Ubicon: Observing Physical and Social Activities

Martin Atzmueller*, Martin Becker†, Stephan Doerfel*, Andreas Hotho†, Mark Kibanov*,
Bjoern-Elmar Macek*, Folke Mitzlaff*, Juergen Mueller*, Christoph Scholz* and Gerd Stumme*

*Knowledge and Data Engineering Group, University of Kassel, Germany
Email: {atzmueller, doerfel, kibanov, macek, mitzlaff, mueller, scholz, stumme}@cs.uni-kassel.de

†Data Mining and Information Retrieval Group, University of Würzburg, Germany
Email: {becker, hotho}@informatik.uni-wuerzburg.de

***Abstract*—The connection of ubiquitous and social computing is an emerging research area which is combining two prominent areas of computer science. In this paper, we tackle this topic from different angles: We describe data mining methods for ubiquitous and social data, specifically focusing on physical and social activities, and provide exemplary analysis results. Furthermore, we give an overview on the UBICON platform which provides a framework for the creation and hosting of ubiquitous and social applications for diverse tasks and projects. UBICON features the collection and analysis of both physical and social activities of users for enabling inter-connected applications in ubiquitous and social contexts. We summarize three real-world systems built on top of UBICON, and exemplarily discuss the according mining and analysis aspects.**

I. INTRODUCTION

Connecting the social and the physical world is one of the challenges in ubiquitous and social computational systems: Such applications usually involve the users' contexts in their physical and social manifestations, that is, in the offline and online world. In the following, we discuss data mining and analysis aspects in ubiquitous and social contexts. Additionally, we outline the UBICON[1] platform for enabling ubiquitous and social networking. Furthermore, we describe exemplary data mining techniques and components that were evaluated using the systems, from localization to user recommendation methods. For sketching the capabilities and potential of observing physical and social activities, we also present exemplary analysis results using real-world data collected by the three example applications: CONFERATOR[2], MYGROUP[3] and WIDENOISE[4].

The structure of this paper is as follows. Section II describes the applications currently implemented on the UBICON platform. Section III illustrates the data mining technics which are implemented in UBICON. Section IV provides examples of analysis that can be conducted using data gathered by UBICON applications. Section V discusses related work, and Section VI concludes the paper with a summary and promising options for future work.

[1]http://ubicon.eu/
[2]http://conferator.org
[3]http://ubicon.eu/about/mygroup
[4]http://cs.everyaware.eu/event/widenoise

II. UBICON: SUPPORTING UBIQUITOUS SOCIAL NETWORKING

In the following, we describe the applications MYGROUP, CONFERATOR and WIDENOISE which are implemented using the UBICON platform. Using one platform for different ubiquitous applications has several advantages: It is not necessary, for example, to implement the same components that are typical for ubiquitous and social systems again and again. Furthermore, the data from different systems can be combined for enhancing the functionality of the system.

MYGROUP and CONFERATOR are maintained and developed by the University of Kassel, at the Knowledge and Data Engineering group in the inter-disciplinary context of the VENUS research cluster,[5] which is concerned with the design of social, legal and technological networking issues in situated ubiquitous systems. The WIDENOISE web application is jointly developed by the University of Wuerzburg and the University of Kassel in the context of the EU project *EveryAware*.[6]

From a technical point of view, the UBICON platform consists of the application logic, components for privacy management and database management, a (customizable) set of data processors that process the incoming (raw) data, a set of data processors for more subsequent sophisticated processing, and a storage architecture based on a MySQL database[7]. The set of data processors include, e. g., the localization component for determining the location of RFID tags. The system is implemented with a model-view-controller pattern using the Spring framework[8]. UBICON can be deployed using a standard servlet container, e. g., Apache Tomcat.[9] Figure 1 shows a conceptual overview of the system's architecture.

A. MYGROUP*: Social Networking in Working Groups*

MYGROUP aims at supporting members of working groups. It employs active RFID tags for localizing the members and for monitoring their social contacts. Additionally it provides profile information including links to (external) social software, e. g., BibSonomy [6], Twitter, Facebook, or XING.

