management system, the data structure and its behavior can be adjusted and adapted to specific needs.

Adaptivity The use of adaptive space partitioning schemes like octrees [SW88] for the storage of the geometry information greatly enhances scalability compared to simple uniform voxel grids.

Functionality Modular tools for the generation, modification, processing, and conversion of hybrid geometry information are provided for convenient development of hybrid algorithms. Examples are tools for the rasterization of objects, the extraction of meshes from a hybrid representation, requisite distance and intersection computations, and morphological operations.

3 Hybrid Geometry Processing

Boolean Operations A very common method used in CAGD and CAM is the creation of complex models out of simple basic objects by combining these in a proper way by Boolean operations. We developed a novel hybrid algorithm that performs such operations on solids represented by triangle meshes [PCK09]. The framework is employed to manage the hybrid surface information and the provided functionalities and tools are used to process it. By exploiting this hybrid information the algorithm is able to perform the required operations robustly as well as quite accurately.

Our method is hybrid in two different aspects. First, we use a hybrid geometry information data structure to represent the objects. This constitutes an extensive foundation for robust and accurate computations. Second, the output mesh is composed of parts created by volumetric extraction methods and parts preserved from the explicit input representation.

Figure 1 visualizes the main processing steps of our method on an example.

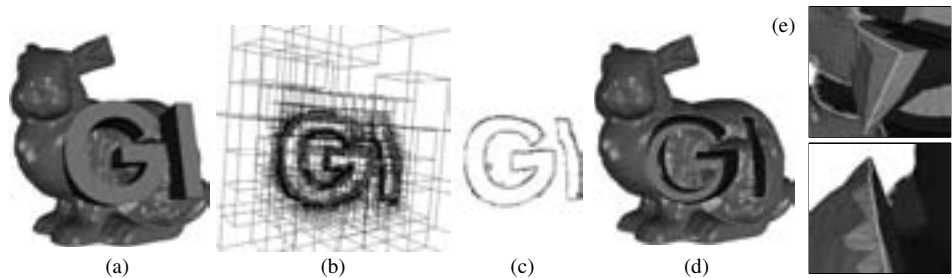


Figure 1: Two letters are to be subtracted from the bunny (a). First an octree is built, adapting to the intersection areas (b). The Boolean expression is evaluated on the hybrid representation and critical areas (voxels with ambiguous geometry information) are determined (red). New mesh parts are generated in these critical areas by volumetric extraction [JLSW02] (c). Input meshes are clipped and connected to these parts in order to obtain the output mesh (d). Sharp and thin geometric features (yellow) are respected by application of a feature-sensitive sampling technique (e).

By relying on the voxel topology, we are able to evaluate expressions robustly and without complex case distinctions, generating manifold and watertight output. By exploiting the explicit information, we achieve high accuracy, preserve the input structure and respect sharp geometric features that are newly introduced during the processing. Assuming r^3 to be the effective voxel resolution, runtime scales almost linearly with r due to adaptivity.

Further Methods A number of recently presented geometry processing algorithms can be classified as hybrid algorithms due to their exploitation of both types of geometry information in some way. Examples are algorithms which compute *offset surfaces* [PK08] and algorithms that *repair inconsistent meshes* [BK05, BPK05]. We adopted these approaches and transferred them to our hybrid framework. By further exploitation of the hybrid information, we are able to enhance the produced results, especially the quality of sharp geometric features in the output surface.

References

- [BK05] Stephan Bischoff and Leif Kobbelt. Structure Preserving CAD Model Repair. *Computer Graphics Forum*, 24(3):527–536, 2005.
- [BPK05] Stephan Bischoff, Darko Pavić, and Leif Kobbelt. Automatic restoration of polygon models. *ACM Trans. Graph.*, 24(4):1332–1352, 2005.
- [Hof01] Christoph M. Hoffmann. Robustness in Geometric Computations. *J. Comput. Inf. Sci. Eng.*, 1(2):143–155, 2001.
- [JLSW02] Tao Ju, Frank Losasso, Scott Schaefer, and Joe Warren. Dual contouring of hermite data. In *SIGGRAPH '02 Proceedings*, pages 339–346, New York, NY, USA, 2002. ACM Press.
- [PCK09] Darko Pavić, Marcel Campen, and Leif Kobbelt. Hybrid Booleans. *submitted*, 2009.
- [PK08] Darko Pavić and Leif Kobbelt. High-Resolution Volumetric Computation of Offset Surfaces with Feature Preservation. *Computer Graphics Forum*, 27(2):165–174, 2008.
- [SW88] Hanan Samet and Robert E. Webber. Hierarchical Data Structures and Algorithms for Computer Graphics. Part I. *IEEE Comput. Graph. Appl.*, 8(3):48–68, 1988.

Analyse und dynamische 3D-Visualisierung des Blutflusses von zerebralen Gefäßstrukturen unter Verwendung von 3D- und 4D-Magnetresonanztomographie-Bildfolgen

Nils Daniel Forkert

Universitätsklinikum Hamburg Eppendorf
Institut für Medizinische Informatik
n.forkert@uke.uni-hamburg.de

Art der Arbeit: Diplomarbeit am Department Informatik der Universität Hamburg

Betreuer der Arbeit: Dr. Dennis Säring, Prof. Heinz Handels und Prof. Dietmar Möller

Abstract: Im Rahmen dieses Beitrages wird eine Methode zur dynamischen 3D-Visualisierung des zerebralen Blutflusses (Hämodynamik) mittels Kombination von 3D- und 4D-MRA-Bilddaten präsentiert. Hierzu wird in einem ersten Schritt das zerebrale Gefäßsystem aus den hochaufgelösten 3D-MR-Bilddaten mittels Fuzzy-Methoden extrahiert und ein individuelles 3D-Oberflächenmodell erzeugt. In einem weiteren Schritt wird für jeden Voxel aus den räumlich-zeitlichen MRA-Bilddaten mittels linearer Kurvenanpassung ein Einflusszeitpunkt bestimmt. Nach einer nicht-linearen Registrierung der MRA-Datensätze können die extrahierten Einflusszeitpunkte direkt auf das Oberflächenmodell übertragen und dort farbkodiert dynamisch visualisiert werden. Die entwickelte dynamische 3D-Visualisierung des zerebralen Blutflusses auf einem interaktiv-beweglichen 3D-Modell des Gefäßsystems wurde von erfahrenen Neuro-radiologen als Gewinn für die diagnostische Unterstützung bei der Beurteilung von zerebralen Gefäßmissbildungen gegenüber konventionellen Methoden eingestuft.

1 Einleitung

Der Schlaganfall stellt mit einem Anteil von 11,4% die dritthäufigste Todesursache in Deutschland dar. Dieser wird zu 80% durch eine Minderdurchblutung des Gehirngewebes und zu 20% durch die akute Hirnblutung hervorgerufen. Die Hirnblutungen lassen sich in den meisten Fällen auf Rupturen von zerebralen Blutgefäßmissbildungen, wie z.B. Aneurysmen und arteriovenöse Malformationen, zurückführen. Bei der Risikoeinschätzung und Therapieplanung von Gefäßmissbildungen ist ein genaues Verständnis der individuellen Gefäßanatomie und des Blutflusses (Hämodynamik) von essenzieller Bedeutung. Zur Beurteilung der individuellen Hämodynamik ist die kontrastmittelbasierte digitale Subtraktionsangiographie (DSA) derzeit der Goldstandard. Diese Technik ermöglicht die Aufnahme einer Folge von 2D-Projektionsbildern mit einem zeitlichen Abstand von bis zu 0,2 Sekunden. Neben der hieraus resultierenden fehlenden Tiefeninformation für die Ana-

lyse, stellen die verwendeten ionisierenden Röntgenstrahlen, sowie Komplikationen bei der invasiven Aufnahme-prozedur eine Gefahr für den Patienten dar. Neue nicht-invasive 3D- und 4D-Magnetresonanz-Angiographien (MRA) liefern räumliche (3D) und räumlich-zeitliche (4D) Bildinformationen, welche eine Einschätzung von hämodynamischen Prozessen ermöglichen und das Risiko für den Patienten erheblich reduzieren.

2 MR-Bildsequenzen

Mit Hilfe von neuen parallelen Bildgebungs- und Echo-Sharing-Techniken, können zeitaufgelöste kontrastmittelgestützte 4D-TREAT-Sequenzen (time-resolved echo-shared angiography technique) erzeugt werden, welche die Grundlage für die zeitliche Analyse der Hämodynamik bilden. Die visuelle Bildqualität ist bei den 4D-TREAT-Bilddaten mit einer zeitlichen Auflösung von zwei 3D-Bildsequenzen pro Sekunde und einer Voxelgröße von $1,875 \times 1,875 \times 5,0$ mm eher gering. Daher werden zusätzlich räumlich hochauflösende 3D TOF-MRA-Bildsequenzen (time-of-flight) aufgenommen, welche durch einen verbesserten Blut-zu-Hintergrund-Kontrast und eine geringe Voxelgröße von $0,469 \times 0,469 \times 0,5$ mm eine detaillierte Extraktion des Gefäßsystems ermöglichen.

3 Methoden

3.1 Extraktion des Gefäßsystems

Zur Extraktion des zerebralen Gefäßsystems aus den hochauflösenden TOF-Bildsequenzen wird zunächst aus dem TOF-Datensatz ein Vesselness- [SNS⁺98] und Maximum-Parameterbild berechnet. Diese werden dann unter Verwendung eines Fuzzy-Inferenzsystems mit dem TOF-Datensatz kombiniert. In dem daraus resultierenden Datensatz heben sich sowohl kleine als auch pathologisch veränderte Gefäße deutlicher vom restlichen Gewebe ab. Abschließend wird das Gefäßsystem unter Verwendung des Fuzzy-Connectedness-Regionenwachstums [US96] extrahiert (Abb. 1(a)). Eine detaillierte Beschreibung und Evaluation der Methodik kann [FSW⁺09, For08] entnommen werden. Das segmentierte Gefäßsystem kann dazu verwendet werden, um mit Hilfe des Marching Cubes Algorithmus [LC87] ein interaktiv bewegliches 3D-Oberflächenmodell (Abb. 1(b)) zu erzeugen.

3.2 Analyse der Hämodynamik

Für die Definition eines Einflusszeitpunktes anhand zeitlicher Intensitätsverläufe wurden viele Methoden vorgestellt [SCVD00]. Diese Methoden (z.B. Gamma Variate Funktionen) basieren jedoch vorwiegend auf röntgenbasierter Bildgebung und führten bei den vorliegenden 4D-TREAT-Bildfolgen zu unrealistischen Zeitabständen zwischen den Einfluss-

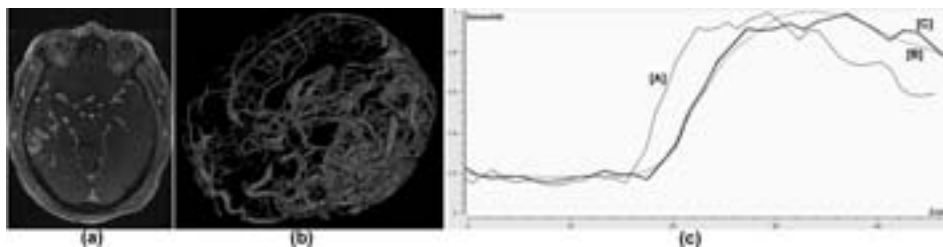


Abbildung 1: TOF-Schicht mit extrahiertem Gefäßsystem (a), 3D-Gefäßmodell (b), zeitlicher Intensitätsverlauf [A], Referenzsignalverlauf [B], angepasster zeitlicher Intensitätsverlauf [C] (c)

zeitpunkten zweier direkt benachbarter Voxel. Um diese Artefakte zu reduzieren wurde daher eine Methode entwickelt, die auf der Kurvenanpassung eines zeitlichen Signalverlaufs und einer Referenzkurve basiert [FSE⁺07, For08]. Hierbei wird zunächst ein für Blutgefäße typischer Referenzsignalverlauf patientenindividuell extrahiert. In einem nächsten Schritt werden die zeitlichen Intensitätsverläufe der 4D-MRA-Bildsequenz analysiert und voxelweise eine Transformation bestimmt, unter der sich der gegebene Intensitätsverlauf optimal, unter Berücksichtigung von Verschiebung und Skalierung, an den extrahierten Referenzsignalverlauf anpasst (Abb. 1(c)). Die Verschiebung dient als relativer Einflusszeitpunkt. Mittels einer entwickelten B-Spline basierten Registrierung der MRA-Datensätze [FSF⁺09, For08] können die Einflusszeitpunkte direkt auf das Oberflächenmodell übertragen und dort farbkodiert dynamisch visualisiert werden (Abb. 2).

4 Ergebnisse

Die Ergebnisse der entwickelten Methode zur Definition eines Einflusszeitpunktes anhand zeitlicher Intensitätsverläufe mittels Kurvenanpassung wurden von erfahrenen Neuroradiologen qualitativ besser beurteilt, als die mittels konventioneller Methoden erzielten. Wesentlicher Vorteil der Methode ist, dass die Anzahl von starken zeitlichen Differenzen zwischen den Einflusszeitpunkten zweier benachbarter Voxel deutlich verringert werden konnte. Die entwickelte dynamische 3D-Visualisierung des zerebralen Blutflusses auf einem interaktiv beweglichen Modell des Gefäßsystems trägt wesentlich zur diagnostischen Unterstützung bei der Beurteilung von zerebralen Gefäßmissbildungen bei.

5 Diskussion

Die entwickelte Methode zur Analyse des Blutflusses durch lineare Kurvenanpassung wurde bereits in ersten medizinischen Studien verwendet (z.B. [FIP⁺08] und [IFS⁺08]) und kann in weiteren geplanten Studien neue Erkenntnisse für die Bewertung von zerebralen Gefäßkrankheiten liefern. Die dynamische 3D-Visualisierung des zerebralen Blutflusses erlaubt eine verbesserte und intuitive Darstellung der hämodynamischen Verhältnisse.

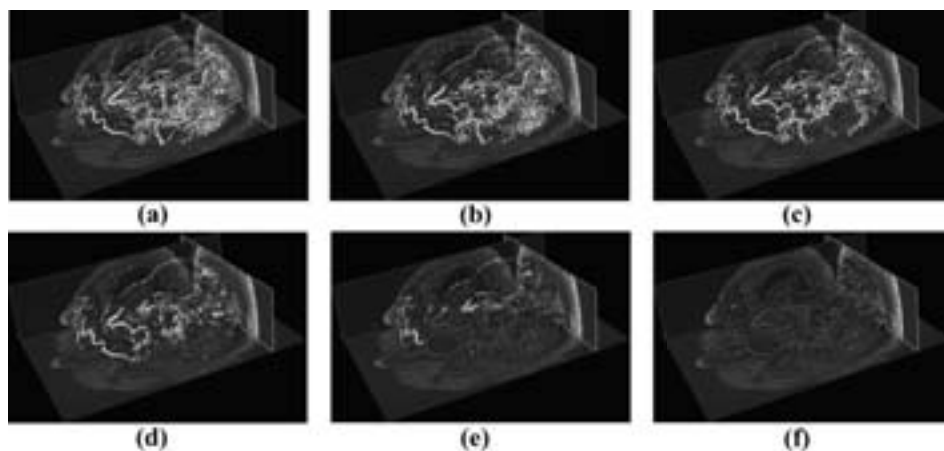


Abbildung 2: Ausgewählte Frames der dynamischen 3D-Visualisierung des zerebralen Blutflusses

Literatur

- [FIP⁺08] J. Fiehler, T. Illies, M. Piening, D. Säring, N. Forkert, U. Grzyska, H. Handels und J. Byrne. Territorial and microvascular perfusion impairment in patients with cerebral arteriovenous malformations. *American Journal of Neuroradiology*, 2008. (akzeptiert).
- [For08] N. Forkert. Analyse der Hämodynamik und Segmentierung von zerebralen Gefäßstrukturen durch Kombination von 3D- und 4D-MRT-Bildsequenzen, Diplomarbeit, Universität Hamburg, Department Informatik, 2008.
- [FSE⁺07] N. Forkert, D. Säring, J. Ehrhardt, J. Fiehler, M. Piening und H. Handels. Analyse der Hämodynamik von zerebralen arteriovenösen Malformationen in 4D-MR-Bilddaten mittels referenzbasierter Kurvenpassung. *Biomedizinische Technik*, 52, 2007.
- [FSF⁺09] N. Forkert, D. Säring, J. Fiehler, T. Illies, D. Möller und H. Handels. Analysis and Dynamic 3D Visualization of Cerebral Blood Flow Combining 3D and 4D MR Image Sequences. *Proceedings of SPIE Medical Imaging, Orlando*, 2009. akzeptiert.
- [FSW⁺09] N. Forkert, D. Säring, K. Wenzel, T. Illies, J. Fiehler, D. Möller und H. Handels. Automatische Segmentierung der zerebralen Gefäße aus 3D-TOF-MRA-Bildsequenzen mittels Fuzzy-Methoden. In *Bildverarbeitung für die Medizin 2009*, 2009. (akzeptiert).
- [IFS⁺08] T. Illies, N. Forkert, D. Säring, J. Regelsberger, M. Westphal, U. Grzyska und J. Fiehler. Haemorrhage Risk in Brain Arteriovenous Malformations (AVMS) is Related to High Flow in Deep Venous Structures. *Neuroradiology*, 50 Supplement:68, 2008.
- [LC87] W.E. Lorensen und H.E. Cline. Marching cubes: A High Resolution 3D Surface Construction Algorithm. *Computer Graphics*, 21(4):163–169, 1987.
- [SCVD00] S. Shpilfoygel, R. Close, D. Valentino und G. Duckwiler. X-ray videodensitometric methods for blood flow and velocity measurement: a critical review of literature. *Med. Phys.*, 27(9):2008–2023, 2000.
- [SNS⁺98] Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig und R. Kikinis. Three-dimensional multi-scale Line Filter for Segmentation and Visualization of Curvilinear Structures in Medical Images. *Med Image Anal*, 2(2):143–168, 1998.
- [US96] J. K. Udupa und S. Samarasekera. Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation. *Graphical Models and Image Processing*, 58(3):246–261, 1996.

Ein Open Source Framework für iterative Bildsegmentierung mit Graph Cuts

Stephan Jonas
stephan.jonas@rwth-aachen.de

1 Einleitung

Interaktive Bildsegmentierung, in diesem Falle die Extraktion des Vordergrundes von beliebigem Hintergrund, ist ein oft komplizierter und aufwendiger Prozess. Werkzeuge wie beispielsweise die “Intelligente Schere“ oder der “Zauberstab“, die in aktuellen Bildbearbeitungsprogrammen enthalten sind, liefern in den meisten Fällen erst nach einer Vielzahl von Benutzerinteraktionen ein befriedigendes Ergebnis.

Ein Ansatz um interaktive Segmentierung mit minimaler Interaktion durchzuführen, wurde 2001 von Boykov und Jolly [BJ01] eingeführt und von Rother et al [RKB04] weiterentwickelt. Bei diesem sogenannten Graph Cut Ansatz wird ein Modell anhand markierter Bildbereiche erstellt und mit diesem der noch unmarkierte Teil vollständig segmentiert. Im besten Fall kann die Benutzerinteraktion auf nur zwei Mausklicks reduziert werden.

2 Aufgabenstellung

Die zentrale Aufgabe, der sich diese Arbeit widmet, ist die Implementierung eines Frameworks für Bildsegmentierung mittels Graph Cuts sowie ein Border Matting um die Übergänge von Vorder- zu Hintergrund zu glätten. Das Framework soll eine graphische Benutzeroberfläche (GUI) für interaktive Segmentierung einzelner Bilder sowie eine Möglichkeit zur Stapelverarbeitung bieten. Die Software soll eine möglichst geringe Interaktion erfordern, gleichzeitig aber erfahreneren Benutzern die Möglichkeit geben verschiedene Modelle für die Segmentierung auszuwählen und die Ergebnisse maximal zu beeinflussen. Die Software wurde für den medizinischen Bereich entwickelt und soll möglichst universell einsetzbar sein, also sowohl bei Graubildern (Röntgen, MRT, CT) als auch bei normalen Fotografien (Haut, Personen, etc.) gute Ergebnisse erzielen. Aus diesem Grunde soll kein Vorwissen über mögliche Segmentierungsaufgaben in die Modelle eingehen.

Die Umsetzung soll plattformunabhängig und leicht erweiterbar implementiert werden und als Open Source Werkzeug zugänglich gemacht werden.

3 Theorie

Der Graph Cut Ansatz basiert auf dem *Max-Flow=Min-Cut* Theorem. Dieses Theorem besagt, dass ein Graph durch einen maximalen Fluss von einem Quell- zu einem Senkenknoten an einen entstehenden Flaschenhals in zwei Teile geteilt wird. Der Quell- und der

Senkenknoten können sich hierbei nicht im selben Teil des Graphen befinden. Der Graph - bzw. das sog. Flussnetzwerk - auf dem ein maximaler Fluss/minimaler Schnitt berechnet wird, hat gewichtete Kanten, die der maximalen Kapazität eines Flusses durch diese Kante entsprechen. Das Problem des maximalen Flusses kann mit einem Röhrennetzwerk mit einem Ein- und einem Ausgang verdeutlicht werden, innerhalb dessen verschiedene Röhren verschiedene Durchmesser und somit verschiedene Flussgeschwindigkeiten (Kapazitäten) haben. Versucht man einen maximalen Durchfluss durch das Röhrennetzwerk herzustellen, so entsteht zwischen dem Ein- und Ausgang ein Flaschenhals. Dieser Flaschenhals entspricht dem minimalen Schnitt durch das Netzwerk. Hierbei bedeutet ein höheres Gewicht auf einer Kante, also eine höhere Maximalkapazität, im Allgemeinen, dass der minimale Schnitt an dieser Stelle weniger wahrscheinlich auftritt.

