COLLA 2015

The Fifth International Conference on Advanced Collaborative Networks, Systems and Applications

ISBN: 978-1-61208-436-7

October 11 - 16, 2015

St. Julians, Malta

COLLA 2015 Editors

Pascal Lorenz, University of Haute-Alsace, France

Pierre Leclercq, University of Liege, Belgium

COLLA 2015

Forward

The Fifth International Conference on Advanced Collaborative Networks, Systems and Applications (COLLA 2015), held between October 11 - 16, 2015 - St. Julians, Malta, continued a series of events dedicated to advanced collaborative networks, systems and applications, focusing on new mechanisms, infrastructures, services, tools and benchmarks.

Collaborative systems became a norm due to the globalization of services and infrastructures and to multinational corporation branches. While organizations and individuals relied on collaboration for decades, the advent of new technologies (Web services, Cloud computing, Service-oriented architecture, Semantics and Ontology, etc.) for inter- and intra-organization collaboration created an enabling environment for advanced collaboration.

As a consequence, new developments are expected from current networking and interacting technologies (protocols, interfaces, services, tools) to support the design and deployment of a scalable collaborative environments. Innovative systems and applications design, including collaborative robots, autonomous systems, and consideration for dynamic user behavior is the trend.

The conference had the following tracks:

- Collaborative architectures and mechanisms
- Cooperation and collaboration mechanisms
- Collaborative applications

Similar to the previous edition, this event attracted excellent contributions and active participation from all over the world. We were very pleased to receive top quality contributions.

We take here the opportunity to warmly thank all the members of the COLLA 2015 technical program committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to COLLA 2015. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the COLLA 2015 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope COLLA 2015 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of collaborative networks, systems and applications. We also hope that St. Julians, Malta provided a pleasant environment during the conference and everyone saved some time to enjoy the beauty of the city.

COLLA 2015 Chairs

COLLA Advisory Committee
Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Mohamed Eltoweissy, Virgina Military Institute/Virginia Tech, USA
Howard Spoelstra, Open University in the Netherlands, Netherlands
Pascal Salembier, University of Technology of Troyes, France
Atsuo Hazeyama, Tokyo Gakugei University, Japan
Henric Johnson, Blekinge Institute of Technology, Sweden
Christophe Thovex, University of Nantes, France
Alessandro Campi, Politecnico di Milano, Italy
Anna Kocurova, University of West London, United Kingdom

COLLA Industry/Research Chairs
Anna Divoli, Pingar Research, New Zealand
Jin-Hee Cho, U.S. Army Research Laboratory (USARL) - Adelphi, USA
Holger Kett, Fraunhofer IAO - Stuttgart, Germany
Raymond A Liuzzi, Raymond Technologies - Whitesboro, USA
Jérôme Dantan, Esitpa, France

COLLA Publicity Chairs
Dapeng Dong, University College Cork, Ireland
Benjamin Knoke, BIBA – Bremer Institut für Produktion und Logistik, Germany
Mateusz Adamczyk, University of Warsaw, Poland
Kenji Takagi, Tokyo Metropolitan University, Japan
Yann Veilleroy, Université Catholique de Lille, France

COLLA 2015

Committee

COLLA 2015 Advisory Committee

Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Mohamed Eltoweissy, Virgina Military Institute/Virginia Tech, USA
Howard Spoelstra, Open University in the Netherlands, Netherlands
Pascal Salembier, University of Technology of Troyes, France
Atsuo Hazeyama, Tokyo Gakugei University, Japan
Henric Johnson, Blekinge Institute of Technology, Sweden
Christophe Thovex, University of Nantes, France
Alessandro Campi, Politecnico di Milano, Italy
Anna Kocurova, University of West London, United Kingdom

COLLA 2015 Industry/Research Chairs

Anna Divoli, Pingar Research, New Zealand
Jin-Hee Cho, U.S. Army Research Laboratory (USARL) - Adelphi, USA
Holger Kett, Fraunhofer IAO - Stuttgart, Germany
Raymond A Liuzzi, Raymond Technologies - Whitesboro, USA
Jérôme Dantan, Esitpa, France

COLLA 2015 Publicity Chairs

Dapeng Dong, University College Cork, Ireland
Benjamin Knoke, BIBA – Bremer Institut für Produktion und Logistik, Germany
Mateusz Adamczyk, University of Warsaw, Poland
Kenji Takagi, Tokyo Metropolitan University, Japan
Yann Veilleroy, Université Catholique de Lille, France

COLLA 2015 Technical Program Committee

Imad Abugessaisa, Karolinska Institutet - Stockholm, Sweden
Marcelo Armentano, ISISTAN Research Institute (CONICET/UNICEN), Argentina
Liz Bacon, University of Greenwich, UK
Thierry Badard, Laval University, Canada
Mostafa Bassiouni, University of Central Florida, USA
Khalid Benali, LORIA - Université de Lorraine, France
Lasse Berntzen, Vestfold University College - Tønsberg, Norway

Table of Contents

Enhancing Mobile Devices with Cooperative Proactive Computing 1
Gilles Irenee Fernand Neyens, Remus Alexandru Dobrican, and Denis Zampunieris

Instrumented Analysis Method for Collaboration Activities 10
Samia Ben Rajeb and Pierre Leclercq

Online Learning Community Software to Support Success in Project Teams 17
Brian Thoms and Evren Eryilmaz

ShareLab, Support for Collective Intelligence. 1 deadline, 11 designers, 1 project. 27
Samia Ben Rajeb, Alexandru Senciuc, and Irene Pluchinotta

Agile Knowledge Networking - A Novel Approach to Research Collaboration Between Industry and Academia 34
Bjorn Erik Munkvold

Exploiting Argumentation Content and Structure to Advance Collaboration through Hybrid Recommendations 40
Spyros Christodoulou, Nikos Karacapilidis, and Costas Pappis

A Policy Specification Language for Composite Services 46
Muhammad Asim, Susana Gonzalez Zarzosa, Qi Shi, and Bo Zhou

SilentMeet - A Prototype Mobile Application for Real-time Automated Group-based Collaboration 52
Remus-Alexandru Dobrican, Gilles Irenee Fernand Neyens, and Denis Zampunieris

An Experience on Leadership Identification in Social Cognocracy Network 57
Alberto Turon, Juan Aguaron, Jose Maria Moreno-Jimenez, and Maria Teresa Escobar

An Attempt to Evaluate Chances and Limitations of Social Information Retrieval 60
Christoph Fuchs and Georg Groh

Enhancing Mobile Devices with Cooperative Proactive Computing

Gilles I. F. Neyens , Remus-Alexandru Dobrican , Denis Zampunieris
University of Luxembourg
Luxembourg, Grand-Duchy of Luxembourg
Email: gilles.neyens.001@student.uni.lu , {remus.dobrican, denis.zampunieris }@uni.lu

***Abstract*—With the increasing popularity of smartphones and with the fact that they are connected to the Internet most of the time, people manage to stay online everywhere they go. They can access online services remotely at any time they want, using their mobile devices. However, in order to make the best out of these circumstances, the users have to use sophisticated mobile applications. These applications do not have to only address key aspects like collaboration and cooperation between various devices but have to deal also with the involvement of the users in order to achieve the desired outcome. The main contribution of this paper is to present a solution, i.e., Proactive Engine for Mobile Devices (PEMD), together with its implementation for Android-based systems, for enhancing mobile devices with proactive properties. The model serves as a basis for developing smart applications that are able to perform complex real-world tasks. Furthermore, it provides a method for achieving cooperation, coordination and collaboration of multiple smart devices. Finally, we provide the performance experiments and we discuss the results and the effects of using PEMD on different devices.**

Keywords–Mobile devices; Collaborative architectures and mechanisms, Proactive Computing.

I. Introduction

Smartphones and tablets have become more and more popular and more powerful in terms of their computing capacity [1]. This opens numerous possibilities for developing new types of software applications and, for the existing ones, to expand, to incorporate new features and to provide better services to their users. Also, due to the latest advances in computer hardware technology [2] [3], mobile devices are more than ready to be enhanced with proactive features and properties. These properties are essential for enhancing mobile devices with collaborative methods for their applications.

A great challenge of today's mobile world is to provide services and applications that support collaboration and distributed tasks. Middleware systems are seen very often as solutions for supporting the construction of mobile collaborative applications [4]. We investigated how to implement a more specific category of middleware systems, i.e., rule-based middleware systems, because they offer various advantages over other middleware systems such as the possibility of having rules where the developers can easily insert their instructions. These rules, among others, can deal with unforeseen situations, can predict future events and can be programmed to take appropriate actions in most of the situations. The main difficulty consists of implementing this kind of systems on mobile devices because they are very complex and they need many resources, depending on which kind of applications they are executing. Having desktop rule-based systems executing all the tasks of the mobile devices is not a valid solution as the communication between them is not very stable because of the mobility of the mobile devices. The probability of having successful collaborative tasks increases when each mobile device has an integrated middleware architecture, capable of initiating collaborations and performing distributed jobs. The major advantage of our middleware architecture is that it is capable of communicating with other mobile devices and desktop computers without any compatibility issues.

The contribution of this paper is three-fold. First, it proposes a rule-based engine capable of performing complex automated and distributed tasks on mobile devices on behalf of the user. Second, it offers details about the implementation of our model on Android-based mobile devices. And third, it shows how the model was tested on real devices and provides an evaluation of their performance.

The rest of the paper is organized as follows: Section II describes the state of the art related to this study. Section III provides the main characteristics of our model in terms of interaction with the operating system, communication mechanisms, information sharing strategies and data storage. In Section IV, we present the tests that were done to evaluate our model, what methodology was used, then we provide an analysis of these results and, at the end of the section, we discuss the future implications of the results. Finally, Section V concludes the paper and indicates the next research directions that will follow after this work.

II. State of the Art

Proactive Computing was introduced in [5] and was described as a new computational way in which software systems can operate and can perform different tasks. It was initially seen as a solution for removing humans from the computational loop, for moving from human-centred computing to human-supervised computing. Proactive Systems were characterised as those systems that can work for the user and that can take decisions on their own initiative, without necessarily involving the user in this process [6].

More recent studies [7][8] describe Proactive Systems as being aware of the changes, which occur in their surrounding environment, being able to react to foreseen events and of adapting their behaviour in order to address the increasing needs of their users. Recent empirical investigations [9][10] support these properties and provide additional examples of the advantages that Proactive Systems have to offer in numerous domains of Computer Science. For instance, enhancing Learning Management Systems (LMSs) with Proactive Computing led to an increase of the online participation of the students [9], helped improve the assignment system [11] and raised the chances of student to obtain higher grades at their final exams [10].

 ISBN: 978-1-61208-436-7

```
Rule 102
Description: This Rule was activated by Rule 101 and started
creating a Community of Practice for new cities detected from
the student's profile.
data acquisition
    String groupName = cityName;
    String [] students = getStudentFromSameCity (cityName);
activation guards
    return groupExists(cityName);
conditions
    return true;
actions
    foreach student in students []
        if(userIsNotPartOfGroup(student.ID, groupName)))
            inscribeUserInGroup(student.ID, groupName);
        end if
    end foreach
rules generation
    if (activationGuard());
        createRule103(groupName());
    end if
    cloneRule(Rule102);
```

Figure 1. A Proactive Rule in pseudo-code

More precisely, using a Proactive Engine (PE) aside the LMS transformed the LMS from a static system into a proactive system. The PE was created as a structure capable or executing Proactive Rules [12], which are simple mechanisms for executing specific actions, like, for instance, sending emails to the users. They were designed for offering developers the possibility of implementing proactive actions in a simple way, without having to have advanced knowledge about proactive systems and their implementation. In Figure 1, an example of a Proactive Rule is given in pseudo-code. The rule was implemented in Java and was used as part of a predefined scenario for automatically creating certain Communities of Practice inside a LMS [13].

Multiple collaborative middleware systems for mobile devices are available on the market, e.g., [14]-[16]. They differ by their type: event-based architectures, e.g., for supporting location aware-mobile applications [17], or publisher-subscriber architectures, e.g., for [4] systems. These frameworks offer communication services for the mobile devices, which use them for performing collaborative tasks. Our framework does not only achieve message passing from one device to another but also permits more complex actions like remote rule activations, parallel commands or complex reasoning algorithms.

Rule-based systems, like Java Expert System Shell (JESS) [18] or C Language Integrated Production System (CLIPS) [19], are powerful desktop-based general-purpose tools, which give the possibility of programming expert systems. With the help of their scripting language facts and rules can be uniformly defined and described. For example, JESS employs the Rete [20] algorithm for compiling and executing forward-chain rules. These rules are simple statements, composed of a left side, i.e., the IF portion, and of a right side, i.e., the THEN portion. Unfortunately, they represent solutions only for the server side and are not suitable for mobile devices as these devices are quite limited in terms of computing power. We do not only propose an engine for mobile devices capable of executing complex rules but we propose a technique, called Global Proactive Scenario (GPaS) [21], for breaking down complex scenarios or applications in multiple sets of Proactive Rules. More about this technique is explained in subsection III-D.

Existing rule-based systems [22][23][24] for mobile devices solve only simple tasks and do not provide methods for achieving more complex tasks like distributed reasoning, task distribution, data sharing, acquiring global context information or/and collaborative filtering. And most of all, these systems do not take advantage of the global information that can be built with information from each particular mobile device. IF THIS THEN THAT [24] is a mobile application that realizes automation for small tasks between Internet-connected services. The user can write simple rules, also called recipes, in order to achieve different goals like adding the photo of a user to the cloud-based archive if the user has been tagged in that particular photo on Facebook [25]. These rules, however, are just simple conditional statements. HeaRT [22], a lightweight rule-based inference engine designed for mobile devices, was used in [26] for providing simple tasks like online reasoning, part of a bigger plan to develop context-aware mobile applications. The rules that are written for this engine can achieve local reasoning only based on the internal sensors of a mobile device and do not explore the possibility of having multiple engines performing global reasoning. Minimal Rule Engine (MiRE) [23], a context-aware processing engine, was implemented in order to obtain an engine capable of processing rules on mobile devices. However, the rules are written in Extensible Markup Language (XML) and, due to their structure, are not capable of integrating more complex logics. The above approaches, [22]-[24], try to address the growing demand of using rule-based tools on mobile platforms and manage to do it but for a very narrow type of applications.

Until now, Proactive Engines were only used on desktop computers [13][27][28], which limited a lot their usability. Analysing the advantages and functionalities of mobile devices makes it clear that mobile devices offer new possibilities for Proactive Systems, as well as the other way around. Proactive Computing can help mobile device become smarter in terms of how they make use of the data coming from all the sensors, of how they exchange information, of how they execute complex tasks and in terms of how they provide services to their users. We therefore developed a PE for mobile devices.

III. THE ARCHITECTURE OF PROACTIVE ENGINES FOR ANDROID-BASED MOBILE DEVICES

In order to better understand the structure of the new engine we will now first explain parts of the already existing server-based engine. The Rules Engine is the core piece of the engine. It consists of two First In First Out (FIFO) lists: the Current Queue, which contains the rules that are currently being executed and the Next Queue, which contains the rules that were created during the current iteration of the Rules Engine and that will be executed at the next iteration. During an iteration, the Rules Engine will execute the rules that are stored in the Current Queue one by one. Proactive Rules can perform different operations like checking for special conditions or constraints, saving events or relevant context information into the local database, or cloning themselves in order to run at the next iteration. They can also generate other Proactive Rules during the *rules generation phase*, which is one of the five phases that compose a Proactive Rule, as shown in Figure 1. The new rules, created during the *rules generation phase* are stored in the Next Queue. An iteration finishes its execution if there are no more rules in the Current Queue or if there were already N rules executed during the current iteration or if the execution time exceeded the time limit F, an internal parameter of the PE, which represents the frequency of activation periods. At the end of the iteration all rules contained in the Next Queue are added to the Current Queue and the Next Queue gets cleared. The Current Queue is then saved into the database and the Rules Engine will continue with the next iteration. This Rules Engine is the basis for the implementation of our PE for Mobile Devices (PEMD). Other components like the database also had to be adapted in order to fit the requirements of mobile devices.

A. The Rules Engine as background service

The Rules Engine was designed to run constantly in the background in order to allow the user to interact with different applications. On Android this can be achieved with the help of background services. However, these services cannot run constantly as they might be killed or stopped by the Operating System if they use too many resources over an extended period of time. Therefore, the existing solution for desktop computers needed to be adapted in order to still allow the PE to execute Proactive Rules. This was done using an Alarm Manager that activated the background service every F seconds and that executed one iteration of the Rules Engine. The Alarm Manager was triggered by the *onBoot* event.

The added essential functionalities are the communication of PEs, explained in detail in Section III-C, and the possibility of notifications to the user from within rules, briefly explained at the end of section III-B.

B. Data Storage

Data storage is an important part of the application. The rules in the Current Queue of the Rules Engine need to be saved at every iteration so that the PEMD can recover in case of failure by using a previous state. This is particularly important as the Android Operating System (OS) may decide to kill the background service that contains the Rule Engine, if the system is low on memory. If the Current Queue is not saved at every iteration, the Rules Engine would recover the state it had when it first started, meaning that every progress in the execution of the Rules is lost. Also, the rules need to be saved when the device shuts down so that the engine can recuperate and continue executing them when the device is turned on again. The saving process of the Current Queue is performed in a transaction, both for performance and failure recovery reasons.

Additionally, the sent and received messages need to be saved to allow the engine to resend lost messages and also the notifications displayed to the user. As there will likely be different types of rules and notifications depending on the purpose of the rules engine, the proposed solution will automatically create the appropriate database table upon installation or update of the application, which is achieved through the Object Relational Mapping Lite (ORMLite) [29] package framework for Android. The ORMLite package is a lightweight package for persisting Java objects to Structured Query Language (SQL) databases. To add new types of rules or notifications to the PEMD, one has only to create the appropriate file with the correct annotations and the PEMD will take care of creating the correct tables and of the saving. Notifications use the internal notification system provided by Android and can be further customised and modified by the applications developers if needed.

C. Communication of Proactive Engines

The most important part of the PEMD architecture is the communication between several devices as this allows the devices to exchange information. There are several technologies available for smartphones to achieve this, each with their own advantages and disadvantages. After analysis of alternatives, we decided that the solution, which suits our needs best and which was used in our implementation of the PEMD is Google Cloud Messaging (GCM). There are another few alternatives on the market like Parse [30], PubNub [31] or UrbanAirship [32]. Nevertheless, while these services make it easier to develop push notifications for iOS and Android, they are still using GCM. One alternative, which does not use GCM at all, is Pushy [33]. However, Pushy's architecture is very similar to GCM, maintaining its own background socket connection, to receive push notifications [34].

1) Google Cloud Messaging: GCM for Android is a service provided by Google, which allows the sending and reception of data between a server and Android-based smartphones. The GCM service handles the delivery process of the messages, meaning that it takes care that the messages are delivered to the correct device. In the case where a message is sent to an offline device, GCM temporarily stores the message until the receiving device comes back online. In order to use GCM for device-to-device communication, we proposed an architecture, illustrated in Figure 2 and Figure 3. The registration process of two devices is shown in Figure 2, while Figure 3 is showing the communication between already registered devices. The devices first have to register at the GCM and get a registration ID. They then communicate this ID along with a username to the server, which stores them in a database. If an already registered device now wants to communicate with another registered device, it sends its message along with the receiving device's username to the server, which retrieves the correct ID and pushes the message along with ID to the GCM, which then takes care of the delivery of the message.

2) Message Structure: The messages exchanged between PEMDs need to follow the same standards. The messages are encoded with JavaScript Object Notation (JSON) and have different attributes depending on the type of message. The only attribute that is common to all message types is the instruction attribute that allows the receiving PEMD to parse incoming messages correctly. There are currently two types of messages, the activate Rule message type and the confirmation message type. The activate Rule message type has the following attributes:*instruction*,*msgID*, *senderID*, *receiver ID*, *ruleName*, *parameterList*.

 ISBN: 978-1-61208-436-7

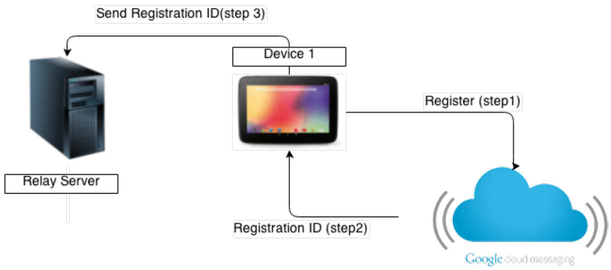

Figure 2. The registration process for one device, passing through Google's Cloud Messaging Server

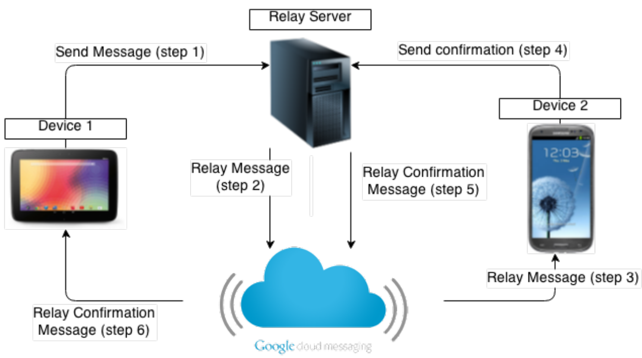

Figure 3. The communication process between 2 registered devices

The *msgID* along with the *senderID* allows the receiving PEMD to keep track of already received messages. This will be explained in more detail in the error handling section. The *ruleName* and *parameterList* attributes are the core of the communication process. They allow the creation of a Rule on another PEMD. The *ruleName* attribute contains the name of the rule that will be created on the receiving device and the *parameterList* contains the parameters necessary to create this rule dynamically. After the rule is created, it will be added to the Next Queue of the Rules Engine. An example of a message exchanged between PEMDs, which contains a command to active a Proactive Rules, can be seen in Figure 4. The confirmation message only contains three attributes: *instruction*, *msgID*, *receiverID*, where *msgID* is the ID of the message whose delivery is confirmed.

3) Error Handling: In a distributed environment, messages are not guaranteed to arrive, as they can be lost along the communication process. In order to prevent the loss of messages, PEMDs keep track of sent and received messages,

{ 'instruction':'activate rule',
'msgId':'messageID ',
'senderID':'ID ',
'receiverID':['registrationID '],
'ruleName':'R004 ',
'parameterList ':['param1 ', 'param2 '] }

Figure 4. Example of a message exchanged between 2 PEMDs, which contains a command to activate a Proactive Rule

including their senders and receivers, by saving them locally in a database table. After sending a message, the PEMD saves it to the database. If no confirmation message is received within a given time period, the message is sent again. The time period varies depending on the priority level of the message, which can be set when sending it. It is also possible to set no priority level at all so that the message does not have to be confirmed. This is quite useful to do broadcasts in order to find other devices with the same preferences, where it is not important that really every device receives the broadcast. Upon receiving the confirmation for a specific message, this message is deleted from the database.

Similarly after receiving a message, the message is saved in the database and a confirmation message is sent to the sender of the message. In the case the confirmation message does not arrive, the sending device resends the message and as the receiving device has saved the message in its database, it ignores the second message. The stored messages are deleted after a fixed time period. In order to take care of the resending and ignoring of messages, there are two rules constantly running on the engine; one rule that takes care of the messages that were sent and one rule that takes care of the received messages.

4) Limitations of GCM and Workarounds: GCM has two main disadvantages, a size limit on the messages and a limit on the number of devices a message can be simultaneously sent to. The size limit on the messages is of 4 Kb. A lot of messages of our application can be delivered using the direct method as they are smaller than 4Kb. If the size limit is exceeded, our server will store the complete message in its database and just send a small message to the receiving device to notify it that a message is available. The device will then download the message directly from the server and add the rule to the queue.

The second restriction of GCM is that a message can only be sent to 1000 devices simultaneously. In our application this can only happen if a broadcast is sent to all devices. In this case, the server just splits the list of all devices into packets of 1000 and pushes the same message for every packet to the GCM.

D. Information Sharing Strategies between Proactive Engines

As explained in Section I, GPaSs [21] were proposed in order to address complex situations or tasks like the collaboration and collaboration between multiple Proactive Engines. Depending on their complexity, GPaSs contain one or more Proactive Rules from different categories of rules like Adaptation Rules, Cooperation Rules, Coordination Rules, Communication Rules and Notification Rules. For example, an application that would automatically form social groups of people based on common interests or activities would use Cooperation Rules to exchange information between the devices, Context-Awareness Rules to obtain details about the user's context and Notification Rules to keep the users up-to-date with the application's latest actions. The developers have to focus more on how to decide what functionalities they want to include into the application and how to transform it into Proactive Rules then to take care of how the PE executes the rules or how the information is exchanged between multiple PEs.

IV. RESULTS AND DISCUSSIONS

Performance is an important aspect of software development, especially when designing and implementing smart applications for mobile devices. It was thus necessary to investigate how PEMD perform on different devices and what are the factors that influence the overall performance of GPaSs.

In this section, we present the hardware and software specifications of the devices on which PEMD was tested, we explain which type of experiments were performed, why we chose to carry out these experiments and we discuss the results that we collected after the tests.

A. Hardware and Software specifications

Three different computing systems were included in our tests: a smartphone, a tablet and a desktop computer. In Figure 5 we provide, for each device, the most important software and hardware specifications at the moment the tests were performed. These specifications include the versions of the Operating System and of the Kernel, the types of central processing units (CPU) and graphics processor units (GPU) used, the amount of random access memory (RAM), the available sensors and the types of communication protocols that could be used by the devices. The smartphone and the tablet have been available on the market since a couple of years. They were chosen for the tests to show how PEMD behave on devices that are used by the majority of the existing applications on the market and which contain stable version of the OS and of the other frameworks used. The PC was included in our tests to check if there are any specific factors on mobile devices that are influencing the performance of PEMD.

B. Methodology

Our main goal was to see approximatively what is the highest number of Proactive Rules PEMDs are able to run on various mobile devices without affecting the overall performance of these devices. More precisely, we focused on evaluating the how many Proactive Rules can be executed by the PEMD in a reasonable amount of time in order to still be able to provide real-time services to the user.

Our method for analysing the performance of the PEMD involved measuring the time between two consecutive *iterations* of the Rules Engine. An *iteration* is an executing instance composed of a set of Proactive Rules in the Rules Engine and can be measured in terms of duration. Ten different sets of Proactive Rules were considered for the tests: starting from small sets containing 100 rules until large sets with 1000 rules. From a technical point of view, an iteration is composed of two main operations: *the execution phase* and the *saving phase*.

During the execution phase, the Rules Engine runs each instruction, part of each Proactive Rule, which can contain possible actions like acquiring data from the sensors or from the local database, sending notification to the users or even the generation of other rules for the next iteration. The saving phase is mainly used for saving the set of rules that are to be executed during the next iteration. A list of rules, together with their parameters are saved into the local database. This phase was created as a safety measure in case a crash occurs or the Rules Engine is stopped. After a crash, when the Rules Engines restarts, it reads the last list of rules that was saved in the database and it will start executing them.

Multiple rounds of tests were performed, each round of tests containing 10 evaluations for each device. The execution time averages for all the tests were computed for obtaining more accurate values. Other applications and services running on the devices involved in the tests were not explicitly closed because we wanted to analyse the performance of the PEMD in the same circumstances that a common user would be using his/her device.

