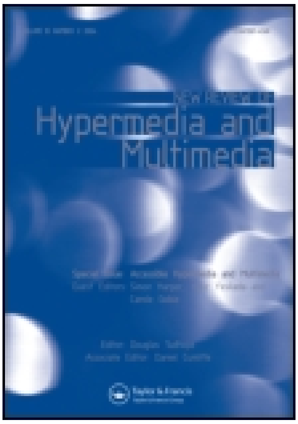

New Review of Hypermedia and Multimedia
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tham20

Ubicon and its applications for ubiquitous social computing

Martin Atzmueller[a], Martin Becker[b], Mark Kibanov[a], Christoph Scholz[a], Stephan Doerfel[a], Andreas Hotho[b], Bjoern-Elmar Macek[a], Folke Mitzlaff[a], Juergen Mueller[a] & Gerd Stumme[a]

[a] Knowledge and Data Engineering Group, University of Kassel, Kassel, Germany

[b] Data Mining and Information Retrieval Group, University of Würzburg, Würzburg, Germany

Published online: 18 Mar 2014.

To cite this article: Martin Atzmueller, Martin Becker, Mark Kibanov, Christoph Scholz, Stephan Doerfel, Andreas Hotho, Bjoern-Elmar Macek, Folke Mitzlaff, Juergen Mueller & Gerd Stumme (2014) Ubicon and its applications for ubiquitous social computing, New Review of Hypermedia and Multimedia, 20:1, 53-77, DOI: 10.1080/13614568.2013.873488

To link to this article: http://dx.doi.org/10.1080/13614568.2013.873488

New Review of Hypermedia and Multimedia, 2014
Vol. 20, No. 1, 53–77, http://dx.doi.org/10.1080/13614568.2013.873488

Ubicon and its applications for ubiquitous social computing

MARTIN ATZMUELLER[†]*, MARTIN BECKER[‡], MARK KIBANOV[†], CHRISTOPH SCHOLZ[†], STEPHAN DOERFEL[†], ANDREAS HOTHO[‡], BJOERN-ELMAR MACEK[†], FOLKE MITZLAFF[†], JUERGEN MUELLER[†] and GERD STUMME[†]
[†]Knowledge and Data Engineering Group, University of Kassel, Kassel, Germany
[‡]Data Mining and Information Retrieval Group, University of Würzburg, Würzburg, Germany

(*Received 23 February 2013; final version received 5 December 2013*)

The combination of ubiquitous and social computing is an emerging research area which integrates different but complementary methods, techniques, and tools. In this paper, we focus on the Ubicon platform, its applications, and a large spectrum of analysis results. Ubicon provides an extensible framework for building and hosting applications targeting both ubiquitous and social environments. We summarize the architecture and exemplify its implementation using four real-world applications built on top of Ubicon. In addition, we discuss several scientific experiments in the context of these applications in order to give a better picture of the potential of the framework, and discuss analysis results using several real-world data sets collected utilizing Ubicon.

Keywords: Social computing; Ubiquitous computing; Data mining; Social sensing; Applications

1. Introduction

The idea of ubiquitous computing introduced by Weiser (1991, 1993) as omnipresent, unobtrusively, and invisibly functioning information systems continues to shape information systems in our daily lives. Similarly, social computing (Parameswaran and Whinston 2007, Wang *et al*. 2007) has a continuing impact on system and service development. Both involve the users' contexts. Applications use sensor networks to collect context information for anticipating the goals and intentions of the users. In consequence, the ubiquitous system adapts automatically to changing interests and integrates several services to increase its use. Additionally, devices are becoming increasingly smaller and are integrated into everyday objects. Connecting the social and the physical world is then one of the challenges in ubiquitous and social computational systems, for example, the integration of social connections, communities, and collaborative applications.

*Corresponding author. Email: atzmueller@cs.uni-kassel.de

In this paper, we focus on Ubicon[1] and its applications. Ubicon is a software platform for implementing ubiquitous and social applications; it aims at supporting applications at the intersection of ubiquitous and social computing, integrating functionalities of both environments. Ubicon provides a number of components for data collection, processing, and serving. At its core, Ubicon provides the means for creating and hosting ubiquitous and social applications. Grounded by several main principles of the lambda architecture (Marz and Warren 2013), Ubicon features flexible ways for adaptations and extensions in the respective applications.

Our contribution can be summarized as follows: we present the Ubicon platform and summarize its basic architecture and components. We then provide a detailed view on four real-world applications that have been successfully implemented using Ubicon; Conferator[2] and MyGroup[3] focus on enhancing ubiquitous social interactions, while WideNoise[4] and AirProbe[5] are applications for collaborative noise collection and air quality monitoring, respectively. In these respective application contexts, we discuss data mining and analysis methods for comprehensively sketching the capabilities and the potential of the combination of ubiquitous and social computing. These methods range from social interaction networks of face-to-face contacts to participatory open-sensing in environmental application contexts. The presented data mining techniques and components of the respective applications include methods from localization to user recommendation.

The remainder of this paper is structured as follows. Section 2 provides an overview on the Ubicon platform. After that, Section 3 summarizes four example applications and exemplifies different data mining and analysis techniques implemented in the respective applications. Next, Section 4 provides analysis results using real-world data from the presented systems. Finally, Section 5 discusses related work, and Section 6 concludes the paper with a summary and promising options for future work. This article is an extended version of Atzmueller *et al*. (2012a).

2. Ubicon—an overview

In the following, we first briefly provide an overview on the Ubicon platform. From an architecture perspective, we outline the different layers and core modules provided by Ubicon. Since applications typically have specialized and unique requirements, Ubicon offers support by providing an application framework with several core components that can be utilized and extended by applications as needed. Using these core components, it is not necessary, for example, to implement the same components that are typical for ubiquitous and social systems again and again. In addition, the data from different applications built on Ubicon can be easily combined for enhancing the overall functionality of the different system.

From a data-centric view, Ubicon implements data storage, processing, and serving pipeline similar to the *lambda architecture* (Marz and Warren 2013). In that way, core concepts such as immutability and recomputation are transparently enabled by the platform. Figure 1 shows a conceptual overview of the system's architecture. The data flow is organized in three layers shown in the left side of the figure. The functionality for each of these layers is backed by the Ubicon core (shown in the right), which provides canned functionality, i.e. framework classes and interfaces, which can be utilized throughout different applications.

The functionality of the layers and core components is summarized in the following:

- Immutable data storage layer: built on the principle of data immutability, the first layer receives and stores the *input data*. The data storage layer provides a general mechanism for storing ubiquitous data that are directly stored in the *input data table* without further processing for enabling a low-complexity overhead and for avoiding data corruption issues.
- Data-processing layer: data processing aspects involve scalability, extensibility, and debuggability. Such aspects are handled by the data-processing layer. It utilizes a set of data processors that use the input data table and generate application-specific *processed data* tables. The application of the data processors is repeatable, as a mapping from input data to processed data, providing a general and extensible mechanism. In addition, scalability issues can be addressed, for example, by utilizing frameworks such as map/reduce (Dean and Ghemawat 2008).
- Data-serving layer: the data-serving layer features data access to the processed data. On the framework level, it provides access, for example, via database connectors. Using these, generic data access can be easily implemented.
- Ubicon core: This core framework functionality includes basic privacy components, data analysis methods, a generic data logic—specifically on the database level, a query API, social connectors, a basic user management system, and core web components. These are implemented with a model-view-controller pattern using the Spring framework[6]. Using these, Ubicon can be deployed using a standard servlet container, e.g. Apache Tomcat.[7]

Typically, applications utilize the provided core components, interfaces and classes, and extend the overall workflow according to their individual application

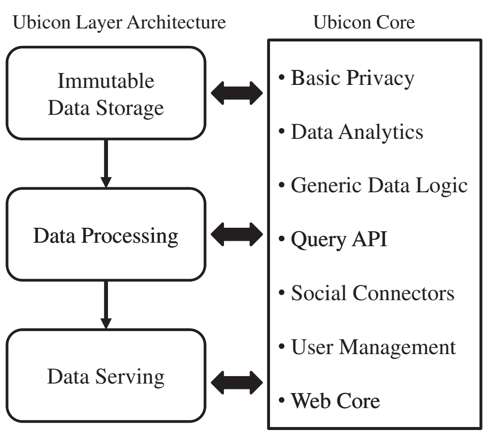

Figure 1. Conceptual overview on the architecture of the Ubicon software platform.

requirements. We will discuss specific examples in the next section when describing several applications built on top of Ubicon.

3. Applications

MyGroup and Conferator—social computational systems for enhancing interactions in working groups and at conferences, respectively—are maintained and developed by the Knowledge and Data Engineering (KDE) Group[8] at the University of Kassel in the context of the inter-disciplinary VENUS research cluster,[9] which is concerned with the design of social, legal, and technological networking issues in situated ubiquitous systems. The WideNoise and AirProbe web applications for environmental sensing with respect to noise and air pollution are jointly developed by the Data Mining and Information Retrieval (DMIR) Group[10] at the University of Würzburg and the KDE group at the University of Kassel in the context of the EU project *EveryAware*.[11]

3.1. MyGroup: social networking in working groups

MyGroup aims at supporting members of working groups. It employs active RIFD (radio-frequency identification) tags for localizing the members and for monitoring their social contacts. These so-called *proximity tags* have been developed by the SocioPatterns collaboration[12] and are able to detect face-to-face proximity of individuals wearing them. The face-to-face proximity of two persons usually implies that they are engaged in a conversation. The proximity tags send out two types of radio signals: proximity-sensing signals and tracking signals. Proximity-sensing signals are emitted at a low power level and are used for the detection of face-to-face proximity. For localization purposes, the proximity tags send out tracking signals at different power levels that are received by RFID readers at fixed positions in the target area (typically a room in a building). A technique that allows the detection of individuals at room-level basis is presented in Scholz *et al*. (2011). Furthermore, the system provides profile information including links to (external) social software, e.g. BibSonomy[13] (Benz *et al*. 2010), Twitter, Facebook, or XING.

