Gesellschaft für Informatik e.V. (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
• seminars
• proceedings
• dissertations
• thematics
current topics are dealt with from the vantage point of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure high quality contributions.

The volumes are published in German or English.

Information: http://www.gi.de/service/publikationen/lni/

ISSN 1617-5468
ISBN 978-3-88579-298-7

In 2012, the annual International Conference on Innovative Internet Community Systems (I²CS) comes to the knowledge and technology capital of Norway, Trondheim. By its motto "Future requires Interdisciplinarity" the organizers continue to combine contributions on foundations, technology, applications and socialization of virtual communities. Besides paper presentations, discussions, brainstorming and other, creative activities are enabled. The present LNI volume contains 18 carefully selected contributions of this conference.

G. Eichler, L.W.M. Wienhofen, A. Kofod-Petersen, H. Unger (Eds.): I²CS 2012

204

GI-Edition

Lecture Notes in Informatics

Gerald Eichler, Leendert W. M. Wienhofen, Anders Kofod-Petersen, Herwig Unger (Eds.)

12th International Conference on Innovative Internet Community Systems (I²CS 2012)

June 13 - 15, 2012
Trondheim, Norway

Proceedings

Gerald Eichler, Leendert W. M. Wienhofen,
Anders Kofod-Petersen, Herwig Unger (Eds.)

12th International Conference on Innovative Internet Community Services (I^2CS 2012)

June 13 - 15, 2012
Norwegian University of Science and Technology
Trondheim, Norway

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-204

ISBN 978-3-88579-298-7
ISSN 1617-5468

Volume Editors
Dipl.-Ing. Gerald Eichler
Deutsche Telekom AG, Laboratories
D-64295 Darmstadt, Germany, Email: gerald.eichler@telekom.de
Leendert W. M. Wienhofen
NTNU, Department of Computer and Information Science
NO-7491 Trondheim, Norway, Email: leendert.wienhofen@idi.ntnu.no
Anders Kofod-Petersen
SINTEF, ICT
NO-7463 Trondheim, Norway, Email: anders.kofod-petersen@sintef.no
Prof. Dr. Herwig Unger
Fernuniversität Hagen, Faculty for Mathematics and Informatics
D-58084 Hagen, Germany, Email: herwig.unger@fernuni-hagen.de

printed by Köllen Druck+Verlag GmbH, Bonn

Foreword

The Innovative Internet Community Systems (I^2CS) has grown successfully from a national German workshop, founded in 2001, to a small but valuable international conference. The conference alternates between international and German locations, which are selected by the steering committee. The annual three-day event takes place around the second half of June. Its name is its mission: to make an innovative community comprising of scientists, researchers, service providers and vendors. This year, the community goes north from the German capital towards the polar circle.

The close corporation with the Gesellschaft für Informatik e.V. (GI) to publish the presented papers as Lecture Notes in Informatics (LNI) proceedings is continued. A strong peer review by two to three members of the program committee guarantees the high quality of contributions. The present proceedings are structured in six topics, covering the selection of the best submitted papers from 2012. Please refer to LNI P-148, P-165 and P-186 for the previous I^2CS proceedings. To receive a copy of a dedicated volume, please email your request to lni@i2cs-conference.org.

The width of I^2CS topics for 2012 spans a unique choice of aspects, bundled into the three classical areas: “foundations”, “technology”, “applications and socialisation”. Ambient Assisted Living (AAL) is given a strong focus in this year.

Foundations – Theories, models, algorithms for communities

- Distributed algorithms and simulation models
- Game theory, graph theory and cost models
- Innovative communication protocols
- Self organization and self stabilization
- Security and privacy protection
- Swarm intelligence and collaborative behaviour

- Small world models and cloud computing

Technology – Distributed architectures and frameworks

- Service oriented architectures and QoS for communities
- Peer-to-peer, grid and innovative architectures
- Distributed community and middleware in the cloud
- Software agents and adaptive systems
- Web x.0, Enterprise x.0 and Wikis
- Community management in ad-hoc environments
- Information retrieval, text processing and ontologies

Applications and socialization – Communities on the move

- Mobile Internet applications and user experience
- Context and location awareness
- eHealth challenges and ambient assisted living
- Personalization, social media and search
- Social and business aspects of user generated content
- Recommender solutions and expert profiles
- Domain specific languages for semantic design

Many thanks to the members of the program committee and all the volunteers for their flexible support during the preparation phase of I^{2}CS 2012.

The 13th I^{2}CS conference, hosted by the FernUniversität in Hagen, will take place in Iserlohn, Germany from June 19th through 21st, 2013. Please check the permanent conference URL http://www.i2cs-conference.org/ regularly for more details. Any new ideas and proposals for the future of the I^{2}CS are welcome to request@i2cs-conference.org.

With kind regards on behalf of the steering committee and the editors' board

Darmstadt, June 2012

Gerald Eichler, volume editor in chief
Deutsche Telekom AG, Laboratories, Berlin & Darmstadt, Germany

Preface to the 12th I²CS 2012, Trondheim, Norway

It is a great pleasure for us to welcome all participants of the 12th International Conference on Innovative Internet Community Systems – I²CS 2012 – to Trondheim. On behalf of the organizing and program committee, we hope that you will have nice and unforgettable days in the knowledge and technology capital of Norway, Trondheim. "Future Requires Interdisciplinarity" is the motto of the 12th I²CS. The selected venue, the Natural Science Building at the Norwegian University of Science and Technology (NTNU) in Trondheim, is a modern university building, containing all the facilities required for both tuition and conferences.

Following the initial concept of keeping the conference small and make it possible for all participants to get to know each other, we selected 18 full papers for presentation. It is the culture of I²CS to bring participants together, not only in order to share their latest results and experiences, but also to get to know and to learn about and of each other.

We will not contribute to the tendency that conferences just serve the purpose of increasing the authors' publication count. We offer an inspiring and unique atmosphere in which everybody will feel comfortable. As the past I²CS events have shown, this guarantees an excellent platform to lift the community idea. Hence, participants will not only attend our scientific sessions, but will also spend as much time as possible together to discuss, share opinions or define projects ideas in brainstorming meetings.

Last, but not least, a well-chosen cultural program around our sessions will give the participants the possibility to get in contact with both the historical and modern Trondheim. We are pleased to see that, over the years, many colleagues join our conference series over and over again, but in the same manner we appreciate and welcome all people who participate in I²CS for the first time and hope they will remain in contact with our community. Next year, the I²CS conference will be held in Iserlohn, Germany, and we hope we meet again there.

Yours I²CS 2012 organizing team Trondheim, June 2012

Leendert W. M. Wienhofen
Norwegian University of Science and Technology, Norway

Gerald Eichler
Deutsche Telekom AG, T-Labs, Germany

Anders Kofod-Petersen
SINTEF ICT, Norway

I²CS Program Committee 2012

Gilbert Babin	HEC Montreal	Canada
Thomas Böhme	Technical University of Ilmenau	Germany
Gerald Eichler	Telekom Innovation Laboratories, Berlin	Germany
Christian Erfurth	University of Applied Science, Jena	Germany
Hacène Fouchal	Universiy of Reims	France
Wolfgang Halang	FernUniversität in Hagen/GI	Germany
Hagen Höpfner	Bauhaus-University of Weimar	Germany
Philippe Hunel	Université des Antilles et de la Guyane	Martinique (F)
Janusz Kacprzyk	Polish Academy of Science	Poland
Peter Kropf	University Neuchâtel	Switzerland
Kyamakya Kyandoghere	University of Klagenfurt	Austria
Andreas Landmark	NTNU, Trondheim	Norway
Ulrike Lechner	Universität der Bundeswehr, Neubiberg	Germany
Franz Lehner	University of Passau/GI	Germany
Ernesto William De Luca	University of Applied Sciences Potsdam	Germany
Bjørn Magnus Mathisen	SINTEF ICT, Trondheim	Norway
Phayung Meesad	King Mongkut's UoT North Bangkok	Thailand
Marius Mikalsen	SINTEF ICT, Trondheim	Norway
Armin Mikler	University of Northern Texas, TX	U.S.A.
Anders Kofod-Petersen	SINTEF ICT, Trondheim	Norway
Christian Prehofer	FhG ESK/LMU Munich	Germany
Davy Preuveneers	Distrinet, KU Leuven	Belgium
Lior Rokach	Ben Gurion University, Beer Sheva	Israel
Wilhelm Rossak	Friedrich Schiller University Jena	Germany
Jörg Roth	Georg-Simon-Ohm-Hochschule Nürnberg	Germany
Harald Sack	HPI, Potsdam University	Germany
Volkmar Schau	Friedrich Schiller University Jena	Germany
Holger Schilder	nexum AG/GI	Germany
Kit-Sang Tang	City University Hong Kong	Hong Kong
Pieter Toussaint	NTNU, Trondheim	Norway
Herwig Unger	FernUniversität in Hagen	Germany
Leendert W. M. Wienhofen	NTNU, Trondheim	Norway

Additional Reviewers

Heinz Erretkamps	Neda Petreska	Martin Radzuweit
Nadine Steinmetz	Michael Stiller	

I²CS Steering Committee 2012

Thomas Böhme	TU Ilmenau	Germany
Gerald Eichler	Telekom Innovation Laboratories, Berlin	Germany
Herwig Unger	FernUniversität in Hagen	Germany
Leendert W. M. Wienhofen	NTNU, Trondheim	Norway

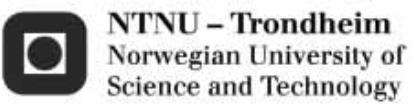
I²CS Organizing Committee 2012

Leendert W. M. Wienhofen, NTNU, Trondheim (Conference chair)
Gerald Eichler, Telekom Innovation Laboratories, Berlin (Conference co-chair)
Anders Kofod-Petersen, SINTEF ICT, Trondheim (Program chair)

Cooperation

Hosted, organized and sponsored by

NTNU
http://www.ntnu.edu/



Co-Organized by

SINTEF
http://www.sintef.no/

Telekom Innovation Laboratories
http://www.laboratories.telekom.com/

Supported by

Gesellschaft für Informatik e.V.
http://www.gi-ev.de/

Table of Contents

KEYNOTE

Making Money out of Innovative Communities on the Internet

Tom Arnøy & Per Gunnar Auran

CEO & Senior Data Scientist at Zedge
Ladebekken 17
NO-7041 Trondheim, Norway
tom@zedge.net, pauran@zedge.net

Abstract: Online advertising has been around for more than a decade, and today most Internet users are familiar with the contextual ad targeting in online search, of which Google AdWords is a dominant example. The pay per click model (PPC) introduced by Goto.com (later Overture) in 1998 has been the main revenue model since Google adopted it in 2002, but the advertising industry is under increasing pressure to find a better alternative model as the conversion rate for ad clicks to actual purchase is very low.

There is a huge potential for improving the experience as seen from both, advertiser, publisher and end user point of view. Enter the cost per install model where advertising cost is directly tied to the response of potential customers. While this reduces risk for advertiser spending, it demands precise customer targeting to create added value for the end user and secure revenue for the publisher.

The authors present a system for content-based targeting where the user behaviour within Zedge's massive mobile community is the core signal for segmenting users into preference groups that are used for targeted advertisements. Large scale clustering using Hadoop and machine learning is combined to optimize the performance.

Biography: Tom Arnoy is the CEO and Co-founder of Zedge. He has been intimately involved in overseeing product development, strategy and technology in Zedge, and has since 1998 been involved in multiple Internet and mobile start-ups.

Tom is a passionate product centric entrepreneur, always believing that things can change and improve to the better.

Biography: Dr.-Ing. Per Gunnar Auran received his PhD on 3D sonar imaging and object analysis in 1996 from the Norwegian University of Science and Technology, and his MSc on parallel image processing using transputers in 1991 from the Norwegian Institute of Technology, Trondheim. He has worked with large scale search relevancy since early 2000 and recommendation and personalization systems since 2004 for Fast Search and Transfer AS, Overture and Yahoo! Search.

In 2009 he co-founded a start-up company specializing in recommender systems whose technology was acquired by Zedge in 2010 where he now is the Data Analysis team lead.

INVITED TALK

Urban Traffic System, Computer Networks and PowerGrids: Different Systems – Same Solutions?

Herwig Unger

FernUniversität in Hagen, Chair of Communication Networks
Universitätsstraße 11
D-58097 Hagen, Germany
herwig.unger@fernuni-hagen.de

Abstract: In the last decade, technical and logistic systems became more and more specialised, complex and in most cases globally distributed. Already today, it is not possible to oversee or control them from any centralised instance.

Nevertheless, from the modelling point of view and their mathematical background those networks seem to have a lot in common. The author gives a comparison of urban traffic systems, peer-to-peer computer networks and power grids. In a set of examples will be shown, how an interdisciplinary application of methods and principles from one system may contribute to the progress of other ones.

Last but not least, it is intended to show that self-organisation will be the key issue to guarantee the efficient work of those systems in the future. We argue that self-organisation not only come along with structure building but also demonstrate that a suitable structure building may support the emergence of work division and cooperation.

Biography: Prof. Dr.-Ing. habil. Herwig Unger received his PhD with a work on Petri Net transformation in 1991 from the Technical University of Ilmenau/Germany and his habilitation with a work on large distributed systems from the University of Rostock/Germany in 2000. Since 2006 he is a full professor at the FernUniversität in Hagen and the head of the Chair for Communication Networks.

His research interests are in self-organization, adaptive and learning systems, Internet algorithms and simulation systems.

INVITED TALK

Semantic Search – Adaptation of Knowledge Models

Stein L. Tomassen

Program Manager II at Microsoft
Abelsgate 5
NO-7030 Trondheim, Norway
steint@microsoft.com

Abstract: Search applications have for many become part of their daily tools, empowering the users with easy access to vast amount of information. Despite the commonality with these tools finding quality information can still be both tedious and frustrating often involving multiples queries and analysis of extensive amounts of information. Semantic search is believed to solve many of the shortcomings found with current search applications, but still not commonly used. Semantic search has been a research area for a while, yet there are many issues to solve and improvements to make.

In this talk, the author will present a variety of semantic search systems that focus on adaptation of knowledge models. The author will present different approaches for how knowledge models can be enriched or created for a domain and how they can be used in search to increase the user experience.

Biography: Stein L. Tomassen finalized his PhD in 2011 at Norwegian University of Science and Technology (NTNU). His work was titled 'Conceptual Ontology Enrichment for Web Information Retrieval'. Prior to his PhD he worked three years for SINTEF ICT where his main focus was on mobile applications in relation to user context in particular.

In 2012, he started working for Microsoft Development Center Norway as a Program Manager. His research interests include semantic technologies, text mining, user context, ontology and information quality evaluation, in general; and improvement of information retrieval quality using ontologies, in particular. His publications include more than 20 articles in international refereed books, journals, conferences and workshops.

SESSION 1

Social Networks and Mobility

Sharing Personal Symbolic Locations between Friends – A Location Service for Small Communities

Jörg Roth

Department of Computer Science
Univ. of Applied Sciences Nuremberg
Kesslerplatz 12
90489 Nuremberg, Germany
Joerg.Roth@Ohm-Hochschule.de

Abstract: In this paper we introduce a novel community service based *on personal symbolic locations*. Users can share their location with other people using simple textual distribution mechanisms without the need of a geographic map display service. Small communities can set up a light-weight infrastructure and do not need to send private locations to big central service centres.

1 Introduction

People often want to know *where* their friends or family members currently are. Many public services such as *Google Latitude* or *facebook* support communities if they want to share their locations. Such services are mainly based on so-called *physical positions* (i.e. geographic coordinates) and usually show locations on a geographic map.

People have some difficulties with physical positions. Only few people know their current physical co-ordinate or the co-ordinate of their home. Instead they use terms such as 'at home' to indicate their location. Locations that are described as texts are called *symbolic locations*. Some community services already support symbolic locations when users check in e.g. into a restaurant. Such symbolic locations are defined by a global directory that contains publicly relevant points of interests.

We go into another direction: our approach is based on the idea of *personal symbolic locations* which are defined by users and only have a meaning for small communities. Typical personal symbolic locations are 'home', 'office' or 'shopping'. In contrast to physical locations, they can be distributed and displayed using purely textual mechanisms.

2 Related Work

Location-based services such Google Latitude allow users to track friends. Small icons represent the friends' locations painted on a background map that is loaded from, e.g.,

Google Maps. Users define a bidirectional *friend* relation and can decide to temporarily sign off from the service if they want to become 'invisible'. Friends automatically see updated position without any interaction of the tracked user. A manual 'send my location' function allows a user to distribute her or his location to any other user (not only friends) by e.g. email or SMS. The receiver gets an URL that opens a Google Maps view in which a marker shows the user's position.

Often, social services are combined with so-called *check-in services*. Places to check in are, e.g. restaurants, cinemas or public places. *Google Latitude* uses its social network *Google+*; services such as *facebook*, *foursquare* or *Gowalla* follow a similar idea. Users can automatically check-in based on their precise measured position or can manually check in via a web-based front end. The latter approach ensures that the user really resides at a certain location. Typical locations are small compared to position measurements errors, thus due to the measurement jitter, a full automatic system may fail. In addition, GPS fails indoors, but most check-in-locations are indoor locations.

Other projects try to identify symbolic locations from physical positions. In [Ve09] three major contexts are distinguished: 'home', 'work' and 'on the move'. In [BC+08] visited locations are stored in a kind of diary. This can by classified later by a Bayesian network that also takes into account the time of visit. A similar approach follows [LW+06], but it first maps physical locations to semantic locations with the help of a central geo database. [AS03] presents an approach to geometrically cluster GPS locations of multiple users to identify interesting locations. Projects such as *CybreMinder* [DA00], *Nexus* [HH+10] or [SH+06] provide a general platform to capture and to distribute location information. Usually they base on a centralized infrastructure that collects information and executes the rules to trigger events. *ContextPhone* [RO+05] primarily provides a system to plug-in context components on a mobile device. If a platform relies on central services, privacy issues become more important.

Connecto [BB+08] assigns names to physical locations that can be shared with friends. Each name still represents a single point in space. In contrast, our personal symbolic locations represent (potentially large) areas, also cities, city centres or districts. This difference has implications on privacy issues, as friends only know that users are *somewhere* inside an area, not the exact position. Privacy can also be achieved by rules that control how to share locations in groups. In [JH+02] the 'Principle of Minimum Asymmetry' is suggested that is in short: users that provide useful location information to the group will get the same amount of information back from other users. We do not support such mechanisms as they are often difficult to manage and to understand by users. Especially, different users have certain roles in communities and often cannot achieve the required symmetry between receiving and providing data. We strongly believe that our notion of personal symbolic locations is much easier to understand and allows communities to control privacy issues themselves without 'magic' and potentially intransparent control mechanisms.

Besides symbolic and physical locations there is a third type to define friends' locations: the *mailing address*. Actually, it could be summarized under symbolic locations, as it also can be transmitted via purely textual channels. Mobile platforms have easy-to-use

APIs to get the mailing address from a physical location. The drawbacks: only users that are close to buildings with a mailing address can be located. Second: often mailing addresses do not have a meaning for other users, thus the recipient again has to use a map service to locate an address.

3 The Zonezz Location Service

The *Zonezz Location Service* picks up the idea of *personal symbolic locations*. It can be installed on a smart phone (currently an Android phone). It is deployed as app, but mainly provides a background tasks that runs without any user intervention. It periodically detects the user's current symbolic location that is distributed to community members.

The Zonezz Location Services is based on the *Zonezz Core* as presented in [Ro11]. The Core was created to provide a local provider of personal symbolic locations for other apps that run on the respective smart phone. This is useful, if, e.g. the smart phone's home screen should change between a Job and Hobby view or if a calendar app should present job-related dates more intrusively when at work. The original Zonezz Core only locally communicates with components on the same smart phone. Our new extension, the Zonezz Location Service now communicates across the network and allows users to share their current location with other users.

As described in the related work section, the basic idea of sharing locations is realized in many platforms. However, the Zonezz Location Service has some important differences to other services:

- We use symbolic locations instead of physical positions. The locations are transmitted as text string such as 'job' or 'home' and not as physical location such as 49.465778 North, 11.158749 East.

- The symbolic locations are user-defined and *not* globally unique. Whereas other platforms map physical to symbolic locations with the help of a central database of *Points of Interests*, our symbolic locations are individually defined. Each user defines his own set of locations that are only stored on his smart phone. There is no central repository of locations at all. Location names may be *non*-disjoint, e.g. two users may define a location 'home', but with a different meaning.

- We only consider small communities. Each community creates its own distribution infrastructure that is *not* shared with other communities. This means, there is no big central server that holds locations for thousands or millions of people.

- We incorporated the notion of *trigger events* into the concept. Not only to *be* somewhere is important – to *change* the location is also important.

- The mobile user has full control over the distributed data and can adjust the published location information in a fine-grained manner. She or he can decide which locations to be presented to other users and which not.

We believe that especially the distribution of user-defined texts provides a significant benefit compared to existing map-based systems that only share physical locations.

As smart phone platforms do not provide a fully stable environment, there cannot be any guarantee for location information delivery. Thus, Zonezz is *not* designed for mission-critical applications, payment or access control. An example for the intended usage is a family that wants to get overview about each other's locations. Useful events could be: 'Child John enters school at 7:55', 'Dad leaves the office at 17:33', 'Ma enters the city center at 12:05'.

3.1 The Architecture

Zonezz starts without user intervention whenever the smart phone boots up. It installs itself as a background service, i.e. it periodically runs its code even though the smart phone is in stand-by. Usually, after a short configuration, user can 'forget' this program as it never requests any attention.

It is important that the service does not consume too much CPU power and battery load. The background task wakes up every 2 minutes to check the location. Zonezz especially supports the 'network' positioning system that consumes considerably less battery than GPS. Between the checks, the phone is switched to low power mode. The check as such does not significantly affect battery lifetime. Note that on typical smart phones a lot of background services periodically work.

Fig. 1 presents the architecture with the following components:

Position Providers are components of the smart phone operating system that measure the current physical position. Typical smart phones have a GPS provider and a 'network' provider which detects the current position with the help of signal strength fingerprints (of WLANs and Cell phone systems).

A *Position State* register contains information about the last measurement. State information may be: *positioning system error*, *last fix too old or last measurement is new and valid*.

The component *Hit Zones and Delta Zones Computation* maps a physical position to personal symbolic locations called *zones* (see section 3.2). It geometrically checks if one or more user-defined zones cover the current physical position. That leads to a set of *hit zones*. The *delta zones* reflect changes of the hit zone set. If we have two hit zone sets hit_n, hit_{n+1} of consecutive time stamps t_n, t_{n+1}, the delta zones contain the two sets $entered_{n+1}=hit_{n+1} \setminus hit_n$ and $left_{n+1}=hit_n \setminus hit_{n+1}$.

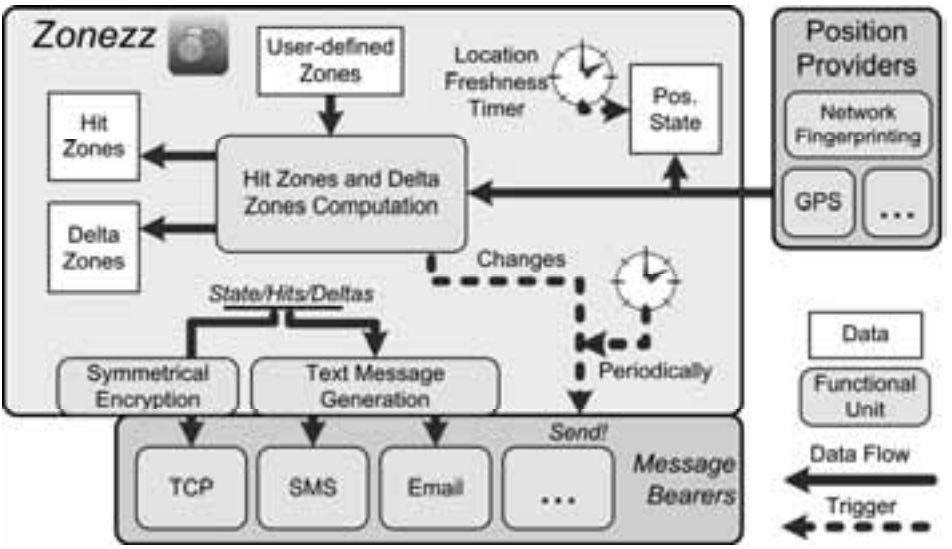


Figure 1: Service architecture

Periodically or whenever $entered_{n+1}$ or $left_{n+1}$ are not empty, the position state and the sets of hit zones and delta zones are transmitted using the *Message Bearers*. The bearers can easily be extended. A bearer has to implement a software interface that requests two functions: one to send location data and one to open a configuration dialog, to set e.g. the target address.

The transmission of location information can be done in two ways: *textual* or *raw*. Textual means: the transmitted message is human-readable without any further processing. This requires a mapping of the location information into a readable text such as 'Jörg enters the office at 8:03'. This is performed on the sending smart phone with the component *Text Message Generation*. Examples for this usage are emails or SMS messages that Zonezz can automatically emit. If in contrast the receiver again is a software component (especially our *Aggregator*, see section 3.3), the bearer uses a raw mode that only transmits the location data without any additional explainable texts. In raw mode, the location information is symmetrically encrypted using a community password concatenated with an initialization vector. This provides a basic security.

3.2 Expressing Personal Symbolic Locations with Zones

In [Ro06, Ro08] we already presented the benefits of symbolic locations for communities. In this paper we now endorse *personal* symbolic locations as a perfect tool to share locations in small communities such as families or between friends. Each user defines a set of *zones* which are two-dimensional circular regions on the Earth's surface with arbitrary center and radius (fig. 2). The zone geometry is only stored on the user's device and never published. Each zone has an arbitrary zone name that is the only data transmitted to other users. Usually a user defines zones such as 'home', 'job' or 'office', 'shopping', or hobby zones such as 'gym'.

Figure 2: Definition of zones

The model is easy to understand by the mobile user and allows the execution of efficient algorithms. To define a new zone, a user points on a map and specifies name and radius. For the radius, there exist discrete values 50m, 100m, 200m,…, 50km. It is important to note that the map is only required during zone configuration.

One could argue that circles usually do not precisely represent location borders, but more complex areas such as polygons are difficult to enter and to administrate on a mobile device. In addition, for circular regions the system can easily suggest new zones based on a position history (fig. 2 right). The user can query the system for position clusters that can be added to the zone list.

Another benefit of circular regions: Position measurements are Gaussian distributed around a real position. Areas that include all position measurements of a certain area tend to circles anyway, even though the original areas are not circles. Thus, the actual benefit for more complex geometries would be small.

Every time when a measured position is inside such a zone circle, the corresponding zone is considered as part of the current location. As zones may overlap, a location is defined by a set of hit zones. We expect the hit zone with the smallest radius as most specific for the context, thus the set of hit zones is ordered ascending by their radius.

Crossing the border of a zone is important information, sometimes more meaningful than the static set of hit zones. If between two consecutive timestamps the delta zones are not empty, the user changed at least one of his zones. This always means movement. Movement often is more interesting for community members. E.g., to be in the office at a working day is obvious. But leaving the office at 18:10 is an important information, e.g. for family members. Here, our concept has a huge advantage compared to services that are based on physical positions: as physical positions virtually change permanently due to measurement fluctuations real movement detection is more difficult.

3.3 Aggregating Personal Symbolic Locations

Transmitting Emails or SMSs is an easy way to send locations and change triggers without any additional processing component. However, if multiple users should be informed, this type of transmission is not suitable. For this scenario a so-called *Aggregator* can collect locations of a certain community and serves as a central information point.

The Aggregator performs simple access control, stores the positioning states, hit zones and delta zones of all community members and emits trigger events. In addition, it creates formatted information pages (e.g. HTML) that reflect the current situation.

From the network point of view, an Aggregator in fact is a server. But compared to, e.g. a Web server, the aggregator is very light-weight with only 18 kB of Java code. Thus, it could even run on small or embedded devices such as a router, any PC or on the *Infowall* screen (see below). In principle, the Aggregator could also run on a user's smart phone, but currently, mobile phone networks do not support server sockets on mobile devices.

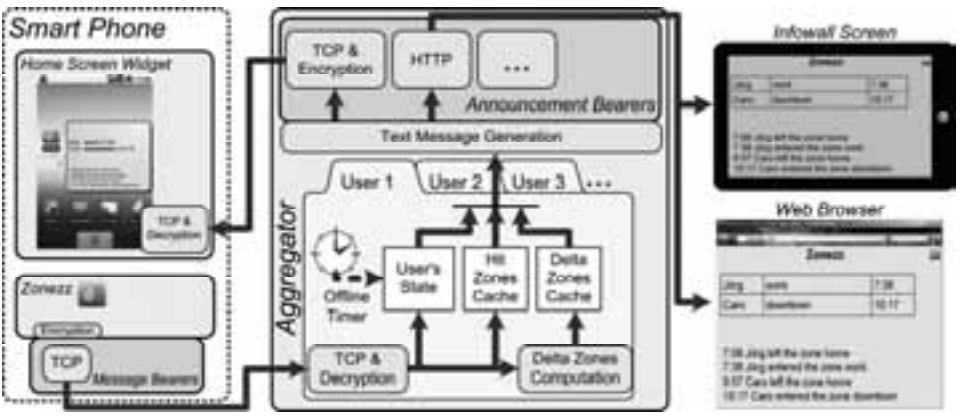


Figure 3: Distributing locations with the Aggregator

Fig. 3 shows the flow of data. After decrypting incoming messages, all location information (state, hit zones and delta zones) are cached. An *Offline Timer* detects if the last received message was too old. As each smart phone periodically sends a message (and not only in case of a change), the Aggregator can detect devices that are offline. Offline devices may have no wireless connection or may be completely switched off. In the current version, the transmission period is 30 minutes and the offline timeout 90 minutes.

Even though the smart phone already computes the delta sets, the Aggregator additionally computes them. This is necessary because messages may be lost due to network problems and the Aggregator caches may run out of sync.

Table 1: User states

No Position	The user is online but the smart phone cannot measure a position, because, e.g. no GPS satellites can be received.
Old Position	The user is online but transmitted zone sets are based on an old position (older than 10 minutes).
No Zone	A new position is available, but the hit zone set is empty.
Zones Available	A new position is available that resides at least in one zone.
Offline	The last update message is too old. The device is considered as offline.

Table 1 shows the possible user states. Note that **No Position** and **No Zone** have a different meaning. Even though **No Position** also leads to an empty hit zone set, it indicates a failure of position measurement, whereas **No Zone** indicates a properly running system.

3.4 Privacy

The approach provides some means to respect the user's privacy:

- Zones are mainly expressed by user-defined strings, thus provide a basic obfuscation. The *meaning* of the same string (e.g. 'job') is different for different users, thus cannot easily mapped to geographic positions.

- As each community sets up its own Aggregator, there is no big organization that can access location information of all users. Thus, users do not have to be afraid of large companies running data mining over all users.

- Encrypted communication between smart phones and Aggregator prevents attacks against the message transfer.

- If a special location is too 'private', the user is free not to model it as a zone. As zones have a certain geometric extent, other user cannot reveal the exact position.

Users can be members of multiple communities and may share different locations with each community. Let Z be the set of user-defined zones. For each Aggregator A, the user defines a set $Z_A \subseteq Z$ that indicates all publishable zones. Consistently, the sets *hit*, *entered* and *left* are mapped to hit_A, $entered_A$ and $left_A$ that are published to an Aggregator A: $hit_A = hit \cap Z_A$, $entered_A = entered \cap Z_A$ and $left_A = left \cap Z_A$.

As an example: a professor is member of the communities 'family' and 'university'. With the latter he wants to run a service for students that always indicates, whether he currently is at the university or not. For his family, he sets $Z_{family} = Z$, but for its students, he only sets $Z_{university} = \{\text{'Office'}\}$. Thus, the aggregator for students is not able to distinguish any zone outside the university. It is important to note that the filter Z_A is executed on the smart phone, not on the Aggregator. Thus, this mechanism is not vulnerable against Aggregator attacks.

4 Presenting Zonezz Locations

We support different tools to present the Aggregator's results (fig. 4).

We first tested to present the locations on our *Infowall service* (fig. 4 top). The *Infowall* screen is a fixed mounted wall screen or a bed table device that displays daily information, e.g. the local weather report, news messages, the message of the day or a calendar sheet with important dates. It can also display instant messages or reminders for family members. The screen periodically changes through all available information pages without user interaction.

Our idea was to integrate the Zonezz locations as a separate page inside the Infowall service (fig. 4 bottom left). The upper half of the page displays the most important (i.e. smallest) hit zone of each user and the time when the user entered this zone. The lower half shows the delta messages as a message log.

Figure 4: Different ways to present the location output

The Infowall decides when an information page is not interesting any more. Zonezz pages are automatically removed when they are not changed for two hours. If a Zonezz page changes, the Infowall service automatically speaks aloud the last delta message using Text-to-Speech function. Thus, a user that is in range of the Infowall screen can *hear* the most important message without reading the text.

Mobile users require an additional tool to view Zonezz messages (fig. 4 right). A user thus can install a home screen widget on the smart phone that presents the information page.

5 Issues and Open Problems

Even though all described components are fully implemented and run as desired, we had to face some problems. Here, we describe the two most important ones:

1. *Android Background Issues*: Android supports periodic background execution and services that are started at boot time. However, the main goal of Android is to run usual apps and *not* background tasks. As a result, the stability of background code is not as desired. To give some examples:

- Network connections are not always established in the background. E.g., if the phone enters a known WLAN, sometimes it does not automatically connect.
- Background tasks are sometimes removed for memory reasons. Even though they are re-instantiated at the next cycle, sometimes irreproducible null pointers occur.
- Some operating system components are not fully error-free. If errors occurred in usual apps, the user simply would restart the app, but for background services this is a critical problem as the user does not even notice a problem.
- Sometimes the positioning systems do not properly work in background. On some devices, e.g., GPS is switched off in background.

2. *Location Fluctuations*: The concept of zones already is robust against positioning errors. A zone may be large enough to cover imprecise measurements at this location. However in rare cases, the current measurement jumps too far. We can distinguish two problems.

- *Enter fluctuation*: the measured location jumps inside another zone that is not the actual zone.
- *Leave fluctuation*: the measured location jumps outside the current zone.

In both cases, trigger events are emitted that may confuse other users. From the cases above, the leave fluctuation is more critical. As an example: if my position wrongly jumps outside my job zone, a delta message 'Jörg leaves the office' will be distributed, thus friends assume that I leave off work.

We could address fluctuation problems with location filters on the smart phone. E.g. a zone change has to be stated by multiple measurements before distributed to others. This however, would delay each trigger event for a certain time.

6 Conclusion

This approach to share locations in small communities is based on *personal symbolic locations*. They can be presented by purely textual mechanisms without the need of a map service. The underlying light-weight service can even run on small devices. In addition to the current location, we identified the change of locations as important information for community members. We integrated the idea into an information screen that could e.g. used by a family to know each other's current location.

References

[AS03] Ashbrook D., Starner T. 2003. Using GPS to Learn Significant Locations and Predict Movement across Multiple Users. Personal and Ubiquitous Computing archive Vol. 7 No. 5, Oct. 2003

[BB+08] Barkhuus L., Brown B., Bell M., Sherwood S., Hall M., Chalmers M. 2008. From awareness to repartee: sharing location within social groups, in Burnett M., Costabile M.F., Catarci T., de Ruyter B., Tan D., Czerwinski M., Lund A. (eds.): Proceeding of the 26th SIGCHI Conference on Human Factors in Computing Systems, ACM, 497–506

[BC+08] Bicocchi N., Castelli G., Mamei M., Rosi A., Zambonelli F. 2008. Supporting location-aware services for mobile users with the whereabouts diary. 1st intern. conf. on Mobile Wireless Middleware, Operating Systems, and Applications

[DA00] Dey A., Abowd G. 2000. CybreMinder: A Context-Aware System for Supporting Reminders, Proc. of the Handheld and Ubiquitous Computing 2000, Bristol, UK, Springer LNCS 1927, 172-186

[HH+10] Häussermann K., Hubig C., Levi P., Leymann F., Simoneit O., Wieland M., Zweigle O. 2010. Understanding and designing situation-aware mobile and ubiquitous computing systems. Proc. of intern. Conf. on Mobile, Ubiquitous and Pervasive Computing, March 2010, 329-339

[JH+02] Jiang X., Hong J.I., Landay J.A. 2002. Approximate information flows: socially-based modeling of privacy in ubiquitous computing, in Borriello G., Holmquist L.E. (eds.): UbiComp 2002: Ubiquitous Computing, Springer, Vol. 2498, 176–193

[LW+06] Liu J., Wolfson O., Yin H. 2006. Extracting Semantic Location from Outdoor Positioning Systems. 7th intern. conf. on Mobile Data Management

[RO+05] Raento M., Oulasvirta A., Petit R., Toivonen H. 2005. ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications. IEEE Pervasive Computing, April-June 2005, Vol. 4 No. 2, 51-59

[Ro06] Roth, J. 2006: Detecting Identifiable Areas in Mobile Environments, 21st Annual ACM Symp. on Applied Computing, 23.-27. April 2006, Dijon, (France), ACM Press, 986-991

[Ro08] Roth, J. 2008. A Probabilistic Approach for Context Reasoning. Use In Context, GI Informatik 2008, Munich, Germany, Sept. 12. 2008, Proceedings 134, Vol. 2, 802-807

[Ro11] Roth, J. 2011. Context-aware Apps with the Zonezz Platform, ACM MobiHeld 2011, Proc. of the 3rd ACM SOSP Workshop on Networking, Systems, and Applications on Mobile Handhelds, Cascais (Portugal), 23. Oct. 2011

[SH+06] van Sinderen, M. J., van Halteren, A. T., Wegdam, M., Meeuwissen, H. B., Eertink, E. H. 2006. Supporting context-aware mobile applications: an infrastructure approach. IEEE Communications Magazine, Sept. 2006, Vol. 44, No. 9, 96-104

[Ve09] Verkasalo H. 2009. Contextual patterns in mobile service usage. Personal and Ubiquitous Computing. Vol. 13, No. 5, 331-342

Social Network for Elderly

Simen Kind Gulbrandsen*, Eirik Fikkan*, Emil Grunt*, Kjetil Mehl*,
Safoura Shamsolketabi*, Jaspeet Singh*, Miso Vrucinic*, Bjørn Magnus Mathisen†,
Anders Kofod-Petersen†

*Department of Computer and Information Science,
Norwegian University of Science and Technology (NTNU), Trondheim, Norway
simgul@gmail.com, efikkan@gmail.com, emilg@stud.ntnu.no,
kjetime@stud.ntnu.no, safouras@stud.ntnu.no, j_singh_89@hotmail.com,
miso88ni@gmail.com
†SINTEF ICT, Trondheim, Norway
BjornMagnus.Mathisen@sintef.no, akof@sintef.no

Abstract: Elderly's health is closely related to their level of social activity. Maintaining an active social life contributes to peoples' ability to live at home longer and enjoy an active and healthy life. Yet, as age progresses maintaining and expanding a social network can be challenging. The work presented here demonstrates a design and implementation of a social network system for elderly, including a recommender system, which will recommend relevant cultural and social events, and friends and acquaintances to enjoy these events with.

1 Introduction

The expected development in demographics in the western world, "... can be characterized as a gradual shift from a society with quantitatively dominant younger cohorts to a society in which the elderly form a solid majority" [Mue07, p. 5]. The current projected scenario estimates that the median age of the population in the European Union will have risen to 48 years, up from 38.5 in 2005. The ratio between the older and the younger population is expected to change from 100 people in the working age range 15–65 to 25 people of age 65 and up in 2005 towards 100 younger people to 51 older in 2050.

Gaymu et al. examine issues of care needs of dependent older adults for nine European countries based on projections for future development until 2030 [GEB07]. They expect that the proportion in the ageing population requiring professional care will grow slower than those with at least the potential of being supported by family. These trends will make the need for children to care for their parents more pronounced. Consolvo et al. classify caretakers based on the impact that carers allow care giving to have on their own lives: the drastic life changer will often sacrifice career or hobbies, the significant contributor will see a profound impact, but with his own life still in focus, and the peripherally involved will provide sporadic care [CRS04].

The field ambient assisted living attempts to counter this fast growing need for care personnel by augmenting the living environment with computerised artefact enabling elderly to live healthy and good life at home. Ambient assisted living applications range from smart home scenarios (see e.g. [SML+09]) to mobile wellness applications (see e.g. [TH07]). Recently, there has been a growing interest in stimulating social connections.

Living alone and isolated is a common problem among the elderly part of the population, in particular in the western world [DW96, RP03]. Tackling this isolation and loneliness among elderly is increasingly being recognised as an important aspect of improving elderly's wellbeing and health. Loneliness is often associated with being disconnected from society, lacking social relations, such as "...nobody asking for me and nobody to ask for...", and not being appreciated [HK10].

Preventing social isolation and maintaining elderly's health and functional capabilities are critical premises for realising the ageing-at-home vision. While social media can play a key role in helping people feel connected to the outside world, conventional solutions are typically not designed to cater for the elderly population. Solutions such as Facebook or Google+ are designed without the specific focus that is required when working with elderly. Thus, only certain types of individuals might benefit from these types of services [MFM08].

The work presented here describes the development of a personalised mobile social recommender system to help promote active living among elderly, both in a physical and social sense. The Social Network for Elderly (SNE) system focuses on developing a system specifically for the elderly population. SNE will inform the elderly about various events and engage them in joining social activities through smart phones. The system was designed and implemented as a student project as part of a project course for 4^{th} year engineering students in the cooperation with the ongoing AAL project *co-living*[1].

The rest of the paper is organised as follows: Section 2 describes the design and implementation of SNE along with some of the design considerations; Section 3 details the functional testing, the limited acceptance tests carried out and the preliminary results obtained; Finally, Section 4 ends the paper with a summary, a brief overview of ongoing testing and and outlook on future work.

2 System Design and Implementation

The SNE system is based on the Service Oriented Architecture (SOA) principle. The architecture consists of three main parts: the content provider, in this case the municipality; the end-users, who are primarily using an Android client; and the SNE server (see Figure: 1). The SNE server consists of the *content data base*, the *user data base*, the *user model*, the *social graph* and at the core the *recommender*.

The content provider's interface is a straight forward web-interface that allows users to add events along with classification information, such as the type of event, when and where,

[1] http://www.project-coliving.eu/

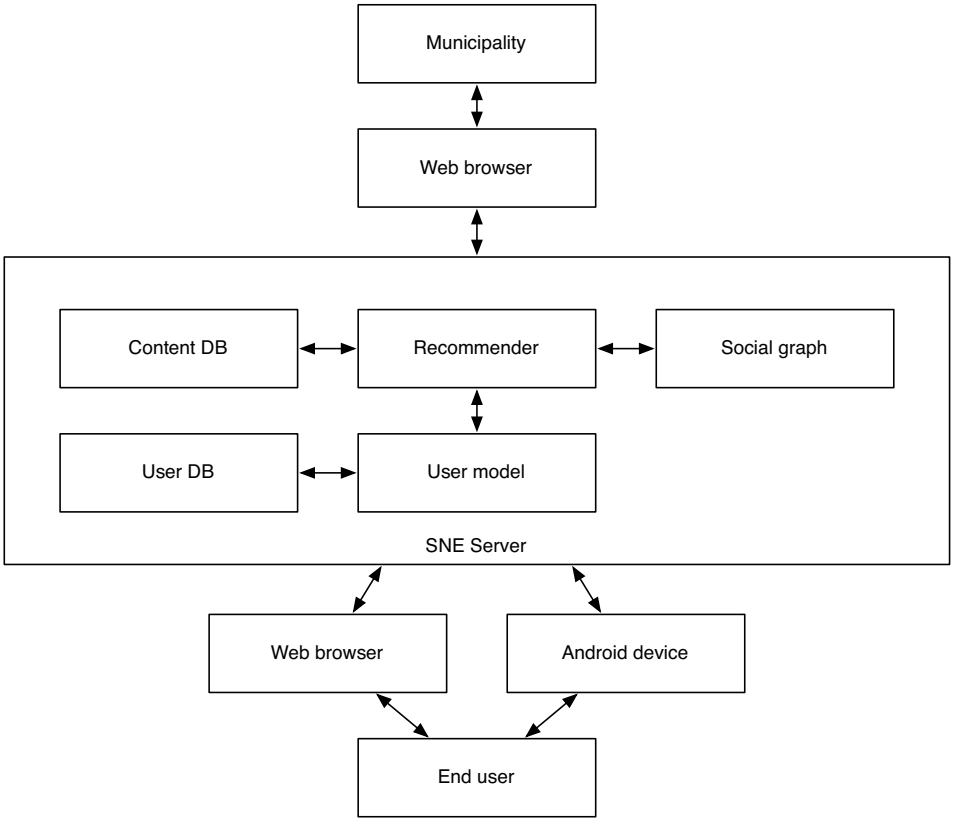


Figure 1: SNE architecture

and price. The content and the user data bases are persistent storages for content and user information, respectively. These parts are more or less run-of-the-mill implementations and will not be detailed further. The remaining parts of this section will describe the user and event modelling, how this is used in the recommender; the social graph and its affect on the system and the end-user application and interface.

2.1 User and Event Modelling

When choosing a specific way of modelling users for adaptive systems some considerations as to the nature of the users and system in question are required. The three most important aspects to consider are: i) if users are homogenous or heterogeneous; ii) whether users are permanent or not; and iii) if their interests are persistent [Lie95].

When working in domains where users are a homogeneous group it is likely that canonical

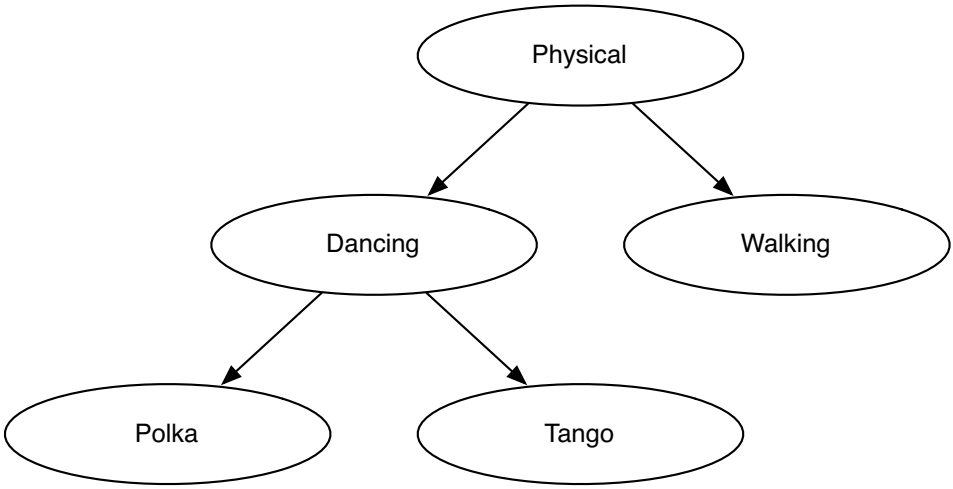


Figure 2: Example of tag taxonomy

user models are to be preferred. Whilst in domains consisting of heterogeneous users special user models would be the preferred option. In domains where user are permanent the system has the option to learn specific user models over time. Whereas, in domains with non-permanent users, canonical models are most useful.

In general, systems with homogeneous and non-permanent users will use very limited and standardised user models. Compared to systems with heterogeneous and permanent users, where very personalised user models will be used. Systems like SNE will typically contain permanent and heterogeneous users; one can also assume that users' interests are persistent. Thus, the option of learning user models exist. However, boot strapping recommender systems with new items and/or new users is a well known problem even in these types of domains. Not having sufficient information is know as the *cold start* problem.

The cold start problem can, to a certain degree, be mitigated by employing user models. Yet, building these models require sufficient knowledge about the specific users. Acquiring this knowledge suffers from the *knowledge bottleneck* problem. That is, it is time consuming (for the user) and not necessarily easily accessible.

The work presented here uses stereotype modelling, in the tradition of Rich [Ric83], to mitigate both the cold start and the knowledge bottleneck problem. Using stereotypes is a quick and efficient way of building user models. A stereotype contains characteristics about a user that fits particular stereotypes. The main advantages of using stereotypes are that they are easy to build and quick to use (mitigating the bottleneck problem).

Stereotypes contain information about stereotypical characteristics, which are known as facets. A facet will traditionally be represented by a value ranging from -5 to 5. Further, each facet has a certainty assigned to it, ranging from 0 to 1000. The rating gives an estimate on how certain the system is about the value assigned. Thus, a high facet value

tells us that the users is very interested in the particular facets, while a high certainty rating tells us that we are very certain of the rating.

Stereotypes are organised as a directed acyclic graph (DAG), where the root node, *anyperson*, contains all the facets with average values, and more descriptive values for the facets are given for the more specialist stereotypes. When a user first registers to SNE, a series of questions are asked that will match the user to a set of stereotypes. Whenever possible the most specific stereotype will be chosen. The sum of matched stereotypes is know as the *user synopsis*. This model is fed into the recommender whenever a recommendation is required, and based on the result of the recommendation the user synopsis can be adapted over time to the user's own idiosyncrasies.

Events are tagged with facets that correspond to the ones used in the user model. These tags are also ordered in a DAG, or more specifically a taxonomy. Figure 2 gives an example of the facets related to physical activity. As depicted, *dancing* and *walking* are both types of physical activity; and *Polka* and *Tango* are both types of dances.

The specific user's model describes, by giving a value to, the relationship between the facets ascribed to the user and the tags available to the system. Using these values the recommender is capable of estimate the fit between a specific user and a specific event.

2.2 Social Graph

The main purpose of the SNE system is not to be yet another event recommender, rather to recommend not only events that are to the user's liking but also people to attend these events with. Thereby (hopefully) increase sociability of the elderly users. The social graph follows much the same approach as the one by e.g. Facebook[2].

The users are organised in the graph (see Figure 3 for an example) and when recommending events to a user the system can traverse the graph in several dimensions. For an example, finding other users in user's neighbourhood, in the graph or physically, who has similar preferences and suggesting that these two participates in an event together. Alternatively, suggesting users with similar preferences, but outside the user's graph neighbourhood to expand the user's social network.

2.3 Recommender

Recommender systems are designed to help users cope with vast amounts of information, and they do so by presenting only a certain subset of items that is believed to be relevant for the user. Traditionally, these systems recommend items like books (`amazon.com`) or films (`movielens.com`), but may just as well recommend events relevant to users.

Recommender systems are usually grouped into two categories: *Content-based* systems and *collaborative filtering*. Content-based systems make recommendations based on a

[2]It could actually be preferable to use Facebook as the supplier of the graph in future incarnations

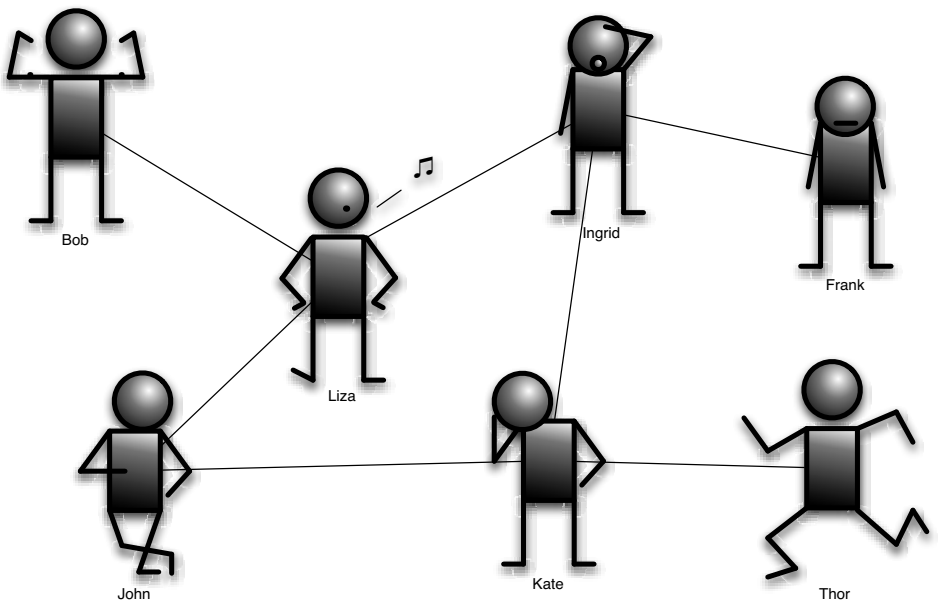


Figure 3: Example of social graph

user preference model that combines the user's ratings with, e.g., content information and structured descriptions of events. On the other hand, collaborative filtering uses the "ratings" of like-minded users to make recommendations for the user in question. Ratings can be explicit (e.g., a number of stars given to an event), or implicit (e.g., the amount of time attended an event that was recommended).

Over the last decade recommender systems based on collaborative filtering have enjoyed a great deal of interest. Collaborative filtering systems are often characterised as either being *model-based* or *memory-based* [BHK98], although hybrid systems have also been developed [PHLG00]. Roughly speaking, memory-based algorithms use the whole database of user ratings and rely on a distance function to measure user similarity. On the other hand, model-based algorithms learn a model for user preferences, which is subsequently used to predict a user's rating for a particular item that he or she has not seen before.

The work presented here approaches the idea of recommending events through a content and model-based approach. Currently the recommender calculates the distance between a user's preferences and the corresponding tags. The collaborative filtering algorithm starts by calculating the ÓPrefValueÓ for each tag, and then traverses the graph from the bottom, since the bottom-most tags are estimated to describe the event best. The ÓPrefValueÓ of the bottom-most tags are multiplied by a factor of 1.0, then their parents 0.5, and grandparents 0.25. Then all these values are added together for a combined preference value, if there are multiple paths of tags in the graph, then the combined preference value is the mean of these paths.

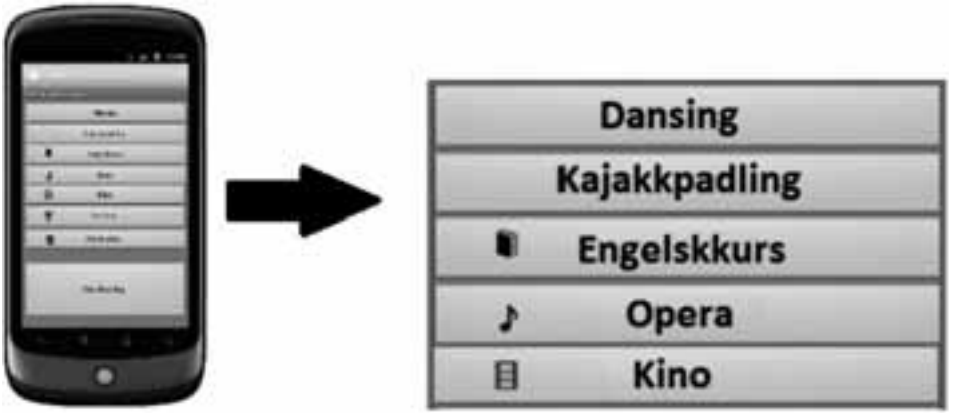


Figure 4: Initial user interface

2.4 End-user Interface and Universal Design

When designing applications for elderly or people with disabilities, the rules of universal design must be adhered to. Universal design is a set of concepts and ideas to use when designing (universally) accessible solutions. The term *universal design* was coined by Ronald L. Mace[3]. Selwyn Glodsmith pioneered the concept of free access for disabled people [Gol97, Gol00].

In the context of computer systems, universal design is used to ensure that web-pages are accessible and usable for all people, regardless of life situation, abilities or disabilities, age and education. Disabilities might include impaired vision, light sensitivity, physically being unable to handle equipment of certain kinds or learning difficulties. Since June 2011 all new IT solutions for the public in Norway has to comply with universal design standards. The most important reference for this work in the Web Content Accessibility Guidelines (WCAG) from W3C [CCRC08].

The implementation of SNE conforms with the main principles of WCAG:

Perceivable Information and user interface components must be presentable to users in ways they can perceive.

Operable Information and the operations of the user interface must be understandable.

Robust The content must be robust enough to be interpreted reliably by a wide variety of user agents and assistance tools.

The user interface on the Android device is designed and implemented following the universal design philosophy. This means that there are large buttons, large font sizes and good

[3] `http://www.udinstitute.org/`

Figure 5: Android interface for specific event

colour contrast to improve the usability for the end users. Further, the system also has a clean and simple layout, which the elderly users are able to handle. Figure 4 shows the initial user interface on an Android device. The buttons lists five categories of events that the user can choose from.

3 Testing and Results

Testing of the SNE system has been done, and indeed is being done, continuously. Testing can be divided into two main parts: *verification*, the code actually works the way it is supposed to do; and *validation*, the program actually fulfils the requirements. The SNE system was designed, implemented and tested during the autumn of 2011. The process was carried out through nine sprints following the Scrum approach.

The verification tests covered 20 different test cases relating to the components for the content providers and end-users, as well as functionality testing of the databases. With respect to the validation testing, the user interface and functionality was tested by two employes at the municipality and two elderly subject.

For the municipality, the test persons were satisfied with the system and it lived up to the requirements put on it. The system was fairly simple and easy to use, something that they liked. They suggested some additional features like the ability to add groups of people with companion and have a printable version of the list of participants.

The end-users appeared to have a bit more technical knowledge that what could be ex-

pected on average. One of the users were quite happy with the system, whereas the other one was sceptical; he thought that elderly would rather call their friends than having the system do it for them. In general, the idea of social graphs and not having total control of friends appeared to be a bit intimidating. The result of the graphical user interface testing was that some of the layout should be changed. It was mostly a problem with the buttons name and layout, that made some choices not seem intuitive (Figure 5 shows the user interface as it looks after the initial testing).

4 Summary and further work

This is work in progress. So far the SNE system has been implemented in a working prototype. Functional verification of the core functionality has been carried out. Initial validation has also been carried out on a limited number of persons. Parts of the SNE system has been implemented in the Co-living system.

There are currently some limitations compared to the original envisioned functionality, The social graph has rather a limited functionality. Currently the recommender does not offer social recommendations. This is currently under implementation.

As part of the ongoing Co-living project, long term and extensive test and validation is currently being carried out. We are currently recruiting and training end-users in the municipality of Trondheim.

Acknowledgements

Parts of this work has been carried out in context of the co-living project, which is supported by the EU commission and the Research Council of Norway through the Ambient Assisted Living program. Parts of this work has been carried out as part of the "customer driven project" course at the Department of Computer and Information Science, the Norwegian University of Science and Technology (NTNU). The authors would like to thank the organisers of this course.

References

[BHK98] John S. Breese, David Heckerman, and Carl Kadie. Empirical Analysis of Predicitive Algorithms for Collaborate Filtering. Technical Report MSR-TR-98-12, Microsoft Research, 1998.

[CCRC08] Ben Caldwell, Michael Cooper, Loretta Guarino Reid, and Gregg Canderheiden, editors. *Web Content Accessibility Guidelines (WCAG) 2.0*. W3C, 2008. http://www.w3.org/TR/WCAG/ [last accessed 2012-03-01].