Um diesen Ansatz für die Segmentierung von Vorder- und Hintergrund verwenden zu können, muss ein zu segmentierendes Bild in ein Flussnetzwerk umgewandelt werden. Zuerst wird jedem Pixel ein Knoten zugeordnet und dem Graphen ein Quell- und ein Senkenknoten hinzugefügt. Der Knoten jedes Pixels wird dann mit den Knoten der Vierzelnachbarschaft des Pixels (also die Pixel links, rechts, oben- und unterhalb des Ausgangspixels) und den Quell- und Senkenknoten durch eine Kante verbunden. Die erstgenannte Art von Kanten wird im folgenden Pixel-Kanten und die zweite Q/S-Kanten genannt. Nach einem minimalen Schnitt durch das Netzwerk werden alle Knoten, die zu dem Quellen-Teil des Netzwerkes gehören als Vordergrund und alle anderen als Hintergrund segmentiert. Alle Pixel, die zum Hintergrund gehören können dann ausgeblendet werden.

Um die Pixel-Kanten zu gewichten, wird ein Ähnlichkeitsmaß für zwei Pixel benötigt, da zwei identische benachbarte Pixel wahrscheinlich auch zu demselben Bildsegment (also Vorder- oder Hintergrund) gehören, und unähnliche benachbarte Pixel im Gegenzug wahrscheinlicher zu zwei verschiedenen Segmenten. Als Ähnlichkeitsmaß \hat{D} zur Gewichtung der Pixel-Kanten zwischen zwei Pixeln p_1 und p_2 wird in dieser Arbeit eine invertierte Euklidische Distanz verwendet. Die Invertierung ist wie in Formel 1 als der negative Exponent der Euklidischen Distanz definiert. Der Faktor f normalisiert die Distanz zwischen 0 und 1.

$$\hat{D}(p_1, p_2) = e^{-f \times \text{euk}(p_1, p_2) * f} \quad (1)$$

Die Gewichtung der Q/S-Kanten erfolgt durch ein statistisches Modell. Diese Modelle werden durch eine Vormarkierung des Bildes trainiert und berechnen dann für jeden Pixel die Wahrscheinlichkeit, dass selbiger Teil des Vordergrundes ist. Diese Wahrscheinlichkeit wird als Gewicht der Kante zur Quelle verwendet und die inverse Wahrscheinlichkeit für die Gewichtung der Kante zur Senke. In dieser Arbeit werden drei verschiedene Modelle verwendet, die auf den unterschiedlichen Verfahren *Nächster-Nachbar-Klassifikator*, *Gauß-Mischverteilungen* und *Neuronale Netzwerke* basieren. Die Funktionsweise der einzelnen Verfahren kann hier nicht näher beschrieben werden, für weitere Informationen sei beispielsweise auf Duda und Hart [DHS01] verwiesen.

Um die Ränder zwischen Vorder- und Hintergrund der durch den Graph Cut ermittelten harten (binären) Segmentierung zu glätten wird ein so genanntes Border Matting verwendet. Bei diesem Verfahren wird ein Korridor um die Ränder der Segmentierung gelegt und

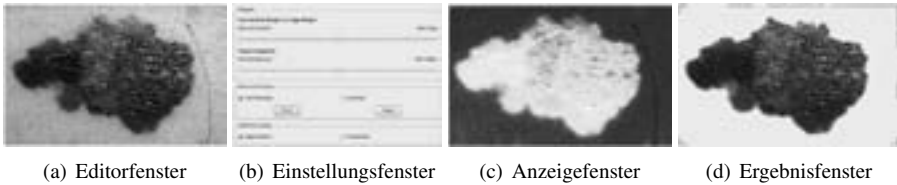


Abbildung 1: Unterfenster der GUI

in diesem Korridor die Segmentierung aufgehoben. Anschließend wird ein Modell auf den noch vorhandenen Vorder- und Hintergrundpixeln geschätzt. Anhand des Modells wird für jeden Pixel innerhalb des Korridors eine Vordergrundwahrscheinlichkeit berechnet. Diese wird in Transparenz umgerechnet, d.h. je wahrscheinlicher ein Pixel zum Vordergrund gehört, desto weniger transparent ist er.

4 Implementierung

Die Software wurde mit C++ implementiert. Portabilität wird durch die Erstellung der GUI mit QT [Tro09] gewährleistet. Für zwei der Modelle und die Berechnung des minimalen Schnittes werden Open Source Lösungen eingesetzt [Des09, Nis09, Kol09].

Die GUI teilt sich in vier Unterfenster, wobei die Interaktion mit dem Benutzer lediglich in den beiden oberen stattfindet, die beiden unteren dienen der Anzeige. Abbildung 4 zeigt die vier Unterfenster mit einem Beispiel. Im Editorfenster - abgebildet in Teil a) - wird die Segmentierung erstellt und verfeinert. Um eine erste Segmentierung zu berechnen muss ein Teil des Bildes als Vordergrund und ein Teil als Hintergrund markiert werden. Dies kann entweder mit einem Pinsel oder durch Ziehen eines Rechteckes um den Vordergrund durchgeführt werden. Anhand dieser Benutzereingabe wird ein Modell erstellt und trainiert. Wenn ein Rechteck über das Objekt gezogen wird, werden die Wahrscheinlichkeiten für alle Pixel neu berechnet, ansonsten ausschließlich für die unmarkierten Pixel, da die anderen durch die Markierung bereits eindeutig Vorder- oder Hintergrund zugeordnet sind. Wenn ein Ergebnis berechnet ist, wird dieses als Overlay über das vorhandene Bild im Editorfenster gelegt. In diesem kann dann mit dem Pinsel das Ergebnis weiter verbessert und die Modelle mit den zusätzlichen Markierungen weiter trainiert werden. Alternativ zu dem iterativen Markierungsprozess kann das Ergebnis auch über Veränderung der Kantengewichte im Flussnetzwerk des Graph Cut Algorithmus verbessert werden. Dazu dienen die Schieberegler im Unterfenster b). Der erste Regler gewichtet das Verhältnis der Gewichte der Q/S-Kanten zu denen der Pixel-Kanten. Wenn die Q/S-Kanten stärker gewichtet werden wird der Vordergrund weniger zusammenhängend und die Pixel werden aufgrund ihrer Ähnlichkeit zu Vorder- und Hintergrund zugeordnet. Wenn die Gewichte der Pixel-Kanten erhöht werden, wird der Vordergrund zusammenhängender und Löcher werden geschlossen. Der zweite Regler ändert das Verhältnis von Vorder- zu Hintergrundwahrscheinlichkeiten, also tritt der stärker gewichtete Teil häufiger auf. Im Unterfenster b) lassen sich zusätzlich die Eingabemethode für das Editorfenster und die Anzeige im Unterfenster c) steuern. Dort kann entweder die Segmentierung als Heatmap oder die für die

Q/S-Kanten verwendeten Wahrscheinlichkeiten angezeigt werden. Das letzte Unterfenster d) zeigt das Segmentierungsergebnis, also nur den Vordergrund, an.

5 Experimente und Diskussion

Um die Software zu testen wurden Bilder nach verschiedenen Kriterien ausgesucht und anhand der Software segmentiert. Die Universalität der Software wird beispielsweise überprüft, indem Bilder mit vielen gleichen Objekten gewählt, von denen lediglich eines als Ergebnis gewünscht war (bspw. ein einzelner Fingerknochen in einem Röntgenbild). Auch die Gegenprobe wurde durchgeführt, indem überprüft wurde, ob alle Objekte gefunden werden können, wenn nur ein einzelnes markiert wird. Weitere Bilder wurden zum Teil der Medizin entnommen (z.B. Segmentierung von Gewebe in MRT Bildern, Leberflecken auf Haut), aber auch aus alltäglichen Aufnahmen aus dem Internet (Fotos von Tieren, Menschen, etc.). In vielen Fällen konnte die Software innerhalb weniger Iterationen eine gute Segmentierung finden. Lediglich bei manchen schweren Fällen, wie beispielsweise der Segmentierung eines Tigers vor heterogenem Hintergrund konnte kein zufriedenstellendes Ergebnis erzielt werden. Da bisher keine Datenbank zur Evaluierung von "weicher Segmentierung" zur Verfügung steht, fehlt jedoch eine objektive Bewertung.

Verbesserungspotential für die Software ergibt sich in verschiedenen Bereichen. Da der Schwerpunkt dieser Arbeit im Bereich der Programmierung auf der Erweiterbarkeit und Wartbarkeit der Software lag, sind Speicher- und Laufzeitoptimierungen möglich. Auch die Auswahl der Modelle und der für die Modelle verwendeten Features kann erweitert werden.

Als wichtigste Verbesserungsmöglichkeit des Graph Cut Algorithmus wird jedoch die Erweiterung auf eine Segmentierung im 3D Raum gesehen. Diese Erweiterung ermöglicht sowohl eine Volumensegmentierung, als auch die Segmentierung einzelner Szenen in Videosequenzen vergleichbar zur Blue-Screen Methode.

Literatur

- [BJ01] Y. Boykov und M.P. Jolly. Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in n-D Images. *ICCV*, 2001.
- [Des09] Thomas Deselaers. Flexible Image Retrieval Engine, Januar 2009. www-i6.informatik.rwth-aachen.de/deselaers/fire.
- [DHS01] R. Duda, P. Hart und D. Stork. *Pattern Classification*. Wiley-Interscience, second edition. Auflage, 2001.
- [Kol09] V. Kolmogorov. Max-Flow/Min-Cut V2.2, Januar 2009. www.adastal.ucl.ac.uk/vladkolm/software.html.
- [Nis09] Steffen Nissen. Fast Artificial Neural Network, Januar 2009. leenissen.dk/fann.
- [RKB04] C. Rother, V. Kolmogorov und A. Blake. "GrabCut Interactive Foreground Extraction using Iterated Graph Cuts. In *Siggraph*, 2004.
- [Tro09] Trolltech/Nokia. QT, Januar 2009. trolltech.com/products.

Automatische Detektion primärer Neuronen in Fluoreszenzmikroskopbildern

Danny Misiak

Martin-Luther-Universität Halle-Wittenberg, Institut für Informatik
danny.misiak@informatik.uni-halle.de

1 Motivation und Zielsetzung

In den biomedizinischen Wissenschaften kommen aufgrund des heutigen Technikstandards immer mehr bildgebende Verfahren zum Einsatz, deren Ergebnisse verarbeitet und ausgewertet werden müssen. Der zeitliche Aufwand spielt dabei eine sehr große Rolle, da es bei dem heutigen Stand der Technik keine Seltenheit mehr ist, enorme Mengen an Daten zu produzieren. Das erschwert die manuelle Auswertung zusätzlich und macht den Einsatz halbautomatischer Programme fast unmöglich. Um diese Aufgaben zu erfüllen müssen Anwendungen und Algorithmen entworfen werden, die in der Lage sind, mit den neuen Problemen umzugehen und den Wissenschaftlern und Forschern eine effiziente, anwendungsbezogene und vollautomatische Auswertung der Bilder ermöglichen.

Für eine große Reihe von Anwendungen in Biologie und Medizin existieren Programme, die eine Bearbeitung, Verarbeitung und Auswertung der Bilder unterstützen. Eines dieser Programme, das gerade in den biomedizinischen Wissenschaften zum alltäglichen Werkzeug geworden ist, ist das Java basierte *ImageJ* [Ima]. Das Programm verfügt über eine graphische Benutzeroberfläche, die eine Reihe von nützlichen Plugins und Werkzeugen mitbringt, speziell im Hinblick auf wissenschaftliche Analysen. ImageJ ist *Open Source* und somit für jeden frei verfügbar. Die große Verbreitung der Software begründet sich, neben der Plattformunabhängigkeit, vor allem in der Erweiterungsmöglichkeit der Funktionalität durch die Einbindung eigener oder frei verfügbarer Plugins.

In Zusammenarbeit mit der Nachwuchsgruppe „Posttranskriptionelle Genregulation und Karzinogenese“ von Dr. Stefan Hüttelmaier¹ soll in dieser Arbeit ein automatisches Verfahren auf Grundlage eines ImageJ-Plugins entwickelt werden, mit dem es möglich ist, die Konturen primärer Neuronen in Fluoreszenzmikroskopbildern automatisch zu detektieren. Anhand dieser Konturen können dann weitere biologische oder biochemische Analysen erfolgen, wie z.B. Untersuchungen zum Wachstumsverhalten der Neuriten oder die Quantifizierung von Proteinen innerhalb der Neuronen. Da es sich um eine vollautomatische Anwendung handeln soll, sind im Gegensatz zu aktuellen Anwendungen, wie [MJS⁺04] oder [Lon], keine Interaktionen von Seiten der Wissenschaftler erwünscht.

¹Zentrum für angewandte medizinische und humanbiologische Forschung an der Martin-Luther-Universität

2 Hierarchischer Detektionsansatz

Ausgangspunkt für diese Arbeit stellen Fluoreszenzaufnahmen von primären Neuronen dar, in denen verschiedene Proteine immunhistochemisch mittels Antikörpern angefärbt wurden. Als Ergebnis und Grundlage für die weitere Analyse ergeben sich 4-Kanal-Bilder, in denen jeweils eine fluoreszenzmarkierte Komponente dargestellt ist. Für die automatische Konturdetektion wird ein hierarchischer Ansatz aus mehrere bildverarbeitende Methoden gewählt.

Binarisierung. Der erste Schritt ist die Binarisierung der Neuronen und ihrer Zellkerne, also die Trennung der Objekte im Bild vom Hintergrund der Mikroskopaufnahme. Diese Transformation des Grauwertbildes in ein Binärbild geschieht durch die Anwendung sogenannter Schwellwertoperationen, mit deren Hilfe ein Schwellwert für das Intensitätshistogramm der Grauwerte festgelegt werden kann. Die Ermittlung der Schwellwerte erfolgt für die Zellkerne über das globale Schwellwertverfahren von Otsu [Ots79], während beim Cytoskelett zur Binarisierung der Neuronen das lokale Schwellwertverfahren von Niblack [Nib86] zum Einsatz kommt.

Segmentierung. Anhand der Binarisierungsergebnisse können in einem zweiten Schritt Segmentierungen vorgenommen werden, die die Zellregion und die inneren und äußeren Konturen dieser Region ermitteln. Aufgrund auftretender Abweichungen durch Rauschen, schwacher Fluoreszenz oder Überstrahlungseffekten, stimmen diese Regionen jedoch nicht exakt mit den Regionen der realen Zelle überein (vgl. Abb. 1, unteres Bild). Um später korrekte Aussagen, z.B. bei der Quantifizierung von Proteinen, machen zu können, ist eine möglichst genaue Bestimmung der Kontur von Nöten.

Aktive Konturen. Um eine Verbesserung der lokalisierten Konturen zu erzielen wurde ein dynamisches Modell einer parametrisierten Kurve angewendet, sogenannte *aktive Konturen* oder *Snakes*. Ziel dieser Snakes ist das Auffinden von Kanten oder Konturen eines dargestellten Objektes. Das aktive Konturmodell besitzt ein dynamisches Verhalten und versucht schrittweise, seine Energie zu minimieren. Dies führt dazu, dass die Kontur in Bezug auf die Gesamtenergie ein lokales Minimum anstrebt. Die zu minimierende Snake-Energie besteht aus zwei unabhängigen Energien, einer internen Energie, die sich aus der Kurve selbst ermitteln lässt, und einer externen Energie, die sich aus dem betrachteten Bild ergibt:

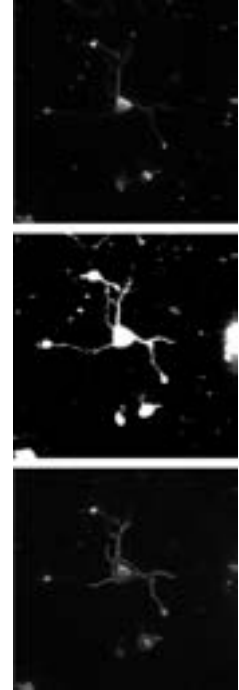


Abbildung 1: oben Fluoreszenzbild eines Neurons, mittig Binarisierungsergebnis, unten Ergebnis der Kontursegmentierung

$$E_{snake}^* = \int_0^1 \underbrace{E_{int}(v(s)) + E_{ext}(v(s))}_{E_{snake}(v(s))} ds . \quad (1)$$

Die interne Energie $E_{int}(v(s))$ beeinflusst die Glattheit und die Elastizität der Kurve $v(s)$, während die externe Energie $E_{ext}(v(s))$ dafür sorgt, dass sich die Snake in ein lokales Minimum bewegen kann. Die externe Energie wird meist als negativer Gradientenbetrag $(-|\nabla I|^2)$ des Bildes I dargestellt, oder als negativer Gradientenbetrag eines gaußgefilterten Bildes $(-|\nabla (G_\sigma * I)|^2)$. Der Nachteil an diesen externen Energien ist deren geringe Reichweite, was eine Initialisierung der Snake nur sehr nah an der wahren Kontur zulässt. Um diesen Nachteil zu beseitigen, wurde eine alternative externe Energie benutzt. Diese Energie beruht auf den sogenannten *Gradient Vector Flow*-Feldern [XP98] - kurz GVF - die es ermöglichen, nicht nur den Gradienten unmittelbar in der Nähe einer Kontur zu berechnen, sondern ihn räumlich nach außen Richtung Bildrand zu propagieren. Folglich wird dadurch der Einzugsbereich der Kontur erhöht und kann sich bis zum Rand des Bildes erstrecken.

Zur Lösung des Optimierungsproblems aus Gleichung (1) muss eine Kurve $\tilde{v}(s)$ gefunden werden, so dass die Snake-Energie E_{snake}^* minimal wird. Als Initialisierung für die Snake dienen die Konturen, die mittels Segmentierung im zweiten Schritt ermittelt wurden. Durch die Anwendung der Variationsrechnung erhält man aus der Snake-Energie zwei unabhängige Euler-Lagrange Gleichungen (jeweils eine für die x- und die y- Richtung), die es zu lösen gilt. Die Lösung erfolgt mittels Diskretisierung, wodurch sich die Euler-Lagrange Gleichungen in Matrixform darstellen lassen und daraus ein neues Gleichungssystem hervorgeht. Da dieses System nicht direkt lösbar ist wurde ein iterativer Lösungsansatz gewählt, der unter Verwendung des impliziten und expliziten Euler-Verfahrens eine direkte Lösung liefert.

Zweistufige Konturextraktion. Zur Detektion der Neuronenkontur wurde ein hierarchisches Verfahren angewendet, welches auf einem zweistufigen Optimierungsansatz beruht. Zweistufig deshalb, weil sich die GVF-Felder je nach Iterationszahl bei ihrer Berechnung unterschiedlich weit ausdehnen und mit steigender Iterationszahl alle Gradientenvektoren auf Punkte mit hohen Intensitätswerten zeigen. Im hierarchischen Ansatz werden zwei GVF-Felder berechnet, eines mit 5 Iterationen und eines mit 100 Iterationen. Das Feld mit 5 Iterationen zeigt die gesuchten Konturen des Neurons auf, allerdings treten in der Nähe dieser Kontur viele lokale Minima durch Hintergrundrauschen auf, in denen die Snake stehen bleibt und die Kontur der Zelle nicht erreichen kann. Das GVF-Feld mit 100 Iterationen glättet dieses Rauschen fast vollständig und propagiert das Feld nach außen in Richtung Bildrand. Jedoch mit dem Nachteil, dass die Konturen der sehr feinen Neuriten verloren gehen und zu einer Konturlinie verschmelzen. Eine Kombination beider Ansätze, in der man zu Beginn auf dem Feld mit 100 Iterationen detektiert und anschließend das Konturergebnis als Initialisierung für das Feld mit 5 Iterationen einsetzt, führt zur gewünschten Detektion der Neuronenkontur. Mit Hilfe dieser Hierarchie von Methoden ist eine automatische Initialisierung der Snakes gewährleistet. Durch die folgende Optimierung der Snake-Energien können die Konturen der Neuronen im Bild ermittelt werden.

3 Ergebnisse

Das entwickelte Plugin wurde auf 15 Testbildern der Arbeitsgruppe getestet und die Ergebnisse wurden anschließend manuell ausgewertet. Erste Ergebnisse der automatischen

Detektion sind auszugsweise in Abb. 2 dargestellt. Wie zu erkennen ist, kommen die aktiven Konturen (weiß) sehr gut auf den wahren Konturen der Neuronen zum liegen. In den ersten beiden Bildern a) und b) erkennt man, dass die Kontur nicht immer bis in die Spitzen der eigentlichen Neuronenausläufer (Neuriten) reicht. Trotzdem kann mit Hilfe des



Abbildung 2: Darstellung der Detektionsergebnisse unter Anwendung aktiver Konturen auf unterschiedlichen Fluoreszenzaufnahmen. Die Initialisierung der aktiven Kontur ist in schwarz, das Ergebnis der Konturbestimmung in weiß dargestellt.

Modells der aktiven Konturen ein gutes Ergebnis bei der Detektion von Neuronen in Fluoreszenzbildern erreicht werden. Das Verfahren ist somit gut für die automatische Detektion der Neuronen geeignet, und kann diese, bis auf wenige Ausnahmen, korrekt lokalisieren.

4 Zusammenfassung und Ausblick

Die Ergebnisse lassen ganz klar erkennen, dass eine automatische Detektion der Neuronen möglich ist und keine Interaktionen durch den Nutzer erfordern. Im Gegensatz zu den aktuell verfügbaren Programmen müssen keine Pfade der Neuriten von Hand nachgezeichnet werden oder eine Vielzahl an Parametern individuell an das Bild angepasst werden. Gerade die Detektion der Neuriten spielt bei der Quantifizierung und Wachstumsuntersuchung eine große Rolle. Auch wenn die hier dargestellte Detektion der Neuriten noch nicht perfekt ist, stellt sie doch eine solide Grundlage für ein integriertes Analysewerkzeug dar.