[5]http://www.iteg.uni-kassel.de/venus
[6]http://everyaware.eu
[7]http://mysql.com/
[8]http://springsource.org/
[9]http://tomcat.apache.org/

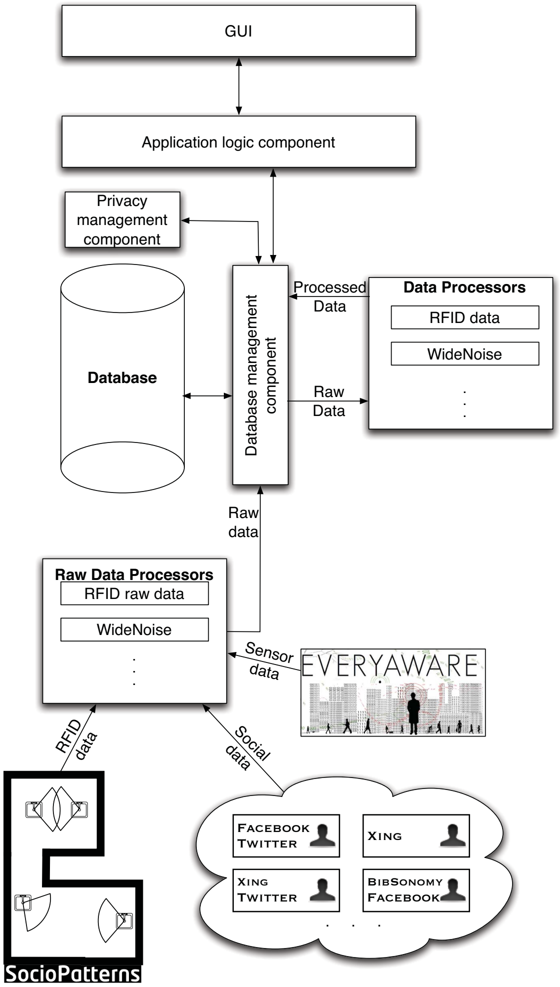

Figure 1. The overview of the architecture of UBICON software platform.

MYGROUP applies a new generation of cost-effective RFID devices. These so-called *proximity tags* make use of active RFID technology and have been developed by the SocioPatterns project[10] and the company Bitmanufaktur.[11] The setup requires a number of RFID readers at fixed positions in the target area; the participants are then equipped with RFID tags. The technology allows for the localization of tags and for detecting tag-to-tag proximity. When the tags are worn on the chest, tag-to-tag proximity is a proxy for face-to-face communication, since the range of the signals is approximately 1.5 meters if not blocked by the human body. In this way, one can detect and analyze individual face-to-face contacts. For more details, we refer to Barrat et al. [5].

MYGROUP has been applied at a number of different events: It is being used by the Knowledge and Data Engineering Group (KDE) of the University of Kassel research group, and it currently is also being extended towards a larger research cluster. In addition, MYGROUP has also been utilized at a large student party,[12] for supporting organizational processes, at the First International Changemaker-Camp at the University of Kassel for profiling group processes, and within the VENUS project at a CodeCamp for supporting software development.

[10] http://sociopatterns.org/
[11] http://bitmanufaktur.de/

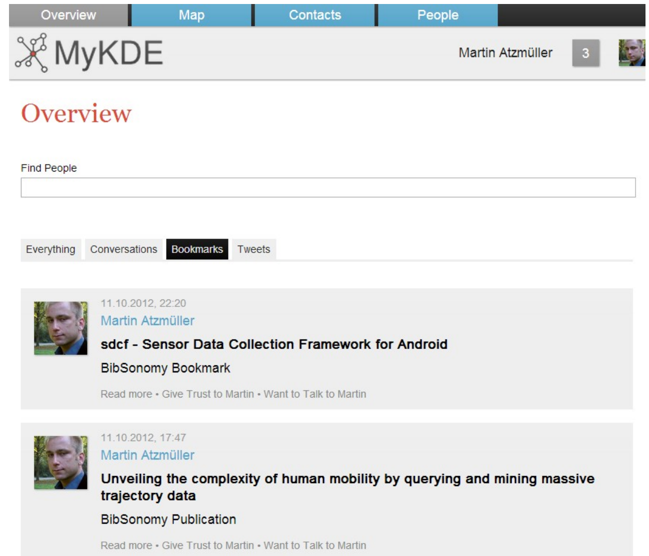

Figure 2. A screenshot of the *timeline view*: The screenshot shows two recent BibSonomy posts – enabled by different filters.

MYGROUP provides several functions for improving interactions and discussions in working groups: 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, which will be introduced in the next section below, see Figure 4) provide detailed information, for example, about position and interests of a group member.

