

CAMPUS NEWS - Artificial Intelligence Methods Combined for an Intelligent Information Network

Markus Maron, Kevin Read, Michael Schulze

Department of Computer Science, Artificial Intelligence Research Group
University of Koblenz-Landau, Universitätsstr. 1, 56070 Koblenz
[maron|kread|schulzemich]@uni-koblenz.de

Abstract: In this paper we describe a network for distributing personalised information with the usage of artificial intelligence methods. Reception of this information should be possible with everyday mobile equipment. Intelligent filtering and spam protection aim at integrating this technology into our environment. Information on the system architecture and usage of the installation are also presented.

1 Introduction

At the start of each semester much information is presented to the freshmen students. Ranging from basic questions like the location of the registrar's office up to course-specific data, each morsel of information must be looked up in a different part of the campus. Senior students look for office opening hours, announcements for specific courses or extracurricular events. Our concept of developing a campus information system seeks to answer these questions in a personalised way, at any time, at any location. The idea is to enable the user to find and access all information that is of relevance to her. All she needs is a Bluetooth enabled mobile device, like a PDA or mobile phone. All used techniques are in themselves not new or unique, but the combination of instant messaging, Bluetooth-centric transport and profile based information narrowcast is novel. Likewise the involved AI methods are well-known but combine to a complete and complex result. On top of that, the system based on a platform made purely for research is in the stage of evolving into a product and is even now being utilised as a public service on-campus.

This information network is only one piece of the puzzle of our view of a ambient intelligence information network. Previous steps done on the Koblenz campus: in a series of projects funded by the EU (Trial Solution) and BMBF (In2Math, [BGOM04]) we developed "Living Books" [BFGHS03], personalized, intelligent teaching material, which is also available for PDAs and smart mobile phones. There is also an approach to use mobile devices for interaction during classroom teaching¹. Altogether we find a situation on Campus, where students use their mobile device for learning and interacting and for location-based, personalized information.

¹www.mobilelearn.org

Research done in the area of mobile information systems include the projects SmartWEB [Wah07] and SmartKOM [Wah06], which allow queries from anywhere, at any time, utilizing natural language and speech recognition. These technologies require state-of-the-art hardware and broadband mobile network connections. Other groups researching applications on mobile personal computers for ambient intelligence have come to the same conclusion as we have, that the main attention with pervasive applications has shifted from a "use anytime, anywhere" perspective to a location-based, personalized view [RMM05]. A lot of work is happening in this area at the moment. Using a Bluetooth mesh for positioning to send data over non-local wireless links like GSM or GPRS is one avenue to take [AGKO04]. In our approach we opted for positioning and transmission over the same channel. The local wireless link can also be skipped completely, which leads to different usage models [FV02]. A bit closer to our usage scenario of an intelligent university environment than these mentioned projects is the project "mobile cafeteria menu"², although there are neither location-based nor personalized aspects involved. Our approach is in a certain sense a reduction of all mentioned projects as we do not install software on the phone, which is a fragile process, nor do users incur additional costs.

2 Campus News – Concept

The Campus News System is based on results of the research project IASON³, funded by the "Stiftung Rheinland-Pfalz für Innovation". Motivated by the development of powerful mobile devices and the semantic web, we defined a *Semantic Mobile Environment*. In such an environment, so-called service nodes are installed at chosen points of interest. These service nodes broadcast messages to nearby mobile users using Bluetooth wireless technology. For example a bookshop could send its latest offers, or the University restaurant could present its menu or a faculty presents the schedule of events to the students. Each message is annotated with a profile, so that end users will only receive messages that are of interest to them.

This semantic annotation is a logical concept in Description Logic (DL) [BCM⁺03, BHS03]. We also gave the users the opportunity to build their individual interest profile. The first usable prototype of the project (see [Mar05]) was implemented in J2ME, such that the user profile and the inference engine for the personalization was stored in the mobile device.

During several tests in the University and in the City of Koblenz within the framework of the EU project Spatial Metro⁴ it turned out that most mobile phones did not yet fulfill our system requirements. They could not access the Bluetooth wireless functions from Java. Apart from that we learned that the barrier to install software on mobile phones or PDAs is higher than with computers. As of today, users aren't used to application installation on

²<http://www.studentenwerk-dresden.de/mensen/handy.html>

³www.uni-koblenz.de/~iason

⁴www.spatialmetro.org

these devices and as such distrust the idea. To overcome both the technical shortcomings of mobile devices and the need for application installation, we chose to move the decision process (the "reasoning engine") from the mobile phone onto a server, thus eliminating the application. The profile of the user now needs to be entered centrally on a web page. The following describes this solution, which we call "Campus News Information System".