Literatur

- [Ima] ImageJ. *Image Processing and Analysis in Java*. <http://rsb.info.nih.gov/ij/>.
- [Lon] Mark Longair. Simple Neurite Tracer. <http://homepages.inf.ed.ac.uk/s9808248/imagej/tracer/>, (Stand 15.01.2009).
- [MJS⁺04] E. Meijering, M. Jacob, J.-C.F. Sarria, P. Steiner, H. Hirling und M. Unser. Design and Validation of a Tool for Neurite Tracing and Analysis in Fluorescence Microscopy Images. *Cytometry*, 58A(2):167–176, April 2004.
- [Nib86] Wayne Niblack. *An Introduction to Digital Image Processing*. Prentice Hall, 1986.
- [Ots79] Nobuyuki Otsu. A Threshold Selection Method from Grey-Level Histograms. *IEEE Trans. Systems, Man and Cybernetics*, 9(1):62–66, Januar 1979.
- [XP98] Chenyang Xu und Jerry L. Prince. Snakes, Shapes, and Gradient Vector Flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.

Deferred Shading

Georg Tamm

Sheffield Hallam University, Reutlingen University
Faculty of Arts, Computing, Engineering and Sciences
gilga1@gmx.net

Art der Arbeit: Abschlussarbeit, 2 Monate, Computergrafik, Computerspiele
Betreuer der Arbeit: Paul Parry

Abstract: Beleuchtung ist ein wesentlicher Aspekt, um realistische und hochqualitative Umgebungen in visuellen Anwendungen wie Game Engines zu ermöglichen. Deferred Shading ist eine Rendertechnologie, die es ermöglicht, eine sehr hohe Anzahl an Lichtern innerhalb einer virtuellen Welt einzusetzen. Dieses Projekt stellt die Entwicklung eines leistungsstarken Deferred Renderers auf Basis von DirectX 9 vor.

1 Einleitung

Beleuchtung ist ein wichtiger Bestandteil in grafischen Anwendungen, die eine virtuelle Welt so realistisch und lebendig wie möglich darstellen wollen. Beleuchtungsmodelle wie das Phong-Modell sind rechenintensiv und werden per Pixel angewendet, um die geforderte Qualität zu erreichen. Entsprechend kann das Verwenden von Lichtern in einer Szene großen Einfluss auf die Performance haben.

In diesem Zusammenhang ist es ein wesentliches Problem von konventionellen Forward Renderern, dass Beleuchtungsberechnungen unmittelbar auf Pixel angewendet werden, während die Rasterisierung noch nicht abgeschlossen ist. Dies kann dazu führen, dass schon voll beleuchtete Pixel von davor liegenden Fragmenten überschrieben werden.

Deferred Shading [Har04] [Pol05] umgeht dieses Problem, indem die Beleuchtung verzögert wird, bis die Sichtbarkeit aller Fragmente im Rasterisierungsprozess festgestellt wurde. Auf diese Weise werden teure Beleuchtungsberechnungen nur auf Pixel angewendet, die auch tatsächlich auf dem Bildschirm zu sehen sind.

Hierzu werden beim Rendern beleuchtungsrelevante Geometrie- und Materialinformationen auch als Erscheinungsparameter bezeichnet in verschiedenen Renderzielt Texturen (geometrischer Buffer G-Buffer) abgelegt anstatt die Beleuchtung sofort durchzuführen und Farben auszugeben. Diese Texturen dienen im Anschluss als Eingabe für einen Post-Prozess, welcher die Informationen nutzt, um die Beleuchtung für jedes Licht vollständig in der Bildebene zu berechnen.

2 Deferred Renderer

Das Ziel dieses Projektes war es, einen Deferred Renderer und eine Beispielszene zur Demonstration der Qualität des Renderers zu entwickeln. Außerdem sollen die Vorteile der Technologie gegenüber Forward Rendering aufgezeigt und somit die Relevanz von Deferred Shading für die Zukunft der Spieleentwicklung untermauert werden.

Der Deferred Shading Prozess lässt sich in zwei Kernschritte einteilen – das Füllen des G-Buffers und die anschließende Beleuchtung anhand der im G-Buffer enthaltenen Informationen. Die sorgfältige Implementierung beider Schritte entscheidet über die Qualität des Deferred Renderers.

2.1 G-Buffer

Im G-Buffer werden beim Rendern der Szene alle Informationen abgelegt, die benötigt werden, um die Beleuchtung nach der Rasterisierung zu berechnen. Der G-Buffer besteht aus mehreren Texturen in Bildschirmauflösung. Auf diese Weise können für jeden Pixel Erscheinungsparameter gespeichert werden. Die im G-Buffer enthaltenen Informationen hängen von den Anforderungen der Anwendung ab. Mindestens enthalten sein müssen Farb- und Positionsdaten sowie Normalen, um die Beleuchtung für ein Licht hinzufügen zu können. Da das Füllen der Texturen teuer ist, ist es notwendig, das Format der Texturen sorgfältig auszuwählen und die erforderlichen Erscheinungsparameter in so wenigen Texturen wie möglich zu platzieren.

Die aktuelle Implementierung des G-Buffers verwendet drei Renderziele, welche Farb-, Normalen- und Tiefeninformationen sowie einen Reflektionsfaktor für jeden Pixel enthalten. Anhand der Tiefe kann die Position eines Pixels in der Welt rekonstruiert und somit die Relation zu einem Licht hergestellt werden.

2.2 Beleuchtung

Nachdem der G-Buffer gefüllt wurde, werden die darin enthaltenen Informationen als Eingabe für den Beleuchtungsprozess herangezogen. Um einen Zugriff auf die Informationen im Pixel-Shader zu ermöglichen, kann für jedes Licht ein Rechteck in Vollbildgröße gerendert werden. Dies ist allerdings ineffektiv, da unabhängig von der Charakteristik des Lichts grundsätzlich jeder Pixel berücksichtigt wird. Um eine hohe Performance zu erzielen, ist es notwendig, möglichst viele Pixel auszusortieren, die nicht von einem Licht betroffen sind. Dieser Culling-Prozess in der Bildebene ist der essentiellste Bestandteil des Deferred Shading.

Fünf verschiedene Culling-Verfahren wurden entwickelt. Je nach Charakteristik der in der Szene enthaltenen Lichter bestehen Vor- oder Nachteile für bestimmte Verfahren.

Das sogenannte Mixed Culling wählt für jedes Licht individuell eine der vier übrigen Culling-Techniken aus. Dies geschieht auf Basis der Eigenschaften des Lichts. Mixed Culling ist in den meisten Fällen die beste Wahl, da grundsätzlich eine hohe Performance erzielt wird.

Z-Culling macht sich die Eigenschaften des Z-Buffers zu Nutze, um verdeckte Bestandteile eines Lichtes auszusortieren. Zu diesem Zweck wird ein Lichtvolumen gerendert, welches den Wirkungskreis eines Lichtes repräsentiert (z.B. Kugel für ein Punktlicht, Zylinder für Scheinwerfer). Verdeckte Bereiche dieses Volumens fallen durch Z-Tests weg. Zusätzlich werden alle Pixel außerhalb des Volumens nicht berücksichtigt.

Stencil-Culling macht sich die Funktionen des Stencil-Buffers zu Nutze, um Überschneidungen von Lichtvolumen mit der Geometrie zu markieren und alle Pixel außerhalb abzuweisen. Diese Methode ist am präzisesten, aber mit erhöhtem Aufwand bei der Durchführung verbunden, weswegen Z-Culling in vielen Fällen bessere Resultate verspricht. Aus diesem Grund wurde eine weitere Variante des Stencil-Cullings entwickelt, in der alle Lichter zusammengefasst werden. Wiederholtes Füllen und Leeren des Stencil Buffers sowie Renderzustandsänderungen für jedes Licht fallen somit weg. Das sogenannte Kombinierte Stencil-Culling erreicht oftmals die höchste Leistung aller Culling-Techniken.

Scissor-Culling macht sich die Scissor-Funktionalität der Hardware zu Nutze, um alle Pixel außerhalb eines Rechtecks in der Ebene des Scissors zu entfernen. Für jedes Licht wird das kleinstmögliche Rechteck, welches das Licht voll beinhaltet, berechnet und als Scissor Rectangle verwendet. Scissor-Culling erreicht nicht die Performance der anderen Verfahren, eignet sich aber zur Optimierung der Erzeugung von Schatten mit der Schattenvolumentechnik [Nea05].

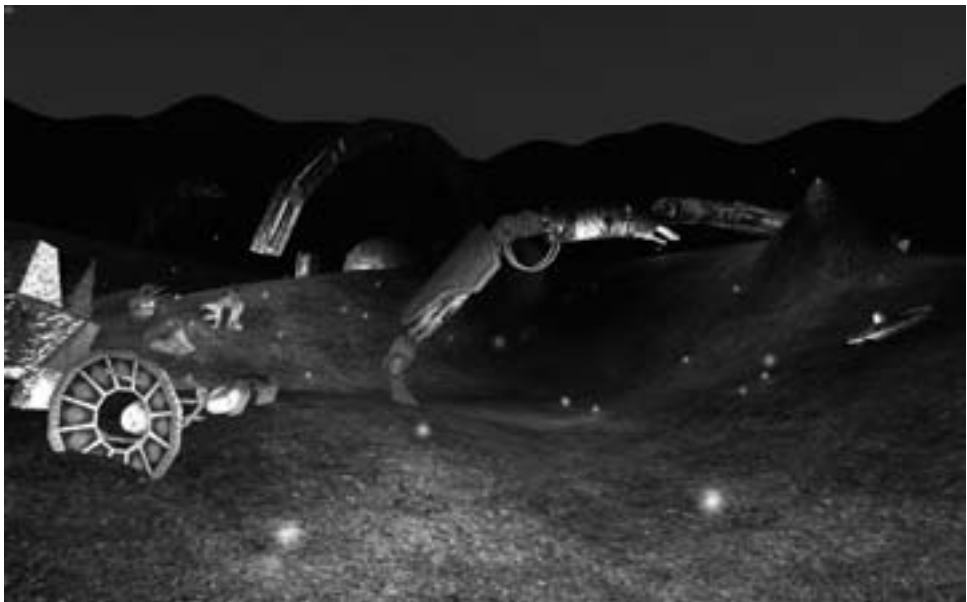


Abbildung 1: Beispielszene mit 50 großen, beweglichen Lichtern

3 Resultate

Zum Vergleich der Performance wurde ein zuvor entwickelter und erfolgreich innerhalb eines Studienprojektes eingesetzter Forward Renderer herangezogen. Der Wechsel zwischen den beiden Renderern ist zur Laufzeit jeder Zeit möglich. Mehrere Beispielszenen wurden zusammengestellt, um den Deferred Renderer in verschiedenen Szenarien zu untersuchen.

Der Deferred Renderer ist dazu in der Lage, mühelos eine große Anzahl an Lichtern innerhalb einer Szene und gleichzeitig auf dem Bildschirm darzustellen. Abgesehen von Situationen mit sehr wenigen großen Lichtern ist die Technologie dem Forward Renderer überlegen. Abhängig vom Aufbau der Szene (Anzahl und Art der Lichter, Anzahl und Komplexität der Geometrie, Anteil schattenwerfender Lichter und Geometrie, Anteil tatsächlich beleuchteter Geometrie) können substantielle Performanceunterschiede verzeichnet werden. Szenen, welche den Forward Renderer nahezu zum Einfrieren bringen (< 5 FPS), werden vom Deferred Renderer problemlos berechnet (> 45 FPS). Auf der Testmaschine (CPU: Intel Core 2 Duo 2.4Ghz, GPU: NVIDIA 8800M GTX) konnte eine Szene mittlerer Geometriekomplexität mit mehr als 1000 beweglichen Lichtern in hoher Bildschirmauflösung (1920x1200) und akzeptabler Framerate (35 oder höher) dargestellt werden.

Die Ergebnisse untermauern die Relevanz, die Deferred Shading und die darin enthaltenen Konzepte für die weitere Entwicklung der Computergrafik und Spieleentwicklung haben. Mittlerweile existieren hoch-qualitative Beispiele aus der Spieleindustrie, welche sich die Technologie zu Nutze machen (z.B. Cryteks CryEngine 2 [Wen06]).

Literaturverzeichnis

- [Har04] Hargreaves, S, Harris, M.: Deferred Shading. 6800 Leagues Under the Sea, NVIDIA developer conference, 2004

Hargreaves, S.: Deferred Shading. Game Developers Conference März 2004, San Jose, USA
- [Pol05] Policarpo, F., Fonseca, F.: Deferred Shading Tutorial SBGames, November 23-25 2005, Sao Paulo, Brazil
- [Nea05] Nealen, A.: Shadow Mapping and Shadow Volumes Januar 2005, online bei www.devmaster.net
- [Wen06] Wenzel, C.: Real-Time Atmospheric Effects in Games. International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH 2006 Courses Session: Advanced Real-Time Rendering in 3D Graphics and Games 2006, Boston, Massachusetts ISBN: 1-59593-364-6

Ein Rahmenwerk für Genetische Algorithmen zur Lösung erweiterter Vehicle Routing Problems (VRPSPDMUTW+)

Sebastian Magnus

Universität Koblenz-Landau, Koblenz
smagnus@uni-koblenz.de

Abstract: Zur Kostenminimierung und zur Erfüllung der Kundenwünsche müssen zur Bestimmung optimaler Tourenpläne eine Vielzahl an Nebenbedingungen berücksichtigt werden, wofür nur Heuristiken in Frage kommen. Klassische Verfahren berücksichtigen oft nur eine geringe Anzahl an Nebenbedingungen, wichtige Sachverhalte der realen Welt werden im mathematischen Modell nicht betrachtet. Dieses Paper beschreibt ein Rahmenwerk für Genetische Algorithmen zur Lösung von Tourenplanungsproblemen unter zusätzlichen Nebenbedingungen bezüglich Zeit, Fahrzeugen, Ladung und Rampen. Die Algorithmen können aus verschiedenen Modulen kombiniert werden. Ausgewählte Algorithmen werden am Beispiel einer 20-Kunden-Instanz verglichen.

1 Motivation

Tourenplanungsprobleme befassen sich mit der Minimierung der Fahrstrecken einer Fahrzeugflotte und damit der Kosten eines Speditionsunternehmens. Das Grundproblem ist bereits sehr alt, allerdings verschärfen sich in jüngerer Zeit die Nebenbedingungen aufgrund großer, heterogener Fahrzeugflotten, eng planender Supply Chain- und Lagerhaltungsstrategien, rechtlicher und tariflicher Regelungen sowie wachsender Planungsgebiete und Kooperationen. Steigender, globaler Konkurrenzdruck erschwert insbesondere kleinen und mittelständischen Unternehmen die Erhaltung der Marktfähigkeit.

Eine manuelle Tourenplanung führt fast immer zu suboptimalen Lösungen, in denen Kunden nicht bedient, Fahrzeuge angemietet oder ein Vielfaches der benötigten Kilometer gefahren werden.

Exakte Tourenplanungsverfahren sind bereits bei recht kleinen Probleminstanzen nicht in angemessener Zeit lösbar, sodass Heuristiken herangezogen werden. Das Institut für Wirtschaftswissenschaften der TU Clausthal untersucht verschiedene dieser Näherungsansätze, von denen im Folgenden ein Framework für verschiedene Genetische Algorithmen (GA) vorgestellt wird. Der GA wird dabei in mehrere Phasen unterteilt, die modular implementiert und in verschiedenen Kombinationen ausgeführt und verglichen wurden.

2 Problemformulierung

Die vollständige Formulierung des vorliegenden VRPSPDMUTW+ als gemischt-ganzzahliges lineares Optimierungsproblem (MILP) umfasst 39 Formeln, die genauer in [Mag08] erklärt werden. Als Grundlage wurde das MILP von Rieck [RZG07] für den mehrfachen Einsatz desselben Fahrzeuges erweitert.

Das Vehicle Routing Problem (VRP) ist ein zum Traveling Salesman Problem (TSP) sehr ähnliches Problem, bei dem eine Menge von Kunden genau einmal besucht werden muss. Im Gegensatz zum TSP werden diese von mehreren Fahrzeugen mit begrenzter Kapazität bedient, die ihre Touren in einem gemeinsamen Depot starten und beenden müssen. Die Fahrt zwischen den Knoten sowie die Ladevorgänge benötigen Zeit pro Strecken- bzw. Ladungseinheit. Im vorliegenden Problem können Fahrzeit und Kosten, die durch die Zielfunktion zu minimieren sind, voneinander abweichen. Das Problem geht von einer heterogenen Fahrzeugflotte aus, in der zwar alle Fahrzeuge die gleichen Kosten verursachen und mit gleicher Geschwindigkeit fahren, aber über unterschiedliche Kapazitäten verfügen. Eine maximale Tourdauer, die sich in praktischen Szenarien z.B. aus Lenk- und Ruhezeiten ergibt, darf nicht überschritten werden.

Bei der Erweiterung „Simultaneous Pickup and Delivery“ (SPD) können Kunden zugleich über eine Auslieferungsmenge verfügen, die vom Depot zum Kunden gebracht werden muss, sowie eine Einsammelmenge, die vom Kunden zum Depot geliefert wird. Ein- und Auslieferung können getrennt, aber jeweils in einem einzigen Ladevorgang erfolgen.

„Time Windows“ (TW) sind zusätzliche Zeitrestriktionen bezüglich frühester und späterer Zeitpunkte, an denen ein Ladevorgang bei einem Kunden oder im Depot beginnen darf. Hierfür ist der Zeitaufwand für die Fahrt und die Ladevorgänge zu berücksichtigen. Die Ladezeit berechnet sich aus der Anzahl der Ladungseinheiten.

Die Fahrzeuge werden als „Multiple Use“ (MU) mehrfach eingesetzt. Das ursprüngliche MILP von Rieck [RZG07] beschränkt sich auf einen zweifachen Einsatz eines Fahrzeuges, in [Mag08] wird auch der mehrfache Einsatz modelliert.

Das Modell bildet zusätzliche Restriktionen bezüglich der Rampenbelegung ab (+). Das Depot verfügt über eine oder mehrere Rampen, an denen zu einem bestimmten Zeitpunkt höchstens ein Fahrzeug laden darf. Die Ladezeit berechnet sich erneut aus der Anzahl der Ladungseinheiten. Be- und Entladung kann nacheinander an unterschiedlichen Rampen stattfinden. Rampen bei den Kunden werden nicht berücksichtigt, da jeder Vorgang vollständig von einem Fahrzeug bedient wird.

3 Rahmenwerk für Genetische Algorithmen

Genetische Algorithmen lösen Optimierungsprobleme nach einem mehrphasigen Schema, welches hier in die Phasen Elternselektion, Crossover, Mutation und Umweltselektion unterteilt wird. Für die Selektionsphasen ist eine Bewertung anhand einer Fitnessfunktion notwendig, für die abweichend von klassischen Ansätzen eine Minimierungsfunktion verwendet wird. Ein in VS-C++ implementiertes Rahmenwerk ermöglicht die Auswahl unterschiedlicher Algorithmen zu den einzelnen Phasen.

In einer Konfigurationsdatei werden in festgelegter Notation die gewünschten Methoden der einzelnen Phase gewählt. Zusätzliche Parameter lassen sich mit geringem Aufwand in die Parameterdatei integrieren. Startverfahren zur Erzeugung einer Initialpopulation können mit einer zweiten Konfigurationsdatei definiert werden. Zur phasenübergreifenden Datenthaltung verwendet das Framework ein gemeinsames Populations-Objekt.

Insgesamt wurden 6 Fitnessfunktionen, 8 Crossover und 5 Mutationsoperatoren integriert. Eine generalisierte Samplingstrategie kann durch 5 Parameter eine Vielzahl bekannter Strategien wie Elitist, (μ, λ) , $(\mu + \lambda)$ und Truncation nachbilden oder kombinieren, wobei mehrere stochastische Modelle zur Auswahl stehen.

4 Crossover, Mutation und Selektion zur Lösung von VRPSPDMUTW+

Aus früheren Arbeiten von Glasgow, Rieck und Zimmermann [Gla07][RZG07] lagen auf einer TSP-Repräsentation basierende Nachbarschaftsoperatoren für Swap, 2-Opt, DB-4-Opt und Lin-Kernighan vor, die an das Framework angepasst und darin integriert wurden. Die ersten drei Ansätze basieren auf dem Vertauschen von Kanten oder Knoten, der letzte implementiert eine verkürzte Heuristik. Zusätzlich können die Touren in eine Giant-Tour zusammengefügt und anschließend über eine auf einem Bellman-Algorithmus basierende Konstruktionsheuristik neu den Fahrzeugen zugeteilt werden [RZ08].

Als Crossover wurden die Operatoren OX, PPX und PSX [Gla07][RZG07] aus früheren TSP-orientierten Versuchen sowie an das Problem angepasste Operatoren ERX, SBX und RBX [YN97][MF00][Pad03][PB96] integriert. Die TSP-Operatoren führten bereits bei kleinen Instanzen zu unzulässigen, weit vom Optimum entfernten Lösungen. Bei den an das spezielle VRP angepassten Operatoren werden bereits während der Erstellung der Tour nur zulässige Lösungen erstellt. Ist für einen Kunden keine zulässige Einplanung möglich, wird er zunächst auf eine Liste gesetzt und in anschließenden Reparaturschritten der unvollständigen Lösung hinzugefügt.

5 Ergebnisse

Zur Auswertung wurden verschiedene Konfigurationen miteinander verglichen. Tabelle 1 zeigt die Ergebnisse verschiedener Crossover-Operatoren mit Roulette-Wheel-Selection, 30% Wahrscheinlichkeit zu einer Mutation über eine Bellman-Heuristik und einer Umweltselektion, die garantiert mindestens zu 50% Kindindividuen und die 10% besten Individuen zieht. Weitere Ergebnisse können in [Mag08] nachgeschlagen werden.

Die beste Lösung konnte mit einem randomisierten Savings-Algorithmus [CW64][RZG07] bei 570 gefunden werden. Während die TSP-basierten Operatoren OX, PSX und PPX deutlich davon abweichen, fanden ERX, SBX und RBX nahezu gleichwertige Lösungen. Bei ERX stellte sich heraus, dass die Kreuzung von mehr als zwei Elternteilen das Ergebnis verbessern kann.