[CRS04] Sunny Consolvo, Peter Roessler, and Brett E. Shelton. The CareNet Display: Lessons Learned from an In Home Evaluation of an Ambient Display. In *Ubicomp*, pages 1–17, 2004.

[DW96] J. M. Donaldsom and R. Watson. Loneliness in elderly people: an important area for nursing research. *Journal of Advanced Nursing*, 25(5):952–959, November 1996.

[GEB07] Joëlle Gaymu, Peter Ekamper, and Gijs Beets. Who will be caring for Europe's dependent elders in 2030? *Population (english edition)*, 62(4):675–706, 2007.

[Gol97] Selwyn Goldsmith. *Designing for the Disabled: The New Paradigm*. Architectural Press, 1997.

[Gol00] Selwyn Goldsmith. *Universal Design*. Architectural Press, 2000.

[HK10] Solveig Hauge and Marit Kirkevold. Older Norwegians' understanding of loneliness. *International Journal Qualitative Studies on Health and Well-being*, 5(1), February 2010.

[Lie95] Henry Lieberman. Letizia: An Agent That Assists Web Browsing. pages 924–929, Montreal, Quebec, Canada, 1995. kaufmann.

[MFM08] David Mellor, Lucy Firth, and Kathleen Moore. Can the Internet improve the well-being of the elderly? *Aging International*, 3(1):25–42, March 2008.

[Mue07] Rainer Muenz. *Aging and Demographic Change in European Societies: Main Trends and Alternative Policy Options*. Number 0703 in SP Discussion Paper. Social Protection Advisory Service– The World Bank, Washington, DC, USA, 2007.

[PHLG00] David M. Pennock, Eric Horvitz, Steve Lawrence, and C. Lee Giles. Collaborative Filtering by Personality Diagnosis: A Hybrid Memory- and Model-Based Approach. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, pages 473–480. Kaufmann, 2000.

[Ric83] Elaine Rich. Users are Individuals: Individualizing User Models. *International Journal of Man-Machine Stuides*, 18:199–214, 1983.

[RP03] Pirkko Routasalo and Kaisu H. Pitkala. Loneliness among older people. *Reviews in Clinical Gerontology*, 13(4):303–311, August 2003.

[SML+09] Andrew Sixsmith, Sonja Meuller, Felicitas Lull, Michael Klein, Ilse Bierhoff, Sarah Delaney, and Robert Savage. SOPRANO âĂŞ An Ambient Assisted Living System for Supporting Older People at Home. In Mounir Mokhtari, Ismail Khalil, JÃľrÃľmy Bauchet, Daqing Zhang, and Chris Nugent, editors, *Ambient Assistive Health and Wellness Management in the Heart of the City*, volume 5597 of *Lecture Notes in Computer Science*, pages 233–236. Springer Berlin / Heidelberg, 2009.

[TH07] Barnabás Takács and Dávid Hanák. A Mobile System for Ambient Assisted Living with Ambient Facial Interfaces. *IADIS International Journal on Computer Science and Information Systems*, 2(2):33–50, 2007.

Co-Living social Community for Elderly

Bjørn Magnus Mathisen[†] Anders Kofod-Petersen[†]
Idoia Olalde[‡]

[†]SINTEF ICT, Trondheim, Norway
bjornmagnus.mathisen@sintef.no, akof@sintef.no
[‡]Andago, Bilbao, Spain
idoia.olalde@andago.com

Abstract: With a rapidly increasing elderly population, keeping elderly healthy is more important than ever. Physical and social activity is closely related to people's well-being. This paper describes some of the challenges faced when activating elderly people in an organised fashion. We propose a Virtual Collaborative Social Living Community for Elderly; utilising friends, family and care-professionals connected to the elderly in achieving this goal. Challenges related to this work in progress is then presented. The paper presents experiences made in an ongoing development and deployment of a socio-technical system.

1 Introduction

The western world is scrambling to prepare for a rapidly increasing average age. Without changing the way we care for our ageing population, more and more of our workforce will need to work within the caregiver profession. This will have a dramatic effect on the output of the western economies. As the western governments realise this they seek technological and social solutions to this situation.

Researches show that people that have larger social networks (and social supports) and are better integrated into the social fibre of their communities have longer life expectancy. [AA07]

In this paper we will present some of the challenges that are present themselves when developing systems designed to activate elderly. The work presented in this paper is part of the Co-LIVING EU project[1]. The main objective of this project is as follows:

The development of an ICT-based Virtual Collaborative Social Living Community for Elderly (Co-LIVING) people, aiming to stimulate and prolong their independent and active living in an outward environment through an advancement in elderly people social interaction, contributing thus positively to their wellbeing.

The main goal of the Co-LIVING project is to improve the health and quality of life of the elderly using the system, and as a consequence, lengthen the amount of time they can

[1] `http://www.project-coliving.eu/`

live in their own home. The way we want to achieve this goal is by activating the user and creating a social network around the user that can support him/her in the daily tasks.

The following lists the subgoals of the Co-LIVING project:

- In Co-LIVING, innovation emerges not only from the area of new technologies but also from the *development and support of an innovative elderly social practice-oriented community model.*
- Design and develop the Co-LIVING solution by scaling up the successfully developed and piloted IST FP6 mPower[MHF09] open source platform and make it applicable to the field of elderly social interaction context.
- Develop and integrate in the mPower platform the Co-LIVING ICT based services.
- Operation and evaluation of two pilots in Netherlands and Norway.

2 Related Work

Previous EU projects that have worked within this domain; PERSONA[Per], NETCARITY[Net], SOPRANO[Sop], MonAMI[Mon], Companiable[Com], OASIS[Oas] and SMILING[SMI] projects targeting in-house "independent" living.

The Co-LIVING solution uses and scales up the successfully developed IST FP6 mPower open source middleware platform. mPower is an open platform to simplify and speed up the process of developing services for elderly and people with cognitive disabilities. This platform provides reusable, flexible and interoperable service specifications and implementations.

The overall architecture of mPower is an adaptation of the layered model specified in the IBM SOA Reference Architecture[Ars04]. As depicted in Figure 2, the mPower architecture consists of five main layers and three sidecars. Each layer comprises a set of components that conform to specific rules and requirements [Gam95]. This enables us to easier adapt mPower's services to the Co-LIVING project. [FOC+11]

Existing mPower services will be integrated with SoCo-net and the ICT-based services and will be used for the development of new, specific parts of the virtual community network. Co-LIVING will expand mPower to offer services in the social domain.

3 Method

For the development of an innovative elderly social practice oriented community model, two communities of end-users participates in the design and evaluation of Co-LIVING through workshops and pilots. Co-LIVING is not only centred on the design and development around the elderly's needs but it includes the elderly in the process to identify their

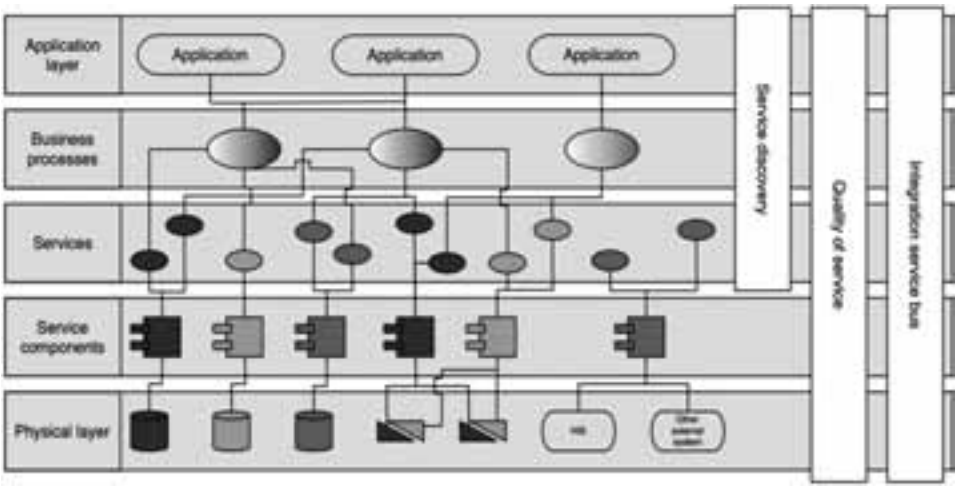


Figure 1: The mPower layer model [FOC+11]

wishes and needs in terms of what factors, relationships and communication issues are meaningful and generate the greatest impact on their social life.

The two end-users communities are:

ORBIS Hoogstaete

Hoogstaete is a unique integrated living village concept in Sittard, the Netherlands. The Orbis medical and healthcare group has developed an integrated neighbourhood where sheltered housing, care, wellness, leisure and education services are provided in an integrated fashion. The complex, around the Highveld Park, aims to encourage the cooperation within the district and integrate young and old communities together to stimulate an active live of the latter. Caregivers and elderly people who live in these facilities (with different ICT skills levels) are involved in the design and testing of Co-LIVING in order to create a system that is fully adapted to the real needs of an elderly community.

Trondheim Kommune

On the other hand, Trondheim represents the physical neighbourhood trial. Trondheim is the third largest Norwegian municipality with a population of 176,348 inhabitants. A wide range of services are offered by the municipality in the health sector. One of these initiatives is the Service Centre for Seniors established in 2003 to ”enable seniors to improve/maintain their functionality level (social and physical) so that they can live an independent life as long as possible”. The Services Centre for Seniors is used for the Norway pilot. Elderly (with different ICT skills levels) and workers of the health system are part of the testbeds.

3.1 Trial Structure

The formal interaction between the system development and the trial subjects is structured into several (at least three) development cycles and their subsequent trials. Before the first development cycle, the project conducted a pre-trial that identified the needs and wishes of the users that had signed up for the trials.

A first version of the Co-LIVING system has already been tested with a sample of real end-users involving elderly and caregivers. The feedback from these trials will be used for the redesign and improvement of the Co-LIVING architecture and services in the second phase of the project.

4 Results

Based on the the input we received in the pre-trials (see section 3.1.) the system is designed to activate the user in two ways:

- By reminding the user about interesting (determined by previous history and the social network) events and asking the user if he or she has been active today.
- By drawing more directly on the social network; if your friend is planning to attend an event (or taking a walk) the system will facilitate your partaking in the event.

There are several challenges related to setting up a Virtual Collaborative Social Living Community for Elderly.

- Motivating the elderly to be active.
- Getting the elderly to use the system.
- Create usable and easy-to-use interface adapted to the needs of the elderly.

Mobile wellness applications can motivate people to exercise more because of their ubiquitous presence and suitable technological possibilities. Wellness applications utilise several factors that can increase physical activity levels of users, e.g., data tracking, social sharing and playfulness. [AH10]

The Co-LIVING solution is a collaborative and social platform for the elderly that fosters the community life and encourages them to interact socially and participate in outward activities. Co-LIVING takes into account that the networks of an aged person consist of people of different ages (young and old) and roles (relatives, friends, neighbours, care professionals, etc.).

Co-LIVING is based on an innovative Social Community network (SoCo-net), an elderly centric web based network that manages and constitutes social care teams around the elderly in order to provide them with personalised care considering that these groups are:

Figure 2: A screenshot of the Co-LIVING system, displaying a list of possible activities to the user

- Virtual, they assist and provide care to the elderly without being together physically.
- Dynamic, the teams are dynamically adjusted, due to availability, preferences and status of the users.
- Collaborative, as combining different people together as needed and combine their knowledge to provide effective care.

Detecting changes in the behaviour of the elderly is a key factor to anticipate their physical and psychological deterioration and take proactive actions to prolong their active life. SoCo-net is provided with adaptive user profiling techniques and intelligent adaptive interfaces to detect these changes and update user's profile consequently. These adaptive mechanisms make use of user's context and user's behaviour data to identify deviations in elderly's daily activities as they age (Behaviour Analysis component). The Behaviour Analysis component will adapt, based on historical data, the elderly's profile (preferences, capabilities) to provide appropriate services for the current user's status. This compo-

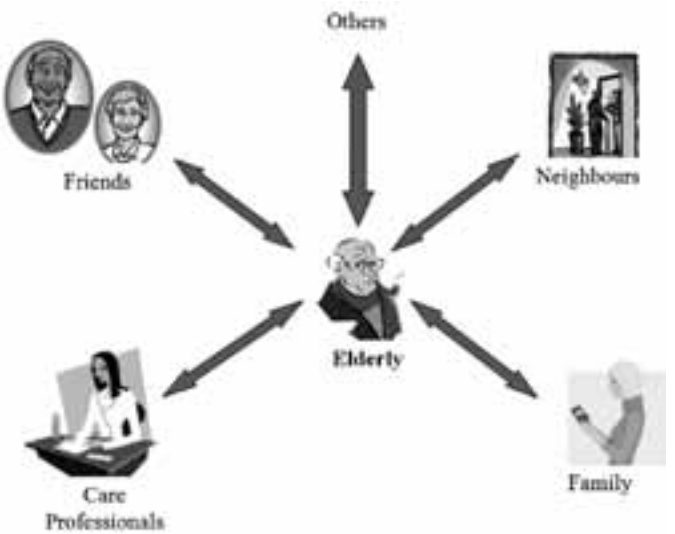


Figure 3: Virtual Care Teams build around the elderly

nent will suggest and motivate elderly to participate and maintain an active social life if a deterioration of their social participation is detected.

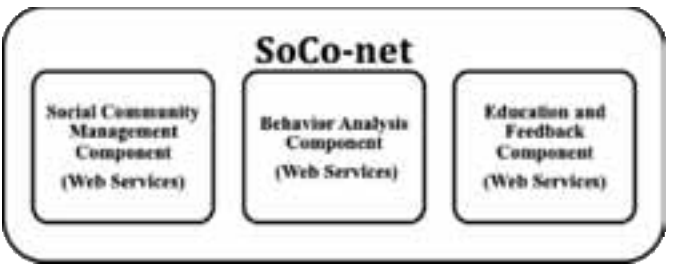


Figure 4: SoCo-net can be thought of as three different components; Social Community Management Component, Behaviour Analysis Component and Education and Feeeback Component.

SoCo-net will also detect behaviour patterns regarding the use of the Co-LIVING system and it will try to stimulate elderly to use it in order to maintain an active social life. SoCo-net is provided with intelligent decision making techniques on user's context and historical data and intelligent explanation generation systems to train and maintain the interest of the users (Education and Feedback component). The Education and Feedback component will encourage the elderly to make use of the Co-LIVING system if a decrease of the activity is detected and assist them in the use of services, adapting the interfaces to their current needs.

Co-LIVING aims at the development of different ICT-based services, which will make use of the virtual social care team management and organisation tool provided by SoCo-net,

to stimulate the elderly to maintain an independent and active life for longer. The Co-LIVING ICT-based services address the three main areas of the elderly social interaction context:

- Care & Wellness
  The **physical activity** service investigates how information on the individual's actual physical and psychological status can be optimally combined to define a challenging yet realistic physical activity schedule. Instead of adapting the physical activity schedule only based on the physical status or progress in performance, the person's psychological status is also included. For example when the person gets less motivated, the system may invite the user to select and do other exercises (targeted towards the same physical goal), or to slightly increase (or decrease) the intensity or duration of the exercises.
  The **group leisure activities** service uses SoCo-net to create groups to share activities with the elderly, by taking into consideration the members' preferences and capabilities. The elderly creates an invitation for a specific activity and shares it with the members of his/her Social Community Network.
  The **competence/knowledge/skills exchange** service allows the elderly to register his/her expertise in the system and other members of the Social Community Network can make requests for support.

- Guidance
  Daily tasks assistance provide to the elderly direction indications, explanations on how to perform different tasks, or even instructions on how to call for human assistance by making use of SoCo-net.
  Cognitive failure assistance provides memory help reminders i.e., accessories such as stick, eye-glasses, medication, planned activities or appointments, directions indications to a place, etc. The services are designed not to be a blind guide, providing all the instructions, but rather as assistance with increasing levels of social elderly care provision.

- Mobility Monitoring
  Services for the **early detection of limitations** are based both on wireless sensors providing real time monitoring of mobility and activity of the elderly, like GPS and accelerometer sensors and on physical status information entered through developed questionnaires. All information is analysed and services are developed as regards the provision of care to the elderly by providing direct feedback to him or by informing his care givers. Additionally localisation based monitoring services not only identify the position of the user in real time, but also obtain and correlate information regarding the surrounding environment of the user.
  Services for **Daily activity follow up** enable the elderly to set up his daily schedule with various activities. The time, place and group members that may be involved in each activity are defined. The daily timetable is transmitted to the group member that is responsible for the follow up of the activities. At any time during the day the group member may contact the elderly and enquire variations in the schedule (i.e., delay or absence from a meeting).

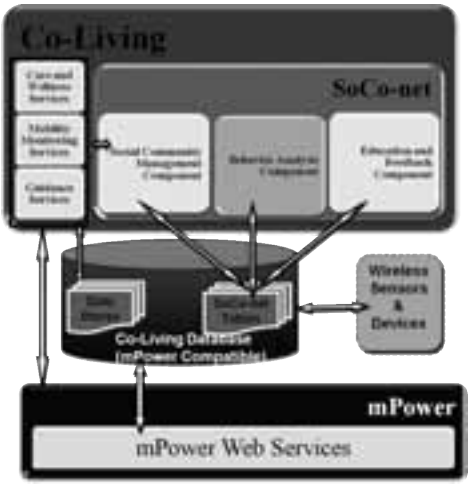


Figure 5: Co-LIVING Architecture

5 Conclusions and Future Work

In section 2 we listed challenges relating to activating elderly. We proposed that several of these challenges could be addressed through an ICT-based Virtual Collaborative Social Living Community (Co-LIVING).

Trials have tested the positive outcome of Co-LIVING in terms of the increase of the socialisation of the elderly and Co-LIVING consortium keeps on working in the development of the different components to reach a complete solution that will foster an active life through the networks with friends, family and care-professionals that older people maintain.

As mentioned in section 3, Co-LIVING is currently in the first of three planned stages of iterative development. The feedback we received during the development and testing in this first phase will be included when entering the next stage.

Acknowledgements

Parts of this work have been carried out in the context of the Co-LIVING project[2], which is supported by the EU commission and the Research Council of Norway through the Ambient Assisted Living program.

[2] http://www.project-coliving.eu/

References

[AA07] C Antonucci and H Akiyama. *Social networks, support and integration*. Elsevier Inc., 2007.

[AH10] Aino Ahtinen and Pertti Huuskonen. Let's All Get Up and Walk to the North Pole : Design and Evaluation of a Mobile Wellness Application. *October*, 2010.

[Ars04] A Arsanjani. Service oriented Modelling and Architecture: How to Identify, Specify, and Realize Services for Your SOA, 2004.

[Com] CompanionAble. http://www.companionable.net/.

[FOC+11] Paulo Freitas (Ed), Idoia Olalde, Christophoros Christophorou, Eleni Christodoulou, Christiana Tsiourti, Jorge Dias, Paul Koster, John Bernsen, Anders Kofod-Petersen, Ståle Walderhaug, Dimosthenis Georgiadis, and George Samaras. D2.1 Overall system design. Technical report, Co-living consortium, 2011.

[Gam95] E Gamma. Design patterns: elements of reusable object-oriented software, 1995.

[MHF09] Marius Mikalsen, Sten Hanke, and Thomas Fuxreiter. Interoperability Services in the MPOWER Ambient Assisted Living Platform. *Medical Informatics*, (7465):366–370, 2009.

[Mon] MonaMI. http://www.monami.info/.

[Net] Netcarity. http://www.netcarity.org.

[Oas] Oasis. http://www.oasis-project.eu/.

[Per] Persona. http://www.aal-persona.org/.

[SMI] SMILING. http://www.smilingproject.eu/.

[Sop] Sopranos. http://www.soprano-ip.org/.

SESSION 2

Transport Systems and Architecture

Context-Awareness and Real-Time Information in an Intelligent Smartphone Application

Christoffer Jun Marcussen **Lars Moland Eliassen**
{chrimarc,larsmola}@stud.ntnu.no

Rune Sætre **Björn Gambäck**
{satre,gamback}@idi.ntnu.no

Department of Computer and Information Science
Norwegian University of Science and Technology

Abstract: With the constant increase in smartphone sales, integrated sensors and map navigation have now become available to the average user. This allows for mobile applications to use context-awareness to provide more relevant information. An interesting use-case for such applications is a route information system for buses.

The paper describes an application which interfaces over a mobile phone to BusTUC, a natural language-based reasoning-system for bus routes in Trondheim. By combining user context, BusTUC reasoning and real-time data from the bus service provider, the user-interaction is simplified, compared to a standard information system. We discuss issues on supporting context-awareness and real-time information in this system, comparing it to other available route information systems. Feedback from beta-testers indicates that the application suits the needs of typical bus travellers well.

1 Introduction

Smartphones today are pervasive and personal: they are everywhere, almost always turned on, and customised to each user. Hence smartphones are well suited for context-aware applications [ROPT05], and indeed *context-awareness is at the core of location-aware computing* [HSK]. Location-aware systems utilise information about the user's location, through location-sensing technology such as GPS or Wi-Fi, or by short range transmission technology (e.g., Radio Frequency Identification, RFID). An example of RFID usage in a context-aware application is to place tags in doorways, to track passing people.

Context-awareness can be utilised in several types of applications. For example, the tour guide for the Nidaros Cathedral, implemented in the context-aware *Nidaros Framework* [WSB+05] can use location information to display zones and nearby objects. The domain of route guidance is particularly well suited for context-awareness: the systems can, e.g., monitor the user's behaviour through sensor input, and use this data to provide route suggestions or other information. With the recent progress in smartphone technology, several route guidance applications have become available. Here, we present TABuss, an intelligent application developed to explore new possibilities and to utilise more smartphone

capabilities, within the bus route information domain. To this end, some important design decisions will be discussed, in particular which specific smartphone capabilities are most relevant to the users in this type of setting, how the user interface best should be designed, and whether developing for one particular type of phone using native code is the optimal choice or if cross-platform development through a web application is a better option.

TABuss is based on BusTUC [Amb00], a natural-language query system which became publicly available to the inhabitants of Trondheim in 1998, supporting them in getting information on the time tables of "Trondheim Trafikkselskap", the city's bus company as the time. BusTUC was commercialised by LingIT AS in 2001, and approximately one million queries have been posed to the system every year since then. Now hosted by the current public transportation provider in Trondheim, AtB (`www.atb.no`), BusTUC can be accessed both through the web and a Short Message Service (SMS).

The main topic of the paper is context-awareness as implemented in TABuss. The system also incorporates real-time capabilities by providing information on the actual arrival times of all buses. Queries are sent using a SOAP[1] interface to a server hosted by AtB. The only necessary input parameter is a bus stop's real-time ID which can be retrieved from the same server as a list mapping each bus stop ID to a real-time ID. AtB updates this list from time to time, so an updated list is necessary in order to query the correct real-time data.

The rest of the paper is laid out as follows: Section 2 discusses the concept of context awareness, and its application within a bus route information system. Section 3 describes some other relevant systems. Section 4 moves on to the design decisions taken in the TABuss application, including the tradeoff between native and web-based applications, while Section 5 details the actual TABuss application. Section 6 reports the results of system testing and initial user feedback. Finally, Section 7 further discusses the experiences of the development of TABuss and suggests ways in which the work could be extended.

2 Context Awareness

To describe and implement a context-aware bus route information system, the concepts *context* and *context-awareness* will be used quite specifically: context only uses location information, while context-awareness introduces time and destination as additional factors.

Several researchers have defined context and context-awareness. Pascoe (1998) defined context as "a subset of physical and conceptual states of interest to a particular entity" (e.g., a person), where the importance of the involved states has to be determined [Pas98]. Schilit and Theimer (1994) introduced three factors necessary to define context: location, descriptions of people in the immediate surroundings, and objects (with the changes these objects go through) [ST94]. Ryan *et al.* (1997) added time as a factor, defining context as the user's location, environment, identity, and time. They generalised context by including a number of physical and logical attributes, assumed to affect the user's environment [RPM98]. Dey and Abowd (1999) gave an even more general definition of context.

[1] `http://www.w3.org/TR/soap/`

Rather than enumerating a list of factors needed to be matched, they wrote: "context is any information that can be used to characterise the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and application themselves" [DA99]. This definition simplifies declaring functionalities and theories as context. If any piece of information can describe or help the user at a given time, it can be called context.

Dey and Abowd's definition will be adopted here, partially since it is the most general, stating that a factor or an entity is a part of the context, as long as it concerns the user. It also allows for context to be either implicit or explicit. This means that context information both can be provided by the user and automatically detected by the system. Dey and Abowd further define context-awareness as the system's responses to changed context, without having to determine whether the system should initiate an action automatically or not. For example, the system should not automatically start downloading the real-time data for one or more bus stops, only because the user has changed his/her location. However, in an intelligent application, it is still an advantage to have the option to do so.

Pascoe's definition cannot be used here, because of its lack of modularity. If adding additional factors to determine context, the individual importance of the factors would have to be determined. This is difficult, as situations change (what is rated as an important factor in one setting, might be less important in another). The definition by Schilit and Theimer as well as the one by Ryan *et al.* would also — considering the limited and changing factors — not be adaptable to our purposes.

As mentioned above, the concepts of context and context-awareness are used in specific ways in a bus route information system. Context is used during the general location tracking of the user, which triggers the loading of nearby bus stops. This happens dynamically as the user moves and can be seen on the map while moving. Clickable Bus stop icons are added or removed as the user changes his/her location. This is an example of implicit context: the user does not provide the location information manually (the location is automatically detected and tracked by the system). However, location technology, such as GPS, Wi-Fi or 3G must be enabled on the phone. Context-awareness in TABuss, which uses more factors (e.g., time and day), is described in Section 5.1.

3 Related Work

Several intelligent route information systems include natural language interfaces. The *Let's go* system [RLBE03] uses speech from phone calls as input, and returns route information. The system was developed for "elderly and non-native English speakers", providing information for the city of Pittsburgh. Speech is recognised by comparison and retrieval of the closest match, with the emphasis on creation of a grammar model for spoken language, and on including an overall generality regarding different structuring of sentences with the same meaning. Several challenges with speech processing and route information were identified, with the main challenge being that different users use different phrases, when referring to bus stops or places.

TravelMan was developed for the city of Tampere, Finland [THK+07]. The input consists of locations or addresses, provided to the system as text or speech. The user can also set personal preferences, such as exclusion of specific transportation options, as *TravelMan* covers metro and tram in addition to bus transportation. Guidance functionality for visually impaired users was implemented, to provide what was referred to as "an unbroken trip chain". That is, a successful trip should be complete with full system guidance. An interesting feature in *TravelMan* is the use of context and user location. The real-time guidance relies on location information, which can also be used to infer departure addresses.

Access to real-time bus information has previously been addressed in, e.g., the *MyBus* system [MD], which predicted real-time arrival of buses based on historical data, the bus schedule and a prediction algorithm. The algorithm produced estimates, where traffic and passenger information was used as noise, affecting the original schedule. A mobile version of the system, *Mobile MyBus* [MD01] used WAP communication. *MyBus* gave users real-time information based on input consisting of destination and a route number. Both inputs were provided as digits, since devices at the time did not have input mechanisms similar to a computer keyboard. The functionality can be compared to retrieving a real-time ID for a bus stop in TABuss, before sending a SOAP-request to fetch the real-time data. Today, buses have GPS-trackers on-board, where location information is continuously transmitted to a server. Such real-time data is updated at specific time intervals and is fairly accurate. The real-time service was introduced in Trondheim by AtB early in 2011.

OneBusAway focuses on context awareness in addition to the use of real-time data. *OneBusAway* uses the location of the user to automatically display the closest bus stops on a map, similar to TABuss. Context aware functionality has proven to be useful through user testing, where 93% of the existing users of *OneBusAway* reported that more concise information was provided [FWB10]. What distinguishes TABuss from *OneBusAway* is that the main query functionality in TABuss only needs the user's destination as input. *OneBusAway* needs to know both the departure stop and which route to select in order to reach the planned destination.

There are also several smartphone applications developed for bus transportation in Trondheim, many of which use the BusTUC oracle service. All these applications have been downloaded by several users, indicating that they have attractive functionality. For our project, it was important to investigate what has to be done to move the concept of a bus route information application to a higher level. We also compare the different levels of artificial intelligence in the given apps. Two of the existing solutions are close to what we want to make: *Alf's ByBuss* and *Bartebuss*. They both use BusTUC, maps and the real-time functionality. *Bartebuss*[2] is developed in HTML5 and uses the *BusBuddy*[3] API(Application Programming Interface). *Bartebuss* has the option to store favourites, it can find near-by bus stops or search for specific bus stops, and it can use BusTUC and show maps on the phone. The use of HTML5 makes the map less responsive than in *native applications*. The user interface, on the other hand, is intuitive and easy to navigate, but it might provide too many choices to the user. *Alf's ByBuss*[4] is a native Android application

[2] `http://bartebuss.no/om`
[3] `http://busbuddy.norrs.no/`
[4] `http://bybuss.alfsimen.com`

that also uses the *BusBuddy* API. *Alf's ByBuss* appears more responsive than *Bartebuss* during map navigation, but the user interface is not as polished.

4 Developing Native Applications vs. Web Applications

When developing for mobile platforms, there are several technology decisions to be made. An important choice is whether to develop *native applications* or *web applications*. Native development has been the main choice for platforms such as Android and iOS, as earlier versions of HTML did not provide enough framework possibilities. Lately a new option has emerged, with the release of HTML5. Applications written in HTML5 and JavaScript have now become the new competitors to the *native applications*.

The basic definition of a *web application* is: "having no ties to a specific operating system or device". *Web applications* do not rely on any platform-specific API or SDK(Software Development Kit). The user interface can be designed to resemble *native applications*, but the *web applications* can be deployed on all major platforms. *Web applications* can consist of web code only, or they can be hybrids with both HTML and native code. A *hybrid application* runs the web code in addition to some parts implemented in native code. The native parts can be just simple parts of the user interface, or large amount of back-end code. A user interface written for the web is different from a native one in that the rendering will be done in a web browser instead of by a native graphic component on a specific device.

Native applications are developed using platform-specific SDKs and have (through the underlying operating system) direct access to the device's hardware. The main disadvantage with native development is that the applications are not directly portable to other platforms.

Although *web applications* can perform many of the same functions as *native applications*, the end result is not always as satisfying. *Web applications* (as of today) perform slower than *native applications*, and large background computing is not supported by build systems such as PhoneGap[5] which maps web code to native code. Still, it is debatable how many applications actually need the most optimised performance to function properly.

Hardware access has been a problem for early *web applications*, as the possibilities were limited compared to native code. Today, with libraries that give access to components such as GPS, camera and compass, the limitations are less visible. However, some are still noticeable, e.g., when using maps in iOS. Then, the rendering speed is similar to native code, while on Android the performance is slower.

Both *web applications* and *native applications* have advantages and disadvantages, so the choice depends on the complete context. TABuss is developed as a *native application* (for Android), since *native applications* are ideal for research within mobile development: The end goal for research is seldom mass distribution, but proof-of-concept. This can be realised by having the newest (native) tools available.

[5] `http://www.phonegap.com`

5 The TABuss Implementation

The development goal of TABuss was to make the application as easy to use as possible. TABuss is divided into several *Android activities*, where the top activity defines the main(home) screen. The phone's menu and buttons start *sub-activities*, and the user can choose for whether or not to use the map.

5.1 Context Awareness in TABuss

Context is mainly extracted from the user's location. The application automatically loads the closest bus stops based on the user's location whenever a location change has been detected. Real-time data for these bus stops can be accessed from the map or through a list available in the menu. The closest bus stops also play an important part in the main query functionality, where the user's location determines which bus stops are included as departure stops. TABuss distinguishes itself from other existing solutions by giving the user the option to let the application guess where he/she is going. A simplified version of case-based reasoning is implemented, by logging each query as a case (see [AP94]). The queries are stored locally, in a database, and each case consists of the departing area, the time of day, the day of week and the destination. The departure area is a 500×500 metre square, with a defined area code stored in a separate table. Whenever a new case is created, a new area is created if the origin location is not covered by an existing area.

Queries with similar origin and time are fetched from the database in order to retrieve relevant cases. Similarity is indicated by an overlapping area or by closeness in the time of day for the query (+/- 2 hours). The retrieved cases are rated by the Euclidean distance between the locations, the time-difference and whether they happened on the same day of the week. The best matching destination is presented to the user, based on loose time matching, since exact matches are rare. The stored cases could be extended, for example, to include delayed bus departures, and bus departures from within a time period.

When TABuss suggests a route, the user can respond by validating the result. Positive user feedback currently triggers a query run, while negative feedback has no effect. The level of intelligence is fairly low, but still higher than in approaches based on direct look-ups.

5.2 Language Processing

TABuss has an option to switch between two different natural language modes: the "new" mode which assumes that the user wants to depart from one of the closest located bus stops, and the "standard" mode which allows for user-defined departure stops (and other complete natural language queries about buses), as examplified below, showing queries in the "standard" (1) and "new" (2) modes (where the n represents walking distance, in this case to the bus stop 'Samfundet'). Switching between the two language modes can be done in the home screen menu.

(1) *Når går bussen fra Samfundet til Torvtaket?*
When goes the-bus from Samfundet to Torvtaket?
When does the bus to Torvtaket depart from Samfundet?

(2) *(Samfundet +n, Prinsen +n) til Torvtaket.*
(Samfundet +n, Prinsen +n) to Torvtaket.

TABuss relies on an existing text messaging service when no data network is available.[6] An SMS (text message) query starts with "*rute*" (route), followed by a natural language query. This has been incorporated in two ways. If the "new" mode is chosen from the home screen menu, TABuss uses the closest bus stop to the user's location as the departure stop. If the "standard" mode is chosen, the user has to provide a complete sentence including the place of departure and the destination.

5.3 Real-time Functionality

Real-time data can be accessed from the map by pressing a bus stop icon or through the home screen menu. Both access methods utilise the user's location to retrieve and display the n closest bus stops.

The retrieval of a bus stop's ID is done by comparing the chosen bus stop's location with the locations of each of the n closest bus stops. If matched, the found bus stop ID is used to extract the real-time ID. The real-time ID is then sent via SOAP to the real-time server, which returns the five next bus departures. The user can also search for bus stops that are not among the n closest, by providing a bus stop name as input. This option also lets the user select which direction to retrieve real-time data for (either from or towards the city centre) before the real-time query is sent.

5.4 Answer Display

The TABuss application is best explained through some screenshots. Figure 1a shows the start menu, where text buttons represent shortcuts stored on the SD-card of the device. Figure 1b is the answer screen with route suggestions shown in a list view. The displayed routes with updated departure times are the results of an HTTP query sent to BusTUC. Walking distances to the bus stops are shown within parentheses. "**Overgang**" indicates that the suggested route includes a transfer.

The context-awareness of the TABuss application is shown in a second set of screenshots. Figure 2a shows a map displaying the user's location and the closest bus stops represented by clickable bus stop icons. The map is displayed by selecting an element in the results list. Finally, Figure 2b displays the result of a real-time data query for a specific bus stop. The query is either initiated from the menu or by pressing a bus stop icon on the map.

[6] `https://www.atb.no/spoer-bussorakelet/category228.html`

(a) Start screen

(b) Answer screen

Figure 1: Screenshots from TABuss

(a) Map showing the closest bus stops and the user's current location

(b) Real-time bus arrival information for a specific bus stop

Figure 2: Context-awareness in TABuss

6 Evaluation

A small-scale user test was performed to get feedback on the TABuss' application. An extensive user test was not conducted because of time limitations and delays in the public release of the application. Hence all the test subjects were Trondheim inhabitants and experienced bus travellers.

The general user opinion indicates that the application is easy to use. It is clear that the users appreciate the user interface. Positive feedback has been received on both the colour combinations and the layout. Most users prefer the application functionalities detached from the map, and feedback suggest that the map should only be optional. Users find the query functionality useful. The main functionality with the "new" BusTUC mode is seen as interesting. Users requested the possibility to use the standard BusTUC mode for queries not involving the closest located bus stops, something which was not an option in the early development stage. Another suggestion was to include a "settings"-screen, allowing the user to set preferences (such as the number of bus stops to use in queries). The real-time data functionality for the closest bus stops was easy to access and use. This applies to the real-time functionality on the home screen, as this required less navigation than through the map.

It is difficult to draw concise conclusions after this small user test, but the feedback received from the target users is valuable. The suggestions and error reports give an indication that TABuss suits the needs of bus travellers.

7 Discussion and Future Work

In Section 3, the applications *Bartebuss* and *Alf's ByBuss* were compared to TABuss. TABuss' functionalities are more focused on user location and context-awareness than *Bartebuss*. The level of "intelligence" is what separates TABuss from *Bartebuss*, and also from other apps tested in Trondheim. In order for a bus route information application to "be intelligent", the natural input source has to be context data. Still, whether TABuss can be classified as "better" than *Bartebuss* is an open question: *Bartebuss* has been developed over a longer period of time, and been through more extensive user testing. However, TABuss represents a more complete approach, with a good market potential, since no other application has the exact same functionalities.

Section 4 discussed the advantages and disadvantages of native vs. web development, and stated that native development was the preferred choice for a research prototype. In retrospect, we are satisfied with the choice of technology. Compared to *Bartebuss*, a notable technology difference lies in the storage functionalities. For Bartebuss to work cross-platform, and also through a regular browser, "web storage" through "local storage" is used. The size limit of local storage depends on which browser is used, but it cannot be larger than 10 megabytes (Internet Explorer). TABuss uses the devices' external storage, where the size limit depends on the size of the mounted SD-card, which can normally store gigabytes of information. It is possible that future releases of TABuss will need more stor-

age space than 10 megabytes. The storage limitation of also affects the iOS version of Bartebuss, where the internal storage optimally is used instead (since no external storage is available).

The problems with lagging maps in web applications deployed on the Android platform are avoided in TABuss, where the map is much more responsive. Native development also allows for pinch zooming, which is an important feature when navigating maps.

Although web applications can be deployed on multiple platforms, native applications provide the best user experience for Android and the bus route domain. It is preferable to develop a competitive application for a specific platform, rather than to deploy a "working" solution to multiple platforms (given today's web application performance on Android). Future SDK updates will benefit web application development and improve the browser rendering. One problem is that older devices will not receive these updates (and it will take time for newer devices to get them). The release to newer devices usually happens only after the different manufacturers have adapted their own distributions. Developers will then have a dilemma regarding which SDK versions to target (or which users to exclude).

TABuss uses location data as context input. An extension is to use more sensors than only the location sensor, such as in *ContextPhone* [ROPT05]. *ContextPhone* uses four sensors: location, user interaction, communication behaviour and physical environment. This means that besides from location information, *ContextPhone* monitors what actions the user performs, calls and SMSs, and surrounding devices. Such sensor information could be used to introduce context awareness to the TABuss user interface. The user interface could track the user's actions through sensors, register trends and then adjust visibility and availability accordingly. The tracking of the user's trends could also be used to create better route suggestions. People of different ages have different levels of mobility, and different walking speeds. This has been addressed in *UbiBus* which considers different people's and vehicle's mobility, and other factors that can affect which bus departures that are most "attractive" [VCS11]. An interesting idea is for AtB to contribute to such functionalities in order to improve route suggestions. Buses have cameras installed, which, for example, could be used to monitor how crowded a bus is.

Acknowledgments

Thanks to LingIT AS and AtB AS for allowing the use of the BusTUC and the real-time bus information services, respectively, and to Magnus Raaum for providing the source code of the original Android application. We would also like to thank our beta-testers: Jirka Konietzny, Trond Bøe Engell, Jostein Klakegg, Ola Hast, Marita Gjerde, Håvard Axelsen and Morten Fornes. Runar Andersstuen and Trond Bøe Engell provided valuable ideas and feedback during the project. Special thanks to the late Tore Amble for developing the BusTUC system and for initiating this work.

The work of the third author was partly funded by the Research Council of Norway, through the project UbiCompForAll (`www.sintef.no/Projectweb/UbiCompForAll`) under contract nr. 187865/S10.

References

[Amb00] Tore Amble. BusTUC: A Natural Language Bus Route Oracle. In *6th Conference on Applied Natural Language Processing*, pages 1–6, Seattle, Washington, 2000. ACL.

[AP94] Agnar Aamodt and Enric Plaza. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications*, 7(1):39–59, 1994.

[DA99] Anind K. Dey and Gregory D. Abowd. Towards a better understanding of context and context-awareness. In *1st International Symposium on Handheld and Ubiquitous Computing*, 1999.

[FWB10] Brian Ferris, Kari Watkins, and Alan Borning. OneBusAway: Location-Aware Tools for Improving Public Transit Usability. *IEEE Pervasive Computing*, 9(1):13–19, 2010.

[HSK] Mike Hazas, James Scott, and John Krumm. Location-Aware Computing Comes of Age. *IEEE Computer*, 37(2).

[MD] Stuart D. Maclean and Daniel J. Dailey. In *7th Annual World Congress on Intelligent Transport Systems*.

[MD01] Stuart D. Maclean and Daniel J. Dailey. Real-time Bus Information on Mobile Devices. In *Intelligent Transportation Systems. Proceedings*, pages 988–993, 2001.

[Pas98] Jason Pascoe. Adding Generic Contextual Capabilities to Wearable Computers. In *2nd IEEE International Symposium on Wearable Computers*, 1998.

[RLBE03] Antoine Raux, Brian Langner, Alan W. Black, and Maxine Eskenazi. LET'S GO: Improving Spoken Dialog Systems for the Elderly and Non-natives. In *8th European Conference on Speech Communication and Technology*, 2003.

[ROPT05] Mika Raento, Antti Oulasvirta, Renaud Petit, and Hannu Toivonen. ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications. *IEEE Pervasive Computing*, 4:51–59, 2005.

[RPM98] Nick S. Ryan, Jason Pascoe, and David R. Morse. Enhanced Reality Fieldwork: the Context-aware Archaeological Assistant. In V. Gaffney, M. van Leusen, and S. Exxon, editors, *Computer Applications in Archaeology 1997*. 1998.

[ST94] Bill N. Schilit and Marvin M. Theimer. Disseminating Active Map Information to Mobile Hosts. *IEEE Network*, 1994.

[THK+07] Markku Turunen, Jaakko Hakulinen, Anssi Kainulainen, Aleksi Melto, and Topi Hurtig. Design of a Rich Multimodal Interface for Mobile Spoken Route Guidance. In *10th European Conference on Speech Communication and Technology*, pages 2193–2196, Antwerp, Belgium, 2007. ISCA.

[VCS11] Vaninha Vieira, Luiz Rodrigo Caldas, and Ana Carolina Salgado. Towards an Ubiquitous and Context Sensitive Public Transportation System. In *4th International Conference on Ubi-Media Computing*, pages 174–179, 2011.

[WSB+05] Alf Inge Wang, Carl-Fredrik Sørensen, Steinar Brede, Hege Servold, and Sigurd Gimre. Development of Location-Aware Applications: The Nidaros framework. In *2nd International Working Conference on Mobile Information Systems*, 2005.

Integrating the individual vehicle in the transport system using open services in a distributed systems architecture

Tor K. Moseng, Marit K. Natvig

SINTEF ICT
P.O.Box 4760 Sluppen, NO-7465 Trondheim, Norway
torkjetil.moseng@sintef.no
marit.natvig@sintef.no

Abstract: Road traffic management has traditionally been targeted through control and monitoring of the flow of vehicles, using the same measures towards all vehicles. As vehicles can have very different capabilities and profiles, it is however desired to use different control strategies towards the individual vehicles to meet the environmental, safety and efficiency targets for the future. SMARTFREIGHT has developed a holistic control and monitoring tool for managing the traffic, and individual vehicles in particular. The individual vehicle is integrated with the traffic management and freight distribution management centers by using open service interfaces for an interoperable information exchange between distributed systems across a heterogeneous wireless infrastructure. SMARTFREIGHT realized and successfully demonstrated this integration in its final event in Trondheim.

1 Introduction

Transport is a significant contributor to some of the major global challenges we face today. About 25% of EU's greenhouse gas emission in 2008 was due to transport, where road transport is responsible for 70% of this [EEA11], while 34.500 people were killed on EU roads in 2009 [EC 11a]. In addition, traffic congestion costs Europe about 1% of Gross Domestic Product (GDP) every year [EC 11b], illustrating the importance of improving future transport solutions.

The traffic management functions and systems of today are not able to identify, monitor and control individual vehicles based on their characteristics (e.g. type of vehicle, engine class, propulsion technology, and carried cargo). This is mainly due to the lack of communication possibilities with the individual vehicles in the traffic, but also due to lack of information about each of the vehicles, such as information about their environmental profiles, their destinations (transport plans) and their carried cargo in terms of freight vehicles. Current traffic management addresses traffic flows in general, and equal instructions are provided by traffic signs or through radio broadcasts to all vehicles on the road or in an area. By not targeting the individual vehicle in traffic, highly polluting vehicles can drive in environmental sensitive areas, and dangerous cargo can without any restrictions be transported in areas with different safety risks like crowded urban areas and tunnels.

New emerging information and communication technologies (ICT) brings along new possibilities within the concept of Intelligent Transport Systems (ITS). Short and medium ranged communication technologies like DSRC (Dedicated Short-Range Communication as specified by the European Committee for Standardization (CEN)) [CEN03], which today mainly is used in automatic tolling systems, and the new WiFi amendment for mobility (i.e. IEEE 802.11p), along with long range cellular systems (e.g. 2G and 3G technologies) enable a continuous connected vehicle for information exchange with roadside equipment (RSE) and traffic management centers. These communication technologies offer capabilities to a diverse of ITS services, among others the possibility for more individualized traffic management. The implementation of such transport services are guided by the standardization organizations ISO and ETSI through their communication architectures Communications Access for Land Mobiles (CALM) [ISO10] and European Communication Architecture (ECA) [ETS10], respectively, both which are results of work committed in several European research projects. One of these projects are CVIS [1] that implemented the CALM architecture for interaction between distributed systems such as central systems, personal devices, vehicle systems and roadside systems [A+07].

In the European research project SMARTFREIGHT [2], the CALM and ECA architectures, along with the CVIS implementations, are taken further by developing services that integrates the individual vehicle with both the traffic management center and its freight distribution center [3]. The different systems were integrated by following the transport service development methodology as described by the ARKTRANS framework [NWMV09], which ensures holistic and generic services for usability across a range of different European city requirements with respect to traffic management and ICT infrastructure. The distributed architecture and open services proved to be a good foundation when implementing and demonstrating the vehicle integration at the final SMARTFREIGHT event in Trondheim, Norway.

The rest of the paper is organized as follows; Section 2 presents the knowledge, which the solutions presented in this paper rest on, Section 3 presents how the vehicle is part of the distributed systems architecture in SMARTFREIGHT, and Section 4 describes the access control use case used for testing the approach. Finally, Section 5 concludes the paper.

2 Background

The overall objective of SMARTFREIGHT was to develop and demonstrate generic technology that can benefit the society by making urban freight transport more efficient, environmentally friendly and safe. The more detailed objective was to address new traffic management measures towards individual freight vehicles.

The SMARTFREIGHT objectives meets the transportation challenges stated early in the

[1] CVIS - Cooperative Vehicle-Infrastructure Systems (EC FP6). http://www.cvisproject.org.

[2] SMARTFREIGHT - Smart Freight Transport in Urban Areas (EC FP7). http://www.smartfreight.info.

[3] SMARTFREIGHT's focus was on freight vehicles in the urban area, for more efficient, safe and environmental-friendly behavior

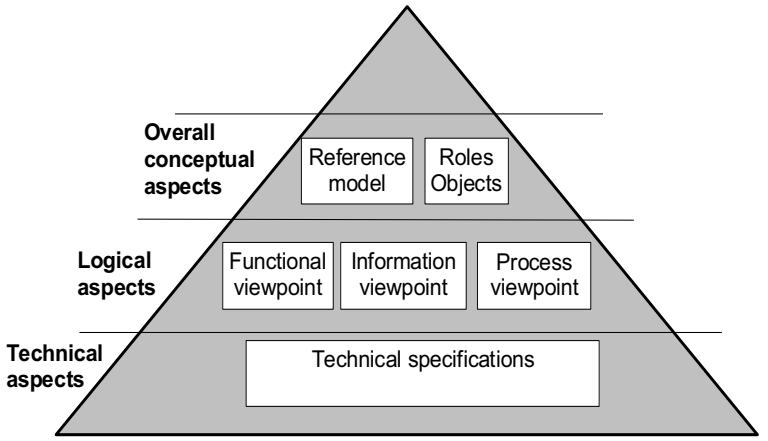


Figure 1: The ARKTRANS content

introduction, and along with the user needs identified through stakeholder consultations (see Section 3.1, the solution presented in this paper has a high problem relevance. The methodology followed in the work to reach the objectives has been on using the process stated as Design Science[4] [HMP04]. The solution is build on using open ICT services, with an emphasis on the interoperability between distributed systems in a heterogeneous communication environment. To obtain this, the SMARTFREIGHT work was rooted on the work made in the CVIS project, which lead to the ETSI and ISO standards on communication architectures for distributed ITS systems. This background knowledge is described in some more detail in the following sub sections.

2.1 ARKTRANS

The transport sector consists of several communicating actors. The objective of the Norwegian ITS architecture ARKTRANS is to gain interoperability when an actor exchanges information with other actors. This is achieved by breaking the transport sector into domains, each with a responsible role, and a set of necessary functions. The functions are further arranged into processes that identifies the required interactions between the different roles. The roles, overall functions and processes are all described by ARKTRANS. Figure 1 shows the ARKTRANS content where the content is gouged into conceptual, logical and technical aspects. The technical specifications are not thorough described – partially due to the focus on implementation independence.

Using ARKTRANS as basis for defining the ITS services in SMARTFREIGHT has pro-

[4]The SMARTFREIGHT work is founded on the transportation challenges of today, and by using established knowledge, innovative solutions are established in an iterative process with development and evaluation.

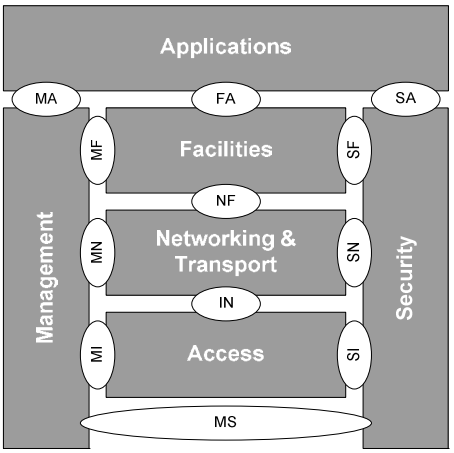


Figure 2: The layered ITS reference architecture [ETS10]

vided holistic, generic and interoperable ITS solutions. However, to accommodate new requirements from the vehicle-to-infrastructure interactions, new and more detailed functionality, processes and information models were defined by SMARTFREIGHT and included in ARKTRANS.

2.2 Distributed Communication Architecture for ITS

The communication architecture, as standardized within both ETSI and ISO, the ECA and CALM, respectively, is based upon a layered reference architecture, which is comparable to the original OSI protocol stack. The reference architecture, shown in Figure 2, defines how ITS services are related to different transport, network and access protocols depending on their communication requirements[5]. Various access technologies like cellular 2G and 3G systems, infra-red (IR), and medium range 5 GHz communication are supported and their use are described in a series of ISO standards from TC204 WG16 [ISO]. One important part of the reference architecture is the management and security support, added as vertical layers. The management layer handles i.a. ITS service advertisements and application lifecycle management, while the security layer handles i.a. authentication, authorization, and certificate management.

ITS systems following the reference architecture interact for information exchange and cooperation; back-end systems (e.g. traffic management systems), handheld personal devices, in-vehicle systems, and roadside systems can thus interact in an ITS domain where ITS routers (i.e. Figure 2 without the Facilities and Applications layer) enable a peer-to-peer communication network. Interaction with external systems (i.e. systems not fol-

[5]The reference architecture is described in more detail in both [ETS10] and [ISO10].

lowing the ITS reference architecture) is handled by ITS gateways [6]. There will typically be many existing services, also services outside the ITS domain, that will interact with systems within the ITS domain. E.g. traffic management and freight distribution centers will have existing traffic and planning services that will co-exist and co-operate with new ITS services, while the vehicle systems must interact with proprietary in-vehicle networks and sensors (e.g. statuses on vehicle components like brakes, engine, etc.). ITS gateways will in such scenarios sustain interoperability with legacy and proprietary systems and networks.

2.3 The Open CVIS Application Framework

The research project CVIS was the first to implement the distributed communication architecture, and developed several communication services, as well as other facility services (e.g. distributed directory service, application lifecycle management service, etc.), to serve end-user ITS applications. These services were bundled within the OSGi framework and made available from there. The facility services have made it easier for application developers to develop new ITS services [B$^+$07, A$^+$07].

The main communication service implemented by CVIS is the CALM Manager service. The CALM Manager handles application requirements and maps these with the availability of the different communication interfaces. E.g. a latency critical application would use a short-range communication interface like CALM M5 instead of CALM 3G.

3 The SMARTFREIGHT Distributed System

The distributed communication architecture, along with the facility services implemented by CVIS, have provided an infrastructure of interconnected systems. SMARTFREIGHT has on top of this infrastructure developed open services that provide information interoperability between the systems, in-vehicle systems included.

3.1 User Needs

A user needs review and stakeholder consultations were undertaken to identify generic user needs and to quantify and qualify the needs for information exchange between urban traffic management systems (UTMS), freight distribution management systems (FDMS) and individual freight vehicles [M$^+$08]. Both the UTMS and FDMS identified the need to exchange information with the individual vehicle (e.g. to give directed information and instructions such as dynamic route guidance). The study in Dublin confirmed that the

[6] Gateway services are part of the Facility layer. Dedicated ITS gateways will thus not need the Application layer.

identified user needs to a large extent cover the requirements that the freight operators have. The user needs were also discussed with the local reference group in both Winchester and Trondheim, and the representatives for both the operators and the city authorities confirmed the user needs collected.

Some UTMS (e.g. Dublin, London) restrict heavy goods vehicle (HGV) access and need enforcement systems to do this. Enforcement systems typically employ automatic license plate recognition (ALPR) cameras and a database containing vehicle registration details for exempt users or for registered users who must pay a fee. The access control solution presented in this paper is a more flexible mechanism to handle vehicle accesses, a solution in where representatives from the road authorities in Trondheim were consulted and involved in during the work.

3.2 SMARTFREIGHT Concepts and Open Services

All cities and regions are different, and the traffic management strategies towards freight distributions also differ. Hence, the cities should be allowed to define their traffic management policies depending on local needs. To support the diversity among cities, SMARTFREIGHT defined a set of generic concepts, which by using these concepts in the information exchange, different traffic management strategies can be handled in a common and generic way [NM11]. Related to the use case presented in Section 4 is the concept *Controlled Area*. It is an area or section of the transportation network that is monitored or has a priority or access restriction schemes (e.g. tunnels, green city areas and parking areas). Other concepts related to the transportation network are *Transportation Network Resource* and *Checkpoints*.

In addition, there are concepts supporting the traffic management, like the *Access and Priority Assignment (APA) policy* (i.e. a formal definition of the traffic management rules for a Controlled Area) and *Access and Priority Offer (APO)* (i.e. priority and access right assigned to an individual vehicle for a Controlled Area). Such area policies will have static rules for normal traffic conditions, while there also might be dynamic rules in case of traffic situations that require specific measures. The access rights and priorities are assigned based on vehicle properties, which arrange for control and monitoring of individual vehicles.

The concepts arrange for generic services that cover many purposes. The concepts are decoupled from underlying communication technology and implementations. The generic services are defined with APIs for information exchange between the systems (see Section 3.3), and to make the solutions transferable across different ICT infrastructures. These APIs also ease the integration of the new services into existing UTMS and FDMS services. Table 1 shows some of the service interfaces defined for interaction with the vehicle.

Table 1: Service APIs for vehicle interaction

Service interface	APIs provided
Resource Management	Req. for/Prov. of resource booking
	Req. for/Prov. of resource booking cancelation or update
	Req. for/Prov. of info on resource availability
Traffic Management	Req. for/Prov. of city/regional policy
	Req. for/Prov. of notifications
	Req. for/Prov. of network and traffic situation information
Vehicle Reporting	Req. for/Prov. of tracking info
	Req. for/Prov. of vehicle info
	Req. for/Prov. of entry/exit notification for controlled area
	Req. for/Prov. of vehicle (safety) status
Route guidance	Req. for/Prov. of route guidance
Goods	Req. for/Prov. of goods tracking
	Req. for/Prov. off gods status
	Req. for/Prov. of goods info

3.3 SMARTFREIGHT System Components

Figur 3 shows the SMARTFREIGHT system components and the communication paths between the system components (strong lines). The vehicle consists of an in-vehicle host and router[7] (i.e. On-Board Equipment (OBE)) and connected cargo (with On-Goods Equipment (OGE)). The in-vehicle host encompasses an application runtime environment[8] for the installed applications. The vehicle communicates with the FDMS and UTMS through service interfaces from Table 1. Both existing and new UTMS and FDMS services can use these service interfaces to take advantage of the SMARTFREIGHT functionality for using the possibilities the new information acquiring brings along.

Roadside stations (i.e. RSE) can function both as a communication relay and as a distributed UTMS; one example is the local control and monitoring in the access control use case in Section 4. Most SMARTFREIGHT scenarios only require existing communication infrastructure like cellular systems as information bearers, while RSEs provide new means for local interaction with vehicles. RSEs use short-range communication technology, which have a limited range, but the higher bandwidth and lower latency enable new possibilities for local control and monitoring of individual vehicles. The Control Centre (CC) and Host Management Centre (HMC), which distributes and manages applications, are included for a complete picture[9].

[7]The mobile router uses the IPv6 Network Mobility (NEMO) for global and ubiquitous connectivity with improved vehicle and session mobility.

[8]The application runtime environment is OSGi with necessary facility services.

[9]Refer to [B+07] for more details on the CC and HMC.

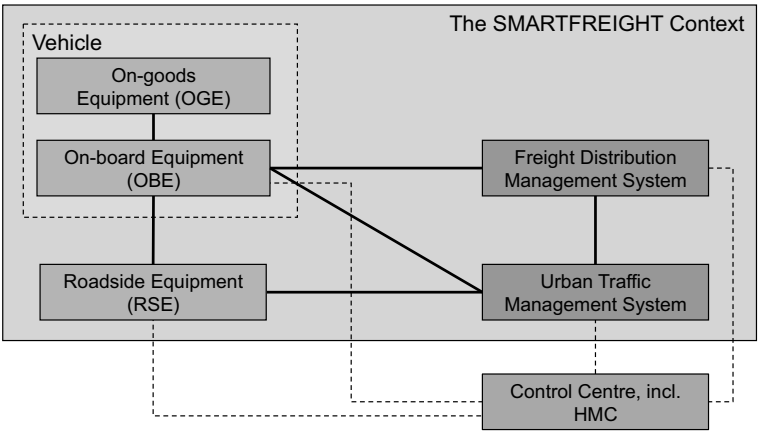


Figure 3: SMARTFREIGHT system component architecture

4 Use Case: Access Control

Vehicle access control can either be performed centrally, distributed, or somewhere in between. The degree of distribution depends on where the individual vehicles' properties are processed and compared to the access policy requirements. Central solutions require a lot of signaling if individual vehicles' properties are taken into consideration (all vehicles will then send their properties for central processing), while the distributed approach is more scalable with respect to processing and communication (a broadcast of the access policy enables distributed processing within the OBEs). The distributed approach also preserves the privacy. As far as we know, SMARTFREIGHT is the first project to specify and implement a complete distributed access control.