Utilizing the system, we can exploit social structures, social contacts, and social knowledge provided by both social networks and social resource sharing systems for supporting complex and structured interactions: We can recommend persons, for example, based on current joint research topics. We apply data mining methods on the collected data to make this information visible to our users as described in Section III. Different trust and privacy settings, e. g., concerning the visibility of contacts and locations, allow a selective distribution of sensitive information. The system is continuously refined according to user feedback and usability studies, in particular in the VENUS project, leading to continuous improvement of the system and implementation of new useful features.

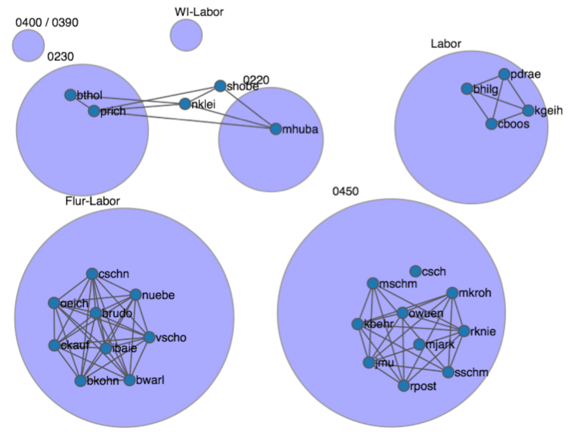

Figure 3. A screenshot of the *map view* of MyGroup. The large circles denote individual rooms, the smaller circles participants; connections between those indicate ongoing conversations.

Figure 4. A screenshot of an exemplary Conferator user profile, with general information, context information, social tag cloud and latest posts.

B. Conferator*: A Social Conference Guidance System*

Conferator [1] 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 using these features.

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., concerning 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 4. Similar to MyGroup, Conferator offers several privacy settings in order to enable privacy protection, e. g., for sharing locations or contact information.

Conferator has been successfully applied at LWA 2010[13] [2], [3] 2011[14] and 2012[15] – conferences for special interest groups of the German Computer Science Society (GI), at the Hypertext 2011[16] conference [16], and at a technology day of the VENUS project.

C. *WideNoise: Collective Observation of Environmental Noise*

The WideNoise web application aggregates, summarizes and illustrates noise-related data collected by the WideNoise smartphone application.[17] This smartphone application is used for monitoring environmental noise levels using the mobile device, for uploading it to the WideNoise web application, and for sharing it on social media.

The user can access several statistics through the WideNoise web application. Some of them are public and shown directly on the front page. Others are only visible to the user personally if he or she has a WideNoise account. On the WideNoise start page, the user is provided with several public views on the data, for example:

- A histogram summarizing the number of measurements for each continent for the last three days.
- A table listing the registered users with the most samples overall as well as a table listing the registered users with the most samples covering the last two months – a stimulus for users to put more effort into taking samples.
- A table showing the latest recordings and a table with average values for the last day, month and year. This helps tracing current measurements and comparing one's own samples with the average.

The second type of statistics can be accessed by the user via his personal page. There he finds comprehensive information on his own measuring behavior. The page also provides a KML[18] (Keyhole Markup Language) export containing all user's measurements to display, offering an alternative to the *map* visualization.

The *map* page, is shown in Figure 5. From a user perspective it probably contains the most interesting features. The map view displays, for example, dynamic clusterings of user measurements, and according detail information on demand [21]. Currently, WideNoise is being extended towards other sensor types, for example, for measuring environmental air quality.

[12] http://wintersause.de/

[13] http://www.kde.cs.uni-kassel.de/conf/lwa10/

[14] http://lwa2011.dke-research.de/

[15] http://lwa2012.cs.tu-dortmund.de

[16] http://ht2011.org/

[17] http://itunes.apple.com/de/app/widenoise/id302052132/ (iOS) and https://play.google.com/store/apps/details?id=eu.everyaware.widenoise.android (Android)