Für Instanzen mit 100 Kunden konnten nur SBX und RBX zulässige Lösungen finden, die

Tabelle 1: Vergleich der Crossover-Operatoren, beste gefundene Lösung: 570

Crossover	Min	Max	Avg	Zulässig	ERX Eltern
OX	670	689	682	3/3	
PSX	755	774	762	3/3	
PPX	666	671	668	3/3	
SBX	585	592	581	3/3	
RBX	576	582	579	3/3	
ERX	580	591	584	3/3	3
ERX	587	592	590	3/3	2

zudem um 10% bis 20% besser als die unzulässigen Lösungen der anderen Operatoren waren. Allerdings blieben sie deutlich schlechter als Lösungen, die mit einem randomisierten Savings erzeugt wurden.

Literatur

- [CW64] G. Clarke und J.W. Wright. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12(4):568–581, Juli 1964.
- [Gla07] Matthias Glagow. Heuristische Lösungsverfahren für die Tourenplanung mittelständischer Unternehmen, 2007.
- [Mag08] Sebastian Magnus. Heuristiken zur Lösung von Tourenplanungsproblemen, insbesondere Genetische Algorithmen, September 2008.
http://www.uni-koblenz.de/~smagnus/Publications/Diplom_Sebastian_Magnus_GA.pdf.
- [MF00] Zbigniew Michalewicz und David B. Fogel. *How to solve it: Modern Heuristics*. Springer Verlag, 2000.
- [Pad03] Jorge Cantón Padilla. *Integrated support system for planning and scheduling of batch chemical plants*. Universitat Politècnica de Catalunya, April 2003.
- [PB96] J.-Y. Potvin und S. Bengio. The vehicle routing problem with time windows — Part II: Genetic search. *INFORMS Journal on Computing*, 8:165–172, 1996.
- [RZ08] Julia Rieck und Jürgen Zimmermann. Koordinierung der Geschäftstätigkeiten von Expeditionen in Netzwerken von Stückgutkooperationen. *Netzwerklogistik, Begleitband zur 13. Magdeburger Logistik-Tagung*, Seiten 123–135, 2008.
- [RZG07] Julia Rieck, Jürgen Zimmermann und Matthias Glagow. Tourenplanung mittelständischer Speditionsunternehmen in Stückgutkooperationen: Modellierung und heuristische Lösungsverfahren. *Zeitschrift für Planung und Unternehmenssteuerung*, 17(4):365–388, Januar 2007.
- [YN97] Takeshi Yamada und Ryohei Nakano. Genetic Algorithms for Job-Shop Scheduling Problems. *Proceedings of Modern Heuristic for Decision Support, UNICOM*, Seiten 67–81, März 1997.

Towards extending the Tracker Workflow Capabilities of SourceForge Enterprise Edition

Johannes Nicolai

johannes.nicolai@hpi.uni-potsdam.de

Abstract: The collaborative development and application lifecycle management platform CollabNet SourceForge Enterprise Edition features a workflow engine to enforce tracker rules for feature requests, bug reports and other customer defined tracker items. To find out whether the currently implemented tracker workflow capabilities are perceived to be sufficient to support advanced change management processes, a survey was created. This paper explains the steps that lead to the final design of the questionnaire that was put online by CollabNet.

1 Introduction

CollabNet SourceForge® Enterprise Edition (SFEE)¹ is an integrated suite of Web-based development and collaboration tools that synchronizes distributed teams and disparate processes to simplify development [Col08]. The focus of this paper lies on the fully-configurable tracker manager of SFEE. This tracker manager is used to create and manage trackers for feature requests, bug reports, user stories, test cases and any other development related artifact type. Since SFEE claims to be a full-fledged ALM platform, CollabNet tries to close the gap between the current tracker workflow functionality and the functionality desired by its customers and prospects to support advanced change management processes.

Section 2 introduces a questionnaire to gather the most desired still missing tracker workflow capabilities, as well as the assumptions and constraints behind this survey. Section 3 outlines references to related research work. Section 4 summarizes the previous sections and gives an outlook on what future fields can be worked on after having evaluated the results of the tools.

2 Survey-based Requirements Engineering for Missing Tracker Workflow Capabilities

Since CollabNet's customers are distributed all over the world and usually only have several minutes to take a survey, the questionnaire mostly uses closed-ended questions: It

¹Whenever the abbreviation SFEE is used within this paper it refers to CollabNet SourceForge Enterprise Edition 5.0.

Concept	Real-world Example
Entry Action Support	Once a tracker item enters into the “Approved Status”, all its children artifacts are automatically set into the “Approved Status” as well.
Invariant Support	If a user is not part of the “Manager Role”, she cannot change the value of the “Estimated Hours” field as long as the tracker item is still in the “Planning Status”.
Triggered Action Support	If a tracker item is in the “In QA Status” and the tester selects the “Approved without Changes Option”, it automatically gets transitioned into the “Preproduction Status”. If the tester selects the “Approved with Changes Option”, it automatically gets transitioned into the “In Development Status”. In both cases, the associated developer receives an e-mail.
Precondition Support	A tracker item in the “Planning Status” can only be changed into the “Ready for Development Status” if the “Estimated Hours” field value is less than 200. If the “Estimated Hours” field value is higher than 200, the tracker item can only be switched to the “Second Approval Status”
Exit Action Support	Once a tracker item leaves the “Testing Status”, the summaries of all associated test reports (children artifacts) are concatenated and stored into a custom field of the parent.

Table 1: Real-world Examples for the Proposed Tracker Workflow Capabilities

introduces conceptually new tracker workflow capabilities. These capabilities can be rated on a balanced likert scale from 1 point *useless feature* up to 5 points *highly desired feature*.

These new semantic concepts were heavily inspired by already existing requirement documents² describing tracker processes, customers and prospects would like to enact using SFEE’s tracker workflow engine. The missing tracker workflow features in order to support the desired tracker processes in these documents were grouped into five categories (resulting in the five introduced semantic concepts):

1. Support of Status Entry Actions: Everytime, a tracker item enters into a certain status, an entry-action associated to this status can be automatically executed.
2. Support of Status Invariants: Everytime, a user tries to change field values of a tracker item, a condition associated to the tracker item’s current status can be evaluated. If this condition is not fulfilled, the user is not allowed to change the artifact.
3. Support of Value-Change Triggered Actions: Everytime, a user has changed the field values or the status of a tracker item, an action associated to the artifact’s current status can be executed. This action can refer to the current field values and status as well as to the previous field values and status of the tracker item.

²These documents are treated confidentially and cannot be cited here.

4. Support of Transition Preconditions: Everytime, a user tries to change the status of a tracker item, a precondition associated to this status transition can be evaluated. If this precondition fails, the user is not allowed to switch the artifact to the anticipated status.
5. Support of Status Exit Actions: Everytime, a tracker item leaves a certain status (i.e. switches into another status), an exit-action associated to this status can be automatically executed.

Abstracting the requirements contained in a collection of customer documents to a set of new tracker workflow capabilities and letting the survey participants rate them is probably a reasonable idea. However, just naming and explaining the new capabilities usually does not lead to meaningful results. In order to be able to judge whether the new semantic constructs are usable for the own organization, customers have to identify their specific use cases with the categories presented in the survey. Therefore, every semantic construct comes with three real-world examples how this feature might be actually used to support the organizations' change management processes.

Table 1 shows an excerpt of the examples used in the survey to illustrate the use of the newly proposed workflow capabilities.

To promote the survey, a blog post on open.collab.net [Nic08a], a newsletter entry to all subscribers interested in CollabNet related news and a landing page have been prepared [Nic08b].

The survey can also be taken online and due to the nature of the questions should be completed in less than quarter an hour³. The online survey was realized with the software *surveymonkey* [Sur08] which automatically ensures the accessibility requirements according to section 508 federal law [IAT08] and already provides a survey results analysis backend.

3 Related Work

CollabNet SourceForge Enterprise Edition (SFEE) is not the only collaborative development and application lifecycle management platform available. Seven different Collaboration platforms are compared within [MV07]. When comparing the tracker workflow capabilities of these products, their workflow engines are either based on an event driven finite state machine as well or do not exist at all. Therefore, SFEE's tracker workflow engine functionality can be seen as an representative example of common tracker workflow functionality available within industry leading collaborative development platforms.

The naming and semantics of the proposed tracker workflow extensions presented in section 2 loosely relate to a small subset of the UML state machine specification [Obj07]. When it comes to modelling workflows with UML, activity diagrams are often preferred because of their higher expressiveness [Esh02]. However, SFEE's workflow engine is optimized to support highly standardized, straight-through, data-centric tracker workflows

³This was the result of a pretest with 10 participants.

and the proposed features should be realizable by just extending the finite state machine processing model. If it turns out that the new workflow capabilities will be actually used by the customers and further functionality is requested, a new survey might concentrate on more sophisticated concepts of business process management as described in [vdAHW03].

4 Summary and Outlook

CollabNet SourceForge Enterprise Edition features a workflow engine to enforce tracker rules for feature requests, bug reports and other customer defined tracker items. To find out whether the currently implemented tracker workflow capabilities are perceived to be sufficient to support advanced change management processes, a survey was created.

The survey mainly focuses on rating five new proposed semantic constructs (status entry actions, status exit actions, status invariants, value-change-triggered actions and transition preconditions) that were heavily inspired by existing customer requirement documents.

CollabNet has decided to put the survey online and start the promotion for it in late summer of 2008. The survey results and a prototype that extends SFEE to support the most desired features is presented in the author's Master Thesis *Suitability of Enterprise Collaboration Software for Advanced Change Management Processes* that will be published in the first quarter of 2009.

References

- [Col08] CollabNet. SourceForge Enterprise Edition. <http://www.collab.net/products/sfee/>, 2008.
- [Esh02] H. Eshuis. Semantics and Verification of UML Activity Diagrams for Workflow Modelling, 2002.
- [IAT08] IATW. Section 508 Federal Law. <http://www.section508.gov/index.cfm?FuseAction=Content&ID=3>, 2008.
- [MV07] Susanne Muehlbauer and Gerhard Versteegen. *Studie Collaboration Software - Zusammenarbeit im Team*. Heise Zeitschriften Verlag, 2007.
- [Nic08a] Johannes Nicolai. Blog post promoting the survey and the usage mining tool. <http://blogs.open.collab.net/oncollabnet/2008/07/help-to-shape-t.html>, 2008.
- [Nic08b] Johannes Nicolai. Landing page for the survey and the usage mining tool. <http://workflow-extractor.open.collab.net>, 2008.
- [Obj07] ObjectModellingGroup. OMG Unified Modeling Language (OMG UML), Superstructure, V2.1.2. <http://www.omg.org/docs/formal/07-11-02.pdf>, 2007.
- [Sur08] Inc. SurveyMonkey. SurveyMonkey. <http://www.surveymonkey.com>, 2008.
- [vdAHW03] W. van der Aalst, A. Hofstede, and M. Weske. Business process management: A survey, 2003.

Performanzanalyse von Web Services für das Netzwerkmanagement mit SNMP

Stefan Solich

Hochschule Furtwangen
Fakultät Informatik
st@solich.de

Art der Arbeit: Bachelor Thesis

Betreuer der Arbeit: Prof. Dr. Friedbert Kaspar

Abstract: Ziel der Arbeit ist es Latenzen, d.h. im Wesentlichen Verarbeitungszeiten, die durch die WebServices-Schnittstelle zur Integration von SNMP Managementdaten entstehen, zu untersuchen. Dazu wurde die Apache Muse Implementierung von Management Using Web Services (MUWS) als Basis für einen Prototypen verwendet. Dieser Prototyp diente zur Durchführung der vorgestellten Messungen. Apache Muse ist ein Werkzeug, das für die Integration heterogener Managementansätze entwickelt wurde. Diese Arbeit untersucht die damit evtl. verbundenen Performancenachteile. Mit drei Versuchsaufbauten wurden praxisrelevante Fälle betrachtet. Die Ergebnisse geben Hinweise für Einsatzbereiche und für den Entwurf von WebServices-Schnittstellen zur Integration von SNMP Managementdaten.

1 Einführung

Heutzutage sind nahezu alle IT-Landschaften heterogen. Daher besteht ebenso Bedarf an heterogenem Netzwerkmanagement und SNMP ist der de facto Standard für Netzwerkmanagement. Im Rahmen dieser Arbeit wird untersucht, in wie weit es performant ist, Web Services als Integrationstechnik, für Systemmanagement zu verwenden. Es wurde ein Prototyp entwickelt, der die Integration am Beispiel von SNMP untersucht. Dafür wurde auf der WebService-Seite die Apache Muse Implementierung von Webservice Distributed Management (WSDM) MUWS verwendet. Apache Muse implementiert auch die dafür benötigten, zugrundeliegenden Basisstandards WSRF und WSN [Ap07]. Die Verwendung von Apache Muse unterscheidet diese Untersuchung von früheren Arbeiten [ADMQ04], [MSSG07]

Der Versuch soll zeigen, ob dies eine relevante Alternative sein kann oder ob ein zu erwartender Performanceverlust so groß ist, dass es sich nicht lohnt, solch eine Variante in einem real eingesetzten IT-System zu verwenden.

2 Versuchsaufbau

Alle benötigten Hosts sind virtuelle Maschinen mit jeweils einem Prozessor und 1536 MB RAM. Das Betriebssystem der virtuellen Maschinen ist Debian 4.0 mit Java 5 EE. Auf allen Maschinen, wo ein *Application Server* notwendig ist, wird Apache Tomcat 5.5.27 verwendet. Es wird Net-SNMP 5.4.2.1 verwendet und die Version 1 des SNMP Protokolls. Im Folgenden werden die einzelnen Versuchsaufbauten beschrieben.

Im **ersten Versuchsaufbau** wird der *SNMP Agent* für jede Operation ein *SNMP Request* erhalten und darauf hin einen *SNMP Response* an den *WS based SNMP Consumer* zurücksenden. Die SNMP Daten werden mittels „Java SNMP Package“ [Se05] verarbeitet. Da es in diesem Aufbau lediglich darum geht, Vergleichswerte zu erhalten, werden keine weiteren Operationen ausgeführt. Der *WS based SNMP Consumer* ist ein Web Service, der ein Servlet enthält, dass bei Aufruf die SNMP Abfragen ausführt und die Ergebnisse in eine Datei speichert.

Im **zweiten Versuchsaufbau** werden für die Messungen, im Gegensatz zu Versuchsaufbau 1, drei Hosts benötigt. Der SNMP Agent soll SNMP-Daten mittels SNMP Response an den *SNMP2XML Translator* senden. Diese können sowohl auf zwei verschiedenen virtuellen Hosts laufen sowie auf demselben. Letzteres wird in Versuch 3 behandelt. Der *SNMP2XML Translator* Webservice läuft in einem Webservice Container. Der *SNMP2XML Translator* überführt die empfangenen Daten in eine SOAP Nachricht und überträgt diese an den WS-Based Consumer. Der Translator verarbeitet die SNMP-Daten mit Hilfe des oben genannten Java SNMP Package und wandelt diese dann mittels JAXB in XML um. Der WS-Based Consumer empfängt lediglich die SOAP Nachrichten und misst die Zeiten. Da nur der Zeitverlust am SNMP2XML Translator gemessen werden soll, ist es nicht nötig, weitere Verarbeitungsmechanismen zu implementieren.

Der **dritte Versuch** läuft analog zu Versuchsaufbau 2 ab, jedoch sind hier der *SNMP Agent* und der *SNMP2XML Translator* auf demselben virtuellen Host beheimatet. Die Hypothese ist, dass dies zu kürzeren Laufzeiten führt als in Versuchsaufbau 2. Außerdem zeigt dieser Versuchsaufbau in Relation zu den ersten beiden Versuchen, wie viel Performanceverlust bei der Übertragung und wie viel durch die eigentlichen Operationen verloren geht.

3. Vorstellung der Messergebnisse

Im Folgenden werden die erhaltenen Messergebnisse vorgestellt und ausgewertet. In Tabelle 1 sind die drei durchgeführten Messreihen zu sehen. Daraufhin folgt eine grafische Darstellung (Abb. 1) der Messergebnisse.

	Messung 1	Messung 2	Messung 3
TCPActiveOpens	13,6	116,2	95,3
TCPAttempFails	16,9	117,8	97,1
TCPCurrEstab	9,4	107,1	100,1
TCPEstabResets	9,6	111,4	96,8
TCPInErr	9,8	114,1	99,7
TCPInSegs	8,9	101,2	99,3
TCPMaxConn	8,9	117,9	95,6
TCPOutRsts	9,7	109,7	101,2
TCPOutSegs	8,4	110,6	94,5
TCPPassiveOpens	11,2	107,2	98,1
TCPRetransSegs	10,1	106,8	93,9
TCPRtoAlgoithm	6,3	101,2	100,2
TCPRtoMax	10,7	103,4	92,7
TCPRtoMin	8	105,2	96,1
Gesamt:	141,5	1529,8	1360,6
TCPConnTable	437,3	1716,3	1549,7

Tabelle 1: Messergebnisse

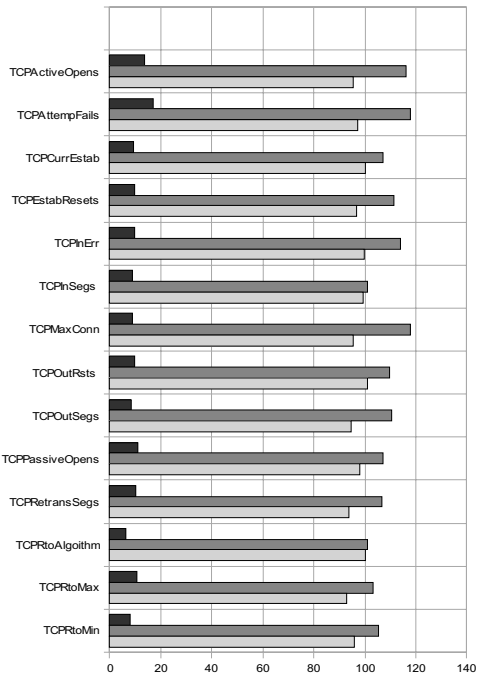


Abbildung 1: Grafische Darstellung der Messergebnisse

In der Grafik (Abb. 1) wird deutlich, dass der Faktor des Performanceverlustes bei den Einzeloperationen (alle außer TCPConnTable) ca. 8 bis 10 betragt. In der Gesamtsumme betragt der Faktor 10,81 fur Messung 2 und 9,62 fur Messung 3. Auerdem betragt die Performanceverbesserung bei Messung 3 im Vergleich zu Messung 2 lediglich 169,2 ms absolut oder 11%. Wenn man nun bedenkt, dass fur die Performanceverbesserung zusatzlich auf der *Managable Resource*¹ immer ein Applikationsserver gestartet sein muss, welcher zusatzlich Performance verbraucht, ist es in vielen Fallen sinnvoll, eher die Variante aus Messung 2 zu bevorzugen. Hinzu kommt noch, dass es in vielen Fallen gar nicht moglich sein wird auf einer Ressource noch zusatzlich einen Applikationsserver zu installieren, da es sich z.B. um ein Gerat handelt, das nicht die erforderliche Leistung bringt, um zusatzlich zu seinen Funktionen noch einen Applikationsserver bereitzustellen.

Allgemein ist zu uberlegen, ob es sich uberhaupt lohnt in einem praktischen Umfeld eine solche Variante zu benutzen, da der Performanceverlust mit dem Faktor 10 in groem Rahmen durchaus ein Kostenfaktor werden kann.

Im Gegensatz zu den Einzeloperationen ist der Performanceverlust bei der Operation TCPConnTable lediglich Faktor 3,9 fur Messung 2 und 3,5 fur Messung 3. Sollten Operationen dieser Art oft angewendet werden, kann es sich evtl. lohnen eine Variante wie in Messung 2 und 3 auch praktisch zu verwenden.

¹ In diesem Fall der SNMP Agent

Durch den Vergleich von Messung 2 und 3 ist deutlich zu erkennen, dass der größte Performanceverlust bei der Umwandlung selbst entsteht und nicht durch die Übertragung.

4. Fazit

In den meisten produktiv eingesetzten Netzwerken wird eine solche Umwandlung kaum eine Alternative sein, da der Performanceverlust zu hoch ist. Sie könnte jedoch dann interessant werden, wenn bestimmte SNMP Agenten auf SNMP festgelegt sind, allerdings alle anderen Verwaltungsmechanismen auf Web Services basieren. Sollte, die in dieser Arbeit untersuchte Form der Integration notwendig sein, so sollte bei der Umsetzung der Umwandlung Wert auf eine sehr performante Implementierung gelegt werden.

Literaturverzeichnis

- [ADMQ04] Pras, A.; Drevers, T; Van de Meent, R.; Quartel, D.: Comparing the Performance of SNMP and Web Services-Based Management. In IEEE electronic Transactions on Network and Service Management, 2004
- [Ap07] The Apache Software Foundation: Apache Muse <http://ws.apache.org/muse/> (Abruf: 13.1. 2009)
- [MSSG07] Moura, G. C. M.; Silvestrin, G., Sanchez, R. N.; Gaspary, L. P.; Granville, L. Z.: On the Performance of Web Services Management. In Proc. Of the 10th IFIP/IEEE International Symposium on Integrated Network Management, 2007
- [Se05] Jonathan Sevy: Java SNMP Package, Drexel University, 2005 <http://gic1.cs.drexel.edu/people/sevy/snmp/snmp.html> (Abruf: 13.1. 2009)

Arbeitsumgebungen in der Softwareentwicklung im Team

Themen- und Expertennetzwerke und deren zeitabhängige Veränderungen

Steffi Döbbelin

Universität Paderborn
Fakultät Informatik
steffid@mail.uni-paderborn.de

Art der Arbeit: Diplomarbeit

Betreuer der Arbeit: Wolfgang Reinhardt, Didaktik der Informatik (DDI)

Abstract: Um "Social Software Engineering" technisch zu begleiten haben sich viele Methoden und Arbeitsumgebungen etabliert. Das Besondere an diesem Prozess ist, dass Netzwerke aufgrund des gemeinsamen Bearbeitens gleicher Komponenten des Projektes automatisch entstehen. Visualisiert man diese Themennetzwerke erhält man eine Liste von Projektmitgliedern, die Experten für die entsprechenden Komponenten sind. In der Diplomarbeit wurden einerseits die Aktivitäten des Softwareentwicklungsprozesses visualisiert, als auch die Liste möglicher Ansprechpartner für Teile des Projektes extrahiert.