4.1 Distributed Access Control

The distributed access control can be used to control and monitor any Controlled Area. It is based on an access policy (i.e. the APA policy) where the traffic management have defined requirements for any vehicle entrance. Figure 4 shows the information elements for the access policy in an UML representation. The figure shows that the entrance will depend on both vehicle properties and timing, while the vehicle can be obligated to report its activities. The reporting is essential for monitoring vehicles' activities (e.g. for statistics or enforcement). Note that identification of vehicles can also be obtained through ALPR cameras. Since each vehicle compares the general access rules with its own vehicle properties (automatically in the OBE), the traffic management is able to reach the individual vehicles with a generic measure. The drivers are then informed on-screen about restrictions and accesses that apply to their specific vehicle; they do not need to be aware of the

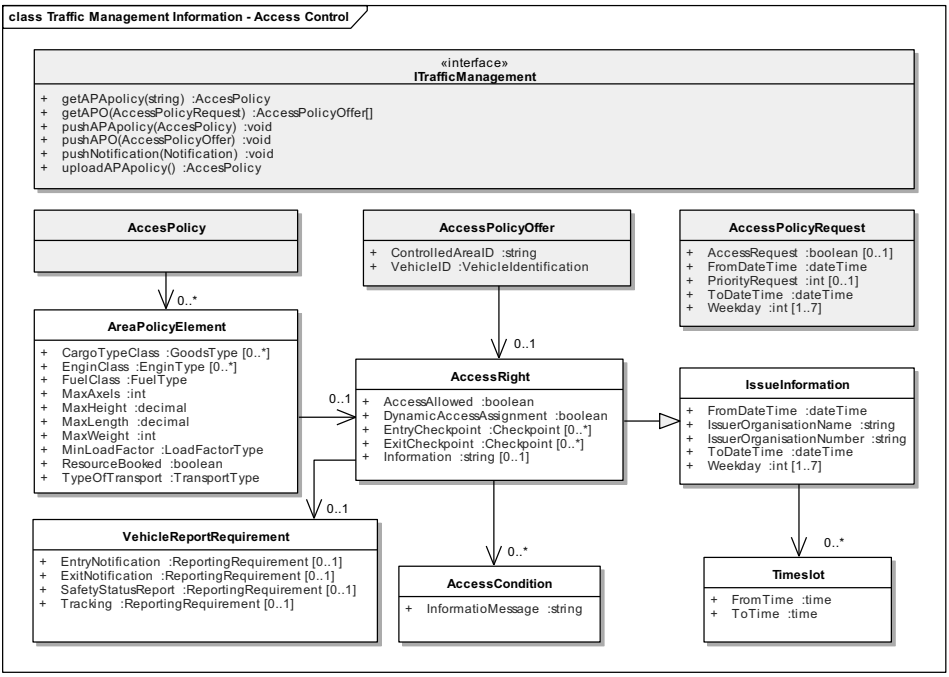


Figure 4: Information elements for the access policy

many informational properties related to their vehicle. Also, the freight management may use the access policy definitions to better plan the transport operations ahead due to their knowledge of their vehicle fleet and cargo to be carried. The traffic management APIs defined in Figure 4 can thus be used by any of the system components in Figure 3.

4.2 Local Access Control

The distributed access control mechanism can also be used for local control and monitoring of a specific Controlled Area. To show this, the mechanism was used as access control for tunnels. In this scenario a tunnel is surrounded by an approaching area, which is guarded by an RSE, for discovering approaching vehicles and communicating with their mandatory in-vehicle tunnel applications. The tunnel access policy contains information about the road network around the tunnel, approach area, holding area, by-pass roads and cargo restrictions that apply to the tunnel, which determines the final access rights of the vehicle[10]. The new element in the tunnel access control compared to the full distributed city/regional access control is the use of a dynamic parameter that gives conditional access until the vehicle actual is present at the Controlled Area (i.e. tunnel) entrance. This is

[10]Please refer to the appendix in reference [Lyk11] for a complete example of a tunnel access policy XML file.

represented by the parameter *DynamicAccessAssignment* in Figure 4. Also, the entrance decision is given by a Tunnel Controller, and not decided by the in-vehicle tunnel application due to the dependency on the necessary tunnel status (e.g. amount of dangerous goods (DG) present, number of vehicles present, etc.).

4.2.1 Realization and Demonstration

Both the city/regional access control application and the tunnel access control application were in SMARTFREIGHT defined as mandatory applications, and download of the applications was triggered by Service Advertisement (SA) messages (a service residing in the Facility layer of the reference protocol in Figure 2) broadcasted by RSEs. When the vehicle enters a RSE coverage area, the vehicle receives a SA message containing an URL of wherefrom the access control application can be downloaded (the URL points to the HMC that manages and distributes applications - see Figure 3). The in-vehicle OBE may choose whatever transmission medium available for the actual download, which was handled by the CALM Manager service. The SA service also ensures that the access control applications are notified about the XML-based access policies.

In the SMARTFREIGHT demonstration, which was held in the city of Trondheim, Norway, a non-DG type of cargo was transported without any access restrictions. Then a reload of DG cargo[11] gave restrictions for entering the inner city area of Trondheim as this area was defined as a Controlled Area with access restrictions regarding DG in the access policy. The tunnel access policy used in the demonstrations allowed only a certain amount of DG classified cargo into the tunnel at a time. *Our* vehicle was therefore instructed to hold and wait at a waiting area before entering the tunnel. Figure 5 shows the driver display where the vehicle is asked to hold before entrance. The communication was based on IPv6 and NEMO, which ensures a fixed contact point towards the vehicle independent of physical point of attachment, while CALM enabled connectivity through 2G/3G and CALM M5 when available. For more information about the demonstration, please refer to the video in [MS10] and the demonstration handout available at the SMARTFREIGHT home page.

The SMARTFREIGHT services are generic and independent of implementation technology. However, some of the services require more from the ICT infrastructure than others. E.g., the download of mandatory applications requires some trigger mechanism like the SA service used in the demonstrations. Possible gate entry technologies to use as triggers can include the WiFi based RSE (as in SMARTFREIGHT), toll collection tags based on DSRC, and geofences based on GNSS information. Also, the introduction of distributed systems where the vehicle is integrated in the information exchange will encompass some challenges with non-supported vehicles without any OBE. Using cameras with ALPR is one alternative, while another is to integrate some level of service support in the forthcoming standardization of electronic registration books that can be integrated with in-vehicle toll collection tags.

[11] The cargo were equipped with communication capabilities on the OGE (i.e. CEN DSRC), informing directly the vehicle's OBE about itself.

Figure 5: Access restricted areas

5 Conclusions

This paper has described how SMARTFREIGHT has relied on a distributed systems architecture, as standardized in both ETSI and ISO as the European Communication Architecture and CALM, respectively, and open services with clearly defined service interfaces for interoperable information exchange between the distributed systems to meet the increasing challenges within transportation today. The use of new emerging ICT enables the integration of individual vehicles in the traffic management for more targeted and effective traffic control and monitoring. Also, the continuously connected vehicle will thus support the freight management with real-time information for improved transport and logistics operations.

The approach has been realized and successfully demonstrated for an use case showing access control on both a wider and a local area. As far as we know, SMARTFREIGHT is the first to define a data structure that can express simple as well as advanced traffic management policies. This opens for exchange of pre-defined policies as well as dynamic policies that can be used in case of abnormal traffic situations. The same structures also allow for use both in a fully distributed case and the more local tunnel access control case. The services are developed independent of existing ICT infrastructure, but an advanced underlying heterogeneous wireless infrastructure will improve the services' functionality and capabilities.

Acknowledgements

The SMARTFREIGHT project is partially funded by the European Commission under the 7th Framework Program for research and development, and their involvement has been highly appreciated. The authors would further like to thank the partners in the SMART-

FREIGHT project for the contribution to the framework architecture and related ITS services – especially the partners SINTEF and Q-Free ASA in the realization and demonstration phases of the access control. In addition, we would like to thank the Norwegian University of Science and Technology (NTNU) for the availability of the roadside infrastructure test site in Trondheim.

References

[A+07] A.Solberg et al. CVIS.D.3.3 Architecture and System Specifications. Technical report, The CVIS project, July 2007.

[B+07] M. Baggen et al. D.FOAM.3.1 Architecture and System Specifications. Technical report, The CVIS project, July 2007.

[CEN03] CEN TC 278. EN 12795:2003 Road transport and traffic telematics - Dedicated Short Range Communication (DSRC) - DSRC data link layer: medium access and logical link control, 2003.

[EC 11a] EC – European Commission. COM(2011) 144 final White Paper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system, March 2011.

[EC 11b] EC – European Commission. SEC(2011) 358 Final Impact Assessment, accompanying document to the 2011 Transport White Paper, March 2011.

[EEA11] EEA – European Environment Agency. Transport emissions of greenhouse gases (TERM 002) - Assessment published Jan 2011, March 2011.

[ETS10] ETSI. ETSI EN 302 665 V1.0.0 (2010-03) Intelligent Transport Systems (ITS); Communications Architecture, March 2010.

[HMP04] A.R. Hevner, S.T. March, and J. Park. Design Science in Information System Research. *MIS Quarterly*, 28(1):75–105, March 2004.

[ISO] ISO TC204 Working Group 16. http://www.isotc204wg16.org.

[ISO10] ISO. ISO 21217:2010: Intelligent transport systems - Communications access for land mobiles (CALM) – Architecture, 2010.

[Lyk11] O. M. Lykkja. D4.3 - Prototype of on-board application. Technical report, The SMARTFREIGHT project, 2011.

[M+08] F. McLeod et al. D2.1 - User Needs Review. Technical report, The SMARTFREIGHT project, 2008.

[MS10] C. Mausethagen and K. Sørensen. Video recording from the SMARTFREIGHT demonstration (20 min). http://www.sintef.no/uploadpages/296176/Smartfreight-Demo.wmv, October 2010.

[NM11] M. Natvig and T. K. Moseng. D5.2 SMARTFREIGHT framework architecture. Technical report, The SMARTFREIGHT project, 2011.

[NWMV09] M. K. Natvig, H. Westerheim, T. K. Moseng, and A. Vennesland. ARKTRANS - The multimodal ITS framework architecture, Version 6. Technical Report SINTEF A12001, ISBN 978-82-14-04444-7, SINTEF, 2009.

A Multiple Platform Approach to Building a Bus Route Information System for Mobile Devices

Runar Andersstuen **Trond Bøe Engell**
{runaa,trondboe}@stud.ntnu.no

Rune Sætre **Björn Gambäck**
{satre,gamback}@idi.ntnu.no

Department of Computer and Information Science
Norwegian University of Science and Technology

Abstract: The paper describes a multiple platform-based approach to creating a bus route information system for mobile devices. The system is context aware: users only need to tell the system (in natural language) where they wish to go, and the system takes care of the rest. The users are presented with a list of possible routes they can take to reach their desired destination. The results are also shown on a map that makes finding the bus stops very easy.

In order to make the system available to as many users as possible, the architecture is client-server-based and relies on technology standards that are widely accepted and implemented, making it easily adaptable to new platforms. The application can be run on multiple platforms, with a minimal amount of calculations needed on the client side. The amount of data transfer between server and client is also kept to a minimum.

The ability to run on multiple different platforms is achieved using technology such as HTML5, PhoneGap and Sencha Touch. The client's functionality includes a search function and a map view, as well as the ability to use bookmarks. The server handles most of the business logic and communicates with external services such as the natural language processing back-end and the server for real-time bus departure information updates.

1 Introduction

The worldwide smartphone market has expanded immensely during the last few years. There were 440 million mobile devices sold by vendors in the 3rd quarter of 2011 [Pet11]. Of these, 115 million were smartphones. This equals a market share of 26.1%. The market share has increased continuously during the last few years. Companies like Apple, HTC and Samsung that have focused on developing smartphones, have gained large parts of the market share. Other companies, like Nokia, that are big on mobile phones, seem to be on a negative trend. As smartphones take over the market, the need for mobile-enabled content and services increases. Unfortunately, the huge amount of available devices has split the market into several mobile technology platforms. Leading platforms like Android,

iOS (Apple), and RIM are all based on different operating systems and code languages.[1] Software developers need to make choices on which platform to support and then learn the native language of that platform. If they want to focus on several platforms and reach out to a larger audience, duplicate efforts are needed to implement specific software on each platform and keep maintaining each code base separately. Consequently, application development time can be immense.

One of the reasons why smartphones are popular is that people always bring their phones with them, wherever they go. The mobility of smartphones opens up for new use-cases where stationary PCs cannot compete. Uncertainty in the highly competitive commercial mobile market space causes questions to appear when approaching the challenges of cross-platform publishing: Which platforms will succeed? What resources and knowledge are needed to make a sensible decision? Fortunately, there is a way around this issue: multi-platform development. The big advantage here is a single code base, which reduces application development time greatly. There are several ways of developing multi-platform software. Determining which strategy that is most suitable can be a challenge, and it is certain that the project requirements must be the primary decision-factor. The main alternative strategies to create multi-platform applications are the following:

Web-based applications make use of HTML5, Javascript and Cascading Style Sheets (CSS) to create mobile websites that aim to look and feel like a native mobile application. Web applications can use JavaScript frameworks, such as Sencha Touch and jQuery Mobile, that are solely designed for mobile development, to replicate mobile user interfaces.[2] Web applications can be conveniently run in a web browser and are therefore already multi-platform, since most mobile device today are equipped with web browsers. One disadvantage for web-based applications is that they have only limited access to device-specific features. Also, they cannot be uploaded to application stores, like the Android Market or the iOS App Store, which will have a negative effect on the availability of the product.

Proprietary Middleware. Applications can also be based on web services such as Red Foundry.[3] Developers get access to a web interface where an application is graphically created by selecting a set of prebuilt modules. When all the modules that provide the necessary functionality have been picked, the service builds a native application which can be submitted to an application store or market. The advantage of this strategy is that the developer does not need any specialist knowledge or programming experience to create applications that look good and perform well. The drawbacks are that the proprietary services often are expensive and that the design and functionalities are limited to what is offered by the service.

Native Applications are written in a specific code language and designed to run in a specific operating system. The main advantage of native applications is that they work as intended by the operating system developers. Device features like sensors, contact lists and storage are easily accessed directly, and the libraries offered by the application programming interface (API) are optimised for the specific operating system.

[1] See `http://www.apple.com/ios`, `http://www.android.com` and `http://www.rim.com`.
[2] Available at `http://www.sencha.com/products/touch` resp. `http://jquerymobile.com`.
[3] `http://www.redfoundry.com/`

Hybrid applications are written as web applications, using coding technologies such as HTML5, CSS and JavaScript. The web applications are then wrapped by one of the available "multiple phone web-based application frameworks" in order to emulate native behaviour. Device features like sensors, contact list and storage are provided by these platforms. Unlike the other alternatives that are confined to browsers and have limited functionality, hybrid solutions form a strong strategy for multi-platform development [Pad11, Chr11]. Developers get a greater control over application design. They use one single code base, but still get access to device features.

In a purely server-based system, all the business logic[4] resides on the server. Having the business logic on a server provides several benefits. First, it saves the client from heavy computations, which is desirable because saved CPU cycles means saved battery power [FZ94]. Second, if all the business logic is handled on the server side, less data needs to be transferred to the client. Finally, since there is one place where all the business logic is handled, optimisation through resource sharing and information caching is easier to implement. Updates are also made easier, for instance, if one of the external service providers decides to alter how their service is accessed, only one central update is needed. If the business logic existed on the client, all existing applications would need to have their code updated to handle the new service change.

In order to reach many users on different platforms, it is essential that the system relies on technology standards that are widely accepted and implemented. The paper introduces a prototype system is called MultiBRIS: a Multi-platform Bus Route Information System. MultiBRIS gives access to a natural language bus route system for Trondheim. The system is called BusTUC [Amb00] and was previously only available via the web or SMS, but through MultiBRIS, it can now be accessed from multiple different smart-phone platforms. The MultiBRIS work is based on a previous application developed specifically for the Android platform [Raa10]. MultiBRIS extends and generalizes this into an application for multiple platforms, supporting real-time bus route information, location tracking and context awareness. All functionality from the previous native Android application has been implemented, and the multi-platform application gives the user the look and feel "illusion" of being a native application. The hybrid application strategy worked out as planned and gave the benefits from using the multiple platform approach we were looking for, making it possible to successfully create one application which can run on both Android and iOS devices without the use of any platform-specific code.

The paper first describes other related bus route applications in Section 2, and then gives an overview of the multi-platform application in Section 3. The benefits and improvements rendered by the client-server solution are discussed in Section 4, while Section 5 concludes and points to areas of future research.

[4] 'Business logic' is a term relating to the functional algorithms that handle the information exchange between a database and a user interface. In this paper, the term will be used to refer explicitly to the part of the system that makes the actual computations and calls the external services.

2 State of the Art

To give an overview of the technology and functionality available in current smartphone route guidance applications, this section first reviews two topical systems, Google Transit and OneBusAway, and then makes a thorough comparison of the various smartphone-based bus route information applications presently available in the city of Trondheim.

Google Transit is a general public transportation planning tool integrated with Google Maps. It essentially consists of two parts: the *Google Transit Trip Planner* (GTTP) which relates to the consumer of the service, and the *General Transit Feed Specification* (GTFS) which is used by data providers (typically public transportation agencies) to feed the Google Transit Service with data [Mor09]. GTTP lets users pose queries to the transit system in three different forms: by address, by a location name with directional indicators (NE, NW, SE, and SW), or by GPS coordinates. All queries are accompanied by date and time, and query types can be combined. A result from GTTP displays both text and directional lines on Google Maps. Google Transit offers a good way for users to directly interact with the service. The drawback is the lack of APIs for external developers, making it impossible for them to use the data in their systems. However, GTFS provides a good starting point as to what data is needed in order to make a good public transport system.

OneBusAway is a set of transit traveler information tools developed for providing real-time arrival information to Seattle area bus riders [FWB10]. It includes a trip planner, a schedule and route browser, and a transit-friendly destination finder. The project has concentrated on tools for providing real-time arrival information and includes functionality supporting location sensing. An iPhone application was created first, to exploit its localisation framework and built-in multi-touch map support. Later this was generalised in a JavaScript-based experimental multi-platform web application for real-time arrival information. A OneBusAway user study provided valuable user feedback for such applications, including that the users want a bookmark functionality, enabling them to tap a bookmarked destination at any time and receive route suggestions to it from their current location.

To find other bus route information applications in Trondheim, both the web, the Android Market and the Apple App Store were searched. Table 1 compares the functionality of the applications found on October 5th, 2011. The applications range from simplistic to more sophisticated with lots of functionality, as indicated by the comparison chart.

Google Transit differs from the other systems by being purely server-based. All business logic resides on the server. The client, which is a web browser, only handles the display of information. For the Trondheim applications, it is interesting to note that the web-based BarteBuss that has been created by means of HTML5 and JavaScript is the most feature rich and well-working application of them all. This hybrid strategy, using HTML5, JavaScript, CSS, and a deployment technology can be used to easily create a client prototype that works on several platforms. As discussed above, a server-based solution in addition offers several other benefits, and is thus chosen here.

	Buss-Orakel	Barte-Buss	Buss-tider	Alf's Bybuss	Buss-droid	BusApp Tr.heim	Buss-ruter	Buss-øye
Platform	A/iP	web	A/iP	A	A	A	A	iP
Multi-Language	Yes	No	No	Yes	No	Yes	No	No
Cost	No	No	No	No	No	Yes	No	No
Favourites	No	Yes	No	No	Yes	No	No	Yes
History	Yes	Yes	Yes	Yes	No	Yes	No	Yes
Route download	No	No	No	Yes	No	No	Yes	No
Map function	No	OSM	GM	GM	No	GM	No	GM
Closest bus stops	No	Yes	No	No	No	Yes	No	Yes
Uses GPS	No	Yes	Yes	Yes	No	Yes	No	Yes

Table 1: Comparison of the bus route information applications available in Trondheim
A=Android, iP=iPhone, web=HTML5, OSM=OpenStreetMap, GM=GoogleMaps

3 System Overview

This section describes the prototype server developed during this project. First, the system services are introduced and then a description of the technologies is given. The importance of a multi-platform application has been explained earlier. The languages HTML5, CSS and JavaScript make it possible to create a multi-platform solution. In order to reach out to all users, the application also needs to be available from a variety of application stores which provide opportunities for publishing, advertisement and collecting fees. Hybrid applications are designed to give the advantages describes above using deployment technologies. The disadvantages of the hybrid-solutions will dissipate as the deployment technologies, along with the browsers on the devices, mature and become more robust.

The server offers three distinct services. Figure 1 shows a block diagram of the system. The "main service" effectively replaces all the business logic implemented in the previous

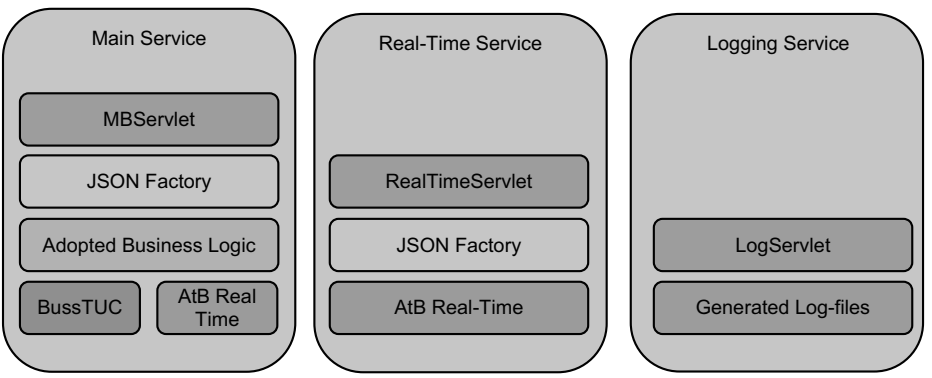


Figure 1: Block diagram of the server system
(MB = MultiBRIS, JSON = JavaScript Object Notation)

Android application [Raa10] and is described in Figure 1. The system also includes a "real-time service" where the client can send a bus-stop ID to the MultiBRIS server and get a list of the exact time for the next five buses arriving at that stop. Finally, the server provides a "logging system service" (for debugging purposes) which is easily accessible through a web browser.

3.1 The Main Service

Figure 2 shows a complete interaction diagram for the main service.

The HTML returned from the BusTUC natural language system contains both a textual answer and a JSON object, but TABuss only uses the JSON object. Calculating the distance between two GPS coordinates is done on the server. The distance calculation is not as trivial as just using the Euclidean distance, because of the approximately oblate spheroid shape of the Earth. The distance (between two GPS locations) is therefore based on Vincenty's inverse formula [Vin75], using the World Geodetic System (WGS 84) standard.

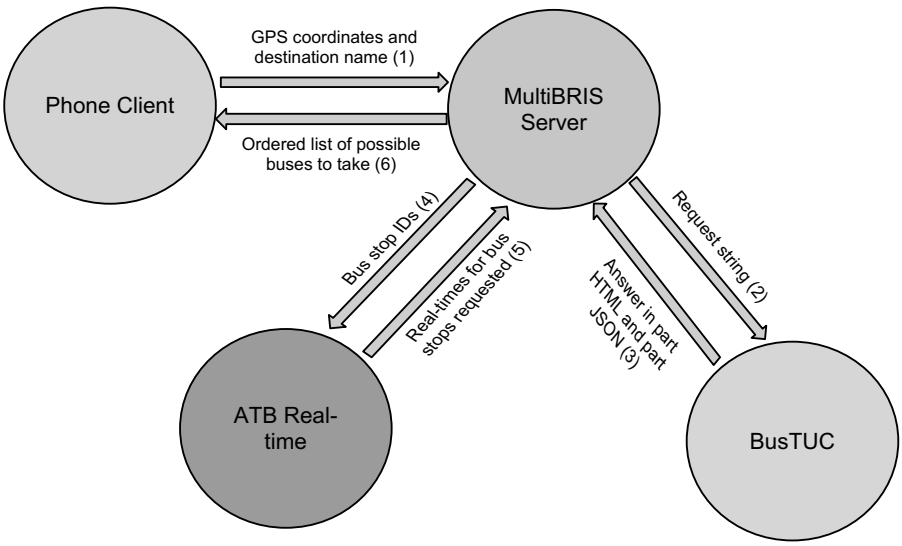


Figure 2: Main service overview.
1: The client sends a desired destination and its current location to the MultiBRIS server. The server looks up a fixed number of bus stops near the client's location.
2,3: When the bus stops are found, a query is sent to the BusTUC web-interface which responds with a set of the next scheduled bus routes from the clients destination to the target.
4,5: The MultiBRIS server then updates these routes with real-time departure times. The real-time update is done by contacting a real-time system provided by AtB, the current public transportation service provider in Trondheim.
6: The MultiBRIS server sends the updated route alternatives back to the client.

3.2 Technologies

The aim of MultiBRIS is to reach as many mobile device platforms as possible. This is the key point when choosing a deployment technology, ruling out solutions like Appcelerator Titanium that only supports iOS and Android (As well as solutions placing a heavy burden on the developer to know a particular technology, e.g. Rhodes which requires knowledge of Ruby and has an extensive API). Another prerequisite is that the framework should be easy to use and be able to collaborate with other frameworks that can make it faster and easier to create well-working GUI-components.

This is easier in PhoneGap[5] than in the other deployment technologies. **PhoneGap** is an open-source mobile development framework enabling software programmers to build applications using JavaScript, HTML5 and CSS3. Our MultiBRIS client application does not require heavy computing or graphics, so the performance weakness of PhoneGap is not a big problem. We use PhoneGap as our deployment technology, together with the Sencha Touch JavaScript library for graphical implementation.

The server technology used is Java Servlets, a technology which now is at version 3.0 and has been around for over a decade. With Java Servlets, business logic can be written in Java and then made available to consumers through servlets [Per04]. A Java servlet can be published through any available servlet container, making it very portable. We chose to use the **Jetty**[6] servlet container which is made up of pure Java code, ensuring portability.

4 Results

Migrating the business logic from the phone to the server resulted in two clear benefits, in addition to making the implementation easier: query time reductions were obtained by introducing a new web-interface for BusTUC (Section 4.1), while both the amount of data transfer and the power usage were reduced (Section 4.2).

4.1 Query Time Reduction

The MultiBRIS server has been optimised in two ways. Sharing the ”bus ID to bus-stop live ID” lookup list between all the clients using the server, and by the introduction of threading in the retrieval of real-time data for multiple bus stops. This resulted in computation times twice as fast as the original, as multiple threads can send queries in parallel, instead of sequentially. However, the time from when the query was posed to the server response was still too long, sometimes up to 30 seconds, since the MultiBRIS server had to wait for answers from the BusTUC server. Hence speeding up the BusTUC server was imperative for the practical usability of the entire system.

[5] `http://www.phonegap.com/`
[6] `http://www.eclipse.org/jetty/`

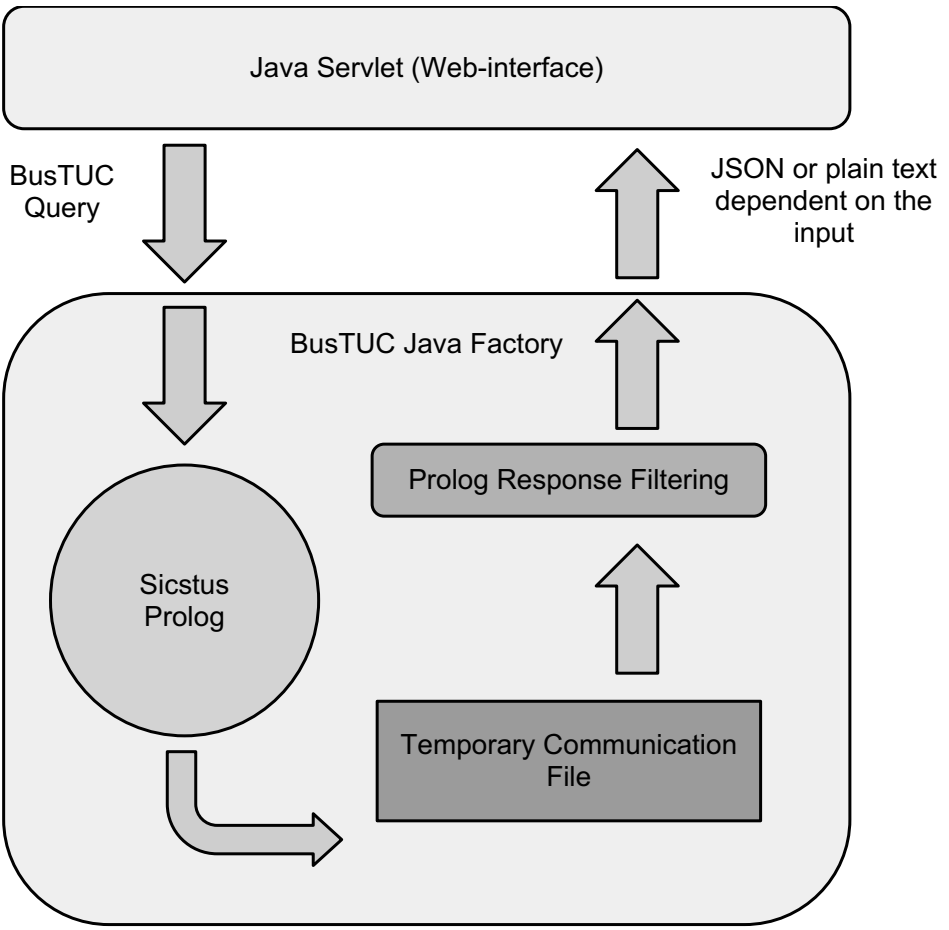


Figure 3: The BusTUC web-interface

As a response to this, a new web-interface for BusTUC was built, using Java Servlet technology with the same Jetty container as the MultiBRIS server. Figure 3 shows an overview of the web-interface. When a query arrives at the Java Servlet, the query is sent to what is called the BusTUC Java Factory. The BusTUC factory poses the query to the BusTUC Prolog code, which, in turn, puts the answer into a temporary file. This file is then read and filtered (in the Prolog Response Filtering Module) before the result is returned to the Java Servlet. By creating a new BusTUC web-interface, a substantial reduction in query time was achieved. The average query time was reduced from 15 to 6 seconds, but the most important improvement was in the maximum query time, which was lowered to just one third of the original (from 30 to 10 seconds).

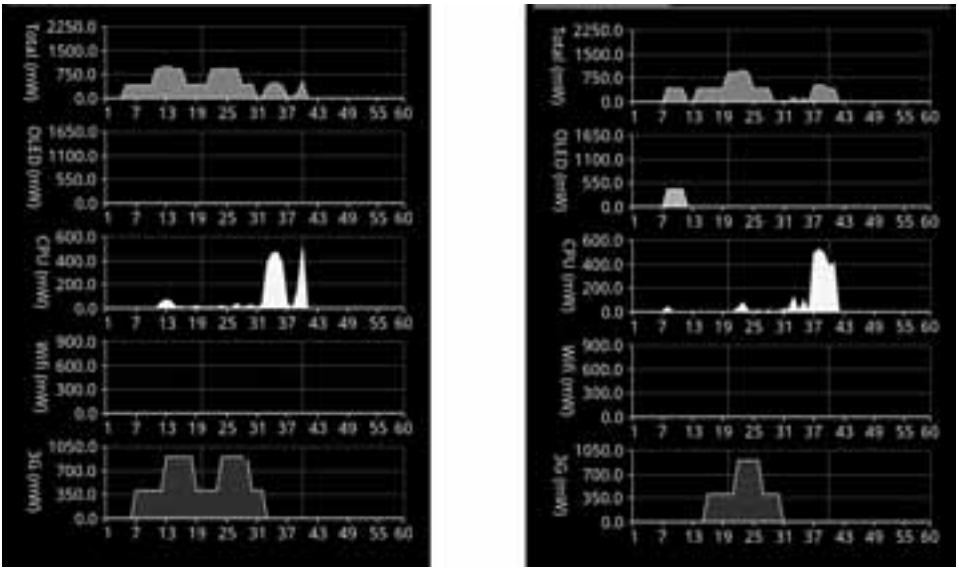
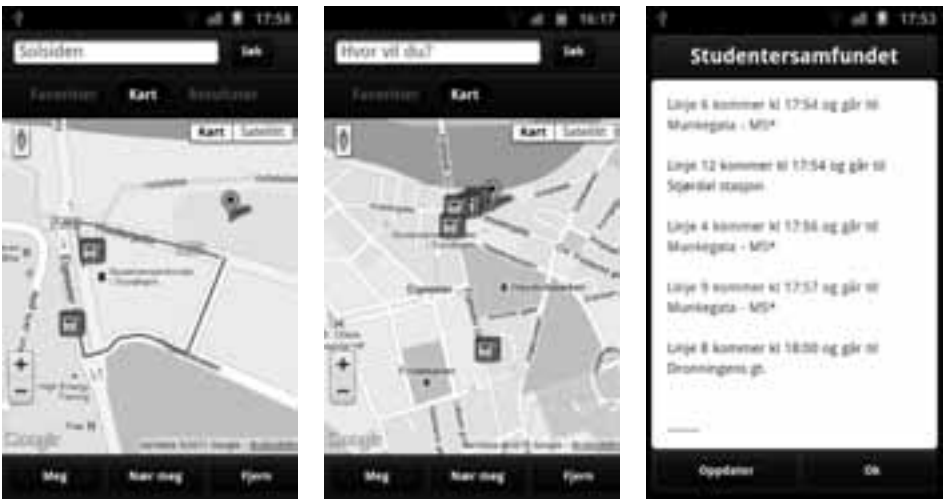


Figure 4: Power usage with business logic on the client (left) vs. on the server

4.2 Server-side Testing

One of the results of moving the business logic to a server is a decrease in the data transfer to the client. If all business logic were to be handled on the client side, all the data would need to be transferred to the client. A lot of data needs to be transferred, since the business logic requires one request to be sent to BusTUC, and several requests to be sent to a SOAP service for the real-time data. By handling the business logic on the server, the server can make all these requests to the various services. Since some of the requested resources can be shared between all the clients using the server, the server can also relieve the external services from heavy request load, by caching the results.

The average data transfer for the client before and after the business logic was moved to the server was reduced from 570 KB to 5 KB (measured with WireShark 1.6.1; the client used was an AppleWebKit 535.2 based browser). The measurement scenario consists of starting the client and making (a bus route alternatives) query, that is, using the "main service" on the MultiBRIS server (Section 3.1). This scenario was chosen as it would represent the most natural usage pattern. About 400 KB of (extra) transferred data comes from downloading the live ID to bus-stop ID list from AtB. Regarding the saved power usage, Figure 4 shows a comparison between power usage for having business logic on the client and on the server. The topmost graphs display the total amount of milliwatts (mW) usage for each application, and clearly show that the solution where the business logic is on the client consumes more power than the server-based solution.



(a) Results shown on the map (b) Close-by bus stops (c) Real-time information

Figure 5: Screenshots

4.3 The Client Application

When the MultiBRIS client application starts it tries to retrieve the user's current location. A search bar always resides on top of the application, giving the user quick access to bus route search functionality from wherever the user has navigated in the application. The search bar has an auto-complete function that suggests bus stops in Trondheim.

As the result of a search query, the user is presented with a list of the five most optimal bus routes, sorted by total travel time. The user can either tap on the most suitable result to switch to the map and view the target bus stop and travel route to that bus stop, or click on a map tab and be presented with all the bus stops in the result. Coloured lines show the user how to get to the respective bus stops from the current location (Figure 5a).

The map of the application consists of a Google Map with added functionality. The current location is centered on the map when the user taps the "*Meg*" (me) button. In the map tab the user can also click on "*Nær meg*" (close to me) to add nearby bus stops to the map (Figure 5b). The user may click on a bus stop to get real-time information on the next five buses passing through it. Buses are marked in red if they have actual real-time values, or black if only scheduled times are available. If any of the buses are among the user's previous search results, they are also marked with an "*" (Figure 5c).

5 Conclusions and Future Work

The implementation of the MultiBRIS server and update of the BusTUC web-interface proved successful. The system as a whole went from having a client-application that transferred up to 500 kB of data for a bus route query and a query time that was around 20 seconds, to transferring only 5 KB of data and having query times at around 10 seconds. Saving both time and data transfer was imperative for the practical viability of the client. When looking at the power usage in Figure 4, a surprising property was revealed: the difference in power usage in a large part comes from the extra data transferred and not from more CPU usage. Normally, the CPU cycles and the display are portrayed as the main battery power consumers; however, the results here indicate that data transfers can consume as much power as CPU cycles in some cases.

As shown, there are a lot of benefits from moving much of the business logic to a server. The client saves battery and it is easier to maintain and update the system. However, there are some possible drawbacks with adding a server as part of the solution. Doing this effectively creates another layer which the information has to pass through to reach the client, adding another point which can potentially fail to work properly. Another aspect is that the server, when used in a production environment, needs proper infrastructure as a foundation in order to be reliable enough for any client to use. Hence, in a commercial solution, the infrastructure cost for a server-infrastructure has to be considered.

An interesting extension would be to look at systems taking intelligent decisions on where to compute, such as Spectra [FPS02] which dynamical decides whether to perform computation on the server or on the client. "Spectra" monitors resource usage both on server and client, and makes an "optimal choice" based on given system parameters. This functionality could be implemented for MultiBRIS so that if, for example, the MultiBRIS server was under such heavy load that it would delay query times, the clients could be instructed to perform the route calculations and contact the underlying services themselves. However, this would make the client code grow substantially, as it would need to contain all the business logic necessary to perform computations and service calls.

Acknowledgments

Thanks to to LingIT AS and AtB AS for allowing the use of the BusTUC and the real-time bus information services, respectively; to Chistoffer Jun Marcussen and Lars Moland Eliassen for their valuable ideas and feedback during the project; and to Magnus Raaum for providing the source code of the original Android application.

Special thanks to Tore Amble: your memory lives with us and with the BusTUC system which you devoted so much of your life to.

The work of the third author was partly funded by the Research Council of Norway, through the project 187865/S10 UbiCompForAll.[7]

[7] http://www.sintef.no/Projectweb/UbiCompForAll/

References

[Amb00] Tore Amble. BusTUC: A Natural Language Bus Route Oracle. In *6th Conference on Applied Natural Language Processing*, pages 1–6, Seattle, Washington, 2000. ACL.

[Chr11] Adam M. Christ. Bridging the Moble App Gap. *Sigma*, 11(1):27–32, October 2011. Special Issue: Inside the Digital Ecosystem.

[FPS02] Jason Flinn, SoYoung Park, and M. Satyanarayanan. Balancing Performance, Energy, and Quality in Pervasive Computing. In *Proceedings of the 22nd International Conference on Distributed Computing Systems*, pages 217–226, Vienna, Austria, July 2002. IEEE.

[FWB10] Brian Ferris, Kari Watkins, and Alan Borning. OneBusAway: Location-Aware Tools for Improving Public Transit Usability. *IEEE Pervasive Computing*, 9(1):13–19, January–March 2010.

[FZ94] George H. Forman and John Zahorjan. The Challenges of Mobile Computing. *IEEE Computer*, 27(4):38–47, April 1994.

[Mor09] Daniel Moraff. Google Transit Feed Specification: A Primer. Website, November 2009. `http://www.eot.state.ma.us/downloads/developers_beta/MassDOT_GTFS_Primer.pdf`.

[Pad11] Richard Padley. HTML5 - bridging the mobile platform gap: mobile technologies in scholarly communication. *Serials*, 24(3):S32–S39, November 2011.

[Per04] Bruce W. Perry. *Java Servlet & JSP Cookbook*. O'Reilly Media, Sebastopol, CA, USA, January 2004.

[Pet11] Christy Pettey. Gartner Says Sales of Mobile Devices Grew 5.6 Percent in Third Quarter of 2011; Smartphone Sales Increased 42 Percent. Website, November 2011. `http://www.gartner.com/it/page.jsp?id=1848514`.

[Raa10] Magnus Raaum. An Intelligent Smartphone Application. Master's thesis, Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway, June 2010.

[Vin75] Thaddeus Vincentry. Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations. *Survey Review*, 23(176):88–93, April 1975.

SESSION 3

Health, Wellness and Emergency Support

Key Technological Success Features for a Domain Specific Open Software Ecosystem for Ambient Assisted Living

Marius Mikalsen[1], Ståle Walderhaug[1], Dario Salvi[2], Geir Kjetil Hanssen[1]

[1]STIFTELSEN SINTEF
NO-7465 Trondheim, Norway
marius.mikalsen@sintef.no
stale.walderhaug@sintef.no
geir.kjetil.hanssen@sintef.no

[2]LST
Life Supporting Technologies
Universidad Politécnica de Madrid
Avenida Complutense, Madrid, Spain
dsalvi@lst.tfo.upm.es

Abstract: Ambient Assisted Living (AAL) is a domain with great potential for economic and societal impact. But market uptake of such solutions is so far limited because of market, standards and technology uncertainty. The businesses that will prevail are those that are cost effective. To support cost effective development of AAL solutions, domain specific open software ecosystems are being established. We developed a survey to investigate key technology success factors for such ecosystems. The survey was sent to 60 developers from a representative selection AAL development projects. 18 responded. Following a qualitative data analysis we found several key factors and features that must be in place to facilitate the success of such ecosystems. We found that given the nature of the AAL domain, characterised by divergent users, software and hardware, developers are seeking for support in three main areas. First, they want the artefacts available in the ecosystem to support relevant standards in the domain. Second, support for tracing artefacts available in the ecosystem to requirements (domain-fit) is needed. Third, they want support for developing, testing and emulating for complex user-software-hardware workflows in this distributed environment. The main obstacle that will scare away developers from the ecosystems is lack of documentation of the artefacts in the ecosystem. Second, not enough decoupled components, and finally, lack of proper search features. Finally, in order to be able to learn to use the artefacts, examples, scenarios and API documentation is necessary.

1 Introduction

As the western population ages, more support is needed with fewer hands to cater for their needs. This has led to a recent attention to ICT for active and healthy ageing, commonly known as Ambient Assisted Living (AAL). AAL is applications and technology that provide support for active daily life, inclusion in the society, and

prevention and management of sickness and disease [VDB10]. ICT for Active and Healthy Ageing (AHA) has a great market potential, estimated to potential annual revenue of 1837 M with moderate take up, and 2576 M with high penetration [KM10].

Adoption of such technology is still limited, because it is not yet clear how the market for AAL will emerge or what kinds of applications, technology and standards are most likely to be successful in the domain. The business that will prevail are those that are flexible enough to deliver technology addressing changing market requirements more cost effectively than the competition. Cost effective in this context is to provide sufficient quality AAL applications at lower costs than the competition.

To support such cost effective development and to spark the breakthrough of AAL applications and services, the European Commission has launched the Integrated Project universAAL[1] that aim to consolidate previous EU funded research and take steps towards a *domain specific open software ecosystem* providing a standardized approach making it technically feasible and economically viable to develop AAL applications.

Hanssen defines the emerging concept of software ecosystems as [Ha11]: "… *a networked community of organizations, which base their relations to each other on a common interest in a central software technology*." The ecosystem created by universAAL is denoted *open*, which implies that it is centred on free/libre open source (FLOSS) technology.

This paper introduces the foundations of the universAAL project and reports on a survey done by the project on key technology success features needed in such an ecosystem.

This paper is structured as follows. First, in the background section, we present the foundation of universAAL in terms of a characterisation of AAL technology, open software ecosystems, as well as introducing the universAAL ecosystem itself. Second, we present the methodology used for performing the survey. Third, lessons learned so far from AAL application developers´ perceptions on what are key technological qualities that are important for the success of the software ecosystem, pointing out important success factors for the ecosystem to succeed. Forth, we discuss the lessons learned in terms of recommendations for building such an open AAL software ecosystem. Finally we summarize our experience with our concluding remarks and identify steps for future work.

2 Background

2.1 Ambient Assisted Living Needs an Ecosystem

In the conclusions of its meeting of 4 February 2011, the European Council endorses the Commission's proposal for an Innovation Union, and in particular the launch of a

[1] http://www.universaal.org/

European Innovation Partnership on Active and Healthy Ageing[2], by stating that: "*Innovation contributes to tackling the most critical societal challenges we are facing. Europe's expertise and resources must be mobilized in a coherent manner and synergies between the EU and the Member States must be fostered in order to ensure that innovations with a societal benefit get to the market quicker*."

Farshchian et al has done an analysis of innovation characteristics in the AAL domain, and discuss the fact that AAL technologies operate in complex settings involving multitude of stakeholders and organizations. The target end users, i.e. the elderly and their next of kin, constitute a varied group of people with a wide range of needs [FHM12]. They conclude that: "*A range of aspects related to innovation in the marketplace are still missing*." This indicates the need for creating an ecosystem.

This need is also evident in the Lecce declaration, which unambiguously calls for the establishment of a well functioning software ecosystem[3]: "*…work should be directed at building sustainable ecosystems through targeted work, e.g., on ecosystem design, ecosystem compliance and interoperability tests, ecosystem marketing, and life-cycle management of products and services.*"

2.2 What is a Software Ecosystem?

As the term of software ecosystem is emerging there is not yet any single authoritative definition of the concept. In addition to the definition presented in the introduction, Jansen et al provides the following [JFB09]: *"… a set of businesses functioning as a unit and interacting with a shared market for software and services, together with the relationships among them."*

Bosch focuses on the common interest in software and its use [Bos09]: *"the set of software solutions that enable, support and automate the activities and transactions by the actors in the associated social or business ecosystem and the organizations that provide these solutions."*

The perhaps most well known example of a software ecosystem is Apple´s iOS and App Store, where developers can build, deploy, market and sell applications in the ecosystem owned and controlled by Apple. Application developers are drawn to the ecosystem by the potential impact their applications can have. We denote it as closed as Apple controls the hardware (iPhone, iPad etc.), the iOS operating system, the development tools, and the distribution channel (App Store).

Another well know eco-system, that is semi-open, is Google´s based on the Android OS. This consists of several hardware vendors that deliver phones and tablets running the open source operating system Android. Although Android is open source, the development of the operating system itself is closed and controlled by Google. Also, Google controls the market place, Android Market.

[2] http://ec.europa.eu/research/innovation-union/index_en.cfm?section=active-healthy-ageing
[3] www.aalforum.eu/group/leccedeclaration

UniversAAL, as we shall see in the next section, adopts a very open approach to the notion of creating an ecosystem for active and healthy ageing applications.

2.3 universAAL: Steps Towards an open Ecosystem for AAL

universAAL is a 4 year 13M European Union project that develops an open technology platform for active and healthy ageing. The goal of the project is to create a domain specific open software ecosystem to answer the needs identified by the AALIANCE roadmap [VDB10], including architectures, software components and development and design tools. All results developed in the project will be available as open source under the Apache 2.0 license. In support of the technology platform, communities are established to maintain the project results. In combination, this makes the project constitute a domain specific open software ecosystem for AAL. The main results of the project are introduced below.

Technological Results:

The universAAL Runtime Support [Taz11] is a set of libraries that must be installed on devices that support universAAL applications. At the lowest level it provides a common interface for abstracting resources, e.g., sensors pushing data through the universAAL platform (Execution Environment).

On top of this hardware level abstractions there are some basic services, such as connectivity, security, context awareness and personalization support. The AAL Platform Services enrich these services realizing functions such as secure access to medication information or shared care plan.

Two versions of the runtime support are developed. One is based on OSGi[4] and is available for download and use in version 1.0 from the universAAL Developer Depot. An Android version of the runtime adhering to the same reference architecture is planned.

The developer depot is a repository where the AAL developer community can browse and download existing as well as publish new services. Developers can reuse existing universAAL AAL platform services in the depot in order to create composite platform services [SW11]. The developer depot is accessible from http://depot.universaal.org/.

Using the Eclipse IDE[5] as the basis, new software (plug-ins) that incorporates the universAAL reference architecture and design concepts (UML Profiles from the universAAL reference architecture) is available. Part of the tools is dedicated to expressing the universAAL architecture and ontologies using model driven software engineering techniques such as model-to-model transformation and model-to-text generation [SW11].

[4] http://www.osgi.org/
[5] http://www.eclipse.org/

The reference architecture represents a detailed specification of functional components and the relationships among them for building systems in the AAL domain. Its most central parts are:

- Reference Use Cases: shows the main stakeholders of the system and how they would interact with it
- Reference Requirements: is a collection of requirements for a generic AAL system
- Reference Detailed Architecture: is a proposal of an architecture with its modules and functionalities that are mapped to the use cases and requirements

The uStore [BBA11]: is marketplace that enables selling and buying of AAL services. The uStore will enable service providers to specify complete AAL services (the required software, hardware and resources) and sell these through a secure and easy to use interface.

Finally, universAAL is producing a set of example services that demonstrate use of the underlying platform capabilities. The services will be used both for giving the adopters of the platform a training material for understanding the functionalities offered by the runtime support, and for reusing software components that can be included into new services. The services include an implementation (i.e. the source code) and documentation [Ibz11].

UniversAAL Community Results:

To sustain the universAAL results after the project ends, focus has been to build a community that can maintain and further develop the results in the ecosystem. To this end, the Ambient Assisted Living Open Association (AALOA) has been established and has currently, over 100 individual supporters. AALOA is available from http://www.aaloa.org/. The main mission of AALOA is to bring together the resources, tools and people involved in AAL in a single forum that makes it much easier to reach conclusions on provisions needed to achieve AAL progress. Second, make sure that all technology providers, service providers and research institutions involved in AAL are either directly involved in AALOA or (as a minimum) aware of decisions it promotes. Third, involve end-user representatives in all work of AALOA. Fourth, identify key research topics in AAL, and reach agreement on prioritization of these. Fifth, design, develop, evaluate, standardize and maintain a common service platform for AAL.

3 Methodology

We developed a survey sent out to a selected set of AAL projects to get insight that will help shape the development and evaluation of the software development tools created in the universAAL project. Developers from five different AAL-related EU projects were asked to participate in the survey (the survey was sent to 60 developers, 18 responded). The reason for asking developers in the selected projects is that our goal is to build tools

that solve key challenges facing developers of AAL software, and we wanted to get an understanding of what key factors that will influence developers' acceptance of such open tools and components. Participation in the survey was motivated by the tools being based on an existing development platform (Eclipse) and that the results will be provided for free and as open source.

In accordance with Norwegian law, the survey was approved by the Data Protection Official for Research, Norwegian Social Science Data Services.

Questionnaire Design

The questionnaire was designed to address what the developers perceived as important aspects of developing software for AAL/e-inclusion projects with a strong focus on developer tools and practices. The questionnaire was designed by the researchers in the universAAL project with the goal of getting answers to the main barriers facing developers of AAL systems, and what they would see as the most beneficial features of the universAAL tool results. The questionnaire was pre-tested by participants of the universAAL project before being sent to make sure that all items were clear and understandable.

The survey had two core items pertaining to the creation of the universAAL software ecosystem:

1) Investigate current development practice, and design and expected benefits of an AAL software ecosystem.

2) Investigate possibilities and obstacles of a dedicated AAL software ecosystem, and learning strategies.

The questionnaire was based on 5-point Likert scales and fields for textual input where respondents could provide more information when and where applicable.

Subjects

Developers from five different AAL-related EU projects were asked to participate in the survey. The projects were Amigo[6], Soprano[7], MPOWER [St11], Persona[8], Genesys[9], and Oasis[10]. The invitation to participate where sent through project internal developer mailing lists. A total of 60 developers were invited. 18 responded.

Data Analysis

The rather low number of respondents (n=18) reduces the effect of doing statistical analysis (like regression analysis etc.). The result does however lend itself quite well to

[6] http://www.hitech-projects.com/euprojects/amigo/
[7] http://www.soprano-ip.org/
[8] http://www.aal-persona.org/
[9] http://www.genesys-platform.eu/
[10] http://www.oasis-project.eu/

data analysis using descriptive statistics and frequencies. The survey also allowed for free text input to allow developers to provide input important to them on selected items. As the survey is meant for gathering input to shape the development and evaluation of the software development tools created in the universAAL project, we see the results as valid for their intended purpose.

For the analysis of the free text input we report the text provided to us as well as the researchers analysis of the provided input. The text were analysed based on the principles of "Grounded Theory" [SC98], which can be used to categorise or find central concepts from textual data. We used an open coding approach where we identified categories of phenomena within the topics selected for the survey.

4 Results

4.1 Findings on Current Development Practice and Expected Benefits of Tools and Components in an AAL Software Ecosystem

Summary of data on programming languages and tool usage

We found from the self-reported programming skills that the highest skills were reported for Java (a mean score of 3.44 reported). Knowledge of OSGi development was rated lower. Design and modelling knowledge (such as UML), gets a lower rating (a mean score of 2.72 reported). The most frequently used developer tool in our sample is Eclipse.

Summary of data on the importance of use of standards

The statement "Using standards is important for my company/organization" got a mean score of 3.94. The statement "Standards are relevant in my work" got a mean score of 3.89.

Summary of data on design and development tasks that would benefit from tools support and reusable components

We found from the categorization of responses that a main thing that developers would benefit from having support in is design, documentation and traceability support. This can be in the form of traceability from requirements, through design and to the components, or by documenting use cases. The respondents also wanted to get support in requirement analysis towards finding AAL specific services, that is, a link between requirements and available components in the ecosystem. This is also reflected in the feedback on the user interface artifacts developers would like to have available in the ecosystem, where the respondents highlight a need for reusable user interface components for divergent devices (e.g. PCs and embedded) and proven user interface designs.

The mean value reported for use of design patterns is quite high, 3.88, indicating that developers use design patterns. Also quite high is the use of code conventions, 3.59. A mean of 3.47 suggests that the development process tends to be repetitive and follows a certain pattern or structure.

Equally important in terms of number of responses is the request for support for deployment to divergent smart devices, sensors and other hardware. Developers also want support for configuring gateways and integration of sensors into their solution. Finally they want to be able to emulate and test on emulators for smart devices/sensors. Two respondents that report need for testing and automated testing, and the reported need for workflow management and messaging in distributed OSGi environments.

4.2 Findings on Possibilities and Obstacles of a Dedicated AAL Software Ecosystem, and Learning Strategies.

Summary of data on how to access to software components

We found that the majority (15 of 18) would prefer to have access to development resources (code, designs etc.) directly through the development tools they use. The alternative is to use a web browser to navigate such repositories.

We also investigated what would be the main features of a dedicated AAL repository. The primary feature required was to have documentation and examples. This includes descriptions of components, manuals, APIs, tutorials and functional code examples.

Second, the components that developers wanted are in three main categories. These are, support for different platforms (hardware and software), libraries for AAL specific reoccurring tasks (including user interface components), and examples real application (that has been in use) that can be used for evaluating new project ideas.

Only two developers mention open source and only one mention the possibility to get user ratings of components in the repository.

Summary of data on current obstacles to code reuse

The primary obstacle to reusing code in terms of number of responses is lack of documentation and bug tracking. The second most important obstacle is that components that are available for reuse are not open enough in that they have to many interdependencies with other components and are not properly decoupled. The third most important obstacle is that it is hard to find appropriate components due to lack of search support ("*Google is not enough*")

Summary of data on assets that would help a developer to implement a successful application

The results show that examples are the primary source of information developers would need. This ranges from example of use of the application (including demonstrations,

screencasts and step by step descriptions) and also examples including code samples. The second most important asset is scenarios, use cases and requirements in the domain. These should also be accompanied by code examples. The third most important asset is API and other documentation including standards and UML models. Other sources of information are feedback from experts in the domain, tutorials and tool support.

This is also supported by the numbers we found when we asked what resources the developers would use when learning new platforms and code components in a new domain (such as AAL). First, the majority (n=16/18) look at example code to learn. Second, they use API documentation to learn (n=15/18). Third, quite few look at designs (n=5) and use project wizards and templates (n=4) to learn. Finally, one respondent lists other sources for learning, namely to ask colleagues.

5 Lessons Learned and Preliminary Recommendations for Building an Open AAL Ecosystem

5.1 Current Development Practice and Expected Benefits from an AAL Ecosystem

Programming language

We found that Java was the programming language where most developers reported their highest skills. However knowledge of OSGi (which is written in Java) is rated lower. Attention therefore must be paid to provide sufficient levels of training material and documentation as more advanced technologies than plain Java APIs are introduced in the ecosystem. Eclipse was the integrated development environment (IDE) that most developers reported using, this was expected, as Eclipse is a very popular open source IDE, it is extendable (plug-in model) and as such should be a good choice for providing tool support in an AAL software ecosystem.

Standards

For standards, we found that a majority of the developers found the use of standards important and relevant for their work. An AAL ecosystem therefore will benefit from providing software artefacts that adhere to relevant standards in the domain, such as for example the Continua standard for communication to medical devices (IEEE 11073), or HL7 or similar CEN TC 251 standards for exchanging information with dedicated health systems. An AAL software ecosystem benefits from providing software components that work with these standards, and should as a very least have a flexible design so as to allow the use of relevant standards, facilitating divergent business models, e.g. both CEN and HL7.

Tools and design support and reusable components

From the tools, developers want support in four main areas. First, support will be appreciated in discovering requirements in the domain (understanding the domain), and then to be able to have traceability from these requirements to the applications created,

so as to be sure to implement the right functionality. Second, as the developers use design patterns, relevant and applicable patterns in the domain should be available. Third, tool support is needed to emulate different kinds of sensors, hardware and gateways and manage complex workflow in distributed environments. The domain of AAL is characterised by a mix of divergent user needs, hardware and software. These findings indicate that developers are seeking for guidance in the form of best-practice AAL specific design patterns (domain specific languages), components, and traceability from requirements to designs and components in this fluctuating reality. Finally, they need emulators and testing facilities to be able to test complex distributed applications.

5.2 Possibilities and Obstacles of a Dedicated AAL Software Ecosystem, and Learning Strategies

Possibilities

Developers would like to have access to certain resources from the ecosystem directly from the development tools. The resources they would like to have access to include documentation, tutorials, and functional code examples for the software artifacts available. The most wanted software artifacts were support for different hardware, for reoccurring tasks, and examples of actual applications that have been in use and that are possible to run out of the box and experiment with. Not so many was concerned with the resources being open source. Again we find that the developers are looking for ways to cope with a domain of divergent user needs, software and hardware, and good example applications that has been in use, that are available for use and tailoring are appreciated. The fact that open source is not a requirement many posed, is presumably that other concerns related to quality is more important, as we shall see next.