C. Performance analysis

Table I contains the averages of the total amount of time in milliseconds that one iteration required for running sets of 100, 300, 700 and 1000 Proactive Rules. The total time is divided furthermore into *Saving Time* and *Execution Time*, which are average values in milliseconds that represent the amount of time needed by the Rules Engine for the *Saving Phase* and for the *Execution Phase*. All the sets of rules contained clones of the same basic Proactive Rule. This rule was designed specially to see how much time does the Rules Engine need to save an instance of all the rules that will run at the next iteration. This explains the relatively low values for the *Execution Time*.

For example, running 100 Proactive Rules on the smartphone took, in average, 436 milliseconds for one iteration. The time needed to save the rules for the next iteration took 329 milliseconds, representing 75% of the total amount of time of the iteration. This case is confirmed by the values obtained from running the same 100 Proactive Rules on the tablet, were the time for saving the rules for the next iteration took 83% of the total time of the iteration. The same results were obtained for *saving the rules in the* was also confirmed by the values obtained from running 300, 700 and 1000 rules. If for 100 rules it took 75% of the total iteration time for saving the rules in the database, for 1000 rules the percentage decreased to 60% of the total iteration time and stabilised around that value.

Significant differences for running an instance with 1000 rules, between the smartphone, table and PC, appear in the last column of Table I. If in the PC's case the total time for finishing the execution of one iteration took 220 milliseconds, which is quite fast, the same operation took almost twice more on the tablet and approximately 5 times more on the smartphone.

The results in Table I also meet our expectations in terms of computing capabilities of the involved systems. For instance, executing one iteration with 100 rules required approximately 4 times more time on the tablet than on the PC and more than 8 times more on the smartphone that on the PC. The difference did not change much when executing 300, 700 and 1000 rules. This is mainly due to the particular hardware configuration of each system as illustrated in Figure 5. The smartphone was equipped with a 1 Gigabyte (GB) of RAM and a quad-core 1.4 Gigahertz (GHz) processor, while the tablet, which had better performance results, was equipped with 2 GB of RAM and a dual-core 1.7 GHz processor. The PC had the best results as it had 8 GB of RAM and an i7 processor with 4 physical cores capable of operating at frequencies up to 3.4 GHz.

In conclusion, saving rules on the database took a lot of time in comparison with executing the rules. This may be

Device	Samsung Galaxy S3	Nexus	PC
Model	GT-I9300	10	
OS	Android 4.3	Android 4.4.4	Window7 64-bit
Baseband Compilation	I9300 XXUGNA8	KTU84P	
Kernel	3.0.31-2429075		
RAM	1 GB	2 GB	8 GB
Chipset	Exynos 4412 Quad	Exynos 5250	Intel 6 C200
CPU	Quad-core 1.4 GHz Cortex-A9	Dual-core 1.7 GHz Cortex A15	Intel Core i7-2600 3.4 GHz
GPU	Mali-400MP4	Mali-T604	
Sensors	Accelerometer, gyro, proximity, compass, barometer	Accelerometer, gyro, proximity, compass, barometer	None
WLAN	Wi-Fi 802.11, a/b/g/n, dual-band, Wi-Fi Direct, DLNA, DLNA, Wi-Fi hotspot	Wi-Fi 802.11, a/b/g/n, dual-band, Wi-Fi Direct, DLNA, DLNA, Wi-Fi hotspot	None

Figure 5. Hardware and software specifications of the devices used in the experiments

TABLE I. Iteration average time on different devices

#rules/iteration		100		300		700		1000	
Smartphone	Iteration Time (ms)	436.7		672.5		1099.8		1294.2	
	Saving Time (ms) / Execution Time (ms)	329 (75%)	107.7 (25%)	475 (71%)	197.5 (29%)	671.4 (61%)	428.4 (39%)	772.2 (60%)	522 (40%)
Tablet	Iteration Time (ms)	130.5		233.2		383.9		495.5	
	Saving Time (ms) / Execution Time (ms)	108.6 (83%)	21.9 (17%)	175.4 (75%)	57.8 (25%)	275.4 (71%)	108.5 (29%)	352.4 (71%)	143.1 (29%)
PC	Iteration Time (ms)	42.5		68.5		127.5		220	
	Saving Time (ms) / Execution Time (ms)	6.5 (15%)	36 (85%)	2 (3%)	66.5 (97%)	24.5 (19%)	103 (81%)	35.4 (16%)	184.6 (84%)

caused by the ORMLite package framework used for data storage on Android. On the other hand, the data storage on the PC was done with the help of MySQL [35] and Hibernate ORM [36] frameworks and needed far less time than for executing rules. A possible solution for avoiding time losses is to remove the feature of saving rules at each iteration and to set up a Saving Phase for the Rules Engine only once, e.g., at the shutdown event of the mobile devices. Another important conclusion is that running a big amount of rules on the smartphone and tablet is a time-consuming process and, for the moment, these devices are able to run only limited sets of Proactive Rules, while still being able to provide real-time services to their users. This amount of rules is directly related to the computing capabilities of each device and the libraries used to ensure different functionalities of the PEMD.

Other series of tests were conducted between the smartphone and the tablet to measure the duration of their communication. In Figure 2 from section III, we illustrated the steps necessary to send a message and to get the confirmation that the message has been successfully received. Both operations need 3 steps. We measured how much time it took for sending and receiving the message, meaning the first 3 steps of the communication process, between the two registered devices. After 10 tests, in average, it took 751.3 milliseconds for the following operations: the creation of the message on device 1, sending the messages to the relay server, sending the message to Google's Cloud Messaging server and then, receiving the message on device 2. The message used in the experiments was the same one as the one as presented in Figure 4. It contained specific instructions for activating a rule on device 2. The Relay Server was running on the PC with the software and hardware configurations shown in Figure 5. The entire communication process also depends on external factors like the network bandwidth and latency, and on the response time of Google's Cloud Messaging Server, which is also not part of our system and cannot be controlled.

Figure 6 and Figure 7 illustrate better the differences of running one iteration between the different devices. They include the average results of all the 10 sets of Proactive Rules that were performed on all 3 devices. The scale for the execution time of one iteration was kept on purpose to show the major difference between amount of time needed for one iteration with the saving phase and without the saving phase. In Figure 6 we can notice quite big fluctuations in the execution time of one iteration on the smartphone. It is not linear like in the tablet's case and the PC's case. If until running iteration of 800 rules the time increased in an expected manner, afterwards, it started to rise up quickly. It means that for iterations with more than 800 rules the smartphone is starting to consider the PEMD quite heavy in terms of the processing resources needed. It can also slow down other applications that need to access the same resources. In Figure 7 , however, the

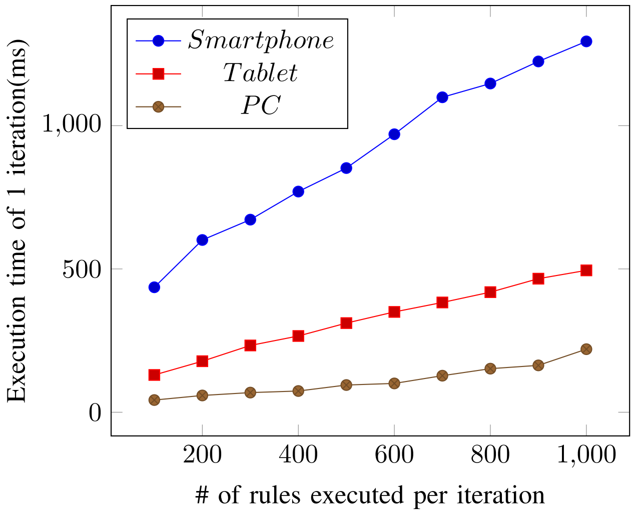

Figure 6. Total time of one iteration on all the devices when saving the rules at the end of the iteration

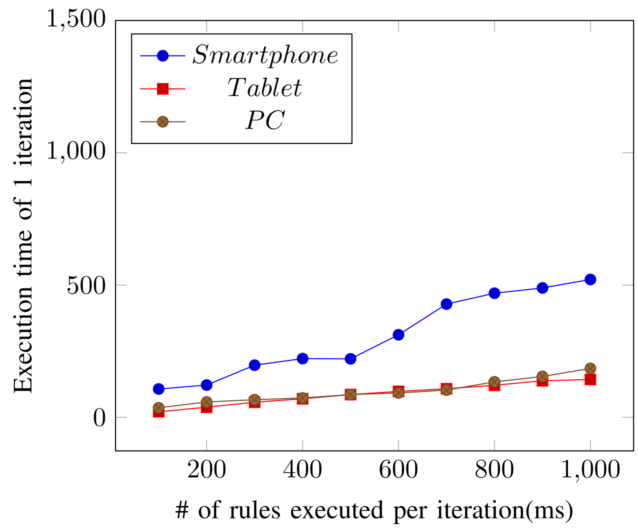

Figure 7. Iteration time on different devices without saving the rules at the end of the iteration

distribution of execution time without saving the rules in the database is increasing constantly, meaning that the duration time can be anticipated for various number of rules.

D. Battery Consumption

Mobile devices obtain the necessary energy for performing complex operations from their batteries ,which implies that analysing power consumption on these devices is very important. Our approach to calculate energy consumption was to measure the battery level on the smartphone using Android's internal system functions calls [37].

1) Benchmarks: We ran three types of benchmarks. The first one was designed for testing only the PEMD alone, which was executing sets of 100 Proactive Rules each 30 seconds. The Proactive Rule used in these tests was designed to simulated rules that would be used in different real-world applications. In the second benchmark, we simulated the interaction of a user with the screen of his/her smartphone by using a wakelock application that woke up the screen of the device, at 100% brightness, for a total of 18 minutes per hour. And, in the last benchmark, both the PEMD and the wakelock application were running simultaneously on the smartphone.

For all three benchmarks, 10 executions of 1 hour each were performed in order to compute their averages. During the tests the smartphone's Global Positioning System (GPS), Wireless fidelity (Wi-Fi), Bluetooth and the other mobile data connections were turned off.

2) Results: The results, shown in Table II, indicate first that executing Proactive Rules on a running PEMD, during 1 hour, takes only 1.5% of the total amount of the battery of the smartphone and second that a standard application consumes significantly more energy than a PEMD. The results of the third benchmark confirm the difference obtained between the first two benchmarks.

TABLE II. Average Battery Consumption

Applications	Average
`PEMD`	1.5%
`Wakelock app`	4.3%
`Wakelock app + PEMD`	5.4%

E. Discussions

Our tests were necessary for estimating the optimal number of Proactive Rules, which can be executed in one iteration by the Rules Engine on a smartphone and on a tablet, without having big delays. This aspect is very important when we want to design applications that can execute Proactive Rules on each device.

Our energy consumption analysis on the smartphone indicates that PEMDs do not take much battery consumption, which is a key aspect when developing models for smart applications.

In the future, the PEMDs should be able to allow multiple applications to execute concurrently Proactive Rules. This opens new perspectives and new challenges. If we take 5 applications with peak periods of approximately 250 rules instances for each we reach easily a number bigger than 1000 rules. And, if for now, running 1000 Proactive Rules on mobile devices does not raise performance issues, for bigger sets of Proactive Rules we could imagine slight problems. Solutions for this problem can be developed either by dividing rules into optimal sets of rules for running at each iteration or setting up a priority mechanism to distinguish between the crucial rules that need to be executed and the rules that can wait a couple of more iterations to be executed.

V. CONCLUSION AND FUTURE WORK

In this paper, we indicated the need of a rule-based engine for mobile devices capable of executing complex tasks like on the desktop computers. We also showed how few and limited possibilities are currently available for having rule-based systems on mobile devices. Then, we introduced a new

ISBN: 978-1-61208-436-7

model capable of enhancing mobile devices with Proactive Computing properties and with support for executing collaborative applications. Our experiments indicate that Proactive Engines can be successfully integrated into mobile devices, that the model is able to run on different mobile devices and that the processes of our model are very efficient from a computational point of view and do not affect the overall performance of a device. Moreover, the performance of the PEMDs on smartphones and tablets was analysed and compared to the performance of PEs on desktop computers. A solution was provided for improving significantly the execution time of Proactive Rules by saving them when shutdown events occur instead of saving them during each iteration of the PEMD.

More tests are to be completed, to check the duration of the communication operation when multiple devices are involved in this process. For instance, we would like to know how much time it would take for a message to arrive to its destination when there are hundreds or thousands of this kind of operations performed at the same time. Also, future tests will include measurements to check how a device handles multiple receiving operations at the same time, like, for example, when more than 100 messages are sent to the same device at the same time. Also, our next evaluations will include services like Pushy instead of GCM for exchanging messages between the Proactive Engines, in order to check if these frameworks affect the overall performance of the mobile devices.

Future work will include developing smart applications capable of collaborating together and of performing joint complex actions. Our current work includes the development of a version of an PE for mobile devices running on iOS, and, on developing modified version of the Proactive Engine for wearable devices, which are capable of performing complex tasks, like the new generation of smartwatches.

REFERENCES

[1] I. D. C. I. W. M. P. Tracker., "Worldwide smartphone shipments edge past 300 million units in the second quarter," Tech. Rep., [Retrieved: May, 2015]. [Online]. Available: http://www.idc.com/getdoc.jsp?containerId=prUS25037214

[2] D. James, "Recent advances in memory technology," in Advanced Semiconductor Manufacturing Conference (ASMC), 2013 24th Annual SEMI, May 2013, pp. 386–395.

[3] A. e. a. Wang, "10.3 heterogeneous multi-processing quad-core cpu and dual-gpu design for optimal performance, power, and thermal tradeoffs in a 28nm mobile application processor," in Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International, Feb 2014, pp. 180–181.

[4] N. Cacho, K. Damasceno, A. Garcia, T. Batista, F. Lopes, and C. Lucena, "Handling exceptional conditions in mobile collaborative applications: An exploratory case study," in Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE '06. 15th IEEE International Workshops on, June 2006, pp. 137–142.

[5] D. Tennenhouse, "Proactive computing," Communications of the ACM, vol. 43, no. 5, 2000, pp. 43–50.

[6] A. Salovaara and A. Oulasvirta, "Six modes of proactive resource management: a user-centric typology for proactive behaviors," in Proceedings of the third Nordic conference on Human-computer interaction. ACM, 2004, pp. 57–60.

[7] R.-A. Dobrican and D. Zampunieris, "Moving towards a distributed network of proactive, self-adaptive and context-aware systems," in ADAPTIVE 2014, The Sixth International Conference on Adaptive and Self-Adaptive Systems and Applications, 2014, pp. 22–26.

[8] D. Shirnin, S. Reis, and D. Zampunieris, "Experimentation of proactive computing in context aware systems: Case study of human-computer interactions in e-learning environment," in Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2013 IEEE International Multi-Disciplinary Conference on. IEEE, 2013, pp. 269–276.

[9] R. Dobrican, S. Reis, and D. Zampunieris, "Empirical investigations on community building and collaborative work inside a lms using proactive computing," in World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, vol. 2013, no. 1, 2013, pp. 1840–1852.

[10] S. Coronado and D. Zampunieris, "Continuous proactivity in learning management systems," in Education Engineering (EDUCON), 2010 IEEE. IEEE, 2010, pp. 199–204.

[11] S. Marques Dias, S. Reis, and D. Zampunieris, "Personalized, adaptive and intelligent support for online assignments based on proactive computing," in Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies, Rome, Italy 4-6 July, 2012. IEEE Computer Society Conference Publishing Services, 2012, pp. 668–669.

[12] D. Zampunieris, "Implementation of a proactive learning management system," in Proceedings of" E-Learn-World Conference on E-Learning in Corporate, Government, Healthcare & Higher Education", 2006, pp. 3145–3151.

[13] R.-A. Dobrican and D. Zampunieris, "Supporting collaborative learning inside communities of practice through proactive computing," in Proceedings of the 5th annual International Conference on Education and New Learning Technologies, Barcelona, Spain 1-3 July, 2013. IATED, 2013, pp. 5824–5833.

[14] J. Janeiro, T. Springer, and M. Endler, "A middleware service for coordinated adaptation of communication services in groups of devices," in Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on, March 2009, pp. 1–6.

[15] H. Chung, "Toward implementing a mobile collaborative system," in Systems and Informatics (ICSAI), 2012 International Conference on, May 2012, pp. 1248–1252.

[16] V. Sacramento, M. Endler, H. Rubinsztejn, L. Lima, K. Goncalves, F. Nascimento, and G. Bueno, "Moca: A middleware for developing collaborative applications for mobile users," Distributed Systems Online, IEEE, vol. 5, no. 10, Oct 2004, pp. 2–2.

[17] M. Caporuscio and P. Inverardi, "Yet another framework for supporting mobile and collaborative work," in Enabling Technologies: Infrastructure for Collaborative Enterprises, 2003. WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on. IEEE, 2003, pp. 81–86.

[18] "Rule-based system jess." [Retrieved: May, 2015]. [Online]. Available: http://www.jessrules.com/jess/

[19] "Rule-based system clips," [Retrieved: May, 2015]. [Online]. Available: http://clipsrules.sourceforge.net/

[20] C. L. Forgy, "Expert systems," P. G. Raeth, Ed. Los Alamitos, CA, USA: IEEE Computer Society Press, 1990, ch. Rete: A Fast Algorithm for the Many Pattern/Many Object Pattern Match Problem, pp. 324–341, [Retrieved: May, 2015]. [Online]. Available: http://dl.acm.org/citation.cfm?id=115710.115736

[21] R. Dobrican and D. Zampunieris, "A proactive approach for information sharing strategies in an environment of multiple connected ubiquitous devices," in Proceedings of the International Symposium on Ubiquitous Systems and Data Engineering (USDE 2014) in conjunction with 11th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2014). IEEE, 2014, pp. 763–771.

[22] S. Bobek, G. Nalepa, and M. layski, "Challenges for migration of rule-based reasoning engine to a mobile platform," in Multimedia Communications, Services and Security, ser. Communications in Computer and Information Science, A. Dziech and A. Czyewski, Eds. Springer International Publishing, 2014, vol. 429, pp. 43–57, [Retrieved: May, 2015]. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-07569-3_4

[23] C. Choi, I. Park, S. J. Hyun, D. Lee, and D. H. Sim, "Mire: A minimal rule engine for context-aware mobile devices," in Third IEEE International Conference on Digital Information Management (ICDIM), November 13-16, 2008, London, UK, Proceedings, 2008, pp. 172–177.

[24] "Ifttt (if this then that)," https://ifttt.com/, [Retrieved: May, 2015].

[25] N. Peers. Your online life made simpler, thanks to ifttt. [Retrieved: May,

2015]. [Online]. Available: http://blog.1and1.co.uk/2014/10/02/your-online-life-made-simpler-thanks-to-ifttt/

[26] G. J. Nalepa, "Architecture of the heart hybrid rule engine," in Proceedings of the 10th International Conference on Artifical Intelligence and Soft Computing: Part II, ser. ICAISC'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 598–605.

[27] S. Coronado and D. Zampunieris, "Towards a proactive learning management system using early activity detection," in SITE08-Society for Information Technology & Teacher Education International Conference. AACE, 2008, pp. 306–311.

[28] S. Marques Dias, S. Reis, and D. Zampunieris, "Proactive computing based implementation of personalized and adaptive technology enhanced learning over moodle (tm)," in Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies, Rome, Italy 4-6 July, 2012. IEEE Computer Society Publications, 2012, pp. 674–675.

[29] "Ormlite - lightweight object relational mapping (orm) java package." [Retrieved: May, 2015]. [Online]. Available: http://ormlite.com/

[30] "Parse push," [Retrieved: May, 2015]. [Online]. Available: https://parse.com/products/push

[31] "Pubnub global data stream network," [Retrieved: May, 2015]. [Online]. Available: https://www.pubnub.com/

[32] "Urban airship," [Retrieved: May, 2015]. [Online]. Available: http://urbanairship.com/

[33] "Reliable push notifications," [Retrieved: May, 2015]. [Online]. Available: https://pushy.me/

[34] E. Nava, "Pushy a new alternative to google cloud messaging," January 2015, [Retrieved: May, 2015]. [Online]. Available: http://eladnava.com/pushy-a-new-alternative-to-google-cloud-messaging/

[35] "Mysql. my structured query language database management system." [Retrieved: May, 2015]. [Online]. Available: http://www.mysql.com/

[36] "Hibernate orm. hibernate object/relation mapping framework." [Retrieved: May, 2015]. [Online]. Available: http://hibernate.org/

[37] "Battery management, android developers." [Retrieved: May, 2015]. [Online]. Available: http://developer.android.com/reference/android/os/BatteryManager.html

Instrumented analysis method for collaboration activities

Samia Ben Rajeb and Pierre Leclercq
University of Liege
LUCID Lab for User Cognition & Innovative Design
Liege, Belgium
e-mail: samia.benrajeb / pierre.leclercq @ ulg.ac.be

***Abstract*— Analysing collective design activities is a difficult task, especially in a context that involves the remote collaboration and/or multidisciplinarity. It is thus necessary to define a dedicated analysis process, instrumented by tools that can facilitate the data acquisition and visualization. The method presented here enables to cross-reference the two aspects of a complex collective activity: the process and the content treated by a group. Our method offers the possibility to analyse different types of collective work configurations (co-attendance or remote / instrumented or not). Its flexibility leaves the possibility to the researcher to update his frame and thus avoid the preconceptions earlier defined before the activity without possibility to reconsider.**

Keywords-collaborative design; methodologies and tools for collaborative activity analysis; visualization of collaborative processes.

I. Introduction

Collective activities have been the object of much research in psychology, ergonomics, and cognitive science that aim to create models for this kind of complex interaction [1]. These are based on two models of synchronisation: the first is cognitive synchronisation, relating to the construction of a context of shared knowledge; the other is operational synchronisation, relating to the division of tasks of the different collaborators [2]. These synchronisations aim to build awareness that will enable collaborators to interact with their environment and with a group of actors [3]. The place of the common ground is primordial because it contributes to the sharing of each one's specific competence and the acquisition of new competences to work in groups [4]. Other research has also highlighted the complexity of these activities because they can be different depending on the number of actors [5], the aim of the activity [6], the space and the time during which these interactions take place [7].

For the diversity of configurations involving several actors, supports (which can be specifications of tools as well as the organisation of on-going work groups) have been proposed in the context of the scientific field of Computer Supported Cooperative Work (CSCW) and, more precisely, of the Computer Support for Cooperative Design (CSCD, focussing on collective design activities).

In this context, the framework of our study only concerns the design meetings, held in co-presence or remotely, between different actors (architects, engineers, or designers) who are collaborating on the same project (and not coordinating, for example). This phase is that of the emergence of ideas that evolve corresponding to the interactions and graphic representations that are produced and shared. It is so much more difficult to observe and analyse when they take place at a distance. The analysis of this complex collaborative activity interests us because it raises several methodological questions for the researcher, in terms of the methods of collection, treatment and analysis of data, which we will develop in this article.

Our goal consists of designing an operational method to analyse the process of any complex collaborative activity, to code efficiently the gathered data and to assist in their analysis. As opposed to classical protocols and existing tools, we aim to use real time data collection (in addition to video support that usually requires from 3 to 10 times the duration of the observed activity), a flexible coding frame, which allows to adapt the coding variables and analysis criteria, and a agile visualisation tool, which can help the interpretation work, using dynamic graphs and diagrams.

This paper will first present a short state of the art on understanding collective activities (section II) and a description of our research context and our application framework (section III). The section IV will describe our process for analysing collaboration activities and the COMMON Tools support. Finally, the conclusions will emphasize the flexibility of our instrumented method and trace some perspectives.

II. State of the art and question of research

A. *Understanding collective activity and gathering data*

Since the 90's, many research projects have aimed to promote and aid collective activity. In a synthetic manner, one can distinguish [6][8]-[11]: those that try to categorise and define collective activity; those that concentrate on the technical aspects of this activity; those that focus on the social aspects; those that deal with developing man-machine interfaces and others man-man interfaces to help collaboration; those that develop methods and tools to analyse this complex activity in their real context or in the laboratory.

Focussing on this last aspect to understand collective activity from the research point of view, one of the main methods of gathering and treating data to analyse situations of collaboration is the "Protocol analysis" [12], which generally takes place in controlled environments. "Protocol analysis" is based on two methods of gathering data that can, separately, produce similar results for coherent

 ISBN: 978-1-61208-436-7

understanding of the problem-solving process and can also be complementary depending on the research objectives [13].

"Retrospective protocols" consist of asking the operator, after having finished his activity, to choose representative elements of his activity and then to describe them in order to better define the specificities of his work, alone or within a group. Thus, it concerns the study of design objects and their components independently of the situations in which they evolve [12]. In our opinion, this approach contributes to changing the designers' point of view of their design object by asking them to conceptualise their activity by calling on their memory. It has nevertheless been shown that if the stored information calls on short-term memory, the cumulative data can provide important details related to the research question [10] and the origins of choice to solve one problem or another [14]. Self-confrontation can also be another approach to analyse a task already completed. It consists of asking an operator to realise a self-examination of his own work process (alone and/or with others) from filmed sequences of his activity [15]. But, in our opinion, this other method demands that the operator make a big investment in terms of time and involvement in the research.

"Concurrent protocols" consist of the operator verbalising orally out loud his thoughts while working on a specific task (like the "think aloud" [16]). His thoughts are then transcribed, coded, and analysed by the researcher. This approach comes from the hypothesis that the verbalisation of thoughts during the process of problem-solving does not affect the process [12]. Other researchers do not agree with this hypothesis and think that the "retrospective protocols" are less intrusive in the process since it is put into play once the activity is finished [17]. But, in the context of collective activity, the actors naturally find themselves in the obligation of speaking and verbalising their thoughts in order to collaborate. In this way, the "concurrent protocols" make sense, in our opinion, because it is closer to the real work conditions and the context. Taking into account the context where the activity takes place reinforces the ecological validity of the observations and does not exclude the social process, the team work and the communication which make up the daily work.

B. *Choose an approach and define its analytical steps*

The methodological approach for the analysis and treatment of data are all the more varied and can result in qualitative or quantitative results. To reach a certain degree of precision in data processing, these methods are generally based on a segmentation system that, according [14], can be slanted according to two approaches.

• "Process-oriented segmentation": this approach cuts the process into several sequences relative to the actors' intentions and identifies the time spent for each of these sequences, as well as the correlation between them. According to [18], the COMET method for example [11], allows one to describe the principal identification phases and argumentation of a problem. As for the coding table developed for the specific analysis of the comparison of the points of view in concurrent engineering, it enables one to draw up a tree diagram of propositions and verbal interactions between the collaborating actors [19]. The analysis by the word-processing software ALCESTE [20] enables, moreover, to structure the information put into play and shared by the actors to solve a problem. Even though all these methods are complementary, this "Process-oriented segmentation" approach is sometimes criticized because it does not look closely enough at the contents, that is, the problem treated by the actors during the activity, the documents and annotations that are produced [21].

• "Content-oriented segmentation": this approach enables one to complete the first as it looks specifically at the visual contents (representations, annotations, references, artifacts, etc.) and examines the cognitive interactions between the designers and the artifacts [15]. One of the best-known methods is that of Gero [22], which is based on a principle of encoding, called FBS, depending on the functionality of the object ("Function"), the behaviour of the actors ("Behavior"), and the structure of the collaboration ("Structure"). In this way, the author formulates the design as a series of transformations of the model's functions. Brassac & Gregori [23] propose a clinical approach that looks at the real activity and its different interactions, by studying the discursive productions, the gestures, the graphical representations and the conversational sequence. In our opinion, this approach enables one not only to hierarchize the acts of language, by breaking them down into sequences and sub-sequences, but also to illustrate the conversational dynamic between the collaborators [23]. In the same way, being based on ethnographical studies, Boujut & Laureillard introduce themselves directly into real industrial context and propose methods of "research-action", analysing this framework and introducing new tools to aid the collaboration [24].