1 Kooperative Softwareentwicklung

Unter dem Schlagwort "Social Software Engineering" verbirgt sich heute einerseits die Entwicklung von "Social Software" als auch den Entwicklungsprozess, der gemeinschaftlich im Team begangen wird (s. [AJLN08]). Das Verständnis von Softwareentwicklung als „soziale Aktivität“ geht zurück auf Gerald Weinberg. Er beschreibt dies in seinem Buch „The Psychology of Computer Programming“ von 1971 ([W71]).

Da die Mitglieder eines Teams oftmals nicht zur selben Zeit und am selben Ort am Projekt arbeiten, benötigt man für die Zusammenarbeit im Hinblick auf Kommunikation, Koordination und Kooperation technische Unterstützung.

Es gibt viele Arbeitsumgebungen, die Teile des Softwareentwicklungsprozesses unterstützen und begleiten. Ein Vertreter davon ist Trac. Ursprünglich hervorgegangen aus einem Ticketsystem hat es aktuell auch eine SVN-Anbindung und ein Wiki-Modul.

Aufgabe der Diplomarbeit ist es neue verbesserte Ansichten für die Trac „Timeline“ und Trac „Browse Source“, sowie eine neue Benutzer-zentrierte Darstellung der Daten eines Trac-Systems zu erstellen. So soll die Listendarstellung der Timeline durch Diagramme ersetzt werden. Und die „Browse Source“-Ansicht wird für jeden SVN-Pfad um eine Liste von Ansprechpartnern, den sog. „DomainExperts“, erweitert.

2 Beschreibung der Ergebnisse der Diplomarbeit

Die entstandene Oberfläche ist kein Prognose-Werkzeug auf Basis von Trac und den Daten aus einem Trac-System. Das bedeutet, dass keine Handlungsanweisungen abgeleitet werden. Dennoch sollen die Daten grafisch so aufbereitet werden, dass sie Aussagekraft besitzen. Beispielsweise sollen im Rückblick auf ein Projekt Probleme oder Missstände aufgedeckt werden.

Mit der Diplomarbeit soll aufgezeigt werden, wie im Projekt gearbeitet wird. Wichtige Fragen wie etwa „Wer arbeitet an welcher Komponente mit wem zusammen?“ und „Was passiert gerade im Projekt?“ sollen schnell beantwortet werden können.

Die Implementierung der Diplomarbeit besteht aus zwei Teilen: Der neuen Oberfläche „InteractiveTrac“, die mit Adobe Flex 3 entwickelt wurde und einem Trac-RPC-Plugin „TracRPC“, mit dessen Hilfe die Oberfläche an die benötigten Daten aus dem Trac-System kommt.

Im Folgenden sollen die drei neuen Ansichten der Oberfläche kurz beschrieben werden.

3 Vorstellung der neuen Visualisierungen

Wie auch in der ursprünglichen Trac „Browse Source“-Ansicht kann man durch das gesamte SVN-Repository des Trac-Projekts navigieren. Dargestellt wird immer die aktuellste Revision eines SVN-Pfades mit zusätzlichen Informationen zu dieser Revision (ID, Autor, Message).

Die in der Diplomarbeit entwickelte Erweiterung dieser Ansicht – „Trac Domain Experts“ genannt - fügt jedem SVN-Pfad eine Liste von Ansprechpartnern für diese Komponente hinzu. Zusätzlich kann zu jedem Pfad die Historie (also die Liste von Changesets, in der der Pfad bearbeitet wurde) eingesehen werden.

Diese Liste von Experten wird automatisiert generiert. Die Bestimmung, wer ein Experte ist und wer nicht, erfolgt dynamisch: Bei dem Hauptmodul „GenerateExperts“ können sich beliebig viele Teilmodule registrieren, die etwas zu der Berechnung beitragen wollen. Beispielsweise gibt es ein Modul, das die Anzahl Changesets pro Autor für eine Datei zählt. Um damit eine Liste von Experten für einen Ordner zu berechnen, wird ein Durchschnitt über alle Elemente des Ordners und gegebenenfalls Unterordner erstellt.

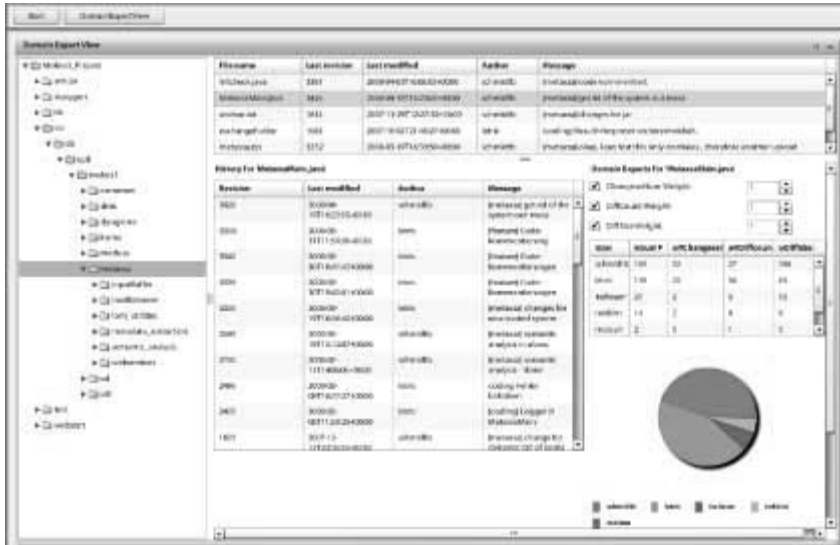


Abbildung 1 - Screenshot der „Trac Domain Experts“-Ansicht

Die alte Timeline von Trac wird vollständig durch die „Trac Event Charts“ (Abb. 2) ersetzt. Alle Timeline-Events für einen Zeitraum werden in einem Zeitstrahldiagramm dargestellt. Dabei ist der betrachtete Zeitraum frei wählbar: über Datumsfelder oder über ein Übersichtsdiagramm, das ein intuitives Zooming und Scrolling erlaubt. Für ein Event erscheinen bei Auswahl Detailinformationen. Zusätzlich zum Hauptdiagramm links gibt es ein Balkendiagramm rechts, das pro Tag die Events pro User aufsummiert, um so die aktiven Teammitglieder schnell zu identifizieren.

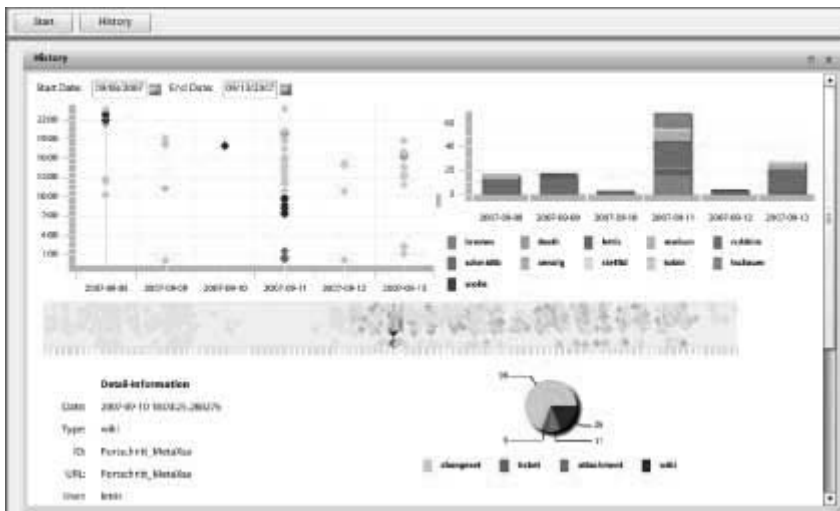




Abbildung 3 - Screenshot der “Trac User Panel”-Ansicht

Die dritte Ansicht ist das „Trac User Panel“ (Abb. 3), die im Rahmen der Diplomarbeit entwickelt wurde. Ziel war es die Beziehungen der Mitglieder eines Teams zu visualisieren, um damit aufzuzeigen wie die Teammitglieder untereinander kooperieren und welche Teammitglieder dabei besonders intensiv zusammenarbeiten. Dazu wurden alle drei Ressourcen eines Trac-Systems herangezogen: Tickets, Wikiseiten und Changesets.

Für Tickets eines Trac-Systems gibt es 4 Rollen, durch die ein Benutzer an einem Ticket teilhaben kann: Reporter, Owner, Author und CC. Im Trac User Panel werden alle Tickets gelistet, in der der Benutzer in mindestens einer Rolle auftaucht. Neben der Tabelle befindet sich eine grafische Auswertung über die Ticketbearbeitung innerhalb des Teams. Damit werden beispielsweise Fragen beantwortet wie: Von wem hat ein Benutzer besonders viele Tickets zugewiesen bekommen? Wem haben die Tickets gehört, an denen ein Benutzer mitgearbeitet hat?

Die zweite Benutzer-zentrierte Auswertung betrifft die Wikiseiten: Es werden zu einem Benutzer alle Wikiseiten gelistet, an denen er mindestens einmal mitgeschrieben hat. Daneben wird grafisch pro Wikiseite ausgewertet, welche Benutzer an diesen Seiten wie oft Editierungen vorgenommen haben.

Um nicht die Auswertungen der Trac Domain Experts zu wiederholen, werden für die Benutzer-zentrierte Auswertung der Changesets die Anzahl Changesets eines Benutzers sowie die Anzahl der dabei geänderten Pfade angezeigt. Zusätzlich wird die durchschnittliche Zeit zwischen zwei Changesets eines Benutzers berechnet. In dem Diagramm daneben wird ausgewertet, wie oft ein Benutzer mit anderen Benutzern an denselben Pfaden gearbeitet hat.

3 Ausblick

Die Trac Event Charts visualisieren die Timeline Events eines Trac-Systems und ermöglichen so einen zusammengefassten Blick auf ein Softwareentwicklungsprojekt. Die Events werden jedoch inhaltlich nicht ausgewertet, daher wird in der Diplomarbeit auch der Ansatz und eine mögliche Implementierung der "Process Behavior Charts" von Florac und Carleton [FC99] diskutiert. Die "Process Behavior Charts" werten die Daten aus der Vergangenheit eines Softwareentwicklungsprozesses aus und erstellen damit Prognosen für die Zukunft des Prozesses.

Desweiteren wird in der Diplomarbeit das Klassifikationssystem von Anita Sarma (vgl [S05]) herangezogen und eine Einordnung sowohl des ursprünglichen Trac-Systems als auch der neuen Oberfläche vorgenommen. Dadurch sollen noch fehlende Funktionen und Charakteristika herausgearbeitet werden, die auf dem Weg zu einer vollkommenen Arbeitsumgebung für kooperative Softwareentwicklung noch fehlen.

Literaturverzeichnis

- [FC99] Florac, W.; Carleton, A.: Measuring the Software Process. (The SEI series in software engineering) Addison-Wesley, Reading, Massachusetts, 1999.
- [AJLN08] Ahmadi, N.; Jazayeri, M.; Lelli, F.; Nescic, S.: A Survey of Social Software Engineering, Appears in: Automated Software Engineering - Workshops, 2008. ASE Workshops 2008. 23rd IEEE/ACM International Conference on
- [W71] Weinberg, G.: The Psychology of Computer Programming: Silver Anniversary Edition Dorset House Publishing Company, Incorporated; Anl Sub edition (September 1998)
- [S05] Sarma, A.: A Survey of Collaborative Tools in Software Development; Institute for Software Research, Donald Bren School of Information and Computer Sciences, University of California, Irvine; ISR Technical Report # UCI-ISR-05-3, March 22, 2005

PocketLogo – Ein Logo-Interpreter für Handhelds

Markus Hunke

Prof. Dr. Marco Thomas

Am Brunnenbach 53
48727 Billerbeck
Markus.Hunke@gmx.de

Didaktik der Informatik
Universität Münster
ddi@uni-muenster.de

1. Was ist PocketLogo?

PocketLogo ist eine Programmierumgebung, die unter dem Motto "A turtle in your pocket" die Programmierung in einem Dialekt der Sprache Logo auf mobilen Geräten (Windows Mobile Pocket PCs bzw. Smartphones) erlaubt. PocketLogo unterstützt alle üblichen Logo-Funktionen und enthält zusätzlich einige Erweiterungen, z.B. können Wände definiert werden und die Turtle kann Objekte aufnehmen bzw. ablegen. PocketLogo übersetzt außerdem Lindenmayer-Systeme in Logoprogramme und der aus einen Logoprogramm erzeugte Syntaxbaum kann angezeigt werden. Nähere Informationen zu diesen Erweiterungen finden sich in Abschnitt 4 und auf u.a. Webseite.

Es existiert außerdem eine Portierung von PocketLogo für PCs (Windows XP/Vista) mit größtenteils identischem Funktionsumfang. PocketLogo wurde für den Einsatz im Informatikunterricht der Mittel- und Oberstufe konzipiert. Ich habe PocketLogo im Rahmen meiner Examensarbeit [Hu08] am Fachbereich Didaktik der Informatik der Westfälischen Wilhelms-Universität Münster im Sommersemester 2008 entwickelt und am Gymnasium Nepomucenum in Coesfeld in einem Informatikkurs der Klasse 10 erprobt.

Viele Bildschirmfotos und -Videos, Programmbeispiele und natürlich Software und Handbuch selbst können bei <http://pocketlogo.sourceforge.net/> heruntergeladen werden.

2. Warum Logo?

Die Sprache Logo ist aus verschiedenen Gründen für einen Einstieg in die (textuelle) Programmierung recht gut geeignet, von denen hier nur einige genannt seien:

- Logo ist eine interpretierte Sprache: Befehle können nacheinander und einzeln eingegeben und direkt ausgeführt werden, anstatt sofort ein vollständiges Programm zu entwickeln.
- Eine wichtige Komponente von Logo ist die sog. Turtlegrafik: eine virtuelle Schildkröte (engl.: turtle) wird durch Kommandos wie "vor", "zurück", "links" und "rechts" bewegt; sie zieht eine Linie hinter sich her und erzeugt so eine Zeichnung.
- Die Verbindung dieser Punkte macht Logo zu einer interaktiven Sprache: Der Benutzer erhält ein sofortiges, motivierendes Feedback bereits nach der Ausführung des ersten (Turtle-)Befehls – nämlich die Bewegung der Turtle.

- Es gibt nur drei Datentypen (Wort, Liste und Array). Zahlen werden wie Worte behandelt; die Unterscheidung geschieht für den Benutzer transparent.
- Variablen müssen nicht deklariert werden. Name und Wert von Variablen werden syntaktisch unterschieden: Es gibt keine Anweisungen wie $x = x + 1$ (z.B. in Java), die mathematisch keinen Sinn ergeben. Logo zwingt in der Anweisung `setze "x :x + 1` zur Unterscheidung zwischen dem Namen und dem Wert der Variablen x .
- PocketLogo erlaubt es dem Benutzer, alle Befehle und Meldungen ins Deutsche (oder andere Sprachen) zu übersetzen bzw. eine fertige Übersetzung zu laden.

3. Logo für Handhelds?

Die Idee zur Entwicklung eines Logo-Interpreters für Handhelds entstand aus dem Projekt "Informatik goes mobile" des Arbeitsbereichs Didaktik der Informatik in Münster. Im Rahmen dieses Projekts wird untersucht und erprobt, wie sich PDAs sinnvoll in den schulischen Informatikunterricht integrieren lassen [Bü07]. In den vergangenen Jahren wurden bereits die visuellen Programmiersprachen PocketCoder [Po06] und Feinstein [Ho07] entwickelt. Für den Einstieg in die textbasierte Programmierung kam PythonCE zum Einsatz. Aus den o.g. Gründen war Logo ein weiterer Kandidat für eine einsteigergerechte, textbasierte Programmiersprache.

Logo-Interpreter (auch freie) für den PC gibt es bereits etliche. Dagegen gibt es für Handhelds bislang offenbar äußerst wenige Logo-Interpreter. Das "Logo Tree Project" [Bo08], das die Zielsetzung hat, alle bekannten Logo-Dialekte aufzuführen, nannte vor der Entwicklung von PocketLogo als einzigen Handheld-Interpreter TinyLogo für Palm OS [Bo08, Nr. 170]. Interpreter für Windows Mobile sind mir nicht bekannt, so dass PocketLogo hier möglicherweise die erste Implementierung ist.

Auch unabhängig vom Einsatz in der Schule sind PDAs und insbes. Smartphones (Mobiltelefone mit PDA-Funktionen) inzwischen weit verbreitete mobile Medien, mit denen man (entsprechende Software vorausgesetzt) lernen, Lösungen entwickeln und ausprobieren kann, egal wo man gerade ist und unabhängig vom stationären Computer(raum). Aber auch in der Schule haben Handhelds eine Daseinsberechtigung, da sie in Verbindung mit geeigneter Software z.B. das Erlernen einer Programmiersprache auch außerhalb des Computerraums ermöglichen bzw. dessen Kapazität erweitern können. Nicht unterschätzt werden sollte auch, dass der Einsatz von Handhelds im Unterricht zur Steigerung der Motivation der Schülerinnen und Schüler beitragen kann.

Die Sprache Logo ist für den Einsatz auf Handhelds vergleichsweise gut geeignet, da die Programmierung nicht sehr "textlastig" ist – man kann schon mit kurzen Logoprogrammen z.B. eindrucksvolle Grafiken erzeugen. Zudem gibt es für die meisten Befehle alternative Kurzformen aus je nur zwei Zeichen, so dass sich der Tippaufwand in Grenzen hält (z.B. `fd/vw` für `forward/vorwärts`). Ich habe bei der Entwicklung von PocketLogo versucht, möglichst viele Einschränkungen, die sich bei der Programmierung auf dem PDA gegenüber dem PC zwangsläufig ergeben, abzuschwächen: PocketLogo bietet beispielsweise eine (konfigurierbare) Funktion, die

durch einfaches Antippen häufig verwendete "Codeschnipsel" an der aktuellen Cursorposition einfügt, etwa das Grundgerüst einer *repeat*-Schleife. Des Weiteren ist eine Verlaufsfunction implementiert, mit der durch die zuletzt getätigten Eingaben geblättert werden kann, um z.B. einzelne Parameter einer Befehlsfolge zu ändern.

4. Eigene Erweiterungen der Sprache Logo

Ich habe in PocketLogo darüber hinaus noch einige Erweiterungen implementiert, die meines Wissens noch kein anderer Logo-Interpreter bietet: Beispielsweise können (durch Programmcode oder per Stift) Wände definiert werden, die die Turtle nicht überschreiten kann. Es gibt einen Befehl, der die Entfernung der Turtle zur nächsten Wand in ihrer Blickrichtung zurückliefert. Außerdem kann die Turtle Objekte (Äpfel) aufnehmen und ablegen sowie die Anzahl der Äpfel in Reichweite prüfen, ähnlich den Kleeblättern des programmierbaren Marienkäfers Kara [Re07]. Diese Funktionen machen die PocketLogo-Turtle zu einer Art programmierbaren virtuellen Roboter, der mit seiner Umgebung interagieren kann – vergleichbar mit dem LEGO NXT-Roboter. Im "Informatik goes mobile"-Projekt erstellten die Schülerinnen und Schüler mit Hilfe der PDA-Software Feinstein [Ho07] ein Programm, das einen Lego-Roboter durch einen (unbekannten) Irrgarten navigieren lässt. Dieser Algorithmus ist mit wenigen Zeilen auch als Programm¹ für den "Turtle-Roboter" umzusetzen. Der real vorhandene Irrgarten kann in PocketLogo übernommen werden, indem er digital fotografiert und als Vorlage zum Nachzeichnen in PocketLogo geladen wird.

Logo kann aber nicht nur im Anfängerbereich eingesetzt werden. Eine der Leitlinien beim Entwurf der Sprache lautete "low threshold and no ceiling" [Te00, "What is Logo?"]. Dies bedeutet, dass Logo einerseits einen einfachen Einstieg in die Programmierung ermöglicht, die Sprache aber andererseits auch ausreichend mächtig für die Realisierung fortgeschrittener Konzepte ist, zum Beispiel...

- ...**Rekursion**: Die Turtlegrafik erlaubt, das Konzept der Rekursion anhand rekursiv definierter Fraktale sehr anschaulich darzustellen. In einem einfachen Binärbaum-Fraktal beispielweise entspricht jeder Ast einem rekursiven Aufruf.
- ...**funktionale Programmierung**, etwa der Entwurf von Funktionen zur Listenbearbeitung. Funktionen höherer Ordnung können ebenfalls definiert werden.

Meine Erweiterungen der Sprache Logo umfassen im Bereich der...

- ...**formalen Sprachen** die automatische Übersetzung² eines Lindenmayer-Systems wahlweise in ein rekursives oder nicht-rekursives Logoprogramm, welches das vom L-System induzierte Fraktal zeichnet. Das Erzeugen von (uninterpretierten) Symbolketten mit Hilfe der Regeln des L-Systems ist ebenfalls möglich.
- ...**theoretischen Informatik** die Möglichkeit, den von PocketLogo zur Interpretation eines Programms erzeugten Syntaxbaum anzuzeigen¹.