[18] http://opengeospatial.org/standards/kml/

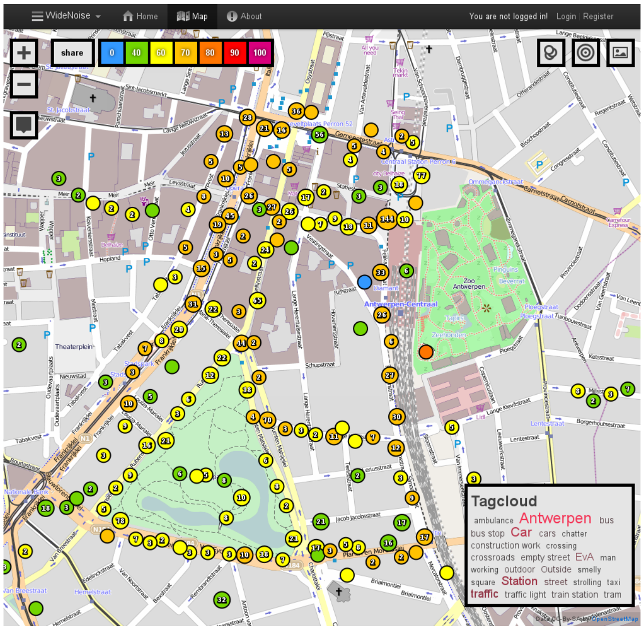

Figure 5. A screenshot of the *map page* of WIDENOISE.

III. MINING PHYSICAL AND SOCIAL ACTIVITIES

The following sections describe several aspects of the presented systems concerning data analysis and data mining. Below, we summarize the data mining techniques that allow us to utilize the collected data for extracting useful information and making it visible to our users: Specifically, we first focus on the dynamics of live user interactions considering the mapping and localization component. After that, we focus on intelligent recommendations concerning user interactions for supporting and enhancing social interactions and communication. Finally, we discuss a collaboration recommender focussing on software development contexts.

A. Mapping Live Interactions

Capturing and visualizing live interactions of individual users is an important task for MYGROUP and CONFERATOR, essentially in order to enable collective intelligence. Therefore, a localization framework is a central component for such a system.

Knowing where colleagues are supports group organization and thus facilitates everyday work processes. The 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 is listening to – apparently interesting – talks, and to potentially recommend those to undecided participants.

The localization framework consists of two parts, the visualization of the individual locations (see Figure 3) and the positioning component. For the calculation of the individual user positions, we focus on localization at room-level. Specifically, we developed an advanced localization framework: We use the information revealed by data mining to improve the localization accuracy. In particular, we exploit the proximity information of other people [20]. Using the social proximity information in the *Social Boosting* algorithm [20], we could improve the localization accuracy from 84 % using a baseline algorithm to nearly 90 %, as evaluated during the poster session at the LWA 2010 conference.

B. 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. We provide two kinds of *user recommendations*: The *Acquaint-O-Matic* for CONFERATOR and MYGROUP and the *Similar Users* section on the profile page of each user. 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 connect to those users. Thus, it provides information and the means to get in touch with new communities. Furthermore, the *Similar Users* section allows to explore other communities as well. While visiting someone's profile page, the displayed list suggests people which are similar to the visited user.

For recommendations, similarity is measured based on previous face-to-face contacts, established links between users, co-authorship relations,[20] 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 [7] algorithm for graph like representations and the cosine similarity measure for calculating similarities between rows in the adjacency matrix is applied.

C. Recommending Software Developers

Locating experts for a given problem is one of the main challenges when working in a large team. In the context of MYGROUP, we focussed on supporting software development groups [15]. 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

[19]Privacy is preserved with certain privacy settings, however, due to the limited space we refer to [1] for more details.

[20]The network of co-authorship, that can be retrieved (mainly for computer science) from DBLP (http://dblp.uni-trier.de/)

is then measured by the number of changed lines of code. We combine this 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 [7] 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 most about a specific source file, or an expertise profile. An example analyzing the Ubicon framework is depicted in Figure 6: 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 that might profit from each other's knowledge are seated together.

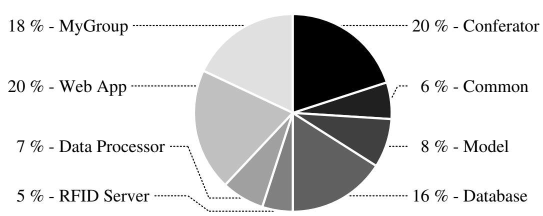

Figure 6. Expertise profile of an exemplary developer generated by the analysis of logs from a revision control system and RFID communication for the Ubicon system.

IV. Case Studies & Exemplary Analysis

In this section, we present exemplary analysis results of the three distinct applications discussed above: The first one is a MyGroup application featuring a long-term data collection of our research group's daily work interactions over a period of six months. The second application example considers the VENUS Technology Day, a one day social networking meeting supported by a combination of MyGroup and Conferator. The third describes exemplarily the application of WideNoise for measuring environmental noise.