Faced with this variety of methods, our approach is clearly placed in the "concurrent protocols" that look at the process, both at the evolution of the process and in the time ("process-oriented segmentation"), and at the different interactions between the actors, as well as the design project that is treated ("context-oriented segmentation").

The difficulty lies in the context of our research. In fact, our analyses concern the collective activities (in co-presence and/or remotely) that take place: (1) either in a professional context that does not allow, for reasons of confidentiality, to gather the audio data and/or videos to treat and analyse them afterwards, (2) or in a pedagogical context of project realisation over an entire semester involving students, teachers, and experts (substantial, complex and difficult-to-analyse data). In our opinion, it is a question of studying the ensemble of these interactions (oral and graphical) with the objective of describing the process of negotiation and making collective decisions. This description takes places, in our method, in a qualitative manner and is also supported by the visualisation of the quantitative data looking into the many criteria that play a role in the specification of the collective activity.

 ISBN: 978-1-61208-436-7

III. RESEARCH CONTEXT AND APPLICATION FRAMEWORK

A. *Research context*

This research project fits into the framework of the ARC (Actions de Recherche concertée) program. This program is financed by the Walloon-Bruxelles Community and involves the multidisciplinary consortium COMMON (Natural Multimodal Mediatised Collaboration), which groups about fifteen researchers from five departments of the University of Liège: engineering sciences (LUCID, Lab for User Cognition and Innovative Design, that coordinated the scientific program), linguistics and semiotics (Science of Language and Rhetoric), work psychology, and cognitive sciences (LECIT, Laboratory of Cognitive Ergonomics and Intervention in Work), architecture (Architecture and Society) and medicine (Systematic Human Anatomy) [25]. Being spread over four years (2011-15) the objective of this research project focussed on the analysis of multimodal characteristics of collaboration and verbal and non-verbal exchanges in complex activities. To answer this question, the consortium put in place a method of analysis of collaborative practices that are presented in this article based on multiple observations and real practices, articulating quantitative with qualitative ones.

B. *Application Framework*

We have applied this method to analyse different configurations involving co-presence and remote meetings that bring together varied actors in "training by projects" contexts and in professional contexts (cf. Figure 1).

Figure 1. Examined configurations: freehand collaboration vs instrumented work / remote collboration vs co-attendance meeting.

Focussing on complex collective activities and, more specifically, on collaborative design activities, we have examined fields of architecture as well as engineering, design and ergonomics. The variety of analyses on the quantity of gathered data were privileged because it is difficult to observe complex collective activities in detail, even taking place in a limited lapse of time. Our gathered data vary between four hours in professional contexts to several months in the "training by projects" framework. We have also privileged situations grouping a limited number of participants (between 2 and 5 participants) to get a finer understanding of the activity. Focussing on the integration of new technology being used in these collective activities, certain analysis (mostly the remote ones) involved the use of an innovative system called Collaborative Digital Studio. This system associates 3 elements: video-conferencing (enabling geographically distant collaborators to see each other and to discuss remotely in real time), a digital table using an electronic pen (by which collaborators can interact graphically via an electronic pen), and a graphic interaction software called SketSha [26]. Developed in the LUCID Laboratory of the University of Liège, this allows remote collaborators to share documents (sketches, plans, pictures, technical drawings and texts) and to interact graphically in real time [27].

IV. ANALYSING DATA PROCESS

Fitting into the "concurrent protocols" method, we have tried to grasp the particularities of collective activity put to work in real social contexts to try to help them. To do this, we defined a group of criteria put forward in our state of art and that focus as well on the project as the design object itself. The criteria we examined here concern: the actions put into play by each actor, their typology and their evolution over time, the work spaces involved and the passage from one to another, the documents used and the kinds of annotations produced, the evolution of the shared design object in terms of the degree of abstraction and the degree of grasping. These criteria are important to define before beginning because they contribute to fixing the observation protocol and gathering data as well as the treatment and its analysis that takes into account: (1) time, (2) the role of each actor and his work space, and (3) the implications of these interactions in the evolution of the object to be designed.

A. *Observation*

To carry out the data gathering during an activity, two methods can be put to work. The first concerns the video recording that was applied in the "training-by-projects". These videos were captured according to two focal lengths: a wide angle (centered process) to film all the scene of the interaction between the actors and a narrow angle (content-centered) on the work surface to film all the artifacts and annotations put into play during collaborative work.

The second, which demands more preparation before the observation, concerns rapid note-taking. To do this, observers that were trained in this method in advance (between 3 and 4 researchers per situation) receive different observation methods to which are attached pre-constructed tables according to the theme of the data:

• Theme 1. Observe the collaboration: list established according to time landmarks, interactions of designers and their work spaces (I-space, We-space, Space-between [28]), documents used and representations realised during the process; counting possible emotions explicitly expressed.

• Theme 2. Observe the design: following the design process and the artifacts that are created or shared by the designers (parts of the project concerned by each action, documents used and/or created) and listing analogies and references put to work.

• Theme 3. Observe freely: qualitative tracking, always in relation to time, of the evolution of the object conceived and negotiated in the group, tracking key moments and particular uses of tools used during the design process.

This list of chronological actions and their modalities can be completed in real time by the observers, thanks to this rapid note-taking form.

B. Treatment

During this phase, the temporal point of reference takes on its importance. All gathered data are first synchronized then coded according to the criteria cited above from the note-taking. A common description of the collaborative process is then constructed in the form of actions based on a consensus between the different observers. By putting each action into words, the observers cut the activity into moments of interactions that they then code in relation to the categories defining the collaborative process and that of design. This division is made via a coding frame (cf. Figure 2).

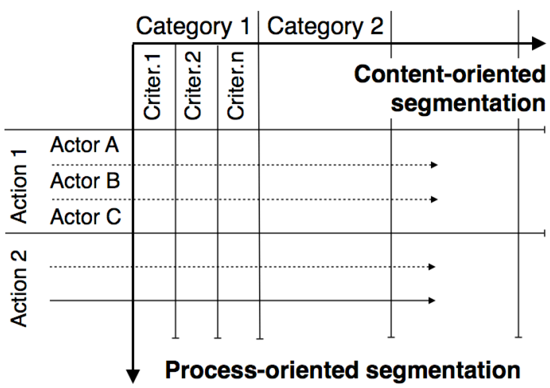

Figure 2. Process-oriented vs content-oriented segmentations.

This frame divides the activity vertically, according to the temporal point of reference in order to describe the process ("process-oriented segmentation"), and horizontally following the predefined categories of analysis in order to describe the content and the specificity of each action ("content-oriented segmentation") defined according to the following categories: type of action, concerned space of work, documents used, representations created, degree of grasping and degree of the object’s abstraction, manifested emotions, etc. The process-oriented segmentation and the data of the content-oriented segmentation are written in a differed time frame, after the observation phase, by mixing the points of view of observers and observed collaborators, in one common and unique coding Excel file.

C. COMMON Tools

After their synchronization and coding, the data is then treated in COMMON Tools (CT). CT is a web platform initiated in the framework of the ARC COMMON project and developed by LUCID of the University of Liège. It was made available to the researchers enabling the transformation of the data from the coding frame (in the .csv or .xls files) into consolidated data then quantified and translated according to different choices of visual formalisms (pie, stacked columns, time line, crossing, clouds, etc. This tool offers researchers a tool for visualizing data in order to analyze the collective design activity in the form of a panel of interactive graphics (generating multiple graphs per analysis-type). It enables one to visualize the crossing of data treated in relation to time, occurrences and the specificities of each actor involved in the collective design process (cf Figure 3).

By comparison to other usual visualisation tools (like Excel for example), CT offers a crossed data representation, which allows to observe concomitance of two variables (cf. Figure 4) and a of timeline representation, with dynamic functions to zoom in a particular duration of time along the observed process (cf. Figure 5).

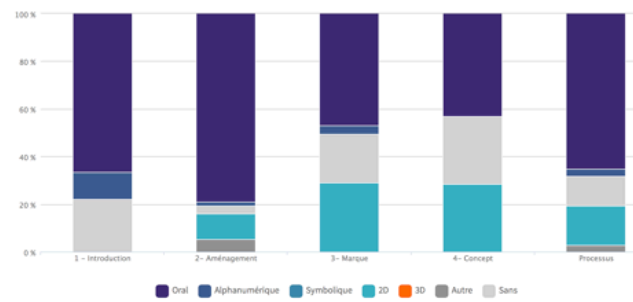

Figure 3. Example of visualizations proposed by the COMMON Tools : types of used representations in five steps of the observed process : oral / written / symbolic / 2D / 3D / others / none.

	Intention	Topol.	Geometry	Charact.	None
Context	0	0	0	1	0
All	5	0	2.33	0.67	0
Part	0.67	5.33	6.67	0	0.33
Detail	14.33	0.67	2	1	0
None	1.33	0.33	0	0	8

Figure 4. Example of visualizations proposed by the COMMON Tools : collective object characterization (crossing level of abstraction / level of grasping)

Let's give two result examples that has been brought thanks to this analysing process using the COMMON Tools.

It has been possible for us to show the importance of space management in the collaborative process between the designers. In fact, it has been shown that the group cohesion is favored by the enhanced spaces offered by the Collaborative Digital Studio, creating intermediary spatiality between the enhanced presence and virtual co-presence. According to our analysis, these augmented spaces participate in helping and giving tools for learning to collaborate for students on one hand. They favor, on the other hand, sharing depending on the empowerment and the creation of private conversations (aparté) [29].

Moreover, the contribution of these augmented spaces has been demonstrated in the production of the group. In

 ISBN: 978-1-61208-436-7

fact, the Digital Collaborative Studio enables remote sharing of different artifacts in real time thus favoring collective production and, in certain observed cases, to realize drawings by two hands in an instantaneous and synchronized manner. These augmented spaces interfere also on the production operations the and interpretation of a drawing thus creating new manners of construction of the shared artifact in pairs.

D. *Analysis*

The analysis that is proposed here focusses on the process as well as the contents, by describing the evolution, in time, of the interactions of the actors and their implications in the common design object. It integrates the relevant descriptive dimensions already released in a qualitative manner during the transcript (communicational strategies, kinds of sequences, forms of collaboration, corporal communication, relational evolution, etc.). This qualitative point of view is then enriched by the interpretation of quantitative visualizations, offered by the COMMON Tools.

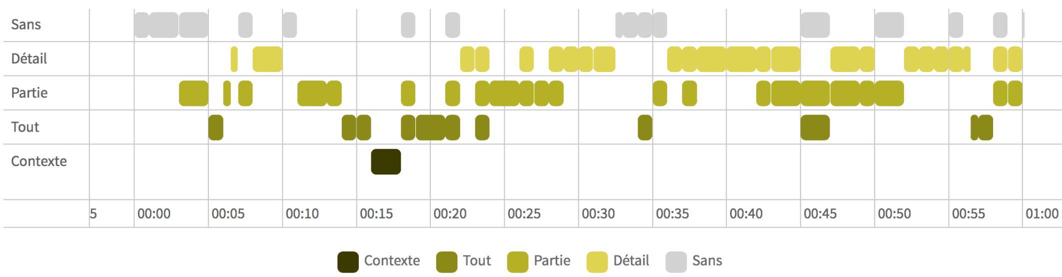

Figure 5. Example of visualization proposed by the COMMON Tools. Time line of the design object : context / whole / part / detail / none.

V. Conclusion, Limits and Perspectives

The method presented here and summarized in Figure 6, enables us to cross-reference the two aspects of a complex collective activity: the process and the content treated by a group. It focusses on the specificity of each actor, his work space, his documents as well as the interactions with other collaborators.

The direct observation method, without the possibility to record video data, moreover enables one to rationalize the rapid note-taking process. This procedure is not as complicated as the treatment of the verbalization but it does not produce only the qualitative observations. With the COMMON Tools, the researcher also has quick and easy access to the graphs during his analysis, with the support of a diversity of visual formalisms among, which he can interactively choose those which prove to be the most pertinent to be useful for his research question.

Contrary to other systems (cf. Section II.B), the flexibility of our method offers the possibility to renew, to call into question and/or add categories during the transcript and coding. This flexibility enables, on one hand, to analyze a substantial corpus of diverse configurations of collective activities involving several actors, and, on the other hand, to leave the possibility to the researcher to update his frame and thus avoid the preconceptions earlier defined before the activity without possibility to reconsider.

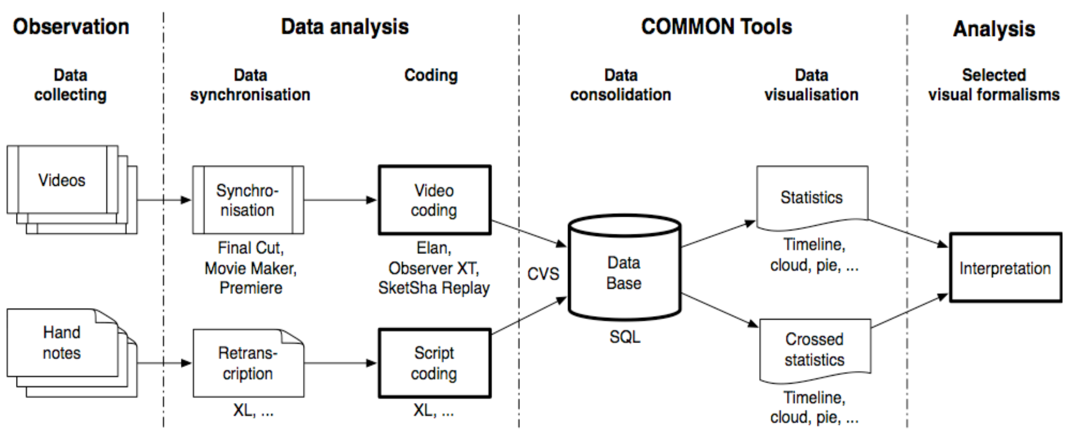

Figure 6. Process for analysing collaboration activities.

The main limit of the "concurrent protocols" is that the observed actors know that they are being observed, and that interferes with their way to work and interact together [14]. That is why it is important that the observers must be perfectly trained, capable and motivated. Thus, it is primordial to prepare them well and, more specifically, in the context where rapid note-taking is essential. In such a context, we recommend a first phase where the observers take note of what happens during the first half hour of observation, according to very specific missions (cf. Section IV.A). Afterwards, the process is temporally suspended: all the observers meet in an isolated room, far from the observed site, to discuss for a few minutes their difficulties in observing. This phase allows them to stabilize and coordinate their strategies in order to start again, afterwards, their note-taking in a more coherent way and better adapted to the observed context.

A second difficulty was raised at the level of data processing (cf. Section IV.B). This difficulty concerns the choice of different criteria corresponding to each action treated. It is sometimes difficult to categorize each action in an exclusive and definitive manner. Nevertheless, the proposed frame makes it possible to cut the actions into sub-actions (vertically, in relation to time). It also offers the possibility to cross two categories (horizontally coded) and, thus, to clearly specify the links between one criterion and another. This flexibility and crossing are necessary to prevent beforehand interpretation by the analyst, who risks to make shortcuts in the conclusions or to slant the coding according to his own preconceptions.

A final difficulty is to be emphasized concerning the multiplication of proposed graphs by the COMMON Tools for visualizing the data (cf. Section IV.C). This multiplication enriches the analyses but makes the job of interpretation more difficult to organize. In fact, it is important to keep this flexible aspect, at the level of choosing the formalism, as well as at the level of the variety of criteria to be crossed. It is nevertheless contradictory to think that simple statistics done automatically by a tool could make sense by themselves. The method put forward in this article, leaning on the COMMON Tools, above all, enables one to build a first quantitative structure of observations to get one's bearings in the qualitative analysis of complex collective activity. It does not pretend to lead directly to interpretations and activity shortcuts by these quantitative data. It orchestrates and facilitates the work of interpretation and thus, enables the researcher/analyst to quantitatively confirm or reject hypotheses made during observations qualitatively ahead of the treated corpus.

ACKNOWLEDGMENT

This research is financed thanks to the grant Research Concerted Actions (ARC COMMON 2011-2015), the Scientific Research Direction of Education, French and Belgian Community, granted by the Wallonie-Europe Academy.

REFERENCES

[1] J. Karen, J. Ostergaard, and J. D. Summers, "Development of a systematic classification and taxonomy of collaborative design activities", in Journal of Engineering Design, 20:1, pp. 57-81, 2009.

[2] F. Darses, P. Falzon, and P. Béguin, "Collective design processes", in Proceeding of COOP'96, Second International Conference on the Design of Cooperative Systems. Juan-les-Pins, Sophia-Antipolis, INRIA, 1996.

[3] J. M. Carroll, D. C. Neale, P.L. Isenhour, M. B. Rosson, and D.S. McCrickard, "Notification and awareness: synchronizing task-oriented collaborative activity," in International Journal Of Human-Computer Studies, 58, pp. 605-632, 2003.

[4] J. S. Gero and U. Kannengiesser, "The situated function-behaviour-structure framework," in Design Studies, 25 (4), pp. 373-391, 2004.

[5] P. A. Gloor, "Swarm Creativity: Competitive Advantage through Collaborative Innovation Networks," Oxford University Press, Oxford, New York, 2006.

[6] Y. E. Kalay, "Architecture's New Media. Principles, Theories, and Methods of Computer-Aided Design," in The MIT Press: Communication, MIT Press, Cambridge, pp. 83-198, 2004.

[7] W. W. Gaver, "The affordances of media spaces for collaboration," in Proceedings of CSCW'92, ACM, New York, 1992.

[8] P. Dillenbourg, et al., "The socio-cognitive functions of community mirrors," in Proceedings of the 4th International Conference on New Educational Environments, Lugano, 2002.

[9] P. Leclercq, "Going collaborative," in Proceedings of 5th International conference CDVE Computer Aided Architectural Design: Experience Insight and Challenges, Calvià, Mallorca, 2008.

[10] J. S. Gero and H. H. Tang, "The differences between retrospective and concurrent protocols in revealing the process-oriented aspecs of the design process," in Design Studies, vol. 21, no. 3, pp. 283-95, 2001.

[11] F. Darses, P. Falzon, and C. Mondutéguy, "Paradigmes et modèles pour l'analyse cognitive des activités finalisées," in Ergonomie, Presses Universitaires de France, Paris, pp. 191-212, 2004.

[12] K. A. Ericsson and H. A. Simon, "Protocol analysis: Verbal reports as data," in MIT Press, Cambridge, MA, 1993.

[13] H. Kuusela and P. Pallab, "A comparison of concurrent and retrospective verbal protocol analysis," in American Journal of Psychology, vol. 113, no. 3, pp. 387-404, 2000.

[14] L. Nguyen and G. Shanks, "A framework for understanding creativity in requirements engineering," in Information and Software Technology 51, pp. 655–662, 2009.

[15] M. Suwa, T. Purcell, and J. Gero, "Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions," in Design Studies, vol. 19 pp. 455-83, 1998.

[16] C. H. Lewis, "Using the "Thinking Aloud" Method In Cognitive Interface Design," in Technical report, IBM. RC-9265, 1982.

[17] P. Lloyd, B. Lawson, and P. Scott, "Can concurrent verbalisation reveal design cognition?," in Design Studies, vol. 16, pp. 237-59, 1995.

[18] G. Gronier, "Méthodes d'analyse des communications fonctionnelles en situation de travail collectif," in Recherches Qualitatives, 9, 153-171, 2010.

[19] F. Détienne, G. Martin, and E. Lavigne, "Viewpoints in co-design: a field study in concurrent engineering," in Design Studies, 26, pp. 215-241, 2005.

[20] M. Reinert, “Alceste, une méthode statistique et sémiotique d’analyse de discours; application aux Rêveries du promeneur solitaire,” in La Revue française de psychiatrie et de psychologie médicale, 49(5), pp. 32-46, 2001.

[21] K. Dorst and J. Dijkhuis, “Comparing paradigms for describing design activity,” in Design Studies 16, pp. 261-274, 1995.

[22] J. S. Gero and T. M. Neill, “An approach to the analysis of design protocols,” in Design Studies, vol. 19, pp. 21-61, 1998.

[23] C. Brassac and N. Gregori, “Co-construction de sens en situation de conception d’un outil didactique,” in Studia Romanic Posnaniensia, no 25/26, pp. 55-66, 2000.

[24] J. F. Boujut and P. Laureillard, “A co-operation framework for product–process integration in engineering design,” in Design Studies 23 (6), pp. 497–513, 2002.

[25] www.lucid.ulg.ac.be/www/research/common/ [retrieved: July, 2015].

[26] www.sketsha.be [retrieved: July, 2015].

[27] S. Safin, V. Delfosse, and P. Leclercq, “Mixed-reality prototypes to support early creative design,” in The Engineering of Mixed Reality Systems, Springer, London, pp. 419-445, 2010.

[28] S. Ben Rajeb and P. Leclercq, “Using Spatial Augmented Reality in Synchronous Collaborative Design,” in Lecture Notes in Computer Sciences, Springer, vol. 8091, pp. 6-15, 2013.

[29] S. Ben Rajeb and P. Leclercq, “ Spatial augmented reality in collaborative design training: articulation between I-space, We-space and Space-between”, in Lecture Notes in Computer Sciences, Springer, LNCS 8526, vol 2, pp. 343-353, 2015.

Online Learning Community Software to Support Success in Project Teams

Brian Thoms
Department of Computer Science and Information Technology
California State University, Channel Islands
California, United States
Brian.Thoms@CSUCI.edu

Evren Eryilmaz
Department of Business Education and Information and Technology Management
Bloomsburg University
Pennsylvania, United States
EEryilma@BLOOMU.edu

***Abstract*—In this research we explore aspects of social interaction and community as they relate to success in project-based courses. Using specialized online community software consisting of social networking technologies and project-based wikis, project teams are able to collaborate and interact as they progress towards project milestones. Our study underscores the importance of sustained engagement as a means for fostering high levels of community and how these levels relate to project motivation and, ultimately, project success. Guided by a theoretical model that explains how individuals collaborate within online communities, we measure member perceptions of the software before and after our intervention. Survey results found that online learning community (OLC) software can successfully support learning and social interaction. These results are supported by a social network analysis (SNA), which shows high levels of individual engagement across the project lifecycle.**

***Keywords*—Social Networking; Online Learning Community; Wiki; Project Management; Capstone Project**

I. INTRODUCTION

It is easy to take the World Wide Web for granted, which turned 25 in March of 2014. And it is easy to forget that, for most of the 21st century, electronic peer-to-peer (P2P) communication occurred through postal mail and land-based telephones. This was true for the average person as well as individuals pursuing advanced education. Today, while these forms of communication are pervasive as ever, they are no longer the primary means of P2P for certain demographics, such as digital natives, or those individuals having grown up with internet technologies.

It is estimated that 86% of digital natives participate in some form of online social networking (OSN), with some estimates as high as 98% [1] [2] [3] [4].

In this research, we leverage this population's technical prowess for social media to implement online learning community (OLC) software, which increases interaction, enhances levels of community and supports learning. More specifically, we implement OLC software in senior-level information technology (IT) capstone courses, which require students to work in project teams and construct a final IT artifact. We measure our interventions through survey research and a social network analysis (SNA) to discover how OLC software, comprised primarily of OSN technologies, provides students with an enhanced project space; one that fosters higher levels of interaction and learning and strengthens course community.

The rest of this paper is structured as follows. In Section II we establish the conceptual background for academic communities emphasizing the role online community software plays within project-based course communities. Section III provides a theoretical framework that bridges constructivism, engagement theory and social presence theory. Section IV focuses on the design of the OLC software. Section V highlights the research methodology. Section VI describes the results of this research. Section VII provides comprehensive discussion and analysis of our results section. Section VIII identifies the limitations of this study. Section IX is the conclusion section followed by references in Section X.

II. BACKGROUND

A. *Academic Communities*

When individuals enter college, they join the college conversation, and intrinsically become a part of the academic community. Academic communities can be classified as one subset of what Lave and Wenger [5] have coined communities of practice. In such communities, individuals work together towards common goals, collaborate on common problems, share best practices, support one another and share in a common identity. Thus, successful academic communities, as suggested by Adams and Freeman [6], help to sustain engagement and collaboration among individuals whereby knowledge building becomes an intrinsic function of the community itself.

B. *Project-based Course Community*

Courses can be considered a more specialized form of academic communities [7]. In this research, we focus our attention on specific types of courses, capstone courses and, in our case, team-based project courses.

Within the IT industry, team-based projects are recognized as a core component of effective undergraduate education [8] [9]. The inclusion of team-based projects into the undergraduate experience is largely influenced by industry expectations that graduates exhibit high-levels of problem solving, oral and written communication, teamwork and project management skills [10] [11].

One way for students to prepare to meet these expectations is through capstone projects. Capstone projects-based courses are valuable ways for students to prepare for

 ISBN: 978-1-61208-436-7

careers in their respective industries. While there are many approaches to this type of experiential learning, one approach considers individuals working within project teams, which allows individuals to build critical team-based skills and learn how cooperation and group dynamics plays-out as they work towards the completion project milestones. This notion is supported in Lainez et al. [12], who state that capstone projects in IT should deliver important skills such as 1) a basic understanding of business processes, 2) a product development with high-quality concerns, 3) know-how to conceive, design, implement and operate medium-size complexity systems and 4) communicative, initiative/leadership teamwork, analytical and problem solving and personal abilities. Furthermore, Ayas and Zeniuk [13] state that learning within project teams can be instrumental in building communities of reflective practitioners.

And these experiences are well received by students. Dunlap [14] discovered that engaging students in learning and problem-solving activities reflects the true nature and requirements of workplace communities and help students feel better prepared to work effectively in their profession, a viewpoint is supported by students [11]. Furthermore, Clarke [15] identified that industry-aligned projects increased student confidence and allowed students to explore areas of IT not covered in the curriculum.

Ultimately, within IT capstone courses, students are presented opportunities to consolidate their understandings of “systems analysis, software development lifecycles, specific software design support tools, entity relationship modelling, entity life histories, database design, web site design, or web server programming” [16]. Furthermore, when students engage in experiential learning, they become active participants in the learning process, constructing their own internal knowledge through both personal and environmental experiences [16] [17]. Lynch et al. [19] found that capstone projects provide students the opportunity to build, not only technical skills of the discipline, but the social aspects of systems development as well.

As a means to facilitate project communities and senior capstone courses, we incorporate an online learning community that allows students to participate across multiple dimensions of the project lifecycle, while also working towards project milestones.

C. Online Learning Community Software

Thoms et al. [7] [19] [20] argue that OLC software, constructed from underlying OSN technologies, offers benefits over traditional learning management software (LMS) within higher education. As touched upon earlier, project-based courses can be seen as niche types of communities. In such communities, individuals take equal ownership in content production and work towards developing a unique voice.

OSN technologies include a large array of Web 2.0 technologies such as asynchronous online discussion boards, blogs and wikis along with peer-to-peer networking and file sharing to name a few. These tools empower individuals to take ownership of their content while also making it easier to communicate and interact with other members of the community. Thus, constructing an OLC that more closely resembles OSN environments makes sense, since the overwhelming majority of individuals within higher education (i.e. digital natives) are already competent with these technologies. With roots firmly planted in OSN technologies, OLC software offers the greatest potential for facilitating a communal space, while also facilitating knowledge construction and learning in a higher education setting [7] [19] [20].

III. Theory

To help guide the design and construction of an OLC to support our project-based needs, we follow a holistic theoretical model represented in Figure 1. The model considers, at its core, 1) the individual, 2) how individuals engage in online media through technology and 3) the overall sense of online community that results from active participants.