MyGroup has been applied at a number of different events: it is being used by the KDE Group at the University of Kassel, and is currently being extended for use in a larger research cluster. MyGroup has also been utilized at a large student party,[14] for supporting organizational processes, at the First International Changemaker Camp at the University of Kassel, and at a CodeCamp for supporting software development processes.

The collected contact and location data can be utilized for different research purposes, as described below. Both raw and processed data are stored, thus allowing to check inconsistencies or incompleteness of the data—utilizing the available raw data—according to the Ubicon layer architecture. Processed data in this context means, for example, a face-to-face contact between individuals or a person's position at the specific point in time. As in Szomszor *et al*. (2010), the data processor identifies a face-to-face contact when the corresponding proximity tags detect each other for more than 20 seconds. A contact ends, when both proximity tags do not detect each other for more than 60 seconds. In the database we store the contact time interval as well as the usernames of those persons who wear these proximity tags. So far, we collected approximately 2 terabytes of raw data that are now the main source for our ongoing research about structures and behavior within research groups.

3.1.1. Overview—functionality. MyGroup provides several functions for improving interactions and the discussions in working groups: each user can inspect his/her face-to-face contacts with other users on an individual basis—complemented by the provided *contact information* for each participant. The *timeline* (cf. Figure 2) is an aggregation of different activities of the group: it provides an aggregated view on the currently active topics published on Twitter or BibSonomy by the members of the group and the conversations that recently happened. The timeline which is displayed on a large LCD screen often stimulates interesting research discussions and enables enhanced dissemination and exchange of knowledge.

The *map view* (cf. Figure 3) enables an easy localization of the group members. Elaborate *user profiles* (similar to those of the Conferator, see Figure 5) provide detailed information, for example, about position and interests of a group member.

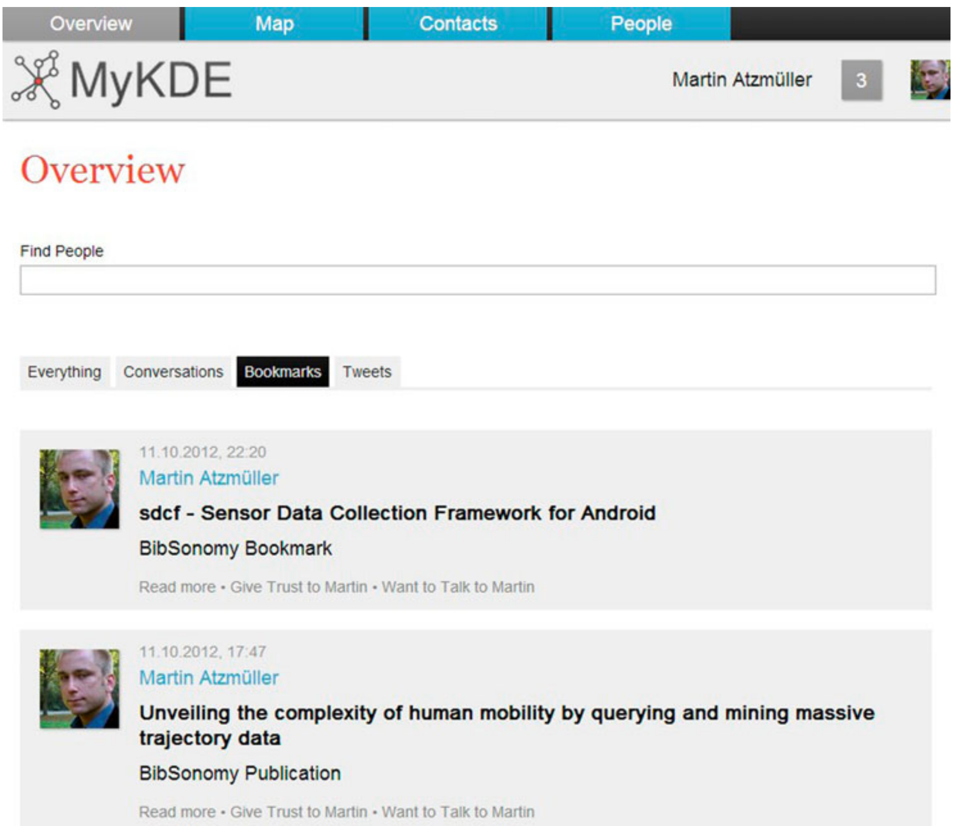

Figure 2. A screenshot of the *timeline view*: the screenshot shows two recent BibSonomy posts.

Utilizing the system, we can exploit the social information for supporting different interactions; we can recommend, for example, persons based on their expertise in (software) projects, as described below. Different trust and privacy settings, for example concerning the visibility of contacts and locations, allow a selective distribution of sensitive information. The system is continuously refined according to the user feedback and usability studies, in particular, in the VENUS project. For more details on a successful evaluation of the applied approach, cf. Geihs *et al*. (2012), and in the context of Ubicon, we refer to Behrenbruch *et al*. (2013).

3.1.2. Recommending software developers. Identifying experts for a given problem is one of the main challenges when working in a large team. In the context of MyGroup, we focused on supporting software development groups (Macek *et al*. 2011). The presented approach can potentially be generalized for any organization using revision control systems, e.g. for recommending collaborators based on changes in documents, papers, or wikis.

In the software development context, we analyze code changes and the structure of the software projects. In this way, we create resource trees resembling the hierarchic organization of source files. The contribution of each developer is then measured by the number of changed lines of code. We combine this

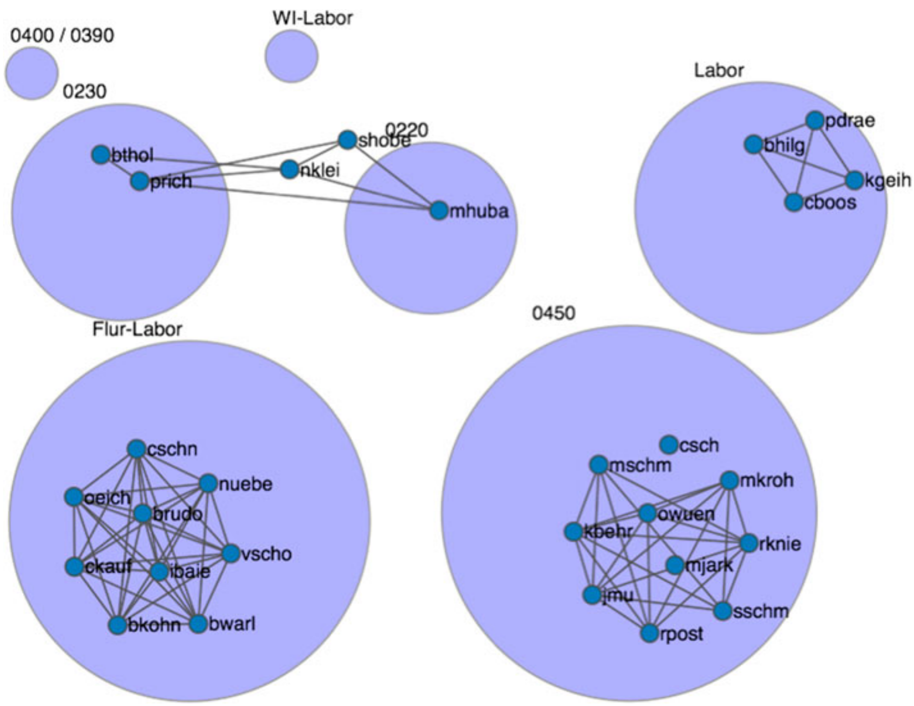

Figure 3. A screenshot of the *map view* of MyGroup. The large circles denote rooms, the smaller circles participants; connections between those indicate ongoing conversations. The screenshot also contains two participants who are changing rooms.

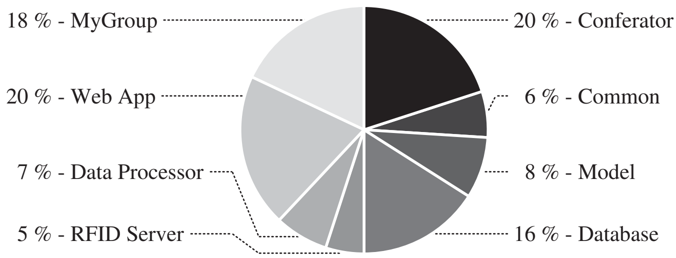

Figure 4. Expertise profile of a developer generated by the analysis of logs from a revision control system and RFID communication for the Ubicon system.

information with the RFID contact graph of the developers: in addition to weighted edges which reflect the relative amount of changed lines of code, we consider edges between the developers. These edges are weighted by the cumulative duration of the face-to-face contacts of the developers within the last eight hours before committing changes to the source code. Thereby, we connect their real-life communication and interaction using the MyGroup system. The resulting structure captures important knowledge of a social group: exogenous information, e.g. developers writing code by themselves, and endogenous information representing the knowledge transfer from one individual to another by means of communication.