A. KDE Research Group

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 II-A. For a first impression of the typical contacts and conversations during a workday, Figure 7 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 on the y-axis. Both axes are scaled logarithmically.

It is easy to see, for example, that 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. It is also interesting to take a look at the distribution of the discussion lengths between the individuals (cf. Figure 8). We focus on the communication behavior of two professors (*prof*), two post-docs (*post*), five PhD candidates (*cand*), and three students (*stud*).

It is unambiguous that professor *prof1* is a very central person. Professor *prof2* is also a major discussion partner but does not have that many connections. This observation can be explained by the fact that *prof2* is affiliated with another university and visiting the KDE group only for a few days per week. The post-docs are involved in many conversations as well since they are project managers and central in project organization. Overall, there is a good connectivity between all PhD candidates, for which the graph shows two extremes – *cand2* with very many (strong) connection, and *cand5* with a low number of connections. The graph shows that the students at that time did not work together and there is only little discussion with the three students at all. This is likely partially due to the part-time nature of their employment, and partially to the construction work that was going on during the considered time frame, which resulted in limited on-site office space for the students.

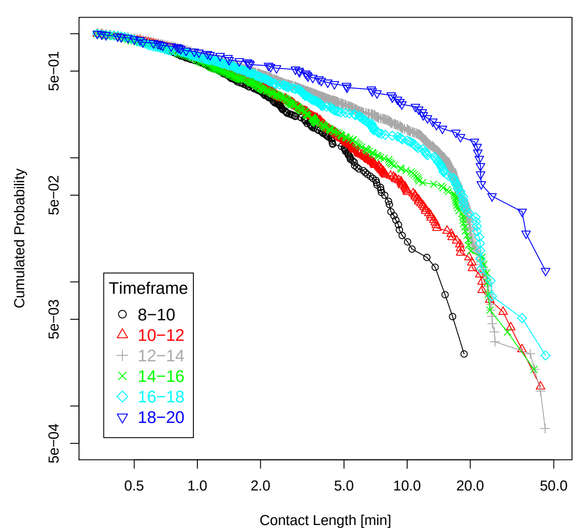

Figure 7. Cumulated probability distribution of face-to-face contact lengths in the KDE group, for different timeframes.

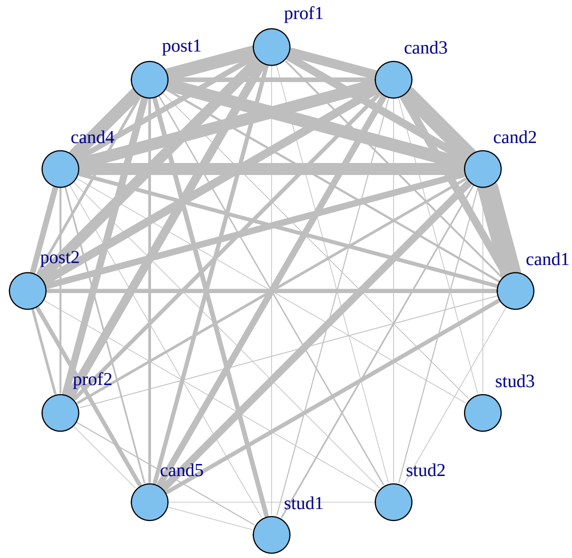

Figure 8. Aggregated contact lengths of the KDE research group.

B. VENUS *Technology Day*

The VENUS Technology Day on May 26th, 2011, was supported by a combined MyGroup and Conferator application. There were 66 participants from the VENUS project and partners from the automotive, medical, chemical, and software industry as well as from research and public institutions. Figure 9 depicts the aggregated contact count graph of all participating groups (without self-edges). As expected, the majority of connections is with the hosting VENUS team, contributing most with 19 connections to Research.

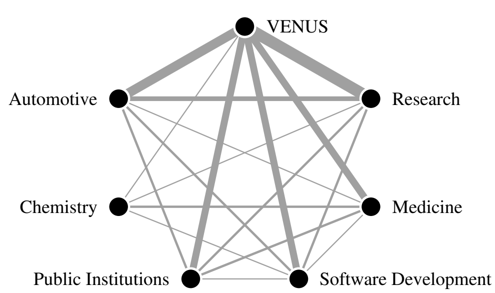

Figure 9. VENUS Technology Day: Aggregated contact count graph of the different participating groups (without self-loops).