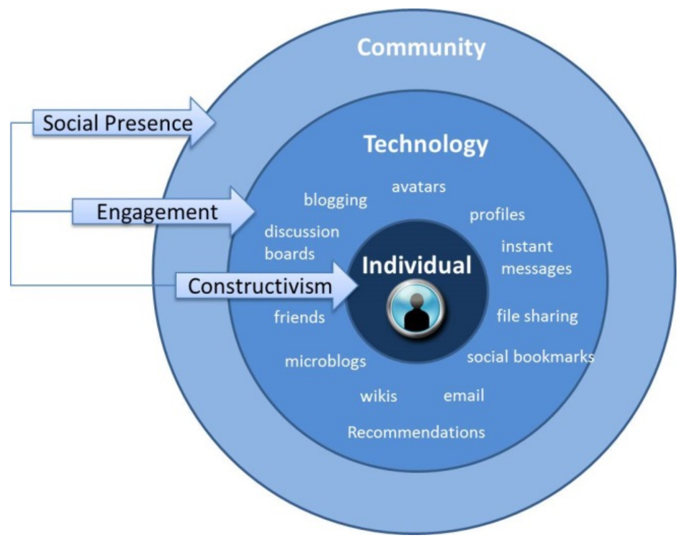

Figure 1. Theoretical Model for Online Learning Communities

A. Individuals (Constructivism)

Prior research has traced the roots of online communities to constructivism [21] [22] [23]. Consequently, at the center of our model we place the individual. Constructivism states that learning stems from the interactions and experiences of the learner [24] [25]. We believe that these interactions and experiences can be directly influenced by a user’s engagement with OSN technologies. OSN technologies can be configured to facilitate many different types of participants. Largely linked to the work of Piaget [25], who first theorized that learning can be based on the interaction and experiences of the learner within a specific context, constructivism provides a holistic view of individual learning and how individuals interact within larger groups. Hagstrom and Wertsch [24] state that constructivism encourages, utilizes, and rewards the unique and multidimensional characteristics of the individual throughout the learning process. Additionally, Squires [26] states that constructivism

 ISBN: 978-1-61208-436-7

focuses on individual control, with individuals making decisions that match their own needs. While constructivism began as a theory of learning, it has progressively been used as a theory of education, of the origin of ideas, and of both personal knowledge and scientific knowledge [27]. Specific to this research, our OLC is comprised of individuals, who occupy a shared space, whereby she or he is influenced heavily by interactions with technology and others. Consequently, it is not unlikely that individuals will experience the community in different fashions.

B. *Technology (Engagement Theory)*

Research has linked student engagement to grades and motivation [28]. Thus, getting participants engaged in course project objectives early on can be tantamount to the project's success. Engagement theory states that individuals must be meaningfully engaged in project activities through interaction with others, which can be facilitated and enabled through specialized technologies [29]. Dalsgaard [30], whose research is supported in Waycott et al. [31], argues that social software can be used to support the constructivist approach set forth in the previous section. Social software engenders a cooperative approach to learning and work towards the establishment of a cohesive community. In this respect, social software can refer to any loosely connected application where individuals are able to communicate with one another, and track discussions across the internet [32]. Consequently, the development of OLC software must consider the individual's point of view in such a manner that they are provided with a certain level of control and autonomy within the community. Once again, social software supports these philosophies and makes participants the locus of control within a self-governing environment.

C. *Community (Social Presence)*

We introduce Social Presence Theory to understand the manifestation of our OLC as a dynamic and vibrant collaborative project space. Social Presence Theory looks at the degree to which an individual's perception of the online community, affects his or her participation [33] [34]. When an individual believes that others are interacting and exchanging information, that individual may be more inclined to engage themselves. As discussed in Garrison et al. [35], alternative methods for enhancing social presence must be explored to help substitute for the lack of visual cues individuals receive in face-to-face settings. Research by Richardson and Swan [36] identified that a student's perceived level of social presence directly relates to their perceived learning. This suggests that increasing levels of community can yield higher levels of learning. OSN technologies work well in this regard and have successfully helped enhance social presence through peer feedback [37] and individual profiles and avatars [38], both of which are implemented within the OLC designs we investigated. Additionally, Thoms [39] discovered that OSN technologies can foster higher levels of course learning through openness and collaboration and can align very well with course learning objectives.

Together, these three theories provide a holistic model that considers course community, individual learning styles and how each can be influenced and enhanced with technology.

IV. Community Software Design

Just prior to the Web 2.0 revolution, Preece [40] stated that OSN developers can control the design of OSN software, but it remains difficult to control social interaction across the software. This statement implies that not all social technologies will yield the desired level of interaction intended by their design. After the Web 2.0 explosion, an influx of new social software presented software developers with a treasure-trove of open-source plug-and-play technologies, which have now become staples of popular OSNs. These technologies include filesharing, blogging, status updates, tagging, social bookmarking, and individual and community profile building.

Elgg is an OSN engine and is currently used as the primary engine across numerous online communities. The software provides a wealth of social technologies and has an easy-to-use and customizable interface, which can mirror the look and feel of any organization. A default installation of Elgg provides features such as discussion boards, blogs, file galleries and peer-to-peer (P2P) networking capabilities. Elgg also offers the ability to create sub-communities, a crucial feature that allows academic instructors to implement multiple communities within a single Elgg installation. Sub-communities allow instructors to separate courses, while allowing students to interact with peers outside their respective course community. Additionally, Elgg allows restrictions across multiple levels, including individual-level, community-level and logged-in-user-level restrictions, making it a choice system for project-based communities, who may wish to limit the availability of project information to the larger course community.

Illustrated in Figure 2 is the user landing page for a typical Elgg community, in this case, the community is the capstone course homepage. Within this environment, members are also able to add customizable modules to the homepage that presents them with real-time community activity. By default, users are presented with active content from across the site, which can be filtered by user or date. Rather than be reactive, our OLC is proactive, and a greater emphasis is placed on content creation, content dissemination and user-interaction.

 ISBN: 978-1-61208-436-7

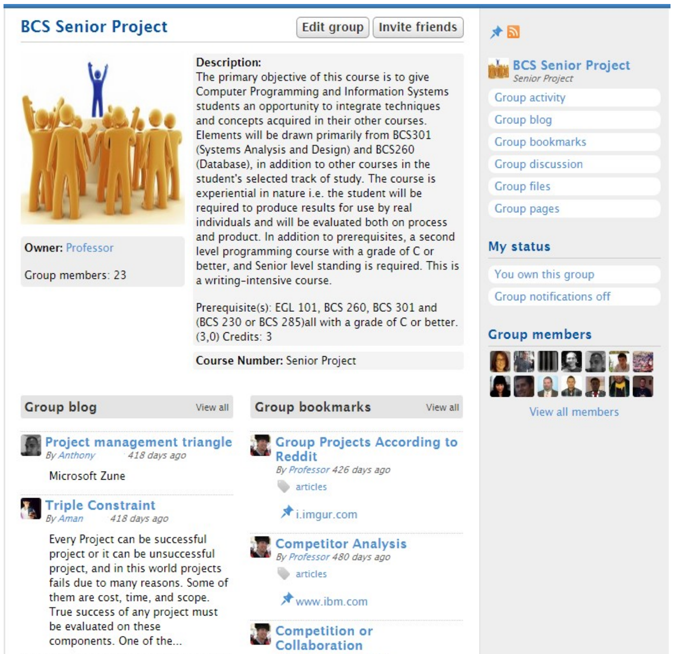

Figure 2. OLC Landing Page

A. *Collaborative Writing Software*

The primary artifact within a project-based OLC centers on the analysis and design of project objectives. As a shared artifact, collaborative writing software functions as a mechanism to support information sharing and group knowledge construction. As one subset of collaborative writing, wiki software utilizes Internet-based technologies to facilitate the collaborative writing process by keeping track of page creation and page edits. Not only are wikis an effective mechanism for obtaining information and knowledge, such as with Wikipedia, the world's largest encyclopedia, they are also an effective technology for facilitating virtual collaboration. Wikis provide a shared dialogue and centralize information among collaborators in group projects. Additionally, wikis allow members to engage in group learning and share in knowledge construction within a virtual community [41]. These notions are paramount for project teams working collaboratively towards shared goals and shared understandings. Additionally, a wiki can provide project teams with a level of coordination and synchronization not afforded by other means.

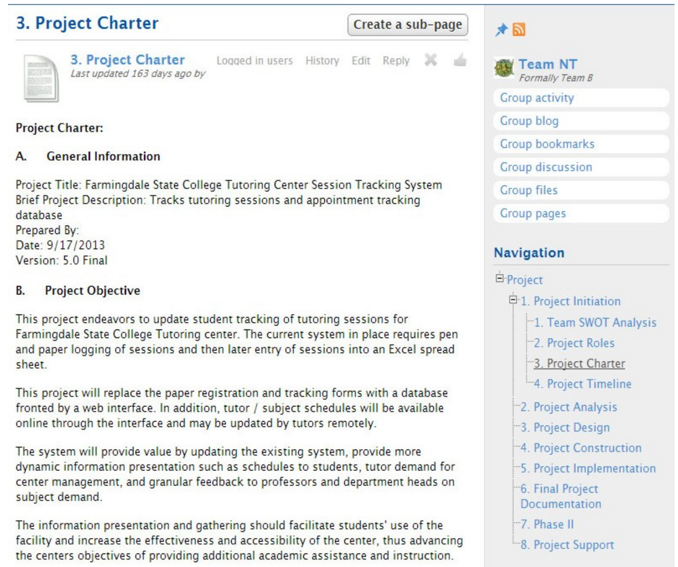

Figure 3. OLC Wiki Page

Wiki technology was developed prior to the Web 2.0 explosion and, thus, limited collaborative writing to early HTML-style markup [42]. Today, wiki-technologies allow collaborators to breathe life into wiki-documents through multi-media and allow editors to embed files. Illustrated in Figure 3, today's wikis are no longer syntax-based, with difficult HTML-style markup notation. Current wikis embed rich-text editors, which allow novice web users to participate, a notion that is particularly important for student groups, many of whom have limited experience with wiki-technology. Recent research by Xu [43] implemented wiki-technology in project-based computer science courses, highlighting how wiki technology helped tcentralize and capture all project activities through wiki pages created by both the instructor and students. Additionally, Popescu [44] discovered that wikis also helped students to find interesting information; by reading other teams' wiki pages, students could check their progress, see how they compare with others teams, look for inspiration and models and discover different ideas and approaches. A limitation identified in He and Yang [45] is that a wiki should not be a tool that aims to supplant communication channels and works best when additional modes exist. This limitation is accounted for in our OLC since the wiki comprises only one component.

V. RESEARCH METHODOLOGY

To measure how OLC software supports levels of project community and leads to project success we targeted capstone courses at our university, a small public university in the United States. Our research can be categorized as a proof-of-concept case study, where we look to measure the effects of a specific intervention on existing populations of university students taking a required capstone group project-oriented course. The IT capstone course consisted of six project teams whose seventeen-week endeavor focused on constructing a fully-functional information technology (IT) artifact. Project teams consisted of three to four students and were formulated by the instructor prior to the start of the semester.

The final IT artifact consisted of 40% of an individual's course grade. An additional 30% included project documentation and collaborative activities relating to project analysis, design and construction. Additionally, each project team was assigned a wiki space with pre-defined templates for project development phases. Groups were encouraged to communicate both synchronously during weekly in-class meetings and asynchronously through each groups' designated project space. Each group was required to present and defend their final IT artifact at the end of the semester.

Data was captured over the six-month intervention and include both pretest and posttest survey data in addition to a social network analysis (SNA).

As exploratory research, we are largely concerned with discovering mitigating factors surrounding learning, social interaction, course community and project success within OLC software. Our theoretical model asserts that, when members are presented with tools to support interaction and community, higher levels of social presence and course community will exist. At a high level, we ask what affordances and constraints OLC software provides in facilitating project collaboration and project success? Specifically, we ask the following research questions:

R1: Will an OLC enhance levels of course community?

R2: Will wiki software enhance levels of project management and collaboration?

R3: If R1 and R2, will an OLC facilitate project success?

To explore these questions, we measure the impact of our customized OLC software within our IT capstone course. The course is experiential in nature and students are required to produce results for use by real individuals and are evaluated both on process and product. While individuals were encouraged to use the OLC to discuss all project-related material, it was not required for project milestones.

VI. RESULTS

To explore how specialized OLC software can support project-based courses and enhance classroom learning, we collected data from multiple sources. Our first point of data collection is through survey research, which measures perceived levels of learning, community and interaction. To support survey findings, we perform a social network analysis and look at in-bound and out-bound interactions among OLC participants.

A. Demographics

Demographic information was captured through a pretest survey. Including the instructor, our total user population was 25. 25% of participants were female and 75% were male. 48% of participants were aged 18 to 25, 36% were aged 26 to 35, 12% were aged 36-50 and 4% were aged 50 and above. All interventions occurred using a capstone IT project-based course, a required upper-division IT course.

B. Survey Data Analysis

Closed-ended survey pretest and posttest questions were distributed to all participants resulting in 25 completed pretests and 23 completed posttest surveys, or a 100% response rate for both sets. To ensure confidentiality, no personally identifiable information was collected, thus linking survey data and SNA data was not achievable.

1) Instrument Reliability

Cronbach Alpha scores for our survey constructs related to items associated with the OLC scored .83 indicating that survey items maintain an adequate level of internal consistency.

2) OLC Perceptions on Interaction (Pretest)

Illustrated in Figure 4, on factors relating to OLC interaction, pretest results identified that individuals agreed (52%) or strongly agreed (32%) that interaction through an OLC would be important, with 40% agreeing and 40% strongly agreeing that the OLC will increase interaction.

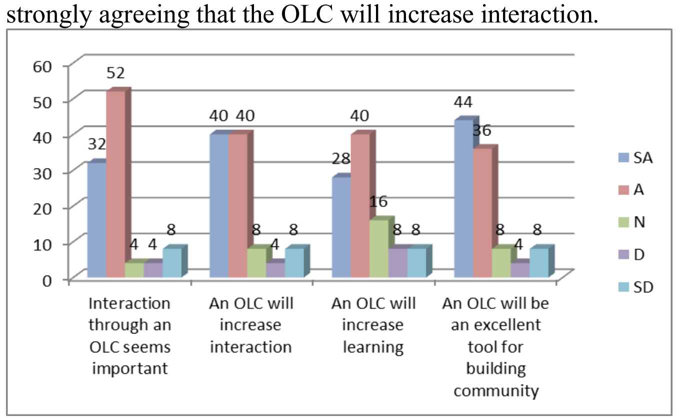

Figure 4. OLC (Pretest)

3) OLC Perceptions on Interaction (Posttest)

Illustrated in Figure 5, on factors relating to OLC interaction, posttest results show higher levels of agreement across these constructs. Individuals agreed (39%) or strongly agreed (56%) that interaction through the OLC was important, with 44% agreeing and 44% strongly agreeing that the OLC increased interaction.

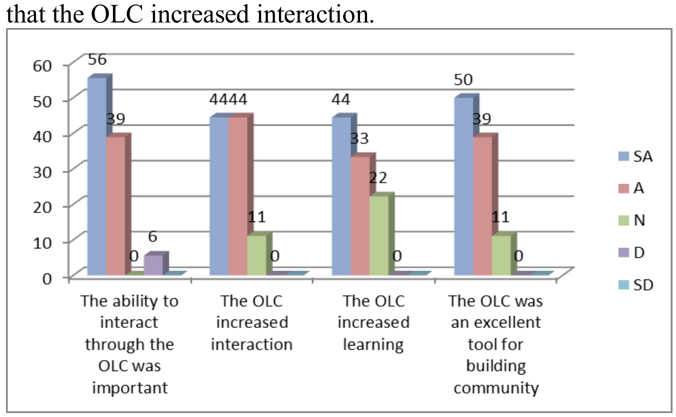

Figure 5. OLC (Postest)

4) Wiki Perceptions on Interaction (Pretest)

Illustrated in Figure 6, on factors relating to wiki interaction, pretest results identified that individuals agreed (48%) or strongly agreed (32%) that a wiki would facilitate group cohesion. Additionally, individuals agreed (40%) or strongly agreed (32%) that a wiki would facilitate group collaboration. Individuals also agreed (40%) or strongly agreed (32%) that a wiki would facilitate group interaction.

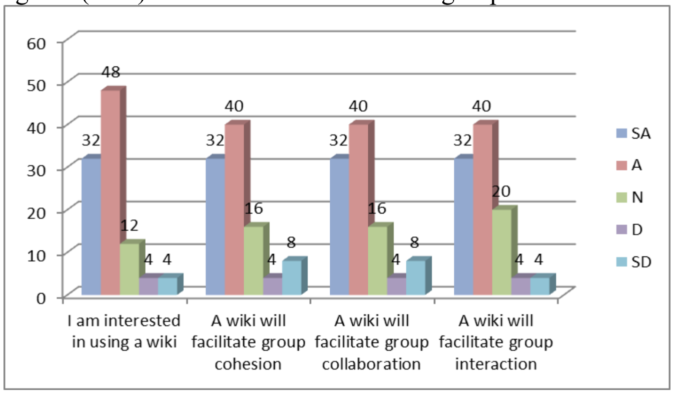

Figure 6. Wiki on Interaction (Pretest)

5) *Wiki Perceptions on Interaction (Posttest)*

Illustrated in Figure 7, on factors relating to wiki interaction, posttest results show higher levels of agreement across these constructs. Individuals agreed (61%) or strongly agreed (28%) that the OLC wiki facilitated group cohesion. Additionally, individuals agreed (50%) or strongly agreed (39%) that the OLC wiki facilitated group collaboration. Individuals also agreed (56%) or strongly agreed (33%) that the OLC wiki facilitated group interaction.

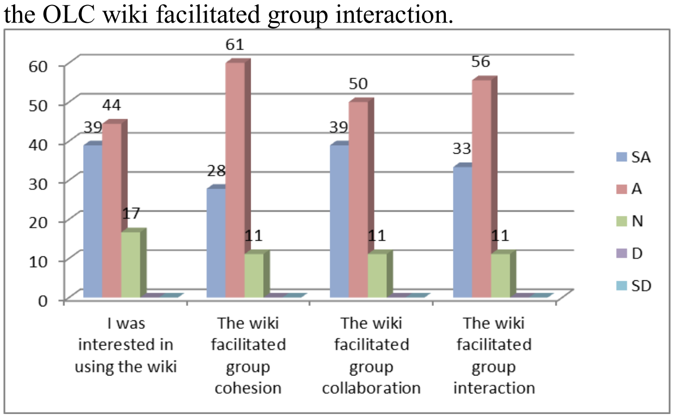

Figure 7. Wiki on Interaction (Posttest)

6) *Wiki Perceptions on Project Management (Pretest)*

Illustrated in Figure 8, on factors relating to project management, pretest results identified that individuals agreed (52%) or strongly agreed (28%) that a wiki would project management. Additionally, individuals agreed (40%) or strongly agreed (36%) that a wiki would facilitate content creation. Individuals also agreed (44%) or strongly agreed (36%) that a wiki would facilitate the organization of project content.

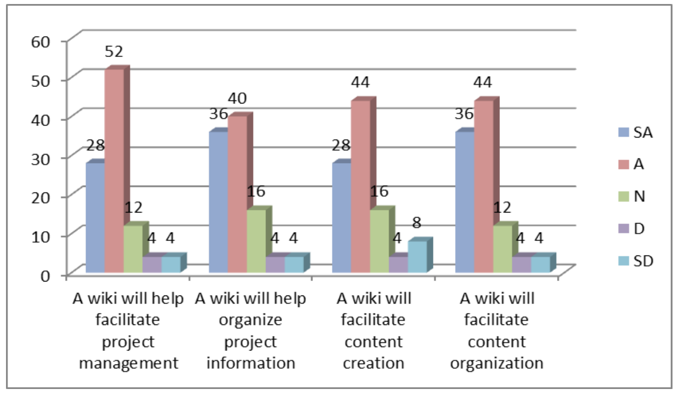

Figure 8. Wiki on Project Management (Pretest)

7) *Wiki Perceptions on Project Management (Postest)*

Illustrated in Figure 9, on factors relating to project management, posttest results show that individuals agreed (50%) or strongly agreed (39%) that the OLC wiki facilitated project management. Additionally, individuals agreed (50%) or strongly agreed (39%) that the OLC wiki facilitated content creation. Individuals also agreed (44%) or strongly agreed (50%) that the OLC wiki facilitated organization of project content.

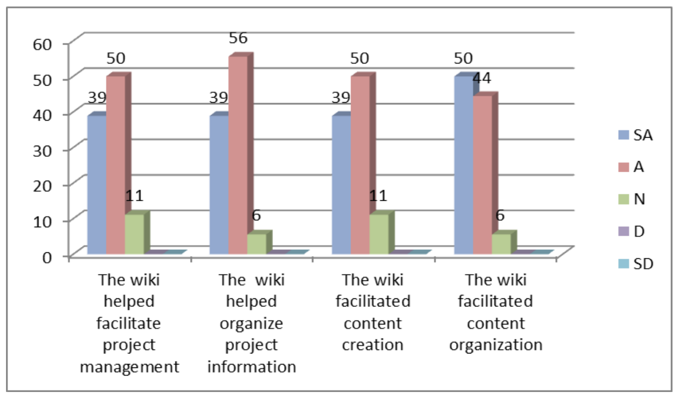

Figure 9. Wiki on Project Management (Posttest)

C. *Social Network Analysis*

1) *SNA Background*

A social network analysis (SNA) is often used to identify interactions that take place within a community. Specifically, SNAs can provide a visual analysis of the social network and allow for a better understanding of all participants in the process of learning and interaction across online environments [46]. The ability to view social graph structure and community evolution is crucial to successful facilitation of a learning design and can serve as an early indicator of its success [47].

2) *SNA Design*

We constructed our SNA graphs using the 2014 NodeXL Template for Microsoft Excel. NodeXL is a free and open source extension that provides a range of basic network analysis and visualization features [48]. Utilizing the Fruchterman-Reingold Algorithm to generate a force-directed layout, we are able to position team members (aka, nodes) in our graph so that all edges are of more or less equal length and there are as few crossing edges as possible. Additionally, each arrow represents a weighted interaction, where larger arrows indicate a greater number of interactions between participants. Furthermore, bi-directional arrows occur when there is interactivity between students, measured in-degree / out-degree values. A high average value for in-degree / out-degree indicates those students more frequently interacting with one another.

3) OLC Sociogram

Illustrated in Figure 10 is the sociogram for our capstone course. Project team members are identified by the group letter and group project grade. For example, B2(92) was the second member of Group B, whose final project grade was 92 out of 100. Overall, the course experienced high levels of interaction with 1299 total aggregate interactions. Average in-degree and out-degree was also high at 8.1 meaning. In other words, on average, each student communicated with 8 other individuals. Additionally, the total number of unique edges, or communications between any two nodes was 186.

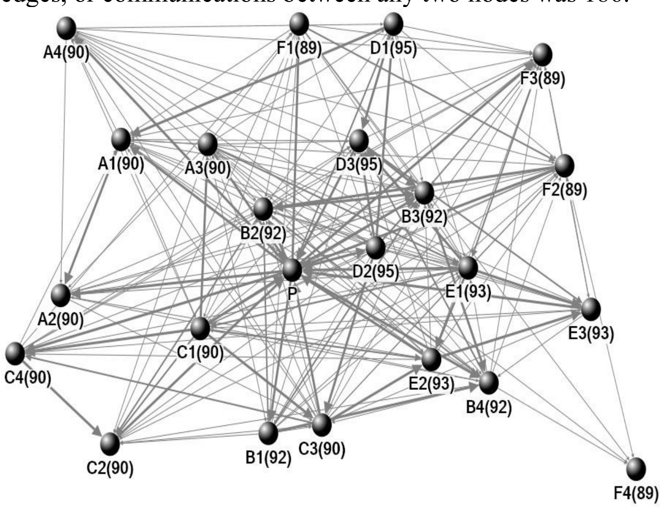

Figure 10. OLC Sociogram

VII. DISCUSSION

A. OLC Supports Community (R1)

Our first research question, R1, asked if OLC software could enhance levels of course community. To answer this question we look at a couple of factors. Community can be measured in terms of social capital, or the common social resource that facilitates information exchange, knowledge sharing, and knowledge construction through continuous interaction [49]. Furthermore, Social Presence Theory looks at the degree to which an individual's perception of an online community, in its entirety, affects his or her participation in that community. Therefore, our first point of reference focuses our attention on survey responses relating to the OLC software's ability to enhance interaction and community. Encouraging pretest results showed that individuals were positive from the start that an OLC could be important (84%). These perceptions continued throughout the lifecycle of the intervention and it was more encouraging to discover higher levels of agreement that the OLC was, in fact, an important factor for facilitating interaction (95%). Similarly, pretest results showed that individuals were positive that an OLC could be an excellent tool for building community (80%). And, again, posttest results supported these perceptions, revealing even higher levels of agreement (89%). The fact that the OLC is an open environment allowed team members to review the progress of their classmates and pose questions and receive responses in an open dialogue.

Our second factor for measuring community focuses on the SNA sociogram. Illustrated in Figure 10, it is evident that the OLC was an active and engaging community. In fact, the average in-degree/out-degree was 8, which indicates that a third of all participants interacted with other members of the OLC. An additional factor considers the large number of unique peer-to-peer interactions (186), which means that many members were communicating with many other members. This factor reinforces the notion that OLC software provides a high affordance for individuals to discover and connect with other members of the community in addition to those members from within their respective project teams. As this community building was exactly the goal of our OLC software, the presented results confirm that the developed OLC software provided affordances for students to cultivate their collaboration skills in team-based IT projects. These findings are particularly important as employers increasingly ask their employees to work in virtual teams [50]. Furthermore, we interpret these important results as a sign that the team-based IT project employed in this study was industry-aligned.

B. OLC Wiki Supports Project Management (R2)

Our second research question, R2, asserted that the OLC project wiki would enhance levels of interaction and facilitate team collaboration. To answer this question we consider a couple of factors. First, we focus our attention on survey responses relating to the OLC wiki's ability to enhance group cohesion, collaboration and interaction. Similar to R1 results, it was encouraging to find that the majority of individuals believed that the wiki could facilitate cohesion (72%), collaboration (72%) and interaction (72%). More so, however, it was very encouraging to discover the higher levels of agreement in the posttest that the wiki actually contributed to higher levels of cohesion (89%), collaboration (89%) and interaction (89%). Engagement theory is concerned with meaningful engagement. This amounts to finding the right tools for the right projects. Wiki software is geared towards collaboration and interaction where individuals bear witness to the evolution of a project's analysis and design. Wiki software also reinforces the notion that projects can be both user-centric and group-oriented, thus facilitating individual ownership and motivation.

Referring back to the sociogram in Figure 10, the proximity of nodes reflects that these nodes interacted with one another more frequently. In other words, the closer a set of nodes are to one another, the more cohesive that group of

 ISBN: 978-1-61208-436-7

nodes are as a unit. As one would expect of our capstone project teams, with the exception of Group F, all groups exhibited high levels of team cohesion. This outcome was largely to be expected since individuals, while functioning as part of the larger course community, were still responsible for working within their own respective project teams in order to accomplish project milestones. Taken together, the quantitative results at hand indicate that students utilized OLC to set achievable project goals, resolve misunderstandings about design decisions, and negotiate deliverables, similar to the way team-based IT projects function in the real world. In this way, collaborating students used one another as a resource for learning, while also working to complete their project milestones.