For recommending software developers, the PageRank algorithm (Brin and Page 1998) is applied to the extended contact graph which combines the resource and developer contact graph. The output can either be an ordered list of developers that are supposed to know mostly about a specific source file or an expertise profile. An example analyzing the source code of the Ubicon framework is depicted in Figure 4; it shows the personal experience (expertise profile) of one developer for a set of modules of the Ubicon system. Such information can be used to support project managers in their task to assign work packages to certain developers or to organize the office structure such that developers who might profit from each other's knowledge are seated together.

As described in Macek *et al*. (2011), an evaluation of the recommendation method in the context of MyGroup outperformed baseline methods based on pure lines-of-code analysis for the recommendation. Due to the limited space, we refer to Macek *et al*. (2011) for more details of the applied algorithm and ranking methods.

3.2. Conferator: a social conference guidance system

Conferator (Atzmueller *et al*. 2011) is a social and ubiquitous conference guidance system, aiming at supporting conference participants during conference planning, attendance, and their post-conference activities. It features the ability to manage social and face-to-face contacts during the conference (based on the same technology as MyGroup) and to support social networking.

3.2.1. Overview—functionality. At its core, Conferator comprises two key functionalities: Conferator helps to manage organizational information like the conference schedule. Furthermore, Conferator provides information about personal social contacts, by providing context-sensitive information, e.g. the location of other conference participants or a contact history using the *timeline* view. Furthermore, the users can browse the list of participants to search for acquaintances or friends. The corresponding *user profiles* provide additional information, cf. Figure 5. Similar to MyGroup, Conferator offers several privacy settings in order to enable privacy protection, e.g. for sharing locations or contact information.

Conferator utilizes the same RFID technology for indoor localization and face-to-face contact detection similar to the MyGroup application, cf. Section 3.1. Conferator also provides information about the conference schedule; it contains information about talks, i.e. the authors, time, and place of the talk. The talks are usually assigned to sessions, which are assigned to tracks. Combining the conference schedule with localization information can deliver interesting information, e.g. "Who visited which talk?" or "Which talks, sessions and tracks were the most popular during the given event?" (cf. Atzmueller *et al*. 2012b, Macek *et al*. 2012).

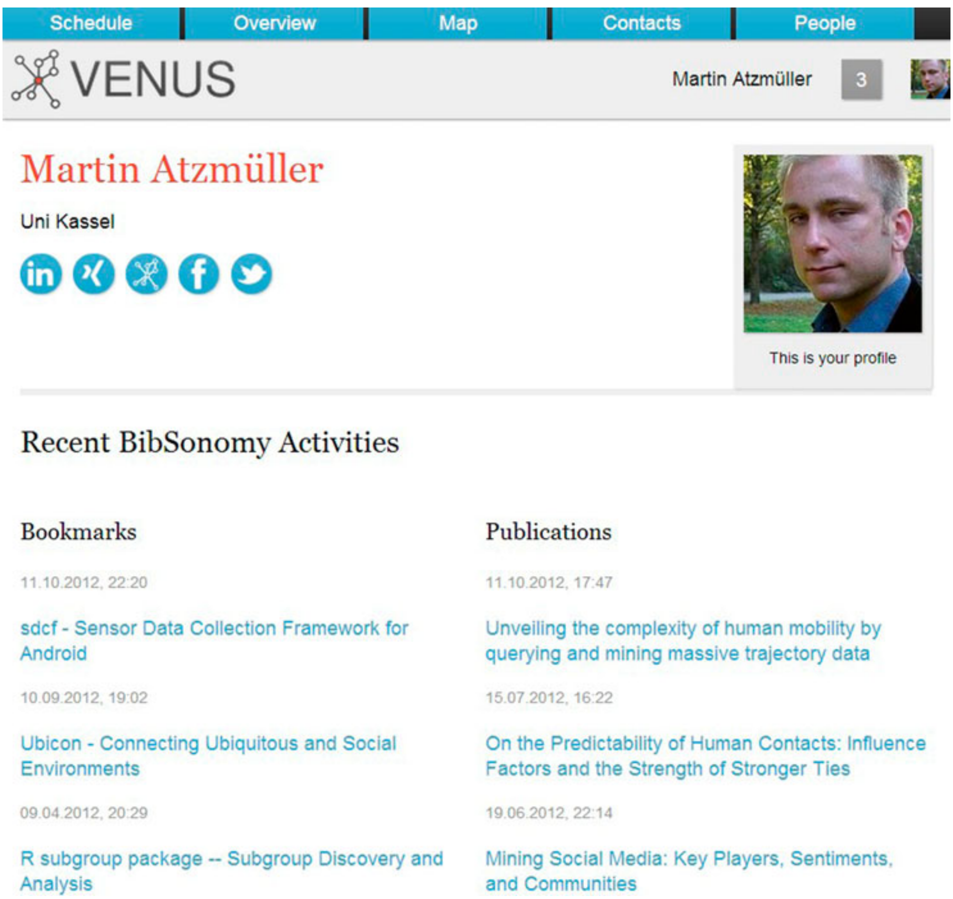

Figure 5. A screenshot of a Conferator user profile with context information and latest posts.

Conferator has successfully been applied at conferences for special interest groups of the German Computer Science Society (GI)—LWA 2010[15] (Atzmueller *et al*. 2010, 2012b), 2011,[16] and 2012,[17] and at the ACM Hypertext 2011[18] conference (Macek *et al*. 2012).

3.2.2. Mapping live interactions. Capturing and visualizing live interactions of individual users are an important task for MyGroup and Conferator. Therefore, a localization framework is a central component for such a system. The indoor localization component provides the location of all the users, and shows where their conversations take place.[19] During conferences, for instance, Conferator offers the possibility of observing who is visiting a given talk, thus facilitating the academic exchange during the subsequent coffee breaks. Furthermore, it is possible to identify hotspots, e.g. conference rooms where a large number of conference participants are listening to—apparently interesting—talks, and to potentially recommend those to undecided participants.

The localization framework consists of two parts, the positioning component and the visualization of the individual locations (see Figure 3). The individual user positions at room level are computed based on the positions of the RFID readers in the target area. Further, it exploits the proximity information of users to improve the localization accuracy. Using the social proximity information in the *Social Boosting* algorithm (Scholz *et al*. 2011), the accuracy could be improved from 84% (baseline algorithm using only the readers' positions) to nearly 90%, as evaluated during the poster session at LWA 2010. For more details, see Scholz *et al*. (2011).

3.2.3. Communities and recommending user interactions. Intelligent recommendations include in particular the suggestion of *interesting* contacts, topics, or context-specific reminders for tasks or people. These often motivate certain *actions*, e.g. contacting other users or working with certain resources. The *Acquaint-O-Matic* for Conferator and MyGroup can then be applied for directly recommending users or for browsing a list of *similar users*. The former generates a personalized list of users that the current user might know or might be interested to know. It provides links to the suggested users' profile pages and encourages to get in touch with those users. The *similar users* section allows to explore a broader view on similar users and the own community of the user.

For recommendations, similarity is measured based on previous face-to-face contacts, establishing links between users, co-authorship relations,[20] and expressed interest in talks as well as actually attended talks within a conference. For calculating recommendations, each of these interaction networks is stored as a weighted (directed or undirected) graph, giving rise to various established similarity metrics which can be used for obtaining personalized and context-specific recommendations. Currently, a personalized PageRank (Brin and Page 1998) algorithm for graph-like representations and the cosine similarity measure for calculating similarities between rows in the adjacency matrix are applied.

In addition, recommendations based on link-prediction measures are also included into the *Acquaint-O-Matic*. Link prediction aims to predict *new* or *recurring* links between participants, i.e. links that are newly established or repetitive links, respectively. Specifically, we consider the face-to-face contact network and additional information, e.g. roles or academic status information about the participants (cf. Scholz *et al*. 2012).

For a more interactive approach, Conferator provides the "My Community" page. It applies automatic community detection methods, e.g. *InfoMap*: Rosvall and Bergstrom (2008) and Rosvall *et al*. (2009); Label Propagation: Raghavan *et al*. (2007); Leading Eigenvector, Newman (2006); and Walktrap: Pons and Latapy (2005). The applied community detection algorithm detects communities based on the face-to-face contact network of the conference participants, e.g. weighted according to the total contact length of participants or by the frequency of their contacts. During the conference, the algorithm is run in periodic intervals to analyze the respective data captured so far.

On the "My Community" page, the user may see members of the community which he/she belongs to. The user may then explore profile pages of the conference participants of the respective community, as inspiration for future conversations and new contacts during the conference.

3.3. WideNoise: collective observation of environmental noise

The WideNoise web application aggregates, summarizes, and illustrates noise-related data collected by the WideNoise smartphone application[21] (cf (Atzmueller *et al*. 2012a). This smartphone application is recording environmental noise levels and allows the user to express certain perceptions about the recorded samples via perception sliders, e.g. love/hate. To further characterize the samples, it is possible to attach custom tags. In order to share samples with friends or the general public, the smartphone application also supports posting results on social media.