Figure 10 shows a key-actor analysis on the aggregated contact network in which only conversations longer than five minutes were considered. Each participant is positioned in the diagram according to his eigenvector centrality and his betweenness centrality. The eigenvector centrality measures the influence of a participant, while the betweenness centrality quantifies the control of a participant on the communication between other participants. The diagram shows that in general the members of the VENUS project and the members of the advisory board ("Beirat") have high centralities.

The higher eigenvector centrality of the project members reflects their overall influence during the technology day, while the higher betweenness centralities of the advisory board members indicate that they play indeed their desired role as bridges. Industrial partners have in general lower centralities, with the exception of "Praxispartner24" and "Praxispartner25".

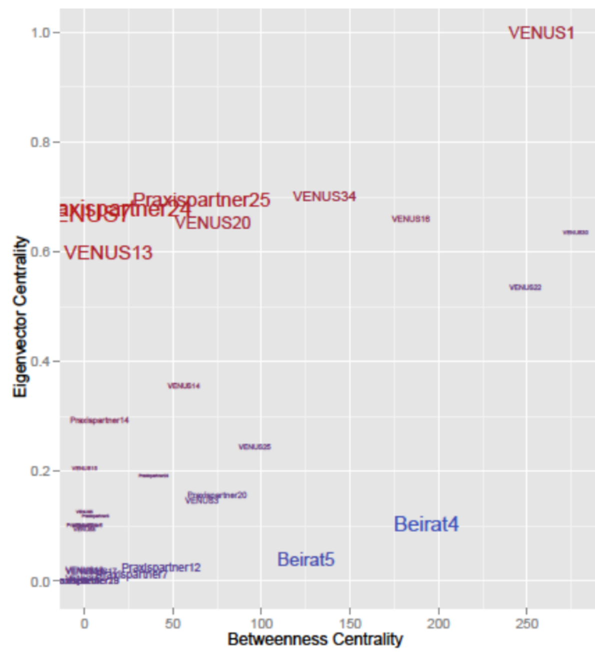

Figure 10. Key-actor analysis for the VENUS Technology Day, showing the betweenness centrality of each participant on the x-axis, and her eigenvector centrality on the y-axis, for conversations that are longer than 5 minutes. Larger labels indicate more "exceptional" centralities.

C. WideNoise *Case Study*

WideNoise is a global ubiquitous system for monitoring environmental noise pollution, for example in cities. The data can then be further analyzed, e. g., by ecologists. In the following, we present first insights about the statistics we collected during the continuous operation of the system (Kassel, London, Paris) as well as during special experiments (Antwerp, Rome): The case study in Antwerp focuses on the 10th of July, 2012 between 9:30 AM and 1:00 PM, roughly covering 2.6 km^2, in which about 1200 measurements were taken. The case study in Rome focused on the 9th of June, 2012 on a square area in Rome roughly covering 0.66 km^2. About 700 measurements were taken in eleven hours. The participants were asked to cover as much of the area with their WideNoise measurements as possible.

Table I shows basic statistics of the collected measurements. We observe, that three big cities (Rome, Antwerp and London) have a higher noise level than the average worldwide and a lower standard deviation. This is especially relevant for London. In contrast, Paris has significantly lower noise level and higher standard deviation. A possible explanation considers different interests of the users participating in the case studies compared to Paris. Additionally, there is a Heathrow-campaign in London, for which the users seem motivated to use the system for measuring higher noise levels.

Table I
BASIC STATISTICS: CITIES/WORLD-WIDE

Case Study	No. of Measurements	Average Noise (dB)	Std. Deviation
Rome	700	68.82	7.15
Antwerp	1160	66.36	10.03
Kassel	698	60.91	17.88
London	3007	73.82	10.64
Paris	1275	60.58	22.69
World-wide	24886	63.94	19.27

When taking measurements with the WIDENOISE app, users can record their perceptions of the actual situation by means of for sliders. Looking at these perceptions for the experiments (Table II) yields additional insights. Compared to the world-wide perceptions, the noise pollution during the experiments was more man-made and more social. Also, cities with more hectic perceptions tend to have higher noise levels. The same can be observed for the 'hate' perception.