C. OLC and Wiki Supports Project Success (R3)

Our third research question, R3, asserted that based on the successful adoption of the OLC as a mechanism for fostering interaction and supporting community along with the successful adoption of the wiki to support project collaboration, the OLC would attribute to project success. Our results indicate that this concept is also well supported.

An important measure of project success stems from a group's ability to establish the parameters of success through analysis of business requirements and the design and construction of the IT artifact. It was clear that the OLC helped to contribute to this success as identified in both survey responses and in each project team's final product. From the pretest survey responses gathered, it was encouraging to discover that the majority of individuals believed that the OLC wiki would facilitate project management (80%), information organization (76%) content organization (80%) and content creation (72%). More so, however, it was encouraging to discover higher levels of agreement in posttest responses, where individuals perceived that the OLC wiki did, in fact, contribute to higher levels of project management (89%), information organization (95%) content organization (94%) and content creation (89%). Each of these factors is an important dimension of project management that promotes a shared understanding of technical requirements, which helps to mitigate expensive and time consuming rework. This concept applies to both short-term and long-term IT projects.

Consequently, the combination of 1) an OLC, which facilitated member interaction and course community and 2) wiki software, which allowed individuals to collaborate towards project milestones, allowed each team to successfully meet capstone project expectations and deliver a final IT artifact that represents understanding of business processes.

D. OLC Supports Technical Learning

Finally, an important consideration should be discussed surrounding the introduction of social software within an academic setting for learning purposes. While the merits of the OLC as a mechanism for project success and/or enhancing levels of academic community may be debated, the introduction of specialized social software, such as an OLC, into team capstone courses provides a number of tangible and intangible benefits not measured completely in this research.

In today's dynamic business world, social software is pervasive across the IT sector. Additionally, capstone courses are one of, if not, the final course for students majoring and graduating with IT diplomas. Consequently, introducing students to how communities of practice engage in information sharing and knowledge construction using such technologies identified in this study may go a long way in preparing those students for similar communication and interaction in the IT industry.

VIII. LIMITATIONS

We acknowledge that a number of limitations exist in this research. One limitation considers using an academic setting as an environment to measure the impact OLC software has on project-based teams. While we acknowledge that this does limit the generalizability of the study, it should be noted that there are numerous similarities between computer supported learning and working teams that make knowledge gained in one setting applicable to another setting [51].

Another limitation considers the relatively small sample size analyzed in this study. While we acknowledge this fact, our primary goal is to showcase OLC software as a proof-of-concept for enhancing collaboration among project-based teams, which we believe we succeed in doing on a number of levels.

IX. CONCLUSION

In this research we leverage the technical prowess of todays' digital natives and measure the impact of specialized OLC software on project success within teams participating in IT capstone courses. Our software allowed individuals to function within close-knit project teams, while also participating within a larger academic community of practice. Through the analysis of survey data and supported through a social network analysis, we discovered the powerful and positive impact OLC software has on supporting project success by facilitating peer-to-peer interaction and enhancing levels of collaboration.

X. REFERENCES

[1] StatisticBrain. (2012). Social networking statistics, Retrieved July 29, 2013, from: www.statisticbrain.com/social-networking-statistics.

[2] Jenks, J. (2011). "The digital world of millennials," *eMarketer*.

[3] Williamson, D. (2007). "Social network marketing: ad spending and usage, social network marketing," *eMarketer*. Retrieved June 02, 2008, from http://www.emarketer.com/Report.aspx?code=emarketer_2000478.

[4] Oblinger, D. and Oblinger, D. (2005). "Is it age or IT: first steps toward understanding the Net Generation," *EDUCAUSE*.

[5] Lave, J. and Wenger, E. (1991). *Situated Learning*: Legitimate Peripheral Participation, Cambridge University Press.

[6] Adams, EC and Freeman, C. (2000). "Communities of practice: bridging technology and knowledge," *Journal of Knowledge Management*, 4(1), pp. 38-44.

[7] Thoms B., Garrett, N. and Ryan T. (2009). “Online learning communities in the New ‘U’. *Intl. Journal of Networking and Virtual Organisations*, 6(5).

[8] Shaw M., Tomayko JE. (1991) "Models for undergraduate project courses in software engineering," *Software Engineering Education in Lecture Notes in Computer Science* 536(1), pp 33-71.

[9] Fincher, S., Petre, M., & Clark, M. (2001). *Computer science project work: principles and pragmatics*. London: Springer-Verlag.

[10] Goold, A.; Horan, P., (2002). "Foundation software engineering practices for capstone projects and beyond," *Proceedings. 15th Conference on Software Engineering Education and Training*, pp.140,146,

[11] Lynch, K., Goold, A., & Blain, J. (2004). “Students’ pedagogical preferences in the delivery of IT capstone courses.” *Issues in Informing Science and Information Technology*, 1, 431-442.

[12] Lainez, M., Deville, Y., Dessy, A., Dejemeppe, C., Mairy, J., & Van Cauwelaert, S. (2014). “A project-based introduction to agile software development.” In L. Alves, P. Ribeiro, & R. Machado, *Project-Based Learning: An Environment to Prepare IT Students for an Industry Career* (pp. 230-249). IGI Global.

[13] Ayas K. and Zeniuk, N. (2001). “Project-based learning: building communities of reflective practitioners,” *Management Learning*, 32(1).

[14] Dunlap JC, (2005). "Problem-based learning and self-efficacy: how a capstone course prepares students for a profession,” *Educational Technology Research and Development*, 53(1), pp. 65-85

[15] Clarke, N. (2005) “Evaluating student teams developing unique industry projects.” *Proceedings of the 7th Australian Conference on Computing Education* (ACE2005) (pp. 21-30). Newcastle, Australia. CRPIT, 42. ACS

[16] Newman, I., Daniels, M., & Faulkner, X. (2003). “Open ended group projects: a 'tool' for more effective teaching.” *Proceedings of the Fifth Australasian Conference on Computing Education*, Adelaide, Australia, 95-103.

[17] Kolb, D. (1984). *Experiential learning - experience as the source of learning and development*. New York: Prentice Hall.

[18] Lynch, K., Heinze, A., & Scott, E. (2007). “Information technology team projects in higher education: an international viewpoint,” *Journal of Information Technology Education*, 6, 181-198.

[19] Thoms B., Garrett B., Soffer M. and Ryan T. (2007), “Resurrecting graduate conversation through an online learning community,” *IADIS Multi Conference on Computer Science and Information Systems*, Lisbon, Portugal.

[20] Thoms B., Eryilmaz, E. (2014). “How media choice affects learner interactions in distance learning classes,” *Computers & Education*, v75.

[21] Johnson, C. (2001). “A survey of current research on online communities of practice,” *Internet and Higher Education*, 4(1), 45-60.

[22] Palloff, R. and Pratt, K. (1999). *Building learning communities in cyberspace: effective strategies for the online classroom*, San Francisco, California: Jossey-Bass.

[23] Savery, J. and Duffy, T. (1996). “Problem based learning: an instructional model and its constructivist framework,” In B. Wilson, *Constructivist Learning Environments: Case Studies in Instructional Design*. Englewood Cliffs, NJ: Educational Technology Publications.

[24] Hagstrom, F. and Wertsch, J. (2004). “Grounding social identity for professional practice,” *Topics in Language Disorders*, 3(24), 162-173.

[25] Piaget, J. (1952). *The Origins of Intelligence in Children*, New York, NY: Intl. University Press.

[26] Squires, D. (1999). “Educational software and learning: subversive use and volatile design,” *Proceedings of the 32nd Annual Hawaii international Conference on System Sciences*, IEEE.

[27] Matthews, M. (2002). “Constructivism and science education: a further appraisal,” *Journal of Science Education and Technology*, 11(2).

[28] Kuh, G., Cruce, T., Shoup, R., Kinzie, J. and Gonyea, R. (2008). “Unmasking the effects of student engagement on college grades and persistence,” *Journal of Higher Education*, 79(1), 540-563.

[29] Kearsley, G. and Schneiderman, B. (1999). “Engagement theory: a framework for technology-based learning and teaching,” Retrieved June 26, 2013, from http://home.sprynet.com/~gkearsley/engage.htm

[30] Dalsgaard, C. (n.d.). “Social software: e-learning beyond learning management systems,” Retrieved 6 26, 2013, from http://www.eurodl.org/materials/contrib/2006/Christian_Dalsgaard.htm

[31] Waycott, J., Sheard, J. Thompson, and C. Clerehan, R. (2013). “Making students' work visible on the social web: a blessing or a curse?,” *Computers & Education*, 68(1), 86-95.

[32] Teeper, M. (2003). *The rise of social software*. netWorker, 7(1), 18–23.

[33] Tu, C. and McIsaac, M. (2002). “The relationship of social presence and interaction in online classes,” *The American Journal of Distance Education*, 16(1).

[34] Short, JA, Williams, E. and Christie, B., (1976). *The Social Psychology of Telecommunications*. John Wiley & Sons.

[35] Garrison, R., Anderson, T., and Archer, W. (1999). “Critical inquiry in a text-based environment: computer conferencing in higher education,” *The Internet and Higher Education*, 2(3), 87-105.

[36] Richardson, JC and Swan, K. (2003). “Examining social presence in online courses in relation to student’s perceived learning and satisfaction,” *Journal of Asynchronous Learning Networks,* 7(1), 68-88.

[37] Ma, A. (2008). “Computer supported collaborative learning and social creativity: a case study of fashion design,” *Journal of Information, Information Technology, and Organizations*, 3(1), 17-39.

[38] Vasileiou, V. N. and Paraskeva, F. (2010). “Teaching role-playing instruction in Second Life: an exploratory study,” *Journal of Information, Information Technology, and Organizations*, 5(1).

[39] Thoms, B. (2011) “A dynamic social feedback system to support learning and social interaction in higher education,” *IEEE Transactions on Learning Technologies*, Issue: 99.

[40] Preece, J. (2001). “Sociability and usability in online communities: determining and measuring success,” *Behavior and Information Technology*, 20(5), 347-356.

[41] Boulos, M., Maramba, I., & Wheeler, S. (2006). Wikis, blogs and podcasts: a new generation of web-based tools for virtual collaborative clinical practice and education. *BMC Medical Education*, 6(41), 1472–1480.

[42] Leuf, B., & Cunningham, W. (2001). *The Wiki Way: Quick Collaboration on the Web*. Boston: Addison-Wesley Longman Publishing Co.

[43] Xu, L. (2007). “Project the wiki way: using wiki for computer science course project management,” *Journal of Computing Sciences in Colleges*, 22(6), 109-116.

[44] Popescu, E. (2014). “Using wikis to support project-based learning: a case study,” *Proceedings of the 14th IEEE International Conference on Advanced Learning Technologies*, Athens, Greece.

[45] He W. and Yang L. (2015). "Is wiki a good tool facilitating team collaboration in undergraduate education?" *Proceedings of the 19th*

Pacific Asia Conference on Information Systems (PACIS 2015), Marina Bay Sands, Singapore.

[46] Freeman, L. (2004). *The development of social network analysis*. Vancouver: Empirical Press.

[47] Bakharia, A., Heathcote, E., & Dawson, S. (2009). “Social networks adapting pedagogical practice: SNAPP,” *Proceedings of ASCILITE* Auckland 2009, (pp. 49-51), Auckland, New Zealand.

[48] Hansen, D., Shneiderman, B., & Smith, M. (2011). *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Burlington: Morgan Kaufmann.

[49] Daniel, B., Schwier, R., & McCalla, G. (2003). “Social capital in virtual learning communities and distributed communities of practice,” *Canadian Journal of Learning and Technology*, 29(3), 113–139.

[50] Majchrzak, A., Malhotra, A., and John, R. (2005) “Perceived individual collaboration know-how development through information technology–enabled contextualization: evidence from distributed teams,” *Information Systems Research* 16(1), pp. 9-27.

[51] Kirschner, P. A. and Erkens, G. (2013) “Toward a framework for CSCL research,” *Educational Psychologist,* 48(1), pp. 1-8.

ShareLab, Support for Collective Intelligence

1 deadline, 11 designers, 1 project

Samia Ben Rajeb
LUCID Lab for User Cognition & Innovative Design
University of Liège, Belgium
e-mail: samia.benrajeb @ ulg.ac.be

Alexandru Senciuc
Independent research
Paris, France
e-mail: alexandru.senciuc @ gmail.com

Irene Pluchinotta
LAMSADE - Paris Dauphine University, France
DICATECh - Technical University of Bari, Italy
e-mail: irene.pluchinotta @ gmail.com

***Abstract*— This paper presents the implementation of a collaborative action research approach aimed to assist in constructing collective intelligence. Named *ShareLab*, this project was implemented as part of a call to an international competition bringing together different skills originating from varying cultures so as to produce a common project in a very short time. What is the origin of *ShareLab*? How was it put into play? What are its advantages and limitations? This article aims to answer these questions thanks to the feedback obtained from this competition experience.**

Keywords-collaborative action-research; case study of collaborative activity; managing collaborative design project; adaptive collaboration; collective intelligence.

I. Introduction

Architectural design is a complex activity that operates in an increasingly coercive regulatory environment and that has to deal with the competition and urgency resulting from ever-shorter deadlines. Under competitive circumstances, the challenge of the designers is not only to comply with all these constraints, but also to propose creative and innovative ideas that can win over a jury. To address these constraints fully, architectural firms (whether small, medium or large) innovate through interdisciplinary approaches to combine various skills needed to carry out the project. Nevertheless, faced with this variety of contributors to the project, certain information related to the constraints and design arguments is lost; the risk of generating misunderstandings and disagreements grows, and managing group cohesion becomes more and more difficult [1].

With regard to this problem, we propose the employment of a novel approach, entitled *ShareLab*. It targets the cohesion of a group in order to gradually bring about a collective intelligence [2]. In this vein, this article first sketches the main scientific contributions that have identified those components promoting collective intelligence. We then propose to connect those findings to create and define our unique approach, while demonstrating its significance thanks to the concrete case of a multidisciplinary, collaborative work carried out as part of an international architectural competition.

Our article is divided into four parts. Section II will introduce the theoretical framework of our study by highlighting the concepts that characterize collaborative activity and the approach that best cultivates it. Section III will describe the context of our study and the design team. Section IV will present the methodology implemented. Finally, the conclusion will address the contributions and limitations of this experience, with possible avenues for improvment.

II. Theoretical Framework

A. Specifying collective activity components before choosing tools

To define our research framework, let us start by distinguishing between what pertains to *cooperation* and what pertains to *collaboration* in collective design activities. We hold that collaboration entails instances in the design process where multiple contributors interact together around the same objective and the same tasks, thus difficult to separate [3]. Conversely, in times when they cooperate, each player carries out his or her own task in parallel with the others. These two distinct stages of the collective design process involve different tools and procedures.

For roughly the past twenty years, the field of Computer Supported Cooperative Work (CSCW) - has been developing many supports for collective activity. However, most of these scientific contributions remain rather technically centered and too often neglect the organizational aspect in addition to personnel management. Most tools proposed therein are intended to assist the cooperation of contributors by facilitating asynchronous, remote exchanges. These cannot, however, be used when the project is sufficiently advanced. The sketch phase, during which the most important design choices are decided upon, enjoys little instrumental aid from the collaborative perspective.

However, meeting up despite long distances, exchanging ideas in real time, maintaining trust between members, managing both *cognitive* and *operational synchronizations*, collectively sharing artifacts, and developing *awareness* and *common ground* are primordial in ensuring a collective design project. On their own, current tools are not enough to manage all these elements of collaborative activity. Our research question lies herein: how can we combine these into an inclusive model that is applicable under real conditions?

B. Orientation towards a collaborative action research approach

To answer this question, we propose here to define, contextualize and integrate all of these elements composing collaborative activity, within the framework of an approach that invites the contributors to reflect upon their situation. To this aim, ShareLab - defined by the LUCID laboratory at the University of Liège and developed as part of this contest with TARTAR, a design team - employs an interventionist approach. This method aims to support work situations that involve multiple collaborators, working together for the first time to design, in this case, an architectural project in a competition with a very short deadline. That is why our approach integrates all of these collaborators and attempts to engage them in "a critical and dynamic reflection about a situation that concerns them" [4]. Inspired by the methods of "collaborative action research" [5], the aim of this approach is to focus the attention of all players both on the results of the competition and on their processes and methods of working with multiple people. This type of method is not opposed to conventional scientific approaches, but rather supplements them by managing the concerns of participants in a situation created by the intervention of researchers wanting to develop a shared understanding of that situation [6]. All contributors (researchers and practitioners, observers and designers) work together to build whole new meanings related to their activity, thanks to the synergy of their views, self- and mutual appraisal of their actions, self- and co-training, and co-evolution of the methods implemented in order to work together [7]. Although the principles of the approach applied here are part of a specific, predefined, theoretical framework, it remains necessary to implement a protocol that integrates the specifics of each work group and their project, as well as the context and the constraints that they will have to manage.

By implementing the concept of ShareLab, our objective is not to impose a sole method, but rather to adopt an integrated approach that encourages group cohesion and that aims to bring about collective intelligence. To adopt an integrated approach while maintaining each designer's individuality: our assumption considers the collective activity in design as complex, difficult to generalize and which result is first thought, negotiated, valued, challenged and co-built by the group before it even exists.

III. A SHARED RESPONSE TO A DESIGN COMPETITION

Cohesion in a work group is not something to be taken for granted, even less so when the group members do not know each other and assemble for the first time, as is often the case. Players might exit the framework of their habitual activity, but their past experiences still shape how they work with others [8]. Their mutual interactions are thus subjected to several emerging and recurring factors, such as trust, culture, language, individual specificities, tools and mediation procedures between them, etc. [9]. Other factors also come into play, like the notion of leadership or motivation to participate in a joint project [10]. As part of the competition involved in this study, the group of designers and researchers had aspired to create an innovative process motivated by the novelty of the situation bringing them together, even before starting the design itself. For the group, this involves going beyond their own experiences and areas of expertise so as to incorporate the interests and fields of others in achieving an unprecedented joint result.

A. Presentation of the competition and project produced

The work team in our study concerns one of three winners of an international architectural design competition that attracted 1.749 applications from 90 different nationalities [11]. This team focused on the problem of "the rise of sea levels around the world." Dubbed TARTAR, it consists of 11 people working together for the first time, whose academic backgrounds are different (graduates of France, Romania, Tunisia or Italy), and whose skills (as practitioners, professors, researchers, and students) also come from diverse backgrounds (architecture, urban planning, engineering and humanities). Such diversity, at the heart of their participation in this contest, and their motivation comprise the specific character of this team. Their premise was that an original idea could only take shape within the diversity of all the points of view made up of each of the project participants. To help them reach their goal, the definition of the organizational protocol itself was put together beforehand along with the collaborators and evolved over time from their feedback. The result of this competition was therefore based on this organization and on this co-constructed process. The project was born from the diversity of skills, on the one hand, and, on the other, from the awareness of problems caused by the rising waters (from inventory of various specialties), and the latest scientific discoveries in the fields of artificial intelligence and synthetic biology.

The project submitted by the group consisted in creating new territories through a cooperative system connecting a digital data system and an evolving material inspired by coral.

B. Organizational setup

The organizational structure created here strives to support collective decision-making and to develop an environment that promotes understanding between group members. The ShareLab complements this structure and is based on principles which are defined enough to be understood by all, but also open enough to be re-appropriated and easily adaptable to changes in the project and how the actors interact.

This organizational protocol, called Collective Intelligence Support Protocol (CISP), aims to manage the team in the collective design of their project, all the while integrating space-time constraints on the one hand, and levels and production capacity of each individual within the group on the other [12]. This protocol ensures the coordination of the team and its *operational synchronization* (relative to the sharing of tasks, according to the definition given by Falzon and Darses [13]. It provides a structure for production level management, deadlines to be met, exchange tools, work produced and the role of each player in the process. Based on a multi-layer system, the work group was initially divided into four interconnected teams: "Organization", "Research" of concept, "Exploration" and "Production" of project. Based on this division, the design stage was then divided into three phases: research phase, exploration phase and production phase. Each team was responsible for its own phase, but all teams were involved in all phases (cf. Figure 1). Wishing to complete the organizational structure, ShareLab was set up to manage the transition from one phase to another and thus ensure the construction of a collective intelligence in the group.

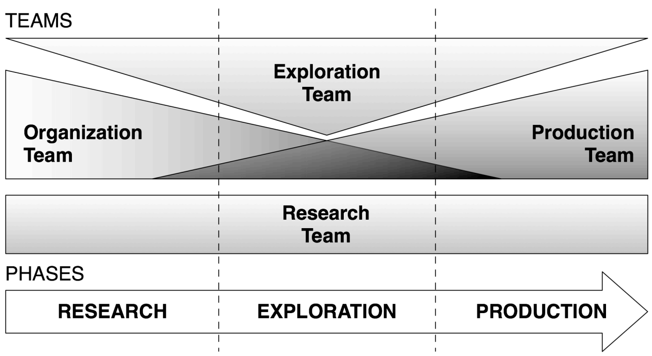

Figure 1. Relationship between each work phase and the involvement level of each team in that phase.

IV. DEFINITION OF SHARELAB

ShareLab arose from reflections built on observing collective activities in action in an architectural, design, and engineering firm setting [14]. Aimed at a more interventionist approach and drawing on methods from "collaborative action research," the LUCID laboratory at the University of Liège has sought to define a process of co-reflection in which the problem definition, analysis and the recommendations come from both researchers and practitioners themselves.

Based on "activity theory model" and following the work of the CRADLE Finnish research team from the University of Helsinki [15], ShareLab attempts to foster common ground, helping synchronization within a group whose work habits, procedures and tools are not yet clearly defined.

Indeed, the "activity theory model" (cf. Figure 2) provides the means to understand collective activity by taking into account the actions and contributions of each individual (*subject*) in the group (*community*) focused on an activity (*object*), to meet a common goal (*outcome*). This model also connects these elements with (1) the explicit or implied rules defined between the subjects and their community, (2) the tools used by the subject to act on the object, and (3) tasks to be performed on the object by each member of the community in order to achieve the final goal collectively [16]. This type of model should be analyzed very carefully because (relative to the 5 principles of "activity theory" as defined by Engeström [17]) it is necessary to incorporate an activity model with other activity systems. While the overall goal of the group is the same, the sub-goals of the individuals can differ. It is therefore essential to see the model as a continuously shifting network of nodes of activities faced with a set of contradictions [18]. These may occur within the same node, between the nodes of a same system, between the existing system and the objective, or between one activity system and another, involved in the production of a common outcome.

In a system where neither instruments nor rules nor even a division of tasks has been defined, this consists in making the three nodes interact throughout the design process. As part of this competition, the definition and implementation of our ShareLab are aimed at better anticipating conflicts and helping the group build its own collective intelligence (see the linking circle in Figure 2).

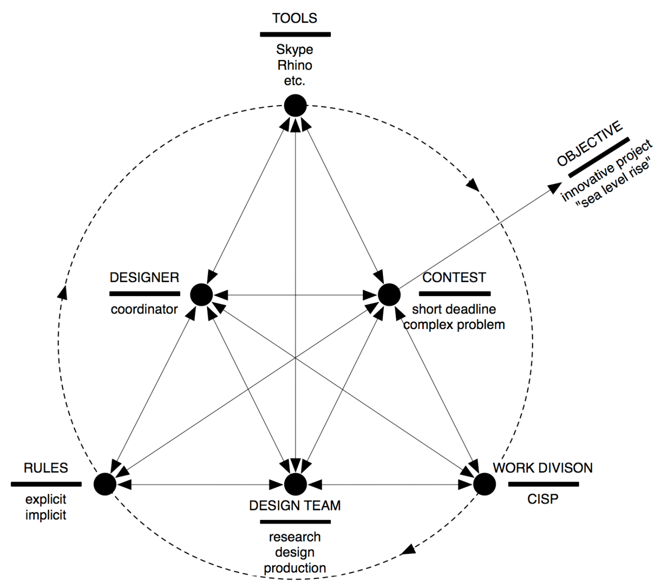

Figure 2. Activity theory model applied to the TARTAR project, according to Engeström, 1987 [19].

"Change Laboratory is a method for developing work practices by the practitioners. Basing on the theoretical conceptions of the dual (double) stimulation (L. Vygotsky) and expansive learning (Y. Engeström) it facilitates both intensive, deep transformations and continuous incremental improvment. The method is developed and registered by the Center of Activity Theory and Developmental Work Research, Universityof Helsinki " [20]. In line with the Change Laboratory, ShareLab rather intervenes upstream, when the participants have not yet built any **awareness**, **trust** between them has not been acquired, the **shared items** still not defined, the **common ground** unincorporated, and cognitive and operational **synchronization** not ensured.

To better know each other, especially to meet a common objective on a short deadline, the ShareLab tends to support, through an iterative process, design collaboration and collective ideation by involving all stakeholders and by integrating the maturity of their thinking and the progress of their joint project.

ShareLab was then imagined to join together these five key concepts, through an iterative process, to foster collective ideation (cf. Figure 3).

Let us examine these notions individually, to elicit the theoretical concepts and to indicate how they are implemented, in conjunction with one another, in our model.

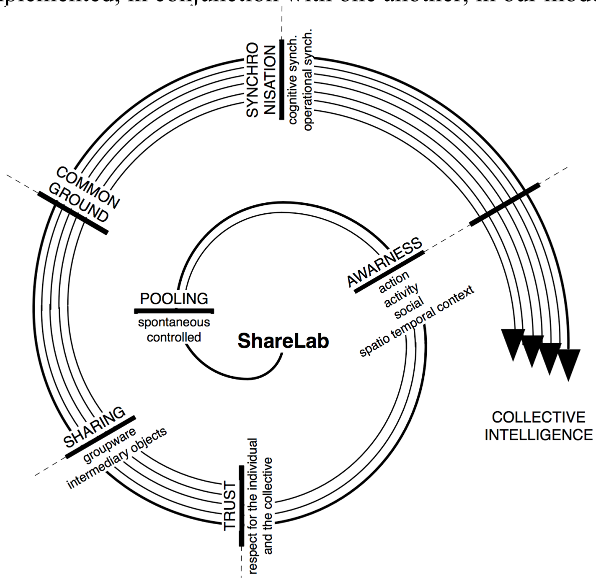

Figure 3. Evolving and growing strategy of building collective intelligence within ShareLab.

A. *Synchronization*

In a collective work, all contributors must know the objectives, processes, project context, topics and tasks to be done. This mutual knowledge improves the efficiency of the participants in their work together and facilitates dialogue between them. Two synchronization modes are distinguished and complement each other [13]: *cognitive synchronization* (on areas of shared skills and knowledge) and *operational synchronization* (relative to the distribution of tasks between collaborators). For its part, cognitive synchronization emerges from a process of discussion, negotiation and evaluation between collaborators. When tasks are new or not clearly divided, the *operational synchronization* plays an important role in the coordination of collective activity. It ensures, in fact, the definition of these tasks and their planning in connection with the common goal of the group. These synchronizations are not acquired but instead emerge from a process of discussion, negotiation and evaluation in which common ground co-evolves between collaborators.

B. *Common ground.*

The involvement of multiple skills requires taking into account the multiplicity of viewpoints [21] via an argumentative and negotiation processes. Everyone tries to ensure that their views are well understood by others [22]. These views are regulated little by little during the process to converge gradually towards a shared understanding of project data. This shared understanding has been described by *common ground* [23]. This *common ground* is critical to collaboration: it helps to pool specific skills and contributes to the acquisition of new skills needed to work in a group. It also participates in the referential interpretation process by increasing the speed with which the referent is identified by collaborators. Common ground is thus not a prerequisite. It follows the same process of collaboration, involving a procedure of pooling and it stems from system rules, negotiations and sharing of artifacts, tools, and conventions [24].