The WideNoise web application provides several data summarization views including the *map view* and several statistics pages. The map is shown in Figure 6. It displays, for example, a clustered view on global and user-specific measurements (providing the corresponding detail information on demand, Shneiderman 1996) or a tag cloud characterizing the summarized data by its semantic context.

The user can access several live statistics about the collected data allowing to trace the current measurement trends or to get an overview of the collected data. Some of these statistics are:

- Latest recordings and recent average values for different time intervals and locations.
- User rankings including users with most samples, the most active users, etc.
- Tag clouds characterizing the semantic context of the measurements.

A second type of statistics can be accessed by users via their personal page, e.g. information on their own measuring behavior. The page also provides a Keyhole Markup Language (KML)[22] export containing all the users measurements—as an alternative to the *map* visualization.

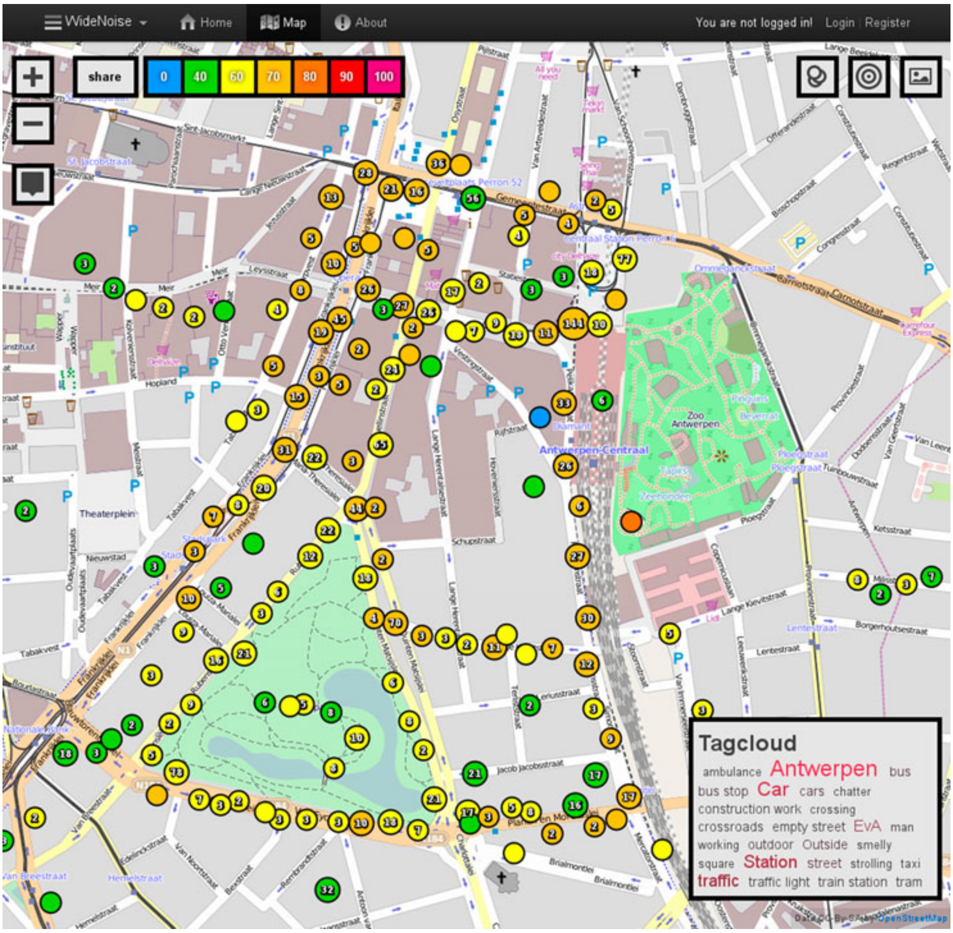

Figure 6. A screenshot of the *map page* of WideNoise.

Interesting analysis options include subgroup analytics on tags, perceptions, and measurements (e.g. Atzmueller and Mueller 2013) or tag-based recommendations (Mueller *et al.* 2013).

3.4. AirProbe: collective measurement of air quality

AirProbe is a system for collaborative air quality monitoring. It consists of three components: a low-cost sensor box for measuring air quality (Elen *et al.* 2012), a smartphone application to communicate with the sensor box, and the web application implemented as part of the Ubicon framework for receiving, aggregating, and visualizing the data.

3.4.1. Overview—functionality. The sensor box produces readings from several air quality-related sensors such as NO_2, CO, O_3, VOC, temperature, or humidity.

Users may add tags concerning their measurements adding semantics or subjective information to the otherwise objective measurements. The web application visualizes the collected data on a map which allows for an easy access to the data as well as for obtaining first insights. The map provides a quantitative view on the data by aggregating the samples using clusters as well as a grid view in order to emphasize the covered area (see Figure 7). The map view also supports the active tracking of currently measured data; tracking these current measurements is further supported by providing data compliant with Google Earth for 3D visualizations. In addition to the map view, the AirProbe web application also provides several statistics like the latest overall measurement activity or air quality averages.

Initial case studies have already been conducted in Antwerp and Torino for sensor calibration. During these case studies, it already became apparent that

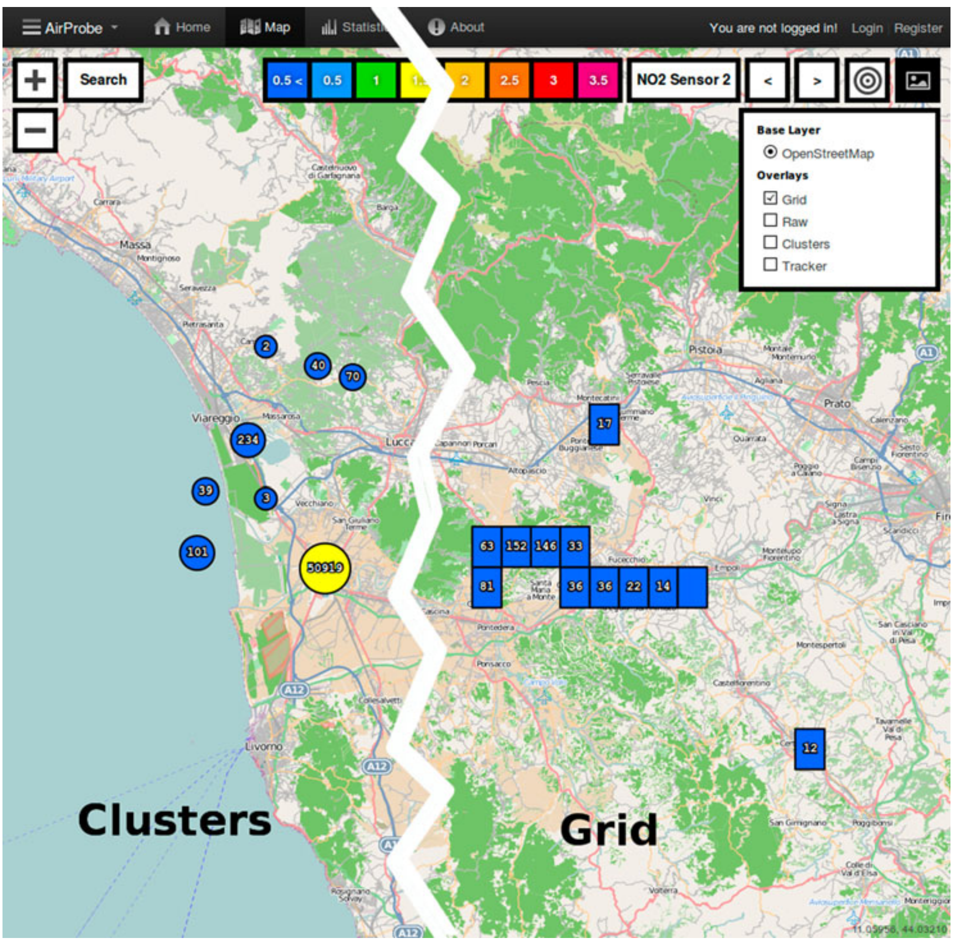

Figure 7. A screenshot of the *map page* of AirProbe. The left side shows the cluster view, the right side shows the grid view.

handling large numbers of measurements will pose many challenges for the web application regarding data processing and visualization. Therefore, mechanisms for speeding up the visualization for the cluster and grid aggregations on the map view have been introduced, as described below.

3.4.2. Clustering large data for fast browsing of annotated maps. One of the main features of WideNoise and AirProbe is the visualization of measurements on a map. In order to motivate the users to explore the collected data and share more of their own, it is essential to provide the best possible user experience when browsing on the map.

Aggregation into clusters and grid cells is one way to allow a quick overview on large amounts of data at the first glance. The clustering is computed on the server and only the aggregated cluster information is transmitted to the client in order to keep things as responsive as possible for the user. However, the aggregation has to be computed for each request (for different viewports, zoom levels, etc.) separately. This is time-consuming, especially as the amount of data increases with continuously recorded measurements. As of February 2013, for instance, the AirProbe database contains about 78,231 measurements amounting to 1,067,916 air quality samples—each measurement can contain several air quality samples.