¹ Siehe <http://pocketlogo.sourceforge.net/bildschirmfotos.html> (Stand: 02.12.08)

² Siehe <http://pocketlogo.sourceforge.net/systeme.html> (Stand: 02.12.08)

5. Literaturverzeichnis

- [Bo08] Boytchev, P.: Logo Tree Project, Rev 1.56, December 2008. Online verfügbar: <http://www.elica.net/download/papers/LogoTreeProject.pdf> (zuletzt abgerufen: 16.12.08)
- [Bü07] Büdding, H.: Mobiles Lernen unter Verwendung von Handheld Computern im Bereich der Schulinformatik. In: Peer Stechert (Hrsg.): Informatische Bildung in der Wissensgesellschaft. Praxisband der 12. Fachtagung "Informatik und Schule - INFOS 2007", Siegen 2007. Online verfügbar: http://www.die.informatik.uni-siegen.de/lehrstuhl/stechert/publikationen/Informatische_Bildung_2007.pdf (24.08.08)
- [CGS99] Chakraborty, A., Graebner, R., Stocky, T.: Logo – A Project History. MIT, 1999. Online verfügbar: <http://web.mit.edu/6.933/www/LogoFinalPaper.pdf> (24.08.08)
- [Ho07] Homann, M.: Entwicklung einer Software für den Informatikunterricht zur Steuerung von Lego-NXT Roboter mittels PDAs. Schriftliche Hausarbeit Lehramt Sek. II Informatik. Münster 2007. Software online verfügbar: <http://www.m-learning.info/projekte/feinstein/> (02.12.08)
- [Hu08] Hunke, M.: Entwicklung eines Logo-Interpreters für Handhelds als Medium für den Informatikunterricht. Schriftliche Hausarbeit Lehramt Sek. II Informatik. Münster, 04.09.2008. Software und Handbuch online verfügbar: <http://pocketlogo.sourceforge.net/> (17.12.08)
- [Po06] Webseiten zu PocketCoder: <http://www.pocketcoder.de/> (02.12.08)
- [Re07] Reichert, R.: Internetseiten zu Kara – Lernumgebungen rund ums Programmieren. <http://www.swisseduc.ch/informatik/karatojava/> (02.12.08)
- [Te00] Tempel, M.: Internetseiten der Logo Foundation. <http://el.media.mit.edu/logo-foundation/index.html> (24.08.08)

Community-Verhalten im Micro-Blogging mit Twitter

Visualisierung zeitabhängiger Veränderungen in personellen & thematischen Netzwerken

David Maicher
Universität Paderborn
dmaicher@mail.upb.de

Daniel Scholz
Universität Paderborn
daniels@mail.upb.de

Abstract: Unsere Bachelorarbeit beinhaltet ein Konzept zur Visualisierung von Kommunikation über den Micro-Blogging-Dienst Twitter und die Umsetzung des Konzepts anhand einer Webseite. Am Beispiel von Twitter werden die Grundlagen von Micro-Blogging und von Visualisierungen erläutert. Darauf aufbauend wurde eine Webseite gestaltet, die in übersichtlichen Graphen eine komfortable Auswertung der Nachrichten und Benutzer von Twitter sowie deren zeitliche Veränderung ermöglicht.

Art der Arbeit: Bachelorarbeit
Betreuer: Wolfgang Reinhardt

1 Einführung

Thema unserer Bachelorarbeit ist die Visualisierung von Kommunikation über den Micro-Blogging-Dienst Twitter¹. Twitter ist ein Echtzeit-Kurznachrichten-Dienst, bei dem die registrierten Benutzer (User) von einer Vielzahl von Endgeräten und Programmen eine bis zu 140 Zeichen lange Nachricht verfassen und für sämtliche andere Nutzer verfügbar machen können. Die Nachrichten (so genannte „Tweets“) erscheinen sowohl in einer öffentlichen als auch in einer persönlichen zeitlichen Übersicht (Timeline). Twitter bietet aber auch die Möglichkeit der Direktnachricht, bei der die Nachricht nur zwischen zwei Personen ausgetauscht wird und somit nur für diese sichtbar ist. Die persönliche Timeline zeigt die Tweets aller User, die mit dem Verfasser ein Netzwerk bilden. Es gibt die Möglichkeit der direkten Antwort an den Verfasser einer Nachricht, sowie die Einordnung einer Nachricht zu einem oder mehreren Themen durch Schlagwörter (Tags).

¹<http://www.twitter.com>

2 Relevanz von Twitter

Durch seine weltweite Verbreitung findet Twitter in immer mehr Bereichen Verwendung. Anfang 2008 gab es eine größere Ausfallzeit während der MacWorld Conference & Expo, als Steve Jobs, CEO von Apple Inc., seine Eröffnungsrede hielt. Viele Zuhörer zitierten Sätze aus der Rede oder schrieben allgemein über die Äußerungen von Jobs. Dabei versahen sie ihre Nachrichten mit einem Tag, so dass andere User sämtliche Tweets über die Rede schnell finden konnten. Durch die Diskussion über die neuen Produkte entstand so viel Traffic, dass die Server von Twitter überlastet waren [Dor08]. Bei der diesjährigen MacWorld wurde Twitter ebenfalls als beliebtes Kommunikationstool benutzt, der Dienst brach dieses Mal aber nicht zusammen. Inzwischen findet Twitter aber nicht nur bei technologischen Events eine Vielzahl an Nutzern. Am 20.12.2008 ging ein Flugzeug in Denver während des Starts in Flammen auf, als es von der Startbahn abkam [Kre08]. Einer der Passagiere schrieb nur sieben Minuten nach dem Unglück einen Tweet [Wil08]. So schnell konnte keine Nachrichtenagentur reagieren. Nach und nach schrieb er in mehreren Tweets, wie er das Unglück erlebt hatte und ließ die Welt daran teilhaben.

Und nicht nur bei einzelnen Katastrophen, auch bei Kriegen wird Twitter zu einer immer stärker genutzten Plattform. Sowohl während des Kaukasus-Krieges als auch seit dem Einmarsch der Israelis in den Gaza-Streifen werden viele Nachrichten pro Minute über dieses Thema verfasst - sowohl von Außenstehenden als auch von Augenzeugen und Betroffenen. In den von Twitter herausgegebenen „Trending topics“ [twi09], in denen die aktuell am Häufigsten benutzten Wörter aufgeführt werden, befinden sich seitdem ständig die Tags „#gaza“ und „#peace“.

Diese neue Art des „Live-Journalismus“ ist nur ein Aspekt, der Twitter zu einem wichtigen Medium gemacht hat. Um in die große und unsortierte Masse der Tweets Ordnung zu bringen, visualisieren wir mit unserer Arbeit die Kommunikation über Twitter, so dass eine genaue Analyse der Nachrichten und User möglich ist.

3 Theoretische Ausarbeitung

Im theoretischen Teil der Bachelorarbeit beschreiben wir, was Micro-Blogging bedeutet und wie Twitter im Speziellen funktioniert. Zu Beginn erläutern wir, woher das Phänomen des Micro-Blogging stammt und welche Denkansätze dahinter stehen. Danach untersuchen wir die Motivation, aus der heraus Micro-Blogging-Dienste benutzt werden, und welche Vorteile daraus für die User entstehen können. Wir gehen darauf ein, wie diese Dienste bei Communities of Practice (CoP) und Communities of Interest (CoI) genutzt werden können. Hierbei behandeln wir im Besonderen den konkreten Einsatz bei Konferenzen und in Projektgruppen.

Weiterhin wird erläutert, welchen Sinn verschiedene Visualisierungen von Daten ergeben können und welche Arten von Visualisierungen es gibt. Der Schwerpunkt liegt dabei auf bereits vorhandenen Verknüpfungen (Mashups) von Visualisierungen mit Twitter. Hier haben wir die Vor- und Nachteile einiger prägnanter Mashups ausgearbeitet und zeigen die

Unterschiede beziehungsweise Gemeinsamkeiten mit unserer eigenen Implementierung auf.

4 Implementierung

Zunächst werden mit Hilfe eines Graphen die Tweets eines Users und seine genutzten Tags in Abhängigkeit zur Zeit dargestellt. Dieselben Daten können auch auf einer Timeline angezeigt werden. Als weitere Funktion ist es möglich diejenigen User zu visualisieren, die einen bestimmten Tag benutzen. Es lassen sich ebenfalls die Tags ansehen, die von einem bestimmten User benutzt wurden. Des Weiteren können die User, die die eigenen Nachrichten in ihrer persönlichen Timeline angezeigt bekommen (Followers), und die Freunde im Netzwerk eines Users (Following) in Abhängigkeit zur Zeit dargestellt werden. Um dieses zu erreichen sind die Verbindungen zwischen den Usern sichtbar gemacht worden. Hierfür nutzen wir Beziehungsgraphen, die Verbindungen zwischen verschiedenen Objekten in einem Netzwerk anzeigen können. Die Veränderungen des Netzwerks werden in einer Tabelle angezeigt. Hier lassen sich die hinzugekommenen und verlorenen Followers und Followings in einem bestimmten Zeitraum darstellen. So lässt sich die Dynamik eines solchen betrachteten Netzwerks gut darstellen.

Wir nutzen die von Twitter.com angebotene Schnittstelle (Application Programming Interface, API), um gezielt Informationen zu Usern, ihren Netzwerken und bestimmten Tags zu erhalten und diese dann graphisch darzustellen. Alle Visualisierungen werden auf einer Webseite angezeigt. Zur schnelleren Auswertung und Erzeugung der Graphen war es notwendig, gelieferte Daten der Twitter API lokal in einer Datenbank zu speichern. Die gezielte Abfrage und Speicherung dieser Daten wurde mit Hilfe von serverseitigen Skripten realisiert, die in bestimmten Zeitintervallen automatisch aufgerufen werden (Cronjob).

In Abbildung 1 sieht man die von uns erstellte Webseite. Links befindet sich die Navigation, in der man aus einer Übersicht über alle in unserer Datenbank vorhandenen User und Tags auswählen kann. Rechts befindet sich die Hauptansicht zum ausgewählten User @wollepb. Hier werden zuerst die zehn letzten Tweets angezeigt. Daneben befinden sich zeitabhängige Statistiken, die benutzten Tags und eine Term Extraction, in der die am häufigsten benutzten Begriffe aus den Tweets dargestellt werden. Unten lassen sich verschiedene Graphen über Tabs aufrufen. Im Bild ist ein Stacked Area Chart, in dem die am häufigsten benutzten Tags dargestellt werden.

5 Ausblick

Mit der von uns entwickelten Anwendung lässt sich die Kommunikation in Twitter leicht verfolgen und analysieren. Nachdem sich alle Teilnehmer einer Konferenz oder einer Projektgruppe auf ein gemeinsames Schlagwort geeinigt haben, kann man diesen Tag in unsere Datenbank aufnehmen. Nach dem nächsten Durchlauf der Cronjobs lassen sich eine Vielzahl von Informationen auf einen Blick erkennen. Es werden die neuesten Tweets, die



Abbildung 1: Die Webseite zu unserer Bachelorarbeit

mit diesem Tag versehen wurden, Statistiken und Graphen angezeigt. In einer Übersicht ist erkennbar, wer schon wie häufig etwas zu dem Thema beigetragen hat. Anhand der Timeline kann nun zum Beispiel ein Redner sehen, wer während seines Vortrags etwas anzumerken hatte, und kann kurzfristig darauf reagieren.

Besonders aktive oder anderweitig interessante Teilnehmer können ebenfalls in die Datenbank aufgenommen werden, um ihre Nachrichten und ihr Netzwerk auszuwerten. Der Beziehungsgraph zeigt anschaulich die Größe des Netzwerks und wie es sich seit dem Start oder einem anderen wichtigen Datum der Konferenz bzw. der Projektgruppe verändert hat. Falls bei den sonstigen Tags des Users ein weiteres relevantes Thema auftaucht, lässt sich leicht herausfinden, ob auch andere Teilnehmer Interesse daran zeigen. Dann könnte man dieses Thema in den Ablauf einer Konferenz einbauen.

Literatur

- [Dor08] Jack Dorsey. MacWorld. <http://blog.twitter.com/2008/01/macworld.html>, 2008.
- [Kre08] Matthias Kremp. Passagier twittet sofort nach Flugzeugunglück. <http://www.spiegel.de/netzwelt/mobil/0,1518,597935,00.html>, 2008.
- [twi09] twitter.com. Trending topics. <http://search.twitter.com/>, 2009.
- [Wil08] Mike Wilson. Tweet from Mike Wilson about the airplane crash. <http://twitter.com/2drinksbehind/status/1069832870>, 2008.

Ein Lernspiel zu grundlegenden Problemen in verteilten Systemen

Gabor Meißner

Friedrich-Schiller-Universität Jena

gabor.meissner@googlemail.com

Wissenschaftliche Hausarbeit zum Ersten Staatsexamen

Betreuer: Prof. Dr. Michael Fothe und Prof. Dr. Brigitta König-Ries

Abstract: Sowohl elektronische Lernspiele als auch der Bereich der verteilten Systeme erfahren in den letzten Jahren zunehmende Bedeutung. Die vorliegende Arbeit zeigt ein Lernspiel, das zentrale Probleme in verteilten Systemen thematisiert und eingesetzt im Unterricht nachhaltiges und motivierendes Lernen ermöglicht.

1 Einleitung

Im Rahmen der wissenschaftlichen Hausarbeit zum Ersten Staatsexamen für das Lehramt am Gymnasium wurde ein Programm entwickelt, das Lehrende dabei unterstützen soll, das Thema verteilte Systeme auf einfache, moderne und effektive Weise zu unterrichten. Dabei war zunächst zu untersuchen, welche Voraussetzungen für die Verwendung von Lernsoftware im Unterricht bestehen, welche Merkmale gute und aktuelle Lernsoftware hat und wie ein Programm sich speziell für das Thema verteilte Systeme modellieren lässt.

Festzustellen ist, dass verteilte Systeme bisher wenig Beachtung im Informatikunterricht finden. Trotzdem kann gezeigt werden, dass das Thema verteilte Systeme die fünf Kriterien für eine fundamentale Idee (für die Informatik ausführlich beschrieben in [SS04]) von Bruner[Bru60] erfüllt, die als Voraussetzung für die Thematisierung in allgemein bildenden Schulen angesehen werden [Mei08, S. 2 f]. Damit sollte die Relevanz für den Unterricht sichergestellt sein. Verbinden ließen sich verteilte Systeme sehr gut mit Lehrplanthemen wie Computernetze oder dem Internet.

2 Didaktische Konzeption

Es wurde eine Lernsoftware entwickelt, mit deren Hilfe vor allem Lernende leichter den Einstieg in die Thematik finden können und die sie zugleich auch besonders motiviert und nachhaltige Lernvorgänge ermöglicht. Die Motivation für die Entwicklung einer interaktiven Lernsoftware war unter anderem durch ermutigende Ergebnisse von Studien gegeben, die Lernerfolge nachweisen konnten. Besonders unter dem Aspekt der Transferleistungen schneidet gute Lernsoftware überdurchschnittlich ab (vgl. z.B. [Ker03]).

Zunächst wurden Kriterien für gute Lernsoftware gesucht, die sich aus einer Metaanalyse zahlreicher Publikationen zu E-Learning ergaben. Darunter zählen u.a. die Herstellung eines Problembezug, also ob es darum geht, ein System im Sinne eines Werkzeuges nur zu bedienen oder ob das Funktionieren eines solches Systems im Vordergrund steht. Mit der Art der Wissensvermittlung ist gemeint, ob sich die ausführenden Handlungen des Lernenden auf das Problem direkt beziehen oder nicht. Ein direkter Bezug ist dann hergestellt, wenn die Methoden zur Bedienung der Software den Methoden des zu simulierenden Problems ähneln.

Ausgehend von der These, dass mit Hilfe von Spielen die Motivation für das Lernen besonders hoch ist, wurde die Software als Spiel mit Bewertungssystem konzipiert. Das Lernen sollte weniger als mit herkömmliche Lernmethoden als solches wahrgenommen und abgelehnt werden, es sollte also versteckt sein (sog. *stealth teaching*). Weitere positive Aspekte von Lernen mit Spielen wurden etwa von Bopp an Hand von kommerziellen Computerspielen eindrucksvoll aufgezeigt (sie sind z.B. *situativ, selbstgesteuert, geplant und konstruktiv*)[Bop06].

Für das Programm wurde eine Spielewelt modelliert, in der die Lernenden in die Rolle eines Technikers schlüpfen, der in einer Stadt dabei helfen soll, die Kommunikationsprobleme von sich feindlich gegenüberstehenden Agenten zu lösen. Dazu wurden zunächst fünf Probleme modelliert, die der Benutzer zu bearbeiten hat.¹ Zum Verständnis der Probleme sollten diese in die Spielewelt inhaltlich eingepasst werden. Weiterhin müssen zur Verdeutlichung der Problemlage stets Animationen entwickelt werden, die das Problem und verschiedene Zustände der Lösung generisch erkennbar machen.

3 Grundlegenden Problemen in verteilten Systemen

Ein typisches Beispiel für ein häufiges Problem in verteilten Systemen ist der Verlust von Nachrichten. Grafisch wird das Problem über Nachrichten gezeichnet, die auf halber Wegstrecke verloren gehen und rot gefärbt sind (vgl. Abb. 1). Zur Lösung des Problems stehen dem Benutzer zwei Werkzeuge zur Verfügung. Eine Sendungswiederholung sendet die ursprüngliche Nachricht immer wieder. Der Stopp der Sendungswiederholung wird über ein weiteres Werkzeug erreicht. Zudem kann eine Nachrichtenbestätigung versendet werden.

Werkzeuge müssen im Programm Ereignissen zugeordnet werden. Ein Ereignis ist stets der Zeitpunkt vor dem Absenden und nach Empfang einer Nachricht von einem bestimmten Typ (vgl. Abb. 2).

Bei vielen Problemen ist es notwendig, dass verschiedene Nachrichtentypen definiert werden (hier: normale Nachricht und Bestätigung). Diese erhöhen die Anzahl möglicher Ereignisse und werden in der Animation mit verschiedenen Farben dargestellt. Das Programm erkennt den Lösungszustand entweder an bestimmten Regeln oder an einer vorgegebenen Konfiguration, anschließend werden die Punkte für den Benutzer errechnet und ausgegeben.

¹Durch den modularen Aufbau des Programms können weitere Probleme entwickelt werden.

PROBLEM	AUSGEWÄHLTER LÖSUNGSANSATZ
Verlorene Nachrichten	Nachrichtenwiederholungen und -bestätigung, Identifizierbarkeit von Nachrichten
Gefälschte Nachrichten	Nachrichtenauthentifizierung im Public-Key-Verfahren
Mitgelesene Nachrichten	doppelte Verschlüsselung im Public-Key-Verfahren
Unterschiedliche Nachrichtenreihenfolge bei verschiedenen Empfängern	Lamport-Zeitstempel, vollständig sortiertes Multicast[TvS03, S. 289 ff], [Mei08, S. 59 ff]
Inkonsistenz einer gemeinsam genutzten Ressource	Quoren-basierte Konsistenz[TvS03, S. 384-392 u. 405 f]

Tabelle 1: Übersicht über die verwendeten Probleme und Lösungsansätze



Abbildung 1: Die Animation zum Problem der verlorenen Nachrichten

Die weiteren Probleme (vgl. Tabelle 3), die bisher entwickelt wurden, sind das Mitlesen von Nachrichten, unterschiedliche Nachrichtenreihenfolgen bei verschiedenen Kommunikationspartnern, das Verfälschen von Nachrichten und das Inkonsistent-werden einer gemeinsam genutzten Ressource (vgl. dazu u.a. [TvS03]).

4 Einsatz im Unterricht

Die Software wurde so konzipiert, dass möglichst wenige Voraussetzungen für Schülerinnen und Schüler bestehen, erfolgreich mit dem Programm zu lernen. Für den Anfangsunterricht eignet sich das Programm nur bedingt, allerdings kann man es verwenden, sobald ein intuitiver Algorithmenbegriff bei den Lernenden aufgebaut wurde. Unumgänglich ist

Machbarkeitsstudie mit Analyse und Bewertung zur Durchführung von Onlinedurchsuchungen in der Praxis

Von Wolfgang Frisch

Wilhelm Büchner Hochschule
Informatik
wolfgang.frisch@mpq.mpg.de

Betreuer der Diplomarbeit:

Michael Best, Dipl. Inform. (FH):

Abstract: Gegenstand der Arbeit ist eine weitestgehend umfassende Analyse der Möglichkeiten einer Onlinedurchsuchung. Dieses Vorgehen wird im allgemeinen Sprachgebrauch auch als „**Bundestojaner**“ bezeichnet.

Die Arbeit befasst sich nicht nur mit den rein technischen Möglichkeiten zur Durchführung der Maßnahme. Zusätzlich werden begleitende Umstände, wie beispielsweise das Ausforschen des sozialen Umfeldes der Zielperson – auch „**Social-Engineering**“ genannt – in das Konzept mit einbezogen.

Ein großer Teil der Arbeit beschäftigt sich mit den Möglichkeiten zur Einbringung einer sogenannten „**Remote Forensic Software**“ – kurz **RFS** – auf ein Zielsystem.

Alle derzeit denkbaren Szenarien wie beispielsweise Exploits, Webapplikationen, SQL-Injektion, Cross-Site-Scripting, HTTP-Response-Splitting, ActiveX, Java, DNS-Spoofing werden auf ihre Tauglichkeit untersucht und kurz deren Wirkungsweise aufgezeigt.

Selbstverständlich werden auch Zugriffsmöglichkeiten via Netzwerk sowie unmittelbar am Rechner des Probanden umfassend beleuchtet.

Zu allen dargestellten Angriffsszenarien werden - soweit vorhanden - auch die entsprechenden Gegenmaßnahmen dargelegt sowie eine Bewertung des Schutzpotentials.

Zuletzt werden auch Möglichkeiten zur Umgehung einer Überwachung vorgestellt.

Wie aus den täglichen Nachrichten zu entnehmen ist, ist das Thema äußerst aktuell, brisant und in vieler Hinsicht breit gefächert..

Die Zugriffsmöglichkeiten und deren Gegenmaßnahmen bilden einen dynamischen Prozess, dessen Ende nicht abzusehen ist.

Nicht ganz außer Acht gelassen wurden deswegen auch die politischen, juristischen und sozialen Aspekte eines solchen Vorhabens.

Interessenkonflikt

Betrachtet man die unterschiedlichen Interessenslagen in einer staatlichen Gemeinschaft bezüglich Privatsphäre, Freiheit und allgemeiner Sicherheit, so sind schnell prinzipielle Unvereinbarkeiten gefunden.