C. *Between the sharing of artifacts and tools*

Mirroring the process of refining ideas, artifacts continue to evolve through the exchange of different views between collaborators. They promote "*reflective conversation*" [25], allowing individual participants to shape their thoughts and share them with others. These artifacts provide a common basis between the participants. They are called *intermediary objects* when not completely fixed but remain changeable throughout the process [26]. These intermediary objects can encourage mediation, translation and/or representation. These artifacts are also called *boundary objects* [27] when they comprise fixed materials for negotiation and collaboration. To manage the sharing of these artifacts, it is equally necessary that the designers agree on the tools to use to work together. This conscious sharing of artifacts and tools is possible, however, only once trust has been established between designers. Indeed, a lack of knowledge of the situation, context, tools and specificities of each member can cause conflicts, leading the group to unsatisfactory decision-making for the project [28].

D. *Management of trust by respect for the individual and the collective*

Within a complex, multidisciplinary group activity, each participant must maintain the distinctiveness of their point of view and develop their own analysis of the problem to solve it. Yet this variety of perspectives may undermine the coherence of the project and may even induce a series of conflicts in the group. Avoiding groupthink while preserving group cohesion is the major challenge of any collective, multidisciplinary activity [3]. That is why it is worth developing methods that can promote integration of this diversity of perspectives, bringing about common knowledge that nurtures the project and participates in its development. Combining these views would, according to Belkadi [28], prevent conflicts and thus foster trust between project collaborators. Note however that trust-building strategies are not the same at the beginning or end of an activity [29]. Early on in an activity, social communications reflecting the enthusiasm of each participant and individual initiatives best promote trust, while, at the end of the process, individuals tend towards a concordance of views and more predictable communications with clear answers given in compliance with deadlines. All these parameters are facilitated by the construction of an awareness among the different members of the group evolving in a dynamic frame of trust,

 ISBN: 978-1-61208-436-7

encouraging each player to circulate their own knowledge, references and experiences with others [30].

E. Awareness

Many cognitive and social science studies have focused on the concept of awareness as a central parameter of any collective activity [31]. Many definitions have been proposed, as well as typologies aimed at specifying these various aspects. One of the best-known typologies is that of Caroll et al. [32], where the authors distinguish three types of awareness: *social awareness* (relating to the consciousness of an activity's social context); *action awareness* (relating to the consciousness of all participants' tasks and contributions in the process); and *activity awareness* (relating to design activity within the group). We add the notion of *spatio-temporal context awareness* which considers the context and the interaction spaces. These various modes of awareness are facilitated by pooling, which may be spontaneous in an informal framework (*spontaneous pooling*) or controlled, in a more planned – but not imposed – one, and that participates in the sharing of knowledge, experience and references between contributors (*controlled pooling*).

V. IMPLEMENTATION

ShareLab was set up as part of this four-month-long competition, with a threefold objective:

- to face the incongruity of managing creativity in a group in a very short timeframe;
- to organize the transition from cooperative circumstances to collaborative ones;
- to ensure a comprehensive group approach while respecting the specificity of each member.

Thus, ShareLab offers participants a collaboration-friendly environment and the emergence of new ideas in a process adapted to the reality of their activity. The principle behind this is that ShareLab is supervised and supported by the intervention of researchers working together with designers. Indeed, each ShareLab, marking the transition from one phase to another, was co-built with the previous phase's team leader. Before meeting with all contributors in the group, the researcher redefines the following with the leader:

- What are the objectives before the start of this phase?
- What are the objectives achieved at the end of this phase?
- What are the requirements, prerequisites and difficulties encountered during this phase?
- What are the goals to be achieved in the next phase?

Based on these concerns, the researcher and the co-leader build the protocol together to be applied in each ShareLab session, which is generally divided into three stages:

1) management awareness and building a climate of trust by:

- presenting each new member of the group: On what are they working now? Do they have other priorities outside of the competition? Do they have references or ideas to share? Do they have difficulty using a given tool, sharing data or answering another member's request?
- co-defining the objectives of this new ShareLab, marking the transition from one phase to another;

2) management of sharing and building a common ground: by the co-design and co-development of ideas from pooling work done in the previous phase and using several methods from Design Thinking and Serious Game (Set, Search, Imagine, Model, Select, Implement);

3) management of operational synchronization: by co-distribution of tasks and co-definition of the objectives of each team in the next phase.

As shown in Figure 4, the ShareLab has been used several times along the design process: during its introduction (when designers meet for the first time), during its finalization (when designers finalize their rendering and synthesize the project together) and during each articulation between phases (for example, when the exploration team seeks to pass on the production team.

Let's take the example of the first ShareLab session: its aim was first to establish trust in the group, since all individuals meet for the first time.

Figure 4. Different modalities established for each ShareLab.

It was then necessary to invite them to participate, as a sub-team, to a common activity on the theme of water, but outside their usual work area and using a new approach they did not applied before (here by taking photos at a water park of any reference concerning the topic of water).

A synthesis of this work had then to be co-built by the group creating a mind mapping with all the collected data. This approach allows to deconstruct their own representations by putting them in a non-standard situation. To force them to quickly collaborate with others in order to produce a common result in a small challenge far from the contest topic, invited them to go through a deconstruction phase prior to co-building a new common perception.

VI. INITIAL FINDINGS

Before even attempting to equip participants to manage their processes and artifacts, one must take into account their interdependence and the context in which their collective activity will evolve. The application prescribed by a coordinator is never just the work carried out by the

 ISBN: 978-1-61208-436-7

collaborator. It is first interpreted and reconstructed by the subject through his or her own internalized psychological instruments [33]. That is why it is relevant to build the objectives of each task together, with its specific characteristics and interdependencies, rather than impose one without trying to share and synchronize tasks performed by one member or the entire group during the previous phase. Throughout the evolution of design and negotiation processes enabled by ShareLab, group maturity can come about. Taking a step back from one's own activities makes it possible to co-/self-assess and to improve the procedures and ways of working together. During ShareLabs, discussions and negotiations, based on their production and their ideas, resonate in each designer, causing various interpretations through the use of post-it, words play, or other types of games helping them to share and to compare their points of view.

Agreeing on the relevance of a particular action helps to clarify the options of each member and encourages iteration, creativity and the emergence of innovative proposals. This way, participants extend their skills and possible fields of inquiry, while taking into account each other's views.

The members of the group evolve together in this way and jointly develop some agility in learning to change. As a catalyst of collective intelligence, the foundations of ShareLab involve:

- participation of all contributors in the project early on and the involvement of a team of researchers, helping the group to build the objectives of each phase together, specifying their activity;
- an approach that aims to be interventionist and scalable via each other's interactions;
- joint definition of tasks, procedures and tools, as well as a co-evaluation of these resources;
- taking into account the space / time necessary for exchanges and for the work to be achieved.

VII. Conclusion and Perspectives

A. Contributions

ShareLab makes it possible to offer a forum for all collaborators while helping them manage their specificities and differences in a very short design time. This approach is intended, firstly, to formalize a collaborative action approach capable of supporting collective intelligence within a work group. This feedback from players also allows the method to be improved, to reconsider it with respect to new application contexts and to help it grow by setting up a dialogue with other approaches simultaneously involving the object, the group and tools to support them in their activity. Although the absence of specific procedures associated with the establishment and definition of ShareLab could be regarded as a limitation, it can also be considered an advantage as it takes into account the specificity of each group, of each phase and of every possible iteration of the design process. We believe that this versatility allows collaborators to better meet their initial objective by taking into account the wide variety of parameters, contexts and approaches.

B. Limitations

While ShareLab reflects the emerging movement and dynamics of collective activity, it is also true that it alone cannot surmount the complexity of the project and of the context in which the players evolve. That is why it is imperative that this approach, even in a rationalized form, should not aim at too systemic a vision of production, at the risk of forgetting the contributors and their specificity in a group.

The motivation that characterizes all members of the group participating in the contest is the incorporation of all areas of expertise, cultures, experiences and views involved in the project. Yet isn't this motivation the cause of ShareLab's success? This parameter should undoubtedly be considered in the construction of intelligence within the group.

Furthermore, the feedback mentioned in this article has been explained on the basis of reports and video recordings of ShareLab conducted throughout the competition. This feedback could have been richer if it had been built on the basis of a reconstitution to participants after the final rendering; unfortunately, such was not feasible here.

C. Future work

A critical perspective, suggested in light of this feedback, is to support this method with the following stipulations:

- by confronting the protagonists with their own activities and by putting them face-to-face with their experiences and contradictions that had occurred in the process;
- by applying this method to other collective design contexts, such as design and engineering;
- by complementing it with other methods, such as analysis protocols or participatory approaches involving users in the study as well.

An epistemological framework is also needed to deal with the concepts that can define collective activity and manage its complexity. Faced with strong levels of inertia in work habits, the real challenge now is to implement strategies to better manage the group cohesion inside the design activity contradictions.

Acknowledgment

The LUCID thanks the eleven members of the TarTar team for their active participation in ShareLab.

References

[1] D. Fay and M. Frese, “Self-starting behavior at work: Toward a theory of personal initiative,” in Motivational psychology of human development: Developing motivation and motivating development, Elservier, Amsterdam, pp. 307-337, 2000.

[2] T. W. Malone, R. Laubacher, and C. N. Dellarocas, Harnessing Crowds: Mapping the Genome of Collective Intelligence,, MIT Center for Collective Intelligence Working Paper, Cambridge, 2009.

[3] J. F. Boujut and P. Laureillard, “A co-operation framework for product–process integration in engineering design,” in Design Studies 23 (6), pp. 497–513, 2002.

[4] M. Bourassa, R. Philion, R., and J. Chevalier, "L'analyse de construits, une co-construction de groupe", in Éducation et francophonie, 35(2), pp. 78-116, 2007.

[5] A. Dolbec, "La recherche-action, dans La recherche en éducation: étapes et approches," in CRP, Sherbrooke, pp. 181-208, 2004.

[6] J. Dubost and A. Lévy, "Recherche-action et intervention," in Vocabulaire de Psychosociologie, Références et positions J. Barus-Michel, E. Enriquez, and A. Lévy, Eds. Éres, Paris, pp. 391-416, 2003.

[7] C. Couture, N. Bernarz, and S. Barry, "Conclusion. Multiples regards sur la recherche participative, une lecture transversale," in La recherche participative, Multiples regards, M. Anadòn and L. Savoie-Zajc, Eds. PUQ, Québec, , pp. 205-221, 2007.

[8] W. J. Orlikowski, "Knowing in practice Enacting a collective capability," in Distributed organizing, Organisation Sciences, 13(3) pp. 249–273, 2002.

[9] F. T. T Phua, "Construction management research at the individual level of analysis: Current status, gaps and future directions," in Construction Management and Economics, 31 (2). pp. 167-179, 2013. ISSN 0144-6193.

[10] L. A. Dechurch and M. A. Marks, "Leadership in multiteam systems," in Journal of Applied Psychology, 91(2), 311, pp. 311-329, 2006.

[11] http://www.jacquesrougeriecompetition.com [retrieved: July, 2015]

[12] A. Senciuc, I. Pluchinotta, and S. Ben Rajeb, "Collective Intelligence Support Protocol: a systemic approach for collaborative architectural design," in Lecture Notes in Computer Science (in press), Springer, 2015.

[13] F. Darses, F. Detienne, P.Falzon, and W. Visser, "A method for analyzing collective design processes", in ECCE 10 : Confronting Reality, P. Wright, E. Hollnagel, and S. Dekker (Eds), Sweden, pp. 47-56, 2001.

[14] S. Ben Rajeb and P. Leclercq, "Instruments for Collective Design in a Professional Context: Digital Format or New Processes ?", in Proceedings of the 8th International Conference on Advances in Computer-Human Interactions, ACHI, Lisbon, pp. 72-79, 2015.

[15] Y. Engeström, et al., "Knotworking in academic libraries: Two case studies from the University of Helsinki," in Liber Quarterly, 21 (3/4), pp. 387-405, 2012.

[16] Y. Engeström, "From individual action to collective activity and back: Developmental Work Research as an interventionist methodology," in Workplace studies, P. Luff, J. Hindmarsh, and C. Heath, Eds. Cambridge, Cambridge University Press, pp.150-168, 2000.

[17] Y. Engeström, "Expansive learning at work: toward an activity theoretical reconceptualization," in Journal of Education and Work, 14 (1), pp. 133- 156, 2001.

[18] K. Kuutti, "Activity Theory as a Potential Framework for Human-Computer Interaction Research," in Nardy, pp. 17-44, 1996.

[19] Y. Engeström, R. Engeström, and T. Vähäaho, " When the center does not hold: the importance of knotworking," in Activity theory and social practice: Cultural-historical approaches, S. Chaiklin, M. Hedegaard, and U-J. Jensen, Eds. Aarhus, Aarhus University Press, pp.345-374, 1999.

[20] J. Virkkunen, R. Vilela, M. Querol, and M. Lopes, "The Change Laboratory as a tool for collaborative transforming work activities: an interview with Jaakko Virkkunen", in Saúde e Sociedade, 23(1), pp. 336-344, 2014.

[21] F. Détienne, G. Martin, and E. Lavigne, "Viewpoints in co-design: a field study in concurrent engineering," in Design Studies, 26, pp. 215-241, 2005.

[22] J. Roschelle and S. Teasley, "The construction of shared knowledge in collaborative problem solving," in Computer-supported collaborative learning, C. O'Malley , Eds. Berlin, Springer, pp. 69-97, 1995.

[23] M. Murer, M. Jacobsson, S. Skillgate, and P. Sundström, "Taking things apart: Reaching common ground and shared material understanding", in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 469-472, 2014.

[24] F. Cabitza, A. Cerroni, A., and C. Simone, C, "Knowledge artifacts within Knowing communities to foster Collective Knowledge", in Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, ACM, pp. 391-394, 2014.

[25] D. A. Schön, The reflective Practitioner: How professionals think in action, NY: Basic Books, New York, 1983.

[26] P. Laureillard, "The Role of Graphical Representations," in Inter- Professional Cooperation, Everyday engineering: An ethnography of design and innovation, D. Vinck, Eds. MIT Press, pp. 178-202, 2003.

[27] S. L. Star, "The structure of III-Structured Solutions: Heterogeneous problem Solving, Boundary Objects and Distributed Artificial Intelligence," in Distributed Artificial Intelligence, M. Huhns, Ed. San Francisco, Morgan Kaufmann Publishers Inc., Vol. 2, pp. 37-54, 1989.

[28] F. Belkadi, E. Bonjour, M. Camargo, N. Troussier, and B. Eynard, "A situation model to support awareness in collaborative design," in International Journal of Human-Computer Studies, Elsevier, 71 (1), pp.110-129, 2013.

[29] S. L. Jarvenpaa and D. E. Leidner, "Communication and Trust in Global Virtual Team", Organization Science, 10(6), pp.791–865, 1999.

[30] M. H. Hsu, T. L. Ju, C. H. Yen, and C. M. Chang, "Knowledge sharing behavior in virtual communities: the relationship between trust, self-efficacy and outcome expectations," in International Journal of Human-Computer Studies 65 (2), pp. 153-169, 2007.

[31] M. L. Maher, "Designers and collaborative virtual environments," in Collaborative Design in Virtual Environments, Intelligent Systems, Control and Automation: Science and Engineering, 48(1), pp. 3-15, 2011.

[32] J. M. Carroll, D.C. Neale, P. L. Isenhour, M. B. Rosson, and D. S. McCrickard, "Notification and awareness: synchronizing task-oriented collaborative activity," in International Journal Of Human-Computer Studies, 58, pp. 605-632, 2003.

[33] Y. Engeström, J.Virkkunen, M. Helle, J. Pihlaja, and R. Poikela, "The Change Laboratory As A Tool For Transforming Work," in Lifelong Learning in Europe, 1 (2), pp. 10 -17, 1996.

 ISBN: 978-1-61208-436-7

Agile Knowledge Networking

A novel approach to research collaboration between industry and academia

Bjørn Erik Munkvold
Department of Information Systems
University of Agder
Kristiansand, Norway
e-mail: bjorn.e.munkvold@uia.no

***Abstract*—The paper presents a novel approach to interaction between industry and academia, in the form of a knowledge networking program as implemented by a telecommunications company. Compared to traditional in-house R&D, this program represents an agile and "lightweight" approach to getting access to scientific knowledge supporting the company's quest for disruptive innovations. Supported by a team of student trainees, the program facilitates interaction with a large number of top academic scholars worldwide. The paper reports experiences and lessons learned from the program so far, and discusses areas of further improvement in terms of managing the collaborative R&D process and cultivating the networking capability of the company.**

Keywords-knowledge networking; industry-academic collaboration; networking capability; collaboration capability.

I. Introduction

The innovation capabilities of an organization are closely related to knowledge creation [1][2]. According to the open innovation concept, rather than developing new products and services through internal R&D processes, the knowledge needed for innovation increasingly resides outside the corporate boundaries [3]. This again places increasing emphasis on the organization's capability for managing external relations and partnerships, often referred to as networking capability [4] or collaboration capability [5]. Collaboration between industry and researchers in academia is an important source for disruptive innovations [6], but managing collaborative R&D projects between industry and academia also implies potential challenges related to cultural differences and goal incongruence [7].

Since the concept of knowledge networking (KN) was suggested in the late 1990s [8], organizations have experimented with different knowledge creation and transfer processes based on network interaction. Still, there is a need for more empirical studies that can contribute to identifying guidelines for innovation through knowledge networking and R&D collaboration.

In this paper, we present an example of an innovative and agile form of knowledge networking undertaken by the Elisa Corporation, a Finnish telecommunications company. Through Elisa Knowledge Networks, interaction with a broad set of leading scholars is facilitated for identifying and supporting disruptive innovations in the company. The concept of agility in this context refers to "a business-wide capability that embraces organizational structures, information systems, logistics processes and, in particular, mindsets" [9, p. 37], with flexibility as a key characteristic. We report on the experiences and lessons learned from the KN initiative so far, and discuss possible areas of improvement including the role of technology support.

The remaining paper is structured as follows. Section II gives a brief overview of related research, and Section III introduces the Elisa Knowledge Networks program. Section IV presents the methodological approach for this study, and Section V presents the key findings in the form of experiences and lessons learned. The findings are discussed in Section VI and Section VII presents conclusions and implications.

II. Related Research

A. Key Definitions

In this section, we present definitions of some key terms that are used in the discussion of the findings from our case study.

Mitegra et al. [4] define networking capability as "the set of activities and organizational routines which are implemented at the organizational level of the focal company to initiate, develop, and terminate business relationships for the benefit of the company" (p. 741). They further detail the concepts into three components referring to the initiation, development and termination of the relationships.

Blomqvist and Levy [5] present the somewhat broader term of collaboration capability for conceptualizing knowledge creation and collaborative innovation in networks. They define the term as "The actor's capability to build and manage network relationships based on mutual trust, communication and commitment" [5, p. 31]. This is presented as a concept for analyzing relational interaction on different levels, including individual, team, intra-organizational, and inter-organizational. Similar, the Global Collaboration Index Model presented by Frost and Sullivan [10] includes collaboration capability as "a forward-looking construct that represents an organization's orientation and infrastructure to collaborate". The organization's culture and structure and its application of collaborative technologies are defined as components of collaboration capability. Collaboration capability in itself facilitates the updating of

 ISBN: 978-1-61208-436-7

old capabilities, and the internal or external development of new ones. It could thus be seen as part of the firm's transformational capacity, i.e., its capacity to continually redefine its product portfolio on the basis of the technological opportunities created within it [5][11].

B. Managing R&D Innovation Processes

Several studies have focused on challenges in managing innovation processes and research collaboration [3][6][7][12]. Among the key challenges identified is providing adequate follow-through of the innovation process so that the mindset of open innovation is also implemented in the existing work routines and daily operations [3][6]. Also, for industry-university collaboration, several potential conflicts have been identified in terms of cultural differences, conflicts over IP rights, and different priorities and time horizons [7].

Almeida and Soares [13] also point to the challenges related to information and knowledge management in project-based R&D institutions, involving different disciplines, cultures and ways of working. They outline recommendations for a digital enterprise information management (EIM) infrastructure, combining Wikis, communication tools (e-mail, Skype) and a central content management system for preserving the project results. In a similar vein, a case study of knowledge networking practices in a large, multinational corporation [3] shows how the implementation of a collaboration platform integrating various social networking tools has been important for supporting the open innovation strategy and knowledge-sharing capabilities of the company.

In terms of suggested solutions for meeting the different challenges, Barnes et al. [7] present an extensive framework for managing collaborative R&D projects. Among the most important factors in this framework is the existence of a collaboration champion, defined as "an individual with great enthusiasm for and commitment to the venture, who is also influential and well-placed within the partner organization" [7, p. 399]. Also, a method of partner evaluation to ensure genuine interest and commitment, ensuring mutual benefit in terms of appropriate balance between academic objectives and industrial priorities, and continuity of personnel, have been identified as critical success factors [12].

III. ELISA KNOWLEDGE NETWORKS

Elisa is a telecommunications, ICT and online service company serving 2.3 million consumer, corporate and public administration organization customers. The company is the market leader in Finland in mobile subscriptions, and in 2014 it employed 4100 persons with a revenue of 1.54 billion euros [14].

The Knowledge Networks program was initiated in 2011, as an initiative by the company's Vice President of Business Development. The company earlier had an internal R&D unit of 14 employees, but the research activity was not seen to be sufficiently targeted towards the needs of the business units.

The Research Collaboration objective of the KN program is stated as to "identify novel disruptive innovations in the scientific community based on selected focus areas". A disruptive innovation is here defined as "A new technology, product, service or business model that will either disrupt the company's existing business or create a new business opportunity by introducing a new domain of offerings that will dwarf some of the existing offerings or totally replace them". The disruptive innovations may be related to four domains: business model, services, products and technology. The current list of disruptive innovations identified by Elisa Knowledge Networks through interviews with academic scholars includes: computer-assisted communication, healthcare co-creation, privacy control, industrial internet, software defined networks, smart device interconnections, and brain-machine interfacing.

The Knowledge Networks team is led by the Vice President of Business Development who has 20 % of his position allocated for this, and a team of 2-3 trainees who are master students recently graduated or in the final stage of their studies. In addition, the program has a steering group of four top level managers in the different business areas of the company, who meet once a month for status updates and approval of new academic contacts identified by the KN team. As all costs related to the initiated collaborative R&D projects are covered by the business units, the costs of the KN team operations only represent a small fraction of the costs for the former R&D unit in the company which amounted to more than 1.5 million euros.

The KN activities are organized in a 'funneling process' where candidate academic scholars are first identified through scouting by the trainees or from internal or external hints. The selected scholars are then invited for an online meeting, and in the case of mutual interest for further collaboration the scholars are suggested to the steering group for initiating collaboration with the business units. The criteria applied in the scouting process are that the scholar should be a world class scientist doing research in one or more of the company's focus areas, having a track record of industry collaboration, and being affiliated with a highly distinguished university or research centre. Further, although as a low priority criterion, the scholar should be well funded. This again is due to the "lightweight" nature of the KN program, where most of the research projects initiated are not funded by the company other than for covering expenses related to travel and data collection, etc.

The KN team develops regular performance reports for the Elisa executive board. In the period from 2011-2014, 743 academic scholars were contacted, of which 142 were approved for further collaboration. Of these, 85 were from institutions outside Finland. In the same period, more than 20 research collaboration projects were initiated.

The KN team uses a Microsoft Sharepoint database and Excel spreadsheets for storing information on the contact with scholars and the projects initiated. Employees in the company can get access to this information by request, but it is so far not made openly available. For online meetings with the scholars, they use the company's desktop video conferencing system. The KN network arranges internal seminars that can be accessed online, and that are also recorded for later view.

IV. Method

Data for this study was collected through interviews with different stakeholders involved in the knowledge networking activities in Elisa, and analysis of different documents from the KN network.

A total of twelve semi-structured interviews with sixteen informants were conducted in the period from May to September 2014. These included:

- Individual interviews with four members of the KN steering committee (including one online interview)
- Group interview with three trainees in the KN team
- Individual interviews with four representatives from the company's business units (including two online interviews)
- Online interviews with five academic scholars in the KN network (including a group interview with three scholars)

Eight of the informants were interviewed at the Elisa headquarters in Helsinki during a one day visit, while the remaining interviews were conducted using the company's desktop conferencing system for the internal employees and Skype for the academic scholars.

All interviews were taped and transcribed in full. The interview transcripts were annotated, and then analyzed for experiences and lessons learned from the KN activities till date.

A preliminary report of the results were also discussed and validated with the KN coordinator and two KN trainees in a meeting in March 2015, then also providing an update on current actions in the KN team.

V. Key findings

In this section, we present key findings from our study in terms of experiences and lessons learned from the KN program so far, and possible areas of improvement. As presented in Section IV, data for this study was collected through interviews with different stakeholders involved in the knowledge networking activities, and analysis of different documents from the KN network. The reported experiences and suggestions thus represent the perspectives from different stakeholder groups.

A. Experienced Benefits from the KN Program

Overall, the company representatives interviewed report positive experiences from the KN program. They regard this way of getting access to cutting edge research as more effective than through the former in-house R&D department, and delivering more benefit to the business units.

While the company does not currently have metrics in place for analyzing the outcomes of the research projects initiated, the informants point to several examples of successful interactions with academia that have led to important input to the company's strategy development in terms of areas to be focused. Several of the informants also point to that even if most contacts with the academic scholars do not lead to any further collaboration it is still valuable to read about and learn from their research work, and that the research articles often give more insight than superficial consulting reports. Also, several of the informants state that it is to be expected that not all contact initiatives give results:

"*In venture capital operation there is the golden rule that based on ten ideas or ten investments there will be one successful and nine failures, and that is very ok. And I think very much this kind of knowledge network operation is like an internal venture capital operation, they are ventures, there is a high risk because they are future-oriented.*"

(Member of KN steering group)

They also point to that the KN initiative in the company can be seen as part of transforming the company to become a more agile and international organization.

The academic scholars interviewed also state that being invited to research collaboration by an industrial company is exceptional, as it is normally the other way round. The initial invitation email is also well prepared in that it refers explicitly to some of the researcher's work, thus managing to make the candidate scholar curious (and perhaps also flattered) and therefore (s)he does not discard this as spam. The willingness of the company to share their data (from surveys, etc.) with researchers and students for further analysis is also emphasized as positive. As an example, the company is currently inviting scholars to conduct big data research on raw data from selected base stations in their mobile network.

The trainee program is also emphasized as a success, with the trainees being praised by both the company informants and the scholars as premium students who are effective in their approach. However, some suggestions were made regarding the length and format of the trainee period, which will be reported in the next section.

B. Challenges and Areas of Improvement

Overall, the informants point to that the KN program is still in an early, ramp-up phase, and that some initial challenges thus can be expected.

The key challenge reported by both the members of the steering group and the business unit representatives concerns the hand-over of academic contacts from the KN team to the business units. It is here considered crucial to create sufficient ownership of the research projects from the business. If not, with the hectic work pace, handling the introduction of the research contact may simply be regarded as extra work from the business persons. As reflected by a business representative on how the hand-over process could sometimes be perceived by the business unit:

"[…] *"this is a good researcher, catch!" (laughter) "Ah, what do I do with this guy? Now I need to use five hours a week to keep it going". And that is often reason enough to say no".*

One of the informants also pointed to the "not invented here" syndrome as a barrier towards taking on the responsibility for new research projects that they have not themselves initiated.