2.1 System architecture

The architecture of the Campus News Information System consists of three components (as shown in figure 1): a web application as the user frontend (blue), a server application (red) in the middle and a freely scalable number of service nodes (green) for delivering the information to the mobile devices. We implemented two different kinds of frontends, one for each group of users. We need an administration interface for the users that want to offer information to the public. We call this frontend the Management Console. We also need a user interface for the recipients of the information, in our case the students. This is called the Userweb; it is depicted in figure 2.

Both Management Console and Userweb access the backend, consisting of a relational database and a server application. The database acts as a central storage for message data, profile data and service node information. Both web frontends store any user-made changes here. The server application also accesses the database, but uses this data to drive the service nodes. As soon as mobile devices are recognized by one of the service nodes, the server looks up the profiles of the corresponding users. This lookup uses a combination of several aspects of the mobile phone including its Bluetooth address to ensure that this matching between mobile phone and user is correct. Using a subsumption check on the annotation of the messages and the users' profile, the server decides on which information conforms to the users' interests. In the next step a history query is made to ensure that no mobile device receives the same information twice. All relevant messages are then transferred to the service node by either wireless or wire-bound networking.

The service nodes scan for mobile devices with activated Bluetooth visibility. After handing this information to the server and receiving the messages, they attempt to transfer this data. After two successive rejections by the mobile device no further attempts will be made for a certain duration, to adapt to students that are not interested in the service.

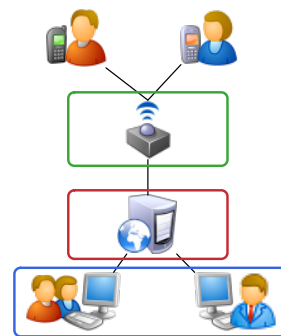


Figure 1: System architecture



Figure 2: Campus News Management Console (left) and Campus News Userweb (right)

2.2 Content entry

Up to now, the system allowed for content to be entered centrally by trusted administrators as described, creating a one-way flow of information. Content will only be received at the end of the flow, i.e. the mobile device, if it fits into the users' need according to a concept filter as outlined above and explained in more detail in 3.1. This has been realized for the Campus News system and is in daily use. Our requirements for an intelligent communication utility required also peer-to-peer or community-centric communication models. This necessitated a bi-directional and decentralized flow of information, which is now being implemented.

The first step is to enable the service nodes to receive messages in a fixed form that is understood by most mobile devices. Using known aspects of the mobile phone and the Bluetooth hardware, the sender can be identified within the network. There are two ways to route the message. It could be a broadcast message in the style of an announcement, which would then be injected into the central database. Example broadcast messages could be *"The car with license plate X is parked with lights on"* or *"The course Y at 1 AM was cancelled"*. Apart from transfer via Bluetooth, this information could then be displayed on public TV screens as a news ticker or on web pages. One problem area here is to filter out inappropriate or hurtful contents, which will be discussed in 3.2.

The other possibility would be to route this message only to a certain set of recipients, thereby enabling personal messaging systems. This platform could then be used for a whole range of applications in the area of ambient intelligence, utilizing the mobile device as a messaging center that is highly portable and ubiquitous. In the context of the intelligent university environment, personal messages with grades, examination appointments and schedule changes could be sent to individuals. In this scenario, as with most peer-to-peer communication, information would not necessarily have to be scanned for malicious content.

3 AI methods

In this section we describe the methods of artificial intelligence which are used within the Campus News environment. As mentioned before we are developing an information system which allows bi-directional flow of information. Each direction requires its own methods of filtering and selection to make sure that only wanted messages pass the system and reach the user.

3.1 Concept filtering

As already mentioned we give the users the possibility to choose the topics they want to be informed about. To achieve this goal we developed a mechanism which is called concept filtering, based on DL profiles and ontologies. A similar server-side reasoning process was already used in the Living Book which is a tool for the management of personalized teaching material. For this the KRHyper [BFGH⁺03], a full first order theorem prover based on the hyper tableau calculus was used. Inspired by the Pocket KRHyper, small version for mobile devices, of the theorem prover, we developed the algorithm for filtering the huge amount of messages which comes within the system.