Obstacles

What we found are best suited to scare away developers from entering the ecosystem is first to not document the resources properly. Second, if there are too many interdependencies in a software artefact, so that it is difficult to start using it, they are simply not going to do it. Finally, if they cannot find the components, they will not be used. These are classical errors the open source projects make, and it is equally true for AAL ecosystems. Resources need to be easy to find, well documented, and decoupled.

Learning

In order to learn to use resources in the ecosystem, developers would prefer to see example use of resources (including demonstrations and walk-throughs) and example code. Scenarios and use cases are necessary to trace to needs in the domain. APIs are another source of information. This is easy to forget, but to be able to use the resources in the ecosystem, these sources of learning are essential, and will contribute to bringing more users into the ecosystem.

6 Conclusion and Future Work

In this paper we have reported the results of a survey that set out to find the key technological success factors that would attract developers to an open AAL software ecosystem. We found that given the nature of the AAL domain, characterised by divergent users, software and hardware, developers are seeking for support in three main areas. First, they want the artefacts available in the ecosystem to support relevant standards in the domain. Second, support for tracing artefacts available in the ecosystem to requirements (domain-fit) is needed. Third, they want support for developing, testing and emulating for complex user-software-hardware workflows in this distributed environment.

The main obstacles that will scare away developers from the ecosystems is lack of documentation of the artefacts in the ecosystem. Second, not enough decoupled components, and finally, lack of proper search features.

In order to be able to learn to use the artefacts, examples, scenarios and API documentation is necessary.

This work gives a synopsis of what a sample of AAL developers rate as technological enablers and barriers for a successful AAL software ecosystem. More work is needed to relate this work to the larger work on theories of software ecosystems, in order to extend the knowledge base. Also, more qualitative studies in the form of e.g. focus groups with developers are needed to find deeper and more concrete insight into the mechanisms that motivates developers are needed to provide even more relevance for the industry.

7 Acknowledgements

This work was partly funded by the EU Seventh Framework Programme, Theme ICT-2009.7.1, ICT & Ageing, Grant agreement no. 247950.

References

[BBA11] Burla, A., Bakalov, I., & Arnaulov, V.,: universAAL Deliverable 3.4 - Tools for Deployment and Management - uStore, 2011 Available at: http://www.universaal.org/index.php?option=com_content&view=article&id=66&Itemid=15

[Bos09] Bosch, J.,: From software product lines to software ecosystems. In Proceedings of the 13th International Software Product Line Conference. SPLC 2009. Pittsburgh, PA, USA: Carnegie Mellon University, s. 111–119.

[FHM12] Farshchian, B. A., Holbø, K., Mikalsen, M., Reitan, J.: AAL Technologies in Rehabilitation – Lessons Learned from a COPD Case Study. In Handbook of Ambient Assisted Living – Technology for Healthcare, Rehabilitation and Well-being. Ambient Intelligence and Smart Environments. IOS Press, 2012.

[Ha11] Hanssen, G. K.: A longitudinal case study of an emerging software ecosystem: Implications for practice and theory. In Journal of Systems and Software, 8699(7465), 2011, s.1–12.

[Ibz11] Ibanez, G., universAAL Deliverable 4.2 – Innovative AAL Applications on universAAL Platform., 2011 Available at: http://www.universaal.org/index.php?option=com_content &view=article&id=66&Itemid=15.

[JFB09] Jansen, S., Finkelstein, A. & Brinkkemper, S.,: A sense of community: A research agenda for software ecosystems. In Proc. 31st Int. Conf. on Software Engineering - Companion Volume, 2009. ICSE, s. 187–190.

[KM10] Kubitschke, L., Müller, S., Gareis, K., Frenzel-Erkert,U., Lull, F.: ICT & Ageing - European Study on Users, Markets and technologies, European Commission, 2010

[SC98] Strauss, A.C. & Corbin, J.M.,: Basics of Qualitative Research: Second Edition: Techniques and Procedures for Developing Grounded Theory 2nd utg., Sage Publications, Inc, 1998.

[St11] Stav, E., Walderhaug, S., Mikalsen, M., Hanke, S., Benc, I.,: Development and evaluation of SOA-based AAL services in real-life environments: A case study and lessons learned. In International Journal of Medical Informatics, 2011.

[SW11] Stav, E., & Walderhaug, S.,: universAAL Deliverable 3.1 A - Tools for design and development, 2011, Available at: http://www.universaal.org/ index.php?option=com_content&view=article&id=66&Itemid=15.

[Taz11] Tazari, S.,: UniversAAL Deliverable 2.1.A - universAAL Execution Environment Installation packages and hardware abstraction layer., 2012 Available at: http:// www.universaal.org/index.php?option=com_content&view=article&id=66&Itemid=15.

[VDB10] Van Den Broek, G, Cavallo, F. & Wehrmann, C.: AALIANCE Ambient Assisted Living Roadmap, 2010.

Probabilistic Event Processing for Situational Awareness

Davy Preuveneers[1], Andreas D. Landmark[2], Leendert W.M. Wienhofen[2]
[1]IBBT-DistriNet, KU Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium
[2]Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU), Sem Sælandsv 7-9, NO-7489 Trondheim, Norway

davy.preuveneers@cs.kuleuven.be
andreala@idi.ntnu.no
leendert.wienhofen@idi.ntnu.no

Abstract: Over the last century the continuous innovation of technology, coupled with a steady increase in the size of the healthcare organizations, has created a need for information systems supporting healthcare professionals with their daily tasks and decisions. Modern hospitals are full of technology producing electronic records of events and activities, with the opportunity of these events culminating in a wealth of information that these semi-autonomous experts can tap into to improve situational awareness, facilitate coordination and take better informed decisions. However, processing these footprints, contextualizing and inferring over them presents several interesting challenges to the current state of Complex Event Processing methods. This article looks at challenges presented by an information system for perioperative process support and how contextualization and adequate tool support can provide the essential backdrop for meaningful inference.

1 Background

Clinical work processes can resemble non-deterministic processes, at least when it comes to the core activities that determine clinical work. As problem solving activities, in which several actors participate, each from their own domain of knowledge, creating detailed models of clinical processes can be difficult. The exact flow of control (the trajectory of actions performed) is determined by the problem to be solved (e.g. the health issue of the patient) and the decisions that individual actors make, based on their assessment of this problem and their knowledge. This means that different problem solving activities can interfere and due to matters of urgency or lack of resources, the ordering of the activities will vary given the local circumstances.

The overall objective of COSTT [1] is to evaluate and design novel ICT support for a selection of clinical processes. Our belief is that efficient support can be achieved not through explicit control of the flow of work, but rather by providing all actors involved an easy accessible, comprehensible overview of the progress of a process and its current status. By making the process transparent to all those involved, the actors can coordinate their own

[1] http://www.costt.no/

work. So, coordination is facilitated but not dictated, which is often favorable in complex organizations [GM01].

The core concept in our visualization of progress and current status of clinical processes is the patient trajectory. We define a patient trajectory as a timeline-oriented representation of what actually has occurred and will happen with the patient during encounters with clinicians. Through inspecting a patient trajectory, a clinician can see how far the plan concerning a patient has progressed, and also whether there have been deviations from the original plan. Based on this information he can decide if he needs to make any adjustments to his own activities.

However, as shown by others [RRvdGK09, HTS06], stringent workflow systems and attempts at creating detailed guidelines or exhaustive process maps for healthcare often break down because of the apparent non-deterministic nature of healthcare processes. Additional work has also uncovered that some process variation can be beneficial for the treatment process and as such, should be supported rather than discouraged by information systems facilitating such work [BFS11].

Capturing these subtleties in software or tool support has shown to be difficult. Several strategies have been explored, often resulting in issues such as increasing the level of abstraction until the semantic value of the end result is diluted, or the numbers of disjoint but similar processes grows almost exponentially to cover all these eventualities. However, attempting to map any exception to any process of non-trivial size often descends into a chaotic representation of reality. Additionally the prospect of mapping out *the unknown* a priori, is daunting at best.

The goal of this paper is to address these challenges by introducing and motivating the concept of situational awareness through event-based information systems. The paper is organized as follows. In Section 2 we introduce the healthcare professionals as a social community with a strong desire to improve self-coordination and cooperation. We show the value of probabilistic event processing in Section 3, and present modeling and tool support in Section 4 demonstrating the importance of this concept. We end with conclusions and further work in Section 5.

2 Motivation: the Professional Healthcare Community

In 'social communities', being 'social' is the driving force behind most of the available platforms, and time is in general of less critical importance, nor is facilitating cooperation. In our domain, time and cooperation are key factors for achieving 'a good day' and hence providing good care. Keeping track of colleagues and what they are doing and how these activities might impact your own work is a difficult task. Operations and associated resource planning can only to a limited degree be captured in an explicit workflow. Due to the rapid changes in the domain (acute patients causing changing priorities and so on), continuous replanning is necessary. The overall activities such as start and finish are quite easy to conceptualize, though for coordination purposes it is also very relevant to know what happens between these markers. Knowing when an operation is expected to end en-

ables for example a coordinator to anticipate, prepare and reschedule following patients as needed.

Not unlike in less critical 'social communities', location and situation awareness can have positive and negative effects. In general people will rarely trust one system completely, yet better overview, provides a better starting point for self-coordination. Still, the 'terrain takes precedence over the map' when reality and plan diverge.

One of the fundamental assumptions in the COSTT-project was that given better transparency of the processes - or *situational awareness*, users of the system would be able to better self-coordinate. Situational awareness is a term used to describe at which level a person has perceived the current situation. In increasing levels of awareness [End95], the first level would be to perceive the current situation. The second level would be to not only perceive the elements in your immediacy, but also to comprehend the meaning of the events occurring. When the highest level of situational awareness is achieved, one is also able to project how the current situation might evolve based on comprehension of the current as well as knowledge about how the current situation *usually* evolves.

Using situational awareness as a concept to describe the comprehension of an environment and its impact on your own goals and objectives, it has also been shown that lack of or inadequate awareness is a contributory factor to 'human errors'. As such, building computer support to help increase and support situational awareness is an obvious extension to the already existing technology for distributing and increasing situational awareness .

3 Situational awareness with events

Situational awareness is a field of study concerned with perception of the environment critical to decision-makers in complex, dynamic areas such as aviation, military operations, and healthcare. The perception of environmental elements with respect to time and/or space, the comprehension of their meaning, and the projection of their status after some variable has changed, such as time. Situational awareness involves being aware of what is happening in the vicinity to understand how information, events, and one's own actions will impact goals and objectives.

Trying to create a software support system for such a process then entails support for aggregating and combining heterogeneous knowledge and events from a variety of sources.

3.1 Events and complex event patterns

In general it is meaningful to distinguish between (at least) two levels of event richness: Simple events and complex events. Simple events are single events that carry slivers of meaning in themselves, without much room for decomposition. Examples of simple events would be stock order placements, atomic bank account transactions, or stock trades (buy/sell-orders being matched). This is in contrast to complex events, which summarize,

represent, or denote a set of single events which combined would denote a ‘pattern of events’ of a specific process. An oft-cited example is the 1929 stock market crash [LS08]. Complex events are formally processes rather than events, but within the domain are often treated as events and subject to the same processing and rules as for simple events.

3.2 Contextualized semantics of non-deterministic event patterns

Tackling the challenge of the non-deterministic nature of healthcare processes is instrumental to realizing a system which can cope not only with the majority of regular cases - but also recognize the minority of cases with deviations in event value. From a practical point of view, the value of any support system is at its highest when it helps support the difficult minority rather than the more streamlined majority.

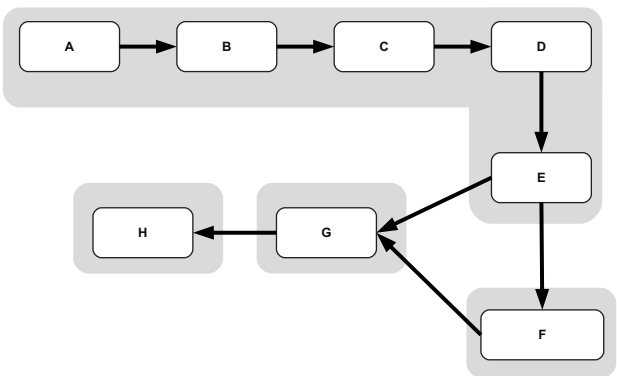


Figure 1: Example patient flow

Figure 1 shows a typical patient flow through several diagnostic activities. In a perfect world, these steps are taken in a pre-determined chronological order. However, in real life deviations to the regular patient flow order are possible. The allowed deviations are driven by dependencies in the information flow dependencies between the healthcare professionals (see Figure 2).

In the following sections, we will refer to these steps as the *situations* a patient can be in. Furthermore, each of these steps is characterized by events, which motivates our approach of situational awareness with events.

One of the main challenges in the perioperative domain is that the systems from which we harvest events were not designed to be part of a larger system and to a large degree live a life of their own. The COSTT-project assumes a bottom-up collection of events wherein we combine physical sensor sources such as indoor positioning systems with events culled from information systems such as planning- and recordkeeping-systems. While this allows for a broad-spectrum of information about the processes and activities in the hospital, it also pushes the issues of temporality, vagueness and uncertainty to the forefront.

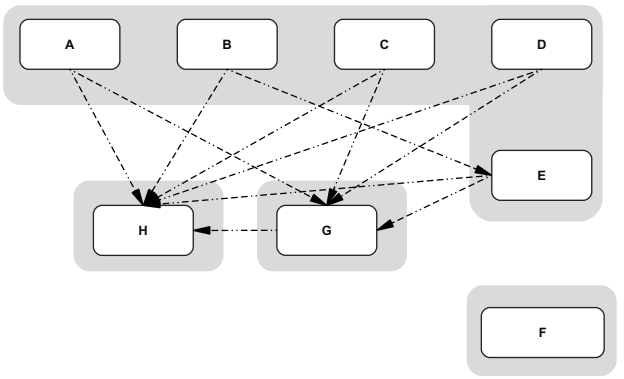


Figure 2: Information flow dependencies

Using sources that rely on manual input as the primary means of gathering information (such as most clinical recordkeeping systems) also means that the variation in the currency of information is, at times, extreme. On the other hand, physical sensors have predictable currency, but inherent shortcomings in sensitivity and specificity.

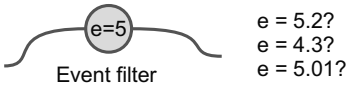


Figure 3: Acceptable deviations in event values

Filtering event values based on medical data can also prove difficult given that the test of clinical significance is often given in ranges rather than specific values. What is acceptable for one patient might be unacceptable to the next patient, even given the same diagnosis. This makes simple filtering as shown in Figure 3 difficult without a rich context for interpretation, which in this case could range from baseline values for this particular patient to encompassing significant parts of the patients medical history. Additionally, certain values have different ranges depending on how the reading was obtained (e.g. body temperature taken rectally, orally or axillary).

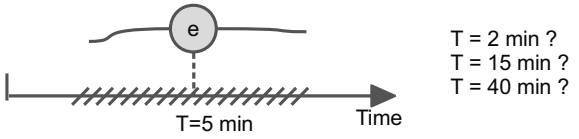


Figure 4: Acceptable deviations in event timestamps

Spatial and temporal relevance (see Figure 4) is just one aspect of the contextual relevance of an event.

Healthcare is rife with examples of procedures where the number of stages depends on the

results of prior steps. Education and experience can also influence whether or not certain stages are skipped or additional ones are performed. The effect from an information system point of view is that some events in a pattern become optional.

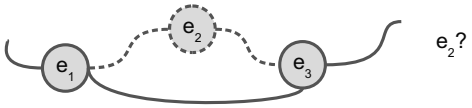


Figure 5: Optional or undetectable events

In terms of defining event patterns, one needs to define which events are optional - see Figure 5 for an illustration - where the non-occurrence of such an event will not make the pattern matching wrongly refute the correct pattern.

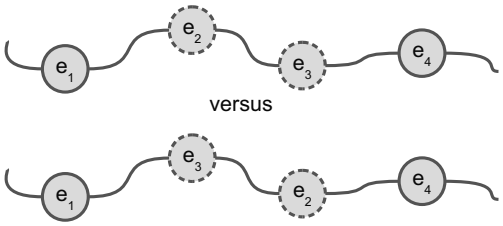


Figure 6: Acceptable deviations in event orderings

The order in which events occur often follows a predefined business logic or workflow path, which could be seen as an event pattern. The order in which events occur can differ based on the workflow path chosen. The ordering of events in the patient workflow may change due to resource constraints or interference with other patients. For example, whereas the logical consequence of events would be e_1, e_2, e_3, e_4, the order of events e_2 and e_3 for a particular patient might be altered, as depicted in Figure 6, if for example there is currently no free slot in the receiving department.

4 Situation Studio: Tool support for event-driven activity recognition

Based on one of the cases in the COSTT project, a pre-operation examination day (for details we refer to [WPL+11]), we designed a flexible event-based workflow system, called the 'Situation Studio'. It is used to model different situations, the probability of their occurrences, the variety of observable and non-observable events in each situation, and the possible partial ordering between situations.

Figure 7 illustrates the graphical user interface of the tool. It shows a concrete instance of the example patient trajectory as shown in Figure 1.

Figure 7: The Situation Studio

4.1 Characterizing situations through events

Each of the situations is characterized by events. For example, the *Blood sampling* situation is characterized by the following observable and mandatory system events:

- `access_lab_system`
- `generate_bar_code`
- `dispatch_blood_sample`

The order of events is typically as listed above. The *Blood sampling* situation is bounded in time by two non-observable events:

- `blood_sampling_started`
- `blood_sampling_finished`

These implicit events mark the beginning and the end of a situation. They can only be inferred from the occurrence of other observed or inferred events. Unless their occurrence matches one-on-one with an observable event, the exact timing of their occurrence is usually uncertain.

Non-observable events are triggered with *rule sets*, i.e. a series of *if-then-else* rules. For example, if the system detects a system event in the *Cardiology outpatient assessment*, we can infer that the previous *Blood sampling* situation has ended and we can trigger the `blood_sampling_finished` event.

With *predicates* we define when a situation can possibly emerge through declaring which (combination of) events should have already occurred, and which ones must not have occurred:

- `patient_id` (this event must have occurred)

- `!blood_sampling_finished` (this event must not have occurred)

For this scenario, when the patient arrives (i.e. he is in the *Patient arrival* situation), the system generates a `patient_id` event with all the personal details of the patient. As long as the patient is not registered, we are sure that the patient cannot be in the *Blood sampling* situation. By also specifying the `!blood_sampling_finished` predicate, we declare that a patient will never go back to this situation if he has completed this step before.

4.2 Probabilistic event processing: a pragmatic approach

With each situation, we associate a probability with each of its events to ascertain the possibility that the patient is still in this situation. For example, the *Cardiology outpatient assessment* situation is characterized by the following observable events:

- `access_epr`: the cardiologist opens the electronic patient record (EPR)
- `dicate_result`: the cardiologist dictates the results of the assessment into speech recognition software

However, healthcare specialists have different working habits. Some may only open the EPR while the patient is sitting in front of them, or dictate the results while the patient is still present, while other ones open all the patient files in the morning or dictate the results after the patient has left. Hence, the occurrence of a particular event is not a guarantee that the patient is still at this location. That is why we associate a prior probability of each event in each situation to characterize the possibility that the patient is at this location when this event occurs. These prior probabilities are derived through discussions with the medical stakeholders. For the *Blood sampling* situation this has led to a prior probability of 100% for the `access_lab_system` and `generate_barcode` events, and a prior probability of 70% for the `dispatch_blood_sample` event. This means that the patient is surely at this location when either of the two first events is recognized. However, there is a slight chance that the patient has already left when the last event is triggered.

Ideally, we would have used proven probabilistic reasoning techniques like Bayes' probability theory, Zadeh's fuzzy logic or Dempster-Shafer's evidence theory. We investigated each of these techniques but none of them turned out suitable due to pragmatic reasons, such as the maintenance of the knowledge for non-technical experts. With Bayes' theorem, we can compute the probability for a situation S given the events E knowing the probability of the events given the situation.

$$P(S|E) = P(S \cap E)/P(E) = P(E|S) * P(S)/P(E)$$

However, each situation is usually characterized by a set of events:

$$P(S|E_1, E_2, E_3, \ldots) = P(E_1, E_2, E_3, \ldots|S) * P(S)/P(E_1, E_2, E_3, \ldots)$$

This means that for any set of events we need to know their probability in every situation, and this is guess work without a proper data set from which we can obtain these probabilities.

Zadeh's fuzzy logic has the advantage that it allows you to express domain knowledge with linguistic terms rather than with crisp values. However, various arbitrary choices have to be made, such as the shape of each fuzzy variable (triangle, trapezoide, bell, ...), the modeling of fuzzy sets and rules, as well as the defuzzification into crisp values. Figure 8

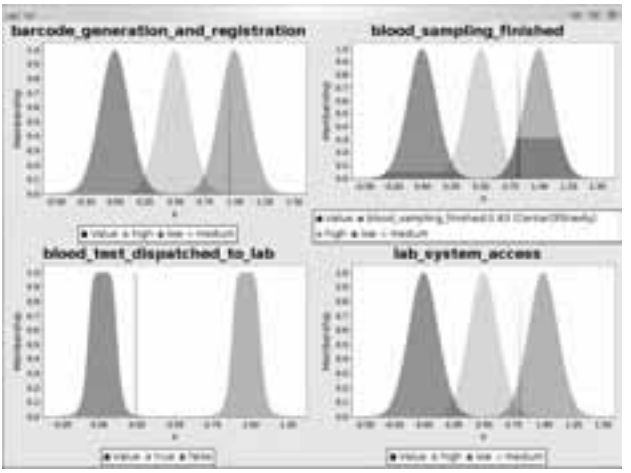


Figure 8: Fuzzy Logic for the Blood Sampling situation

illustrates this concern for inferring the `blood_sampling_finished` event based on the occurrences of the other observable events, based on fuzzy rules like the following:

IF (blood_test_dispatched_to_lab **IS** false) **THEN** blood_sampling_finished **IS** low;

IF (lab_system_access **IS** medium) **AND** (blood_test_dispatched_to_lab **IS NOT** true)
THEN blood˙sampling˙finished **IS** low;

...

The evidence theory from Dempster-Share is a generalization of Bayes based on belief and plausibility, but without going into details, experiments with Dempster's combination rule of evidence have shown that it can sometimes lead to counter-intuitive results. Zadeh himself used the following example to illustrate this concern:

```
Doctor A:      99% brain tumor,     1% meningitis
Doctor B:      99% concussion,      1% meningitis
Dempster's combination rule:     100% meningitis
```

Obviously, this result is very counter-intuitive. Instead, we pursued a more pragmatic approach. Remember that situation *X* means that the patient is at location *X*. Various events pertain to a particular situation (e.g. `access_epr`, `change_ris`, `dictate_result`,

...). Due to the fact that events related to the situation can actually take place before, during or after events, we used prior probabilities to model these uncertainties:

```
P(access_ris | Radiology examination)      = 1.0
P(access_epr | Pulmonary examination)      = 0.6
P(dictate_result | Cardiology examination) = 0.6
```

If predicates of a situation are false, then then particular situation is impossible (likelihood is 0.0). For example, the *Cardiology assessment* cannot take place if the *Cardiology outpatient assessment* has not finished. If all the predicates are true, we compute the probability of the situation based on probability of the last correlated event, and infer the possibility of all the remaining situations. However, this may lead to some mathematical non-sense. Given the likelihoods of the following possible situations:

```
P(Cardiology outpatient assessment) = 0.7    // report_ready
P(Radiology examination) = 0.5               // access_epr
P(Pulmonary assessment) = 0.5                // access_epr
```

We see that the sum of the probabilities is not 1. The reason for this behavior is that the related events do not occur all at the same time. If P(X) would be 1.0, we would be absolutely sure that the patient is at that location. However, if it would be 0.95, then there is room for doubt. To solve this problem, we implemented a function $f(x_i)$ (with x_i being the values above) with the following properties:

- Σ f(x_i) = 1.0
- f(1.0) = 1.0 and f(0.0) = 0.0 *(What is absolutely true or false, remains so)*
- Partial ordering of x_i is the same as partial ordering of f(x_i)

The solution is a value z with $f(x_i) = (x_i)^z$ and z such that $\Sigma(x_i)^z = 1.0$. The value z is not easy to compute directly, so we use an iterative method to find the right value.

```
P(A) = 0.7                              f(P(A)) = 0.494
P(B) = 0.5       with z = 1.980         f(P(B)) = 0.253
P(C) = 0.5                              f(P(C)) = 0.253
```

or

```
P(A) = 0.99                             f(P(A)) = 0.948
P(B) = 0.5       with z = 5.265         f(P(B)) = 0.026
P(C) = 0.5                              f(P(C)) = 0.026
```

The property of the proposed function maintains the weight of the most likely situation while ensuring the transformed values add up to one.

4.3 Qualitative assessment

It is important to realize that the actual value of $(x_i)^z$ is meaningless. However, we used the transformed values in our tool as a baseline for a color coding for the likelihood of a

situation (e.g. red means *impossible*, green means *absolute certainty*, yellow means *possibility*, blue means *prior finished situation*). See Figure 9 for an illustration of a simulated event trace and the color coding of the different situations.

Figure 9: Visualization of a simulated event trace

While testing the Situation Studio, we found that cross-cutting work flows greatly impact the handling of events. For the blood sampling activity for example, for coordination purposes one only needs to know if the sample has been taken and if the patient is done with this activity. However, the outcome of the actual lab results is an input for later activities, but it does not impact the flow of the patient though the day.

We also compared the mathematical output and color coding with the experience of the patient coordinators, and while stepping through the trace of events the likelihood of the outcomes were similar to their expectations.

5 Conclusion

In this paper, we discussed clinical processes as non-deterministic, resembling stochastic processes with healthcare professionals being a social community with a strong desire to improve self-coordination and cooperation. We introduced and motivated the concept of situational awareness through event-based information systems, and presented preliminary tool support to capturing non-deterministic subtleties.

Note that no structured test of the system has been commenced, as one of the most important effects is understanding how putting events in a semi-structured work flow impacts the quality aspect. These effects have been the source for many discussions, which in turn resulted in iterations. This insight is valuable for understanding the effects of compiling events into complex events.

The Situation Studio will be further developed in the frame of the FP7 BUTLER project which investigates the challenges emerging from the recent ICT advances shaping up the

Internet-of-Things (IoT) that will connect sensors, actuators and smart portable devices into an ecosystem reaching an estimated 50 billion devices by 2015-2020. The sheer amount of information and the means to discover and benefit from it will be so vast that humans, no matter how technology-savvy, will not be able to handle it on their own. In BUTLER, the Situation Studio will be used to manage context-aware event processing and pattern recognition techniques to extract meaningful information and transform a stream of events into a representation of human intent for anticipating common user behavior.

6 Acknowledgments

This research is partially funded by the Interuniversity Attraction Poles Programme Belgian State, Belgian Science Policy, by the Research Fund KU Leuven, and partially by the Research Council of Norway (VERDIKT).

References

[BFS11] Berit Brattheim, Arild Faxvaag, and Andreas Seim. Process support for risk mitigation: a case study of variability and resilience in vascular surgery. *BMJ Quality & Safety*, February 2011.

[End95] M.R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):32–64, 1995.

[GM01] S. Glouberman and H. Mintzberg. Managing the care of health and the cure of disease-Part II: Integration. *Health care management review*, 26(1):70, 2001.

[HTS06] Minmin Han, Thomas Thiery, and Xiping Song. Managing exceptions in the medical workflow systems. In *Proceedings of the 28th international conference on Software engineering*, ICSE '06, pages 741–750, New York, NY, USA, 2006. ACM.

[LS08] David C. Luckham and Roy Schulte. Event Processing Glossary - Version 1.1. Online Resource. `http://complexevents.com/2008/08/31/event-processing-glossary-version-11/`, July 2008. Last visited: October 2009.

[RRvdGK09] Hajo A. Reijers, Nick Russell, Simone van der Geer, and Gertruud A. M. Krekels. Workflow for Healthcare: A Methodology for Realizing Flexible Medical Treatment Processes. In *Business Process Management Workshops*, pages 593–604, 2009.

[WPL+11] Leendert W. M. Wienhofen, Davy Preuveneers, Andreas D. Landmark, Pieter J. Toussaint, and Yolande Berbers. A notion of event quality for contextualized planning and decision support systems. In *Proceedings of the 7th international and interdisciplinary conference on Modeling and using context*, CONTEXT'11, pages 307–320, Berlin, Heidelberg, 2011. Springer-Verlag.

Simulation of Rescue Force Communities in Mass Casualty Incident Situations

[1]Volkmar Schau, [2]Kathrin Kirchner, [1]Steffen Späthe, [1]Sebastian Scharf,
[1]Stefan Hellfritzsch, [3]Gerald Eichler, [4]Christian Erfurth, [1]Wilhelm Rossak, [5]Jens Reichel

[1]Friedrich-Schiller-University Jena, Department of Computer Science
Ernst-Abbe-Platz 2-4, D-07740 Jena, Germany
{volkmar.schau|steffen.spaethe|sebastian.scharf|stefan.hellfritzsch|wilhelm.rossak}
@uni-jena.de

[2] Friedrich-Schiller-University Jena, Department of Information Systems
Carl-Zeiss-Strasse 3, D-07740 Jena, Germany
kathrin.kirchner@uni-jena.de

[3]Telekom Innovation Laboratories
Deutsche-Telekom-Allee 7, D-64295 Darmstadt, Germany
gerald.eichler@telekom.de

[4]University of Applied Sciences Jena
Carl-Zeiss-Promenade 2, D-07745 Jena, Germany
christian.erfurth@fh-jena.de

[5]University Hospital Jena
Erlanger Allee 101, D-07747 Jena, Germany
Jens.Reichel@med.uni-jena.de

Abstract: Mass Casualty Incident (MCI) situations require high flexibility from the involved rescue forces, including an increased need for efficient communication in sparse mobile ad-hoc networks. Simulations help to understand the performance of information flooding in these situations, to identify critical gaps in the infrastructure and to develop professional models. Based on resources, described as profiles, roles of community members can then be assigned more flexible. The presented approach and the thereupon aligned simulation provide a technological basis for developing a decision support system. This system can – before or during an MCI situation – provide support in mitigating risks. This will be achieved by executing (simulating) and comparing different alternatives to manage the situation at hand.

Keywords: Community simulation, self-organisation, professional model, agent-based communication, MANET.

1 Introduction

In case of a Mass Casualty Incident situation (MCI) emergency rescue services (fire brigades, medical service and police force) have to face several challenges. At the beginning there is often a discrepancy between the number of injured persons and the number of available emergency units at the place of action. So it is essential to coordinate all involved rescue forces and activities. Moreover, it is necessary to identify tasks and task sequences of all involved forces to support inter- and intra-organizational collaboration. In the light of MCI the *SpeedUp*[1] research project activities are focused on an IT framework to support communication and collaboration between mobile rescue forces [Spe12]. Starting with investigations of organizational structures and strategies for courses of action within various rescue forces, SpeedUp addresses the definition of an IT solution which is acceptable and utilizable for different organizational units in complex situations.

To support development, testing and evaluation of the SpeedUp IT solution we need regular MCI-like-situations. Based on the insight, that it is rather impossible to carry out real MCI exercises again and again, we aimed at finding a solution to replay such scenarios in a cost-efficient way. Therefore, we mapped these complex courses of events within inhomogeneous communities by using simulation techniques.

In this paper, we introduce a tool for simulating heterogeneous communities. Furthermore, we discuss our approach to develop a community simulation after collecting information in a prior analysis process. The approach is explained by using a MCI scenario with rescue worker communities.

The paper is structured as follows: Section 2 explains the scenario in more detail. Section 3 discusses the analysis process that is necessary to build an MCI-like-situation. In Section 4, we explain our agent-based discrete event-based simulation tool. Section 5 provides first results, whereas section 6 gives an outlook for future work.

2 Scenario

Typical examples for a MCI are bus, train or plane accidents. Insured people, rescue forces and equipment are spread over a large area. Structured and coordinated handling requires (1) data collection, (2) information forwarding and (3) hierarchical organization. Paper-based processes are reliable but slow and incomplete. Therefore, SpeedUp relies on electronic support for rescue forces using mobile communication devices (nodes). Known major risks of IT-based communication are inaccessibility as well as lost or broken down nodes. Its major strength is the autonomous and redundant replication of information within the entire ad-hoc network, if it is not partitioned.

[1] The work is part of the SpeedUp project which is funded within the German Federal Government's program "Research for Civil Security" (call "Rescue and protection of people") by the Federal Ministry of Education and Research (duration: 1 May 2009 - 30 April 2012).

In order to act and collaborate in an appropriate way in MCI situations, rescue forces have to aggregate several types of information that form an *information building*. This information building is necessary to allow for fast information flow and filtering of relevant information [SKE+10]. Figure 1 provides an overview of the involved technological fields necessary to cover all information needs.

Instead of using paper to distribute the information to others, a mobile, self-organizing data platform is needed that supports the rescue forces [SSE10]. The information that has to be delivered differs for every organization and rescue worker and is task dependent.

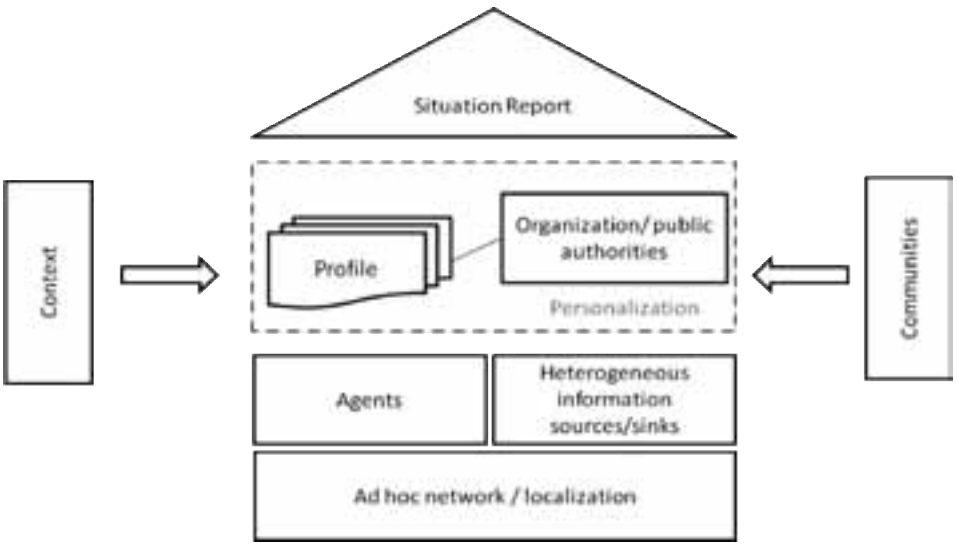


Figure 1: Information Building, adapted from [SKE+10]

At the technical level, ad hoc networks [BFS08] provide the infrastructure for information transmission. Therefore, the localization of rescue workers is important [SEE+11]. Mobile software agents [DCL11, BR05] can be used to integrate heterogeneous data resources and transport information. The share of personalized information for a rescue worker during a MCI in a certain role is restricted by policies [SKE+10] and depends on the context - the involved organization and the tasks to be solved. The context of a rescue worker can be described by his profile. A profile consists of the role the rescue worker plays in a certain mission and associated individual properties, like knowledge, expertise or physical state.

The knowledge about the MCI situation should be shared by the entire community of rescue workers, independent of the organization they belong to. The sharing of experiences helps to improve efficiency in future action situations. Thus, after an MCI situation it is necessary to discuss the event with all involved rescue forces and derive lessons learned. Moreover, the technical as well as the organizational support should be analyzed. For an analytic reflection, it is useful to simulate the entire situation again A simulation gives a cost efficient and a reproducible way to analyze the event. Event

scenarios can be replayed several times with different parameters so that rescue workers can go deeper into concrete situations and can learn from the different perspectives.

Following this general goal we discuss in our paper the technological background of such a simulation tool. As a first step, our tool supports a simulation of prototypical action situations derived from real situations, as well as a what-if-analysis.

3 Simulation modeling approach

For the modeling of a simulation for rescue forces we propose an approach consisting of six phases: (1) data elicitation, (2) data processing, (3) structuring, (4) implementation, (5) simulation and (6) evaluation. The interplay of these 6 phases is shown in figure 2.

The first phase consists of the elicitation and collection of data originating from real action situations or interviews with rescue workers. The data will be processed in a second phase extracting the requirements of the environment and the involved rescue workers. In the third phase, the collected information will be structured for IT processing. In the implementation phase these structured data is transformed for being consistent with the simulation model. Hereby, the emphasis is on the modeling of single community members. The complex course of the action situation is simulated in the fifth phase by enacting the direct interplay of relatively simple entities (community members) and their ongoing interaction with the environment. The evaluation in the last phase helps to analyze the MCI event and may trigger another (simulation) perspective.

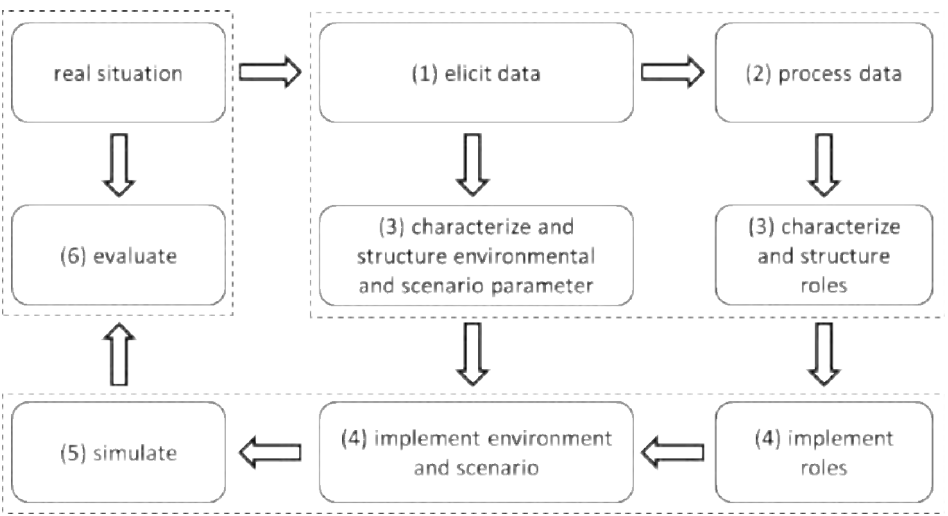


Figure 2: Schematic Procedure for the Development of a Community Simulation

(1) *Data elicitation*: Data collection depends on the given scenario and the available information. In our project, information was elicited from real action situations as well as practice exercises, interviews of rescue force members and from relevant legal regulations [WKM+10, Feu99, Feu06, Feu08, Thu08]. Legal regulations provide rules according to processes, organizational structures and competencies. More concrete

insights come from interviews. Such interviews provide information regarding typical procedures, used strategies, spatial distributions, but also quantitative data about necessary rescue forces, time efforts and problems in concrete situations. Furthermore, exceptions and weaknesses during events can be elicited. The collected data are a basis for future simulation runs and their evaluation.

(2) *Data processing*: The interviews are transcribed, generalized due to several criteria (code, main categories, sub categories, sources, sinks and segments), and extracted into a table [WKM+08]. For analyzing action situations, mind maps were used to look at aspects like location, forces, resources, weather, time and situation. Therefore, information is grouped to detect dependencies and gaps. Moreover, the whole sequence of events is transformed into a time line.

(3) *Structuring*: In every action situation, people are involved in different roles. These different roles are identified using the sequence of events (time line). Thereby, a role abstracts from a concrete person. Prior knowledge, expertise and the state of a person are allocated to a profile. Additionally, commands and activities are extracted from the time line for every role. A role is described by its tasks. Tasks can be divided into subtasks, so that a hierarchical structure is built. From the time line, pre and post conditions are allocated to a task. In table 1, roles are described by a hierarchical decomposition of their tasks to subtasks, combined with suitable pre and post conditions. From the mind map and the time line the necessary data for the environment are extracted.

Task	Action	Communication	Rescue force
Transport <patient> to patient positioning place	*[Patient not transportable]* ensure_transportability <patient> *[Patient transportable]*		
	[<Patient> transportable] organize_orderly <patient> <patient positioning place> *[enough <orderlies> available]*		
	[<Patient> transportable & enough <orderlies> available] give_command <orderly> „ transport <Patient> to <Patient positioning place>“ [<orderly> and <Patient> arrived]	Transfer command, transfer confirmation	paramedic, rescue aid man
escort <patient> to patient positioning place	*[patient can go alone] walk with <Patient> to patient positioning place [arrived]*		

Table 1: Decomposition of roles into tasks

(4) *Implementation*: The description of tasks is realized in the Belief-Desire-Intension (BDI) structure consisting of goals and plans (goal-plan-tree) [HSR12, GPP+99]. In the root of the tree is a goal that can be achieved by the successful execution of one of the subordinated plans (figure 3). Plans may consist of actions or other subgoals. Pre and

post conditions resulting from the structuring phase (3) are expressed by semantic queries to an individual data area (beliefs). Actions in the goal-plan-tree are atomic. They are characterized by their effects on the environment and by time costs.

For more information regarding implementation and simulation we refer to section 4. In the end the evaluation (phase 5) is discussed in section 5.

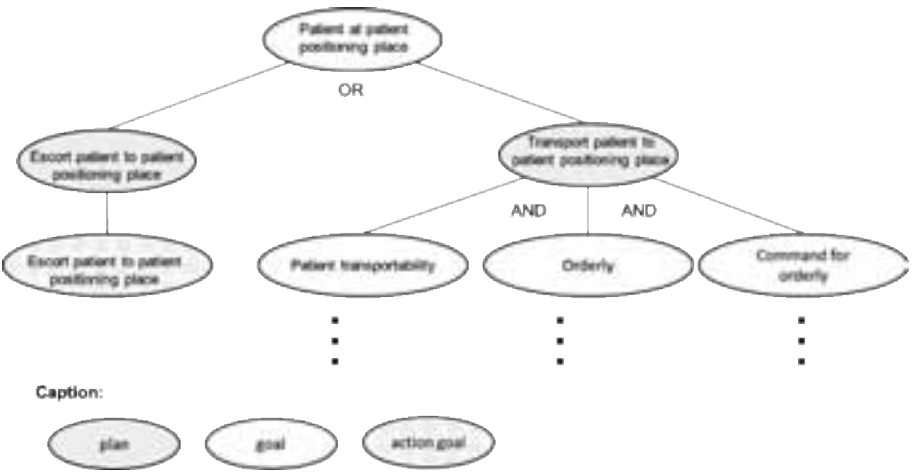


Figure 3: Example of a goal-plan-tree

4 Simulation

Rescue forces, e.g. medical personnel, have a lack formal training in mass causalities. Therefore, their education should be expanded [GJA05]. Simulation modeling is regarded as one of the leading techniques for improving the capabilities in case of incidents [JM06, JML07]. It was used e.g. in training of healthcare personnel [HYH+10] or terror medicine [HM09].

For our simulation, we use an agent-based, discrete event based approach [Gon09]. The simulation of the environment and all rescue workers involved in the scenario is done by agents. The simulated scenario is divided into active and passive entities. Passive entities only react to external influences, while active entities can carry out actions independently and therefore influence their environment. E.g. in our scenario, rescue workers are active entities, while injured people are regarded as passive entities [HSR12].

Figure 4 presents the agents which are necessary for a simulation. The environment agent manages the simulation time and the state of the environment that also includes the passive entities. The starting point for a simulation is a given scenario, which also describes the environment parameters for the scene as well as the number and position of injured people. Furthermore, the scenario includes a number of events that are defined to occur at a certain point in time during the simulation. This event set is managed by the

EnvironmentAgent. It checks in every time step, which events occur at this point in time. Additionally, it calculates how these events influence the state of the environment.

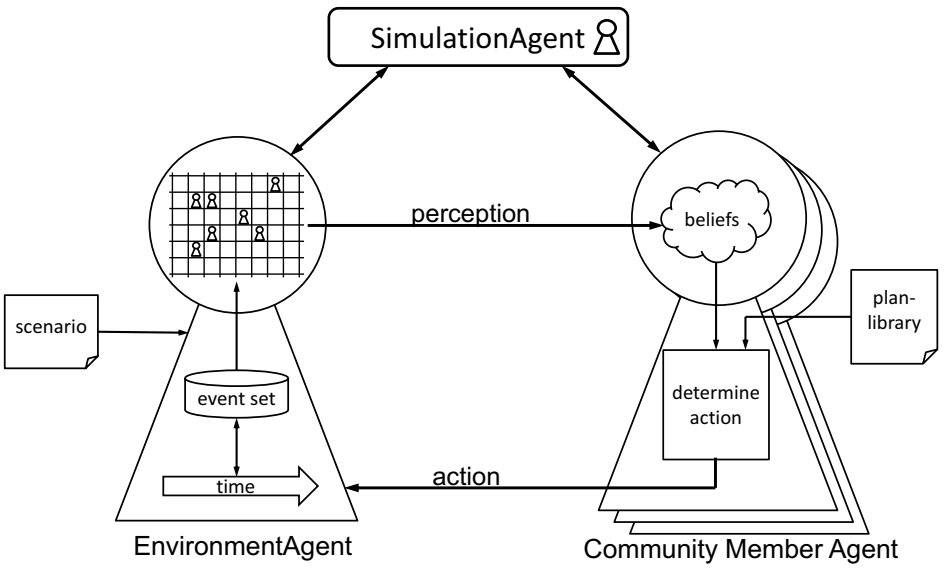


Figure 4: Simulation

An additional task of the EnvironmentAgent is to calculate the perception of every active entity which stands for a rescue worker. The perception is calculated in every time step based on the state of the environment. The results are transferred to an agent that simulates the respective rescue worker. The simulation of a certain rescue worker is thereby implemented by a Belief-Desire-Intention (BDI) agent [GPP+99], called *Community Member Agent* in figure 4. These agents manage their own view on the simulated environment by using their beliefs. The agent reacts to the new perception of the EnvironmentAgent by updating its beliefs, based on the perceptions. Afterwards, the BDI implementation helps to calculate the upcoming action of the rescue worker. This calculation of the next action is based on the current beliefs of the agent as well as on the plan library of the simulated rescue worker.

The processed action is transferred back to the EnvironmentAgent, which calculates the consequences of this action to the environmental state. So in the next time step it is possible to transfer the consequences of the previous action in form of a new perception back to the simulated rescue worker. A simulated time period ends, if all events taking place in this period have been handled, every active entity was sent a perception, and all active entities had the chance to perform their own actions.

For the management of the entire simulation there is another agent called *SimulationAgent*. This agent is able to start and to finish a simulation. Furthermore, it bridges the gap between the simulated system and the underlying agent system ellipsis

[ME11, Sch12]. In the case that more rescue workers are to be involved for a specific scene in a simulation, the SimulationAgent can start more Community Member Agents that can overtake the role of the newly involved rescue workers. In addition, it can link them with the simulated scenario.

5 Results

With our simulation tool, we can map a complex course of events with inhomogeneous communities. By mapping the rescue workers using BDI agents we are able to describe different rescue forces with their specific aims, procedural models and cultures. In a first evaluation, we simulated three different situations: a small, a medium and a large MCI event. We collected the typical information involved in such situations and modeled it.

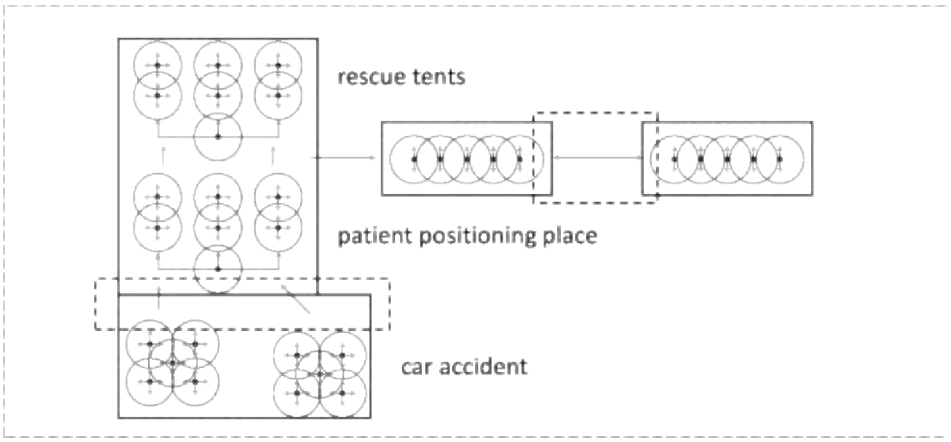


Figure 5: Evaluated small MCI scenario

Figure 5 shows the small MCI scenario that was simulated. Each of the circles defines a rescue force member. The arrows define possible moving directions of a rescue worker. At the bottom, a car accident occurred on the motorway which is the location of accident. Injured people are transported from the accident location to the patient positioning place in the upper left part. From there, patients can be transported in the rescue tents, where medical doctors can treat them. On the right site in figure 5, there are additional medical workers, policemen or fire workers that are waiting for action. All these involved rescue workers have to communicate with one another. If they are near to each other, like in the accident location or the rescue tents, communication is pretty easy using words. In case rescue workers are in different locations there might be interruptions in communication, displayed in figure 5 using squares with dashed lines. In our evaluation we could proof that the simulation approach works very well. For this evaluation we focused on prototypical scenarios with a fairly small number of roles. Therefore, our approach and our current version of the simulation tool need further improvement, refinement, and scaling. Nevertheless, we could show that the current tool implementation forms already a stable basis for a more detailed, professional simulation.

6 Conclusion and Outlook

Our proposed approach and the aligned simulation tool provide the technical basis to develop a simulation based, intelligent decision support system. This system can help to mitigate risks and dangers, as well as to weight different alternatives during preparation and in ongoing MCI situations.

A main objective of our future work will be to develop a complete training and decision support system based on our current, still relatively abstract technical tool. With such a system, real MCI situations will be re-enacted and decisions that were made in the MCI could be analyzed. In this way, experienced rescue workers will be able to reflect the consequences of their actions and simulate possibly better, alternative approaches. Using fictitious scenarios, it will also possible to evaluate new or unusual strategies during MCI events and see whether they are useful in real-life situations.

Bibliography

[BR05] Peter Braun, Wilhelm Rossak: Mobile Agents: Basic Concepts, Mobility Models and the Tracy Toolkit, Morgan Kaufman, 2005.

[BSF08] Remi Badonnel, Radu State, and Olivier Festor. Management of ad-Hoc Networks. In Jan Bergstra and Mark Burgess, editors, Handbook of Network and System Administration, pages 331–360, 2008.

[DCL11] Elton Domnori, Giacomo Cabri, and Letizia Leonardi: Multi-agent Approach for Disaster Management, 2011 International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, pages 311-316.

[Feu99] Feuerwehr-Dienstvorschrift 100: Führung und Leitung im Einsatz: Führungssystem, 1999.

[Feu06] Feuerwehr-Dienstvorschrift 1: Grundtätigkeiten: Lösch- und Hilfeleistungseinsatz, 2006.

[Feu08] Feuerwehr-Dienstvorschrift 3: Einheiten im Lösch- und Hilfeleistungseinsatz, 2008.

[GJA05] Joseph M. Galante, Robert C. Jacoby and John T. Anderson: Are Surgical Residents Prepared for Mass Casualty Incidents? Journal of Surgical Research 132(1), pages 85-91.

[Gon09] Rafael A. Gonzales: Crisis response simulation combining discrete-event and agent-based modeling. In proceedings of the 6th International ISCRAM Conference. Gothenburg, Sweden, 2009.

[GPP+99] Michael Georgeff, Barney Pell, Martha Pollack, Milind Tambe and Michael Wooldridge: The belief-desire-intention model of agency. *Intelligent Agents V: Agents Theories, Architectures, and Languages.* 1999, pages 1-10.

[HM09] Asher Hirshberg and Kenneth L. Mattox: Modeling and Simulation in Terror Medicine. In: Shmuel C. Shapira, Jeffrey S. Hammond and Leonard A. Cole: Essentials of Terror Medicine. 2009, pages 79-94.

[HSR12] Stefan Hellfritzsch, Volkmar Schau and Wilhelm Rossak: Agentenbasierte Simulation und Evolutionäre Optimierung einer Kommunikationsinfrastruktur für einen MANV-Einsatz. Technical Report, to appear.

[HYH+10] LeRoy Heinrichs, Patricia Youngblood, Philip Harter, Laura Kusumoto and Parvati Dev: Training Healthcare Personnel for Mass-Causualty Incidents in a Virtual Emergency Department: VED II. Prehospital and Disaster Medicine 25(5). 2010, pages 424-432.

[JM06] Sanjay Jain and Charles R. MccLean: A Concept Prototype for Integrated Gaming and Simulation for Incident Management. In: Proc. of the 2006 Winter Simulation Conference. 2006, pages 493-500.

[JML07] Sanjay Jain, Charles R. MccLean and Y. Tina Lee: Towards standards for integrated gaming and simulation for incident management. In. Proc. of the 2007 Summer Computer Simulation Conference. 2007, pages 601-608.

[ME11] Multiagentsystem Ellipsis: http://sourceforge.net/projects/masellipsis/ [Online, accessed 15.01.2012]

[Sch12] Volkmar Schau: Ellipsis - Konzeption und Implementierung einer vereinheitlichten und leistungsfähigen Architektur für Multi-Agenten Systeme (MAS) auf Basis frei verfügbarer IT-Infrastrukturkomponenten unter besonderer Berücksichtigung internationaler Standards, Dissertation, Friedrich Schiller University Jena, January 2012.

[SEE+11] Volkmar Schau, Christian Erfurth, Gerald Eichler, Steffen Späthe and Wilhelm Rossak: Geolocated Communication Support in Rescue Management, Proceedings of the 8th International ISCRAM Conference. Lisbon, Portugal, 2011.

[SHS+12] Volkmar Schau, Stefan Hellfritzsch, Sebastian Scharf, Gerald Eichler, Christian Erfurth, Wilhelm Rossak: Simulation of wireless, self-organizing and agent-based dynamic communication scenarios. In proceedings of the 9th International ISCRAM Conference. Vancouver, Canada, April 2012.

[SKE+10] Volkmar Schau, Kathrin Kirchner, Christian Erfurth and Gerald Eichler: Ad-hoc Community Composition of Rescue Forces in Action Situations. In proceedings of the 10th I2CS, LNI P-165, pages 41-52, June 2010.

[Spe12] SpeedUp Consortium: SpeedUp Homepage. http://www.speedup-projekt.de/ [Online, accessed 15.01.2012]

[SSE10] Volkmar Schau, Steffen Späthe, Gerald Eichler: SpeedUp: Mobile und selbstorganisierende Kommunikations- und Datenplattform für komplexe Großlagen, Tagungsband zum 7. GI/KuVS-Fachgespräch „Ortsbezogene Anwendungen und Dienste“, Logos-Verlag, Berlin, 2010.

[Thu08] Thüringer Gesetz über den Brandschutz, die Allgemeine Hilfe und den Katastrophenschutz, Thüringer Brand- und Katastrophenschutzgesetz, ThürBKG, 2008.

[WKM+10] Fabian Wucholt, Andreas Kleinschmidt, Mareike Mähler, Sandy Schuman, Yeliz Yildirim-Krannig, Stefan Strohschneider: Memo VIII, Informationsinhalte beim MAnV/GSL- Feuerwehr-, Interkulturell Wirtschaftskommunikation, SpeedUp Project, 2010.

SESSION 4

Infrastructure and Security

Structured Peer-to-Peer Networks through Distributed Nonmetric Multidimensional Scaling

Hauke Coltzau
Fernuniversität in Hagen
hauke.coltzau@fernuni-hagen.de

Abstract: Multidimensional Scaling (MDS) is a technique for dimensionality reduction widely used in the field of data-mining. For P2P data networks, it holds the potential to provide an intuitive way to browse and explore heterogeneous data distributed over the system. Current MDS approaches are dependent on a centralized instance (controller) and can therefore not be used in decentralized environments. In this article, a distributed algorithm for multidimensional scaling is discussed that does not need any such centralized control.

1 Introduction

This article aims to show that a structured Peer-to-Peer (P2P) network can be built by arranging peers in a 2-dimensional space in such a way that peers storing similar content are located in the same area. The structure of the resulting P2P system provides intuitive ways of navigation and browsing through it's content. Users can search and explore such a network without exactly having to know beforehand, what the result of their search will be. For additional support of orientation and navigation, dynamic maps of such a system could be generated in a distributed manner (see [CU12]), and thus even give an overview of the type of objects stored in the system..

The relations between the objects stored in a peer-to-peer system usually can only be described in a high-dimensional space. Therefore, some distributed dimensionality reduction is needed to obtain the required 2-dimensional projection. In data mining, the term *Multidimensional Scaling* (MDS) denotes a set of such techniques. But these techniques require an omniscient (global) view on all objects, for which the dimensionality reduction shall be done. Hence, to achieve the goal of building a Peer-to-Peer system based on MDS, a distributed MDS technique is needed, which does neither need any centralized instances nor a global view on all participating peers.

The structure of this article therefore is as follows: In section 2, an overview of multidimensional scaling is given, including existing approaches for parallel and distributed MDS. In section 3, an algorithm is proposed to perform distributed MDS and build a structured P2P-system on the resulting configuration. Section 4 shows, how the algorithm performs in an example network. The results gathered from this and other simulations are described in the final section.

2 Multidimensional Scaling

MDS aims to visualize high-dimensional data in a low-dimensional (2D or 3D) target space while preserving (dis-)similarities between the data points. The idea of multidimensional scaling goes back to Torgerson [Tor52] and has been generalized by Kruskal [Kru64a][Kru64b]. All current adaptions and also large-scale approaches like [CPH06] and [BBKY06] are still strongly based on these works, so that a description of the basic idea behind them seems appropriate in the context of this article.

2.1 General Approach for MDS

The input of any multidimensional scaling algorithm is a set N of $n = |N|$ objects, from which the pairwise dissimilarities $\delta_{i,j}$ for (almost[1]) all objects $i, j \in N$ are known. Often, these dissimilarities are given in form of a dissimilarity matrix Δ:

$$\Delta := \begin{pmatrix} \delta_{1,1} & \delta_{1,1} & \dots & \delta_{1,1} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,n} \\ \vdots & \vdots & & \vdots \\ \delta_{n,1} & \delta_{n,2} & \dots & \delta_{n,n} \end{pmatrix} \tag{1}$$

Approaches for executing an MDS can roughly be categorized as *metric* or *non-metric*. Metric approaches expect the dissimilarity values to be in a (ratio-scaled) metric, such that a scalar distance between any to objects can be specified. Non-metric approaches in contrast only expect a rank-order of dissimilarities and any rank-order preserving function to get a scalar representation of each dissimilarity. Hence, non-metric approaches can be seen as generalization of metric approaches.