In addition to recording their perceptions, the app allows its users to add free text words (so-called tags) to their measurements, in order to specify the context of the ongoing measurement more precisely. Tables III, IV, and V show the tags that were assigned most frequently to measurements during the case studies. Most of the users did not utilize tags (only 10% of measurements are tagged in Antwerp and 25% during the Rome experiment). In almost all cases, the tagged measurements tend to have a higher noise level. This observation also holds for the worldwide data (as shown in the table VI): only 20% of all measurements are tagged and the tagged measurements are on average 10% louder than average. The measurements made in Kassel are an exception since they are tagged more frequently (about 64% of measurements are tagged) and do not show the difference of the noise level between the tagged samples and all the samples.

Most of the tags are quite neutral. However, some of them allow an interpretation of the loudness. There are tags, such as *street*, that suggest an unpleasant loudness. On the other hand, tags like *indoor* or *birds* suggest silence or pleasant noise. No one seems to express loudness directly, but there are some users who tag a quiet area as if this is something special.

V. RELATED WORK

Several systems for tracking people, e. g., conference participants have been built using RFID tokens or Bluetooth-enabled devices: Hui et al. [11] describe an application using Bluetooth-based modules for collecting mobility patterns of conference participants. Concerning social interactions, Eagle and Pentland [10] 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. (cf. [18]) in the Jewish Museum Berlin in 2007. Cattuto et al. [8] added proximity sensing in the Sociopatterns project.

Table II
AVERAGES OF THE USER-ASSIGNED PERCEPTIONS DURING THE CASE STUDIES (EXCLUDING ALL DEFAULT PERCEPTION VALUES OF 0.5)

Perception	Rome	Antwerp	Kassel	World-wide
Love / Hate	0.53	0.62	0.49	0.53
Calm / Hectic	0.48	0.67	0.37	0.53
Alone / Social	0.67	0.67	0.57	0.50
Nature / Man-Made	0.69	0.83	0.81	0.59

Table III
TOP TEN TAGS BY COUNT FOR THE ROME DATA.

Tag	Count	Average Noise (dB)
outdoor	34	73.02
street	27	73.05
car	21	74.07
voice	13	71.58
indoor	12	74.27
wind	6	77.78
birds	5	64.99
bookstore	5	71.47
music	4	75.09
quiet	4	69.05
Tagged Samples	*177*	*73.39*
All Samples/Rome	*700*	*68.82*

Table IV
TOP TEN TAGS BY COUNT FOR THE ANTWERP DATA.

Tag	Count	Average Noise (dB)
street	33	69.90
cars	23	64.35
bus	21	76.40
outdoor	18	78.28
traffic	11	75.92
train station	10	59.08
car	8	74.62
bus stop	7	72.11
construction work	7	61.23
traffic light	6	86.14
Tagged Samples	*115*	*74.39*
All Samples/Antwerp	*1160*	*66.36*

Table V
TOP TEN TAGS BY COUNT FOR THE KASSEL DATA.

Tag	Count	Average Noise (dB)
Bashing	70	70.59
home	40	36.59
office	36	60.65
Office Kassel	35	60.45
background noise	33	47.79
At work	31	53.14
Indoor	29	49.77
test	20	55.89
kassel	19	63.92
In the morning	16	47.16
Tagged Samples	*444*	*59.41*
All Samples/Kassel	*698*	*60.91*

Table VI
TOP THIRTEEN TAGS BY COUNT WORLDWIDE DATA.

Tag	Count	Average Noise (dB)
garden	558	72.11
heathrow	342	67.00
aeroplane noise	319	66.30
Antwerpen	249	74.47
Car	215	74.09
street	146	70.17
plane	142	76.22
Station	138	74.25
traffic	135	75.70
office	103	60.08
EvA	101	76.06
route 62 arriva	88	72.84
Indoor	87	56.54
Tagged Samples	*5312*	*70.45*
All Samples/Worldwide	*24886*	*63.94*

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 profile 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. [16].

For improving collaborative group activities there have been several approaches: [9] and [14] propose and examine the GroupScribbles technique for assisting collaborative activities. [13] 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. Also, we take the dynamic structure of the social interactions into account in order to provide instant recommendations and notifications about people and events.

Kanjo presents in [12] the first system for collecting noise data with mobile phones. Maisonneuve et al. [17] present an approach for monitoring the noise pollution by the general public using the NoiseTube[21] system. Aircasting[22] is another platform which allows users to upload the information about surrounding noise using their mobile phones (currently only Android-based devices supported).