The reasoning process itself consists of these steps: before transmitting a message to the user, the server has to decide whether the information fits to the users interest or not. This deduction process called matchmaking [KS05] is done by first order reasoning. The user's interests are called a profile. Each message is annotated by its author with a concept. Both consist of Description Logic concepts and are based upon the same terminology. We also built a small ontology for our semantic environment. The decision whether a message matches a user's profile is based on concept satisfiability and subsumption of the DL in use. The task is done by only two queries.

$$profile \sqcap annotation \neq \perp \tag{1}$$

$$annotation \sqsubseteq profile \tag{2}$$

If the annotation satisfies test (1) the annotation is *compatible* with the profile. Because an unsatisfiable annotation will be subsumed by every profile, the first test prevents any unsatisfiable annotation to be considered as a match. This test avoids spam. Test (2) will give a better *match degree* for those annotations that are subsumed by at least one of the *positive* terms. We call these annotations a *match*. This second test is only performed after successfully testing satisfiability (1).

Example

The example shall illustrate the match decisions with respect to a user that is interested in lectures about philosophy and information about the vegetarian menu, but hates sports.

The profile contains the interests $\exists offer.(lectures \sqcap \forall hasTopic.philosophy) \sqcup$
 $\exists offer.(menu \sqcap \forall hasDiet.vegetarian)$ and the disinterest $\exists offer.sports$.
(Note: a user who is requesting something is interested in offering the same thing.)

On her walk through the campus, the mobile user passes different service nodes and receives the messages listed in Table 3.1.

Message Text	Annotation
The menseria offers the delicious menu <i>Lasagne Bolognese</i>	$\exists offers.(meal \sqcap noon)$
and for the vegetarians a salad	$\exists offers.(meal \sqcap noon \sqcap \forall hasDiet.vegetarian)$
An extra curricula lecture about Descartes is offered	$\exists offers.(lectures \sqcap \forall hasTopic.philosophy)$

Table 1: Example messages

With this profile the messages of the lecture and the vegetarian menu are matched, but the message related to the famous menu *Lasagne Bolognese* is rejected.

3.2 Message filtering

Broadcast messages are displayed on public TV screens or transmitted to all interested mobile users. They are entered by the admins into aforementioned "Management Console". The next step is to enable the users to enter their own messages. Appropriate channels would be either the Campus News web interface or sending messages directly from the mobile device via Bluetooth.

The content entered by the admins is trusted content, since it is a closed and well-known user group. But messages that originated in the user base have to be filtered to ensure that hurtful or malicious content is not displayed, since the admins do not control this content. This problem is exacerbated by the fact that users feel anonymous, although the operator can track identity by means of Bluetooth metadata. These messages therefore cannot be inherently trusted. Admins need tools to classify messages as "spam" or "ham". The system should preferably be able to automatically classify messages after setting a few basic settings and manually classifying a few messages.

Many proven approaches exist for classifying email into the two categories "spam" and "ham". Some of them filter messages according to a list of keywords [Heb01]. If one these keywords appear in the message, it is classified as spam, otherwise as ham. Other methods are based on a statistical approach. A well known method is to use Bayesian filters [Sch04]. Another popular approach is to use decision trees [Bla02]. Instance-based approaches are also in use [Bla02]. When used to filter emails, the layout and email headers can also be analysed. Examples for the use of this metadata would be to categorize

emails based on attached figure files or to look into the route the email took according to information in the header.

Format and mode of display for broadcast messages differ significantly from emails. Compared to the well-known form of email spam, our messages are much shorter. Our tests showed that messages should not be longer than 200 characters, otherwise they will not be readable or intelligible, spread out over many screen widths. Our message format does not include headers or layout instructions, so we cannot use this metadata to aid classification. We have a format that makes classification more difficult, but on the other hand there are less possibilities to trick the filter mechanism. All techniques that rely on header information (as with email) are not an option.

Only a few of these methods are of relevance to our project, like the search for keywords, or rather more "key phrases". The admin can edit this list of keywords. For categorization Bayesian filtering, decision trees and instance-based methods are interesting too. All these approaches have to be tested with respect to our specialized situation.

When users or certain mobile devices abuse the system, the admins can blacklist [Heb01] these entities. There is also a whitelist [Heb01], to specify senders that never need to be filtered.

4 Results and outlook

Now, ten weeks after introducing the Campus News System at the University of Koblenz, we are pleased to say that the usage and acceptance by the students is very high for this short time frame. It will be interesting to see if acceptance will climb even higher in the future. The ratio of found devices to devices that received information was at 12.8% in April 2007. This ratio climbed to 47.1% in June. We consider this to be the number of Bluetooth capable devices owned by users willing to activate Bluetooth functionality, divided by the number of Campus News adopters. We detected over 2200 different mobile devices with Bluetooth activated and served 675 of them. These 675 are comprised of 464 unregistered users that received the cafeteria menu and urgent public announcements, and 211 registered users that got news according to the profile they set. All in all we transmitted over 4078 different messages in this short time frame (see table 2). To put the numbers into perspective, the campus Koblenz has around 5000 students. Taking into account occasional visitors, more than one third of all people on-campus have activated Bluetooth and more than ten percent have received Campus News information.