Das hohe Gut der absoluten Sicherheit kann – wenn überhaupt - nur mittels permanenter Überwachung geschaffen werden. Dies wiederum steht im Widerspruch zu Freiheit und Privatsphäre.

Trotzdem sind die staatlichen Stellen bestrebt sich zumindest in begrenzten Ausnahmefällen möglichst umfassende Informationen von und über eine oder mehrere Personen zu beschaffen.

Eine wichtige Rolle spielt dabei die Überwachung der gesamten Telekommunikation der betreffenden Person(en).

1 Überwachungsmöglichkeiten

Zunächst galt es möglichst alle Formen einer Überwachung zu finden und weiterhin auf ihre Tauglichkeit zur überprüfen. Aufgrund dieser Recherchen fand sich zumindest eine grundsätzliche Unterscheidung bezüglich des Zuganges zum Zielsystem.

1.1 Physischer Zugang

Hier werden die Möglichkeiten zur Erlangung eines physischen Zuganges zum auszuspähenden Rechner untersucht. Desweiteren die technischen Vorgehensweisen bei einem unmittelbaren Physischen Zugang zum Zielsystem.

Im Wesentlichen wurden als Zugangsmöglichkeiten ein unerkannter Einbruch in die Wohnung, das Einsehen der Daten während einer Reparatur, sowie Zugriff durch eine scheinbar vertraute Person (Besucher).

Letztere Möglichkeit ist unter dem Begriff Social – Engineering bekannt und kann in sehr komplexen Beziehungsgeflechten ausarten.

Eine sehr beliebte Möglichkeit bei physischem Zugang ist der Einsatz von sogenannten Keyloggern. Diese sind sowohl als Hardware wie auch als Softwareversion erhältlich. Sie speichern die Daten über einen bestimmten Zeitraum und werden dann im Nachhinein ausgewertet. Die Softwareversion könnte die Daten auch per Internet übermitteln.

1.2 Nichtphysischer Zugang

Die Möglichkeiten des Zugriffes auf die Daten des Probanden ohne direkten Zugang zum System sind ungleich umfangreicher. Gleichzeitig erfordern sie einen hohen technischen Aufwand, welcher aber durch relativ gängige Hard- und Software prinzipiell einfach zur Verfügung steht.

Hier erfolgt der Zugriff entweder mittels Einschleusens eines entsprechenden Spionageprogrammes, welches dem Auszuspähenden z.B. über Disketten, Memory-Stick, CD, E-Mail oder über manipulierte Webseiten untergeschoben wird.

Dieses Wirtsprogramm sammelt dann im Hintergrund die gewünschten Daten und übermittelt diese dann an die entsprechenden Personen zur Auswertung.

In allen Fällen ist es ungemein schwierig, das Vorhandensein einer Überwachungssoftware auf einem infiltrierten System zu tarnen.

2 Fazit

Ein großer Teil der Arbeit zeigt die diversen, teilweise sehr diffizilen Vorgehensweisen um die sogenannte Spyware oder auch Trojaner auf ein Zielsystem unbemerkt zu installieren, zu betreiben, Daten zu sammeln und unbemerkt zu übermitteln.

Dies setzt eine sehr gute Vorbereitung, Kenntnisse über Abwehrmechanismen und sehr gutes Wissen über Betriebssysteme, Abläufe und nicht zu letzt ein detailliertes Wissen über Netzwerke voraus.

Die Arbeit betrachtet alle Aspekte des Vorgehens von der Seite des Eindringenden, wie auch die Möglichkeiten eines Anwenders, sich solcher Datenabgriffe zu erwehren.

Letztlich stellt sich der Zugriff, bzw. das Verhindern desselbigen als ein fortwährendes Katz- und Mausspiel heraus und der Ausgang ist nicht vorhersehbar, weil er von unzähligen und unabwägbaren Faktoren beeinflusst wird. Die Arbeit versucht alle zum Verfassungszeitpunkt relevanten Möglichkeiten gegeneinander abzuwägen. Dabei spielen auch Themen wie Gesetzgebung und technischer Fortschritt eine tragende Rolle.

Als Resümee steht daher die Aussage, daß eine Erfolgreiche Ausspähung von Daten sehr vom Wissen der jeweils Mitwirkenden abhängt. Ein moderner global vernetzter Gesetzesbrecher hat heute vermutlich durchaus die Möglichkeiten und das Wissen, sich gegen das Ausspähen seiner geheimsten Daten erfolgreich zu wehren.

Literaturverzeichnis

- [01] <http://vx.netlux.org/lib/mjk00.html>.
- [02] <http://www.bsi.de/bsi/historie.htm>
- [03] <http://www.it-recht-kanzlei.de/computerkriminalitaet-bekaempfung.html>
- [04] <http://www.heise.de/security/artikel/99749>
- [05] http://www.goldmann.de/firewall-so-funktioniert-eine-firewall_tipp_94.html
- [06] <http://www.luro.de/autostart.html>
- [07] http://en.wikipedia.org/wiki/Option_ROM
- [08] <http://www.littlesister.de/>
- [09] DATA Becker, Viren Würmer & Trojaner, (S 203)
- [10] www.ietf.org
- [11] <http://www.wintotal.de/softw/index.php?rb=33&id=1817>
- [12] <http://www.pc-safety.com/ddlock.html>
- [13] c't 8/03, S. 190
- [14] <http://www.immunitysec.com/downloads/shindig-2-butler-jamie.ppt>
- [15] c't 1/2008 S.92
- [16] <http://www.heise.de/ct/08/01/092/>
- [17] [http://www.heise.de/security/
Wie-Skype-Co-Firewalls-umgehen--/artikel/82054](http://www.heise.de/security/Wie-Skype-Co-Firewalls-umgehen--/artikel/82054)
- [18] <http://www.snort.org/>
- [19] <http://www.owasp.org>
- [20] <http://packetstormsecurity.org/papers/general/SQLInjectionWhitePaper.pdf>
- [21] <http://baseportal.com/baseportal/phishmarkt/de>
- [22] <http://www.blackbone-ev.de/XSS%20Angriffe%20per%20DOM.html>
- [23] [http://www.packetstormsecurity.org/papers/
general/whitepaper_httpresponse.pdf](http://www.packetstormsecurity.org/papers/general/whitepaper_httpresponse.pdf)
- [24] Das Anti-Hacker-Buch, ISBN-10: 3826681673 *Stuart McClure, Joel Scambray, George Kurtz*
- [25] <http://research.eeye.com/html/advisories/published/AD20030820.html>
- [26] <http://www.halcyon.com/mclain/ActiveX>
- [27] <http://www.heise.de/ct/03/21/links/108.shtml>
- [28] <http://www.informatik.uni-bremen.de/~sohr/papers/jit.pdf>

Identifizierung von Webseiten bei verschlüsselten Verbindungen

Dominik Sandjaja
RWTH Aachen - Lehrstuhl für Informatik IV
sandjaja@nets.rwth-aachen.de

Abstract: Benutzt man beim Surfen im Internet eine verschlüsselte Verbindung, wie zum Beispiel einen SSH-Tunnel oder ein Anonymisierungsnetzwerk wie Tor, so geht man davon aus, dass ein Lauscher, welcher den verschlüsselten Inhalt mitliest keine Möglichkeit hat, herauszufinden, welche Webseiten man besucht. Jedoch kann gezeigt werden, dass nur anhand der Analyse der Paketgrößen, Zeitinformationen und Senderrichtungen bereits mit hoher Genauigkeit genau dies möglich ist.

1 Einleitung

Mit der immer alltäglicher werdenden Nutzung des Internets wird auch die Menge an Daten, die jeder einzelne Nutzer beim Surfen hinterlässt, immer größer. Diese Datenmenge erleichtert es Institutionen wie Telekommunikationsunternehmen, Internet-Shops, aber auch staatlichen Autoritäten, immer genauere Aktivitätsprofile einzelner Nutzer zu erstellen. Immer mehr Nutzer möchten bei der Nutzung des Internets daher unerkannt bleiben. Für diesen Wunsch gibt es verschiedene Gründe, unter anderem politische. In vielen Staaten ist es Bürgern nicht erlaubt, das Internet frei zu nutzen und die eigene Meinung kann nur anonym abgegeben werden, ohne Repressionen fürchten zu müssen.

Eine populäre Möglichkeit, im World Wide Web anonym zu surfen, ist die Nutzung von Proxys. Diese verschleiern die IP-Adresse des Anschlusses, über den der Benutzer im Internet unterwegs ist. Dadurch ist der Betreiber eines Webserver, den der Nutzer besucht, nicht in der Lage, anhand der IP-Adresse den Inhaber der Leitung festzustellen, da am Webserver nur die Adresse des Proxys auftaucht.

Häufig verwendete verschlüsselte Übertragungstechniken sind SSH-Tunnel und das Anonymisierungsnetzwerk "The Onion Router" (Tor) [RP04]. Bei letzterem verbindet sich der Benutzer nicht nur zu einem Proxy und von dort aus direkt zum Ziel-Webserver, sondern er verbindet sich über einen Pfad von mehreren Proxys, üblicherweise drei. Der erste Proxy in dieser Kette heißt der Entry-Knoten, der letzte der Exit-Knoten. Der Datenverkehr zwischen diesen Proxys wird verschlüsselt, so dass nur der Entry-Knoten die tatsächliche IP-Adresse des Benutzers kennt und nur der Exit-Knoten weiß, welche Webseite abgerufen wird. Durch diesen Mechanismus, bei dem die Datenpakete wie bei einer Zwiebel in mehreren Schichten verschlüsselt werden, ist mehr Sicherheit im Vergleich zu einem einfachen Proxy wie einem SSH-Tunnel gewährleistet, da man nicht mehr einem einzelnen Proxy vertrauen muss.

Bei beiden vorgestellten Proxys geht ein Nutzer im Allgemeinen davon aus, dass auch ein lokaler Lauscher keine Informationen über die abgerufenen Inhalte erhalten kann. Genau dies ist jedoch, wie gezeigt wird, sehr einfach zu realisieren.

2 Problembeschreibung

Man betrachtet einen Angreifer, der zwischen dem Rechner des Benutzers und dem ersten Proxy die Verbindung mitschneidet. Dabei erhält er über alle IP-Pakete des gesamten Verkehrs folgende Informationen: Zeit, Größe und Senderrichtung. Simuliert der Angreifer nun solch eine Verbindung, kann er für jede Webseite, die er dem Benutzer später nachweisen möchte, eine Sammlung solcher Daten anlegen. Diese Daten kann er dann nutzen, um einen Klassifikator zu trainieren. Anschließend werden die erlernten Informationen genutzt, um im Datenstrom des Benutzers einzelne Webseiten zu identifizieren.

3 Vorgehensweise

Das Vorgehen gliedert sich in vier Abschnitte. Zuerst wird eine Sammlung von identifizierenden Merkmalen der einzelnen Webseiten angelegt. Hierzu wird mit dem Werkzeug TCPDump¹ der Datenverkehr mitgeschnitten. Die resultierenden Daten werden in leichter verwendbare Formate umgewandelt und gespeichert. Im zweiten Schritt wird ein Klassifikator trainiert. Ein solcher Klassifikator entstammt dem Bereich der Mustererkennung. Er dient dazu, bestimmte Objekte (hier die Instanzen, jede ein Abruf einer Webseite) einer bestimmten Klasse (hier die URL) zuzuordnen. Dabei wird der Klassifikator zunächst mit einer Menge von Trainingsdaten trainiert. Das für diesen Zweck eingesetzte Toolkit ist WEKA [Wit05]. Der nächste Schritt besteht darin, tatsächliche Daten des Benutzers zu sammeln. Das Vorgehen entspricht dem, welches bereits im ersten Schritt angewendet wurde.

Am Ende wird der vorher trainierte Klassifikator dazu verwendet, die vom Benutzer gesammelten Daten zu klassifizieren. Dazu wird geschaut, mit welcher Genauigkeit die Datensätze, welche beim Benutzer mitgeschnitten wurden, vom Klassifikator erkannt werden.

In Versuchen wurde der Klassifikator mit Trainingsinstanzen trainiert. Eine Instanz ist der Abruf einer Webseite aus einer Menge von ca. 800 Webseiten über eine bestimmte Proxy-Art. Anschließend wurde mit einer Test-Datenmenge von Instanzen, welche disjunkt zu den Trainingsdaten war, jedoch mit demselben Verfahren gesammelt wurde, überprüft, wieviel Prozent der Instanzen korrekt klassifiziert wurden. Eine Klassifizierung war korrekt, wenn die zur Instanz gehörige URL korrekt erkannt wurde.

Obwohl Proxys, insbesondere aber Anonymisierungsnetzwerke wie Tor, vorgeben, gegen einen solchen lokalen Angreifer, wie er simuliert wurde zu schützen, waren die Erkennungsraten hoch. Insbesondere beim SSH-Tunnel wurden Erkennungsraten von 93,6 %

¹<http://www.tcpdump.org>

erreicht. Dagegen lagen die Erkennungsraten bei einem abgeschotteten Tor-Netzwerk mit 47,55 % um einiges niedriger. Noch schlechter ist das Ergebnis bei Versuchen im echten Tor-Netzwerk. Dort sank die Erkennungsrate auf unter 2 % ab. Die Erkennungsraten sind in Abbildung 1 gegenübergestellt.

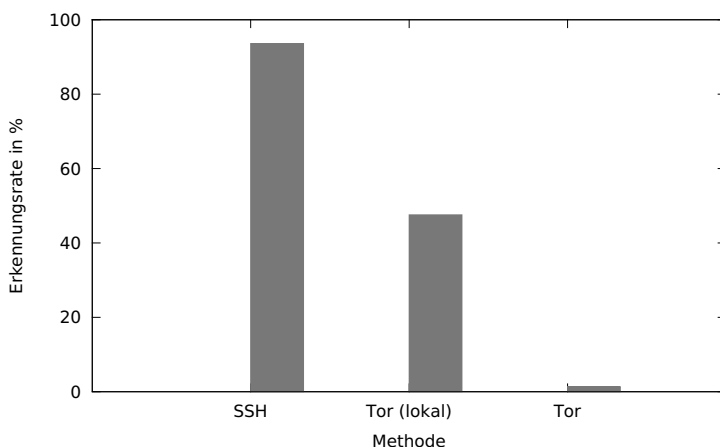


Abbildung 1: Erkennungsraten

Ungeachtet dessen, dass das Problem des vorgestellten Angriffs ein offensichtlich sehr einfach umzusetzendes Verfahren ist, welches keinen starken Angreifer voraussetzt, wurde es bisher nur wenig untersucht. [SSW⁺02] beachtete die grundsätzliche Problematik, Webseiten in verschlüsselten Verbindungen erkennen zu können, das erste Mal. Da die Forschung jedoch bereits 2002 durchgeführt wurde, basierte der Angriff auf der Analyse der Größe der Objekte auf der Webseite. Diese Möglichkeit wurde mit Einführung und Verbreitung von HTTP/1.1 [RJJ⁺99] obsolet, da nun nicht mehr für jedes Objekt eine einzelne TCP-Verbindung aufgebaut wurde. In [Hin02] wurde das erste Mal ein solcher Angriff untersucht, der auf der Analyse von Paketgrößen basiert, es wurde jedoch nur ein einzelner Proxy untersucht. Auch war die Anzahl der untersuchten Webseiten sehr niedrig. Weitere Untersuchungen wurden in [LL06] und [Her08] durchgeführt. Dabei wurden verschiedene Proxy-Verfahren untersucht und die Erkennungsmethoden immer weiter verbessert.

4 Ausblick

Die bisher vorgestellten Methoden, um Webseiten in einem Strom von verschlüsselten Daten zu identifizieren, sind in der Praxis noch nicht anwendbar. Daher wird es interessant sein, inwiefern sich der Angriff soweit verbessern lässt, dass er auch außerhalb kontrollierter Bedingungen erfolgreich ist. Dazu sind Erkenntnisse, welche Klassifikatoren bessere Erkennungsraten ermöglichen, nötig. Desweiteren sollte untersucht werden, ob auch eine

Erkennung möglich ist, wenn über eine Verbindung mehrere Vorgänge zeitgleich stattfinden, es also zu Überlappungen in den Datenpaketen kommt. Es ist weiterhin zu klären, ob das Hinzuziehen zeitlicher Informationen die Erkennungsraten verbessern kann.

Schlussendlich steht auch noch zur Frage, wie sich der Angriff verhindern lässt. Erste Ideen zielen auf eine Veränderung der vom Proxy an den Benutzer gesendeten Daten, zum Beispiel in festen Paketgrößen, ab. Jedoch kann es dadurch zu einer beträchtlichen Erhöhung des zu übertragenden Datenvolumens kommen, wie bereits in [LL06] gezeigt wurde. Die Entwicklung von weiteren Methoden, um dem Benutzer ein höheres Maß an Anonymität zu gewährleisten ohne dabei die Performance der Verbindung zu stark zu verringern, steht daher noch aus.

Literatur

- [Her08] Dominik Herrmann. Analyse von datenschutzfreundlichen Übertragungstechniken hinsichtlich ihres Schutzes vor Datenverkehrsanalysen im Internet. Diplomarbeit, Universität Regensburg, Februar 2008.
- [Hin02] Andrew Hintz. Fingerprinting Websites Using Traffic Analysis. In Roger Dingledine und Paul Syverson, Hrsg., *Proceedings of Privacy Enhancing Technologies workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.
- [LL06] Marc Liberatore und Brian Neil Levine. Inferring the Source of Encrypted HTTP Connections. In *Proceedings of the 13th ACM conference on Computer and Communications Security (CCS 2006)*, Seiten 255–263, Alexandria, Virginia, USA, October 2006.
- [RJJ⁺99] Robert Fielding, Jim Gettys, Jeff Mogul, Henrik Frystyk, Larry Masinter, Paul Leach und Tim Berners-Lee. Hypertext Transfer Protocol: HTTP/1.1. Internet Engineering Task Force: RFC 2616, June 1999.
- [RP04] Nick Mathewson Roger Dingledine und Paul Syverson. Tor: The Second-Generation Onion Router. In *Proceedings of the 13th USENIX Security Symposium*, 2004.
- [SSW⁺02] Qixiang Sun, Daniel R. Simon, Yi-Min Wang, Wilf Russell, Venkata N. Padmanabhan und Lili Qiu. Statistical Identification of Encrypted Web Browsing Traffic. In *S&P '02: Proceedings of the 2002 IEEE Symposium on Security and Privacy*, Seite 19, Washington, DC, USA, 2002. IEEE Computer Society.
- [Wit05] Witten, Ian H. and Frank, Eibe. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2. Auflage, 2005.

GI-Edition Lecture Notes in Informatics – Proceedings

- P-1 Gregor Engels, Andreas Oberweis, Albert Zündorf (Hrsg.): Modellierung 2001.
- P-2 Mikhail Godlevsky, Heinrich C. Mayr (Hrsg.): Information Systems Technology and its Applications, ISTA'2001.
- P-3 Ana M. Moreno, Reind P. van de Riet (Hrsg.): Applications of Natural Language to Information Systems, NLDB'2001.
- P-4 H. Wörn, J. Mühling, C. Vahl, H.-P. Meinzer (Hrsg.): Rechner- und sensorgestützte Chirurgie; Workshop des SFB 414.
- P-5 Andy Schürr (Hg.): OMER – Object-Oriented Modeling of Embedded Real-Time Systems.
- P-6 Hans-Jürgen Appelrath, Rolf Beyer, Uwe Marquardt, Heinrich C. Mayr, Claudia Steinberger (Hrsg.): Unternehmen Hochschule, UH'2001.
- P-7 Andy Evans, Robert France, Ana Moreira, Bernhard Rumpe (Hrsg.): Practical UML-Based Rigorous Development Methods – Countering or Integrating the extremists, pUML'2001.
- P-8 Reinhard Keil-Slawik, Johannes Magenheimer (Hrsg.): Informatikunterricht und Medienbildung, INFOS'2001.
- P-9 Jan von Knop, Wilhelm Haverkamp (Hrsg.): Innovative Anwendungen in Kommunikationsnetzen, 15. DFN Arbeitstagung.
- P-10 Mirjam Minor, Steffen Staab (Hrsg.): 1st German Workshop on Experience Management: Sharing Experiences about the Sharing Experience.
- P-11 Michael Weber, Frank Kargl (Hrsg.): Mobile Ad-Hoc Netzwerke, WMAN 2002.
- P-12 Martin Glinz, Günther Müller-Luschnat (Hrsg.): Modellierung 2002.
- P-13 Jan von Knop, Peter Schirmbacher and Viljan Mahni_ (Hrsg.): The Changing Universities – The Role of Technology.
- P-14 Robert Tolksdorf, Rainer Eckstein (Hrsg.): XML-Technologien für das Semantic Web – XSW 2002.
- P-15 Hans-Bernd Bludau, Andreas Koop (Hrsg.): Mobile Computing in Medicine.
- P-16 J. Felix Hampe, Gerhard Schwabe (Hrsg.): Mobile and Collaborative Business 2002.
- P-17 Jan von Knop, Wilhelm Haverkamp (Hrsg.): Zukunft der Netze –Die Verletzbarkeit meistern, 16. DFN Arbeitstagung.
- P-18 Elmar J. Sinz, Markus Plaha (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2002.
- P-19 Sigrid Schubert, Bernd Reusch, Norbert Jesse (Hrsg.): Informatik bewegt – Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 30.Sept.-3.Okt. 2002 in Dortmund.
- P-20 Sigrid Schubert, Bernd Reusch, Norbert Jesse (Hrsg.): Informatik bewegt – Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 30.Sept.-3.Okt. 2002 in Dortmund (Ergänzungsband).
- P-21 Jörg Desel, Mathias Weske (Hrsg.): Promise 2002: Prozessorientierte Methoden und Werkzeuge für die Entwicklung von Informationssystemen.
- P-22 Sigrid Schubert, Johannes Magenheimer, Peter Hubwieser, Torsten Brinda (Hrsg.): Forschungsbeiträge zur "Didaktik der Informatik" – Theorie, Praxis, Evaluation.
- P-23 Thorsten Spitta, Jens Borchers, Harry M. Sneed (Hrsg.): Software Management 2002 – Fortschritt durch Beständigkeit
- P-24 Rainer Eckstein, Robert Tolksdorf (Hrsg.): XMIDX 2003 – XML-Technologien für Middleware – Middleware für XML-Anwendungen
- P-25 Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Commerce – Anwendungen und Perspektiven – 3. Workshop Mobile Commerce, Universität Augsburg, 04.02.2003
- P-26 Gerhard Weikum, Harald Schöning, Erhard Rahm (Hrsg.): BTW 2003: Datenbanksysteme für Business, Technologie und Web
- P-27 Michael Kroll, Hans-Gerd Lipinski, Kay Melzer (Hrsg.): Mobiles Computing in der Medizin
- P-28 Ulrich Reimer, Andreas Abecker, Steffen Staab, Gerd Stumme (Hrsg.): WM 2003: Professionelles Wissensmanagement – Erfahrungen und Visionen
- P-29 Antje Düsterhöft, Bernhard Thalheim (Eds.): NLDB'2003: Natural Language Processing and Information Systems
- P-30 Mikhail Godlevsky, Stephen Liddle, Heinrich C. Mayr (Eds.): Information Systems Technology and its Applications
- P-31 Arslan Brömmke, Christoph Busch (Eds.): BIOSIG 2003: Biometric and Electronic Signatures