In order to provide cluster or grid data as fast as possible, our solution is to aggregate data for possible requests beforehand. The aggregated data are stored as a collection of spatial objects. These collections are queried instead of calculating aggregations for each request on the fly. This spatial caching mechanism is outlined in Figure 8. For an incoming request, first a spatial object collection is selected based on certain meta-attributes. Then the spatial objects within the specified longitude–latitude bounds are retrieved.

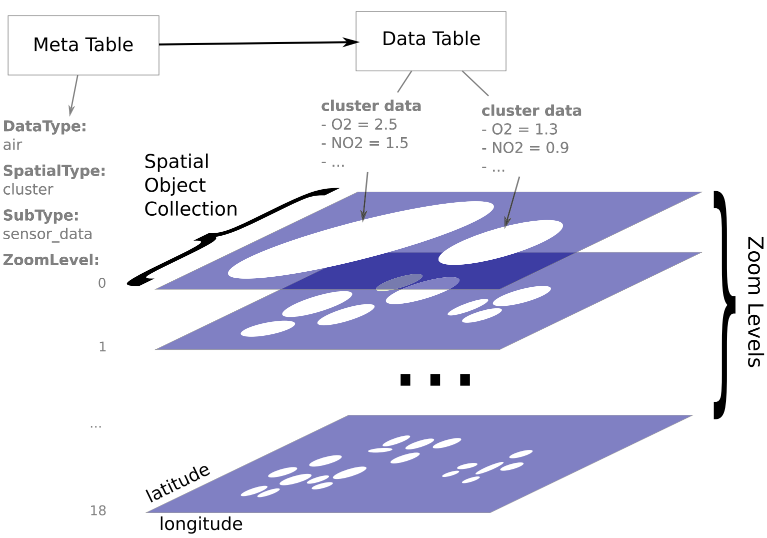

Figure 8. Outline of the spatial object cache concept.

Spatial objects are stored in the *data table*. Each object is described by a collection id, an object id, a longitude–latitude pair, and some object-specific information such as average sensor values or associated tags. The meta-attributes of a spatial object collection are stored in the *meta-table* and are used to retrieve the stored collections. These attributes are as follows:

- *data type*: distinguishes collections of different nature, e.g. air measurements or noise measurements.
- *spatial type*: distinguishes collections by their spatial properties, e.g. clusters, grids, or tracks.
- *sub-type*: distinguishes sub-collections representing the same samples but different corresponding data. This allows to dynamically load additional information (e.g. tags) about spatial objects after transmitting the initial data (e.g. average values) given an object id.
- *zoom level*: spatial objects are aggregated values. Those values are aggregated differently for each zoom level defined by the map.

Using this spatial object cache, we are now running cronjobs on our data in order to update the spatial collections regularly. The rendering time of large clusters was reduced from several minutes to a fraction of a second. This enabled a smooth visualization of large amounts of data on the map. A downside of this approach is that the live characteristics of the aggregated data are reduced. Cronjobs are currently only run every two hours. One run to aggregate grid as well as cluster collections for 18 zoom levels currently takes about 20 minutes in total.

4. Analyses of communication behavior and noise perception

In this section, we present results exemplifying the analysis options provided by the distinct applications discussed in Section 3; first, we focus on the analysis of conference dynamics in the context of the Conferator system. After that, we discuss the utilization of the MyGroup application for a long-term data collection concerning the interactions at a working group. Next, we describe a first analysis of the WideNoise data, focusing on relations between tags and perceived perceptions of the people measuring environmental noise.

4.1. Conference dynamics

Using the conference contact and social data collected by Conferator, we aim to analyze and explore the structure and communication dynamics during conference events. We start with a first analysis of community dynamics, before we discuss communities and associated roles.

4.1.1. Community dynamics. We analyzed the community dynamics across conferences for *LWA 2010* (77 Conferator participants), *LWA 2011* (42 Conferator participants), and *LWA 2012* (42 Conferator participants) and summarized the results in Kibanov *et al*. (2013). For a comparison across conferences, we checked if the participants of different conferences stay in the same communities or if they "switch". For that, we automatically detect communities using the *InfoMap* algorithm (Rosvall and Bergstrom 2008, Rosvall *et al*. 2009) on the contact graph of the participants.

In the following, we associate sets of participants with a community subgroup of the contact graph. We consider the set of participants corresponding to the community with the most participants who visit another conference and compare this with a null hypothesis—that the participants of the community at the first conference would be drawn randomly from all the participants of the second conference.

For a pair of conferences, let G_1 and G_2 denote the respective contact graphs, and N_1 and N_2 the respective sets of participants. For comparing the stability of the communities contained in a pair of conferences, we perform the following steps:

(1) $N = \mathrm{N}_1 \bigcap \mathrm{N}_2$, i.e. the participants who visited both conferences.
(2) Detect sets of communities: C_1 in G_1, and C_2 in G_2
(3) Select the set of participants $N_{1,\max}$ of the community $C_{1,\max} \in C_1$ for which the overlap of the contained set of participants is maximized with respect to N, i.e. the community that contains the largest number of participants in N.
(4) M = set of participants of $\mathrm{N}_{1,\max} \bigcap \mathrm{N}$.
(5) Select the set of participants $N_{2,\max}$ of the community $C_{2,\max} \in C_2$ for which the overlap of the contained set of participants is maximized with respect to M, i.e. the community that contains the largest number of participants in M.
(6) $I := M \bigcap N_{2,\max}$.
(7) $NH = \frac{|M| \cdot |xN_{2,\max}|}{N_2}$ (null hypothesis)

If I (the intersection of both selected communities) is not significantly larger than NH (null hypothesis), then we observe no stability in communities across the conferences.

Table 1 summarizes the results comparing the LWA 2010, 2011, and 2012 conferences. As shown in the table, we can observe the trend that community structure stays relatively stable across conferences, since the number of particular members who stay in the same community is almost every time twice as large as the null hypothesis. For more details, we refer to Kibanov *et al*. (2013).

Table 1. Community dynamics comparing the stability of communities for a pair of conferences against a null hypothesis, that the participants of the community at the first conference would be drawn randomly from all the participants of the second conference.

Conference 1	Conference 2	*I/M*	Null hypothesis
LWA 2010	LWA 2011	9 of 13 participants	4.64
LWA 2010	LWA 2012	6 of 13 participants	3.40
LWA 2011	LWA 2010	9 of 10 participants	4.41
LWA 2011	LWA 2012	2 of 4 participants	1.05

Table 2. Role patterns for a minimal conversation length of 60 seconds.

#	Target	Lift	Share	Size	Pattern
1	Ambassador	1.42	0.63	8	Session chair = true
2	Ambassador	1.14	0.50	12	Affiliation = strong
3	Bridge	2.54	1.0	6	Country = Netherlands AND Presenter = No
4	Bridge	2.18	0.86	7	Country = Netherlands
5	Bridge	0.95	0.37	8	Session chair = true

Note: Min. Contact Length: 60 sec.

4.1.2. Communities and roles. For analyzing communities and roles, we performed an in-depth analysis of the Hypertext 2011 conference with several interesting observations and findings (cf. Macek *et al*. 2012). There is a clear connection, for example, between a community-oriented role of a participant, e.g, an *ambassador* as an important person connecting different communities or a *bridge* as a not-so-visible person bridging different communities, and the academic status of the participant (Professor, PostDoc, PhD Student, and others).

Specifically, we aimed at discovering subgroups using the VIKAMINE system (Atzmueller and Lemmerich 2012), i.e. sets of participants with a specific role as closely as possible using a set of descriptive features, e.g. their country of origin, title, role as session chair, invited speaker, or presenter of a conference paper. Subgroup discovery (cf. Klösgen 1996, Atzmueller and Lemmerich 2009, Lemmerich *et al*. 2012) aims at identifying interesting patterns with respect to a given target property according to a specific interestingness measure. In our context, the target properties of interest are given by the different roles of participants in the contact graph.

Concerning a minimal conversation length of 60 seconds (Table 2), it is easy to see that most of the session chairs serve as ambassadors during the conference (the remaining session chairs are bridges). Furthermore, a *strong* affiliation to Hypertext plays an important role for being an ambassador in the conference. The feature *affiliation* denotes the familiarity with Hypertext, such that authors of at most one Hypertext paper published in 2011 get a *low* affiliation score, authors who published one or two papers before Hypertext 2011 get a *medium* affiliation score, and authors with at least 3 papers before Hypertext 2011 get a *strong* affiliation score. It is also evident that the participants from the Netherlands (including in particular the organizers) are typical bridges, as expected, e.g. subgroup #3 of Table 2 with a target share of 100%.

Furthermore, we analyzed the correlation between different community structures (track, country, conference affiliation, and academic status) and the contact length of participants within these communities. Figure 9 indicates the trend that being in the same track, for example, improves the community quality indicator (p-value, Scripps *et al*. 2007) with increasing minimal conversation lengths. This confirms our intuition that communities tend to be structured according to those properties, e.g. a common track. For more details, we refer to Macek *et al*. (2012).