A *configuration* X constitutes a projection of N into s-dimensional space (where s is very often chosen to be 2 or 3), i.e. each object is assigned to an s-dimensional vector x_i.

$$X = (x_1, x_2, \dots, x_n), x_i \in \mathbb{R}^s \tag{2}$$

The *distance* $d_{i,j}$ between two vectors x_i and x_j in s-dimensional space is commonly measured in euclidian metric, although in some cases, other Minkowski r-metrics as e.g. the Manhattan metric are used.

Especially for non-metric MDS, dissimilarities have to be transformed into values that the distances $d_{i,j}$ in target space can be compared to. These values are called *disparities* $\hat{\delta}_{i,j}$, and can be interpreted as target values for the distances in low-dimensional space. How the disparities are calculated from the dissimilarities, depends both on the goal of the MDS as well as the type of input data. In section 2.2, an example approach for finding the

[1] Kruskal considers missing dissimilarity values as weighted with 0.0, and simply proposes to omit such values in all calculations.

disparities in non-metric MDS is given. In metric MDS, the disparities often are identical to the dissimilarities (i.e. $\hat{d}_{i,j} = \delta_{i,j}$).

To evaluate the fitness of a configuration X as a function of the disparities, a cost function (referred to as *stress*) is used. This function iterates over all object pairs $i, j \in N$ and compares their disparity $\hat{d}_{i,j}$ to the distance $d_{i,j} = |x_i - x_j|$ of their assigned vectors x_i and x_j in X. The better the low-dimensional projection reflects the dissimilarities in high-dimensional space, the lower is the stress and the better the chosen configuration represents the original data. A commonly used stress function is given in section 2.2.

The core task of MDS now is to find a configuration X_{min}, which minimizes the stress. In metric MDS, the minimum stress is achieved, when all distances in target space are as close as possible to the disparities (and therefore to the dissimilarities). In non-metric MDS, the stress is minimized, when the rank order of the distances in target space is equal to the rank order of the disparities (and therefore the dissimilarities).

Due to the complexity of the stress function, the minimization can usually not be done in an analytical manner. Instead, usually the method of steepest decent (i.e. gradient descent), or the approach of majorizing the stress function with a less complex function (*Scaling by majorizing a complicated function*, SMACOF [dL77]) are used, the latter being only applicable in metric MDS.

Usually, when an optimal configuration shall be found, the MDS is executed several times (up to 100) with different random initial configurations to avoid local minima. The one configuration with the lowest stress value is assumed to be the global optimum.

2.2 Kruskal's Approach for Nonmetric Multidimensional Scaling

Since almost all current approaches for MDS go back to the work of Torgeson and Kruskal, and since Kruskal's non-metric MDS can be seen as generalization of metric MDS, a brief description of Kruskal's approach is given. The main properties of the approach are the following:

1. Disparities are calculated in such a way, that the rank order of distances in target space is as close as possible to the rank order of dissimilarities. This is achieved using monotonic regression.

2. A steepest decent is used to approximate the configuration step by step. To do so, at each iteration a gradient is calculated, which is applied to the configuration. The iteration stops, when the first derivative of the gradient is zero, i.e. when a minimum is reached.

2.2.1 Disparity Calculation

A good solution (configuration) of the non-metric MDS is found, when the configuration preserves the rank-order of the dissimilarities, i.e:

$$\delta_{i,j} < \delta_{j,k} < \delta_{k,l} < \ldots \iff d_{i,j} < d_{j,k} < d_{k,l} < \ldots \tag{3}$$

In other words: When the dissimilarities $\delta_{i,j}$ are sorted in ascending order and the distances $d_{i,j}$ are then plotted over the dissimilarities, the resulting curve should be monotonically ascending (see Fig. 1). If this is not the case, then for each pair $(d_{i,j}, \delta_{i,j})$ two cases, in which changes are necessary, can be distinguished

- The distance of the succeeding pair is lower than the current pair's distance. The current pair's distance must be reduced, and the succeeding pair's distance must be increased.
- The distance of the preceding pair is higher than the current pair's distance. The current pair's distance must be increased, and the succeeding pair's distance must be decreased.

In both cases, the target value (i.e. the disparity) for the two distances, which are not in order, is calculated as the average value of the two distances. If more than two succeeding distances do not fulfill the monotonicity constraint, the average over all these distances is used.

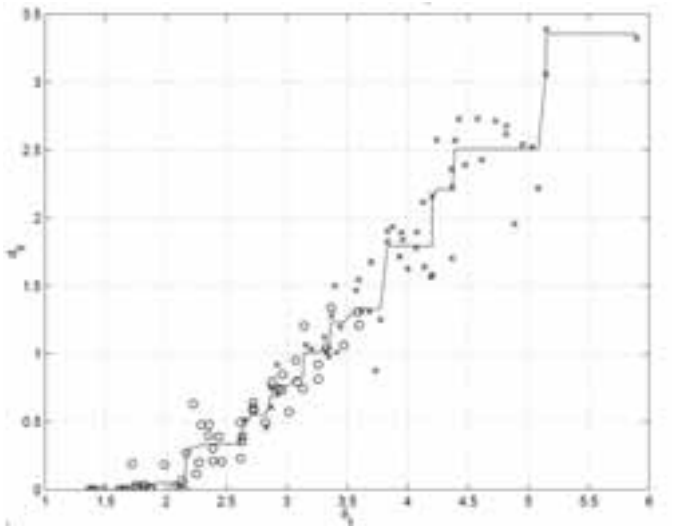


Figure 1: Example scatterplot of distances over dissimilarities (source: [DM00])

With these disparities given, the stress of the current configuration can be calculated and be used to determine the gradient necessary to perform the steepest descent.

2.2.2 Gradient for Steepest Descent

The raw stress over all distances $d_{i,j}$ can be calculated as the sum of the squared deviations of each distance from the according disparity:

$$S' = \sum_{j<i} (d_{i,j} - \hat{d}_{i,j})^2 \tag{4}$$

This value can be normalized with the squared sum of all distances:

$$T' = \sum_{j<i} d_{i,j}^2 \tag{5}$$

such that the normalized and square-rooted stress (see also Eqn. 3) is given by

$$S = \sqrt{\frac{\sum_{j<i}(d_{i,j} - \hat{d}_{i,j})^2}{\sum_{j<i} d_{i,j}^2}} = \sqrt{\frac{S'}{T'}} \tag{6}$$

The gradient is determined using the partial derivatives $\partial d / \partial x_{i,j}$ of the stress function:

$$g_i = \frac{S}{d_{i,j}} \cdot \sum_{(i,j),j<i} \frac{1}{d_{i,j}} \cdot \left[\frac{d_{i,j} - \hat{d}_{i,j}}{S'} - \frac{d_{i,j}}{T'} \right] \cdot \begin{pmatrix} x_{i1} - x_{j1} \\ x_{i1} - x_{j2} \\ \vdots \\ x_{is} - x_{js} \end{pmatrix} \tag{7}$$

This gradient is weighted with a smoothing factor (usually 0.2) and then added to the current position of i in target space. After applying the appropriate gradients to all positions in the current configuration, the resulting stress of the new configuration is determined. When a minimum is reached, the iteration stops and the current configuration is returned as result of the non-metric MDS.

2.3 Parallel and Distributed Approaches

Some approaches exist to execute MDS in a parallel manner. For example, [PD10] proposes a parallel MDS based on a heuristic from particle dynamics. To achieve good parallelisation, the participating nodes (processors) are structured in a stair-like structure as shown in Fig. 2, each node communicating only with few other nodes.

Still, since the arrangement of nodes as well as the collection of results needs to be done by a centralized instance, this approach cannot be applied to Peer-to-Peer systems. This also applies to other parallel implementations as e.g. [CHK+09] and [BBKY06].

To the best knowledge of the authors, the only truly distributed approach for MDS is given in [CPH06]. In this approach, the position of sensors in a sensor network of arbitrary size

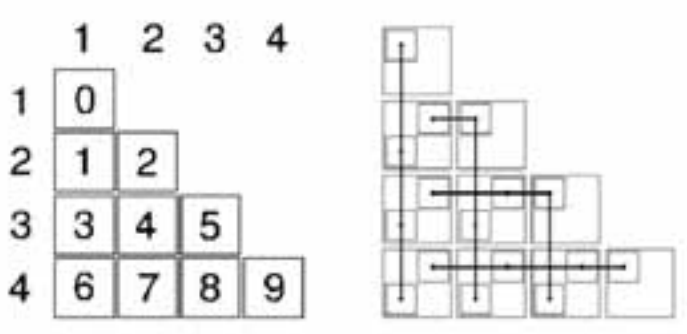


Figure 2: Node arrangement and information flow in [PD10]

is approximated using distance measurements of each single sensor to other sensors in its reach. All other (unreachable) sensors are not considered. The authors show that, if some sensors with fixed (known) positions exist, all other nodes can estimate their one absolute location with good quality. Without having previously fixed positions, still the relative position of each node is found. But, when looking more detailed into the proposed solution, it becomes clear that actually, no real dimensionality reduction is done, because the original data gathered from the sensor can itself already be described in a low-dimension space, even if some additional dimensions may occur due to measurement errors. In their article, the authors assume that their approach could be applied to data of higher dimensions, but provide neither proof nor example for that claim.

3 Algorithm for Distributed Nonmetric MDS

As shown in the previous section, executing an MDS requires global knowledge, i.e. the distances and dissimilarities of all objects need to be known to determine the gradient for each single position. In distributed systems, especially in Peer-to-Peer systems, no single participant has such an omniscient view. It is therefore necessary to run the MDS with only local knowledge.

The basic properties of the proposed approach are:

1. Each peer represents a single object in the high-dimensional space and is able to calculate the dissimilarities to other objects (peers). Since a peer is only a logical entity, a physical peer can still represent several objects.

2. Each (logical) peer is assigned to a 2-D coordinate, representing its position in the current configuration.

3. A peer can only communicate with (and therefore calculate the dissimilarity to) its neighbors. The neighbourhood of each peer is determined using a Voronoi tessella-

tion based on the 2D-coordinates. All peers sharing at least a single Voronoi edge with another peer are set to be a direct neighbour of it. The tessellation is updated locally, whenever a peer's position is changed or a new peer enters the network.

A new Peer P_n is added to the network by successively approximating its position x_P, until an optimum value has been found, at which the peer can be integrated into the network. For the iterative position updates, only the 2D-positions of peers, of whose existence P_n has previously learnt of, are used (sample set). The peer, in whose Voronoi region P_n is currently located, is called the *anchor* peer P_a and may change during iteration. At each position update, P_n gathers the dissimilarity-values and positions of all currently known peers (including the anchor peer as well as P_n itself) and updates its own position such that it is optimal in matters of stress minimization. This can e.g. be done by executing a standard MDS over all known positions, but applying the gradient only to P_n's position and thus leaving all other positions unchanged. If the newly found position lies outside of the Voronoi regions of the current sample set, P_n gets to know additional peers. The positions of these peers are added to the sample set and then taken into account for another update/refinement of x_P. The iteration/refinement stops, when the newly found position lies in the Voronoi-area of any of the already known peers. When this happens, P_n and the current anchor peer reorganize their neighborhood such that P_n is integrated into the network.

The according algorithm goes as follows:

Initialization	1.	Contact any random Peer P_r
	2.	Set initial ($t = 0$) anchor Peer $P_a^0 = P_r$
	3.	Define initial sample set $S^0 = \{P_n, P_a^0\} \cup N(P_a^0)$, where $N(P_a^0)$ denotes the set of all direkt neighbors of P_a^0.
Iteration t	1.	Find position x_n^{t+1} of P_n, such that the stress over all positions in S^t is minimized.
	2.	Set new anchor Peer P_a^{t+1} to be the one peer in the (whole) network, which has the position with the shortest distance to x_n^{t+1}
	3.	If $P_a^{t+1} \notin S^t$
	4.	Add all Peers in direct line between P_a^t and P_a^{t+1} to S
	5.	Continue with next iteration step (goto 1.)
	6.	else
	7.	Final position found, reorganize neighbors with P_a^{t+1}
	8.	Send position update to all neighbors

It is important to point out that finding the anchor peer for any position can easily be done using local algorithms e.g. by means of X-Y routing. This therefore also applies for finding out, which peers are located in direct line between any two positions (see iteration step 4.).

After the new peer has found its final position and has been integrated into the network,

it sends a position update into its neighborhood, so that the neighbors can refine their positions by taking P_n's position into account. This position update can be executed in exactly the way as it is done when adding a new peer.

4 Simulation Example

The proposed algorithm has been tested against different dissimilarity matrices both from publicly available questionnaires, and computer generated test-data. The outcome of these simulations is now discussed using the example of a helix[2] over 4.5 periods as high-dimensional data, of which random points have been chosen and been assigned to the peers. After each iteration, the overall stress has been determined.

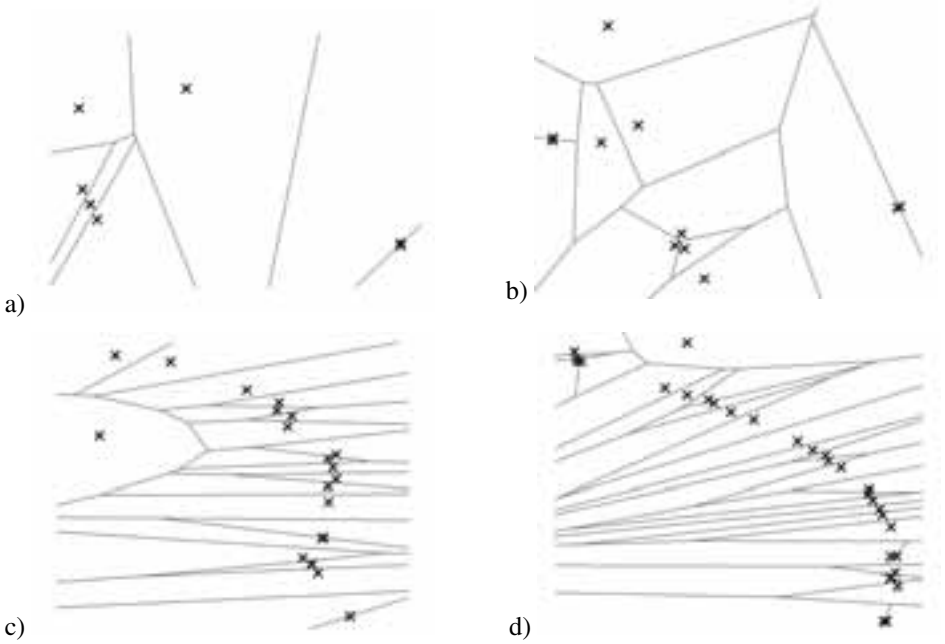


Figure 3: Example network after 7, 11, 22, and 30 peers have been added.

For this example, the proposed algorithm has been run from the first peer on, i.e. no initial global MDS has been executed. Fig. 3a-d) show the configurations after 7, 14, 22, and 30 peers have been added to the network. Each cross represents a peer, while the lines determine the borders of the Voronoi region of each peer. With fewer peers (see Fig3a) and b)), the overall stress assumes values up to $\approx 24\%$, such that the resulting configuration does not allow an intuitive interpretation. Nevertheless, after more peers have been added,

[2]The helix has only been chosen as example, because its projection into 2D is easily understandable, even when a large amount of points are used.

the resulting configurations in c) and d) intuitively seem to be a good 2D-representation of a helix. Indeed, the overall stress of these two configurations is $< 1\%$.

Fig. 4a) shows the development of the stress, when peers are added to the network. It can be seen that when adding only few peers, the stress assumes a large range of values. Nevertheless, with more peers added to the network, good configurations with remarkably low stress values are found.

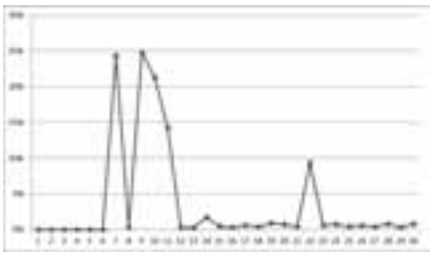
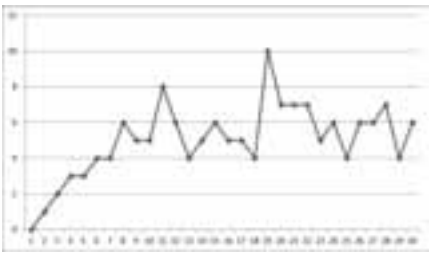
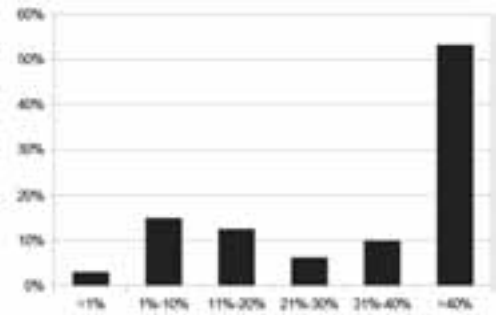
a)  b) 

Figure 4: a) stress as a function of number of peers, b) number of peers considered in add operation

Fig. 4b) shows the number of peers, which have been taken into consideration by a newly added peer to find its optimal position in the configuration. It can clearly be seen, that only a small subset (sample) of all peers is necessary to find the position. In few cases (e.g. when adding peer 12 or peer 19), larger portions of the network have to be explored, until the final position is reached. Since for each considered peer, two messages have to be sent to gather that peer's dissimilarity values, the number of considered peers has the same complexity as the number of messages needed to find the optimal position for the new peer. Since the newly added peer can contact several peers at a time to gather the dissimilarity information, the number of considered peers is also an upper bound for the necessary time.

The simulations have also shown that optimal configurations are not always found. Instead, the resulting final stress strongly depends on the initial seed of the pseudorandom number generator that has been used, and therefore both on a) the initial positioning of a newly added peer, and b) on the results of the stress minimization process during each position update. To demonstrate this effect, the algorithm has been executed on the same dataset repeatedly with varying initial seeds for the pseudorandom number generator.

In Fig. 5 it can be seen that only 3% of all simulation runs actually lead to excellent stress values of less than 1%, while another 15% of the simulation runs at least lead to good results with stress values in the range from 1% to 10%. In all other cases, the algorithm gets stuck in local minima when adding a peer to the network, such that a global optimum cannot be reached. This is a well known problem also in globally executed MDS.



Range	Ratio
$< 1\%$	3%
1% to 10%	15%
11% to 20%	13%
21% to 30%	6%
31% to 40%	10%
$> 40\%$	53%

Figure 5: Distribution of final stress when varying the initial random seed

5 Preliminary Conclusions and Future Work

The various simulations run so far indicate that the proposed algorithm is able to solve the problem of distributed non-metric MDS with results of the same quality as a non-metric MDS with global knowledge. Although further investigation of the algorithm is necessary to provide a profound analysis, the following preliminary conclusions can be made:

1. A good starting configuration found by using a global MDS is necessary to achieve good results and reliable behaviour. Otherwise, the algorithm tends to get stuck in local minima when a single peer updates its position. This leads to unsatisfying overall configurations, when new peers are added.
2. Several initial anchor peers should be used to avoid local minima. This is the distributed analogue to using different initial configurations in global MDS.
3. Position changes (updates) should only be propagated to neighbouring peers, if the distance between old and new position exceeds some threshold. This avoids network load and alternating position updates of neighboured peers.
4. The time and message complexity for add operations have been in $O(\sqrt{n})$ for all simulations.

After the preliminary conclusions have been strengthened through further simulations and analysis, future work will focus on finding robust starting configurations that allow to find globally optimal configurations as good as possible. Additionally, an estimation for the number of bootstrapping peers necessary to find satisfying configurations, shall be found.

References

[BBKY06] M. M. Bronstein, A. M. Bronstein, R. Kimmel, and I. Yavneh. Multigrid multidimensional scaling. *Numerical Linear Algebra with Applications*, 13(2-3):149–171, 2006.

[CHK+09] Raimondas Ciegis, David Henty, Bo Kagström, Julius Zilinskas, and Julius Zilinskas. Parallel Global Optimization in Multidimensional Scaling. In *Parallel Scientific Computing and Optimization*, volume 27 of *Springer Optimization and Its Applications*, pages 69–82. Springer New York, 2009.

[CPH06] Jose A. Costa, Neal Patwari, and Alfred O. Hero, III. Distributed weighted-multidimensional scaling for node localization in sensor networks. *ACM Trans. Sen. Netw.*, 2(1):39–64, February 2006.

[CU12] Hauke Coltzau and Bastian Ulke. Navigation in the P2Life Networked Virtual Marketplace Environment. In Herwig Unger, Kyandoghere Kyamaky, and Janusz Kacprzyk, editors, *Autonomous Systems: Developments and Trends*, volume 391 of *Studies in Computational Intelligence*, pages 213–227. Springer Berlin / Heidelberg, 2012.

[dL77] Jan de Leeuw. Applications of convex analysis to multidimensional scaling. *Recent developments in statistics*, page 133145, 1977.

[DM00] T. Denoeux and M. Masson. Multidimensional Scaling of Interval-Valued Dissimilarity Data. *Pattern Recognition Letters*, 21:21–83, 2000.

[Kru64a] J. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964. 10.1007/BF02289565.

[Kru64b] J. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29(2):115–129, June 1964.

[PD10] Piotr Pawliczek and Witold Dzwinel. Parallel Implementation of Multidimensional Scaling Algorithm Based on Particle Dynamics. In Roman Wyrzykowski, Jack Dongarra, Konrad Karczewski, and Jerzy Wasniewski, editors, *Parallel Processing and Applied Mathematics*, volume 6067 of *Lecture Notes in Computer Science*, pages 312–321. Springer Berlin / Heidelberg, 2010.

[Tor52] Warren Torgerson. Multidimensional scaling: I. Theory and method. *Psychometrika*, 17:401–419, 1952. 10.1007/BF02288916.

Impact of Application Layers over Wireless Sensor Networks

Cédric Ramassamy [†], Hacène Fouchal [*] and Philippe Hunel [†]

[*]CReSTIC Université de Reims Champagne-Ardenne, France
[†]LAMIA, Université des Antilles et de la Guyane, France

Abstract: Applications over Wireless Sensor Networks (WSNs) are various and different. Many routing protocols have been proposed for WSNs since a many years. Medium Access Control (MAC) protocols can differ from a network to another. The transmission range can also be variable from a sensor to another depending on their battery capacities. All these various protocols have an influence on the WSNs. It is worthy to choose the appropriate protocol for each specific situation in order to ensure high security of the network.

In this paper, we show that the type of applications has a real impact on WSN. Indeed, each kind of application with a specific routing protocol, MAC layer and transmission range value impact on security of network in terms of packets lost rate.

We have undertaken a set of experimentations in order to show the importance of an appropriate configuration to deploy a WSN with a high security confidence degree. For the application layer, we have handled three types of applications. For the routing layer, we have handled only with AOMDV protocol. For MAC layer we use 802.15.4 protocols. We have conducted many simulations through the NS-2 simulator in order to analyze one relevant security indicators on WSNs: lost packet rate.

1 Introduction

A WSN is mainly composed of two types of components: resource constrained independent nodes (sensors) and high resourced sink nodes (base stations). A node in a network communicates with each other through a wireless medium, where the communication energy is greater than the computational energy. In addition to that, the energy required to transmit a message is twice higher than the energy required for receiving a message. In the same time, the route of any message sent to base stations is a challenging issue about a network lifetime for many reasons. If we use the shortest routes where some nodes have small capacity batteries, we may reduce the network lifetime. If we use longer routes with many nodes, we will surely increase the network delays. For these reasons, routing objectives are, in general, conducted by higher applications: streaming video applications need longer network lifetime but fire monitoring applications need faster response time. Various routing mechanisms have been proposed for different applications. They are different according to their routing objectives and routing techniques. As said below, the network features will guide to choose adequate routing techniques. The IEEE 802.15.4 standard offers the most common communication protocols for the MAC and the physical (PHY)

layers [BPC+07]. After a deployment, it is not possible to handle nodes, in particular, for military applications, fire monitoring, underwater applications. Then it could be so complicated to replace the batteries [CoSKM04]. Thus, it is necessary to efficiently consume the energy and avoid lost packets when they communicate which will provide more security for WSNs.

We considered the following parameters in WSN: the radio range, the MAC layer, the network size, the routing protocol type, the application type

In this paper, we study the impact of applications on the WSN security.

We have conducted a large number of experimentations using the NS-2 framework. The radio range length has been tested for 25 meters. For the network size, we have used 25, 60, 100 and 300 nodes. But we show impact only for 300 nodes that represents high scalability networks. For routing, we have used AOMDV protocols. For the application type, we have used three kinds of applications having different rates: regular applications, applications with high rate communications (denoted high rate application in the rest of the paper) and applications with some burst communications (denoted burst based applications in the paper).

We observed in each case, the WSN security through the lost packets rate. We have then calculated a security metric which depends on this parameter.

This paper is organized as follows. Section 2 is dedicated to related works. Section 3 presents an overview of usual protocols used over WSNs. Section 4 describes our contribution, experimentations and analysis are also detailed. Section 5 concludes our work and gives some ideas about future work.

2 Related Work

In the literature, a comparative study of routing protocols can be found in [AY05] and [AkK04]. Authors discussed widely properties of many routing protocols. They proposed for WSNs different categories of routing protocols such as Location-based protocols, Data-centric Protocols, Hierarchical Protocols, Mobility-based Protocols, Multipath-based Protocols, Heterogeneity-based Protocols and QoS-based protocols. In [BPSM09], authors analyze the design issues of sensor networks and present a classification and a comparison of these routing protocols. They show that it is no possible to design a routing algorithm which will have the best performances under all scenarios and for all applications.

In [BMJ+98] a useful comparative study about pro-active and reactive routing protocols, where authors have evaluated four ad hoc routing protocols such as DSDV, TORA, DSR and AODV. Many studies and improvements about these routing protocols have been detailed. In [NJ10] and [DP00], authors analyze performance about AODV and DSR routing protocols. In [GS09], authors compare AODV and DSDV with an Optimized-AODV routing protocol, which provides better results than AODV and DSDV. In [MD01], authors present AOMDV, and show that AOMDV always offers better routing performances than AODV in many mobility and traffic conditions. In [Bou04], the authors study and compare

the performance of the following routing protocols AODV, PAODV (Preemptive AODV), CBRP, DSR and DSDV. They show that CBRP has a higher overhead than DSR because of its periodic hello message while AODV's end-to-end packet delay is the shortest when compared to DSR and CBRP, and PAODV has shown some improvements compared to AODV.

In [KKHH06], the authors analyze the IEEE 802.15.4 performance with large-scale WSN applications. They study the CSMA-CA mechanism and the MAC operations in a beacon-enabled cluster-tree structure. They analyze performances in terms of power consumption and goodput of the coordinator. The results are validated using WIreless SEnsor NEtwork Simulator (WISENES). In [KC12] authors deal with two categories of MAC technique: contention based and schedule based. They give a unique performance analysis and comparison of benefits and limitations of each protocol. They show that random topology contention based approach may be helpful and also schedule based approach may be more energy efficient if deployment is not random.

In [BYAH06], authors present X-MAC, a MAC protocol where the main objectives are: energy-efficiency, simple, low-overhead, distributed implementation, low latency for data, high throughput for data and applicability across all types of packeting and bit stream digital radios. They compare X-MAC and LPL. In [SRDV08] authors show that X-MAC provide a power-saving mechanism for routing nodes in Contiki simulator, with this method they show that X-MAC reduces the power consumption for ZigBee routing nodes with up to 90%. Other research studies focus on new methods to transmit data efficiently in order to reduce the energy consumption [YBO09, BOY10] or to guarantee messages arrival [AK05].

[LLW08] proposes a cross-layer handoff ordering scheme. Different quality of service (QoS) requirements of various applications would result in different frame success rate requirements. In order to indicate how critical a handoff request is, both the frame success rate requirement from the application layer and the frame success rate measurement from the MAC layer are taken into consideration in the proposed scheme. The prioritization of handoff requests follows the most-critical-first policy. Performance analysis shows that the proposed scheme effectively reduces the forced termination probabilities.

[ZHL11] investigates energy efficiency of secure communication in wireless sensor networks. It considers link layer security of WSNs which considers cryptographic mechanisms implementation schemes. They evaluate the computational energy efficiency of different symmetric key ciphers and they show that the computational energy cost of block ciphers is less than that of stream ciphers when data are encrypted and transmitted through a noisy channel. They also investigate different factors affecting the communication energy cost of link layer cryptographic schemes (size of payload, the mode of operation applied to a cipher and the quality of the communication channel). They provide a comparison study of many cryptographic schemes in terms of energy consumed by both computation and communication. [BOM10] shows how to increase security in Ad hoc networks by using multipath routing. If one route is damaged or not enough secure, another is chosen. In [GMT12], a smart technique to monitor denial of service (DoS) over WSNs is proposed. It shows how to detect DoS over WSNs using clustering techniques. In [BF10], we have designed an efficient MAC protocol based on reducing useless communications between

nodes.

In all these studies, there is a lack of mixing all layers to consider security performances of any WSN. In our case, we consider all layers and intend to provide mechanisms to tailor with precision the choice of relevant factors in order to reach more secure WSN network.

3 Usual Protocols

This section is dedicated to detail all usual protocols that we use in our comparative scheme.

3.1 IEEE 802.15.4

The IEEE 802.15.4 standard defines the features of the physical and MAC layers for Low-Rate Wireless Personal Area Networks (LR-WPAN) [BPC+07]. It supports short-range operation, reasonable battery life, while maintaining a simple and flexible protocol stack.

3.1.1 Physical Layer

The physical layer uses the Direct Sequence Spread Spectrum (DSSS) access mode in three frequency bands 2450 MHz (with 16 channels), 915 MHz (with 10 channels) and 868 MHz (with 1 channel). Besides radio on/off, the physical layer gives some functionality for channel selection, link quality estimation, energy detection measurement and clear channel assessment to assist the channel selection.

3.1.2 MAC layer

The IEEE 802.15.4 MAC layer defines two types of nodes: FFDs (Full Function Devices) equipped with a full set of MAC layer functions and RFDs (Reduced Function Devices) equipped with a reduced set of MAC layer functions and they can communicate only with a single FFD. The standard considers two main topologies a Star topology and a Peer-to-Peer topology. The MAC layer has two modes for medium access: Non beacon mode is purely based on CSMA/CA and beacon enabled mode operates with a Superframe which contains an active and an inactive portion for energy conservation.

The IEEE 802.15.4 standard has two main topologies: **Star topology** based on a master-slave model. The PAN coordinator represents the master, and slaves are other nodes such as FFDs or RFDs. Slaves can only communicate with the PAN coordinator. A **Peer-to-Peer topology** where the PAN coordinator can communicate with other FFDs, and they relay messages to others outside of their radio coverage. This topology forms a multihop network where the PAN coordinator is an administrator.

The MAC layer has two modes for medium access:

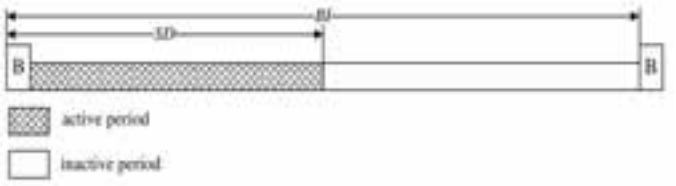


Figure 1: Superframe with SD and BI

1. **Nonbeacon mode:** The medium access is purely based on the CSMA/CA (Carrier Sense Multiple Access) /(Collision Avoidance) mechanism.

2. **Beacon mode:** The PAN coordinator operates with a Superframe. It starts the Superframe with beacon for node synchronization. The Superframe contains an active and an inactive portion where nodes may move to the sleeping status and then save energy. The active portion contains fixed size slots which represent two period: a Contention Access Period (CAP) where nodes use CSMA/CA mechanism, and a Contention Free Period for large packets or time-critical data deliveries assigned by the PAN coordinator. Synchronization and sending (non GTS) operations are executed in the CAP period. Informations for pending delivery are in the beacon frame.

However, the durations of the Superframe are called *beacon interval* (*BI*) ranges from 15ms to 245s. And the durations of an active period are called *SD* (see Figure 1). But the length of *BI* and *SD* can be determined by the parameter *macBeaconOrder* (*BO*) and respectively *SuperFrameOrder* (*SO*). These values can be derived as:

$BI = aBaseSuperframeDuration \times 2^{BO}$

$SD = aBaseSuperframeDuration \times 2^{SO}$

$0 \leq SO \leq BO \leq 14$

where aBaseSuperframeDuration is the minimum duration of a Superframe i.e.

$aBaseSuperframeDuration = aBaseSlotDuration \times aNumSuperframeSlots$

$aBaseSlotDuration = 60 symboles$

$aNumSuperframeSlots = 16$

So in 2400MHz, rate may be $250Kb/s$ or $62.5Ksymb/s$. Thus for $BO = 0$, we have $BI = 960symb$ then $BI = 15.36ms$

The proportion of time where a system can operate represents a duty cycle. For 100% duty cycle, nodes are always on, that represents $BI = SD$ so $BO = SO$. For less than 100%, duty cycle nodes can turn off their transceiver to reduce power consumption. Duty cycle can be calculated as:

$$duty\ cycle = \frac{SD}{BI} = (\frac{1}{2})^{BO-SO}$$

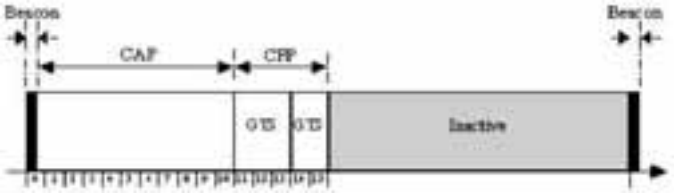


Figure 2: Superframe with active and inactive period

According to [HP07], different pairs values of *BO* and *SO* (in the same duty cycle) will provide various low-power attributes, and will have different impacts on throughput and energy consumption. That means it will definitely influence the network security.

3.2 Routing Layer

Routing protocols are required to ensure multihop communications. Indeed, if the wireless nodes are within the range of each other, a routing protocol is not necessary. But, nodes can move or would communicate with a node out of this range. Then, intermediate nodes are needed to organize the network which takes care of data transmission.

Routing protocols algorithms must choose some criteria to make routing decisions, for instance the number of hops, latency, transmission power, bandwidth, etc. Thus routing protocols are divided into two classes:

- Proactive routing protocols
- Reactive routing protocols

In the following, we present briefly the routing protocol that we used in our study:

3.2.1 AOMDV Protocol

The Ad Hoc On-demand Multipath Distance Vector Routing extends the prominent AODV to discover multiple link-disjoint paths between the source and the destination in every route discovery. It is designed mainly for highly dynamic ad hoc networks where link failures and route breaks occur frequently. The AOMDV protocol has two main features:

- a route update rule to establish and to maintain multiple loop-free paths at each node.
- a distributed protocol to find link-disjoint paths.

Multipath routing protocols, such as AOMDV, try to reduce the high latency of route discovery which can decrease performance.

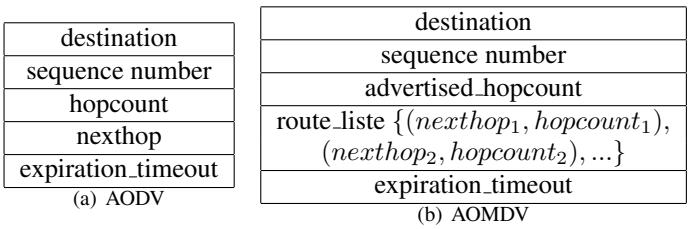


Figure 3: Structure of routing table entries for AODV and AOMDV

Fig. 3 shows the structure of the routing table entries for AODV and AOMDV. In AOMDV $advertised\_hopcount$ replaces $hopcount$ in AODV. A $route\_list$ replaces the $nexthop$, and essentially defines multiple next hops with respective hopcounts. However, all next hops still have the same destination sequence number. The $advertised\_hopcount$ is initialized each time the sequence number is updated.
A node i updates its *advertised_hopcount* for a destination d whenever it sends a route advertisement for d.

3.3 Application Layer

We have classified applications in three types of applications:

1. Regular applications, which characterize those which send low-data with large intervals. In NS-2 we represent this kind of application in a CBR (Constant Bite Rate) stream with 2 seconds intervals. This CBR stream generates UDP traffic according to a deterministic rate. Packets have constant size.

2. High rate applications characterize applications which have large streams. In the NS-2 tool, we use CBR stream with 0.2 seconds to represent the overload of the network.

3. Burst based applications, characterize applications which send data on burst time and sleep during the remaining time. In order to represent this kind of applications in NS-2, we use a Poisson stream with $0.5s$ on period, $2.5s$ for off period and $50kb$ rate. The Poisson stream generates traffic according to an Exponential On/Off distribution. Packets are sent at a fixed rate during a period, and no packet is sent during off periods.

4 Contributions

In this section, we observe the impact of application layer with the different protocols parameters on the network security.

4.1 Working Environment

In the following, we detail the parameters which have been investigated:

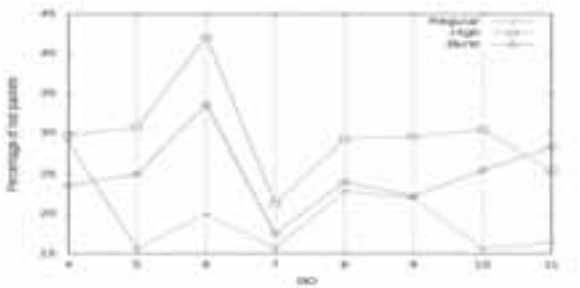
- The network size with 300 nodes. We use a Peer-to-Peer topology with one PAN coordinator.
- The physical layer with all nodes are FFDs using $2400MHz$ band frequency and $250kps$ bandwidth.
- The MAC layer could be in a non-beacon mode and in a beacon-enabled mode.
- For routing, we use AOMDV protocol.

Nodes can reach their neighbors located in the transmission range. We have to follow two steps for simulation. One for synchronization between nodes, and another one for executing the application job. The total duration is different for each simulation because nodes are synchronizing in different duration in beacon-enabled mode. In all simulations, application time is $250\ seconds$. We use a $interval\_sync$ to start nodes at different intervals to reduce the synchronization time, according to BO and SO values, synchronization time varies between 20 and 10000 seconds. For the power consumption, we adopt the data sheet in [SKKW07] where transmission mode is $76.2mW$ and reception mode is $83.1mW$, and nodes start with high energy level ($2.5J$).

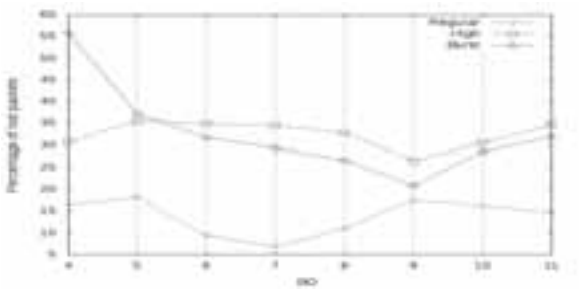
For each application class, we consider some end-device nodes which send messages. The numbers of devices which send data depend on the kind of application load and network load. For example, with 25 nodes, we use 8 nodes to send data to represent regular applications. All messages are sent to the PAN coordinator which is located in the middle of the network.

4.2 Experimental Results

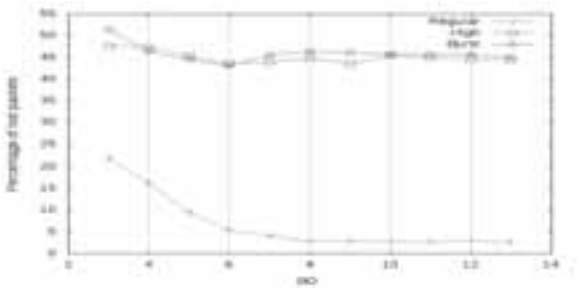
At the end of all simulations, we analyze the rate of lost application packets. We have studied some duty cycle values such as: 6.25%, 12.5%, 25%, 50% and 100% and Non-beacon mode. But we show only 25%, 50% and 100% duty cycle and Non-beacon mode, since under 25%, we have bad results related the complexity of the network to communicate. Indeed, under 25%, the active period is very short between two successive inactive periods. All generated packets will be buffered, and these packets will be transmitted in a burst period during the beginning of the active period, and the collision probability at that time will increase, then the number of received packets decreases. For these values of duty cycle, we use *BO* value between 4 and 11. Values under 4 are not considered in our experiments, because in NS-2, it is harder to simulate scenario with limited bandwidth resources.



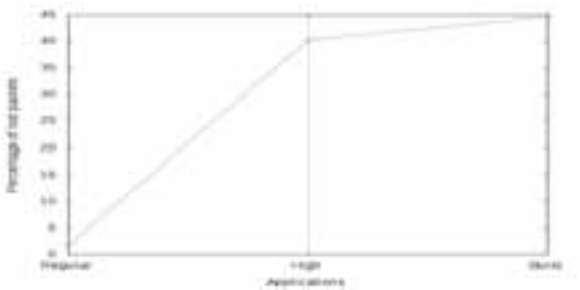
(a) 25% duty cycle



(b) 50% duty cycle



(c) 100% duty cycle



(d) Non Beacon mode

Figure 4: Average rate for all applications with 300 nodes

4.3 Result Analysis

We only comment about 300 nodes results due to the number of graphics. In the first case, we can see that lost packets of burst based and high rate applications is higher than in regular applications in 50%, 100% duty cycle and non-beacon mode. These results show that these applications are more critical than a regular application, which can compromise the security of the network when nodes are mostly on. Indeed, the number of lost packets is higher than the increase of retransmission, collision, and energy consumption. Theses consequences decrease the level of security about the network and may increase attacks of man in the middle.

In the second case, we can see that for all applications the values of MAC layer (BO and SO) impact in a different way. For regular applications, we observe that the best results are obtained when $6 \leq BO \leq 10$ for 100% duty cycle, $6 \leq BO \leq 8$ for 50% duty cycle and when $BO = \{5, 7, 10, 11\}$ for 25% duty cycle. However, the best choice is for non-beacon mode which obtained best results. For high rate applications and burst based applications in 100% duty cycle and non-beacon mode, nodes are always on and lost packets rate is high. Results are similar in these two cases. But when nodes turn off their transmission radio during a period, that reduces lost packets. Indeed for burst based we obtain interesting results when $BO = 9$ for 50% duty cycle and $BO = 7$ for 25% duty cycle. For high rate it is roughly the same results.

So, these results show for a regular application, the security level increases for "on" periods. but in most cases of "off" periods, the security decreases.. Whereas, burst based and high rate applications need off periods to reduce lost packets to improve the level of network security.

5 Conclusion

In this paper, we have shown that the application layer has a real impact on WSN security. We have observed the security of a WSN for many kinds of applications and have measured a metric in terms of packets lost rate.

We have conducted a set of experimentations with 300 nodes executing 3 different types of applications. All simulations have been achieved using the well-know NS-2 simulator. We have presented numerical results of the network performance in each case. We have given what are the best practices for each configuration depending on the application layer type.

We are currently working on extension of this work in order to observe other impacts on WSN security and performances:

- Higher network sizes and much more detailed transmission range values
- Symetric data encryption mechanisms

Our main future work about this study is to justify our results by a theoretical approach

which could be based on stochastic techniques or a qualitative evaluation methods.

References

[AK05] Younis Mohamed Akkaya Kemal. A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks*, 3(5):325–349, 2005.

[AkK04] Jamal N. Al-karaki and Ahmed E. Kamal. Routing Techniques in Wireless Sensor Networks: A Survey. *IEEE Wireless Communications*, 11:6–28, 2004.

[AY05] Kemal Akkaya and Mohamed Younis. A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks*, 3:325–349, 2005.

[BF10] Thibault Bernard and Hacène Fouchal. Efficient Communications over Wireless Sensor Networks. In *Proceedings of the IEEE Global Communications Conference (Globecom-2010)*, pages 1–5, Miami, USA, December 2010 2010.

[BMJ+98] Josh Broch, David A. Maltz, David B. Johnson, Yih chun Hu, and Jorjeta Jetcheva. A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. In *Mobicom 1998*, pages 85–97, 1998.

[BOM10] J. Ben-Othman and L. Mokdad. Enhancing Data Security in Ad hoc Networks based Multipath Routing. *Journal of Parallel and Distributed Computing (JPDC)*, 70(3):309–316, March 2010.

[Bou04] Azzedine Boukerche. Performance evaluation of routing protocols for ad hoc wireless networks. *Mob. Netw. Appl.*, 9:333–342, August 2004.

[BOY10] J. Ben-Othman and B. Yahya. Energy Ecient and QoS based Routing Protocol for Wireless Sensor Networks. *Journal of Parallel and Distributed Computing (JPDC)*, 70(8):849–857, Aug 2010.

[BPC+07] Paolo Baronti, Prashant Pillai, Vince W.C. Chook, Stefano Chessa, Alberto Gotta, and Y. Fun Hu. Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards. *Computer Communications*, 30(7):1655 – 1695, 2007. Wired/Wireless Internet Communications.

[BPSM09] Rajashree V. Biradar, V.C. Patil, S. R. Sawant, and R. R. Mudholkar. Classification And Comparison Of Routing Protocols In Wireless Sensor Networks. *UbiCC Journal*, 4:704 –711, 2009.

[BYAH06] Michael Buettner, Gary V. Yee, Eric Anderson, and Richard Han. X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks. In *Proceedings of the 4th international conference on Embedded networked sensor systems*, SenSys '06, pages 307–320, New York, NY, USA, 2006. ACM.

[CoSKM04] Rachel Cardell-oliver, Keith Smettem, Mark Kranz, and Kevin Mayer. Field Testing a Wireless Sensor Network for Reactive Environmental Monitoring. In *In International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pages 14–17, 2004.

[DP00] Samir R. Das and Charles E. Perkins. Performance comparison of two on-demand routing protocols for ad hoc networks. pages 3–12, 2000.

[GMT12] M. Guechari, L. Mokdad, and S. Tan. Dynamic Solution for Detecting Denial of Service Attacks in Wireless Sensor Networks. In *IEEE International Conference on Communications (ICC'2012)*, June 2012.

[GS09] Dr Aditya Goel and Ajaii Sharma. Performance Analysis of Mobile Ad-hoc Network Using AODV Protocol. *International Journal of Computer Science and Security*, 2009.

[HP07] Yu-Kai Huang and Ai-Chun Pang. A comprehensive study of low-power operation in ieee 802.15.4. In *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*, MSWiM '07, pages 405–408, New York, NY, USA, 2007. ACM.

[KC12] Joseph Kabara and Maria Calle. MAC Protocols Used by Wireless Sensor Networks and a General Method of Performance Evaluation, 2012.

[KKHH06] Mikko Kohvakka, Mauri Kuorilehto, Marko Hännikäinen, and Timo D. Hämäläinen. Performance analysis of IEEE 802.15.4 and ZigBee for large-scale wireless sensor network applications. In *Proceedings of the 3rd ACM international workshop on Performance evaluation of wireless ad hoc, sensor and ubiquitous networks*, PE-WASUN '06, pages 48–57, New York, NY, USA, 2006. ACM.

[LLW08] Po-Chiang Lin, Tsungnan Lin, and Chiapin Wang. Performance analysis of a cross-layer handoff ordering scheme in wireless networks. *IEEE Transactions on Wireless Communications*, pages 5166–5171, 2008.

[MD01] Mahesh K. Marina and Samir R. Das. On-demand Multipath Distance Vector Routing in Ad Hoc Networks. In *in Proceedings of IEEE International Conference on Network Protocols (ICNP*, pages 14–23, 2001.

[NJ10] Amit N.Thakare and Mrs. M. Y. Joshi. Performance Analysis of AODV & DSR Routing Protocol in Mobile Ad hoc Networks. *IJCA Special Issue on MANETs*, pages 211–218, 2010. Published by Foundation of Computer Science.

[SKKW07] D. Schmidt, M. Krämer, T. Kuhn, and N. Wehn. Energy Modelling in Sensor Networks. Bundesrepublik, Deutschland, 2007. Copernicus Publications on behalf of the URSI Landesausschuss.

[SRDV08] Pablo Suarez, Carl-Gustav Renmarker, Adam Dunkels, and Thiemo Voigt. Increasing ZigBee network lifetime with X-MAC. In *Proceedings of the workshop on Real-world wireless sensor networks*, REALWSN '08, pages 26–30, New York, NY, USA, 2008. ACM.

[YBO09] B. Yahya and J. Ben-Othman. Towards a Classication of Energy-Aware MAC protocols for Wireless Sensor Networks. *Wireless Communications and Mobile Computing (WCMC)*, 12(3):1572–1607, Feb 2009.

[ZHL11] Xueying Zhang, Howard M. Heys, and Cheng Li. Energy efficiency of encryption schemes applied to wireless sensor networks. *Security and Communication Networks*, 2011.

On the adoption of usage control technology in collaborative environments

Åsmund Ahlmann Nyre Martin Gilje Jaatun

nyre@idi.ntnu.no martin.g.jaatun@sintef.no

Abstract:

The extensive collaboration across system boundaries facilitated by the Internet is unfortunately also increasing the potential for misuse of shared information. While mechanisms to protect assets from active attackers (such as firewalls, intrusion detection systems and anti-virus software) are commonplace, the availability of commercial software to protect information from misuse remains limited. Businesses employ non-disclosure agreements, but have no means to prevent or detect violations of these. Usage control has been proposed as the means to protect information even after its distribution. However, despite the efforts to develop new usage control technology and the apparent need for it, the industry is less enthusiastic. In this paper we investigate existing theories of technology adoption in order to better understand the industry-perspective and to improve technology development. We base our study well-known general theories on protection motivation, innovation diffusion and technology acceptance. We then utilize these theories and preliminary results of a case study to build a new model for understanding usage control technology acceptance.

1 Introduction

While traditional access control is focused on providing or denying access to resources, it is currently unable to restrict and control how those resources are being used. That is, once the information has left the system any restrictions on its use are solely at the receiver's discretion. For instance, suppose a consultant has been granted access to an incident report from an organisation in order to analyse the cause of the incident. What is there to prevent the consultant from leaking the report to public media or a competitor? In order to remedy this, the research community have proposed *usage control* [PS04], which may be seen as a extension of access control beyond system borders to control usage on the client-side. Returning to our example, usage control could have granted the consultant rights to read the report for the duration of the assignment while ensuring that copying or forwarding was prohibited.

As we have previously shown [Nyr11] there is an ample body of research on technology to specify, organise and enforce usage control polices. Despite the researchers' firm beliefs that businesses should employ usage control technology, there is to the best of our knowledge no empirical support for its appropriateness. This paper represents a starting point to identify why the industry does not share researchers' view on the necessity of distributed

protection against information misuse. To this end, we attempt to combine different technology adoption theories and models as well as preliminary results of an ongoing case study in a collaborative environment in order to explain this discrepancy.

The remaining parts of this paper is organised as follows. Section 2 gives a brief introduction to theories and models to explain the adoption of technology. We provide both general theories and some more specific to protective technologies. Next, we review some of the previous approaches to explaining security technology (or behaviour) adoption in Section 3. In Section 4 we develop and describe our new model for usage control technology adoption. Finally, we provide a brief discussion of the model and its appropriateness and give our concluding remarks in Section 5.

2 Background and theories on technology adoption

Technology adoption and how innovations diffuse in a social system has been a subject of interest for many years. What is it that causes some technology to have a wide uptake while other, seemingly good innovations fail? And, perhaps more interesting for market analysts: is it possible to predict the adoption rate of new technology before it enters the market? Before we return to these questions for usage control technology we provide a brief overview of the main general purpose theories and models that have been developed.

In his theory on *diffusion of innovations*, Rogers [Rog03] outlines five main categories of factors that influence, and possibly can predict, the decision to adopt new technology. A lot of consideration is devoted to factors outside the innovation itself, such as promotion efforts and use of communication channels. We however limit our discussion here to the intrinsic characteristics of the innovation, termed *the perceived attributes of the innovation*.

- Relative advantage
- Compatibility
- Complexity
- Observability
- Perceived risk

Placing attributes within one of these attribute categories may not be straightforward. In a review by Tornatzky and Klein [TK82], they found that there was considerable dispute among researchers regarding the contents of the attribute categories, in part due to the lack of separation between the attributes. In one study *profitability* is seen as a separate factor, whereas in another it is seen as part of the *relative advantage*. Similarly, the *compatibility* attribute may denote both the degree to which it fits with the ideas/norms of the current system and the degree to which it fits with the operational aspects of the current system (i.e. the way things are done). Although these differences made it difficult to assess the dominant attributes in predicting innovation diffusion, the meta-analysis found that *compatibility* and *relative advantage* were positively related to adoption whereas *complexity*

was negatively related. Other attributes did not display significant relation to adoption, however as noted by the authors this might be due to the inconsistency in attribute definition.

From behavioural psychology there are also theories that may be utilised in order to describe, understand and predict how innovations (in our case, technology) may be adopted by users and organisations. The Theory of Reasoned Action (TRA) [FA75] is based on the idea that attitude towards a behaviour and the subjective norm (i.e. others' attitude) affect the intention to behave a certain way, which in turn affects the actual behaviour. The theory was later extended by Theory of Planned Behaviour (TPB) [Ajz91] to include users' perceived control of the behaviour or the degree to which the action is perceived as voluntary. It can be seen that a user's attitude towards a behaviour, or in our case: *use* of technology, is to a great extent affected by Roger's *attributes of the innovation*. Notice however the difference between Rogers' focus on the innovation itself and the theory of planned behaviour's focus on its *usage*. However, as demonstrated by Moore and Benbasat [MB91], a slight change of wording in Roger's attribute definitions would yield a more behavioural theory.

In the realm of information systems the Technology Acceptance Model (TAM) [Dav89] has been one of the main models for predicting technology uptake. It defines two main factors affecting technology adoption: *perceived usefulness* and *perceived ease of use*. It builds on both TRA and TPB as it attempts to explain the actual technology acceptance based on the attitudes towards using the technology. However, the main factors are also very similar to Rogers' perceived relative advantage and perceived complexity of an innovation. Further, Venkatesh and Davis [VD00] argue that the other factors defined by Rogers are mediated by perceived ease of use and usefulness. It is a very simple yet powerful model for explaining and predicting use of information systems. Still, the simplicity has also motivated an extension to TAM, denoted TAM2 [VD00], by extending the original TAM with many of the factors defined in Rogers' innovation model. Thus, social constructs such as *subjective norm*, *image* and *voluntariness* and cognitive constructs such as *job relevance*, *output quality* and *result demonstrability* are included as factors directly or indirectly affecting technology acceptance. In an effort to synthesise the strong points of the various models for technology adoption, Venkatesh et al. proposed the Unified Theory of Acceptance and Use of Technology (UTAUT) [VMDD03].

Another approach to explaining human behaviour is the rational choice theory [Sco00], popularly referred to as the cost-benefit tradeoff. The theory has been extensively used both in economics, criminology and social sciences [Sco00]. Basically, the theory states that when faced with choices, humans will select the option that maximises the utility. Hence, a rational choice is one that adheres to this rule. McCarthy describes the requirements of being "rational" in more detail [McC02]. Referring to the *relative advantage* as defined by Rogers, it is possible to see this as indeed a cost-benefit analysis.

While the above mentioned theories all seem to describe *functional* technology, the situation for *non-functional* technology is somewhat different. For example, the usefulness of a firewall is very much dependent on the perceived risk of not using it. Thus if the risk of intrusions is high, the firewall is perceived to be more useful. Conversely if the risk of intrusions is low, the firewall is not that useful. It is exactly this notion that the Protec-

tion Motivation Theory (PMT) [Rog75] attempts to describe. The essence of the theory is that the intention to behave in a certain manner and consequently the actual behaviour is affected by the risk perception (vulnerability and severity) and the degree to which the coping measure (that mitigates the risk) is efficient or not. Jointly, these two factors may be viewed as the *relative advantage* of the protective technology, and hence the links to Rogers' theory on innovation diffusion as well as TAM are quite clear. Additionally, the tradeoff between perceived risk and perceived efficacy of the coping response may also be viewed in relation to the cost-benefit tradeoffs of the rational choice theory. Protection motivation theory has been extensively used in campaigns for health related issues and traffic safety [lFPDR00].

Deterrence theory has also been used in order to understand adoption of non-functional technology. The general deterrence theory (GDT) [HWF05] is based on the idea that if abusive behaviour is likely to be detected and severely punished, humans will refrain from such activity. Although it is possible to view deterrence as a means to force people to use a certain technology, we instead view deterrence as one of the primary theories behind any audit or monitoring-based usage control technology.

While the above theories on technology acceptance and usage are different both in focus and viewpoint, it is evident that they do share some fundamental concepts. Theories on the adoption of functional technology uses *usefulness*, *relative advantage* or *benefit* to describe the users' need for the technology, whereas the theories on adoption of protective technology use the combination of *threats* and *response efficiacy* to convey the same. All theories do however consider other factors in varying degree of detail such as social complexity, compatibility, social norm, etc.

3 Related work

We have noted earlier that empirical evidence of actual use of information and usage control technology is to a large degree absent [Nyr11]. However, by regarding usage policies as similar to security policies we can draw upon the work on security policy compliance. The problem in this area is to determine and predict the degree to which users adhere to the security policies they are subject to.

Siponen et al. [SPM06] apply a combination of the theory of reasoned action (TRA) [FA75] and the protection motivation theory (PMT) [Rog75] in order to identify the factors affecting information systems security policy compliance. The study was conducted within five companies and a total number of respondents of 919. The results demonstrate that normative beliefs and visibility significantly influences both the threat appraisal and coping appraisal, which in turn affects the intention to comply and the actual compliance with the security policy. Although both the theoretical and empirical evidence are sound, the granularity of the factors makes it somewhat difficult to apply to usage control. For example, vulnerability is defined as *"the employees' assessment of whether their organization is vulnerable to IS security threats"* [SPM06] and the corresponding response-efficacy is to what extent the security policy mitigates the threat. In a follow-up study [SPM07], the

authors also employed the general deterrence theory and demonstrated that sanctions have a significant effect on the actual policy compliance. This is supported by other studies as well [HR09b, HR09a]

A more complex model was developed by Bulgurcu et al. [BCB10] including security policy awareness and general security awareness as important factors affecting the compliance with security policies. By basing their model in part on the rational choice theory, the authors include perceptions of benefits and costs of both compliance and non-compliance with the security policy. The security policy awareness is only self reported, as opposed to a knowledge quiz to determine the awareness. Thus, it is perceived security policy awareness, rather than actual awareness that is found to affect intention to comply.