- P-32 Peter Hubwieser (Hrsg.): Informatische Fachkonzepte im Unterricht – INFOS 2003
- P-33 Andreas Geyer-Schulz, Alfred Taudes (Hrsg.): Informationswirtschaft: Ein Sektor mit Zukunft
- P-34 Klaus Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenberg, Wolfgang Wahlster (Hrsg.): Informatik 2003 – Innovative Informatikanwendungen (Band 1)
- P-35 Klaus Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenberg, Wolfgang Wahlster (Hrsg.): Informatik 2003 – Innovative Informatikanwendungen (Band 2)
- P-36 Rüdiger Grimm, Hubert B. Keller, Kai Rannenberg (Hrsg.): Informatik 2003 – Mit Sicherheit Informatik
- P-37 Arndt Bode, Jörg Desel, Sabine Rathmayer, Martin Wessner (Hrsg.): DeLFI 2003: e-Learning Fachtagung Informatik
- P-38 E.J. Sinz, M. Plaha, P. Neckel (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2003
- P-39 Jens Nedon, Sandra Frings, Oliver Göbel (Hrsg.): IT-Incident Management & IT-Forensics – IMF 2003
- P-40 Michael Rebstock (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2004
- P-41 Uwe Brinkschulte, Jürgen Becker, Dietmar Fey, Karl-Erwin Großpietsch, Christian Hochberger, Erik Maehle, Thomas Runkler (Edts.): ARCS 2004 – Organic and Pervasive Computing
- P-42 Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Economy – Transaktionen und Prozesse, Anwendungen und Dienste
- P-43 Birgitta König-Ries, Michael Klein, Philipp Obreiter (Hrsg.): Persistence, Scalability, Transactions – Database Mechanisms for Mobile Applications
- P-44 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): Security, E-Learning, E-Services
- P-45 Bernhard Rumpe, Wolfgang Hesse (Hrsg.): Modellierung 2004
- P-46 Ulrich Flegel, Michael Meier (Hrsg.): Detection of Intrusions of Malware & Vulnerability Assessment
- P-47 Alexander Prosser, Robert Krimmer (Hrsg.): Electronic Voting in Europe – Technology, Law, Politics and Society
- P-48 Anatoly Doroshenko, Terry Halpin, Stephen W. Liddle, Heinrich C. Mayr (Hrsg.): Information Systems Technology and its Applications
- P-49 G. Schiefer, P. Wagner, M. Morgenstern, U. Rickert (Hrsg.): Integration und Datensicherheit – Anforderungen, Konflikte und Perspektiven
- P-50 Peter Dadam, Manfred Reichert (Hrsg.): INFORMATIK 2004 – Informatik verbindet (Band 1) Beiträge der 34. Jahrestagung der Gesellschaft für Informatik e.V. (GI), 20.-24. September 2004 in Ulm
- P-51 Peter Dadam, Manfred Reichert (Hrsg.): INFORMATIK 2004 – Informatik verbindet (Band 2) Beiträge der 34. Jahrestagung der Gesellschaft für Informatik e.V. (GI), 20.-24. September 2004 in Ulm
- P-52 Gregor Engels, Silke Seehusen (Hrsg.): DELFI 2004 – Tagungsband der 2. e-Learning Fachtagung Informatik
- P-53 Robert Giegerich, Jens Stoye (Hrsg.): German Conference on Bioinformatics – GCB 2004
- P-54 Jens Borchers, Ralf Kneuper (Hrsg.): Softwaremanagement 2004 – Outsourcing und Integration
- P-55 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): E-Science und Grid Ad-hoc-Netze Medienintegration
- P-56 Fernand Feltz, Andreas Oberweis, Benoit Otjacques (Hrsg.): EMISA 2004 – Informationssysteme im E-Business und E-Government
- P-57 Klaus Turowski (Hrsg.): Architekturen, Komponenten, Anwendungen
- P-58 Sami Beydeda, Volker Gruhn, Johannes Mayer, Ralf Reussner, Franz Schweiggert (Hrsg.): Testing of Component-Based Systems and Software Quality
- P-59 J. Felix Hampe, Franz Lehner, Key Pousttchi, Kai Ranneberg, Klaus Turowski (Hrsg.): Mobile Business – Processes, Platforms, Payments
- P-60 Steffen Friedrich (Hrsg.): Unterrichtskonzepte für informatische Bildung
- P-61 Paul Müller, Reinhard Gotzhein, Jens B. Schmitt (Hrsg.): Kommunikation in verteilten Systemen
- P-62 Federrath, Hannes (Hrsg.): „Sicherheit 2005“ – Sicherheit – Schutz und Zuverlässigkeit
- P-63 Roland Kaschek, Heinrich C. Mayr, Stephen Liddle (Hrsg.): Information Systems – Technology and its Applications

- P-64 Peter Liggesmeyer, Klaus Pohl, Michael Goedicke (Hrsg.): Software Engineering 2005
- P-65 Gottfried Vossen, Frank Leymann, Peter Lockemann, Wolffried Stucky (Hrsg.): Datenbanksysteme in Business, Technologie und Web
- P-66 Jörg M. Haake, Ulrike Lucke, Djamshid Tavangarian (Hrsg.): DeLFI 2005: 3. deutsche e-Learning Fachtagung Informatik
- P-67 Armin B. Cremers, Rainer Manthey, Peter Martini, Volker Steinhage (Hrsg.): INFORMATIK 2005 – Informatik LIVE (Band 1)
- P-68 Armin B. Cremers, Rainer Manthey, Peter Martini, Volker Steinhage (Hrsg.): INFORMATIK 2005 – Informatik LIVE (Band 2)
- P-69 Robert Hirschfeld, Ryszard Kowalczyk, Andreas Polze, Matthias Weske (Hrsg.): NODE 2005, GSEM 2005
- P-70 Klaus Turowski, Johannes-Maria Zaha (Hrsg.): Component-oriented Enterprise Application (COAE 2005)
- P-71 Andrew Torda, Stefan Kurz, Matthias Rarey (Hrsg.): German Conference on Bioinformatics 2005
- P-72 Klaus P. Jantke, Klaus-Peter Fährnrich, Wolfgang S. Wittig (Hrsg.): Marktplatz Internet: Von e-Learning bis e-Payment
- P-73 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): "Heute schon das Morgen sehen"
- P-74 Christopher Wolf, Stefan Lucks, Po-Wah Yau (Hrsg.): WEWoRC 2005 – Western European Workshop on Research in Cryptology
- P-75 Jörg Desel, Ulrich Frank (Hrsg.): Enterprise Modelling and Information Systems Architecture
- P-76 Thomas Kirste, Birgitta König-Riess, Key Poustchi, Klaus Turowski (Hrsg.): Mobile Informationssysteme – Potentiale, Hindernisse, Einsatz
- P-77 Jana Dittmann (Hrsg.): SICHERHEIT 2006
- P-78 K.-O. Wenkel, P. Wagner, M. Morgens-tern, K. Luzi, P. Eisermann (Hrsg.): Land- und Ernährungswirtschaft im Wandel
- P-79 Bettina Biel, Matthias Book, Volker Gruhn (Hrsg.): Softwareengineering 2006
- P-80 Mareike Schoop, Christian Huemer, Michael Rebstock, Martin Bichler (Hrsg.): Service-Oriented Electronic Commerce
- P-81 Wolfgang Karl, Jürgen Becker, Karl-Erwin Großpietsch, Christian Hochberger, Erik Maehle (Hrsg.): ARCS'06
- P-82 Heinrich C. Mayr, Ruth Breu (Hrsg.): Modellierung 2006
- P-83 Daniel Huson, Oliver Kohlbacher, Andrei Lupas, Kay Nieselt and Andreas Zell (eds.): German Conference on Bioinformatics
- P-84 Dimitris Karagiannis, Heinrich C. Mayr, (Hrsg.): Information Systems Technology and its Applications
- P-85 Witold Abramowicz, Heinrich C. Mayr, (Hrsg.): Business Information Systems
- P-86 Robert Krimmer (Ed.): Electronic Voting 2006
- P-87 Max Mühlhäuser, Guido Röbling, Ralf Steinmetz (Hrsg.): DELFI 2006: 4. e-Learning Fachtagung Informatik
- P-88 Robert Hirschfeld, Andreas Polze, Ryszard Kowalczyk (Hrsg.): NODE 2006, GSEM 2006
- P-90 Joachim Schelp, Robert Winter, Ulrich Frank, Bodo Rieger, Klaus Turowski (Hrsg.): Integration, Informationslogistik und Architektur
- P-91 Henrik Stormer, Andreas Meier, Michael Schumacher (Eds.): European Conference on eHealth 2006
- P-92 Fernand Feltz, Benoît Otjacques, Andreas Oberweis, Nicolas Poussing (Eds.): AIM 2006
- P-93 Christian Hochberger, Rüdiger Liskowsky (Eds.): INFORMATIK 2006 – Informatik für Menschen, Band 1
- P-94 Christian Hochberger, Rüdiger Liskowsky (Eds.): INFORMATIK 2006 – Informatik für Menschen, Band 2
- P-95 Matthias Weske, Markus Nüttgens (Eds.): EMISA 2005: Methoden, Konzepte und Technologien für die Entwicklung von dienstbasierten Informationssystemen
- P-96 Saartje Brockmans, Jürgen Jung, York Sure (Eds.): Meta-Modelling and Ontologies
- P-97 Oliver Göbel, Dirk Schadt, Sandra Frings, Hardo Hase, Detlef Günther, Jens Nedon (Eds.): IT-Incident Mangament & IT-Forensics – IMF 2006

- P-98 Hans Brandt-Pook, Werner Simonsmeier und Thorsten Spitta (Hrsg.): Beratung in der Softwareentwicklung – Modelle, Methoden, Best Practices
- P-99 Andreas Schwill, Carsten Schulte, Marco Thomas (Hrsg.): Didaktik der Informatik
- P-100 Peter Forbrig, Günter Siegel, Markus Schneider (Hrsg.): HDI 2006: Hochschuldidaktik der Informatik
- P-101 Stefan Böttinger, Ludwig Theuvsen, Susanne Rank, Marlies Morgenstern (Hrsg.): Agrarinformatik im Spannungsfeld zwischen Regionalisierung und globalen Wertschöpfungsketten
- P-102 Otto Spaniol (Eds.): Mobile Services and Personalized Environments
- P-103 Alfons Kemper, Harald Schöning, Thomas Rose, Matthias Jarke, Thomas Seidl, Christoph Quix, Christoph Brochhaus (Hrsg.): Datenbanksysteme in Business, Technologie und Web (BTW 2007)
- P-104 Birgitta König-Ries, Franz Lehner, Rainer Malaka, Can Türker (Hrsg.): MMS 2007: Mobilität und mobile Informationssysteme
- P-105 Wolf-Gideon Bleek, Jörg Raasch, Heinz Züllighoven (Hrsg.): Software Engineering 2007
- P-106 Wolf-Gideon Bleek, Henning Schwentner, Heinz Züllighoven (Hrsg.): Software Engineering 2007 – Beiträge zu den Workshops
- P-107 Heinrich C. Mayr, Dimitris Karagiannis (eds.): Information Systems Technology and its Applications
- P-108 Arslan Brömme, Christoph Busch, Detlef Hühnlein (eds.): BIOSIG 2007: Biometrics and Electronic Signatures
- P-109 Rainer Koschke, Otthein Herzog, Karl-Heinz Rödiger, Marc Ronthaler (Hrsg.): INFORMATIK 2007 Informatik trifft Logistik Band 1
- P-110 Rainer Koschke, Otthein Herzog, Karl-Heinz Rödiger, Marc Ronthaler (Hrsg.): INFORMATIK 2007 Informatik trifft Logistik Band 2
- P-111 Christian Eibl, Johannes Magenheimer, Sigrid Schubert, Martin Wessner (Hrsg.): DeLFI 2007: 5. e-Learning Fachtagung Informatik
- P-112 Sigrid Schubert (Hrsg.): Didaktik der Informatik in Theorie und Praxis
- P-113 Sören Auer, Christian Bizer, Claudia Müller, Anna V. Zhdanova (Eds.): The Social Semantic Web 2007 Proceedings of the 1st Conference on Social Semantic Web (CSSW)
- P-114 Sandra Frings, Oliver Göbel, Detlef Günther, Hardo G. Hase, Jens Nedon, Dirk Schadt, Arslan Brömme (Eds.): IMF2007 IT-incident management & IT-forensics Proceedings of the 3rd International Conference on IT-Incident Management & IT-Forensics
- P-115 Claudia Falter, Alexander Schliep, Joachim Selbig, Martin Vingron and Dirk Walther (Eds.): German conference on bioinformatics GCB 2007
- P-116 Witold Abramowicz, Leszek Maciszek (Eds.): Business Process and Services Computing 1st International Working Conference on Business Process and Services Computing BPSC 2007
- P-117 Ryszard Kowalczyk (Ed.): Grid service engineering and management The 4th International Conference on Grid Service Engineering and Management GSEM 2007
- P-118 Andreas Hein, Wilfried Thoben, Hans-Jürgen Appelrath, Peter Jensch (Eds.): European Conference on ehealth 2007
- P-119 Manfred Reichert, Stefan Strecker, Klaus Turowski (Eds.): Enterprise Modelling and Information Systems Architectures Concepts and Applications
- P-120 Adam Pawlak, Kurt Sandkuhl, Wojciech Cholewa, Leandro Soares Indrusiak (Eds.): Coordination of Collaborative Engineering - State of the Art and Future Challenges
- P-121 Korbinian Herrmann, Bernd Bruegge (Hrsg.): Software Engineering 2008 Fachtagung des GI-Fachbereichs Softwaretechnik
- P-122 Walid Maalej, Bernd Bruegge (Hrsg.): Software Engineering 2008 - Workshopband Fachtagung des GI-Fachbereichs Softwaretechnik

- P-123 Michael H. Breitner, Martin Breunig, Elgar Fleisch, Ley Pousttchi, Klaus Turowski (Hrsg.)
Mobile und Ubiquitäre Informationssysteme – Technologien, Prozesse, Marktfähigkeit
Proceedings zur 3. Konferenz Mobile und Ubiquitäre Informationssysteme (MMS 2008)
- P-124 Wolfgang E. Nagel, Rolf Hoffmann, Andreas Koch (Eds.)
9th Workshop on Parallel Systems and Algorithms (PASA)
Workshop of the GI/ITG Special Interest Groups PARS and PARVA
- P-125 Rolf A.E. Müller, Hans-H. Sundermeier, Ludwig Theuvsen, Stephanie Schütze, Marlies Morgenstern (Hrsg.)
Unternehmens-IT:
Führungsinstrument oder Verwaltungsbürde
Referate der 28. GIL Jahrestagung
- P-126 Rainer Gimnich, Uwe Kaiser, Jochen Quante, Andreas Winter (Hrsg.)
10th Workshop Software Reengineering (WSR 2008)
- P-127 Thomas Kühne, Wolfgang Reisig, Friedrich Steimann (Hrsg.)
Modellierung 2008
- P-128 Ammar Alkassar, Jörg Siekmann (Hrsg.)
Sicherheit 2008
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 4. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)
2.-4. April 2008
Saarbrücken, Germany
- P-129 Wolfgang Hesse, Andreas Oberweis (Eds.)
Sigsand-Europe 2008
Proceedings of the Third AIS SIGSAND European Symposium on Analysis, Design, Use and Societal Impact of Information Systems
- P-130 Paul Müller, Bernhard Neumair, Gabi Dreö Rodosek (Hrsg.)
1. DFN-Forum Kommunikationstechnologien Beiträge der Fachtagung
- P-131 Robert Krimmer, Rüdiger Grimm (Eds.)
3rd International Conference on Electronic Voting 2008
Co-organized by Council of Europe, Gesellschaft für Informatik and E-Voting.CC
- P-132 Silke Seehusen, Ulrike Lucke, Stefan Fischer (Hrsg.)
DeLFI 2008:
Die 6. e-Learning Fachtagung Informatik
- P-133 Heinz-Gerd Hegering, Axel Lehmann, Hans Jürgen Ohlbach, Christian Scheideler (Hrsg.)
INFORMATIK 2008
Beherrschbare Systeme – dank Informatik Band 1
- P-134 Heinz-Gerd Hegering, Axel Lehmann, Hans Jürgen Ohlbach, Christian Scheideler (Hrsg.)
INFORMATIK 2008
Beherrschbare Systeme – dank Informatik Band 2
- P-135 Torsten Brinda, Michael Fothe, Peter Hubwieser, Kirsten Schlüter (Hrsg.)
Didaktik der Informatik – Aktuelle Forschungsergebnisse
- P-136 Andreas Beyer, Michael Schroeder (Eds.)
German Conference on Bioinformatics GCB 2008
- P-137 Arslan Brömme, Christoph Busch, Detlef Hühnlein (Eds.)
BIOSIG 2008: Biometrics and Electronic Signatures
- P-138 Barbara Dinter, Robert Winter, Peter Chamoni, Norbert Gronau, Klaus Turowski (Hrsg.)
Synergien durch Integration und Informationslogistik
Proceedings zur DW2008
- P-139 Georg Herzwurm, Martin Mikusz (Hrsg.)
Industrialisierung des Software-Managements
Fachtagung des GI-Fachausschusses Management der Anwendungsentwicklung und -wartung im Fachbereich Wirtschaftsinformatik
- P-140 Oliver Göbel, Sandra Frings, Detlef Günther, Jens Nedon, Dirk Schadt (Eds.)
IMF 2008 - IT Incident Management & IT Forensics
- P-141 Peter Loos, Markus Nüttgens, Klaus Turowski, Dirk Werth (Hrsg.)
Modellierung betrieblicher Informationssysteme (MobIS 2008)
Modellierung zwischen SOA und Compliance Management
- P-142 R. Bill, P. Korduan, L. Theuvsen, M. Morgenstern (Hrsg.)
Anforderungen an die Agrarinformatik durch Globalisierung und Klimaveränderung
- P-143 Peter Liggesmeyer, Gregor Engels, Jürgen Münch, Jörg Dörr, Norman Riegel (Hrsg.)
Software Engineering 2009
Fachtagung des GI-Fachbereichs Softwaretechnik

- P-144 Johann-Christoph Freytag, Thomas Ruf,
Wolfgang Lehner, Gottfried Vossen
(Hrsg.)
Datenbanksysteme in Business,
Technologie und Web (BTW)
- P-145 Knut Hinkelmann, Holger Wache (Eds.)
WM2009: 5th Conference on Professional
Knowledge Management
- P-146 Markus Bick, Martin Breunig,
Hagen Höpfner (Hrsg.)
Mobile und Ubiquitäre
Informationssysteme – Entwicklung,
Implementierung und Anwendung
4. Konferenz Mobile und Ubiquitäre
Informationssysteme (MMS 2009)
- P-147 Witold Abramowicz, Leszek Maciaszek,
Ryszard Kowalczyk, Andreas Speck (Eds.)
Business Process, Services Computing
and Intelligent Service Management
BPSC 2009 · ISM 2009 · YRW-MBP 2009

GI-Edition Lecture Notes in Informatics – Seminars

- S-1 Johannes Magenheimer, Sigrid Schubert
 (Eds.):
 Informatics and Student Assessment
 Concepts of Empirical Research and
 Standardisation of Measurement in the
 Area of Didactics of Informatics
- S-2 Gesellschaft für Informatik (Hrsg.)
 Informationstage 2005
 Fachwissenschaftlicher Informatik-
 Kongress
- S-3 Gesellschaft für Informatik (Hrsg.)
 Informationstage 2006
 Fachwissenschaftlicher Informatik-
 Kongress
- S-4 Hans Hagen, Andreas Kerren, Peter
 Dannenmann (Eds.)
 Visualization of Large and Unstructured
 Data Sets
 First workshop of the DFG's International
 Research Training Group "Visualization
 of Large and Unstructured Data Sets –
 Applications in Geospatial Planning,
 Modeling and Engineering"
- S-5 Gesellschaft für Informatik (Hrsg.)
 Informationstage 2007
 Fachwissenschaftlicher Informatik-
 Kongress
- S-6 Gesellschaft für Informatik (Hrsg.)
 Informationstage 2008
 Fachwissenschaftlicher Informatik-
 Kongress
- S-7 Hans Hagen, Martin Hering-Bertram,
 Christoph Garth (Eds.)
 Visualization of Large and Unstructured
 Data Sets
- S-8 Gesellschaft für Informatik (Hrsg.)
 Informatiktage 2009
 Fachwissenschaftlicher Informatik-
 Kongress

The titles can be purchased at:

Köllen Druck + Verlag GmbH

Ernst-Robert-Curtius-Str. 14 · D-53117 Bonn

Fax: +49 (0)228/9898222

E-Mail: druckverlag@koellen.de