In contrast to the systems mentioned above, the EveryAware WIDENOISE application collects tags for the individual measurements and aims at the combination of subjective and objective data, e.g., tags and noise measurements.

VI. CONCLUSIONS

In this paper we described the overall architecture of the framework and platform UBICON for ubiquitous applications. We showed that very different applications can be hosted on the platform, and described several components for data mining and analysis that have been evaluated in the systems. Furthermore, we provided three different case studies analyzing the collected data in different application contexts. These also show, that the collected data is relevant for research purposes not only for computer science, but also for psychologists, sociologists, ecologists etc.

For future work, we intend to add more sophisticated community mining components [4] to the system. Furthermore, we aim to refine and improve the Acquaint-O-Matic using results from link prediction [19] and to develop further recommenders, e. g., concerning topics, publications, and locations. We also improve the EveryAware system to enable more relevant noise pollution studies to deliver relevant information for supporting political and environmental decisions.

ACKNOWLEDGEMENTS

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 wish to 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.

[21] http://noisetube.net/

[22] http://www.aircasting.org

REFERENCES

[1] M. Atzmueller, D. Benz, S. Doerfel, A. Hotho, R. Jäschke, B. E. Macek, F. Mitzlaff, C. Scholz, and G. Stumme. Enhancing social interactions at conferences. *it – Information Technology*, 53(3):101–107, May 2011.

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

[3] M. Atzmueller, S. Doerfel, A. Hotho, F. Mitzlaff, and G. Stumme. Face-to-face contacts at a conference: Dynamics of communities and roles. In *Modeling and Mining Ubiquitous Social Media*, volume 7472 of *LNAI*. Springer Verlag, Heidelberg, Germany, 2012.

[4] M. Atzmueller and F. Mitzlaff. Efficient descriptive community mining. In *Proc. 24th International FLAIRS Conference*, pages 459 – 464, Menlo Park, CA, USA, 2011. AAAI Press.

[5] A. Barrat, C. Cattuto, V. Colizza, J.-F. Pinton, W. V. den Broeck, and A. Vespignani. High resolution dynamical mapping of social interactions with active RFID. *arXiv*, 0811.4170, 2008.

[6] 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):849–875, 2010.

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

[8] 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):e11596, July 2010.

[9] C. Digiano, D. Tatar, and K. Kireyev. Learning from the Post-it: Building collective intelligence through lightweight, flexible technology. In *Proc. CSCW*, pages 65–66. ACM, 2006.

[10] N. Eagle and A. S. Pentland. Reality mining: Sensing complex social systems. *Pers. Ubiquit. Comput.*, 10(4):255–268, Mar. 2006.

[11] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot. Pocket switched networks and human mobility in conference environments. In *SIGCOMM*, pages 244–251, New York, USA, 2005. ACM.

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

[13] 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 *Proc. ICDE*, pages 1483–1486, 2009.

[14] C.-K. Looi, W. Chen, and F.-K. Ng. Collaborative activities enabled by GroupScribbles (GS): An exploratory study of learning effectiveness. *Comput. Educ.*, 54(1):14–26, Jan. 2009.

[15] B.-E. Macek, M. Atzmueller, and G. Stumme. Profile mining in CVS-logs and face-to-face contacts for recommending software developers. In *Proc. SocialCom*. IEEE, 2011.

[16] B. E. Macek, C. Scholz, M. Atzmueller, and G. Stumme. Anatomy of a Conference. In *Proc. ACM Hypertext 2012*, 2012.

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

[18] M. Meriac, A. Fiedler, A. Hohendorf, J. Reinhardt, M. Starostik, and J. Mohnke. Localization techniques for a mobile museum information system. In *Proc. WCI*, 2007.

[19] C. Scholz, M. Atzmueller, and G. Stumme. On the Predictability of Human Contacts: Influence Factors and the Strength of Stronger Ties. In *Proc. Fourth ASE/IEEE International Conference on Social Computing (SocialCom)*, Boston, MA, USA, 2012. IEEE Computer Society.

[20] C. Scholz, S. Doerfel, M. Atzmueller, A. Hotho, and G. Stumme. Resource-Aware On-Line RFID Localization Using Proximity Data. In *Proc. ECML/PKDD 2011*, pages 129–144, 2011.

[21] B. Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proc. IEEE Symposium on Visual Languages*, pages 336–343, Boulder, Colorado, 1996.