One of the few studies on security technology adoption utilized Protection Motivation Theory to explain users' attitude and behaviour towards home wireless security [WTL05]. The study used a knowledge quiz to determine respondents' security awareness with respect to wireless security and possible coping measures. Of the factors identified by Rogers [Rog75], perceived vulnerability was the single factor that did not demonstrate a significant effect on determining whether the user actually had enabled the prescribed security measures. A similar approach was taken by Chenoweth et al. [CMG09] in order to study the use of anti-spyware software. In this study, the authors did not only seek to identify factors affecting intention to utilize the coping response, but also maladaptive coping. That is, to identify the factors that lead users to deliberately not utilize the coping behaviour. Perhaps not surprisingly, the response cost was the only factor that significantly affected the maladaptive behaviour. Whereas the threat appraisal was shown to positively affect the intention to use security technology, the response cost was shown to negatively affect it.

Common to all of these approaches is that they consider only the collective threat and the collective response efficacy as perceived by users. For instance in the study by Woon et al. [WTL05], perceived severity is measured on three different threats; identify theft, e-mail eavesdropping and privacy breaches, whereas perceived vulnerability is only measured in total. None of them consider how different attackers may influence both the severity and vulnerability (or probability) of a specific threat. . And in the study by Siponen et al. [SPM06] they only consider the collective threat *"employees' assessment of whether their organization is vulnerable to IS security threats"* [SPM06], which does not consider whether the employees consider the threats that the security policy is supposed to mitigate.

4 A model for usage control technology adoption

In this section we elaborate on our work to apply the existing theoretical models of technology and innovation adoption to the area of usage control technology in collaborative environments. To this end, we have conducted a case study within the oil and gas sector in Norway to identify important factors influencing adoption of usage control technology as seen by the industry itself.

We start by providing a brief background on the case study conducted and a description of characteristics of the oil and gas industry that makes this industry ideal for the type of

collaborative environments we are to study.

4.1 Case study

The motivation for the case study is to better understand the perceptions of risk and coping technology in the IO community context. Through interviews we have sought to identify potential risks, appropriateness of current security measures and the perceived efficacy of usage control technology as the coping response. Currently, we have conducted 9 interviews with key personnel from both oil companies, contractors and consultancy businesses. While we are unable to make definitive statements based solely on these interviews, they do provide insights on the problems and perceptions of the industry that can well be used to build and extend our theoretical models.

The term *Integrated Operations* (IO) has been used by the oil and gas industry on the Norwegian continental shelf to denote a state in which information and services are integrated across onshore and offshore, across different locations, and across organisations. This vision of seamless integration requires collaboration on an unprecedented level. While collaborative efforts are common in many sectors, the special case of the oil and gas sector is the mixture of competition and collaboration. For example: Oil companies are competing to find, produce and ship oil, but at the same time they are collaborating on exploration activities and remote field developments. Thus, it is vital that information shared in one context is not misused in another. Which of course puts heavy demands on the ability to control information flow and information usage. The combination of collaboration and competition coupled with seamless data integration is what makes the oil and gas industry particularly interesting for the adoption of usage control technology.

4.2 The basic model

Our basic model is depicted in Figure 1. As can be seen, it is primarily based on the Protection Motivation Theory but adapted to information security. The model also includes some elements from the theory of innovation diffusion as well as the Technology Acceptance Model.

Risk appraisal is used instead of the original *threat appraisal* in PMT. This is primarily due to different definitions of the concepts. Within information security, *risk* is commonly used to denote the combined likelihood and effect of a threat action performed by a threat agent. Thus, we therefore let system vulnerability be a contributing factor to the likelihood of a threat action.

The *response cost*, *response efficacy* and *self-efficacy* are all taken from PMT. We have however added the *ease of use* concept form the Technology Acceptance Model and the *compatibility* from the diffusion of innovations theory. Since the original PMT model seems to only consider a coping response versus no response, we have introduced the *relative advantage* concept from the diffusion of innovations theory. We have discovered

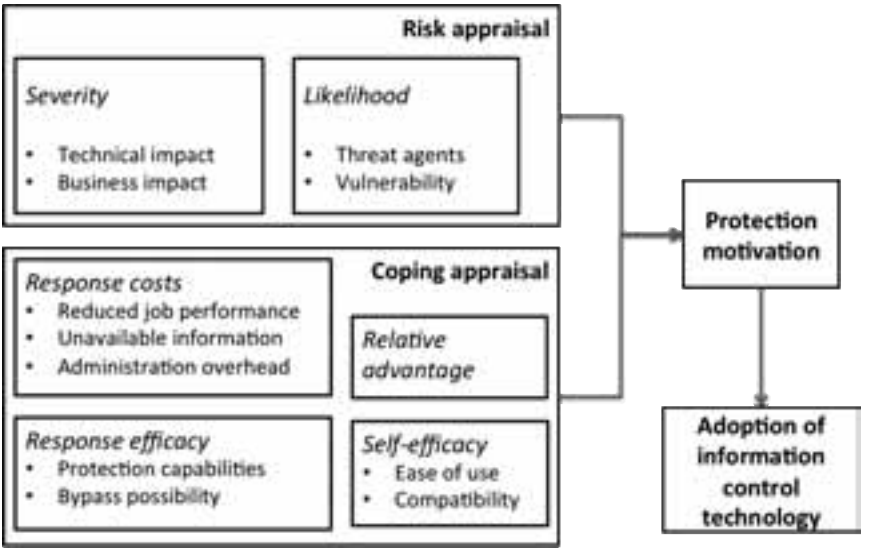


Figure 1: Adoption model for usage control technology

that there are often existing coping responses in use. Therefore the the coping response must not only be effective and efficient, it must also surpass any currently implemented response. In the following subsections we elaborate further on the various parts of our model and provide justifications for them based on the case study.

4.3 Risk perception

The original PMT model only handles threats collectively, both with respect to severity and vulnerability. However, as we discovered through our interviews, there are a wide range of threats to shared information that are virtually impossible rate collectively for severity and vulnerability. Additionally, as already mentioned in Section 4.2, the perceived severity and likelihood of a threat action is greatly dependent on the threat agent (or attacker), something that the original PMT model does not consider. In classical safety thinking, which is where PMT originated, threats are assumed to be unintentional, such as getting cancer from cigarette smoke or being seriously wounded in a car accident. However, within information security, the attacker is quite central when determining both the probability and severity of an attack. If a competitor steals a company's source code, that would be much more severe than if a random teenage hacker did it. We therefore included three new concepts from the OWASP approach to determining risk [OWA08] as part of our model: Technical impact, business impact and threat agents.

The *technical impact* describes the objective effect of the attack, such as compromised source code. However, since the *business impact* of compromised source code may vary considerably, this is treated separately. For some companies it may be devastating while for

Threat	Threat agents
Industrial espionage	competitors, intelligence agencies
Corporate espionage	competitors, suppliers, customers
Economic espionage	suppliers, customers
Terrorism and activism	political, environmental activists
Unintentional disclosure and human mistakes	own and partner employees
Insiders' intentional disclosure	own and partner employees

Table 1: The case study participants most frequently mentioned threats against shared information and their corresponding threat agents

others it may not be that important. On the likelihood, the *threat agent* is the one actively carrying out an attack while the vulnerability is the probability that an attack would be successful.

As part of the interviews we asked the subjects to name the threats to shared information they perceived to be the most severe, the most probable and which actors they considered to pose the greatest threat to their company. In Table 1 we have listed the most frequently mentioned threats (either for severity or probability) with corresponding threat agents. Note however that the lists are unordered with respect the risk they pose.

Espionage is clearly something that the case study participants are worried about. We have separated three different kinds of espionage based on the underlying motive and the kind of information subject to the threat. *Industrial espionage* considers misuse of technical information, Intellectual Property Rights (IPR), innovations and systems. The severity of this kind of threat was perceived to be high since it potentially could reduce the company's competitive advantage towards competitors or foreign nations. *Corporate espionage* considers misuse of corporate information such as strategies, price information, priorities and such. This threat could also reduce the competitiveness of the company as the threat agents could adapt their behaviour. *Economic espionage* considers information on the financial situation of a company such as production volumes and contract negotiations. Misuse of financial information may affect stock market and prices, and hence is of high importance to companies even if threats may not directly harm the company.

Terrorism and activism was mentioned by several as something to be aware of, but all of the interviews regarded successful attacks to be highly unlikely. Still, the severity could potentially be high in terms of damaged reputation.

Unintentional disclosure and human mistakes was also quite frequently mentioned. One of the subjects even claimed that *"if we could only get rid of the human mistakes, we would probably reduce the incident occurences by 80%"*, although not all of these incidents relating to shared information. It is therefore important to consider the amount of information that is compromised due to mishaps by employees. The severity of the disclosure may of course vary considerably, but in worst case it may be very high and pave the way for any type of espionage. *Insiders* may cause similar impact, but are distinguished by the fact that information disclosures are intentional.

4.4 Usage control technology appraisal

Through the interviews we attempted to identify the key factors of the coping response, in our case usage control technology, that would influence its adoption. As mentioned previously, the four main groups of factors are *relative advantage*, *response efficacy*, *self-efficacy* and *response costs*. All but the first are also part of PMT, although with slightly different contents.

Relative advantage was added as a factor since we believed it to be crucial that usage control technology be superior to the mechanisms that are currently implemented in the organisations. Although we acknowledge that some of this could be expressed through response efficacy, it is important to show the perceived superiority of the new approach. As part of the case study we asked respondents to identify existing measures taken to control the use of information they shared with others. Inline with our assumptions, the primary protection measure employed was the use of Non-Disclosure Agreements (NDAs) or other contractual measures to safeguard confidentiality. NDAs were used at very different level of granularity. Some were signed by the management group as a contract between companies, whereas others would be signed by individual employees on individual projects. Thus, a considerable amount of information handled by employees are governed by an NDA they have never seen. Instead, it is assumed that the companies' security policy will ensure proper conduct. One of the companies involved in the case study had implemented the Microsoft Rights Management System [Mic03], but so far it had only been used internally on a hand-full of projects. None of the respondents considered the use of NDAs to be burdensome in any way, and did not think of it as reducing their job efficiency. However, they all noted that relying solely on legal measures may not be adequate to protect against the identified threats. Or as one of the participants stated: *"We have a lot of socio-cultural and legal control, but we realise that we need more technological control"*.

The case study participants did not agree on the *response efficiacy* of usage control technology, but it is evident that this is an important factor affecting the adoption of technology. Particularly whether the technology should be proactive (prevent misuse) or reactive (detect misuse) seemed to be an area of debate. Also, as noted by one of the participants, *"if there is a way to bypass the usage control technology, people will do it and start distributing unprotected versions instead"*. The *response costs* are very much related to the response efficacy. That is, by being too strict, the usage control technology could make important information unavailable and hence greatly reduce the job performance. Particularly since, as one participant noted *"we have currently a lot of problems categorising information, perhaps only 50% of all documents have correct classification"*. Thus, if applying strict preventive usage control to information that has not been accurately categorized would result in a lot of information being unavailable. Further, participants claimed that the administrative burden caused by having to maintain usage rights and monitor usage policy violations would definitely impose additional costs for the companies.

Self-efficacy and the practical issues of day to day use were stressed by several of the participants. *"It is important to strike a balance between protecting information and getting the job done"*. Participants were also concerned that specifying the usage policy would be troublesome, if it would have to be done independently for each document made. This

might be mitigated if the usage control policy somehow could be inferred from the access control policy of the document management systems currently deployed (*compatibility*).

5 Discussion and conclusion

Our adapted model has yet to be verified through proper evaluations. Until then, existing research on protection motivation for security technology at least indicate circumstantial evidence for its fitness of use. Additionally, the protection motivation theory have been proven quite successful in health protection measures [MSO00].

We further argue that it is virtually impossible to speak of any *relative advantage* of security technology without at the same time considering the threats or risks they are to mitigate. Hence, the usage control technology can only be perceived to be useful if there is a threat from which it can protect its users.

One may argue that the entire process of understanding how usage control technology is adopted is very similar to conducting a mere risk analysis. To a certain extent it is true, it does resemble a risk analysis. However, the crucial point in understanding protection motivation lies in the individuals and the focus on users' perception of risk and threats. The risk assessment is often the security professionals opinion, whereas the protection motivation is the lay-men's individual perception.

We have thus far concentrated on the primary elements of protection motivation and the attributes of the innovation as the factors influencing usage control adoption. For future revisions of our model we do however believe factors outside the technology to also be highly relevant when explaining or predicting the adoption of usage control technology. Usage control technology is to certain extent similar to Digital Rights Management - a technology that continues to be controversial - and therefore the concepts of *attitude* and *subjective norm* from TRA could be important. Further, it also seems that previous major changes that have occurred within the oil and gas industry have either been driven by the oil companies or the industry associations, and so the role of *change agents* in the theory of innovation diffusion could provide additional input. Although the case study material contain some information on these aspects, it remains to be further analysed and interpreted. Finally, we are also working on having the model verified through an experiment on a larger population within the oil and gas industry.

This paper has shown the usefulness of technology adoption theories in order to better understand the users' perception of protection technology. In the future we will continue to develop our initial model and will also subject it to evaluation by the industry.

References

[Ajz91] Icek Ajzen. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2):179–211, 12 1991.

[BCB10] Burcu Bulgurcu, Hasan Cavusoglu, and Izak Benbasat. Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness. *MIS Quarterly*, 34(3):523–548, 2010.

[CMG09] T. Chenoweth, R. Minch, and T. Gattiker. Application of Protection Motivation Theory to Adoption of Protective Technologies. In *System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on*, pages 1 –10, jan. 2009.

[Dav89] Fred D. Davis. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3):319–340, September 1989.

[FA75] Martin Fishbein and Icek Ajzen. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Series in Social Psychology. Addison-Wesley Publishing, 1975.

[HR09a] Tejaswini Herath and H. R. Rao. Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness. *Decision Support Systems*, 47(2):154–165, 5 2009.

[HR09b] Tejaswini Herath and H Raghav Rao. Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, 18(2):106–125, April 2009.

[HWF05] George E. Higgins, Abby L. Wilson, and Brian D. Fell. An Application of Deterrence Theory to Software Piracy. *Journal of Criminal Justice and Popular Culture*, 12(3):166–184, 2005.

[lFPDR00] Donna l. Floyd, Steven Prentice-Dunn, and Ronald W. Rogers. A Meta-Analysis of Research on Protection Motivation Theory. *Journal of Applied Social Psychology*, 30(2):407–429, 2000.

[MB91] Gary C. Moore and Izak Benbasat. Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, 2(3):192–222, September 1991.

[McC02] Bill McCarthy. New Economics of Sociological Criminology. *Annual Review of Sociology*, 28:417–442, 01 2002.

[Mic03] Microsoft Corporation. Microsoft Windows Rights Management Services for Windows Server 2003 - Helping Organizations Safeguard Digital Information from Unauthorized Use. Technical report, Microsoft Corporation, 2003.

[MSO00] Sarah Milne, Paschal Sheeran, and Sheina Orbell. Prediction and Intervention in Health-Related Behavior: A Meta-Analytic Review of Protection Motivation Theory. *Journal of Applied Social Psychology*, 30(1):106–143, 2000.

[Nyr11] Åsmund Ahlmann Nyre. Usage Control Enforcement - A Survey. In A Min Tjoa, Gerald Quirchmayr, Ilsun You, and Lida Xu, editors, *MURPBES*, volume 6908 of *Lecture Notes in Computer Science*, pages 38–49. Springer, 2011.

[OWA08] OWASP. OWASP Testing Guide v3. Technical report, The Open Web Application Security Project, 2008.

[PS04] Jaehong Park and Ravi Sandhu. The UCON$_{ABC}$ usage control model. *ACM Trans. Inf. Syst. Secur.*, 7(1):128–174, 2004.

[Rog75] Ronald W. Rogers. A Protection Motivation Theory Of Fear Appeals And Attitude. *Journal of Psychology*, 91(1), 1975.

[Rog03] Everett M. Rogers. *Diffusion of Innovations*. Free Press, 5th edition, 2003.

[Sco00] John Scott. Rational Choice Theory. In A. Halcli G. Browning and F. Webster, editors, *Understanding Contemporary Society: Theories of The Present*. Sage Publications, 2000.

[SPM06] Mikko Siponen, Seppo Pahnila, and Adam Mahmood. Factors Influencing Protection Motivation and IS Security Policy Compliance. In *Innovations in Information Technology, 2006*, pages 1 –5, nov. 2006.

[SPM07] Mikko Siponen, Seppo Pahnila, and Adam Mahmood. Employees' Adherence to Information Security Policies: An Empirical Study. In Hein Venter, Mariki Eloff, Les Labuschagne, Jan Eloff, and Rossouw von Solms, editors, *New Approaches for Security, Privacy and Trust in Complex Environments*, volume 232 of *IFIP International Federation for Information Processing*, pages 133–144. Springer Boston, 2007.

[TK82] Lous G. Tornatzky and Katherine J. Klein. Innovation Characteristics and Innovation Adoption-Implementation: A Meta-Analysis of Findings. *IEEE Transactions on Engineering Management*, 29(1):28–43, February 1982.

[VD00] Viswanath Venkatesh and Fred D. Davis. Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, 46(2):186–204, 2000.

[VMDD03] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3):pp. 425–478, 2003.

[WTL05] I. Woon, G. Tan, and R. Low. A Protection Motivation Theory Approach to Home Wireless Security. In *Proceedings of the Twenty-Sixth International Conference on Information Systems*, pages 367–380, 2005.

SESSION 5

Technology for Cloud Computing and Future Internet

Lazy Disclosure – Mixing Cloud and Local Storage

David Bouck-Standen, Jörg Cassens and Michael Herczeg

Institute for Multimedia and Interactive Systems
University of Luebeck
Ratzeburger Allee 160
D-23562 Lübeck, Germany
{bouck-standen,cassens,herczeg}@imis.uni-luebeck.de

Abstract: The use of multiple computer systems by a single user makes data transfer a necessity. Cloud storage solutions provide synchronisation features and file accessibility across systems but the exact location and type of storage varies between services. Many systems rely on a client-server architecture where the data is held at the data centers of the service provider. This leads to several difficulties for the end user, for example with regard to privacy or data security. This is countered by the convenience of highly available servers. If the user wants to minimise exposure to the challenges of Cloud-based services, then he is required to either host a server infrastructure himself or to rely on inconvenient on-demand synchronisation by means of file transport protocols or even physical storage devices. The approach presented here tries to harvest the positive aspects of using Cloud-based services while keeping the user in control of his data. To this end, we focus on a hybrid system of a client-server and a peer-to-peer architecture which expands the Cloud to the user's system giving him full control over which data he selects to be stored on a server in the Cloud and which data he prefers to keep private on personal computing devices while ensuring unified access to both types of storage over the network.

1 Introduction

Cloud computing is best described as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources like for instance networks, servers, storage, applications and services, that can be rapidly provisioned and released with minimal management effort or service provider interaction [MG09].

Cloud storage solutions allow users to store their data online making files accessible from any location with an Internet connection [AFG+09]. These solutions are also utilisable for synchronisation, collaboration or backup purposes.

Uploading data requires trust in the provider [CKS09] and his abilities to ensure the safety and security regarding both storage of and access to the uploaded data. The data's storage location is concealed within the Cloud and the servers storing the data may not

be run in the same country the user resides in. This also hints at legal issues as laws governing data protection and data security vary between countries.

Most Cloud storage solution providers as i.e. Amazon S3[1], Microsoft SkyDrive[2] or Dropbox[3] offer a software tool which once installed on a client system manages the file uploads. When configuring the tool it is common practise to select a root folder on the client system. For Cloud synchronisation or upload, the user is required to move files and folders into this root folder.

We propose an integrated hybrid solution which allows the user to keep files on his personal computers and in the folder structure he is accustomed to. The user is free to choose whether to synchronise the selected folders with a Cloud storage, giving convenient access in a client-server architecture, or to make the files available to other computers without uploading them into the Cloud. In the latter case, only metadata is synchronised with a server, making it available to other computers hence providing a searchable file index. The actual file content is fetched directly from the client system that makes the file available. This minimises the user's exposure to the Cloud.

2 Cloud Storage

Internet usage continues to increase throughout Europe. Since 2006 the number of households with Internet access grew by 39% to a total of 70% in the EU27 countries, 87% served by a broadband connection [SL10]. Fixed broadband subscriptions also increased in the U.S. by 29% and Australia by 37% since 2005 [Ofc11] indicating a clear international trend toward increased global networking and Internet usage.

In the course of growing Internet usage, a demand for storage solutions is present. This finding is supported by two independent surveys carried out by our research group in 2011. The first survey was concerned with a mobile client for the Network Environment for Multimedia Objects (NEMO) [LCHS10], the second with an augmented reality system. Both systems use the NEMO system for Cloud storage [Sto10].

The first survey shows that 88.4% of the respondents are synchronising files and folders between systems of which 63.0% are using a direct connection to i.e. network storage, external hard drive or a USB-Flash drive [Kha11]. Only 37.0% of the respondents are actively using a Cloud storage service. In the second survey 60.4% of the respondents answered having heard of the term Cloud [Ebe11].

These findings are supported by the latest survey of user habits of Spanish Internet users where 10.6% of 34.096 respondents said they had used FTP for file transfer on the day before the survey [AIM12]. In addition, the latest OfCom report states the increasing interest of Cloud computing to public bodies [Ofc11] even towards a European strategy for Cloud computing [Kro12].

[1] https://aws.amazon.com/s3/
[2] https://skydrive.live.com/
[3] http://www.dropbox.com/

2.1 User-Centered Design

In the following we take a user centred design process as a basis. By presenting a paper prototype to a potential user in an interview, the interviewee has the option to make remarks, changes and additions to the prototype [STG03], [TBBS06b], [TBBS06a].

The concept focuses on users who are accessing their data from multiple systems, are often on the move and cannot take all their data along with them at all times, or need to synchronise their data between systems. In the process of our research, a number of interviews were completed with users representing possible target groups of users [BS11]. The interviews were conducted with a university professor, members of mid- and high-level management as well as university students. The following vision scenario makes use of a persona [CRC07] derived from key characteristics of the interviewee's.

2.2 Scenario

Phil McNeil is a project manager with a company in the oil-business. Being a commuter, Phil has to travel an hour each direction every day. On his way to work he listens to the radio and notices a particular song. Having arrived at his office, he uses a computer system accessing his data stored inside the Cloud. The recent changes he made to a presentation at home yesterday evening were downloaded automatically at system start-up and Phil was able to continue work right where he left it, being also able to see the latest drafts his team finished this morning.

During a break, Phil remembers the song he listened to earlier on. Searching his personal audio library within the Cloud, he notices he had already purchased that album. He then downloads the entire album to his MP3-Player.

Being in a management position bearing responsibility for various projects, he often has to travel, even on short notice. In the afternoon Phil gets summoned to a conference he has to attend. In this case, the data stored inside the Cloud enables him to access all his files regardless whether they have been synchronised and uploaded before. This way Phil can keep track of his files without having to download everything to his laptop, he usually takes to conferences, which has only a limited storage capacity.

Sitting at the gate prior to boarding, Phil chooses a movie from his library stored on his computer at home he might want to watch this evening at the hotel. He notices that his computer at home is switched off and cannot transfer the file immediately but the system saves his request. Having arrived at the hotel, Phil switches on his notebook, which automatically downloads the movie through the Cloud completing his request. The computer at home has in the meantime been switched on by his wife and has, according to his earlier request, automatically uploaded the data into a cache storage location within the Cloud.

3 Concept and Implementation

The outlined scenario describes a hybrid system consisting of a Cloud storage solution enabling a client to up- or download data in a client-server context as well as a peer-to-peer context. In client-server context the data is highly available from the Cloud as all data is uploaded and stored inside the Cloud. Data which is available in a peer-to-peer context may not be obtained at all times, but is mainly stored on the client's system. As the trend of the number of broadband connections is showing an increase, connection speeds are negligible.

3.1 Concept

In our concept the user selects one or more folders on the client system which are enabled for synchronisation or made available. These locations will be referred to as repositories.

The content of each repository is represented by a catalogue containing information about files and folders of the repository, their size, meta-data, file type and preview. Settings regarding repository access and sharing policies are also set within the catalogue. A catalogue of a repository enables the user to browse and search the repository without having to download all its content.

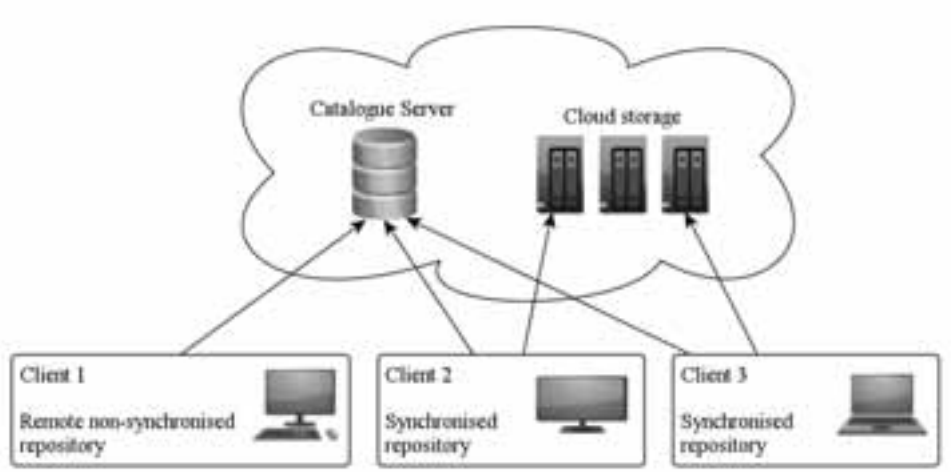


Figure 1: Client 1 holds a remote non-synchronised repository connecting to a catalogue server in the Cloud. Both Client 2 and Client 3 hold a synchronised repository. Catalogue information on the repositories as well as the data is stored inside the Cloud.

As shown in Figure 1, for this concept the Cloud consists of two components: a Catalogue Server and Cloud storage. The latter is primarily serving as storage space for repository contents. This storage is accessible through an Internet connection with high availability. The Catalogue Server represents a central storage for all catalogues. Each catalogue is stored under a unique identification reference identifying the repository.

Two types of repositories are distinguished:

1. **A synchronised repository** is automatically synchronised with the Cloud storage and is therefore available in a client-server infrastructure. Features like for instance version control or multi-user collaboration[4] when working inside a repository are available. Synchronisation between multiple systems accessing the same repository, as shown in Figure 1 is also supported. A synchronised repository is accessible as long as the storage location is available inside the Cloud.

2. **A remote non-synchronised repository** contains data which is available in a peer-to-peer infrastructure. A Catalogue Server stores information on repositories and their location, as they are hosted by other clients which generate, maintain and upload catalogue information. Hence data of remote non-synchronised repositories is not stored in Cloud storage it is available on demand only.

A client's system works with both types of repositories at the same time. As outlined below, the concept is implemented in a client application which serves two purposes.

When working with a synchronised repository, the user's interaction with the application is minimised following initial setup. The user can continue working with his system and accessing files and folders as before. This therefore reduces initial learning effort [MP97]. In-depth interviews with users representing potential groups of users showed that a Cloud storage system should have minimal influence on the daily routine of user interaction. Such systems should therefore not require the user to i.e. change his custom workflow or how he structures his files and folders. These findings are also supported by [LED+ 99].

When working with a remote non-synchronised repository, the client application is used in order to browse, search or download files and folders. In this case the user is required to use the client application. During the in-depth interviews, the interviewees understood the application as a tool for browsing contents of remote systems and expressed the need for an interaction with this storage distinguishable from their files and folders stored locally. For this reason remote non-synchronised repositories are not i.e. mapped into the user's local file system.

3.2 Interaction Design

During the initial setup of either a synchronised or a remote non-synchronised repository, the user selects a directory of his choosing from his local system which serves as repository root folder. In-depth interview findings support the need for flexibility. Most of the systems we have referenced before work with one dedicated root

[4] This concept focuses on the technical sub-system whereas collaboration features can be augmented to extend functionality in the social sub-system [Her09] to serve i.e. as group communication tool reflecting social relationships of the user.

folder forcing the user to move his data into this particular directory whereas our concept allows the selection of any folder[5]. The selection of a directory as repository is indicated by a different folder icon. In-depth interviews prove this to be the most lucid way of displaying the folder's status.

After the initial setup the client application generates and submits a catalogue to the Catalogue Server. Depending on the type of the repository, data is transferred and stored in the Cloud storage, as outlined above.

3.3 Architecture

In order to keep track of the contents of and changes to a repository, the catalogue of a repository needs to be updated in the event of the creation of, changes to, or the deletion of a file. The client application monitors file events, updates the local catalogue and submits catalogue changes to the Catalogue Server which distributes change notifications to other client systems also connected to the respective repository.

With regard to remote non-synchronised repositories, data availability depends on the connection state of the client hosting the remote non-synchronised repository. As the catalogue of a remote non-synchronised repository is stored on a Catalogue Server within the Cloud, an authorised client may browse and search the repository independently of the connection state of the client hosting the data without having to download all its content.

Figure 2 outlines the retrieval protocol for a repository request. Depending on the type of the repository, the data resides either in Cloud storage or on a client's system. Upon a request, the location of the requested data is resolved. Based on the assumption that Cloud storage serves high availability data stored in the Cloud is available directly on request. Cloud storage locations are managed by the Catalogue Server which holds an index and tracks usage quotas of possible storage locations as i.e. local server storage, FTP locations or commercially available Cloud storage locations.

If the request was directed to a file of a remote non-synchronised repository, through catalogue information one or more clients hosting the repository are located. If no system hosting the repository is available through an active Internet connection the request is cached until a host is available, as Figure 2 illustrates. Once one or more clients are available, data is directly transferred to the requesting client system via a peer-to-peer connection. If the requesting client is unavailable in the meantime, the file is cached inside the Cloud and will be downloaded when the client is back online.

The actual concept relies on Cloud storage providing fast and highly available storage space. In the future a peer-to-peer based infrastructure serving distributed storage will supersede today's Cloud storage which not only faces legal regulations but governmental encroachments being the reason for companies as well as individuals not to rely on Cloud storage solutions [JLGS09]. With the Catalogue Server, the outlined concept only

[5] Excluding system folders and folders which are already part of a repository.

stores information on files but not file contents which will remain on the client system only.

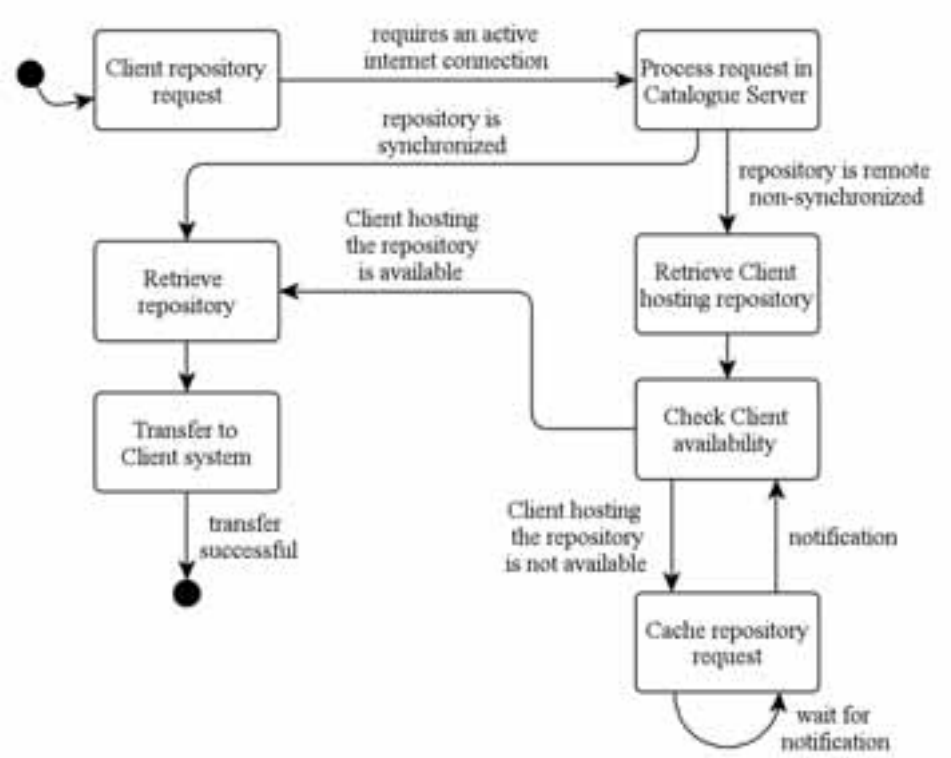


Figure 2: The chart illustrates the retrieval protocol for a repository request. A client repository request is submitted to a Catalogue Server inside the Cloud. The Catalogue Server resolves the repository type and checks the repository's availability. A remote non-synchronised repository is not available from Cloud storage and transferred as long as the Client hosting the respective repository is connected to the Internet.

3.4 Implementation

Our first prototype consists of a client application and a standalone server system written in .NET 3.5. Using MONO[6] allows cross platform deployment of the client application.

The client application was formatively evaluated during in-depth interviews throughout the process of implementation by using a paper prototype as outlined above. In addition, a final summative evaluation was carried out as well. The first component of the client application is a service which runs in the background monitoring the repositories handling catalogue generation and actualisation as well as catalogue and file transfers.

[6] http://www.mono-project.com /

The second component consists of a graphical user interface. When working with synchronised repositories there is no need to use a client application. The user can continue working with his files and folders as he is accustomed to using his preferred file browser. However data of remote non-synchronized repositories is not stored on every client system. To browse the catalogue and thereby the repository, the user is required to use the client application which allows for files to be found or requested for download if the respective host is unavailable and downloaded directly if the host is available. This allows for a lazy and incomplete synchronisation with several locations without having the need for explicit synchronisation points.

The server has been implemented serving a RESTful database-assisted API [FT02], [Rod08] not only offering a stable and flexible service-based infrastructure but an interface for future client development for i.e. mobile devices.

4 Summary and Conclusions

The concept outlines a catalogue-assisted multi-user file storage and collaboration system which is designed as a hybrid between a highly available but from the user's perspective in-transparent Cloud storage solution and a considerable more transparent but possibly less available access to local storage in a peer-to-peer infrastructure.

The surveys conducted reveal a general interest in Cloud storage and furthermore an active demand for storage solutions. The in-depth interviews show the innovative concept presented within this contribution to be understandable, usable and acceptable. A world-wide trend towards broadband connections supports the findings of storage space distributed throughout client systems to supplement highly available Cloud storage.

In context with the scenario presented in this article, various use-cases exist supporting possible applications for this hybrid storage system. Based on our prototypical implementation and our first empirical findings, a long term evaluation is under preparation. Future ideas reflect upon both architectural as well as client-side improvements. For instance improving file search by using the meta-data of files in a faceted context [YSLH03] may prove time-saving and more convenient for the user.

Protecting the user's privacy and data against access by any third party is essential. While the term privacy is ambiguous and can be used in different contexts, we consider the user's personal data to be sensitive. Nissenbaum identifies three concerns towards privacy: (1) monitoring and tracking, (2) dissemination and publication, and (3) aggregation and analysis [Nis09]. These challenges have to be taken into account when designing Cloud-based storage solutions.

Our approach allows the user to decide which data to upload to Cloud-based servers, making the data more vulnerable to be exposed to third parties, and which data to keep on computing devices under his control, better protecting it. The user is made aware of

this difference by the need to use the specific client for remote non-synchronised repositories.

For the time being, data of remote non-synchronised repositories is still cached in the Cloud when transferred to another client. This will not be necessary once we move to a full peer-to-peer solution for data exchange. The catalogue data would be the only data still vulnerable to attacks. One could distribute the catalogue data between clients as well, but there is no guarantee that the information is up to date at all times. This is a trade-off worth investigating further.

Acknowledgments

The authors express their thanks to both Syavoosh Khabbazzadeh and Björn Eberhardt for providing data surveying Cloud storage in different contexts.

References

[AFG+09] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. Above the Clouds: A Berkeley View of Cloud Computing. Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley, Feb 2009.

[AIM12] AIMC. Navegantes en la Red – 14a encuesta AIMC a usarios de Internet. Technical report, Association of Communication Media Research, February 2012.

[BS11] David C.E. Bouck-Standen. The Davebox – A Cloud-Based File Repository System. Bachelor thesis, Institute for Multimedia and Interactive Systems, University of Luebeck, November 2011.

[CKS09] Christian Cachin, Idit Keidar, and Alexander Shraer. Trusting the cloud. SIGACT News, 40:81–86, June 2009.

[CRC07] Alan Cooper, Robert Reimann, and David Cronin. About Face 3: The Essentials of Interaction Design. Wiley Publishing, Indianapolis, IN, USA, 3rd edition, 2007.

[Ebe11] Björn Eberhardt. Augmented Reality an gedruckten Dokumenten. Bachelor thesis, Institute for Multimedia and Interactive Systems, University of Luebeck, 2011.

[FT02] Roy T. Fielding and Richard N. Taylor. Principled design of the modern Web architecture. ACM Trans. Internet Technol., 2:115–150, May 2002.

[Her09] Michael Herczeg. Software-Ergonomie. Theorien, Modelle und Kriterien für gebrauchstaugliche interaktive Computersysteme. Lehrbuchreihe interaktive Medien. Oldenbourg, München, 3., vollst. überarb. und erw. Aufl., 2009.

[JLGS09] Paul T. Jaeger, Jimmy Lin, Justin M. Grimes, and Shannon N. Simmons. Where is the cloud? Geography, economics, environment, and jurisdiction in cloud computing. First Monday, 14(5), 2009.

[Kha11] Syavoosh Khabbazzadeh. Entwicklung eines NEMO-Clients für mobile Geräte mit Android-Betriebssystem. Bachelor thesis, Institute for Multimedia and Interactive Systems, University of Luebeck, 2011.

[Kro12] Neelie Kroes. Towards a European Cloud Computing Strategy. Speech, World Economic Forum, Davos, February 2012. Neelie Kroes is the Vice-President of the European Commission responsible for the Digital Agenda.

[LCHS10] Sebastian Lob, Jörg Cassens, Michael Herczeg, and Jan Stoddart. NEMO – The Network Environment for Multimedia Objects. In Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia – IITM 2010, pages 245–249, New York, NY, USA, December 2010. ACM.

[LED+99] Anthony Lamarca, W. Keith Edwards, Paul Dourish, John Lamping, Ian Smith, and Jim Thornton. Taking the Work out of Workflow: Mechanisms for Document-Centered Collaboration. In Proceedings of the European Conf. Computer-Supported Cooperative Work ECSCW '99, pages 1–20. Kluwer, 1999.

[MG09] Peter Mell and Tim Grance. The NIST Definition of Cloud Computing. National Institute of Standards and Technology, 53(6):50, 2009.

[MP97] Gloria Mark and Wolfgang Prinz. What Happened to our Document in the Shared Workspace? The Need for Groupware Conventions. In Human-Computer Interaction INTERACT '97, pages 412–420. Hall, 1997.

[Nis09] Helen Nissenbaum. Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press, 2009.

[Ofc11] Ofcom. International Communications Market Report 2011. Technical report, Ofcom, December 2011.

[Rod08] Alex Rodriguez. RESTful Web services: The basics. Technical report, IBM developerWorks, New Orchard Road, Armonk, New York 10504, USA, 2008.

[SL10] Heidi Seybert and Anna Lööf. Internet Usage Statistics 2010, eurostat, 2010.

[STG03] Reinhard Sefelin, Manfred Tscheligi, and Verena Giller. Paper prototyping - what is it good for?: a comparison of paper- and computer-based low-fidelity prototyping. In CHI '03 extended abstracts on Human factors in computing systems, CHI EA '03, pages 778–779, New York, NY, USA, 2003. ACM.

[Sto10] Jan Stoddart. Entwicklung einer Serverkomponente für das ambiente System NEMO. Master's thesis, Institute for Multimedia and Interactive Systems, University of Luebeck, 2010.

[TBBS06a] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. Getting the right design and the design right. In Proceedings of the SIGCHI conference on Human Factors in computing systems, CHI '06, pages 1243–1252, New York, NY, USA, 2006. ACM.

[TBBS06b] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. User sketches: a quick, inexpensive, and effective way to elicit more reflective user feedback. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles, NordiCHI '06, pages 105–114, New York, NY, USA, 2006. ACM.

[YSLH03] Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. Faceted metadata for image search and browsing. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '03, pages 401–408, New York, NY, USA, 2003. ACM.

Drizzle: The RAIN Prototype

Martin Gilje Jaatun[1], Christian Emil Askeland[2] and Anders Emil Salvesen[2]

[1]Department of Software Engineering, Safety and Security
SINTEF ICT
Trondheim, Norway
Email: Martin.G.Jaatun@sintef.no

[2]Dept. of Telematics, NTNU
Trondheim, Norway

Abstract: Internet communities are moving to the Cloud, but in addition to the advantages regarding cost and convenience, this also means that cloud service providers are increasingly in a position to aggregate large amounts of personal data, which means that it is becoming prudent to develop mechanisms that can contribute to limiting the information available to providers. In this paper we present a prototype cloud security solution for protection against an "honest but curious" cloud provider. The solution is based on splitting up data and distributing it to multiple cloud providers, without encrypting the individual pieces. Our initial tests indicate that our solution is sufficiently efficient for normal use.

1 Introduction

Internet communities are increasingly employing Cloud Computing; social networks such as Facebook may have started out as hosted applications, but are now moving in the direction of offering cloud-based Software-as-a-Service (SaaS) and even Infrastructure-as-a-Service (IaaS) to external customers [RNJ12].

Cloud Computing is a tantalizing concept that promises flexible computing solutions at an affordable price, without requiring large up-front investments in computer hardware. However, since Cloud Computing can be considered the ultimate evolution of outsourcing, there is lingering concern in the target groups when it comes to security – we trust cloud providers to take care of our holiday snapshots, but many stakeholders are not ready to "bet their business on the cloud". To quote Microsoft's John Howie: "...there are things that will never go into Azure, for example, our SAP back end." [GHR+10]

Since internet community members are predestined to disseminate a lot of (not necessarily sensitive) information through community interaction, it becomes even more important to have an improved mechanism for storing files in the cloud, without adding to the data aggregation problem. Many people have tried to tackle the problem of how to securely store files in an untrusted location. Proposed solutions include fully homomorphic encryption [Gen09], secure multiparty computation [BCD+09] and trusted computing using Trusted

Platform Modules [SGR09].

Our path is on a somewhat different tangent, in that we seek to explore a solution that can offer sufficient security through the splitting and dispersion of data in the cloud. The ultimate goal is to achieve this without the use of encryption, both to save processing cost on mobile devices with limited processing power, but also to avoid having to administrate a large number of encryption keys.

2 Related Work

Surprisingly many researchers have ventured to solve secure distributed storage by splitting up data and spreading it in the wind. Singh et al. [SKZ11] present a scheme for *n-out-of-m* secret sharing of data [Sha79], but do not provide an algorithm for the actual splitting of the data to be stored. Another *n-out-of-m* scheme is proposed by Parakh and Kak [PK09], but they do not discuss why their scheme should be better than e.g. the one proposed by Rabin [Rab89]. Furthermore, they do not discuss that cryptographic keys based on passwords tend to be insecure, and that such schemes become even more vulnerable in the event of password reuse. From the field of Grid computing, Luna et al. [LFMB08] present yet another solution based on Rabin's *n-out-of-m* scheme, but add an additional concept of *Quality of Security* (QoSec) to rate individual storage providers. In this manner, they show that the number of partitions (or the n-to-m ratio) can be adjusted depending on the aggregated trustworthiness of the providers.

The RACS system [ALPW10] aims at preventing vendor lock-in and data loss through failures by performing striping of data (in RAID-5 fashion) across multiple cloud providers, but has no privacy or confidentiality goals.

The Mnemosyne [HR02] system offers steganographic storage which "gives a user strong protection against being compelled to disclose (all) its contents." The idea here is not only to hide data from prying eyes, but also to prevent anyone from determining that there is anything hidden in the first place. Mnemosyne encrypts each block, and uses Rabin's Information Dispersal Algorithm [Rab89] to replicate data and avoid data loss. To protect against traffic analysis, random other blocks are also read/written periodically, only to be discarded.

3 RAIN

The Redundant Array of Independent Net-storages (RAIN) was proposed by Jaatun et al. [JNAZ11, ZJV+11, JZA11], using multiple cloud storage and data processing providers in a fairly complex hierarchy. For our prototype we have adopted an architecture very much like the one seen in Figure 2b in the original paper [JNAZ11], except that we have focused solely on the storage part. The resulting architecture is seen in Figure 1.

For our solution, the Command & Control (C&C) node and the RAIN Dashboard are both

installed on a server within an organisation, maintaining the only database of where in the cloud the file parts are stored. This server runs on Apache with SSL for secure uploading of files within the organisation. Placing it in the cloud will allow for better scaling, and will also permit access from multiple locations.

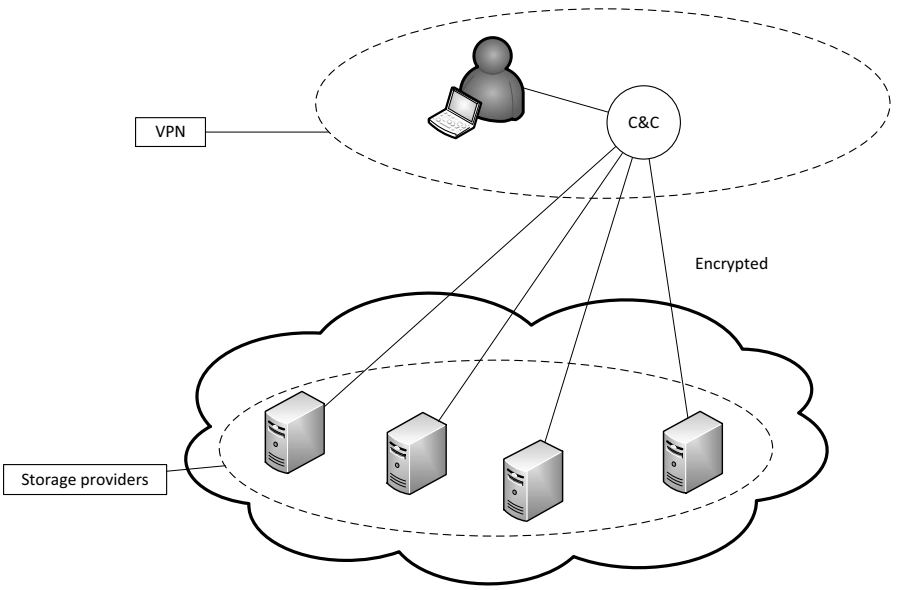


Figure 1: The C&C within a single organisation, communicating with different storage providers.

4 Implementing the C&C

The C&C is the single point of entry for users of RAIN. Users interact with the C&C to split and merge, upload and download files to and from the cloud. It maintains a database, storing all information on files uploaded through it, mainly: *What files were uploaded when by whom?* and *To which providers are the different chunks uploaded?*

The C&C is developed in Python using Eclipse with the Pydev-extension, including the following packages:

- Python 2.7.1
- SQLAlchemy 0.7.3
- SQLite3
- python-cloudfiles 1.7.9.1
- boto 2.1.0

- Django 1.3.1

The C&C consists of the following components:

- The Receiver, providing the API for different User Interfaces.
- The CompleteFile- and Chunk-objects which are mapped to the database.
- A splitter-class, providing the splitting and merging of the files.
- Interfaces to interact with the different storage providers.

4.1 The Receiver

The receiver provides an API for the user interfaces to interact with the C&C. The main functions are:

- showAllFiles – returns a list with all the CompleteFile-objects in the database.
- sendFile – takes a file as argument, splits the file and uploads the chunks.
- getFile – takes the filename as argument, initiates the download of all the chunks belonging to the CompleteFile.
- decider – analyses the argument and decides whether to upload or download, and checks if the file already exists in the database.

4.2 The Splitter

The splitter splits and merges a CompleteFile. The split process read a file, and splits it into a predefined number of chunks, which are stored temporarily on the disk. It returns a list of all these chunks. The merge process does the exact opposite, takes in all chunks of a CompleteFile, assembles them, and stores the result to disk. It then returns the path to the file.

The splitting process is a critical part of the system. To split the data we use a simple algorithm which loops through the files and samples every n-th byte into a new file (chunk) where n is the number of chunks. We store the chunks in temporary files in the C&C before we send them to the cloud storage providers.

4.3 The Storage Provider Interfaces

Most cloud storage providers offer a solution to interact with the platform in different programming languages. Currently we've implemented interfaces for Swift, the storage

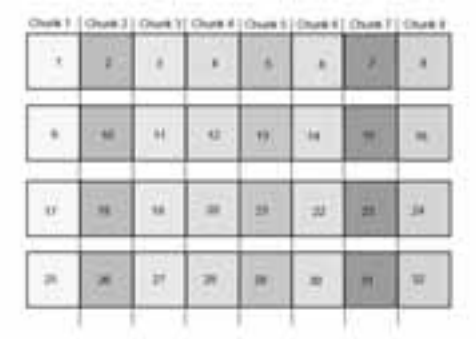


Figure 2: A file is split into 8 chunks. All parts with the same color goes in the same chunk.

solution in OpenStack, and Amazon S3. These implement the following functions:

- sendChunksInList
- getChunksInList

5 Process Description

The process of uploading a file is illustrated in Figure 3.

1. A user uploads a file.
2. The dashboard makes sure it is stored in the correct folder, and sends the file through to the C&C-receiver.
3. The C&C-decider first checks if the file exists in its database. If it doesn't, the file is handed to the sendFile-method.
4. In sendFile, all info of the file is serialized to a CompleteFile-object.
5. The desired splitting algorithm is used to split the file into the desired number of chunks.
6. The chunks are dedicated their cloud storage provider, and are uploaded in threads by each provider.
7. When all chunks are transferred successfully, the CompleteFile-object gets a signal, whereupon it saves all info about itself and its chunks in the database.
8. A message is sent to the dashboard and displayed to the user real time.

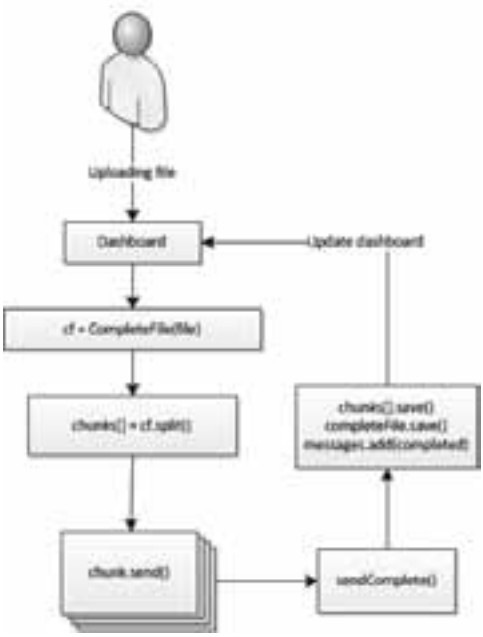


Figure 3: The process of uploading the file

6 Results

The C&C performs IO operations on the file when splitting and merging, and we therefore wanted to measure the IO overhead when uploading and downloading. The preferred protocol for data transfer is HTTP. We have performed measurements with different file sizes and varied the number of chunks to find the best conditions for using the C&C. Overhead is reported on top of 100%, i.e. if sending a complete file takes 100 seconds, and sending all chunks of a certain size takes 150 seconds, the overhead is 50%.

We first uploaded files with different sizes without splitting them, followed by downloading these files again. The uploading and downloading without splitting is the reference we measured the overhead against. We then repeated the procedure with split files. We started by splitting into chunks, then downloading and merging these chunks into a file. Then we proceeded by uploading the same files split into different numbers of chunks.

1. Upload all files in the set to the cloud providers, without splitting.
2. Download these files.
3. Repeat steps 1-2, with splitting and merging the files in a varying number of chunks.
4. Repeat steps 1-3 until the sample space is large enough.

Each upload and download was recorded. After the completion of the benchmark, the average upload and download duration and upload and download bitrate was calculated. The results are shown in Figure 4 and 5 below. The time taken to split and merge was also recorded. Since splitting is done asynchronously with uploading, the split time is not of much value. The merge time, however, is very important since we need to download all chunks before merging them.

Every value presented is a result of a set of minimum ten uploads or downloads, where each value is within 90% of the median value in the result set. When analyzing the data, the values seemed to be so consistent that the result set obtained was deemed large enough to draw conclusions.

6.1 Uploading Results

Figure 4 shows the duration of uploads in seconds. From the graphs we see that more chunks leads to faster uploads. Control measurements have been performed, with 30 and 50 chunks, without including these in the graphs, and they show that the tendency continues.

When looking at the overhead in percent, see table 1, we see that the overhead is off the charts when regarding the 7MB file. If we exclude that file, we notice that the overhead is almost consistent per the number of chunks. We conclude that the average overhead is a little above 140% when the number of chunks is between 12 and 20.

	4 chunks	8 chunks	12 chunks	16 chunks	20 chunks
7 MB	214%	237%	260%	289%	342%
62 MB	151%	148%	143%	145%	147%
175 MB	155%	151%	139%	144%	139%
323 MB	157%	154%	142%	141%	139%
486 MB	-	153%	147%	143%	141%
Average	169%	169%	166%	173%	182%
Average w/o 7MB	154%	152%	143%	143%	141%

Table 1: The overhead in percent when uploading files.

6.2 Downloading

Figure 5 shows the duration of downloads in seconds. From the graphs we see that optimum number of chunks is between 12 and 20 chunks, represented by 16 chunks here. From table 2 we see that the average overhead is 73% when using 16 chunks, so we will operate with this number as the overhead when downloading.

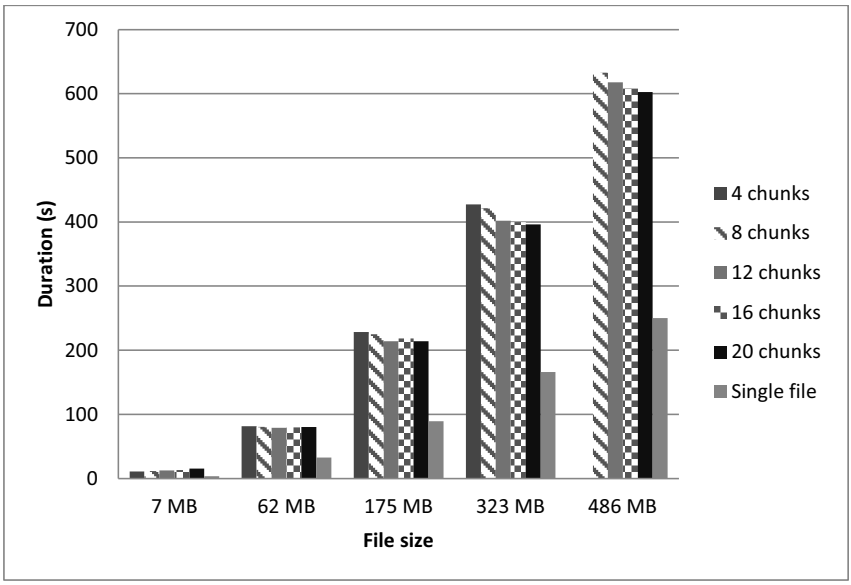


Figure 4: The duration of uploading files of different sizes and different number of chunks to the provider.

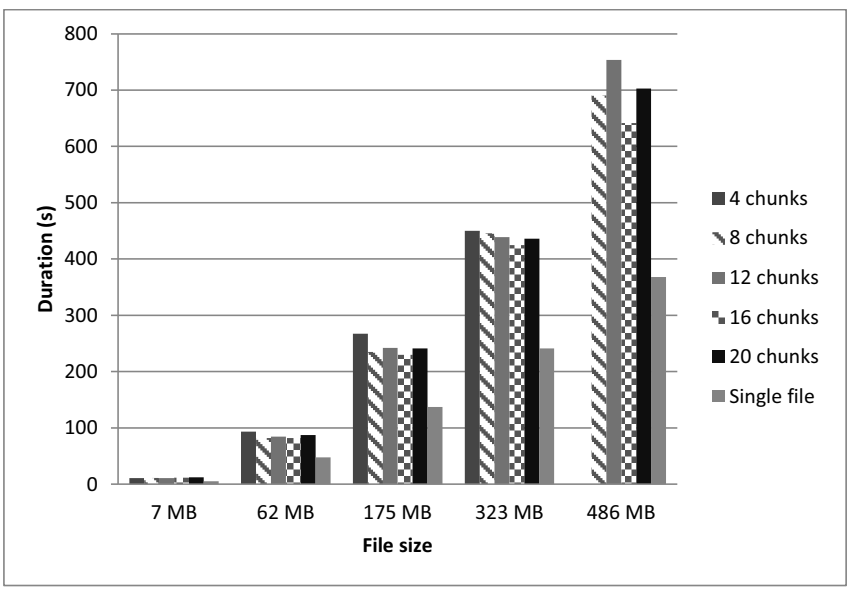


Figure 5: The duration from starting download to all chunks are merged.

	4 chunks	8 chunks	12 chunks	16 chunks	20 chunks
7 MB	103%	106%	102%	122%	130%
62 MB	96%	73%	78%	72%	84%
175 MB	95%	71%	76%	68%	76%
323 MB	87%	85%	82%	76%	81%
486 MB	-	88%	105%	75%	91%
Average	95%	84%	89%	83%	92%
Average w/o 7MB	93%	79%	85%	73%	83%

Table 2: The overhead in percent when downloading files.

7 Discussion

The solution we have presented here has a number of limitations. When we split up the data and send it to different cloud providers, we also lose the possibility to use the cloud for processing of the data. The data has to be merged together in the C&C before it can be processed. We don't have any redundancy in our system, which makes it vulnerable if one of the cloud providers loses a chunk. If the chunk can't be recovered, the data in the original file will be lost. The splitting algorithm is naïve, without any mathematical proofs. In real life, a more sophisticated splitting mechanism would be necessary to prevent providers from extracting interesting information from a single chunk.

In the following we will discuss some possible attacks on our solution.

7.1 Eavesdropping on the Network

If an attacker eavesdrops on the network connection out from the C&C it will be possible for him to collect all the data that are transferred. Collecting the chunks from one file can be done by assuming that chunks sent at approximately the same time, belong together. If the attacker does this, he still has to put the chunks together. If the attacker knows the splitting algorithm and have collected all the chunks from the file, he also needs to know the order of the chunks. If the number of chunks is low, this can be done easily with some trying and failing.

To make it a little harder for the attacker to reassemble the file, there are some countermeasures that can be used. First the C&C can randomize the number of chunks for each file. The randomization will make it harder for attackers to know the exact number of files to reassemble since they don't know how many chunks each file consists of.

Another method is to send file chunks in batches, i.e. let the C&C send chunks from two different files at the same time. If the attacker can't distinguish the chunks from each other it will be a lot harder to reassemble the original files.

A third method is to introduce garbage chunks that are transferred together with the real chunks. Even if the attacker knows that the data contains garbage chunks, he needs to sort

out *which* chunks are garbage in order to be able to obtain the original file.

7.2 Attacking the C&C

The C&C is definitely the most valuable component in the system from an attacker's point of view. If the C&C is exposed, either physically or digitally, the attacker will be able to access the data or perform a denial of service (DoS) attack.