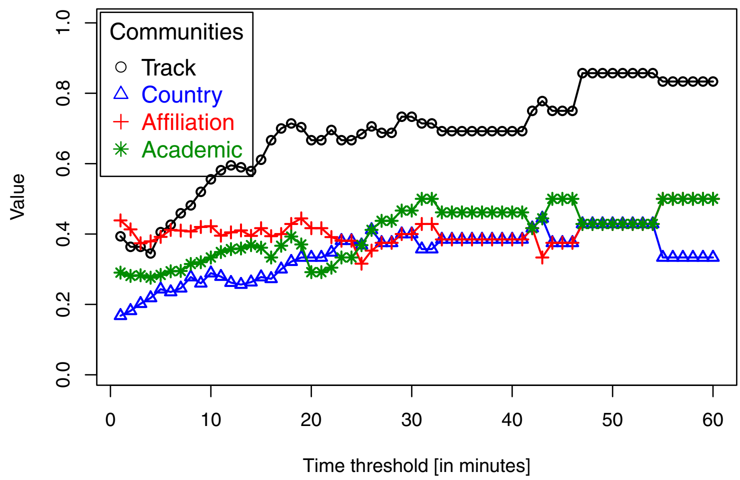

Figure 9. Time-based analysis showing a community quality indicator, i.e. the *p*-value (Scripps *et al.* 2007) for the partitionings track, country, affiliation, and academic status using different minimal conversation lengths. The higher the *p*-value, the higher the probability for a contact within a community. See Macek *et al.* (2012) for more details.

4.2. Communication patterns in working groups

In order to test, apply, and improve our own systems, all members of our group (KDE) wear RFID tags during their daily work, as explained in Section 3.1. For a first impression of the typical contacts and conversations during a workday, Figure 10 displays the cumulated probability distribution of the contact length of the data we collected between October 2010 and March 2011: each line in the graph denotes one of the time slots. We divided a working day into six two-hour slots from 8:00 to 20:00. In the graph, the *x*-axis shows the duration of face-to-face contacts (conversations) in minutes, while the probability that this duration is exceeded (for the particular time slot) is shown in the *y*-axis. Both axes are scaled logarithmically.

It is easy to see that, for example, longer conversations are more likely during the evening hours than during the morning hours. Long discussions (i.e. more than 20 minutes) are not held in the early morning at all. Furthermore, short discussions (i.e. less than a minute) are very likely during the whole day.

4.3. Noise perception in crowd-sourced sensing

The WideNoise application and its corresponding back-end are continuously running since December 2011.

In the following, we present analysis and statistics with respect to different locations and case studies in comparison to the overall worldwide statistics (33,168 measurements). One location to focus on is the Heathrow Airport area near London (6,055 measurements) where a long-running campaign against noise caused by landing of planes and starting from the Heathrow Airport took place. Furthermore, two case studies are examined: one was an all-day event taking place in Rome on 9 June 2012 (830 measurements) and the other took place in

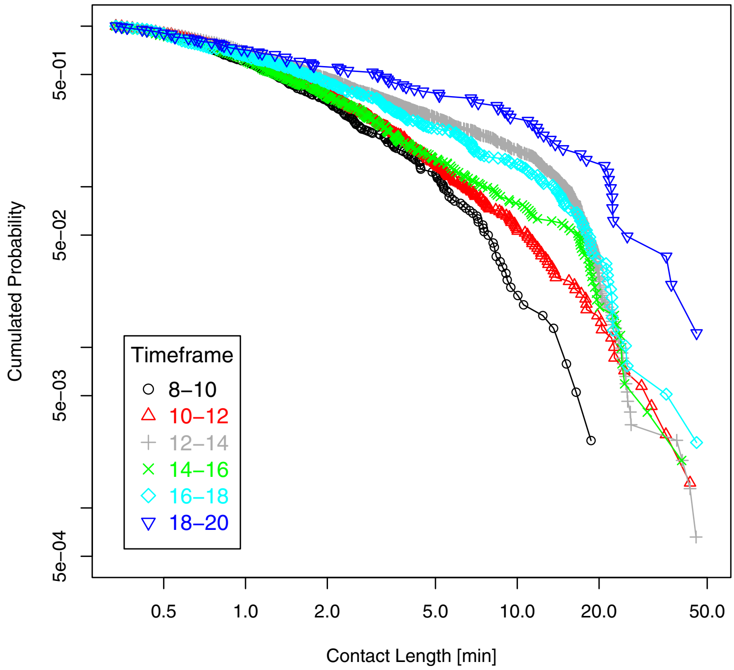

Figure 10. Cumulated probability distribution of face-to-face contact lengths in the KDE group, for different timeframes.

Antwerp on 10 July 2012 from 9:00 AM to 1:30 PM (1,943 measurements). The areas of the corresponding events were roughly estimated using bounding boxes.

As shown in Table 3, we observe that the case studies in Rome and Antwerp have a higher noise level than the average worldwide. This is also and especially true for Heathrow. While we do not list any specific case studies in London, the corresponding measurements include samples which were taken during a Heathrow campaign, for which the users are motivated to use the system for measuring higher noise levels in order to prove their point regarding noise pollution caused by aeroplanes leaving and approaching the Heathrow Airport. This explains the exceptional average noise-level values for Heathrow.

4.3.1. Perceptions. When taking measurements with the WideNoise application, users can record their perceptions of the actual situation using four sliders. Looking at these perceptions for the experiments, Table 4 yields additional insights. Overall, more "man-made" noise was recorded than natural noise and people tend to associate noise with "hate" rather than "love". In the Heathrow area, mostly aeroplane noise is recorded which is usually associated with the "man-made" aspect resulting in the high average value. On the other hand, people in the Heathrow area are mostly sampling alone whereas during the Antwerp and the Rome case study participants were measuring in groups. This explains the strength of the "social" aspect compared to Heathrow.

Table 3. Count of noise samples and average noise level (dB) for different locations and case studies overall and for a chosen set of tags. It is important to note that the noise levels are uncalibrated values as measured by the different types of mobile phones.

	Worldwide		Heathrow area		Antwerp		Rome	
	Count	dB	Count	dB	Count	dB	Count	dB
Tagged samples	4,431	70	1,809	69	721	74	110	72
All samples	33,168	64	6,055	73	1.943	67	830	68
	Worldwide		Heathrow area		Antwerp		Rome	
Tag	Count	dB	Count	dB	Count	dB	Count	dB
aeroplane noise	522	63	522	63	–	–	–	–
Antwerpen	249	81	–	–	218	81	–	–
car	221	74	8	77	148	81	21	63
Esterno	548	76	–	–	–	–	–	–
Heathrow	546	76	546	76	–	–	–	–
indoor	96	71	–	–	2	70	21	50
office	109	69	–	–	–	–	16	63
plane	150	83	127	83	–	–	–	–
station	138	85	1	85	115	76	–	–
traffic	140	83	9	78	100	81	10	81

4.3.2. Perception and tag associations. In addition to recording their perceptions, WideNoise allows its users to add words (the so-called tags) to their recordings for specifying the context of the ongoing measurement more precisely. Table 3 shows the 10 tags that were assigned most frequently. Most of the users did not utilize tags (only about 13.3% of the measurements are tagged in Rome but more than 37.1% in Antwerp). In almost all cases, the tagged recordings tend to have higher noise levels. This observation also holds for the worldwide data: only 13.4% of all measurements are tagged and those measurements are on average 8.5% louder than the mean.

In order to analyze the tag–perception relations in more detail, we applied subgroup analytics, as outlined above, using the VIKAMINE system (Atzmueller and Lemmerich 2012). We aimed at identifying subgroups of measurements tagged with certain tags that are as unusual as possible compared to the mean of a certain parameter, e.g. given by one of the perceptions *love/ate*, *calm/hectic*, *alone/social*, and *nature/man-made*.

Table 4. Averages of the user-assigned perceptions for locations and case-studies (excluding measurements with only default perception values of 0.5).

Location/case study	Love/hate	Calm/hectic	Alone/social	Nature/man-made
Worldwide	0.60	0.57	0.50	0.78
Heathrow area	0.76	0.68	0.37	0.94
Antwerp	0.61	0.72	0.82	0.86
Rome	0.52	0.49	0.59	0.68

In summary, the results support our findings shown in Table 3 for the decibel (dB) values: *heathrow*- and airplane-related tags are among the top relations, while the subgroup containing the combination of the tags *indoor* and *small plane* yields the highest increase for a more complex pattern. The same can be observed for the perceptions *hate*, *hectic*, *social*, and *man-made* where airplane-related tags from the London Heathrow area dominate the overall perceptions in the whole population. Examples of tags associated with *love* include, for example, *silence*, *quit*, *music*, or *sleep*. On the other hand, tags associated with *hate* include, e.g. *aeroplane* or *rooftop*. For the perceptions *calm* and *hectic*, we observe similar relations. For the perception *social*, strong associations are given by the tags *airport*, *commute*, *office*, whereas *alone* is described, for example, by the tags *home*, *quiet*, or *sleep*. Finally, the perception *man-made* is exemplified by the tags *piazza*, *chatter*, or *telco*. Thus, we observe that the perceptions and associated tag assignments correspond to intuition rather well.