Some argued that the trainees could take on a stronger coordinating function for the research collaboration projects, and thus support the business units in this process. But a challenge to this is the relatively short duration of the trainee program, typically lasting 3-4 months. Several informants

 ISBN: 978-1-61208-436-7

thus expressed concern that during this relatively short period of time, the trainees did not get the possibility to take on more challenging tasks beyond administering the contact with the scholars. However, based on his experience, one business representative also pointed out that one should be careful about delegating too challenging tasks to the trainees, such as negotiating contracts. It should here be noted that since the time of this interview study, the length of the trainee period has been somewhat extended, to four months as KN trainee followed by four months as a trainee in one of the business units.

The turnover of trainees was also raised as a challenge by the academic scholars interviewed. As projects could typically span 1-2 years, this would imply contact with 4-5 trainees during the course of the project. While the trainees were generally perceived to be well prepared, this was seen to result in a lack of continuity and to cause some confusion about who was the current point of contact in the company. The scholars also expressed some frustration with the process of establishing the project agreement taking too long, resulting in delay in the project schedule which again could lead to conflicts with other commitments and deadlines for the academics. This is also supported by one of the company representatives, who expressed concern that the company could lose face towards the scholars.

The fact that most of the projects were unfunded was also stated to affect the scholars' priority and commitment:

"*Because they are not a client, they are more like a beneficiary, so I think then it changes a little bit the power structures. Because we don't really owe them too much. I mean, of course we want to help them, but obviously they are not paying, and we are doing the work, so…*".

This was seen in contrast to the rather strict liability clause enforced by the company in the project agreement, involving a penalty of 50.000 euros in the case of any confidentiality breaches.

Also, some of the informants in the company stated that unfunded projects did not tend to receive the same focus and expectations as funded projects:

"*When you pay for something the quality is usually a bit better, and you get committed a lot more than to something which is for free*".

However, the KN coordinator still argues in favor of non-funded projects as the norm, in line with the lightweight nature of the program and regarding the contact with the scholars as the main focus rather than the projects as such.

The general impression among the informants is that the KN program is not yet very visible in the company. This is despite regular online presentations and workshops conducted. However, the KN team is currently working on improving this, through establishing a dedicated site on the company intranet and also running monthly online seminars. A challenge regarding intranet presence was stated to be that the company has too many sites, thus making it difficult to get an overview.

VI. DISCUSSION

Through their Knowledge Networks program, Elisa has transformed their R&D activities from an internal, resource-demanding operation not perceived to fully meet the demands of the business units, to a lightweight and agile operation facilitating flexible interactions with a large number of world-leading researchers in different domains. As pointed to by the informants, this can be seen as part of an overall transformation towards a more internationally oriented company.

While it is still early to measure the output of the KN program, the accounts provided by the informants of projects and ideas initiated through the contact with the research scholars indicate that the program fulfils the overall intention of this partnership, i.e., to engage in research collaborations that could not otherwise be justified in-house [11]. Also, the Elisa employees interviewed were generally positive about the KN program, characterizing it as a "valuable asset" for the company. The scholars were somewhat more mixed in their feedback, pointing out some challenges related to contractual arrangements and project management. Still, they were also positive towards the partnership program.

In comparison with the guidelines for managing R&D collaboration suggested in the literature [7], several of the critical success factors are in place in the KN program. The KN coordinator definitely serves the role of collaboration champion, in terms of commitment to the program. And as a Vice President, he also meets the definitional criteria of being influential and well-placed within the company [7]. The only potential challenge related to this role is the high dependency on one person, making it somewhat vulnerable. In effect of this, the company has now decided to dedicate one more person to the KN team, to assist the coordinator with some of his tasks.

Further, the process for scouting and selecting scholars works well, and the KN team has succeeded in engaging a large number of researchers who bring expert knowledge on focus areas related to potential disruptive innovations. However, as pointed out by the informants, timing is a critical issue, as even though a scholar may possess relevant knowledge for the company there may not be sufficient basis to initiate a project at this exact time. While the informants still regard the academic contact to be of possible value regardless of this, for the trainees in the KN team this represents more of a challenge in 'rejecting' a scholar after the initial contact. This also makes it difficult to contact this scholar again later.

According to the conceptualization of networking capability by Mitegra et al. [4], the Elisa Knowledge Networks program can be seen to demonstrate good relationship initiation capability in terms of attracting valuable partners. But in terms of relationship development capability, the program still experiences some of the challenges frequently mentioned related to establishing effective R&D partnerships. The main challenge here is related to the hand-over of the academic contact from the KN team to the internal business unit, and ensuring ownership in this. If sufficient follow-through is not provided, time pressure and the "not invented here" syndrome may act as barriers towards the further engagement of the business units. As the internal visibility of the KN program was still considered low, this also represents a challenge in creating buy-in for the research collaboration initiatives.

Also in managing the relationship with the scholars, some concerns were raised by both the scholars and company employees about the process of initiating the projects taking too long, and the company not being 'professional' enough in their handling of the scholars. The turnover of trainees was also pointed to by the academics as a challenge in maintaining a focal point of contact in the company. Given that most of the research projects are not funded by the company, these challenges could easily lead to some demotivation among the researchers or at least to a lack of prioritization.

As documented in previous studies, an ICT collaboration infrastructure is regarded an important element for supporting knowledge networking and open innovation [3, 13]. The KN team also uses several tools in for managing information and communicating with scholars, e.g., SharePoint databases and desktop videoconferencing. However, a recommendation would be to make the database of scholars and projects searchable for all employees. Even if this may not be considered useful for all, marketing this opportunity can still contribute to the internal awareness of the KN program and to facilitating an open innovation mindset among the business units [3]. Further, there is a potential for increasing use of social networking tools, facilitating knowledge exchange among the Elisa employees. Integrating this in an enterprise information management infrastructure can facilitate organizational learning through sharing results and best practices across the research projects [13].

With reference to the collaboration capability construct, the KN program in Elisa can be considered strong in its orientation and organizational infrastructure for collaboration, while the application of collaboration technologies is still at an early stage.

Finally, it should be reminded that the KN program is still in an early stage of development, and that it is yet too early to document extensive output measures. Also, several of the issues raised here are currently being addressed by the company, such as routines for project initiation and follow-through, more systematic assessment of completed and disbanded projects, extended trainee period, stronger presence on the company intranet, and use of social media (Twitter) for informing about projects.

In terms of further development of the KN program, a possibility could be to extend the current dyadic relationships between Elisa and each academic scholar to a real network, by connecting scholars with similar or complimentary research interests and skills to form research teams focusing on joint topics. This could then even further support an open innovation strategy for the company. However, this would also entail additional coordination challenges, and is so far not part of the company's further plans.

VII. Conclusion and Implications

The case study presented in this paper illustrates an agile approach to knowledge networking that has demonstrated several benefits over the former in-house R&D organization. The cost effective and flexible KN operation has facilitated contact with a large number of leading academic researchers, resulting in important insights for the company and over 20 collaborative research projects initiated so far. In addition, the KN operation includes a trainee program that gives valuable experience for master student graduates, also resulting in further employment in the company for several candidates.

The study has also pointed out several challenges in the management of the collaborative R&D process, especially regarding the hand-over of the academic contact from the KN team to the business unit to create internal ownership of the projects. Further, effective follow-through and coordination of the research partnership is important for ensuring continued commitment from the partners.

The study supports the findings from previous research on the importance of a collaboration champion, and balancing industry and academic objectives.

With reference to the concept of networking capability, the Elisa Knowledge Networks demonstrate strong relationship initiation capability but can still improve further its relationship development capability. In terms of collaboration capability, the KN program scores high on orientation towards collaboration but does have a potential for further utilization of collaboration technology and social software to support the knowledge networking. It is also recommended that the company provides shared access to their KN database, to facilitate extended knowledge sharing in the company.

The findings reported in this paper can serve as inspiration and benchmark for other companies seeking to develop their knowledge networking capability in partnership with academia. However, as a caveat it should be noted that this form of collaborative R&D partnership requires that influential persons in the company take a genuine interest in the potential contribution of academic research, as is the case with the Elisa KN coordinator and the members of the steering group. Further, the industry partner should be willing and interested in sharing their experiences and data to support the research activities, of which this article represents an example.

Further research should conduct more systematic assessment of the results from the R&D collaboration for both the company and the academic partners, and contribute to develop suitable metrics for this. Also, in-depth studies of how different collaboration technologies can support knowledge networking activities are needed to develop this practice further.

Acknowledgment

I am grateful to the Elisa Knowledge Networks coordinator Kimmo Pentikäinen and his team for facilitating and participating in the study, and to the Elisa employees and the academic scholars interviewed for sharing their insight and experience.

References

[1] S. Popadiuok and C. W. Choo, "Innovation and knowledge creation: How are these concepts related?," International Journal of Information Management, vol. 26, pp. 302-312, 2006.

[2] D. Esterhuizen, C. S. L. Schutte, and A. S. A. du Toit, "Knowledge creation processes as critical enablers for innovation," International Journal of Information Management, vol. 32, pp. 354-364, 2012.

[3] D. Schneckenberg, "Open innovation and knowledge networking in a multinational corporation," Journal of Business Strategy, vol. 36, no. 1, pp. 14-24, 2015.

[4] M. Mitrega, S. Forkmann, C. Ramos, and S. C. Hennenerg, «Networking capability in business relationships – Concept and scale development,» Industrial Marketing Management, vol. 41, pp. 739-751, 2012.

[5] K. Blomqvist and J. Levy, "Collaboration capability – A focal Concept in Knowledge Creation and Collaborative Innovation in Networks," International Journal of Management Concepts and Philosophy, vol. 2, no. 1, pp. 31-48, 2006.

[6] M. Assink, "Inhibitors of disruptive innovation capability: a conceptual model," European Journal of Innovation Management, vol. 9, no. 2, pp. 215-233, 2006.

[7] T. A. Barnes, I. R. Pashby, and A, M, Gibbons, «Managing collaborative R&D projects: development of a practical management tool," International Journal of Project Management, vol. 24, pp. 395-404, 2006.

[8] A. Seufert, G. von Krogh, and A. Bach, "Towards knowledge networking," Journal of Knowledge Management, vol. 3, no. 3, pp.180-190, 1999.

[9] M. Christopher, "The Agile Supply Chain: Competing in Volatile Markets," Industrial Marketing Management, vol. 29, no. 1, pp. 37-44, 2000.

[10] Frost and Sullivan, "Meetings Around the World: The Impact of Collaboration on Business Performance," Frost & Sullivan White Paper, Silicon Valley, Palo Alto, CA.

[11] B. B. Tyler, "The Complementarity of Cooperative and Technological Competencies. A Resource-Based Perspective," Journal of English Technology Management, vol. 18, pp. 1-27, 2001.

[12] T. Barnes, I. Pashby, and A. Gibbons, "Effective University-Industry Interaction: A Multi-case Evaluation of Collaborative R&D Projects," European Management Journal, vol. 20, no. 3, pp. 272-285, 2002.

[13] M. V. Almeida and A. L. Soares, "Knowledge sharing in project-based organizations: Overcoming the informational limbo," International Journal of Information Management, vol. 34, pp. 770-779, 2014.

[14] corporate.elisa.com

Exploiting Argumentation Content and Structure to Advance Collaboration through Hybrid Recommendations

Spyros Christodoulou and Nikos Karacapilidis
University of Patras
Greece
e-mail: shristod@cti.gr; nikos@mech.upatras.gr

Costas Pappis
University of Piraeus
Greece
e-mail: pappis@unipi.gr

***Abstract*—Contemporary collaborative environments involve a flood of collected and exchanged data and require advanced techniques to enhance data processing, allow data transformation in actionable insights and reduce the subsequent cognitive overhead. In line with these requirements, this paper presents a hybrid recommender engine that builds on the synergy of content-based and collaborative filtering techniques to provide recommendations in argumentative collaboration settings. The proposed engine has been integrated in a web-based collaboration support system and exploits the content and structure of the underlying argumentation. Through a scenario of use, we demonstrate the application of our approach and discuss its usefulness in terms of advancing collaboration and augmenting the quality of decision making.**

Keywords-collaboration; argumentation; decision making; hybrid recommendations.

I. INTRODUCTION

Current data-intensive collaboration and decision making settings require efficient and effective techniques that provide personalized support, enhance the collaboration process and, ultimately, improve the quality and accuracy of the decisions to be made [1]. In this direction, recommender (or recommendation) systems [2], a type of information filtering systems that focus on predicting user responses to options, aim to assist users in processing large amounts of information, by reducing the subsequent cognitive overhead and supporting their decision making tasks [3]. Recommender systems have been proven to be valuable for coping with information overload and have become one of the most powerful and popular tools in diverse areas. Consequently, many applications have integrated recommendation techniques to provide users with helpful suggestions.

A variety of recommendation techniques have been already proposed, each one having certain strengths and weaknesses [4]. Besides, much attention is being lately paid to the exploitation of argumentation towards offering more valid suggestions. Argument-based recommender systems [5], as these tools are usually referred in the literature, are tools aiming to better support users by providing recommendations on the basis of associated arguments. For instance, a prototype of a group argumentation support system that applies frame-based information structure and argumentation to support group decision task generation and identification is presented in [6]; an approach to enhance practical reasoning capabilities of recommender system technology by incorporating argument-based qualitative inference is proposed in [7]; finally, ArgueNet [8] was designed as a recommender system based on a defeasible argumentation framework to classify Web search results according to preference criteria that have been declaratively specified by the user.

In line with the above, this paper presents a hybrid recommender engine that builds on the synergy of content-based and collaborative filtering techniques. The novelty of our approach lies in its meaningful exploitation of the content and structure of an ongoing argumentation in order to provide actionable recommendations. The approach presented in this paper assumes that the collaboration taking place adheres to a classic formal argumentation model, namely *Issue-Based Information System (IBIS)* [9]. Adopting IBIS, an ongoing collaboration is structured as a graph, whose basic elements are *issues* (questions to be answered), each of which are associated with alternative *positions* (possible answers); in turn, these are associated with *arguments* which support or object to a given position or another argument. In any case, the approach described in this paper can be easily adjusted to accommodate alternative argumentation models.

The remainder of this paper is structured as follows: Section II reports on related work in the area of recommender systems. Sections III and IV present in detail the proposed hybrid recommender engine and its integration with an already implemented collaboration support system that adopts the abovementioned model. Through an illustrative example scenario, Section V demonstrates how the recommendations produced by the proposed engine may advance an ongoing collaboration and enhance the quality of collective decision making. Section VI concludes the paper and discusses related remarks.

II. RECOMMENDER SYSTEMS

A recommender system can be viewed as a personalized information agent aiming to assist the natural social process of making choices (suggestions on items a user is likely to be interested in) without sufficient personal experience of the existing alternatives. The development of these systems has been based on diverse techniques, which can be classified in four main categories [10]: (i) *collaborative:* the generated item recommendations for a specific user are based on items rated by other “similar” users; (ii) *content-based:* recommendations for a specific user are generated according to each item’s features and the user’s preferences (i.e., the aggregation of items the user likes or dislikes); (iii) *knowledge-based:* recommendations follow the inferences about one’s needs and preferences, and (iv) *demographic:* the demographic profile of the user is exploited to provide recommendations.

Even though a great number of recommender systems belonging to the above categories (referred as “simple”) have evolved since the mid 90’s, all “simple” recommender techniques have certain strengths and weaknesses [11]. For

 ISBN: 978-1-61208-436-7

instance, all "learning-based" techniques (i.e., collaborative, content-based and demographic) suffer from the "cold start" problem (i.e., the difficulty in handling new items or new users); the collaborative and content-based techniques suffer from the "portfolio effect" (i.e., an item similar to an item that a particular user has rated before would be never recommended to that user).

Hybrid recommendation approaches try to mitigate the above drawbacks and, at the same time, exploit the advantages of "simple" recommendation techniques by combining two or more of them in a uniform approach. Depending on the particular method applied to combine the "simple" recommendation techniques, hybrid recommender systems may be classified in seven main categories [11]: (i) *weighted:* each item gets a number of partial scores (as many as the number of the "simple" recommendation techniques) reflecting the value of this item with respect to each recommendation technique. The total item score results from the linear combination of the partial scores (weights are used to state the importance of a "simple" recommender technique over another); (ii) *switching:* based on the evaluation of the recommendation situation, the system selects among a number of "simple" recommender techniques to apply. The selection of a reliable criterion to conduct this method is a critical task and remains an open research issue [12]; (iii) *mixed:* the output of two or more recommendation techniques is presented and it is up to the user to select the best items among the different items' lists returned; (iv) *feature combination:* features of one source are injected into an algorithm that was initially designed to perform data processing of a different source; (v) *feature augmentation:* a recommendation technique is applied to extract a number of features, which are then used as input to another recommendation technique; (vi) *cascade:* a "weak" recommendation technique is applied to refine (but not overturn) the decisions made by a "strong" recommendation technique; (vii) *meta-level:* the model resulting from one recommendation technique is used as input to another.

As described in detail in the next section, our approach integrates the collaborative and the content-based filtering techniques by adopting the switching method.

III. A HYBRID RECOMMENDATION ENGINE

A. *The need for recommendations in a collaboration setting*

In a data-intensive and cognitively-complex argumentative collaboration setting, users often need help in spotting those parts of an ongoing argumentation that can really advance collaboration and augment the quality of decision making. In such settings, a recommender engine could enable users in:

- locating already existing argumentation items that are similar to a new item they have just contributed to an ongoing collaboration; such recommendations may trigger the creation of meaningful interrelations between the new item and the existing ones;
- spotting users with similar profiles, in order to catch up with their argumentation items;
- tracking popular argumentation items, which receive much attention and may influence the evolution of the collaboration;
- gaining insights about the probable outcome of the collaboration.

In the context of an argumentative collaboration support system, an efficient recommender engine should not only rely on the content of the collaboration; it should also exploit the structure of the associated discourse graphs that involve multiple stakeholders. Such a hybrid approach is described in the following, where a content-based recommender exploits features of specific collaboration items, while a collaborative filtering recommender considers the users' rating profiles and the total structure of the argumentative discourses to generate meaningful and helpful recommendations (hereafter, the terms 'collaboration item' and 'argumentation item' are used interchangeably).

B. *Content-based recommendations*

Generally speaking, content-based recommender systems rely on the users' rating profiles to provide recommendations; items sharing similar features with the items a particular user has liked in the past are recommended to the user [13]. In the context of an argumentative collaboration support system, the proposed procedure of providing users with content-based recommendations breaks up into two distinct tasks: (i) calculating a rating profile for each user, and (ii) spotting similar collaboration items (with compatible features) to each user's rating profile.

With respect to the first task, a user's *Z* rating profile *RP(Z)* is defined as the set of all collaboration items rated by her. Collaboration items that have not been rated by user *Z* are not included in *RP(Z)*. As far as the second task is concerned, spotting similar collaboration items to a user's *Z* rating profile requires comparing each argumentation item of the collaboration space with each argumentation item included in *RP(Z)* to decide about their *degree of similarity*. As a prerequisite, we need a definition of an appropriate *degree of similarity DoS(x,y)* function to reflect how similar two collaboration items *x* and *y* are. As we focus on content-based recommendations in this step, *DoS(x,y)* should be based on items' *x* and *y* contents (i.e., their titles and bodies).

`MoreLikeThisHandler` (from the Apache Solr open source library, `http://lucene.apache.org/solr/`) offers a suitable to our purposes implementation of a *DoS(x,y)* function to compare two documents and decide on their degree of similarity. According to it, *DoS(x,y)* corresponds to an increasing function (i.e., the more similar two documents *x* and *y* are, the larger the *DoS(x,y)* value is) and can be easily applied to perform all the necessary comparisons between pairs of argumentation items.

As our basic target is to spot similar argumentation items to a user's *Z* rating profile *RP(Z)*, we have to compare each argumentation item *x* in the collaboration space with each argumentation item *y* in *RP(Z)* and calculate the respective *DoS(x,y)* values (excluding the items the user *Z* has rated). To decide about how similar an argumentation item *x* is to the set of argumentation items of *RP(Z)*, we define the cumulative degree of similarity *CDoS(x,Z)* of an argumentation item *x* to the rating profile of user *Z* as:

$$CDoS(x,Z) = \sum_{y \in RP(Z)} \mathrm{DoS}(x, y) \quad (1)$$

Taking into account that *DoS(x,y)* is an increasing function, the larger the *CDoS(x,Z)* is, the more similar an argumentation item *x* is to the rating profile of user *Z*. The calculation of *CDoS(x,Z)* for each item *x* and user *Z* is straightforward (by using Eq. (1)). The argumentation items recommended to user *Z* are the ones with the larger values of *CDoS(x,Z)*.

C. Collaborative filtering based recommendations

As already stated, the central idea of collaborative filtering is to provide a user with recommendations based on the rating history of similar users (i.e., users with similar rating profiles to the active user). In such systems, the recommendation procedure involves two major steps. The first step involves the construction of the utility matrix containing, for each user-item pair, a value that represents what is known about the degree of approval of that user for that item. The respective values reflecting the degree of approval either come from an ordered set or are scalar. Most entries of the utility matrix are usually unknown, i.e., we have no explicit information concerning the users' approval on the full set of items.

In our approach, the utility matrix and the related degrees of approval for each *(item_x, user_Z)* pair are calculated by taking into account two parameters:

- *User's Z rating on the argumentation item x* (denoted as *R(x,Z)*). We assume that a user *Z* may rate each argumentation item using the 1-5 stars rating scale.
- *The argumentation approval score* (denoted as *AAS(x,Z)*) reflecting a user's Z approval of a particular item *x* (as this approval has been expressed through the argumentation process).

We consider that *AAS(x,Z)* is directly related to the number, type (i.e., in favour or against) and structure of arguments that are linked to the specific item *x* (taking into account only the argumentation items put forward by user *Z*). Intuitively, a large number of arguments (created by user *Z*) in favour of a specific argument *x* expresses a larger approval (concerning user *Z*) on item *x* than a small number of arguments in favour of it. What is needed at this point is a method to measure how supportive (or adverse) to a specific item *x* the arguments posed by user *Z* are.

In the direction of assessing user's *Z* collaboration attitude on item *x*, we define a *user's Z argumentation graph for item x* (denoted as *G(x,Z)*) as the aggregation of all paths (denoted as *p(x,Z)*) leading to the item *x*, under the condition that all paths have been created by *Z*:

$$G(x,Z) = \bigcup p(x, Z)$$

In other words, *G(x,Z)* results from pruning the argumentation graph by removing:

- all relations of the argumentation graph that have been not created by *Z*, and
- all "isolated" items (i.e., items not belonging to any *p(x,Z)*).

As results from the above, the value of *AAS(x,Z)* is directly related to the evaluation of the associated argumentation graph. To calculate the value of the argument on the root of the argumentation graph, we follow the "global" approach (tuple-based valuation [14]) stating that the value of an argument is equal to the algebraic summation of the corresponding argumentation scores of argument paths leading to this argument. The argumentation score of an argument path is directly related to the number of *"against"* and *"in favour"* arguments forming the argument path. If e_A is an "against" relation and $|e_A(p(x,Z))|$ is the number of "against" relations along an argument path *p(x,Z)*, then, following the abovementioned "global" approach, the argumentation approval score *AAS(x,Z)* is:

$$AAS(x,Z) = \sum_{\substack{p(x,Z) \in G(x,Z) \\ e_A \in p(x,Z)}} (-1)^{|e_A(p(x,Z))|} \quad (2)$$

The degree of approval of a user *Z* on a collaboration item *x*, denoted as *DA(x,Z)*, is calculated by combining the two partial scores (user's *Z* rating on item *x*, *R(x,Z,)* and the corresponding argumentation approval score *AAS(x,Z)*). It is:

$$DA(x,Z) = a_1 * R(x,Z) + a_2 * AAS(x,Z),$$

where a_1 and a_2 are user-defined weights to reflect the relative importance of the two associated scores.

The utility matrix results from the calculation of *DA(x,Z)* for each *(item_x, user_Z)* pair. After calculating the utility matrix, the second step includes feeding a collaborative filtering based recommender with all degrees of approval, so as the implemented algorithm to provide recommendations on demand.

For the collaborative filtering algorithm, a modified version of an *Alternating Least Squares* algorithm to factor matrices has been integrated in our approach. An implementation of this algorithm is offered in the open source `Myrrix` recommender engine [15] (which is currently part of the `Oryx` open source project, see details at: `https://github.com/cloudera/oryx`). According to Myrrix creators, the implemented recommender engine is based on large matrix factorization, tries to learn a small number of features in order to explain users' and items' observed interactions, is nearly immune to the "cold start" problem and can provide quality recommendations for very new users or items.

D. Hybrid recommendations

In the settings under consideration, it would be expected that embedding a collaborative filtering based recommender would be enough to provide effective recommendations. However, especially in the early stages of a collaboration process, the limited users' contribution (in terms of the number of argumentation items added, the number of the relationships created and the ratings of the above items) may not be able to provide accurate recommendations. In such a case, where the scores of the collaborative filtering based recommendations provided are pretty close, content-based recommendations are also exploited to discretize among recommendations of which the value is almost equal.

We follow a cascade hybrid recommender approach to return the appropriate list of recommendations to a user. The proposed hybrid recommender includes the following steps (note that *score(Ri, List1)* is a function returning the score of the *i*-th recommendation on *List1* and *T* is a user-defined parameter):

```
1. Apply the content recommender technique and get
   the top-N1 content recommendations (List1)
2. Apply the collaborative recommender technique and
   get the top-N2 collaborative recommendations
   (List2)
3. Parse List1 and compare each two recommendations
   (Ri, Rj) on List1
4. If (score(Ri, List1) - score(Rj, List1)) < T)
       If (score(Ri, List2)- score(Rj, List2)) > T)
              Interchange(Ri, Rj) on List1
5. Repeat Step 4 until no interchange on List1 has
   taken place.
6. Return List1
```

IV. THE DICODE COLLABORATION SUPPORT SYSTEM

The proposed recommendation engine has been fully integrated in a web-based collaboration and decision making support system, namely *Dicode*, and exploits the content and structure of the underlying argumentation. Dicode follows an IBIS-like argumentation model and aims to augment collaboration in diverse data-intensive and cognitively-complex settings [16][17]. To do so, it builds on prominent high-performance computing paradigms and large scale data processing technologies to meaningfully search, analyse and aggregate data existing in diverse, extremely large, and rapidly evolving sources. The Dicode approach brings together the reasoning capabilities of the machine and the humans and enables the meaningful incorporation and orchestration of a set of interoperable web services to reduce the data-intensiveness and complexity overload in collaborative decision making settings.

In particular, the aim of *Dicode's collaboration and decision making services* is twofold: (i) to exploit the reasoning abilities of humans through the creation, management and use of innovative workspaces that augment synchronous and asynchronous collaboration, and (ii) to intelligently support stakeholders in decision making activities by enabling the use and exploitation of appropriate reasoning mechanisms. These services build on an appropriate formalization of the collaboration and exploit a series of reasoning mechanisms to support stakeholders in their daily decision making processes. Dicode implements alternative visualizations of the collaboration space (called *"views"*), each one offering a varying degree of formality.

In the context of this paper, we will focus on the *"mind-map view"* of the collaboration workspace. In this view (Figure 1), the collaboration workspace is displayed as a mind map, where users can upload and interrelate diverse types of items. This view uses a spatial metaphor to organize items, allowing users to select and freely move around any item. Item types supported include *ideas*, *notes* and *comments*. Ideas stand for items that deserve further exploitation; they may correspond to an alternative solution to the issue under consideration. Notes are generally considered as items expressing one's knowledge about the overall issue. Finally, comments are items that usually express less strong statements and may include some explanatory text or potentially useful information. Users can customize the set of available item types by creating additional ones, thus better annotating a particular collaboration workspace.