In our system design we propose that the C&C should be within the company's own infrastructure. This means that they have to secure the C&C physically by making sure it is located in an locked room where only authorized people can access it. They also have to make sure that the machine is not remotely accessible through the network. This means that they have to have a proper firewall, authentication, updated software and disabled unused services.

7.3 Attacking the Cloud Provider

An attack on a single cloud provider will not compromise any data, since the chunks don't contain any information by themselves. An attacker has to attack all the cloud providers and find the chunks that belongs together, e.g. by looking at the file creation timestamps. This attack is subject to the same countermeasures as the eavesdropping attack mentioned above.

7.4 Alternative Approaches

Instead of storing unencrypted chunks, it is possible to generate a keystream, XORing the plaintext with this keystream, and then storing both ciphertext and keystream as chunks at disjunct storage providers. This will double the upload/download time and storage space required, and also add some time for generating the keystream for each storage operation. Alternatively, only the ciphertext could be stored, but then this degenerates into an encrypted cloud storage, and all the splitting etc. is pointless.

8 Further Work

Since this is little more than a proof-of-concept prototype, there are abundant opportunities for further work. One obvious area is the splitting algorithm, which needs to be more advanced if an attacker with access to even a few chunks is to be thwarted. One specific improvement may be to randomize the number of chunks for each file. The randomization will make it harder for the cloud provider or attacker to know the exact number of files to

reassemble since they don't know how many chunks each file consist of.

There is also currently no protection against data loss and corruption. One way to approach this might be to employ automatic duplication of cloud storage providers as suggested by Zhao et al. [ZRJS10]. Another solution could be to duplicate the Redundant Array of Independent[1] Disks (RAID) concept of striping blocks across providers, e.g. dividing content among 4 providers, and placing "parity chunks" on a fifth provider, thus compensating from the failure of any one of the other providers.

9 Conclusion

We have shown that implementing a simplified version of the RAIN design is feasible and working. We have also shown that our solution integrates well with open source cloud solutions such as OpenStack, as well as a commercial version like Amazon's S3. This is a great advantage in preventing vendor lock-in, and makes migration from one cloud provider to another easier. We have tested and measured the prototype and we have pointed out some improvements that can be applied when it comes to performance and security.

We believe that privacy concerns may be a major stumbling block for further development of internet communities, and our hope is that the RAIN concept may contribute to alleviating some of those privacy concerns in the future.

Acknowledgements

This paper is partly based on a minor thesis project at NTNU [AS11]. We thank the anonymous reviewers for useful comments.

References

[ALPW10] Hussam Abu-Libdeh, Lonnie Princehouse, and Hakim Weatherspoon. RACS: a case for cloud storage diversity. In *Proceedings of the 1st ACM symposium on Cloud computing*, SoCC '10, pages 229–240, New York, NY, USA, 2010. ACM.

[AS11] Christian Askeland and Anders Emil Salvesen. Secure Use of Public Cloud Services. NTNU minor thesis (fall project), 2011. `http://sislab.no/drizzle`.

[BCD+09] Peter Bogetoft, Dan Christensen, Ivan Damgård, Martin Geisler, Thomas Jakobsen, Mikkel Krøigaard, Janus Nielsen, Jesper Nielsen, Kurt Nielsen, Jakob Pagter, Michael Schwartzbach, and Tomas Toft. Secure Multiparty Computation Goes Live. In Roger Dingledine and Philippe Golle, editors, *Financial Cryptography and Data Security*, volume 5628 of *Lecture Notes in Computer Science*, pages 325–343. Springer Berlin / Heidelberg, 2009. 10.1007/978-3-642-03549-4_20.

[1] Some prefer the variant "Redundant Array of Inexpensive Disks".

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 169–178. ACM, 2009.

[GHR+10] Eric Grosse, John Howie, James Ransome, Jim Reavis, and Steve Schmidt. Cloud Computing Roundtable. *Security Privacy, IEEE*, 8(6):17 –23, nov.-dec. 2010.

[HR02] Steven Hand and Timothy Roscoe. Mnemosyne: Peer-to-Peer Steganographic Storage. In *Revised Papers from the First International Workshop on Peer-to-Peer Systems*, IPTPS '01, pages 130–140, London, UK, 2002. Springer-Verlag.

[JNAZ11] Martin Gilje Jaatun, Åsmund Ahlmann Nyre, Stian Alapnes, and Gansen Zhao. A Farewell to Trust: An Approach to Confidentiality Control in the Cloud. In *Proceedings of the 2nd International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless Vitae Chennai 2011)*, 2011.

[JZA11] Martin Gilje Jaatun, Gansen Zhao, and Stian Alapnes. A Cryptographic Protocol for Communication in a Redundant Array of Independent Net-storages. In *Proceedings of the 3rd IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2011)*, 2011.

[LFMB08] Jesus Luna, Michail Flouris, Manolis Marazakis, and Angelos Bilas. Providing security to the Desktop Data Grid. In *Parallel and Distributed Processing 2008 IPDPS 2008 IEEE International Symposium on*, pages 1–8, 2008.

[PK09] Abhishek Parakh and Subhash Kak. Online data storage using implicit security. *Information Sciences*, 179(19):3323 – 3331, 2009.

[Rab89] Michael O. Rabin. Efficient dispersal of information for security, load balancing, and fault tolerance. *Journal of the ACM (JACM)*, 36(2):335–348, 1989.

[RNJ12] Chunming Rong, Son Nguyen, and Martin Gilje Jaatun. "Beyond Lightning: A Survey on Security Challenges in Cloud Computing". *Computers & Electrical Engineering*, to appear 2012.

[SGR09] Nuno Santos, Krishna P. Gummadi, and Rodrigo Rodrigues. Towards Trusted Cloud Computing. In *HOTCLOUD*. USENIX, 2009.

[Sha79] Adi Shamir. How to share a secret. *Commun. ACM*, 22:612–613, November 1979.

[SKZ11] Yashaswi Singh, Farah Kandah, and Weiyi Zhang. A secured cost-effective multi-cloud storage in cloud computing. In *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, pages 619 –624, april 2011.

[ZJV+11] Gansen Zhao, Martin Gilje Jaatun, Athanasios Vasilakos, Åsmund Ahlmann Nyre, Stian Alapnes, Qiang Ye, and Yong Tang. Deliverance from Trust through a Redundant Array of Independent Net-storages in Cloud Computing. In *Proceedings of IEEE Infocom*, 2011.

[ZRJS10] Gansen Zhao, Chunming Rong, Martin Gilje Jaatun, and Frode Sandnes. Reference deployment models for eliminating user concerns on cloud security. *The Journal of Supercomputing*, pages 1–16, 2010. 10.1007/s11227-010-0460-9.

The Role of Modeling in Future Innovative Business and Community Information Systems

John Krogstie

IDI, NTNU
Trondheim, Norway
krogstie@idi.ntnu.no

Abstract: Future internet systems have a number of properties supporting new innovative business and community systems. Event-driven architectures (EDA) providing varied information to support collaborative decision-making enable more decisions to be made closer to the problem owner. The "Internet of Things" (IoT) enables the Internet to reach out into the real world of physical objects. Mobile and collaborative applications and services utilizing information processing and process support enabled by sensor data from a vast numbers of connected and cheap devices and directly and indirectly from humans in control of these devices will change a number of markets. Future innovative business and community information systems will need to take this situation into account, addressing technological, methodological and conceptual challenges. This paper will focus on the latter, discussing in particular the potential role of model-based techniques and how to assess and improve the quality of models and modelling approaches in this setting using the SEQUAL framework.

1 Introduction

ICT has over the last 30 years gone from obscure to infrastructure. All organizations are dependent on an application systems portfolio supporting its current and future tasks, and newcomers in any area are dependent on establishing a similar application portfolio quickly in a way that can evolve with changed business needs, technological affordances and expectations among co-operators, competitors and customers, these days especially to address a rising number of (in particular) mobile delivery channels.

Over these 30 years, one has regularly investigated how IT-systems are developed and used. In [LST78], results from a 1977 survey on distribution of work on IT in organizations were published. Somewhat surprisingly at that time, one found that almost half of the time was used on maintenance (i.e. changing systems that were already in production). Similar investigations have been done in Norway in 1993 [KS94], 1998 [HKS00], 2003 [KJS06], and 2008 [DK10]. For some areas, there are large differences for instance which programming languages that are used. Other areas are remarkably stable. As illustrated in Table 1, the split of the time use for development vs. maintenance is relatively similar to how it was 30 years ago. (In absolute terms a smaller proportion is used for both development and maintenance, since most of the time used by IT-departments these days are on user-support and keeping systems operational as we

see in the above part, the last two lines in the table show the numbers when we do only look at the relative distribution of development and maintenance time). Over the last 20 years the percentage of new systems that are in fact replacement systems, being installed basically to replace an old system, has stayed above 50%, rising slowly.

Table 1. Comparisons of maintenance figures across investigations

Category	**Nor. 2008**	**Nor. 2003**	**Nor. 1998**	**Nor. 1993**	**Nosek/ Palvia [NP90]**	**Lientz/ Swanson 1977**
Development	21%	22%	17%	30%	35%	43%
Maintenance	35%	37%	41%	40%	58%	49%
Other work	44%	41%	42%	30%	7%	8%
Disregarding 'other work'						
Development	34%	34%	27%	41%	38%	47%
Maintenance	66%	66%	73%	59%	62%	53%

One example of a notable change is *where* systems are developed, maintained and operate compared to 20 years ago. In 1993, 58% of the systems were developed by the IS-organization, and only one percent was developed *in* the user organization. In 1998, however, 27% of the systems were developed by the IS-organization and 27% as custom systems *in* the user organization. In 2003, 23% of the systems are developed in the IS-organization, whereas in 2008 only 12% were developed in the IS organization. The percentage of systems developed by outside firms is higher in 2008 (40% vs. 35% in 2003, vs. 22% in 1998 vs. 12 % in 1993). The percentage of systems developed based on packages with small or large adjustments is also increasing (41% in 2008 vs. 39% in 2003 vs. 24% in 1998 vs. 28% in 1993). The new category introduced in 1998, component-based development (renamed "use of external web services" in 2008) is still small (5% in 2008) although increasing (1.0 % in 2003, 0.4% in 1998). Whereas earlier the main way of developing systems where to do it in-house or using temporary consultants, outsourcing of all type of IT-developments has been on the rise over the last 10 years, although you also find examples lately on insourcing. In the 2008 investigation around a third of the IT-activity was outsourced (32.9% in private sector, 24.1 in public sector). Whereas only two of the respondents reported to have outsourced all the IT-activities, as many as 84% of the organizations had outsourced parts of their IT-activity. Whereas the public organizations have outsourced more of the development (40% in public, 29% in private) and maintenance (34% in public, 30% in private) than the private organizations, they have outsourced less of the operations (31% in public, 41% in private) and user support (21% in public, 29% in private). Other important aspects to take into account are the rise of new delivery models (e.g. OSS) not only for infrastructure applications, and the trend towards using cloud infrastructures for the operations of ICT-solutions.

A number of needs can be identified in this landscape that must be addressed in future business information systems. We will in this article discuss these issues highlighting the

application and possible changed role of *modelling techniques.* In section 2 we describe the role of modelling in information systems development in general. In section 3 some traits of future business and community information systems are described, and section 4 describes a vision of the role of modelling in this landscape. Section 5 concludes the paper.

2 The Role of Modeling and Quality of Models

Conceptual modelling is usually done in some organizational setting. One can look upon an organization and its information systems abstractly to be in a certain state (the current state, often represented as a descriptive 'as-is' model) that are to be evolved to some future wanted state (often represented as a prescriptive 'to be' model). The state includes the existing processes, organization and computer systems. These states are often modelled, and the state of the organization is perceived (differently) by different persons through these models. This open up for different usage areas of conceptual models as described e.g. in [NK06], illustrated in Fig. 1:

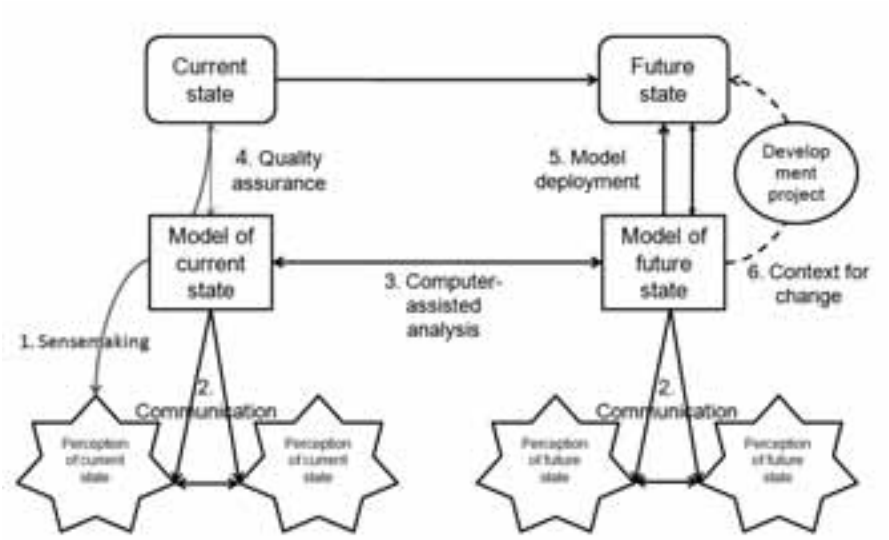


Figure 1: Application of conceptual modelling

1. Human sense-making: The model of the current state can be useful for people to make sense of and learn about the current situation as it is perceived.
2. Communication between people in the organization: Models can have an important role in supporting human communication [W11].
3. Computer-assisted analysis: To gain knowledge about the organization through simulation or deduction, often by comparing a model of the current state and a model of a future, hopefully better state.

4. Quality assurance, ensuring e.g. that the organization acts according to a certified process achieved for instance through an ISO-certification process.
5. Model deployment and activation: To integrate the model of the future state in an information system directly. Models can be activated in three ways:
   a. Through people, where the system offers no active support.
   b. Automatically, where the system plays an active role, as in an automated workflow system.
   c. Interactively, where the computer and the users co-operate to bring the process forward [KJ02].
6. To give the context for a traditional system development project, without being directly activated.

SEQUAL is a generic framework for assessing quality of models [K12, KSJ06, KS03]. The framework has earlier been used for evaluation of modelling and modelling languages of a large number of perspectives, including data, object, process, enterprise, and goal-oriented modelling. Quality has been defined referring to the correspondence between statements belonging to the following sets:

- ***G,*** the set of goals of modelling.
- ***L,*** the language extension, i.e., the set of all statements that are possible to make according to the rules of the modelling languages used.
- ***D***, the domain, i.e., the set of all statements that can be stated about the situation.
- ***M***, the externalized model itself.
- ***K***, the explicit knowledge relevant to the domain of the audience.
- ***I***, the social actor interpretation, i.e., the set of all statements that the audience interprets that an externalized model consists of.
- ***T***, the technical actor interpretation, i.e., the model as 'interpreted' by modelling tools.

The main quality types, illustrated in Fig. 2 are:

- Physical quality: The basic quality goal is that the externalized model ***M*** is available to the relevant social and technical actors for interpretation (***I*** and ***T***).
- Empirical quality deals with comprehensibility and predictable error frequencies when a model ***M*** is read or written by different social actors.
- Syntactic quality is the correspondence between the model ***M*** and the language extension ***L***.
- Semantic quality is the correspondence between the model ***M*** and the domain ***D***. This includes validity and completeness.
- Perceived semantic quality is the similar correspondence between the social actor interpretation ***I*** of a model ***M*** and his or hers current knowledge ***K*** of domain ***D***.
- Pragmatic quality is the correspondence between the model ***M*** and the actor interpretation ***(I*** and ***T)***. One differentiates between social pragmatic quality (to what extent people understand the models) and technical pragmatic quality (to what extent tools can be made that can interpret the models).
- The goal defined for social quality is agreement among social actor's interpretations.
- The deontic quality of the model relates to that all statements in the model ***M*** contribute to fulfilling the goals of modelling ***G***, and that all the goals of modelling ***G***

are addressed through the model ***M***. In particular, one include under deontic quality the extent that the participants after interpreting the model learn based on the model (increase ***K***) and that the audience are able to change the domain ***D*** if this is beneficially to achieve the goals of modelling.

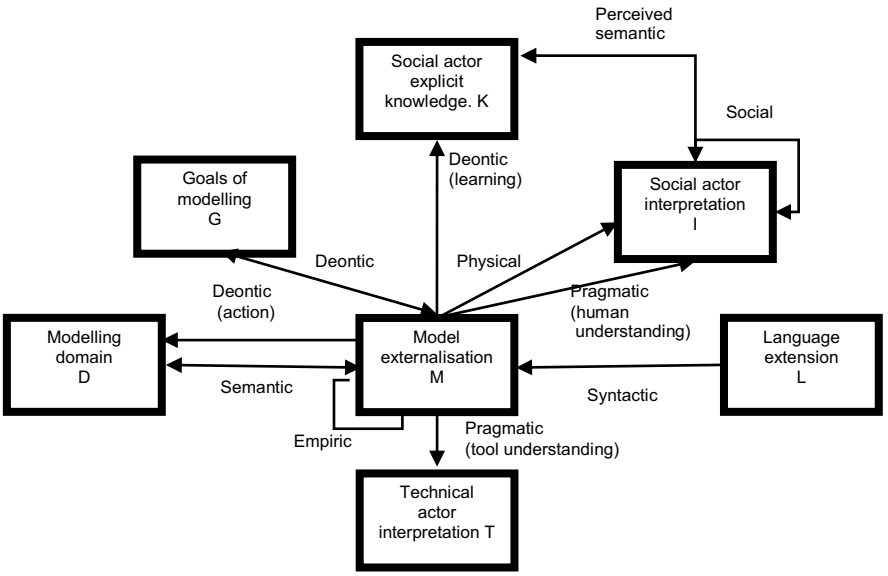


Figure 2: SEQUAL framework for discussing quality of models

3 Future Business and Community Information Systems

The future business environment must take into account a stronger integration across traditional organizational borders and countries. In the FiNES position paper [L11], a number of organizational trends are highlighted:

- *The disappearing boundaries of the enterprise*: Virtual teams, extended value chains, dynamic value networks and a range of digital ecologies [VZUM10] such as digital ecosystems will change the contours of the business world.
- *Everybody is an Enterprise*: The role of SMEs and especially micro enterprises and start-ups is expected to become even more important
- *The WhatYouSenseIsWhatYouGet (WYSIWYG) Enterprise*: With massive quantities of real-time information becoming available one can get new real-time enterprise applications.
- *A Knowledge Commons for Enterprises:* There is a need for an open and accessible "Knowledge Commons" facilitated by the Internet for all to freely, responsibly and legitimately exploit

- *The advent of intelligent virtual reality*: With increasing dematerialization blurring the physical with the virtual, enterprises are not only becoming digital, they are acquiring multiple roles, cyber presences and distinctive virtual identities

Combined with societal trends (e.g. demographic, legislative, financial, geo-political, and environmental) that demands change in current practice there is a number of issues that must be addressed in future business information systems to support these organizational and societal needs. Future business information systems will take advantage of in particular future internet to

- Be empowered by a new participative Web hosting a new wave of services and using user-friendly technologies;
- Create value by leveraging the Internet as the platform through which knowledge is manipulated dynamically, experienced in the business context and represented in a radically different way;
- Have the required capability that enables and supports collaboration with other enterprises, new dynamic relationships, discovery of partnerships, new opportunities and markets, and the management of the new risks involved
- Operate in a new set of business environments that provide support for quality measures, guarantees, persistence, safety, trust, arbitration and other mechanisms for reducing risks on both the customer and the provider side
- Become the WYSIWYG enterprise, where Web-based applications become as rich as the desktop.

A number of partly related needs can be identified in this landscape that must be addressed in future business and community information systems:

- The support of end-to-end engineering process (including full life-cycle support)
- Integration across organizations and nations
- Systems being provided by ecosystems of providers
- Event-oriented systems utilizing the internet of things

We will below go into more detail on each of these areas, discussing in particular the application of modelling techniques.

3.1 The support of end-to-end design and engineering process (including full life-cycle support of products)

Even with the shift to a service economy, manufacturing and engineering remains important for the economy. On the other hand we witness new trends such as sustainable manufacturing and mass customization. Consequently, the manufacturing industry is facing significant changes.

The key enabler for coping with these changes will be IT, due to its strong impact on innovation and productivity. Current IT for manufacturing is characterized by scattered data formats, tools and processes dedicated to different phases in the product lifecycle. Moreover, the flow of information is closely aligned to the product lifecycle i.e., information from the design phase goes into the manufacturing phase, but not in the

opposite direction. In addition, user feedback is often neglected in design. Due to the diversity of tools and data formats, manufacturing struggles to cope with new trends in this area. For example, both the trend to mass customization and the demand for increased sustainability require a tight integration of the design, manufacturing and usage phases of a product, which is currently not in place. What is missing is an integrated, holistic view on information, persons and processes across the full product lifecycle. As experiences of the past show, a tight integration of all tools used throughout a product lifetime is not feasible. A model-based approach to address this in a more federated manner is the application of AKM (Active Knowledge Modeling) [LK08].

3.2 Integration across organizations and nations

Over the last 10 to 15 years the number of varying organizational forms has increased a lot, from strictly structured supply-chains to loosely structured communities of interest. Co-operation across traditional organizational boundaries is increasing, as outsourcing and electronic business is enabled by the Internet and IT in general. When such co-operation moves beyond the buying and selling of goods and well-defined services, there is a need for a flexible infrastructure that supports not only information exchange, but also knowledge creation and sharing.

An extended enterprise (EE) (also often termed virtual enterprise) is defined as a dynamic networked organization, being developed ad-hoc to reach a certain goal based on the resources of several existing co-operating enterprises [KJ04]. The EE-partners often comes from different countries, using different languages and having different cultural background. The EE-partners wants to harvest knowledge from the EE to be reused in the originating organization, or in other EE's. An approach that needs to be extended to support this situation is process-support environment like workflow modelling and BPM [HVLAK10]. Based on globalization trends, also other challenges pop up. [KDS08] provides an example of the needs when a multi-national company is to coordinate their local business units in order to serve other multi-national companies in an integrated fashion. As reported in [KDS08] there was in this case a need to standardize the processes of the company's national branches in order to build a common image of the organization (both inward and outwards), and to support the certification of the cross-national processes of their multi-national customers, but at the same time adhere to national and cultural rules and expectations for how to do business. Similar examples from other businesses (e.g., car rental industry) are presented in [A11]. Such aspects are not only of importance in business, but also in the public administration area. Public administrations and service providers face growing challenges linked to the application of new IT-solutions and are forced to rethink traditional administrative structures and functions and adapt their services to meet new societal demands with reduced budgets. In the public sector there is increasing emphasis in particular in Europe on cross-department and cross-national applications, to increase reuse of solutions across traditional borders [AVJK11], as described for instance in the last EU Ministerial Declaration on eGovernment [EU09]. For instance, even in small countries like Norway and the Netherlands, there are around 430 local municipalities which in principle execute variants of the same processes [A11, AVJK11] which are different due to political, historical and demographic reasons.

3.3 Systems being provided by ecosystems of providers

As described in the introduction, the trend for a long time has been that systems are developed and evolved further and further away from the core users of the system. Rather than being provided by distinct entities, we see a development in the direction of systems to a larger degree being supported by *virtual communities* of human/organizational actors, co-working on partially shared digital artefacts [BDLV09]. The term *'digital ecosystem'* is being used to generalize such communities, emphasizing that their actors constantly interact and cooperate with other actors in both local and remote ecosystems. Such systems are characterized by self-organization, scalability and sustainability, providing both economic and social value as a specialization of the more generic concept digital ecologies. Examples are communities for Open Source (OSS) and Creative Commons, Knowledge Commons, social media networks, or voluntary groups of citizens or academics.

However, the existing digital ecosystems have limited scope, various degree of transparency, insufficient capabilities for search and evaluation of useful, high quality artefacts from the huge and ever-evolving Internet, and none does fully support a wide range of shared artefacts from a wide range of actors. There are two main variants of digital ecosystems; *content ecosystem and software ecosystems. Content ecosystems* are networks that deal with creation and sharing of artistic and intellectual artefacts**.** The impact of IT on participative and democratic processes and on creativity is already here, and will continue to grow with the increasing diffusion of web-based social networking and user generated content and services. S*oftware ecosystems* are defined as ”*a set of businesses functioning as a unit and interacting with a shared market for software and services, together with relationships among them. These relationships are frequently underpinned by a common technological platform and operate through the exchange of information, resources, and artefacts*” [JFB09].

3.4 Event-oriented systems utilizing the internet of things (IoT)

What previously was termed convergence (between the telecommunications world, IT, media, and later also the power systems in what is now popularly called smart grids) is now emerging in practice, propelled by several simultaneous trends.

1. The continuing miniaturization of computing resources making it possible to perform computing at some level everywhere, by any device.
2. The availability of high-bandwidth access to computing resources in an increasing number of places.
3. The infrastructure being built up for utilizing remote computing resources (these day often presented under the term 'cloud computing')
4. More power-efficient solutions. Many battery-operated devices can last more than a year, and passive solutions used by RFID and NFC (Near Field Communication) are powered by the readers. Parasitic energy harvesting devices that extract small amounts of energy from the environment can power sensors where normal power solutions are not economically or practically available.

The emergence of the IoT will lead to a world in which countless everyday objects are interconnected and have their own computational power. These active and interactive objects will be able to monitor and change their environment via sensors and actuators, and to interact and collaborate with each other and with the people around them. Although more accessible, to quickly produce innovative applications and services for this setting is not an easy task. Information systems need to combine transaction-orientation with event-orientation, using event-driven architecture (EDA) [CS09] reflecting the nature of IoT applications. These applications are characterized by being:

- Collaborative, that is including numerous active components that behave concurrently, running on distributed devices, and collaborates to provide a desired functionality.
- Event-driven and reactive, that is continuously reacting on a large number of events from the physical environment, from users and from other parts of the system.
- Dynamic and adaptive, meaning that the applications, associations and collaborations are dynamically established and configured depending on the current context at run-time. The information gathered is often used to support human decision-making rather than to only automate a predefined process. Interaction with human users is essential both to use the systems efficiently, and to effectively improve the systems over time through capturing the experiences from the use of the system.

4 Quality of Models for Future Business and Community Information Systems

There are two main scenarios for the future use of modelling. What we term the steady state scenario, where modelling continues to be a somewhat esoteric activity for a limited number of experts is of course one possibility. The more optimistic view in the light of the above is that abstractions techniques such as modelling is taken into use in an increasing number of areas, to make it possible to at all be able to manage the development. One striking aspect is that the number and variety of stakeholders that will need to relate to models of some sort will increase. Given the increased educational level in most countries, it is not unlikely that also more people will be able to relate to these types of abstractions, given that one of the thing that you are exposed to in a master study of most types, is training to deal with abstractions. Looking at the sets in SEQUAL, we can thus foresee the following under this scenario:

- G: The same list of goals and applications of modelling that is described in Section 2 will still be relevant, but emphasis on less formal, interactive approaches will increase to be able to support the more federated landscape needed to address future IT-driven environments.
- D: The range of relevant domains is increasing, also within what can be classified as the 'same' system, given that systems to an increasing degree ranges across and is expected to integrate a number of areas.

- K: In many areas one need to deal with a more varied set of stakeholders, with a more varied set of skills and knowledge. Not only do you need to align IT-experts with 'business'-experts, but also people across a large range of business expertise, being used to express their knowledge in a wide variety of notations.
- L: Using domain specific modelling, the possibilities of tailoring the language to fit the domain, and the knowledge of the stakeholders have increased. To bring more people into (semi-) formal modelling these possibilities will have to be exploited to a larger degree. Thus rather than having a consolidation of modelling languages like the one we saw in object-oriented design with UML, there will be an increasing number of variants of modelling languages. We will also see a mix of richer media components being integrated with the more traditional "box and arrow"-conceptual modelling notations, thus supporting also richer meta-meta models defining and limiting the type of constructs to include in models in the first place.
- T: An increasing number of tools will be available to extract model information from raw data, e.g. in the area of process mining [A11]. In addition, tools for meta-meta modelling and meta-modelling will be more common.
- M: Models will be pervasively available being coordinated in a federated manner. Models will be across meta-levels in an increasing degree (compared to the models in traditional software engineering being primarily on the type level). Models, in particular interactive models will have a larger value in themselves, acting to a larger extend as knowledge commons and open models (http://www.openmodels.at/).

We believe the core dimensions of SEQUAL will be relevant also for models as used for future systems. On the other hand, a number of specializations might be envisaged. We will briefly discuss some main aspects here.

- Physical quality: Rather than being based on central repositories, more distributed, federated storage of model fragments must be available, utilizing standard interchange formats and supporting model mash-ups.
- Empirical quality: Support for empirical quality will be more built in, e.g. in tools that build up models from raw data in process mining, thus integrating information visualization tools and modelling tools utilizing generic knowledge on good visualization tactics to a larger degree. Note that different meta-meta models can induce the need for rethinking guidelines for achieving empirical quality [NK09].
- Syntactic quality: Syntactic quality can be looked upon as trivial in a sense, since adherence can be enforced. On the other hand, one often sees that one extend languages with new aspects in an (not always conscious) attempt to turn semantic problems into syntactic problems. New tools based on meta-modelling makes this easier to do, and then makes in even more important to do right in the sense that one do not end up with too restricted languages for the different islands of modelling.
- Semantic quality: The federated approach needed for modelling will bring new challenges as for how we look upon the semantic quality of the overall model. Whereas semantic quality in smaller domains would be followed up much as before (i.e. looking at the feasible (perceived) completeness and validity), one would to a larger degree need to be able to live with inconsistencies across federations. In

connection to this, it would be important to be able to identify those aspects of the models across domains that need to be consistent.

- Pragmatic quality: Given that more types of stakeholders are involved, this is of increasing importance. Different techniques can be used for different types of stakeholder, supporting multiple views for different stakeholder types on the same model to ensure individual comprehension.
- Social quality: This will be important in smaller communities, and in interfaces between communities, but less needed across federation. Note on the other hand that since different stakeholder groups might see different views of the overall model, possible visualized in radically different ways, the effort to assure that they comprehend the models equally will potentially have to increase, as seen in e.g. [KDS08].
- Deontic quality: Models will be more important, in particular in organized conduct across traditional organizational boundaries and needs to be handled in a more professional manner [W11].

5 Concluding Remarks

From the above descriptions, we see that the technical challenges and opportunities with the future internet gives new challenges and opportunities for business information systems on many levels, all of which give new challenges and opportunities for model-based techniques such as BPM, MDA, enterprise modelling, value modelling and AKM. In a way many of the core problems are not new, e.g. how to deal with event-based systems have been investigated in the conceptual modelling area for quite some time. Even if the use of modelling need to be extended and improved, general categories underlying discussions on quality of models as described in [KSJ06, K12] remains relevant, although need to be adapted to e.g. quality of interactive models [KJ02,LK08].

References

[A11] van der Aalst WMP (2011) Intra and Inter-Organizational Process Mining: Discovering Processes within and between Organizations) In: PoEM 2011, Oslo Norway Nov 2-3

[AVJK11] Aagesen G; van Veenstra AF; Janssen M, Krogstie J. The Entanglement of Enterprise Architecture and IT-Governance: The cases of Norway and the Netherlands. Paper presented at the HICCS 2011

[BDLV09] Ali Babar M, Dingsøyr T, Lago P, van Vliet H (ed) (2009) Software Architecture Knowledge Management - Theory and Practice Springer

[CS09] Chandy K, Schulte, W (2009) Event Processing: Designing IT Systems for Agile Companies McGraw-Hill

[DK10] Davidsen, M. K., Krogstie, J. (2010). A longitudinal study of development and maintenance. Information and Software Technology, 52, 707-719

[EU09] EU (2009) Ministerial Declaration on eGovernment, Malmø, Sweden 5th Ministerial eGovernment conference

[HKS00] Holgeid, K.K., Krogstie, J., Sjøberg, D.I.K. (2000). A study of development and maintenance in Norway: Assessing the efficiency of information systems support using functional maintenance. Information and Software Technology. 42, 687-700.

[HVLAK10] Houy C, Fettke P, Loos P, van der Aalst WMP, Krogstie J (2010) BPM-in-the-Large – Towards a higher level of abstraction in Business Process Management. Paper presented at GISP under WCC 2010 Brisbane Australia September

[JFB09] Jansen S, Finkelstein A, Brinkkemper S (2009) A sense of community: A research agenda for software ecosystems. In: Proc. 31st International Conference on Software Engineering (ICSE), New and Emerging Research Track - Companion Volume, May 16-24, 2009, Vancouver, Canada, pp. 187-190

[K12] Krogstie, J (2012) Model-based Development and Evolution of Information Systems: A Quality Approach. Springer.

[KDS08] Krogstie J, Dalberg V, Jensen SM (2008) Process modeling value framework. In: Manolopoulos, Y., Filipe, J., Constantopoulos, P., Cordeiro, J. (eds.) ICEIS 2006. LNBIP, vol. 3, pp. 309-321. Springer, Berlin

[KJS06] Krogstie, J., Jahr, A., & Sjøberg, D.I.K. (2006). A longitudinal study of development and maintenance in Norway: Report from the 2003 investigation Information and Software Technology 48, 993-1005

[KJ02] Krogstie, J., Jørgensen, H. D. (2002). Quality of Interactive Models. Paper presented at the First International Workshop on Conceptual Modelling Quality (IWCMQ'02)

[KJ04] Krogstie, J., Jørgensen, H. D. (2004). Interactive Models for Supporting Networked Organisations. Paper presented at the 16th Conference on advanced Information Systems Engineering (CAiSE 2004), Riga, Latvia, June 9-11

[KSJ06] Krogstie, J., Sindre, G., Jørgensen, H. D. (2006). Process Models as Knowledge for Action: A Revised Quality Framework. EJIS 15(1), 91-102

[KS03] Krogstie, J., Sølvberg, A. (2003). Information systems engineering - Conceptual modeling in a quality perspective. Trondheim, Norway: Kompendiumforlaget

[KS94] Krogstie, J., Sølvberg, A. (1994) Software maintenance in Norway: A survey investigation. In H. Muller & M. Georges (Eds.) Proceedings of the International Conference on Software Maintenance (ICSM'94) (pp. 304-313), IEEE CS Press.

[L11] Li, M.-S et al. (eds) (2011) Position Paper on Orientations for FP8 A European Innovation Partnership for European Enterprises, Version 3.0 21. Future Internet Enterprise Systems

[LST78] Lientz, B.P., Swanson, E.B., Tompkins, G.E. (1978). Characteristics of application software maintenance. Communications of the ACM, 21 (6) 466-471.

[LK08] Lillehagen, F. Krogstie, J. (2008) Active Knowledge Modeling of Enterprises, ISBN 978-3-540-79415-8 Springer

[NP90] Nosek, J.T., Palvia, P. (1990) Software maintenance management: Changes in the last decade. Journal of Software Maintenance, 2, 157-174.

[NK06] Nysetvold AG, Krogstie J (2006) Assessing Business Process Modeling Languages Using a Generic Quality Framework. In: Siau K (ed) Advanced Topics in Database Research Series vol 5. Idea Group, pp 79-93

[NK09] Nossum A, Krogstie J (2009) Integrated Quality of Models and Quality of Maps. Paper presented at the EMMSAD 2009

[VZUM10] Vyatkin V, Zhabelova G, Ulieru M, McComas D (2010) Toward Digital Ecologies: Intelligent Agent Networks Controlling Interdependent Infrastructure. IEEE Conference on Smart Grid Communications, Gaithersburg, October

[W11] Wesenberg H (2011) Enterprise Modeling in an Agile World PoEM 2011, Proceedings of the 4th conference on Practice of Enterprise Modeling, Oslo, Norway, November 2-3

SESSION 6

Information Retrieval and Modelling

Open Up Cultural Heritage in Video Archives with *Mediaglobe*

Christian Hentschel, Johannes Hercher, Magnus Knuth,
Johannes Osterhoff, Bernhard Quehl, Harald Sack,
Nadine Steinmetz, Jörg Waitelonis, Haojin Yang
{givenname.lastname}@hpi.uni-potsdam.de
Hasso Plattner Institute for Software Systems Engineering
Prof.-Dr.-Helmert-Str. 2–3, 14482 Potsdam, Germany

Abstract:
Film, video, and TV have become a predominant medium, but most audiovisual (AV) material being part of our cultural heritage is kept in archives without the possibility of appropriate access for the public. Although digitalization of AV objects in conjunction with AV analysis is making progress, content-based retrieval remains difficult because of the so called semantic gap. The *Mediaglobe* project is focussed on digitalization, indexation, preservation and exploitation of historical AV archives. In this context, we show how traditional AV analysis is complemented with semantic technologies and user-generated content to enable content-based retrieval exposing contentual dependencies to promote new means of visualization and explorative navigation within AV archives.

1 Introduction

Audiovisual media such as film, video, and TV have become the predominant medium of the 21th century. From its beginning in the early 1900s until today, all kind of events all around the globe have been captured on celluloid, magnetic tape, or digital hard discs. By accessing these recordings, we are able to turn back the clock and take a peek on people and places of bygone time. But, most AV material being part of our cultural heritage is kept in archives without the possibility of appropriate access for the public audience. Even worse, celluloid is subject to an aging process and deteriorates over time. Also extensive (analog) preservation technology cannot guarantee long time storage and access. Thus, more and more AV material from the early age of film will be irrecoverably lost. Therefore, governments worldwide are funding preservation and digitalization projects, such as NESTOR[1] – the German competence network for digital preservation – to prevent the loss of cultural heritage. Everyday, the stock of AV material is growing. Recently also internet and world wide web (WWW) are used for media distribution, sometimes also providing random access and online search. But the number of online available AV material is only the tip of the iceberg, while its basis remains hidden in multitudinous analogue archives and libraries. The opening of these archives for online access requires

[1] http://www.langzeitarchivierung.de/Subsites/nestor/EN/

digitalization of various analogue media formats of different quality first. But, to access AV materials by its content, metadata for the description of its structure and content are mandatory.

In this paper, we present the *Mediaglobe* semantic video search engine and introduce its workflow including video analysis, metadata generation, semantic analysis, and video search. The primary goal of the *Mediaglobe* project[2] is to develop a generally applicable and commercially efficient infrastructure for digitalization and retrieval of AV archives with an emphasis on historical documentaries. Here, we are focussing on the retrieval process and its prerequisites, leaving out the entire process of digitalization and restoration. Most important entities to be retrieved are persons, locations, organizations, and events. We describe automated AV analysis techniques and how they are complemented by manual and collaborative annotations. Text-based metadata are endorsed by formal semantic ontologies to enable cross linking of data as well as explorative search. Instead of presenting search results in a traditional linear way, we show how semantic relationships can be utilized to enable efficient visualization and navigation.

The paper is structured as follows: Section 2 gives a short overview of related work, while Section 3 introduces the *Mediaglobe* system architecture and its analysis workflow. Visual analysis including scene cut detection, text recognition and visual concept detection in video is presented in Section 4. Section 5 focusses on the semantic analysis within the *Mediaglobe* project including an approach for named entity recognition. In Section 6 we demonstrate the semantic search facilities of the *Mediaglobe* search engine and Section 7 provides an overview of the newly developed semantic search interface. Section 8 concludes the paper with a short outlook on ongoing and future work.

2 Related Work

Current web-based video platforms such as YouTube allow their users to upload content and to assign basic metadata (title, keywords, description). Although, YouTube enables the annotation of temporal video segments using fragment identifiers up to now there is no support for semantic annotation as in *Mediaglobe*. In [SH10] Steiner and Hausenblas exemplify the processing of YouTube's metadata with natural language processing tools (NLP) in order to tag existing video metadata with linked data entities, but this approach is limited to the video as a whole, whereas in *Mediaglobe* semantic annotation is time based and automatically generated via sophisticated video analysis technologies. With regard to time based semantic video annotation for enabling semantic and exploratory video search we refer to previously published in depth results [WKW+10, LS11, WS11]. In [KSR09] Kobilarov et al. present the BBC's approach to join up all of its resources using linked data principles and a tailored ontology for BBC's program data. Later on, the NoTube project[3] harnessed BBC's program information by analyzing it with a NLP-tool to extract named entities and to map them to linked data resources. Again, both approaches are limited to

[2]The *Mediaglobe* project is part of the THESEUS research program funded by the German Federal Ministry for Economics and Technology, http://www.projekt-mediaglobe.de/

[3]http://notube.tv/

program data, and also consider the annotated asset as a whole rather than its temporal segments. In contrast to the BBC program data and the NoTube project in *Mediaglobe* there is no access to rich editorial materials [vAARB09]. Consequently, metadata in *Mediaglobe* has to be generated by automated video analysis techniques such as scene- and overlay-text recognition or automated speech recognition.

Mature projects including Prestospace, Europeana, and THESEUS/CONTENTUS build foundations for semantic content analysis within the cultural heritage domain. The aim of Prestospace[4] was to develop preservation technology for broadcasters. In [MBS08] Messina et al. describe tools for content-based analysis, e. g., scene-cut detection, speech to text transcription and keyframe-extraction. Moreover, named entities and categories from the BBC's program web-pages were also extracted and mapped to the proton upper-ontology [DTCP05]. *Mediaglobe* complements Prestospace's efforts, as high-level abstraction layer analysis technologies, i. e. visual concept detection was not in the scope of Prestospace. In regards of semantic analysis, video assets in *Mediaglobe* rarely come with rich accompaniment metadata as it is the case of tv programs. Therefore, background knowledge has to be aggregated that is maintained independently of the original source, i. e., the film archive. In contrast to [DTCP05, MBD06] we access the DBpedia as multilingual knowledge base to allow crosslingual search and annotation with entities that are steadily curated by the Wikipedia community. The Europeana[5] aggregates video data including supporting documents from various film archives in Europe. Prominent examples of video platforms that supply the Europeana search interface include the European Film Gateway (EFG), the videoactive platform, and EU-screen. However, these projects concentrate on content alignment and crosslingual search by harmonizing existing but heterogeneous metadata originating from broadcasters, film archives, and cinematheques [SET+12]. To our best knowledge, they do not apply sophisticated content-analysis technology to leverage metadata nor perform additional time-based semantic video analysis. Within the CONTENTUS usecase of the German THESEUS research program technologies supporting the media library of the future have been developed [NLFH12]. CONTENTUS targets multimedia in all its diversity whereas *Mediaglobe* is specialized on video only. *Mediaglobe* complements the semantic video analysis technology of CONTENTUS by applying semantic graph analysis for entity mapping as well as by applying high-level abstraction layer analysis technologies, such as visual concept detection.

3 System Architecture

To process large amounts of video content, many requirements have to be met including high scalability, customizable workflow design, reusability, versioning, and compatibility with state-of-the-art standards. Since traditional video processing workflow systems focus on transcoding, delivery, publishing, and metadata management, they still lack sophisticated and semantic search to open up its content. Therefore, *Mediaglobe* is designed not only to perform standalone, but also to complement existing workflow systems by bringing

[4] http://prestospace.org/
[5] http://europeana.eu/

Figure 1: The overall architecture of the *Mediaglobe* system

in flexible video content analysis and semantic search features.

The *Mediaglobe* system consists of three main parts: a *Media Asset Management System* (MAM), the *Analysis Framework* and the *Video Search Engine*. The MAM maintains the video ingest, transcoding, play out streaming, rights management, and commercial exploitation. The analysis framework builds upon the Apache Unstructured Information Management Architecture (UIMA) that enables the compilation of various Analysis Engines (AE) and orchestrates the data flow between them. Work intensive components can be deployed plurally for parallel and/or distributed execution. The video search engine provides a semantic search index and serves the Graphical User Interface (GUI) to efficiently browse and search within the video data. Fig. 1 gives an overview on the main components and the overall architecture. After ingesting a new video from the MAM the workflow engine starts with scene cut detector and keyframe extractor as preprocessing for subsequent video Optical Character Recognition (OCR) and visual concept detection. The auditive tracks are processed by third-party Automatic Speech Recognition (ASR). All extracted textual information including metadata provided by the users is processed by the semantic analysis. This employs the temporal context dependent Named Entity Recognition (NER). Finally, the semantically enriched data is made available in an RDF triple store to query via SPARQL and a search engine for end user access. All essential components are introduced in the following sections, starting with the temporal decomposition of video with scene cut detection.

4 Visual Analysis

The low-level analysis of the visual content of video data provides additional source for metadata. This section describes visual analysis techniques employed in *Mediaglobe*.

4.1 Scene Cut Detection and Keyframe Selection

Structural video analysis by scene cut detection is the first step of video content analysis, indexing, and classification. Video shots are defined as a sequence of consecutive frames taken by a single camera act. Shot boundary detection provides useful information about the temporal structure of a video, which is a prerequisite to the extraction of representative keyframes or frame candidates for video OCR, visual concept detection, and to define temporal contexts for NER. We distinguish two different types of shot boundary transitions: abrupt scene changes (*hard cuts*) and gradual transitions that extend to more than a single frame (*soft cuts*) [AAM10].

Most common gradual transitions are *fade-ins* and *fade-outs*, *wipes* and *dissolves*. As the video data under consideration in the *Mediaglobe* archive restrains to hard cuts and fades, we do not consider wipe and dissolve detection here. Our approach for hard cut detection is based on a statistical method expanded on the idea of pixel differences [ALBK09]. We compute the the $L2$-norm between every 5 consecutive frames and consider a frame a hard-cut candidate if the gradient computed on the first derivative of this metric exceeds an empirically determined (adaptive) threshold. Fade-ins and fade-outs exhibit a slow decrease or increase change in illumination usually ending or starting with a black frame. Detection is based on computing the image entropy – a monotonously increasing entropy value indicates a fade-in, beginning and end frame for this transition are defined by the local minimum and maximum respectively.

Next to shot-boundary detection scene cut detection provides input data that is qualified for further content-based visual analysis of a video. In particular, the selection of meaningful keyframes per segment provides efficient means to reduce the temporal redundancy within the visual data of a video and thus the computational effort required for subsequent analysis steps, e. g., video OCR and visual concept detection. While the video OCR runs on up to 5 frames selected equidistantly per shot, the visual concept detection assumes that a single frame provides enough information about the visual content of an entire video segment, and therefore has to be chosen in a way to contain the most meaningful content.

4.2 Video OCR

Text embedded in video data is a valuable source for indexing and searching in video content. In order to retrieve textual data from video, standard OCR approaches, which focus on high resolution scans of printed documents, need to be extended to meet the requirements for OCR in video data. Our approach consists of two steps: keyframes that contain textual data are identified and subsequently, the location of text within these frames is determined in order to separate text pixels from background and to deliver this preprocessed data to a standard OCR engine. In our approach, we classify text candidates from the video data stream by applying a fast edge-based multi-scale detector and subsequently a projection-profiling algorithm on each keyframe. Frame sections containing no text are rejected by identifying regions that exhibit low edge density. Further refinement is achieved

by a Stroke Width Transform (SWT) based verification procedure on each detected text candidate section. We adapted the original SWT algorithm [EOW10] in order to improve the performance [YQS12b].

Typically, video text is embedded in very heterogeneous background with low contrast, which makes it difficult to be recognized by standard OCR engines. Hence, text pixels need to be separated from background by applying appropriate binarization techniques. We have developed a novel skeleton-based approach for video text binarization [YQS12a], which consists of three steps: First, we determine the text gradient direction for each text line object by analyzing the content distribution of their skeleton maps [CPW93]. We then calculate the threshold value for seed-selection by using the skeleton map which has been created with the correct gradient direction. Subsequently, a seed-region growing procedure starts from each seed pixel and extends the seed-region in its north, south, east, and west orientations. The region grows iteratively until it reaches the character boundary.

Text binarization is concluded by converting text region data to black text pixels on white background, which is a prerequisite for the efficient application of a standard OCR software (*Tesseract*[6]) for text recognition. Moreover, an adapted version of the open source *Hunspell*[7] spell checker is applied to improve the quality of the achieved OCR results. While traditional spell checking software presumes typing errors based on a specific keyboard layout, an adapted OCR spell correction rather cares for visual similarity of the characters.

4.3 Visual Concept Detection

Next to textual data, the depicted visual concepts within a video scene provide additional information for search and classification. We employ methods for content-based image classification to automatically classify video scenes into predefined visual concept classes. Our approach follows the well-known *bag of keypoints* [CDF+04] model, where local image patterns are used to describe the low-level visual features of a keyframe. In our approach, we extract Scale Invariant Feature Transform (SIFT) [Low04] features at a fixed grid on each channel of a keyframe in RGB color space. These features are used to compute a visual codebook by running a k-means clustering that provides us with a set of representative codewords. By assigning each extracted RGB-SIFT feature to its most similar codeword (or cluster center) using a most simple nearest neighbor classifier we compute a histogram of codeword frequencies that is used as keyframe descriptor. The problem of visual concept detection is approached by means of machine learning techniques. Kernel-based Support Vector Machines (SVM) have been widely used in image classification scenarios (cf. [CDF+04, SW09, ZMLS06]). In order to be able to predict the category of an unlabeled keyframe the SVM classifier needs to be trained using labeled data.We consider the classification task a one-against-all approach – one SVM per given visual concept is trained to separate the keyframes from this concept from all other given concepts. Hence, the classifier is trained to solve a binary classification problem, i. e.,

[6] http://code.google.com/p/tesseract-ocr/
[7] http://hunspell.sourceforge.net/

whether or not a keyframe depicts a specific visual concept. We use a Gaussian kernel based on the χ^2 distance measure, which has proven to provide good results for histogram comparison. Following Zhang et al. [ZMLS06] we approximate the kernel parameter γ by the average distance between all training image histograms. Therefore, the only parameter we optimize in a cross-validation is the cost parameter C of the support vector classification. New keyframes can be classified using the aforementioned bag-of-words feature vectors and the trained SVM model.

The advantage of this approach is its simplicity, and its various invariances. The combination of SIFT for local image description and the bag-of-words model makes it invariant to transformations, changes in lighting and rotation, occlusion, and intra-class variations [CDF+04]. By simply counting object features present in an image, missing or occluded object parts do not affect the classification accuracy in the same way as it would be the case for model-driven classification. Moreover, the approach is very generic. It can be extended to additional concepts simply by training another SVM classifier. Since the features are not adapted to the classification task they can be reused.

5 Semantic Analysis

In order to enable an explorative search within video data the textual metadata derived from visual analysis and the authorative metadata (provided by the archive) are mapped to semantic entities. For *Mediaglobe* authoritative metadata comes as persons, places, organizations, other keywords, and freetext, e. g., title, description text, while visual analysis technologies provide time-based OCR and ASR text.

For semantic annotation in *Mediaglobe* DBpedia[8] entities serve as reference mapping entities for NER. Initially, for every entity all DBpedia labels including redirects and disambiguation links are collected. Next, for every label a distance measure to the original (main) label of the entity is calculated to determine a relevance ranking for entity candidates. For instance, the label 'DEBER' of the entity 'Berlin'[9] receives a lower distance score than the labels 'Berolina' or 'City Berlin', where 'Berlin' is the original (main) label. In case a term to be mapped can be assigned to several DBpedia entities, the term is ambiguous and further context data is needed for disambiguation According to their provenance metadata are of different reliability. Therefore, we have determined a context ranking according to metadata provenance. First, the authorative metadata items are disambiguated in the order [persons, places, organizations], [keywords], [freetext], whereas already disambiguated metadata terms serve as additional context for the currently processed metadata text item. Subsequently, the time-based metadata text items are disambiguated within the context boundaries a video segment. Every entity candidate of a metadata term receives a score according to its relevance within the defined context. In the next step, cooccurrence of context text terms and the entity candidates of the currently processed metadata text term is identified with the help of the Wikipedia articles of the

[8] http://dbpedia.org/About
[9] http://dbpedia.org/resource/Berlin

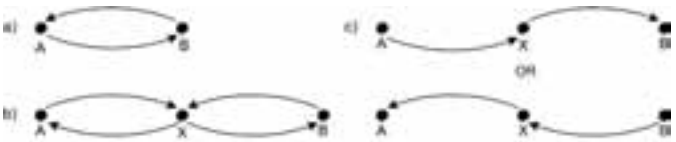


Figure 2: Three different types of Wikilinks: a) direct links, b) symmetric links through same node, c) links through a node, but not symmetric.

referenced entity candidates. The entity candidate with the highest involvement (score) within the context will be chosen as mapped entity [LS11].

In addition to cooccurrence analysis link graph analysis based on the Wikipedia page link graph is applied to identify connected components within a context. Three different types of links are considered (cf. Fig. 2). Similarly to the cooccurence analysis score the link graph score for every entity candidate is calculated according to the total count of links, but also to the count of links to different context terms taking into account how many different nodes are between a pair of entity candidates of different terms.

One of *Mediaglobe*'s key features is the collaborative and time-based annotation of video segments. To this end a customized video annotation tool has been developed, where the user is able to watch the video, stop it anytime and tag it with semantic entities or text keywords. An autosuggestion service provides the user with relevant semantic entities chosen from DBpedia for a typed text term. Once an entity has been selected the user can position the semantic tag anywhere within the video frame to achieve spatio-temporal annotation.

6 Semantic Search

Every search process starts with formulating a search query that will best express the user's information need. The type of query dictates the extent of expressiveness and how the search engine has to parse and interpret the query. Traditional search engines represent the query as independent *keywords* of interest to the user. More advanced systems allow to query with *natural language* or *formal query languages*. The first is the most expressive but also most difficult to parse and interpret by machines, the latter is hard to formulate by non-expert users but well processable by machines.

Mediaglobe combines the best of both worlds in a *hybrid* approach. It enables to search for distinct semantic entities instead of keywords and eliminates the ambiguities caused by polysemy and synonymity of natural language. Disambiguation is enforced while entering the search query via type-ahead suggestions. When the user starts to type, the indexed semantic entities matching the entered text best are all displayed. The user has to choose the desired entity from the presented candidate list. Complementing also the traditional keyword-based search, searching for distinct semantic entities improves precision and recall as well as the user experience. To enable efficient entity based search the generated semantic information has to be indexed appropriately. This implies not to store keywords

Position	1	2	3	4	5	6	7
Term Text	john	f	kennedy johnfkennedy dbp:John_F._Kennedy	said	I	love	berlin dbp:Berlin
Type	word	word	word word URI	word	word	word	word URI
Start- / End-Offset	0 / 38	0 / 38	0 / 38 0 / 38 0 / 38	39 / 43	46 / 47	48 /52	53 / 75 53 / 75

Table 1: Search index terms with URIs and offsets after tokenizing and filtering the annotated text

and normalized terms (e. g. via stemming) only, but also to store the URIs representing the semantic entity bound to the text position the entity is determined at. Therefore, we have implemented an extension of the Lucene search engine, which includes support for annotation aware string tokenizing, word-delimiter filters, payload token filters, and efficient snippet highlighting. Before the process of indexing starts, text data is transformed into semantically annotated text containing the URIs of semantic entities being found via NER. Therefore, a simple markup *'(Label){ URI }'* is used, as e. g.:

Text to annotate:	*John F. Kennedy said: 'I love Berlin!'*
Text with markup and URIs:	*(John F. Kennedy){dbp:John_F._Kennedy} said: 'I love (Berlin){dbp:Berlin}!'*

Table 1 shows the final index terms after the annotated text has been processed by annotation aware tokenizer, word-delimiter filter, Lucene standard filter, lowercase filter and finally the payload processing filter, to store the URIs in payload attributes along with the index term. Searching for the URI 'dbp:John_F._Kennedy' would result in returning all documents containing only the specified entity. Traditional full text search is supported as already existing Lucene feature. The ranking of documents is based on the standard TF/IDF-based Lucene scoring.

7 The Semantic Search Interface of *Mediaglobe*

While most keyword-based search systems are optimized to narrow down huge data spaces to the most suitable results and present them in 'ten blue links', semantic exploratory search in *Mediaglobe* additionally aims at finding results, which are not considered to be related at first glance. Semantic exploratory search is based on facet entities and content-based recommendations, enabling the user to better refine and broaden search queries [WKW+10]. Thus, the result space has to distinct from the classical result list of keyword based information retrieval systems. Our interface design objective was to support its users with a quick feedback on selected facet entities that encourages the exploration of the provided results. Considering Fig. 3, the layout of the *Mediaglobe* user interface is arranged in a *Search Area* and a *Result Area*. The *Search Area* contains an input field, and a list of confirmed semantic entities to its right. These entities determine the current

Figure 3: The user interface of *Mediaglobe*: When *brushing* over a facet on the left, the *linked* result tiles and their pagination are updated instantly

search result in the *Search Area* below, where the search results are represented via content snippets. A vertical pagination on the right represents all videos of the current result space. The facet list on the left of the search result contains the facet filters in categories for Persons, Events, Places, Organizations, Concepts, and visual Genres.

The result tiles in the middle of the layout contain the individual videos; they indicate the title of the video, a video thumbnail, and the *Snippet Container*. In case, results derived from ASR, OCR, and visual concept detection, the respective container is captioned with tabs, each of which switching to the corresponding sections and containing extracted ASR or OCR results or a list of detected visual concepts. A timeline below this container exposes the automatically generated temporal segmentation of the videos with highlighted segments that fit the currently selected entities. When users hover over the timeline, the thumbnail of the result tile is updated with an image from the respective segment. This quick video preview provides feedback regarding the structure of a video and leads to simple and concise content-based comparability of different videos. In order to achieve an interaction principle known as 'Brushing and Linking' [OWJS11] the search facets on the left are *linked* to the search results in the main area of the layout and the pagination on its right: when the user *brushes* over one of the facet filters, the video result tiles and their equivalents in the pagination that do not match the criteria are grayed out. By making the facets available for brushing, and linking them to the representative result tiles with the pagination, the proposed interface paradigm provides valuable feedback on how the result set will be changed according to the next distinct selection by the user.

8 Conclusion & Future Work

In this paper, we have presented the *Mediaglobe* semantic video search engine, which provides efficient access to AV archives focussing on historical documentaries. We have shown how we combine methods for visual analysis into a single workflow for automatic metadata generation. We have furthermore presented sophisticated approaches for semantic analysis supported by formal semantic ontologies to enable cross linking of data and explorative search. Finally, we have introduced a novel search interface, which joins the developed technologies and enables efficient visualization and navigation to support the user in his task of retrieving relevant data as well as exploring the archive. Due to copyright regulations on the underlying AV data it is unfortunately not possible to grant public access to the *Mediaglobe* demonstrator. However, a screencast showing the core features and presented technologies is available online[10].

Future work will improve the search ranking by incorporating semantic relations between entities. This enables to rank documents based on *semantic similarity* and not on syntactic similarity as done by traditional search engines. Our future work will also focus on enhanced semantic analysis techniques comprising more ontological information from different sources. This will increase the precision of entity assignments and improve the semantic relatedness between archive items. Furthermore, we plan to interconnect additional elements of the interface via 'brushing and linking'.