We also did a questionnaire on user wishes and opinion regarding Campus News. We sent out a code and asked the students to enter that code on the answer sheet. On top of that we asked about mobile phone brand and model, opinion of the system in general and wishes or suggestions for future work. Of the 97 students that replied, 12 could not receive the code. Using the stated information about the mobile phone brand we got insight into the workings of Samsung and Motorola brand phones and could increase the compatibility in this area. The opinions varied from general vague acceptance of the concept up to enthusiasm. The most wanted feature was a higher density of service nodes and up-to-date

	April	May	June	since Roll-out (16/04)
found devices	1079	785	1103	2286
served devices	139	163	520	675
transmitted data	828	903	2347	4078

Table 2: Usage of the Campus News System

information in the system for course schedule changes.

The next step is building a pervasive community by extending the system for reception of messages. Combined with the concept of intelligent categorization and thus personalization, this community enables ubiquitous mobile devices to become intelligent information centers.

Acknowledgments

Campus News has been carried out by the cooperative work between the University of Koblenz, Studierendenwerk Koblenz, Arbeitsgruppe Künstliche Intelligenz and wizAI Solutions GmbH and is based on the research work within the IASON project, which is funded by the "Stiftung Rheinland-Pfalz für Innovation".

References

- [AGKO04] Lauri Aalto, Nicklas Göthlin, Jani Korhonen, and Timo Ojala. Bluetooth and WAP push based location-aware mobile advertising system. In *MobiSys '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 49–58, New York, NY, USA, 2004. ACM Press.
- [BCM⁺03] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003.
- [BFGH⁺03] Peter Baumgartner, Ulrich Furbach, Margret Gross-Hardt, Thomas Kleemann, and Christoph Wernhard. KRHyper Inside — Model Based Deduction in Applications. In *Proc. CADE-19 Workshop on Novel Applications of Deduction Systems*, 2003.
- [BFGHS03] Peter Baumgartner, Ulrich Furbach, Margret Groß-Hardt, and Alex Sinner. 'Living Book': -'Deduction', 'Slicing', 'Interaction'. In *CADE*, pages 284–288, 2003.
- [BGOM04] Peter Baumgartner, Barbara Grabowski, Walter Oevel, and Erica Melis. In2Math - Interaktive Mathematik- und Informatikgrundausbildung. *Softwaretechnik-Trends*, 24(1):36–45, 2004.
- [BHS03] Franz Baader, Ian Horrocks, and Ulrike Sattler. Description Logics as Ontology Languages for the Semantic Web, 2003.

- [Bla02] Kai Blankenhorn. Spam-Filterung mittels Maschinellern Lernen. Master's thesis, Fachhochschule Furtwangen, 2002.
- [FV02] Alois Ferscha and Simon Vogl. Pervasive Web Access via Public Communication Walls. In *Pervasive '02: Proceedings of the First International Conference on Pervasive Computing*, pages 84–97, London, UK, 2002. Springer-Verlag.
- [Heb01] Wolf-Guenter Hebel. Spam mail classification system. Master's thesis, University of Georgia, 2001.
- [KS05] Thomas Kleemann and Alex Sinner. Description logic based matchmaking on mobile devices. In Joachim Baumeister and Dietmar Seipel, editors, *1st Workshop on Knowledge Engineering and Software Engineering - KESE2005*, pages 37–48, 2005.
- [Mar05] Markus Maron. IASON Mobile Application - Konzept und Realisierung einer mobilen Anwendung für profilbasiertes Matchmaking von Nachrichten. Master's thesis, Universität Koblenz-Landau, 2005.
- [RMM05] G. Roussos, S. Maglavera, and A. Marsh. Enabling Pervasive Computing with Smart Phones. *IEEE Pervasive Computing*, 4(2), 2005.
- [Sch04] Bernhard Schüler. Aggregator - A Tool to Bootstrap Text Classification. Master's thesis, Universität Koblenz-Landau, 2004.
- [Wah06] Wolfgang Wahlster. *SmartKom: Foundations of Multimodal Dialogue Systems*, volume XVIII of *Cognitive Technologies*, chapter Dialogue Systems Go Multimodal: The SmartKom Experience, pages 3–27. Springer, Heidelberg, Germany, 2006.
- [Wah07] Wolfgang Wahlster. SmartWeb - Ein multimodales Dialogsystem für das semantische Web. In B. Reuse and R. Vollmar, editors, *40 Jahre Informatikforschung in Deutschland*, Heidelberg, Berlin, 2007. Springer.