5. Related work

From a software architecture point of view, there are various frameworks and toolkits for supporting ubiquitous and/or context-aware applications. The Context Toolkit (Salber *et al*. 1999, Dey *et al*. 2001), for example, provides a conceptual framework for the rapid development of context-aware applications. Similarly Bannach *et al*. (2008, 2010) and Kunze and Bannach (2012) present the context-recognition network toolkit/toolchain for building context-aware pervasive applications. Compared to these toolkits addressing mainly context-aware applications, Ubicon has a different focus: on the one hand, its application focus is different. It aims at supporting applications that consider *both* ubiquitous and social aspects. In addition, Ubicon is no general toolkit for rapid prototyping, but aims at providing a general framework support for implementing and hosting ubiquitous and social applications in high-availability online scenarios. This is achieved by providing a layered template architecture with an efficient and effective data storage and processing chain. Then, applications implement this template using the modules provided by the Ubicon core components. In addition, applications can also make use of the same platform components, such that they are hosted on the same server for potentially sharing data and providing an integrated user experience across applications.

Concerning the Ubicon applications, several systems for observing social behavior, for example, at conferences have been built using RFID tokens or Bluetooth-enabled devices; Hui *et al*. (2005) describe an application using Bluetooth-based modules for collecting mobility patterns of conference participants. Concerning social interactions, Eagle and Pentland (2006) present an approach for collecting proximity and location information using Bluetooth-enabled mobile phones. One of the first experiments using RFID tags to track the position of persons on room level was conducted by Meriac *et al*. (2007) in the Jewish Museum, Berlin, in 2007. Cattuto *et al*. (2010) added proximity sensing in the SocioPatterns project. For MyGroup and Conferator, we are using the SocioPatterns hardware as a technological basis. In addition, we increased the precision of the localization component and linked the RFID tag information with further information, e.g. about the working group members or the schedule of a workshop week. This provides for new insights into the behavior of all participants (cf. Macek *et al*. 2012). Similarly Chin *et al*. (2013) describe the Find and Connect system and analysis concerning the relations between physical and social interactions at conferences.

In the context of MyGroup, there have been several approaches for improving collaborative group activities: Digiano *et al*. (2006) and Looi *et al*. (2009) propose and examine the GroupScribbles technique for assisting collaborative activities. Lin *et al*. (2009) present the SmallBlue system to operationalize (generated) social networks for expert finding and connecting people. In contrast to these systems, we do not only aim to improve the collaboration between people and to provide helpful information for networking but also we take the dynamic structure of the social interactions into account in order to provide instant recommendations and notifications about people and events.

While MyGroup and Conferator target the area of social interactions themselves, the WideNoise and AirProbe address the area of participatory sensing on a more sensor-based level. Both are implemented in the EveryAware project,[23] which aims at enhancing environmental awareness through social information technologies. There are approaches like the Partisan architecture (Burke *et al*. 2006) aiming to provide a framework of basic building blocks to support participatory sensing on a 'grassroots' level.

Participatory sensing allows to approach many research questions. One such area is understanding patterns, semantics, and dynamics of social behavior and its interaction with the sensor data collected by corresponding applications. In this context, Pan *et al*. (2013) defines research areas in the trace analysis realm. In this area, applications like WideNoise and AirProbe aim at contributing strongly in the future.

Kanjo (2010) presents the first system for collecting noise data with mobile phones, and discusses its implementation on a technical level. There are several existing platforms dedicated to specialized sensor data types. Maisonneuve *et al*. (2010) present an approach for monitoring the noise pollution by the general public using the NoiseTube[24] system. Aircasting[25] is another platform which allows users to upload information about surrounding noise using their mobile phones. Aircasting also supports air quality measurements. There is other research about participatory air quality sensing as in Hasenfratz *et al*. (2012). The AirProbe application and corresponding research go into the same direction but provide a more advanced sensor box and put additional focus on behavioral change as well as on the combination of objective and subjective data like, for example, perceptions or tags. This also distinguishes the approaches taken by EveryAware from other open-sensing platforms like Eye on Earth, a "global public information network" for creating and sharing environmentally relevant data and information online through interactive map-based visualizations.[26] Further platforms in this context include COSM (Pachube) (cf. Atzori *et al*. 2011) or the OpenSensors platform for earth observation using a sensor web (see Andrae and Simonis 2011).

6. Conclusions

In this paper, we described the overall architecture of the Ubicon platform for ubiquitous and social applications. We showed that very different applications can be implemented using the platform, and described several application-specific components for data analysis within the respective applications. Furthermore, we provided analysis results using the collected data in different application contexts. The results also show that the collected data are relevant for research purposes not only for computer science but also for psychologists, sociologists, ecologists, etc.

For future work, we aim to integrate generic recommendation options into the platform for an easier setup of such solutions. In addition, social connectors to further social services would enable a more comprehensive coverage of the context of the users. Furthermore, the integration and utilization of linked open data is another interesting direction for embedding more context and semantics into the collected data. Finally, we aim to further evaluate the presented components, e.g. tag recommendations (Mueller *et al*. 2013), as well as coverage and data quality of the WideNoise and AirProbe applications.

Acknowledgments

This work was supported by the VENUS research cluster at the interdisciplinary Research Center for Information System Design (ITeG) at Kassel University and by the EU project EveryAware. MyGroup and Conferator apply active RFID technology which was developed within the SocioPatterns project, whose generous support we kindly acknowledge. We also thank Milosch Meriac from Bitmanufaktur in Berlin for helpful discussions regarding the RFID localization. Our special thanks go to Ciro Cattuto, who enabled access to the SocioPatterns technology, and who supported us with valuable information concerning the setup of the RFID technology.

Notes

[1] http://ubicon.eu/
[2] http://ubicon.eu/about/conferator/
[3] http://ubicon.eu/about/mygroup/
[4] http://cs.everyaware.eu/event/widenoise/
[5] http://cs.everyaware.eu/event/airprobe/
[6] http://springsource.org/
[7] http://tomcat.apache.org/
[8] http://www.kde.cs.uni-kassel.de/
[9] http://www.iteg.uni-kassel.de/venus
[10] http://www.is.informatik.uni-wuerzburg.de/dmir
[11] http://everyaware.eu
[12] http://sociopatterns.org/
[13] http://www.bibsonomy.org
[14] http://wintersause.de/
[15] http://www.kde.cs.uni-kassel.de/conf/lwa10/
[16] http://lwa2011.dke-research.de/
[17] http://lwa2012.cs.tu-dortmund.de
[18] http://ht2011.org/
[19] Privacy is preserved with certain privacy settings; due to the limited space, we refer to Atzmueller *et al*. (2011) for more details.

[20] The network of co-authorship can be retrieved (mainly for computer science) from DBLP (http://dblp.uni-trier.de/)
[21] http://itunes.apple.com/de/app/widenoise/id302052132/ (iOS) and https://play.google.com/store/apps/details?id=eu.everyaware.widenoise.android (Android)
[22] http://opengeospatial.org/standards/kml/
[23] http://www.everyaware.eu
[24] http://noisetube.net/
[25] http://www.aircasting.org
[26] http://www.eyeonearth.eu/en-us/Pages/Learn-More.aspx

References

S. Andrae and I. Simonis, "Opensensors: A community platform to enable the sensor web and foster earth observation research", in *Proceedings of IST-Africa Conference*, Botswana: Gaborone, pp. 1–10, 2011.

M. Atzmueller, D. Benz, A. Hotho and G. Stumme, eds., *Proceedings of LWA*, Technical report (KIS), 2010-10 Department of Electrical Engineering/Computer Science, Kassel University, 2010.

M. Atzmueller, *et al.*, "Enhancing social interactions at conferences", *IT – Information Technology*, 53(3), pp. 101–107, 2011.

M. Atzmueller, M. Becker, S. Doerfel, M. Kibanov, A. Hotho, B.-E. Macek, F. Mitzlaff, J. Mueller, C. Scholz and Gerd Stumme, "Ubicon: Observing social and physical activities", in *Proceedings of 4th IEEE International Conference on Cyber, Physical and Social Computing (CPSCom 2012)*, Washington, DC, USA: IEEE Computer Society, pp. 317–324, 2012a.

M. Atzmueller, S. Doerfel, A. Hotho, F. Mitzlaff and G. Stumme, *Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles. Modeling and Mining Ubiquitous Social Media, Vol. 7472 of LNAI*, Heidelberg, Germany: Springer Verlag, 2012b.

M. Atzmueller and F. Lemmerich, "Fast subgroup discovery for continuous target concepts", in *Proceedings of 18th International Symposium on Methodologies for Intelligent Systems (ISMIS), Vol. 5722 of LNCS*, Heidelberg, Germany: Springer, pp. 1–15, 2009.

M. Atzmueller and F. Lemmerich, "VIKAMINE – Open-source subgroup discovery, pattern mining, and analytics", in *Proceedings of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), Vol. 7524 of LNCS*, Heidelberg, Germany: Springer, pp. 842–845, 2012.

M. Atzmueller and J. Mueller, "Subgroup analytics and interactive assessment on ubiquitous data", in *Proceedings of the International Workshop on Mining Ubiquitous and Social Environments (MUSE2013)*, Prague, Czech Republic, pp. 19–27, 2013.

L. Atzori, A. Iera and G. Morabito, "SIoT: Giving a social structure to the internet of things", *Communications Letters, IEEE*, 15(11), pp. 1193–1195, 2011.

D. Bannach, O. Amft and P. Lukowicz, "Rapid prototyping of activity recognition applications", *IEEE Pervasive Computing*, 7(2), pp. 22–31, 2008.