References

[AAM10] A. M. Amel, A. B. Abdelali, and A. Mtibaa. Video shot boundary detection using motion activity descriptor. *Jrnl. of Telecommunications*, 2(1):54–59, 2010.

[ALBK09] D. Adjeroh, M. C. Lee, N. Banda, and U. Kandaswamy. Adaptive edge-oriented shot boundary detection. *Jrnl. Image Video Process.*, pages 5:1–5:13, January 2009.

[CDF+04] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, C. Bray, and D. Maupertuis. Visual Categorization with Bags of Keypoints. In *WS. on Statistical Learning in Computer Vision, ECCV*, pages 1–22, 2004.

[CPW93] C.H. Chen, L.F. Pau, and P.S.P. Wang. *Handbook of Pattern Recognition and Computer Vision*. World Scientific, 1993.

[DTCP05] M. Dowman, V. Tablan, H. Cunningham, and B. Popov. Web-assisted annotation, semantic indexing and search of television and radio news. In A. Ellis and T. Hagino, editors, *WWW*, pages 225–234. ACM, 2005.

[EOW10] B. Epshtein, E. Ofek, and Y. Wexler. Detecting Text in Natural Scene with Stroke Width Transform. In *Proc. of Computer Vision and Pattern Recognition (CVPR)*, pages 2963–2970, 2010.

[KSR09] G. Kobilarov, T. Scott, and Y. Raimond. Media Meets Semantic Web — How the BBC Uses DBpedia and Linked Data to Make Connections. In *Proc. of the 6th European*

[10] http://bit.ly/mediaglobe

Semantic Web Conf. on The Semantic Web: Research and Applications, ESWC 2009 Heraklion, pages 723–737, Berlin, Heidelberg, 2009. Springer-Verlag.

[Low04] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Int. Jrnl. of Computer Vision*, 60(2):91–110, November 2004.

[LS11] N. Ludwig and H. Sack. Named Entity Recognition for User-Generated Tags. In *Proc. of the 8th Int. WS. on Text-based Information Retrieval*. IEEE CS Press, 2011.

[MBD06] A. Messina, L. Boch, and G. Dimino. Creating Rich Metadata in the TV Broadcast Archives Environment: The PrestoSpace Project. In *Int. Conf. on Automated Production of Cross Media Content for Multi-Channel Distribution*, pages 193–200, Los Alamitos, CA, USA, 2006. IEEE Comp. Soc.

[MBS08] A. Messina, W. Bailer, and P. Schallauer. Content Analysis Tools – Deliverable 15.4. Technical report, RAI, 2008.

[NLFH12] J. Nandzik, B. Litz, and N. Flores-Herr. CONTENTUS—technologies for next generation multimedia libraries. *Multimedia Tools and Applications*, pages 1–43, 2012. 10.1007/s11042-011-0971-2.

[OWJS11] J. Osterhoff, J. Waitelonis, J. Jäger, and H. Sack. Sneak Preview? Instantly Know What To Expect In Faceted Browsing. 2011.

[SET+12] N. Simou, J.-P. Evain, V. Tzouvaras, M. Rendina, N. Drosopoulos, and J. Oomen. Linking Europe's Television Heritage. In J. Trant and D. Bearman, editors, *Proc. Museums and the Web 2011*. Archives & Museum Informatics, Toronto, 2012.

[SH10] T. Steiner and M. Hausenblas. SemWebVid – Making Video a First Class Semantic Web Citizen and a First Class Web Bourgeois. In *9th Int. Semantic Web Conference (ISWC 2010)*, 2010.

[SW09] C.G.M. Snoek and M. Worring. Concept-Based Video Retrieval. *Foundations and Trends in Information Retrieval*, 2(4):215–322, 2009.

[vAARB09] C. van Aart, L. Aroyo, Y. Raimond, and D. Brickley. The NoTube Beancounter: Aggregating User Data for Television Programme Recommendation. In *Social Data on the Web (SDoW2009)*. CEUR WS. Proc. Vol. 520, 2009.

[WKW+10] J. Waitelonis, M. Knuth, L. Wolf, J. Hercher, and H. Sack. The Path is the Destination – Enabling a New Search Paradigm with Linked Data. In *Proc. Linked Data in the Future Internet at the Future Internet Assembly, Ghent, Belgium, CEUR WS. Proc.*, volume 700, dec 2010.

[WS11] J. Waitelonis and H. Sack. Towards exploratory video search using linked data. *Multimedia Tools and Applications*, pages 1–28, 2011. 10.1007/s11042-011-0733-1.

[YQS12a] H-J. Yang, B. Quehl, and H. Sack. A skeleton based binarization approach for video text recognition. In *Proc. IEEE Int. WS. on Image Analysis for Multimedia Interactive Service (to appear)*, Dublin, Ireland, May 2012. IEEE Comp. Soc.

[YQS12b] H-J. Yang, B. Quehl, and H. Sack. Text detection in video images using adaptive edge detection and stroke width verification. In *Proc. of 19th Int. Conf. on Systems, Signals and Image Processing (IWSSIP)*, Vienna, Austria, April 11–13 2012. IEEE Comp. Soc.

[ZMLS06] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study. *Int. Jrnl. of Computer Vision*, 73(2):213–238, September 2006.

Detecting Source Topics by Analysing Directed Co-occurrence Graphs

Mario Kubek, Herwig Unger
Faculty of Mathematics and Computer Science
FernUniversität in Hagen
Hagen, Germany
kn.wissenschaftler@fernuni-hagen.de

Abstract: This paper describes a new method to determine the sources of topics, that influence the main topics in texts, by analysing directed co-occurrence graphs using an extended version of the HITS algorithm. Additionally, this method can be used to identify characteristic terms in texts. In order to obtain the needed directed term relations the notion of term association is introduced to cover asymmetric real-life relationships between concepts and it is described how they can be calculated by statistical means. In the experiments, it is shown that the detected source topics and the characteristic terms can be used to find similar documents and documents that mainly deal with them in large corpora like the World Wide Web. In doing so iteratively, it is possible to easily follow topics by analysing documents from these corpora using this method. This way, users can be offered this new search function in interactive search systems that goes beyond a simple presentation of similar documents. This application will be elaborated on as well.

1 Introduction and Motivation

The selection of characteristic and discriminating terms in texts through weights, often referred to as keyword extraction or terminology extraction, plays an important role in text mining and information retrieval. In [KU12] it has been pointed out, that graph-based methods for the analysis of co-occurrence graphs are well suited for keyword extraction and deliver comparable results to classic approaches like TF-IDF [SWY75] and difference analysis [HQW06]. Especially the proposed extended version of the PageRank algorithm, that takes into account the strength of the semantic term relations in these graphs, is able to return such characteristic terms and does not rely on reference corpora. In this paper, the authors extend this approach by introducing a method to not only determine these keywords, but to also determine terms in texts that can be referred to as source topics. These terms strongly influence the main topics in texts, yet are not necessarily important keywords themselves. They are helpful when it comes to applications like following topics to their roots by analysing documents that cover them primarily. This process can span several documents.

In order to automatically determine source topics of single texts, the authors present the idea to apply an extended version of the HITS algorithm [Kle98] on directed co-occurrence

graphs for this purpose. This solution will not only return the most characteristic terms of texts like the extended PageRank algorithm, but also the source topics in them. Usually, co-occurrence graphs are undirected which is suitable for the flat visualisation of term relations and for applications like query expansion via spreading activation techniques.

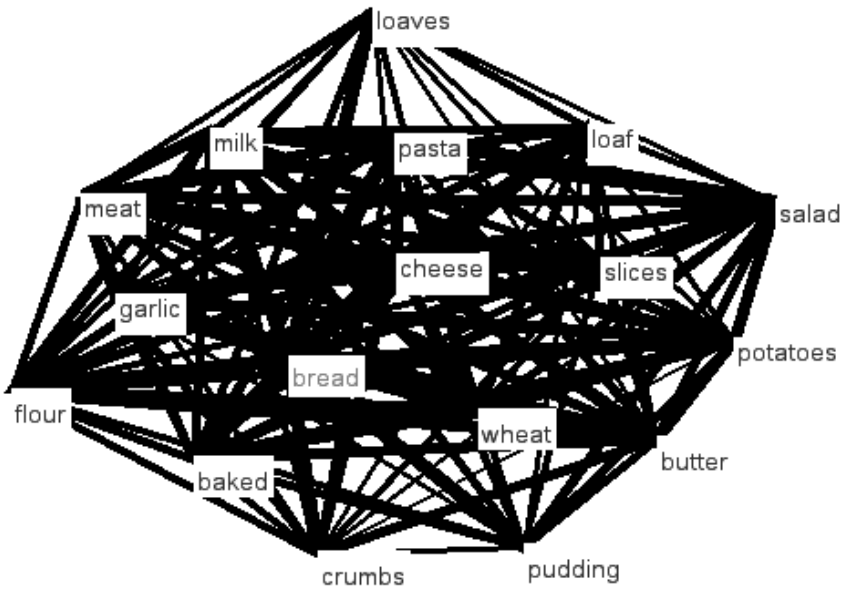


Figure 1: A co-occurrence graph for the word "bread" (http://corpora.informatik.uni-leipzig.de/)

However, real-life associations are mostly directed, e.g. an *Audi* is a *German car* but not every *German car* is an *Audi*. The association of *Audi* with *German car* is therefore much stronger than the association of *German car* with *Audi*. Therefore, it actually makes sense to deal with directed term relations.

The HITS algorithm [Kle98], which was initially designed to evaluate the relative importance of nodes in web graphs (which are directed), returns two list of nodes: authorities and hubs. Authorities, that are also determined by the PageRank algorithm [PBMW98], are nodes that are often linked to by many other nodes. Hubs are nodes that link to many other nodes. Nodes are assigned both a score for their authority and their hub value. For undirected graphs the authority and the hub score of a node would be the same, which is naturally not the case for the web graph. Referred to the analysis of directed co-occurrence graphs with HITS, the authorities are the characteristic terms of the analysed text, whereas the hubs represent its source topics. Therefore, it is necessary to describe the construction of directed co-occurrence graphs before getting into the details of the method to determine the source topics and its applications.

Hence, the paper is organised as follows: the next section explains the methodology used. In this section it is outlined, how to calculate directed term relations from texts by using co-occurrence analysis in order to obtain directed co-occurrence graphs. Afterwards, sec-

tion three presents a method that applies an extended version of the HITS algorithm that considers the strength of these directed term relations to calculate the characteristic terms and source topics in texts. Section four focuses on the conducted experiments using this method. It is also shown that the results of this method can be used to find similar and related documents in the World Wide Web. Section five concludes the paper and provides a look at options to employ this method in solutions to follow topics in large corpora like the World Wide Web.

2 Methodology

Well known measures to gain co-occurrence significance values on sentence level are for instance the mutual information measure [Büc06], the Dice [Dic45] and Jaccard [Jac01] coefficients, the poisson collocation measure [QW02] and the log-likelihood ratio [Dun93]. While these measures return the same value for the relation of a term A with another term B and vice versa, an undirected relation of both terms often does not represent real-life relationships very well as it has been pointed out in the introduction. Therefore, it is sensible to deal with directed relations of terms. To measure the directed relation of term A with term B, which can also be regarded as the strength of the association of term A with term B, the formula 1 of the conditional relative frequency can be used, whereby $|A \cap B|$ is the number of times term A and B co-occurred in the text on sentence level and $|A|$ is the number of sentences term A occurred in:

$$Assn(A \rightarrow B) = \frac{|A \cap B|}{|A|} \tag{1}$$

Often, this significance differs greatly in regards of the two directions of the relations when the difference of the involved term frequencies is high. The association of a less frequently occurring term A with a frequently occurring term B could reach a value of 1.0 when A always co-occurs with B, however B's association with A could be almost 0. This means, that B's occurrence with term A is insignificant in the analysed text. That is why it is sensible to only take into account the direction of the dominant association (the one with the higher value) to generate a directed co-occurrence graph for the further considerations. However, the dominant association should be additionally weighted. In the example above, term A's association with B is 1.0. If another term C, which more frequently appears in the text than A, also co-occurs with term B each time it appears, then its association value with B would be 1.0, too. Yet, this co-occurrence is more significant than the co-occurrence of A with B. An additional weight that influences the association value and considers this fact could be determined by

- the (normalised) number of sentences, in which both terms co-occur or
- the (normalised) frequency of the term A. The normalisation basis could be the maximum number of sentences, which any term of the text has occurred in.

The association $Assn$ of term A with term B can then be calculated using the second approach by:

$$Assn(A \rightarrow B) = \frac{|A \cap B|}{|A|} \cdot \frac{|A|}{|n_{max}|}, \text{where } 0 \leq Assn \leq 1. \quad (2)$$

Hereby, $|n_{max}|$ is the maximum number of sentences, any term has occurred it. A thus obtained relation of term A with term B with a high association strength can be interpreted as a recommendation of A for B. Relations gained by this means are more specific than undirected relations between terms because of their direction. They resemble a hyperlink on a website to another one. In this case however, it has not been manually and explicitely set and it carries an additional weight that indicates the strength of the term association. The set of all such relations obtained from a text represents a directed co-occurrence graph. The next step is now to analyse such graphs with an extended version HITS algorithm that regards these association strengths in order to find the source topics in texts. Therefore, in the next section the extension of the HITS algorithm is explained and a method that employs it for the analysis of directed co-occurrence graphs is outlined.

3 The Algorithm

With the help of the knowledge to generate directed co-occurrence graphs it is now possible to introduce a new method to analyse them in order to find source topics in the texts they represent. For this purpose the application of the HITS algorithm on these graphs is sensible due to its working method that has been outlined in the introduction. The list of hub nodes in these graphs returned by HITS contain the terms that can be regarded as the source topics of the analysed texts as they represent their inherent concepts. Their hub value indicates their influence on the most important topics and terms that can be found in the list of authorities.

For the calculation of these lists using HITS, it is also sensible to also include the strength of the associations between the terms. These values should also influence the calculation of the authority and hub values. The idea behind this approach is that a random walker is likely to follow links in co-occurrence graphs that lead to terms that can be easily associated with the current term he is visiting. Nodes, that contain terms that are linked with a low association value however should not be visited very often. This also means that nodes that lie on paths with links of high association values should be ranked highly as they can be reached easily.

Therefore, the formulas for the update rules of the HITS algorithm can be modified to include the association values $Assn$. In fact, this step is a necessity when dealing with co-occurrence graphs because otherwise less important associations would be treated like more important associations by the HITS algorithm. This is a major difference between such word nets and the World Wide Web, in which the links exist or do not exist at all.

The authority value of node x can then be determined using formula 3:

$$a(x) = \sum_{v \to x} (h(v) \cdot Assn(v \to x)) \tag{3}$$

Accordingly, the hub value of node x can be calculated using formula 4:

$$h(x) = \sum_{x \to w} (a(w) \cdot Assn(x \to w)) \tag{4}$$

The following steps are necessary to obtain a list for the authorities and hubs based on these update rules:

1. Remove stopwords and apply stemming algorithm on all terms in the text. (Optional)
2. Determine the dominant association for all co-occurrences using formula 1, apply the additional weight on it according to formula 2 and use the set of all these relations as a directed co-occurrence graph G.
3. Determine the authority value $a(x)$ and the hub value $h(x)$ iteratively for all nodes x in G using the formulas 3 and 4 until convergence is reached (the calculated values do not change significantly in two consecutive iterations) or a fixed number of iteration has been executed.
4. Order all nodes descendingly according to their authority and hub values and return these two ordered lists with the terms and their authority and hub values.

Now, the effectiveness of this method will be illustrated by experiments.

4 Experiments

4.1 Detection of Authorities and Hubs

The following tables show for two documents of the English Wikipedia the lists of the 10 terms with the highest authority and hub values. To conduct these experiments the following parameters have been used:

- removal of stopwords
- restriction to nouns
- baseform reduction
- activated phrase detection

Table 1: Terms and phrases with high authority and hub values of the Wikipedia-article ”Love”:

Term	Authority value	Term / Phrase	Hub value
love	0.54	friendship	0.19
human	0.30	intimacy	0.17
god	0.29	passion	0.14
attachment	0.26	religion	0.14
word	0.21	attraction	0.14
form	0.21	platonic love	0.13
life	0.20	interpersonal love	0.13
feel	0.18	heart	0.13
people	0.17	family	0.13
buddhism	0.14	relationship	0.12

Table 2: Terms and phrases with high authority and hub values of the Wikipedia-article ”Earthquake”:

Term	Authority value	Term / Phrase	Hub value
earthquake	0.48	movement	0.18
earth	0.30	plate	0.16
fault	0.27	boundary	0.15
area	0.23	damage	0.15
boundary	0.18	zone	0.15
plate	0.16	landslide	0.14
structure	0.16	seismic activity	0.14
rupture	0.15	wave	0.13
aftershock	0.15	ground rupture	0.13
tsunami	0.14	propagation	0.12

The examples show that the extended HITS algorithm can determine the most characteristic terms (authorities) and source topics (hubs) in texts by analysing their directed co-occurrence graphs. Especially the hub list for each text provides useful information to find suitable terms that can be used as search words in queries when background information is needed to a specific topic. However, also the terms found in the authority lists can be used as search words in order to find similar documents. This will be shown in the next subsection.

4.2 Search Word Extraction

The suitability for these terms as search words will now be shown. For this purpose, the five most important authorities and the five most important hubs of the Wikipedia article ”Love” have been combined as search queries and sent to Google. Empiric experiments

have shown that at most five terms and phrases should be used for this purpose. A larger number would limit the search results too much, while too few terms would return too many and possibly irrelevant results. The results of this test can be seen in figure 2 and 3.

Figure 2: Search results for the authorities of the Wikipedia article "Love"

The search results clearly show, that they primarily deal with either the authorities or the hubs. More experiments confirm this correlation. Using the authorities as queries to Google it is possible to find similar documents to the analysed one in the Web. Usually, the analysed document itself is found among the first search results, which is not surprising though. However, it shows that this approach could be a new way to detect plagiarised documents. It is also interesting to point out the topic drift in the results when the hubs have been used as queries. This observation indicates that the hubs of documents can be used as a means to follow topics across several related documents with the help of Google. Hereby, it is desirable that the hubs of the analysed documents are the authorities of the found documents to obtain a chain of documents that are indeed topically depending. This possibility will be elaborated on in more detail in the next and final section of this paper.

Figure 3: Search results for the hubs of the Wikipedia article "Love"

5 Conclusion

In this paper, a new graph-based method to determine source topics in texts based on an extended version of the HITS algorithm has been introduced and described in detail. Its effectiveness has been shown in the experiments. Furthermore, it has been demonstrated that the characteristic terms and the source topics that this method finds in texts, can be used as search words to find similar and related documents in the World Wide Web. Especially the determined source topics can lead users to documents that primarily deal with these important aspects of their originally analysed texts. This goes beyond a simple search for similar documents as it offers a new way to search for related documents, yet it is not impossible to find similar documents when the source topics are used in queries. This functionality can be seen as a useful addition to Google Scholar (http://scholar.google.com/), which offers users the possibility to search for similar scientific articles.

Additionally, interactive search systems can employ this method to provide their users functions to follow topics across multiple documents. The iterative use of source topics as

search words in found documents can provide a basis for a fine-grained analysis of topical relations that exist between the search results of two consecutive queries. Documents found in later iterations in suchlike search sessions can give users valuable background information on the content and topics of their originally analysed documents. In this context, it is also sensible to let users interactively evaluate the topical dependencies of the found documents. Highly relevant results could be marked and act as a suggestion for other users to be further examined. This function would be useful for groups or communities of like-minded people e.g. scientists in a certain field that often deal with domain-specific knowledge whose aspects have topical dependencies. A fast and correct presentation of topically depending documents would be a great help for them, especially when they have a specific information need.

Another interesting application for this method can be seen in the automatic linking of related documents in large corpora. If a document A primarily deals with the source topics of another document B, then a link from A to B can be set. This way, the herein described approach to obtain directed term associations is modified to gain the same effect on document level, namely to calculate recommendations for specific documents. These automatically determined links can be very useful in terms of positively influencing the ranking of search results, because these links represent semantic relations between documents that have been verified in contrast to manually set links e.g. on websites, which additionally can be automatically evaluated regarding their validity by using this approach. Also, these automatically determined links provide a basis to rearrange returned search results based on the semantic relations between them. These approaches will be examined in later publications in detail.

Bibliography

[Büc06] M. Büchler. Flexibles Berechnen von Kookkurrenzen auf strukturierten und unstrukturierten Daten. Master's thesis, Leipzig University, July 2006.

[Dic45] L. R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3):297–302, July 1945.

[Dun93] T. Dunning. Accurate methods for the statistics of surprise and coincidence. *Comput. Linguist.*, 19:61–74, March 1993.

[HQW06] G. Heyer, U. Quasthoff, and T. Wittig. *Text Mining: Wissensrohstoff Text: Konzepte, Algorithmen, Ergebnisse*. W3L-Verl., 2006.

[Jac01] P. Jaccard. Étude Comparative de la Distribution Florale dans une Portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:547–579, 1901.

[Kle98] J. M. Kleinberg. Authoritative Sources in a Hyperlinked Environment. In *Proc. of ACM-SIAM Symp. Discrete Algorithms, San Francisco, California*, pages 668–677, January 1998.

[KU12] M. Kubek and H. Unger. Search Word Extraction Using Extended PageRank Calculations. In Herwig Unger, Kyandoghere Kyamaky, and Janusz Kacprzyk, editors, *Autonomous Systems: Developments and Trends*, volume 391 of *Studies in Computational Intelligence*, pages 325–337. Springer Berlin / Heidelberg, 2012.

[PBMW98] L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report, Stanford Digital Library Technologies Project, 1998.

[QW02] U. Quasthoff and C. Wolff. The Poisson Collocation Measure and its Applications. In *Second International Workshop on Computational Approaches to Collocations*. IEEE, 2002.

[SWY75] G. Salton, A. Wong, and C.S. Yang. A vector space model for automatic indexing. *Commun. ACM*, 18:613–620, November 1975.

Creation and Management of Community-specific Knowledge Domain Taxonomies

[1] Gerald Eichler, [2] Roland Schwaiger

Telekom Innovation Laboratories
Internet & Services: Information Relevance
[1] Deutsche-Telekom-Allee 7, D-64295 Darmstadt, Germany
gerald.eichler@telekom.de

[2] Ernst-Reuter-Platz 7, D-10587 Berlin, Germany
roland.schwaiger@telekom.de

Abstract: The semantic understanding of domain-specific terminology typical for communities, like a scientific conference, requires both, the analysis of daily spoken and written language in documents. A living taxonomy can help to cluster knowledge and classify questions against it, as well as to build up profiles of expertise. Text analysis modules support the parsing of training materials to extract a significant set of key phrases. However, the creation of a balanced categorisation tree requires the experience of community members. The *Spree* project introduces an OntoEditor™ to optimize automatic and manual processes to setup an expert finding & communication application, following well-known Web 2.0 paradigms.

Keywords: Web 2.0, semantic analysis, domain-specific taxonomy, knowledge domain design.

1 Community- and Domain-specific Terminology

When talking about a dedicated technical project, the used language follows a specific terminology. The applied vocabulary is often characteristic for a set of different projects within the same domain, called a *knowledge domain*. Clear definitions are required to create a common understanding among project partners, who form a loose *knowledge community* [Sch00]. Efficient text analysis assists the classification of new text-based materials and the identification of the best expert of an online community [MBA10].

It is also characteristic that written language, used e.g. in specifications, documentations, presentations and manuals, is different from the daily spoken language. Due to self organization processes, the latter is growing with the community and requires much more flexibility and tolerance in interpretation [Ei11].

Starting with a discourse on ontology and taxonomy, their application for knowledge domains is investigated in chapter 1. Chapter 2 illustrates the creation of a new taxonomy, taking the I^2CS as example. Basic linguistic methods are discussed in chapter 3, while chapter 4 introduces the OntoEditor™ as a Java application for knowledge domain management. Finally, in chapter 5 practical experiences are summarized.

1.1 Taxonomy vs. Ontology

There are different methods to preserve knowledge of a domain. Classical document management systems (DMS) support the structured storage of formal documents. However, a much more extensive part of information is exchanged via email, wiki, chat and blog in a highly informal manner. To combine these sources, both, linguistic and semantic tools, are required as known from machine learning [WFH11].

Linguistic modules (LM) help to normalize terminology and extract knowledge from unstructured texts. *Semantic modules* (SM) support to handle complexity, scalability, relations and object types to create a representative semantic network for a certain knowledge domain.

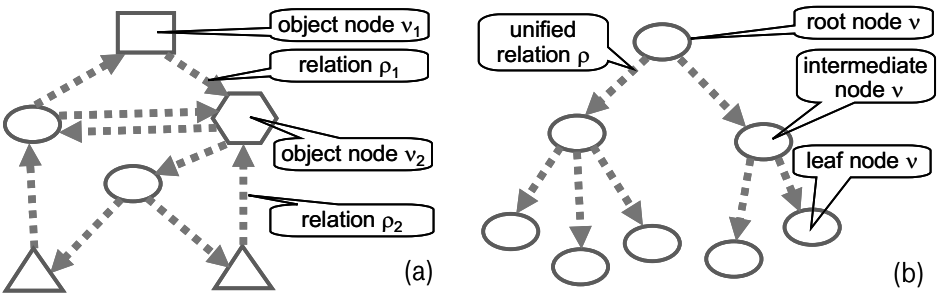


Figure 1: Ontology (a) and taxonomy (b)

As a directed graph, an *ontology* fulfils these requirements excellently, especially the representation of different object types and relations among them. However, it would be very difficult to create the one representative ontology for a certain knowledge domain. The knowledge of a domain grows evolutionary. Therefore, the metaphor of a tree offers key requirements for a hierarchical structure:

(1) New branches and leafs can be added {create}
(2) Existing branches can be split {fork}
(3) Old branches and leafs can die {delete}
(4) Branches or leafs can be bundled {merge}
(5) Situations can be compared and evaluated on similarity {match}
(6) Situations can be represented by multiple branches and leaves {classify}

Requirements (1) to (5) are perfectly met by a *taxonomy*. All edges represent the same relation ρ e.g., "consists of" or "is derived from", depending on the direction of view. All nodes, respectively root node, intermediate node and leaf node, are of the same object type ν, see Figure 1(b). By adding the multi-classification option (6), taxonomies are a simple but powerful means to describe a knowledge domain. Starting from a root, which represents the knowledge domain itself, thematic first level nodes are designed. In semantic context such nodes are called *concepts* and single words and phrases which represent the node, are called *terms*. A concept $\mathfrak{C}$ is represented by:

(A) a **unique identifier** as self-describing and compact node name[1] [mandatory]
(B) an **implicit list** of a set of weighted, machine extracted terms [mandatory]
(C) an **explicit list** of manually added terms as keywords/-phrases [optional]
(D) **sub-concepts**, representing derived lower level (child) nodes [optional]
(E) an **explicit description**, to detail the node context[2] [optional]

The implicit description (B) is also known as *term statistics*. Terms are sorted by its term frequency. The size of a taxonomy is not limited, neither in number of levels nor in number of nodes per level. It is recommended to design the main branches orthogonally in their context to minimize tree overlapping and to maximize independency.

1.2 Exploitation of Taxonomies

The initial node structure of a taxonomy is usually generated by a knowledge domain expert manually. The simplest way is to setup a file system tree, where folders represent the concept identifier (A) and files contain comma or line separated terms (B, C). Once a taxonomy has been created, it is be trained by expanding and re-calculating the term statistics and extended by adding new concepts (1, 2) or modifying them (3, 4).

The knowledge specific taxonomy is exploited in several cases by multi-classification:

Text classification. Any kind of text document including web pages, wikis, chats, emails and simple questions is classified against an existing taxonomy by identifying the nearest concepts.

Profile storage. A community member indicates his or her expertise by selecting the most appropriate concepts from the given taxonomy. Such a profile is represented as an instance of the taxonomy.

Object matching. For a given text fragment, similar text fragments or associated profiles are identified. The degree of similarity can be quantified by an average numeric parameter, calculated over all classified nodes.

Node balancing. By normalization of node weights, a well-adjusted topic assignment is reached. This is required to overcome asymmetries of the availability of training materials.

2 The I²CS Taxonomy

There is a wide range of domain specific knowledge areas. Beside scientific communities, even an event can frame such a knowledge domain. Here, as a domain

[1] The identifier is considered implicitly to be keyword.
[2] This text supports GUI implementations i.e., can be displayed on the mouse-over event.

specific example, the annual International Conference of Innovative Internet Community Systems[3] (I^2CS) is selected.

2.1 Main Branches

Following the history of I^2CS from the very beginning in 2001, three main topic areas were announced. Taking the official Call for papers (CfP) of 2012, they are good candidates for the first level concepts of the I^2CS taxonomy:

- Foundations – Theories, models, algorithms for communities
- Technology – Distributed architectures and frameworks
- Applications and socialization – Communities on the move

The concept identifier (A) should be compact, self-describing and of the same grammar type, singular nouns are preferred. To avoid losing sub-titles, they can be either moved as keywords to the explicit terms (B) or placed as second level nodes. Part (B) is enriched by general applicable descriptors and used synonyms, indicated in italic in Table 1.

To reach a unique and balanced appearance, it is good practise to divide the third bullet into two main branches. So far, there are four main branches with regard to contents: "Foundation", "Technology", "Application" and "Socialization" in Figure 2.

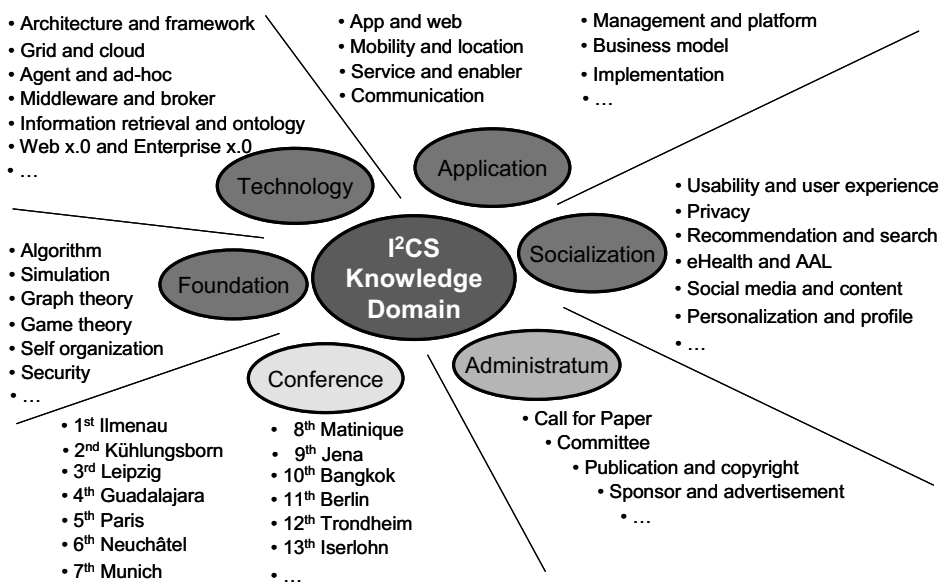


Figure 2: Initial I^2CS knowledge domain taxonomy

[3] URL: http://www.i2cs-conference.org/

The second important source is the event itself, in our case the "Conference". Relevant descriptors here are years and locations, countries and companies as well as the names of participants and chairs.

Semi-permanent information e.g., committees, organisation support go a third type of orthogonal main branch. Let's call it "Administratum".

Table 1: Initial result of the manual I²CS taxonomy approach

1st level node ID	Explicit terms keywords/-phrases	2nd level node ID candidates	Training materials
(A)	(C)	(D), see Figure 2	Source for (B)
Foundation	theory, model, algorithm, *basics, simulation, graph, security*	cfp sub-bullets	related abstracts, authors, submissions, talks, keynotes
Technology	architecture, framework, *middleware, cloud, grid, agent, SOA*	cfp sub-bullets	related abstracts, authors, submissions, talks, keynotes
Application	*service, implementation, search, applet, app, SaaS*	cfp sub-bullets	related abstracts, authors, submissions, talks, keynotes
Socialization	community, *wiki, blog, social media, mobility, privacy, society, profile*	cfp sub-bullets	related abstracts, authors, submissions, talks, keynotes
Conference	*I2CS, IICS, event, venue, program, best paper, panel, session, presentation, talk*	1st, 2nd, 3rd, …, 12th	annual documents (web page, cfp, flyer, poster, program, proceedings foreword)
Administratum	*information, contact, FAQ, call, proceedings, chair, submission, org, registration*	committee, publication	general documents, names, affiliations

2.2 Training Materials

Training materials are any kind of electronically stored documents which contain concept relevant text information for both, the scientific branches and the administrative branches. I²CS related examples are given in Table 1.

There is no need to provide structured materials, as the Language Analysis Module (LAM) will be in charge of extracting text and calculate the term statistics. Sources can be either text files (.txt, .doc, .pdf), or web based URLs (.html, .url, .lnk). Unfortunately, the quality of training materials varies. Therefore, it is preferred, to identify a huge quantity of documents from different sources per node. Missing materials on public topics can be filled by public glossaries e.g., Wikipedia[4], Wikibooks[5], Wiktionaries[6].

[4] URL: http://en.wikipedia.org/
[5] URL: http://en.wikibooks.org/
[6] URL: http://en.wiktionary.org/

3 Automatic and Manual Classification

In the linguistic context, *classification* means, to discover the closest concepts from a provided taxonomy for a given text in a given language.

3.1 Language Analysis Module (LAM)

The Language Analysis Module (LAM) is the key component for the linguistics. It is dictionary-based and therefore language-specific. Dictionaries for English and German were taken from the Weka toolset [We12].

Document type specific converters are responsible for extracting any context relevant text as unstructured stream from the source documents. Beside text documents, any type of Microsoft Office™ documents are supported by specialised converters. By stemming and n-gram building, the document is given a unique footprint with the creation of its term statistics. Figure 3 shows this step-by-step process.

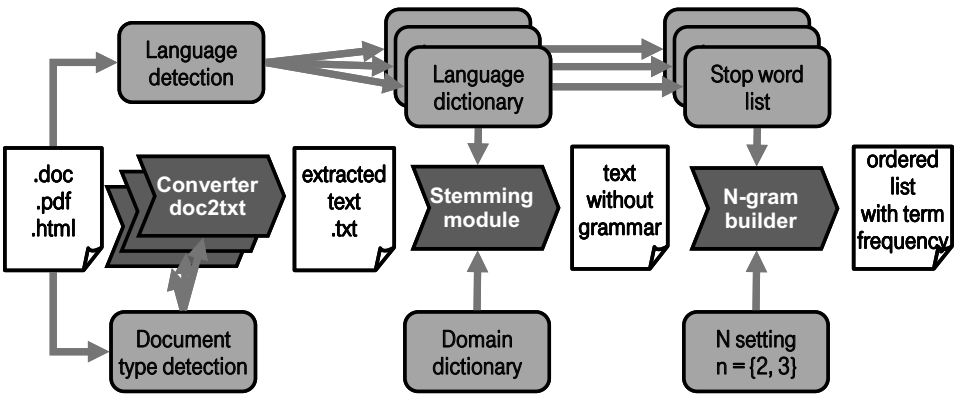


Figure 3: Text analysis process by LAM

Stemming compasses the replacement of nouns by its nominative singular, of verbs by its infinitive present and of adjectives by its neutral singular form. Any capitalisation is ignored and replaced by lower case letters. *N-gram building* comprises the splitting of sentences into semantic phrases of n words. Useful splitting indicators are stop words e.g. conjunctions and prepositions. Meaningless words e.g., articles are dropped. A typical n-gram is the combination of adjective plus noun.

3.2 Design Methodology

The creation phase of a new domain specific taxonomy (see section 1.2) is followed by the training phase (see section 2.2) and applied in the operation phase for matching. Figure 4 illustrates this process.

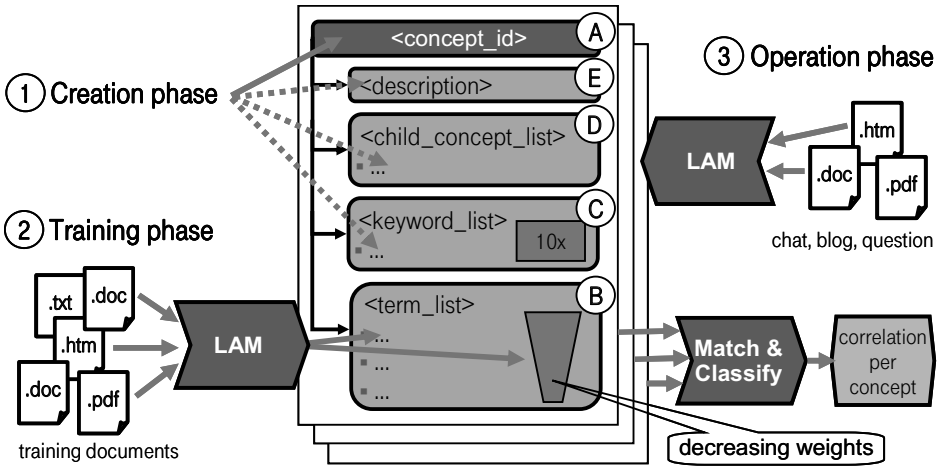


Figure 4: Taxonomy design methodology

Four steps lead to the starting point for the application of a taxonomy to reach the goals of section 1.2:

1. initial structuring of knowledge by a domain expert
   ⇒ basic tree skeleton
2. development of a common understanding by a well-prepared experts' workshop
   ⇒ rich keyword collection
3. manual document classification by single experts
   ⇒ balanced text to node assignment
4. automatic document analysis by machine learning algorithms
   ⇒ ordered, weighted term statistics

4 The Spree™ OntoEditor™

Within the Spree™ project (see section 5.3) an intuitive taxonomy management tool was developed. Even though it handles knowledge taxonomies, in honour to the initial approach it is still called OntoEditor™.

4.1 OntoEditor™ Concept

To allow non-IT community experts to manage and maintain the initial taxonomy, a GUI-based taxonomy editor, the OntoEditor™, was designed. To meet the KISS[7] paradigm, an explorer-like Java application was implemented.

[7] KISS … keep it short and simple

The OntoEditor™ can access any running Spree™ instance on the web by specifying its IP address and providing valid credentials for the role of an “ontology administrator” for that instance. Chances are invoked after an export of the new taxonomy by the “ontology administrator” followed by a restart of the application by the “system administrator”.

Figure 5: Layout of the OntoEditor™

The left column shows the concept tree and allows adding, moving or deleting concepts (nodes) from the taxonomy. Name, description (mouse over event) and explicit keywords for the selected concept are displayed on the upper right. They main window is tab-controlled with tabs for document association, term list, term frequency and the term frequency diagram as shown in Figure 5.

4.2 OntoEditor™ Implementation and Application

If not yet locally available, the Java application can be downloaded from a web site by invocation of a Java start script, including certification. It comes as a zipped set (17 MB) of .jar files.

For the storage format of a taxonomy, several ideas were discussed. The Resource Description Frameworks (RDF)[8] is a common method to define and store ontologies. The Web Ontology Language (OWL and OWL2)[9] by W3C is even more powerful. With Protégé an excellent editing tool is available [KnMu04]. However, to follow KISS again, a standard file system was preferred for the implementation. There are three types of objects:

- Hierarchical folders: representing the concepts
- Document files: containing the trainings materials of the folder above
- Index files: containing IDs, keywords and extracted term statistics

5 Lessons Learnt and Future Dimensions

Both, the generation of knowledge domain specific taxonomies and the development of the Spree™ community application, generated many interesting insights and experiences. This chapter addresses very compact: the community view, the ontology administrator view and the application view.

5.1 People and their Needs for Information

Enterprise and web communities. They have different requirements. Therefore on the one hand, the GUI of an expert search and communication application needs specific adaptation (Figure 7) [REK+11] while the matching core will be the same. On the other hand, the introduction of a community application requires different methodology [Ei11].

Time and duration horizon. The support of a single event requires different organisational and expert competencies than to a long living expert forum. While a single event focuses on management and administrative stuff (How to/FAQ), an expert forum concentrates on contents (a stable knowledge base). The I²CS taxonomy will fill a gap to support a regular event series.

Simplicity. Enterprise communities tend to reflect their administration overhead. However, communities have none or flat hierarchies by its nature. Therefore, the number of roles in an expert system should be strictly limited. As depicted in Figure 6, two roles, the questioner and the expert, are sufficient. This is simplified even more by a single rule: a questioner becomes an expert by creating a personal profile.

Language tolerance. The correctness of language in paper work is decreasing. Moreover, the style of chat and emails is even worse regarding linguistics. Therefore, spelling, grammar and semantic tolerant LAMs are required.

---

[8] URL: http://www.w3.org/TR/rdf-concepts/ (2004)
[9] URL: http://www.w3.org/TR/owl-semantics/ (2004) and http://www.w3.org/TR/owl2-overview/ (2009)

Language redundancy. Firstly, the terminology of experts is different from the language of questioners, as examples in medical domains illustrate. Secondly, the preciseness differs in the comparison of scientific world vs. daily topics.

Re-classification. Community members are asked to re-classify already automatically indexed materials. This will sharpen the selected classifiers.

5.2 Parameters to play with

To optimize the exploitation of a taxonomy, there are many tuning options. However, most of them are only applicable as the size increases. It is important to keep the balance of the size of the taxonomy and the size of the community. The following bullet items summarize shortly practical findings on issues and their potential solution:

- Creation of new nodes: differentiate only if different experts are to be addressed.
- Number of taxonomy levels: follow the Rule of Seven[10].
- Poor training documents: compensate by increased weight of keywords[11].
- Unbalanced set of training documents: taxonomy needs weight normalization.
- Profile fine tuning: introduce skill levels only in large systems.
- Term inheritance: decrease weights of terms derived from child to parent node.

5.3 LAM Implementation in Spree™

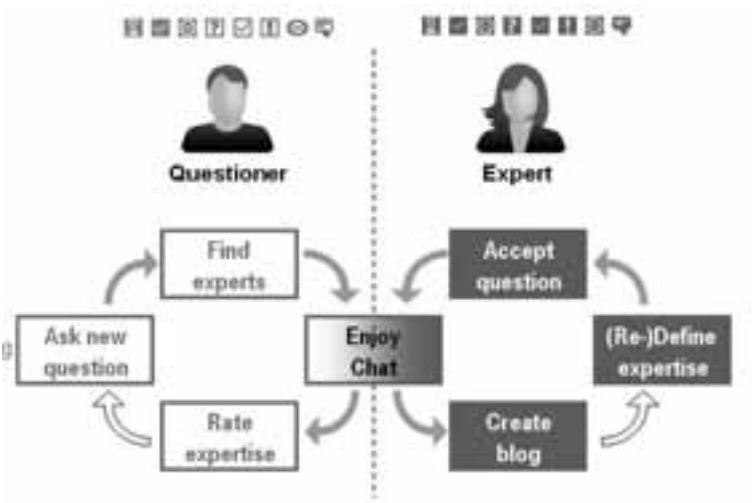


Figure 6: The Spree roles, basic principle and control loops

[10] The Rule of Seven says: limit any human-readable enumeration to a maximum of seven bullet items.
[11] To assign them a term frequency of ten, has been a good practise.

Spree™ is a web application which supports its registered participants in establishing a public (Web 2.0) or corporate (Enterprise 2.0) community [REK+11]. Members will be able to "meet" new colleagues depending on their expert profiles, get the possibility to chat with them and thus extend their contacts. Spree combines linguistic analysis, expert matching, notification, chat, blog and rating functions to initiate and archive asynchronous online discussions.

Spree™ was developed in close corporation of Telekom Innovation Laboratories[12] and the CC IRML of DAI-Labor at Technische Universität Berlin[13]. The matching core was applied in the "Spree – the knowledge exchange network" application for several domains e.g., TechnlogyRadar, system integration testing, customer care services, media recommendation and medicine. A representative example for a scientific community is the implementation for the Lindau Nobel Laureate Meetings, see Figure 7(a) [Ei10]. Figure 7(b) shows the login page of a Spree™ instance for pain management. To optimize the user expectations, a public web survey was established[14].

(b) German skin: information & login page

(a) English skin: personal home page

Figure 7: Spree™ GUI for LNLM (a) and pain management (b)

[12] URL: http://www.laboratories.telekom.com/
[13] URL: http://www.dai-labor.de/en/irml/spree/
[14] URL: http://schmerzlotse.de/

A Spree™ instance for networking among I²CS participants over multiple years is under development and will be soon part of the permanent conference URL at http://www.i2cs-conference.org/ for the upcoming events. Therefore, the presentation of this contribution will contain a discussion (experts' workshop) on future I²CS contents. Any proposals can be sent by email to taxonomy@i2cs-conference.org.

Acknowledgements

Many thanks to Susanne Wieczorek (Foundation Lindau Nobelprizewinners, Lindau), Thomas Strecker (adesso AG, Cologne), Robert Wetzker (aklamio UG, Berlin), Winfried Umbrath (GameDuell GmbH, Berlin), Ernesto William de Luca (University of Applied Sciences Potsdam) and Andreas Heck (CEO DIAmed GmbH, Bensheim), who supported intensively the transfer activities of several Spree™ implementations after the research and development phase.

References

[Ei10] Gerald Eichler: SPREE@Lindau – The Lindau Nobel Laurate Meetings 2010 scientific community. URL: http://www.spree-lindau.de/downloads/Spree-Manual-2S.pdf, June 2010.

[Ei11] Gerald Eichler: Self Organization in Enterprise 2.0 Communities: To introduce a new Experts' Exchange Application. In Autonomous Systems: Development and Trends. Springer Studies in Computational Intelligence, Vol. 391, pp. 189-200, Heidelberg Berlin, November 2011.

[KnMu04] Holger Knublauch, Mark A. Musen: Editing Description Logic Ontologies with the Protégé OWL Plugin. In Proceedings of the 2004 International Workshop on Description Logics (DL2004), pp. 70-78, Whistler, June 2004.

[MBA10] Florian Metze, Christian Bauckhage, Tansu Alpcan: Social and expert search in online communities. In Semantic Computing, P. Sheu, H. Yu, C. V. Ramamoorthy, A. K. Joshi, and L. A. Zadeh, Eds., pp. 341-356. IEEE Press/Wiley, Hoboken, NJ, USA, May 2010.

[REK+11] Andreas Rederer, Gerald Eichler, Thomas Kury, Roland Schwaiger: Spree – Trial-based Improvements transferring an Enterprise into a Web 2.0 Expert Community Application. In proceedings of the 11th I²CS 2011, LNI P-186, pp. 93-104, Koellen, Bonn, June 2011.

[Sch00] Michael Peter Schmidt: Knowledge Communities. 1st edition. Addison-Wesley, Munich, 2000.

[We12] Open Source Sourceforge Project: Weka – Machine Learning Software in Java. URL: http://sourceforge.net/projects/weka/ Last accessed: April 2012.

[WFH11] Ian H. Witten, Eibe Frank, Mark A. Hall: Data Mining: Practical Machine Learning Tools and Techniques. 3rd edition. Morgan Kaufmann, January 2011.

Index of Authors

GI-Edition Lecture Notes in Informatics

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 Lan-guage 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 Magenheim (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 Magenheim, 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 – Er-fahrungen 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 Brömme, Christoph Busch (Eds.): BIOSIG 2003: Biometrics 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.): Persistance, 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, Wofgang 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 Adhoc-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 inforrmatische 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 ist 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 Kowalcyk, 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 Fähnrich, 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 Pousttchi, Klaus Turowski (Hrsg.): Mobile Informationssysteme – Potentiale, Hindernisse, Einsatz

P-77 Jana Dittmann (Hrsg.): SICHERHEIT 2006

P-78 K.-O. Wenkel, P. Wagner, M. Morgenstern, 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 Rößling, 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 Magenheim, 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 manegement 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 Speciel 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 Dreo Rodosek (Hrsg.)
1. DFN-Forum Kommunikations-technologien 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 Informations-systeme (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

P-148 Christian Erfurth, Gerald Eichler, Volkmar Schau (Eds.)
9th International Conference on Innovative Internet Community Systems
I²CS 2009

P-149 Paul Müller, Bernhard Neumair, Gabi Dreo Rodosek (Hrsg.)
2. DFN-Forum Kommunikationstechnologien
Beiträge der Fachtagung

P-150 Jürgen Münch, Peter Liggesmeyer (Hrsg.)
Software Engineering
2009 - Workshopband

P-151 Armin Heinzl, Peter Dadam, Stefan Kirn, Peter Lockemann (Eds.)
PRIMIUM
Process Innovation for Enterprise Software

P-152 Jan Mendling, Stefanie Rinderle-Ma, Werner Esswein (Eds.)
Enterprise Modelling and Information Systems Architectures
Proceedings of the 3rd Int'l Workshop EMISA 2009

P-153 Andreas Schwill, Nicolas Apostolopoulos (Hrsg.)
Lernen im Digitalen Zeitalter
DeLFI 2009 – Die 7. E-Learning Fachtagung Informatik

P-154 Stefan Fischer, Erik Maehle Rüdiger Reischuk (Hrsg.)
INFORMATIK 2009
Im Focus das Leben

P-155 Arslan Brömme, Christoph Busch, Detlef Hühnlein (Eds.)
BIOSIG 2009:
Biometrics and Electronic Signatures
Proceedings of the Special Interest Group on Biometrics and Electronic Signatures

P-156 Bernhard Koerber (Hrsg.)
Zukunft braucht Herkunft
25 Jahre »INFOS – Informatik und Schule«

P-157 Ivo Grosse, Steffen Neumann, Stefan Posch, Falk Schreiber, Peter Stadler (Eds.)
German Conference on Bioinformatics 2009

P-158 W. Claupein, L. Theuvsen, A. Kämpf, M. Morgenstern (Hrsg.)
Precision Agriculture Reloaded – Informationsgestützte Landwirtschaft

P-159 Gregor Engels, Markus Luckey, Wilhelm Schäfer (Hrsg.)
Software Engineering 2010

P-160 Gregor Engels, Markus Luckey, Alexander Pretschner, Ralf Reussner (Hrsg.)
Software Engineering 2010 – Workshopband
(inkl. Doktorandensymposium)

P-161 Gregor Engels, Dimitris Karagiannis Heinrich C. Mayr (Hrsg.)
Modellierung 2010

P-162 Maria A. Wimmer, Uwe Brinkhoff, Siegfried Kaiser, Dagmar Lück-Schneider, Erich Schweighofer, Andreas Wiebe (Hrsg.)
Vernetzte IT für einen effektiven Staat
Gemeinsame Fachtagung Verwaltungsinformatik (FTVI) und Fachtagung Rechtsinformatik (FTRI) 2010

P-163 Markus Bick, Stefan Eulgem, Elgar Fleisch, J. Felix Hampe, Birgitta König-Ries, Franz Lehner, Key Pousttchi, Kai Rannenberg (Hrsg.)
Mobile und Ubiquitäre Informationssysteme
Technologien, Anwendungen und Dienste zur Unterstützung von mobiler Kollaboration

P-164 Arslan Brömme, Christoph Busch (Eds.)
BIOSIG 2010: Biometrics and Electronic Signatures Proceedings of the Special Interest Group on Biometrics and Electronic Signatures

P-165 Gerald Eichler, Peter Kropf, Ulrike Lechner, Phayung Meesad, Herwig Unger (Eds.)
10th International Conference on Innovative Internet Community Systems (I^2CS) – Jubilee Edition 2010 –

P-166 Paul Müller, Bernhard Neumair, Gabi Dreo Rodosek (Hrsg.)
3. DFN-Forum Kommunikationstechnologien
Beiträge der Fachtagung

P-167 Robert Krimmer, Rüdiger Grimm (Eds.)
4th International Conference on Electronic Voting 2010
co-organized by the Council of Europe, Gesellschaft für Informatik and E-Voting.CC

P-168 Ira Diethelm, Christina Dörge, Claudia Hildebrandt, Carsten Schulte (Hrsg.)
Didaktik der Informatik
Möglichkeiten empirischer Forschungsmethoden und Perspektiven der Fachdidaktik

P-169 Michael Kerres, Nadine Ojstersek Ulrik Schroeder, Ulrich Hoppe (Hrsg.)
DeLFI 2010 - 8. Tagung der Fachgruppe E-Learning der Gesellschaft für Informatik e.V.

P-170 Felix C. Freiling (Hrsg.)
Sicherheit 2010
Sicherheit, Schutz und Zuverlässigkeit

P-171 Werner Esswein, Klaus Turowski, Martin Juhrisch (Hrsg.)
Modellierung betrieblicher Informationssysteme (MobIS 2010)
Modellgestütztes Management

P-172 Stefan Klink, Agnes Koschmider Marco Mevius, Andreas Oberweis (Hrsg.)
EMISA 2010
Einflussfaktoren auf die Entwicklung flexibler, integrierter Informationssysteme
Beiträge des Workshops der GI-Fachgruppe EMISA (Entwicklungsmethoden für Informationssysteme und deren Anwendung)

P-173 Dietmar Schomburg, Andreas Grote (Eds.)
German Conference on Bioinformatics 2010

P-174 Arslan Brömme, Torsten Eymann, Detlef Hühnlein, Heiko Roßnagel, Paul Schmücker (Hrsg.)
perspeGKtive 2010
Workshop „Innovative und sichere Informationstechnologie für das Gesundheitswesen von morgen“

P-175 Klaus-Peter Fähnrich, Bogdan Franczyk (Hrsg.)
INFORMATIK 2010
Service Science – Neue Perspektiven für die Informatik
Band 1

P-176 Klaus-Peter Fähnrich, Bogdan Franczyk (Hrsg.)
INFORMATIK 2010
Service Science – Neue Perspektiven für die Informatik
Band 2

P-177 Witold Abramowicz, Rainer Alt, Klaus-Peter Fähnrich, Bogdan Franczyk, Leszek A. Maciaszek (Eds.)
INFORMATIK 2010
Business Process and Service Science – Proceedings of ISSS and BPSC

P-178 Wolfram Pietsch, Benedikt Krams (Hrsg.)
Vom Projekt zum Produkt
Fachtagung des GI-Fachausschusses Management der Anwendungsentwicklung und -wartung im Fachbereich Wirtschafts-informatik (WI-MAW), Aachen, 2010

P-179 Stefan Gruner, Bernhard Rumpe (Eds.)
FM+AM`2010
Second International Workshop on Formal Methods and Agile Methods

P-180 Theo Härder, Wolfgang Lehner, Bernhard Mitschang, Harald Schöning, Holger Schwarz (Hrsg.)
Datenbanksysteme für Business, Technologie und Web (BTW)
14. Fachtagung des GI-Fachbereichs „Datenbanken und Informationssysteme“ (DBIS)

P-181 Michael Clasen, Otto Schätzel, Brigitte Theuvsen (Hrsg.)
Qualität und Effizienz durch informationsgestützte Landwirtschaft, Fokus: Moderne Weinwirtschaft

P-182 Ronald Maier (Hrsg.)
6th Conference on Professional Knowledge Management
From Knowledge to Action

P-183 Ralf Reussner, Matthias Grund, Andreas Oberweis, Walter Tichy (Hrsg.)
Software Engineering 2011
Fachtagung des GI-Fachbereichs Softwaretechnik

P-184 Ralf Reussner, Alexander Pretschner, Stefan Jähnichen (Hrsg.)
Software Engineering 2011
Workshopband
(inkl. Doktorandensymposium)

P-185 Hagen Höpfner, Günther Specht, Thomas Ritz, Christian Bunse (Hrsg.)
MMS 2011: Mobile und ubiquitäre Informationssysteme Proceedings zur 6. Konferenz Mobile und Ubiquitäre Informationssysteme (MMS 2011)

P-186 Gerald Eichler, Axel Küpper, Volkmar Schau, Hacène Fouchal, Herwig Unger (Eds.)
11th International Conference on Innovative Internet Community Systems (I^2CS)

P-187 Paul Müller, Bernhard Neumair, Gabi Dreo Rodosek (Hrsg.)
4. DFN-Forum Kommunikations-technologien, Beiträge der Fachtagung 20. Juni bis 21. Juni 2011 Bonn

P-188 Holger Rohland, Andrea Kienle, Steffen Friedrich (Hrsg.)
DeLFI 2011 – Die 9. e-Learning Fachtagung Informatik
der Gesellschaft für Informatik e.V.
5.–8. September 2011, Dresden

P-189 Thomas, Marco (Hrsg.)
Informatik in Bildung und Beruf
INFOS 2011
14. GI-Fachtagung Informatik und Schule

P-190 Markus Nüttgens, Oliver Thomas, Barbara Weber (Eds.)
Enterprise Modelling and Information Systems Architectures (EMISA 2011)

P-191 Arslan Brömme, Christoph Busch (Eds.)
BIOSIG 2011
International Conference of the Biometrics Special Interest Group

P-192 Hans-Ulrich Heiß, Peter Pepper, Holger Schlingloff, Jörg Schneider (Hrsg.)
INFORMATIK 2011
Informatik schafft Communities

P-193 Wolfgang Lehner, Gunther Piller (Hrsg.)
IMDM 2011

P-194 M. Clasen, G. Fröhlich, H. Bernhardt, K. Hildebrand, B. Theuvsen (Hrsg.)
Informationstechnologie für eine nachhaltige Landbewirtschaftung
Fokus Forstwirtschaft

P-195 Neeraj Suri, Michael Waidner (Hrsg.)
Sicherheit 2012
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 6. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)

P-197 Jörn von Lucke, Christian P. Geiger, Siegfried Kaiser, Erich Schweighofer, Maria A. Wimmer (Hrsg.)
Auf dem Weg zu einer offenen, smarten und vernetzten Verwaltungskultur
Gemeinsame Fachtagung Verwaltungsinformatik (FTVI) und Fachtagung Rechtsinformatik (FTRI) 2012

P-198 Stefan Jähnichen, Axel Küpper, Sahin Albayrak (Hrsg.)
Software Engineering 2012
Fachtagung des GI-Fachbereichs Softwaretechnik

P-199 Stefan Jähnichen, Bernhard Rumpe, Holger Schlingloff (Hrsg.)
Software Engineering 2012
Workshopband

P-200 Gero Mühl, Jan Richling, Andreas Herkersdorf (Hrsg.)
ARCS 2012 Workshops

P-201 Elmar J. Sinz Andy Schürr (Hrsg.)
Modellierung 2012

P-202 Andrea Back, Markus Bick, Martin Breunig, Key Pousttchi, Frédéric Thiesse (Hrsg.)
MMS 2012:Mobile und Ubiquitäre Informationssysteme

P-203 Paul Müller, Bernhard Neumair, Helmut Reiser, Gabi Dreo Rodosek (Hrsg.)
5. DFN-Forum Kommunikations-technologien
Beiträge der Fachtagung

P-204 Gerald Eichler, Leendert W. M. Wienhofen, Anders Kofod-Petersen, Herwig Unger (Eds.)
12th International Conference on Innovative Internet Community Systems (I2CS 2012)

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