D. Bannach, K. S. Kunze, J. Weppner and P. Lukowicz, "Integrated tool chain for recording and handling large, multimodal context recognition data sets", in *UbiComp (Adjunct Papers), ACM International Conference Proceeding Series ACM*, J.E. Bardram, M. Langheinrich, K.N. Truong, and P. Nixon (Eds), New York: ACM Press, pp. 357–358, 2010.

K. Behrenbruch, O. Kieselmann, M. Schuldt, M. Söllner and L. Schmidt, "System evaluation", in *Socio-Technological Design of Ubiquitous Computing Systems*, K. David, K. Geihs, J. M. Leimeister, A. Roßnagel, L- Schmidt, G. Stumme and A. Wacker (Eds), Berlin/Heidelberg: Springer Verlag, pp. 295–318, 2014.

D. Benz, A. Hotho, R. Jäschke, B. Krause, F. Mitzlaff, C. Schmitz and G. Stumme, "The social bookmark and publication management system BibSonomy", *VLDB*, 19(6), pp. 849–875, 2010.

S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine", *Computer Networks ISDN*, 30(1–7), pp. 107–117, 1998.

J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing", in *Workshop on World-Sensor-Web (WSW'06): Mobile Device Centric Sensor Networks and Applications*, Boulder, CO: SenSys '06, pp. 117–134, 2006.

C. Cattuto, W. v. den Broeck, A. Barrat, V. Colizza, J.F. Pinton and A. Vespignani, "Dynamics of person-to-person interactions from distributed RFID sensor networks", *PLoS ONE*, 5(7), p. e11596, 2010.

A. Chin, B. Xu, H. Wang, L. Chang, H. Wang and L. Zhu, "Connecting people through physical proximity and physical resources at a conference", *ACM Transactions Intelligent Systems and Technology*, 4(3), 50:1–50:21, 2013.

J. Dean and S. Ghemawat, "Mapreduce: Simplified data processing on large clusters", *Communications of the ACM*, 51(1), pp. 107–113, 2008.

A.K. Dey, G.D. Abowd and D. Salber, "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications", *Human-Computer Interaction*, 16(2), pp. 97–166, 2001.

C. Digiano, D. Tatar and K. Kireyev, "Learning from the post-it: Building collective intelligence through lightweight, flexible technology", in *Proceedings of CSCW ACM*, New York: ACM Press, pp. 65–66, 2006.

N. Eagle and A.S. Pentland, "Reality mining: Sensing complex social systems", *Personal and Ubiquitous Computing*, 10(4), pp. 255–268, 2006.

B. Elen, J. Theunis, S. Ingarra, A. Molino, J. Van den Bossche, M. Reggente and V. Loreto, "The everyaware SensorBox: A tool for community-based air quality monitoring", in *Proceedings of the Sensing a Changing World Workshop*, 2012.

K. Geihs, L. M. Leimeister, A. Roßnagel and L. Schmidt, "On socio-technical enablers for ubiquitous computing applications", in *2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet (SAINT)*, Izmir, Turkey: IEEE, pp. 405–408, 2012.

D. Hasenfratz, O. Saukh, S. Sturzenegger and L. Thiele, "Participatory air pollution monitoring using smartphones", in *Proceedings of 1st International Workshop on Mobile Sensing: From Smartphones and Wearables to Big Data*, New York: ACM Press, 2012.

P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft and C. Diot, "Pocket switched networks and human mobility in conference environments" in *SIGCOMM*, Philadelphia, PA, USA and New York, NY, USA: ACM, pp. 244–251, 2005.

E. Kanjo, "NoiseSPY: A real-time mobile phone platform for urban noise monitoring and mapping", *MONET*, 15 (4), pp. 562–574, 2010.

M. Kibanov, M. Atzmueller, C. Scholz and G. Stumme, "On the evolution of contacts and communities in networks of face-to-face proximity", in *Proceedings of IEEE CPSCom 2013*, Boston, MA, USA: IEEE Computer Society, 2013.

W. Klösgen, "Explora: A multipattern and multistrategy discovery assistant", in *Advances in Knowledge Discovery and Data Mining*, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (Eds), AAAI Press, pp. 249–271, 1996.

K. Kunze and D. Bannach, *Towards Dynamically Configurable Context Recognition Systems*. [online]. 2012.

F. Lemmerich, M. Becker and M. Atzmueller, "Generic pattern trees for exhaustive exceptional model mining", in *Proceedings of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, *Vol. 7524 of LNCS*, Heidelberg, Germany: Springer Verlag, pp. 277–292, 2012.

C.-Y. Lin, N. Cao, S. X. Liu, S. Papadimitriou, J. Sun and X. Yan, "SmallBlue: Social network analysis for expertise search and collective intelligence", in *Proceedings of ICDE*, Boston, MA: IEEE Computer Society, pp. 1483–1486, 2009.

C.K. Looi, W. Chen and F.K. Ng, "Collaborative activities enabled by GroupScribbles (GS): An exploratory study of learning effectiveness", *Computers & Education*, 54(1), pp. 14–26, 2009.

B.E. Macek, M. Atzmueller and G. Stumme, "Profile mining in CVS-logs and face-to-face contacts for recommending software developers", in *Proceedings of Social Computing IEEE*, Boston, MA: IEEE Computer Society, pp. 250–257, 2011.

B.-E. Macek, C. Scholz, M. Atzmueller and G. Stumme, "Anatomy of a conference", in *Proceedings of 23rd ACM Conference on Hypertext and Social Media*, New York, NY, USA: ACM Press, pp. 245–254, 2012.

N. Maisonneuve, M. Stevens and B. Ochab, "Participatory noise pollution monitoring using mobile phones", *Inform Polity*, 15(1), pp. 51–71, 2010.

N. Marz and J. Warren, *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*, Greenwich, CT: Manning Publishers, 2013.

M. Meriac, A. Fiedler, A. Hohendorf, J. Reinhardt, M. Starostik and J. Mohnke, "Localization techniques for a mobile museum information system", in *Proceedings of WCI* (Wireless Communication and Information) Berlin, October 2007, 2007.

J. Mueller, S. Doerfel, M. Becker, A. Hotho and G. Stumme, "Tag recommendations for SensorFolkSonomies", in *Proceedings Workshop on Recommender Systems and the Social Web*, New York, NY, USA, 2013.

M.E.J. Newman, "Finding community structure in networks using the eigenvectors of matrices", *Physical Review E*, 74(3) Cite arxiv:physics/0605087Comment: 22 pages, 8 figures, minor corrections in this version, 2006.

G. Pan, G. Qi, W. Zhang, S. Li, Z. Wu, and L.T. Yang, "Trace analysis and mining for smart cities: Issues, methods, and applications", *Communications Magazine, IEEE*, 51(6), pp. 120–126, 2013.

M. Parameswaran and A.B. Whinston, "Social computing: An overview", *Communications of the Association for Information Systems*, 19, 2007.

P. Pons and M. Latapy, "Computing communities in large networks using random walks", *J. Graph Algorithms Appl.*, 10(2), pp. 191–218, 2006.

U.N. Raghavan, R. Albert and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks", *Physical Review E*, 76(3), pp. 036106, 2007.

M. Rosvall, D. Axelsson and C.T. Bergstrom, "The map equation", *The European Physical Journal Special Topics*, 178, pp. 13–23, 2009.

M. Rosvall and C.T. Bergstrom, "Maps of random walks on complex networks reveal community structure", *Proceedings of the National Academy of Sciences*, 105(4), pp. 1118–1123, 2008.

D. Salber, A.K. Dey and G.D. Abowd, "The context toolkit: Aiding the development of context-enabled applications", in *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, *CHI '99*, Pittsburgh, Pennsylvania, USA and New York, NY, USA: ACM, pp. 434–441, 1999.

C. Scholz, M. Atzmueller and G. Stumme, "On the predictability of human contacts: Influence factors and the strength of stronger ties", in *Proceedings of Fourth ASE/IEEE International Conference on Social Computing (SocialCom)*, Boston, MA, USA: IEEE Press, 2012.

C. Scholz, S. Doerfel, M. Atzmueller, A. Hotho and G. Stumme, "Resource-aware on-line RFID localization using proximity data", in *Proceedings of ECML/PKDD 2011*, Heidelberg, Germany: Springer Verlag, pp. 129–144, 2011.

J. Scripps, P.N. Tan and A.H. Esfahanian, "Exploration of link structure and community-based node roles in network analysis", in *Proceedings of 7th IEEE International Conference on Data Mining (ICDM)*, Washington, DC, USA: IEEE Press, pp. 649–654, 2007.

B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations", in *Proceedings of IEEE Symposium on Visual Languages*, Boulder, Colorado, Boston, MA, USA: IEEE Press, pp. 336–343, 1996.

M. Szomszor, C. Cattuto, W. Van den Broeck, A. Barrat and H. Alani, "Semantics, sensors and the social web: The live social semantics experiments", in *Proceedings of ESWC 2010*, Heidelberg, Germany: Springer Verlag, pp. 196–210, 2010.

F.-Y. Wang, K. Carley, D. Zeng and W. Mao, "Social computing: From social informatics to social intelligence", *Intelligent Systems*, 22(2), pp. 79–83, 2007.

M. Weiser, "The computer for the 21st century", *Scientific American*, 265(3), pp. 66–75, 1991.

M. Weiser, "Hot topics – Ubiquitous computing", *Computer*, 26(10), pp. 71–72, 1993.