Two collaboration items can be explicitly connected using directed edges (relations). Visual cues are used to express semantics: for instance, a user may appropriately choose the width and colour of an edge to express a specific semantic relationship between two collaboration items (e.g., a red edge denotes an "against" relation, a green one stands for an "in favour" relation). Additional functionalities offered include the creation of adornments (a grouping mechanism to aggregate items related to a particular alternative), a "like/dislike" mechanism to express a user's acceptance/rejection concerning a collaboration item, rating of collaboration items, calculation of workspace statistics, and a replay mechanism that helps a user review the evolution of a workspace over time. The mind-map view builds on the reasoning capabilities of humans to support ease-of-use and expressiveness, as well as individual and group sense-making, by supporting stakeholders in locating, retrieving and meaningfully interacting with relevant information; moreover, in monitoring and comprehending the evolution of collaboration.

V. SCENARIO OF USE

To better illustrate the proposed approach (and its integration in the Dicode system), this section presents an illustrative real-world scenario from the area of prostate cancer research. A physician (George), an urologist (John) and a biomedical researcher (Jane) aim to investigate which is the best alternative treatment for the prostate cancer. Initially, they set up a Dicode collaboration workspace and start using it in the mind-map view (Figure 1).

John suggests that one of the best and most popular treatments for the prostate cancer (Figure 1(a)) is the "active surveillance". He adds an alternative to make his statement (Figure 1(a)). Jane is not in favour of this option, because it requires close monitoring (regular digital rectal exams, PSA tests, and prostate biopsy) to monitor for signs of progression, so she adds her "against" position on the collaboration workspace (Figure 1(b)). Contrary to Jane, George supports John's opinion ("in favour" position supporting the alternative suggested by John (Figure 1(c)), in the sense that active surveillance avoids site effects from radiation therapy or prostatectomy. On the other hand, he is skeptical as with Active Surveillance there is no post-treatment staging information ("against" position - Figure 1(d)).

Jane argues that "Brachytherapy" has been also used to treat tumors in many body sites and this could be one option (Figure 1(e)). One of its major advantages is that this procedure does not need hospitalization ("in favor" position, Figure 1(f)) and, furthermore, there are no surgical risks involved. John is not convinced by her arguments as Brachytherapy requires close monitoring ("against" position, (Figure 1(g)), which may even include hospital visits. To support his consideration against the Brachytherapy, John denotes that there is no post-treatment staging information which is also an important factor ("against" position, (Figure 1(h)).

George argues that the best alternative, in his opinion, is "radical prostatectomy" (Figure 1(i)) as it is quite common with very good results. John is in favour of this option ("in favour" position) as this solution is proven to reduce prostate cancer death rates (Figure 1(j)). Moreover, the removed tissue allows accurate stating (Figure 1(k) – "in favour" position),

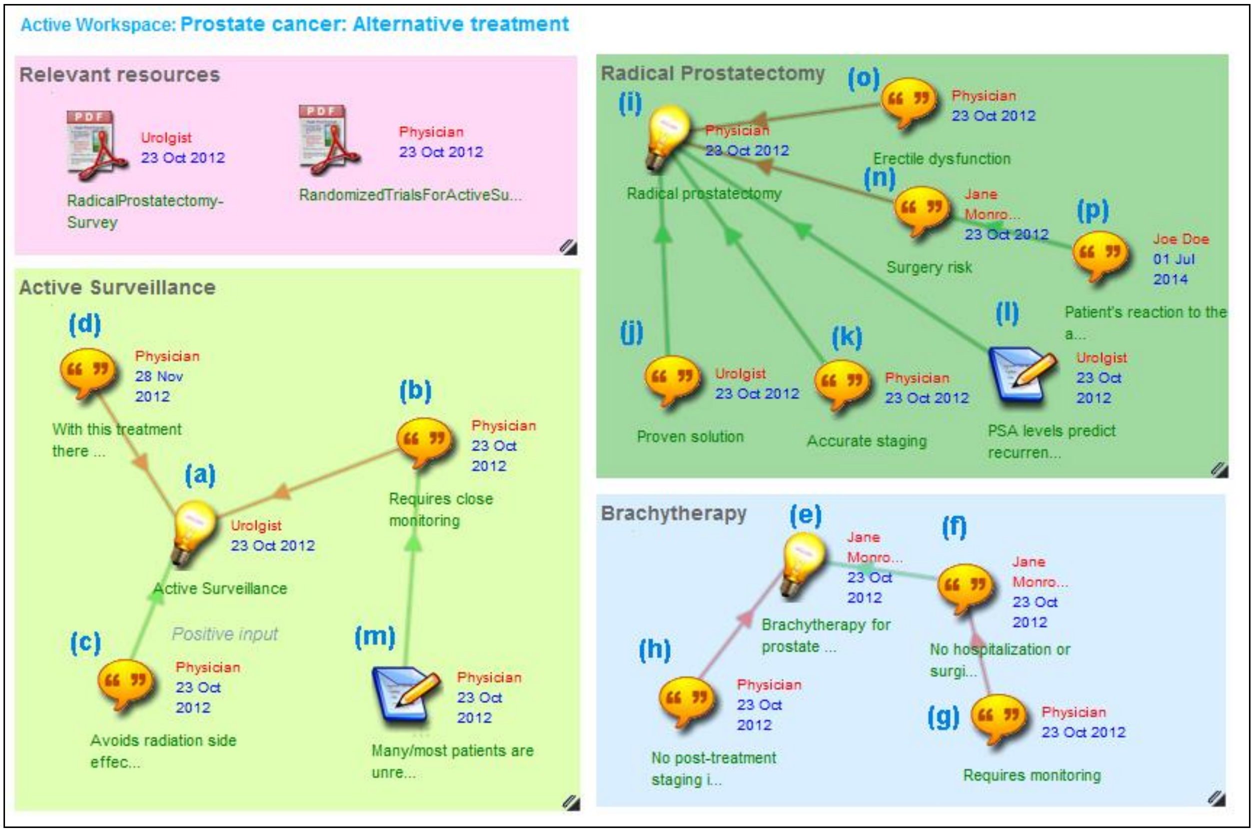

Figure 1. An instance of a real-world scenario concerning collaboration in the area of prostate cancer research.

which is very important and the PSA levels may reliably predict the recurrence (Figure 1(l) – “in favour” position).

Based on the collaboration on the mind-map view so far and his expertise on the field, John is convinced that he is able to contribute on the ongoing collaboration process; however, he is not absolutely certain about the most appropriate collaboration item he should react on (e.g., by creating an argument in favor or against it). He decides to invoke the hybrid recommender to get some insights. As a result, a list of recommended items (Figure 2) is returned. It is noted that these items are compatible to his rating profile and the rating profiles of similar to him Dicode users.

Collaborative recommended collaboration items:

Ranking	docTitle	score
1	With this treatment there .	15.84
2	Requires monitoring	15.78
3	No post-treatment staging	15.67
4	Proven solution	6.40
5	RadicalProstatectomy-Sur	6.40
6	RandomizedTrialsForActiv	6.40
7	No hospitalization or surgi	3.20

Page 1 of 1 10 Viev

Figure 2. The output of the collaborative recommender for John.

Having elaborated the output of the collaborative recommender, John selects the second recommendation as the one closest to his knowledge profile and contributes to Jane’s comment ((Figure 1(b))). He is contradicting to her point of view because, according to his experience, most patients are unreliable as many (most) of them neglect to visit doctors (Figure 1(m)).

Content recommended collaboration items:

Ranking	docTitle	score
1	Requires close monitoring	37.66
2	Erectile dysfunction	16.74
3	No hospitalization or surgi	12.79
4	No post-treatment staging	9.97
5	Active Surveillance	8.14
6	Accurate staging	8.14
7	PSA levels predict recurrer	2.63
8	Proven solution	1.86

Page 1 of 1 10 View 1 -

Figure 3. The output of the content recommender for user “John”.

Jane does not share the enthusiasm for the radical prostatectomy alternative as, due to surgery, a certain amount of risk is involved (Figure 1(n) – "against" position). Apart from this, an erectile dysfunction is expected at the level of 30-50% in 5 years (Figure 1(o)). Joe, who has just joined the collaboration, adds a new collaboration item to support Jane's opinion on the surgery risk involved stating that the danger of a patient's reaction to the anesthesia drugs should be taken into account (Figure 1(p)).

As he is new to the collaboration process and his rating profile is relatively poor, using the hybrid recommendation mechanism to get recommendations invokes the content-based recommender algorithm (due to the collaborative recommender's failure to provide accurate results), which returns a list of collaborative items with similar content to the collaborative item he has just added (Figure 3). Exploiting the recommendations of the content-based recommender, he is now in a better position to contribute to the ongoing collaboration process.

VI. Discussion and Conclusion

The proposed approach builds on the content and structure of an evolving argumentative collaboration, as well as on the rating profiles of similar users, to provide hybrid recommendations in platforms following the IBIS model of argumentation. Concerning the collaborative filtering based recommender, the major benefit of the proposed approach lies in the fact that, in order to provide accurate collaborative recommendations, it exploits the structure of the associated argument trees to estimate the value of the user's inferences on each argumentation item. The application of the proposed hybrid recommender has been demonstrated in the case of Dicode, a collaboration and decision making support platform. Following a similar method of integration, the proposed approach may be easily integrated to any IBIS-like system.

Dicode collaboration support services (including the recommendation support engine presented in this paper) have been thoroughly evaluated in three real-life contexts (clinico-genomic research, medical decision making, and opinion mining from Web 2.0 data). Generally speaking, the feedback received was positive, which clearly points out that the overall approach is promising (a comprehensive description of the evaluation process appears in [18]). Evaluators indicated that our approach reduces the data-intensiveness and overall complexity of real-life collaboration and decision making settings to a manageable level, thus permitting stakeholders to be more productive and concentrate on creative activities [19].

Future work directions include the application of the proposed hybrid recommender in diverse real-life collaborative settings. Through such efforts, we first plan to fine-tune our approach as far as the various parameters of the open-source libraries exploited are concerned. In addition, since there is a number of alternatives to integrate the results of the proposed recommenders (for instance, by using an appropriate switching criterion, the effective selection of which remains an open research issue), more tests have to be conducted in order to decide about the most appropriate method (per collaborative setting) to combine the outputs of the content-based and the collaborative filtering based recommenders.

References

[1] M.J. Eppler and J. Mengis, "The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines". The Information Society, Vol. 20, pp. 325–344, 2004.

[2] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems Handbook". In: Ricci, F., Rokach, L., Shapira, B. and Kantor, P. (Eds), Recommender Systems Handbook, Springer, pp. 1-35, 2011.

[3] M. Aljukhadar, S. Senecal, and C.E. Daoust, "Information Overload and Usage of Recommendations", Proceedings of the ACM RecSys 2010 Workshop on User-Centric Evaluation of Recommender Systems and Their Interfaces (UCERSTI), Barcelona, Spain, pp. 26-33, Sep 30, 2010.

[4] D. Almazro, G. Shahatah, L. Albdulkarim, M. Kherees, R. Martinez, and W. Nzoukou, "A survey paper on recommender systems", arXiv:1006.5278, 2010.

[5] C. Chesnevar, A.G. Maguitman, and M.P. Gonzalez, "Empowering Recommendation Technologies Through Argumentation", In: I. Rahwan and G. Simari (Eds.), Argumentation in AI, Springer, pp. 403-422, 2009.

[6] P. Zhang, J. Sun, and H. Chen, "Frame-based argumentation for group decision task generation and identification". Decision Support Systems, vol. 39, pp.643-659, 2005.

[7] C.I. Chesñevar, A.G. Maguitman, and G. Simari, "A first Approach to Argument-Based Recommender Systems based on Defeasible Logic Programing", Proc. of the 10th Intl. Workshop on Non Monotonic Reasoning (NMR 2004), Whistler, Canada, pp. 109-117, June 2004.

[8] C.I. Chesñevar and A.G. Maguitman, "Arguenet: An argument-based recommender system for solving web search queries". In: Intelligent Systems, 2004. Proceedings. 2004 2nd International IEEE Conference. vol. 1, pp. 282-287, 2004.

[9] W. Kunz and H. Rittel, "Issues as Elements of Information Systems", Working paper No. 131, Studiengruppe für Systemforschung, Heidelberg, Germany, July 1970.

[10] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", IEEE Transactions on Knowledge and Data Engineering, vol.17, no.6, pp. 734-749, June 2005.

[11] R. Burke, "Hybrid web recommender systems". In: P. Brusilovsky, A. Kobsa, and W. Nejdl, (Eds.), The adaptive web, Lecture Notes In Computer Science, Vol. 4321. Springer-Verlag, Berlin, Heidelberg, pp. 377-408, 2007.

[12] S. Gündüz-Öğüdücü, Web Page Recommendation Models: Theory and Algorithms, Morgan & Claypool, 2010.

[13] P. Lops, M. de Gemmis, and G. Semeraro, "Content-based recommender systems: State of the art and trends", Recommender Systems Handbook, pp. 73-105, Springer, 2011.

[14] C. Cayrol and M.C. Lagasquie-Schiex, "Graduality in Argumentation", Journal of Artificial Intelligence Research, vol. 23, pp. 245-297, 2005.

[15] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative Filtering for Implicit Feedback Datasets", Proc. of ICDM '08 - The 8th IEEE International Conference on Data Mining, pp. 263-272, Dec. 2008.

[16] N. Karacapilidis, M. Tzagarakis, N. Karousos, G. Gkotsis, V. Kallistros, S. Christodoulou, C. Mettouris, and D. Nousia, "Tackling cognitively-complex collaboration with CoPe_it!", International Journal of Web-Based Learning and Teaching Technologies, vol. 4, no 3, pp. 22-38, 2009.

[17] N. Karacapilidis (Ed.), Mastering Data-Intensive Collaboration and Decision Making: Cutting-edge research and practical applications in the Dicode project, Studies in Big Data Series, Vol. 5, Springer, 2014.

[18] G. Tsiliki, N. Karacapilidis, S. Christodoulou, and M. Tzagarakis, "Collaborative Mining and Interpretation of Large-Scale Data for Biomedical Research Insights", PLoS ONE 9(9): e108600, 2014, doi: 10.1371/journal.pone.0108600.

[19] N. Karacapilidis, M. Tzagarakis, and S. Christodoulou, "On a meaningful exploitation of machine and human reasoning to tackle data-intensive decision making". Intelligent Decision Technologies Journal, vol. 7, no 3, pp. 225-236, 2013.

A Policy Specification Language for Composite Services

Muhammad Asim[a], Susana González Zarzosa[b], Qi Shi[a] and Bo Zhou[a,]
[a]School of Computing & Mathematical Sciences, Liverpool John Moores University
email:{m.asim, q.shi, b.zhou} @ljmu.ac.uk

[b]Atos Spain S.A., Madrid, Spain
email: susana.gzarzosa@atos.net

***Abstract*— Creating complex systems by combining smaller component services is one of the fundamental concepts in Service Oriented Architecture. Service compositions are built by combining loosely coupled services that are, usually, offered and operated by different service providers. While this approach offers several benefits, it makes the implementation and representation of the security requirements difficult. This paper reviews several requirement specification languages and analyses their suitability for composite services. A set of requirements is identified and a comparison between different specification languages is presented along with some conclusion on the suitability of each language in expressing security requirements for composite services.**

Keywords- policy languages; composite services; security; service-oriented computing

I. INTRODUCTION

Service-based applications are a new class of software systems that allow enterprises to offer their software systems as services by following the principal of *Service Oriented Architectures* (SOA). A service itself is a unit that offers certain functionality. If no single service can satisfy the functionality required by the user, then SOA allows multiple services to be composed to form a larger application in order to fulfil the user requirements. A SOA platform provides a foundation for modelling and composing multiple services in an ad hoc manner [1] [2].

Aniketos is an EU research project [3] that addresses trustworthy and secure service compositions with run-time monitoring and adaptation of services. One important task in the Aniketos project is to choose a specification language that is able to express security requirements, properties or policies for composite services. Also, it is a suitable policy language to specify what we need to monitor at runtime. Besides, the specifications should be able to be generated by both humans and software. In general, this language should serve to other purposes as well, e.g., it should specify the security requirements for a service (either desired by a consumer or advertised by a service provider). Naturally, we may use one language for requirements specification and another one for monitoring these requirements, but then there is a need for a transformation engine. Thus, one language for both purposes significantly reduces the complexity.

This paper reviews several security requirement specification languages and analyses their suitability for a modern, flexible, secure service platform. A set of requirements is identified and a comparison between different specification languages is presented along with some conclusion on the suitability of each language in expressing security requirements for services that are composite in nature. We use the Aniketos Platform as a reference point to discuss these languages and there suitability for composite services.

The paper is organized as follows: The next section presents the requirements for a specification language. Specification languages are discussed in Section 3. The suitability of the language ConSpec for the project Aniketos is discussed in Section 4. Section 5 presents the conclusion on specification language choice.

II. SPECIFICATION LANGUAGES REQUIREMENTS

In the context of the Aniketos Platform development, we are mainly looking for specification languages which are able to address the following requirements. The selected list of requirements is a result of analysis that has been carried out on more than fifty scenarios coming from three different domains (air traffic management, telecommunication and e-government) [4] [5].

- *(Rec-01) Cross-composite- The language for contract specification shall be able to express the properties for a hierarchical service. It should support both atomic and composite services.* Complex services often have a complex hierarchical structure. Thus, the contract specification language should be able to describe the desired and provided properties, taking into account that some parts of the service are provided by the services at the lower end of the hierarchy.

- *(Rec-02) Generalizable and Unambiguous- The language for contract specification shall be general enough to express requirements of various kinds.* Security requirements, which one would like to express with the language could be very different. These requirements may include presence of some countermeasures, various access control policies, well-known security properties, or a numerical security target (e.g., Risk level).

- *(Rec-03) Intelligible- There shall be no difference whether the set of policies is created by a human or* *software.* The language should be easily interpretable both by humans and through automated means.

TABLE 1: CONTRACT/POLICY SPECIFICATION LANGUAGE REQUIREMENTS

Rec-01-01	The specification language should be able to express the scope of the policies to determine if it applies to a single or multiple executions of the same service.
Rec-02-01	The specification language should have unambiguous and restricted semantics to improve its clarity and simplicity.
Rec-02-02	The specification language should be able to represent state transitions.
Rec-03-01	The specifications should be able to be developed for integration with computer programs, i.e., Java.
Rec-03-02	The learning of the language should not require too much technical training in order to be able to express new requirements, properties or policies.

We could make these requirements even more specific as listed in Table 1.

III. SPECIFICATION LANGUAGES

In the literature, we can find a huge amount of work on policy specification languages as well as several taxonomies of these languages. We will start discussing some of these existing classifications that will help us in the search for a suitable specification language to be used in Aniketos and to choose the main potential candidate languages.

First, Bonatti et al. [6] differentiate the following groups of rule-based policy specifications performed by the REWERSE (Reasoning on the Web with Rules and Semantics) Project [7]. They differentiate the following groups of rule-based policy specifications:

1) *Logic-based policy languages:* focused on those languages with unambiguous semantics that enhance clarity, simplicity and modularity. The main advantages of these logic languages are: (i) they are very suitable for validation and verification; (ii) their declarative nature makes them expressive enough to formulate a wide range of policies with simplicity.In this category we find for example the eXtensible Access Control Markup Language (XACML) that is the standard for policy specification developed by the OASIS consortium.
2) *Action languages:* including those languages that can be used to represent actions, changes and their effects. Most of them describe dynamic situations according to a so-called state-action model. One of the most popular logic-based approaches of action languages is EventCalculus.
3) *Business rules:* based on those languages that are more concerned in the formulation of statements about how a business must be done or in other words, the guidelines and restrictions that apply to states and processes in an organization. They distinguish here three categories of rules: reaction rules (“ON event IF condition is fulfilled THEN perform action”), derivation rules (each rule expresses the knowledge that if one set of statements happens to be true, then some other set of statements must also be or become true) and integrity constraints (assertions that must be satisfied in all evolving states). One of the more relevant business process description languages is the Business Process Execution Language for Web Services (BPEL4WS).
4) *Controlled natural languages:* which are defined like “*subsets of natural languages whose grammars and dictionaries have been restricted in order to reduce or eliminate both ambiguity and complexity*”. Therefore, this category would be included in what it is called “semantic languages”. An example is PROTUNE that is the name of the policy language and meta-language developed in the REWERSE Project.

To summarize, from the analysis performed by REWERSE, we select the following potential languages for a further study taking into consideration the requirements indicated above for Aniketos:

- XACML
- Event Calculus
- Web Service Description Language (WSDL) /BPEL4WS
- PROTUNE (and other relevant semantic web languages)

In the PrimeLife Project [8], they define three types of policies that they considered important parts of any privacy policy that have to be covered by any policy language: (i) data handling; (ii) access control; and (iii) trust policies. The languages selected from the PrimeLife study are:

- XACML
- The Platform for Privacy Preferences (P3P)

Finally, we are going to analyse the Contract Specification Language (ConSpec) that is an automata-based policy specification language presented in the literature [9] as a potential language for specifying both policies and contracts in various security enforcement related tasks of the application lifecycle.

In the next subsections, we discuss in more detail each one of the selected policy languages that we have considered as candidates in Aniketos.

Copyright (c) IARIA, 2015. ISBN: 978-1-61208-436-7

A. eXtensible Access Control Markup Language

eXtensible Access Control Markup Language (XACML) [22] is an Extensible Markup Language (XML) based language used to express and interchange access control policies. It is designed to express authorization policies in XML against objects that are themselves identified in XML. XACML is a general purpose policy language and it can be used to protect any resource type (i.e., not just data), but it is difficult to write XACML policies and even more difficult to reason over (i.e., it is unsatisfactory regarding requirement Rec-03-02). Therefore we could use this language in Aniketos project since it would allow encoding most of security properties that will be included into the Contract (requirement Rec-01), but we would need to "misuse" the constraint part of XACML policies since XACML is tailored towards Access Control policies.

B. Event Calculus

Event Calculus (EC) [10] is a first-order temporal logical language for representing actions and their effects that can be used to specify properties of dynamic systems, which change over time. Such properties are specified in terms of events and fluents. An event in EC is something that occurs at a specific instance of time (e.g., invocation of an operation) and may change the state of a system. Fluents are conditions regarding the state of a system, which are initiated and terminated by events. A fluent may, for example, signify that a specific system variable has a particular value at a specific instance of time or that a specific relation between two objects holds.

ecXML [11] is an XML formalisation of the Event Calculus that is used to describe how a contract's state evolves, according to events that are described in the contract. The main advantage of this language for Aniketos is that it is very suitable for runtime monitoring and can be used to represent properties, policies and contracts in a dynamic environment (Rec-01). But it is more oriented towards states and actions than services, and the syntax could become too complicated for compound services and expression of hierarchies (Rec-02). Moreover, it would require a big effort to accomplish requirement Rec-03 to automate the generation and runtime monitoring of these rules in Java code.

C. Web Service Definition Language / Business Process Execution Language for Web Services

The WSDL [12] is the World Wide Web Consortium (W3C) standard language for web service descriptions. It is an XML format used to create a flexible Service Level Agreement (SLA) for web services defining mutual understandings and expectations of a service between the service provider and service consumers. It uses a very limited syntax that defines services as collections of network endpoints or ports.

The Business Process Execution Language for Web Services (BPEL4WS) [13] is a language used for specifying business process behaviour based on Web Services, which was created to overcome the limits of WSDL. It allows building definitions of a business process (that can be either an executable itself or a business protocol) where both the process and its partners are modelled as WSDL services. The language is layered on top of several XML specifications (WSDL 1.1, XML Schema 1.0, and XPath1.0) but makes no use of semantic information.

This language is a service-oriented composition language that forms the base of Aniketos, but we want to express also security properties and trustworthiness (Rec-01). Consequently we need something that provides more information than BPEL4WS.

D. PROVisional TrUst Negotiation

PROvisional TrUst Negotiation (PROTUNE) [14] is a natural language for the specification of rule based policies on the semantic web defined by REWERSE [7]. It is a logic-based and declarative policy language that includes logical axioms to constrain the behaviour and how the web resources must be used. But the main feature of PROTUNE that makes it different from the previously discussed languages is that it is a semantic web language.

The semantic web languages are developed to allow intelligent agents in the semantic web to reason and make decisions policy-driven based on the knowledge it is provided by the semantics. Therefore, one of the main advantages of these semantic web languages for Aniketos is that it facilitates greater automatic machine interpretability of conditions, taking decisions and performing tasks (covering the requirement Rec-03). Besides, this kind of language provides an enormous expressivity and can be used to represent complex knowledge in a distributed environment and support classification in hierarchies (requirements Rec-01 and Rec-02). But this last feature is also a big drawback (high complexity) due to which it cannot be considered in the project Aniketos. Reasoning with a semantic web language is difficult and it requires a well-defined semantic that should be developed specifically for Aniketos. Furthermore, its high expressiveness can lead to non-standard formalism and sometimes to complexity in the reasoning.

The semantic web languages standardized by the W3C are (i) Resource Description Framework (RDF) [15] and (ii) Web Ontology Language (OWL) [16].

OWL includes more vocabulary and consequently extends the facilities offered by XML, RDF and RDF Schema (RDF-S) for expressing meaning and semantics what makes it easier to represent machine interpretable content on the Web. In turn, OWL provides three increasingly expressive sublanguages: OWL Lite, OWL Description Logic (OWL DL), and OWL Full.

In the case of PROTUNE, the syntax is based on normal logic program rules. Finally, we can take into consideration two prominent semantic web languages based on OWL, which appear in much of the literature: Rei and KAoS [17][18]. Rei is a policy language based on OWL-Lite that includes logic-like variables to provide more flexibility in the specification of relationships that are not possible in OWL. For example, it is possible to define individual and group based policies that could be useful in large scale distributed environments for saving time. They are associated with agents, called subjects, by means of the *has* construct: *has(Subject, PolicyObject).*

KAoS is another policy language based on OWL with the following distinguishing features: (i) it does not assume that the policies are applied in homogeneous components: (ii) it supports dynamic runtime policy changes; (iii) the framework is extensible to different execution platforms; (iv) the KAoS framework is intended to be robust and adaptable in continuing to manage and enforce the policy of any combination of components.

E. Platform for Privacy Preferences

The P3P [19], published by the W3C, enables web sites to express their privacy practices in a standard format that can be retrieved automatically and interpreted easily by user agents. P3P user agents will allow users to be informed of site practices (in both machine and human readable formats) and to automate decision-making based on these practices when appropriate. But this option has been discarded for Aniketos because a report [20] on the assessment of P3P and Internet privacy finds that P3P fails to comply with baseline standards for privacy protection. It is a complex and confusing protocol that also fails to address many of the privacy problems. The report concludes that there is little evidence to support the industry claim that P3P will improve user privacy citing the widely accepted Fair Information Practices.

F. ConSpec Language

The ConSpec [9] language with its syntax shown in Fig. 1 is strongly inspired by the policy specification language PSLang, which was developed by Erlingsson and Schneider in [21] for runtime monitoring. However, even though ConSpec is a more restricted language than PSLang, it is expressive enough to write policies referring to multiple executions of the same application, as well as to executions of all applications of a system, in addition to policies about a single execution of the application and of a certain class object lifetime according to the *scope* of the policy.

Effectively, a ConSpec contract specifies a set of guards each with an associated set of reactions. A guard is defined as a method prototype. A reaction is a set of expressions specifying state changes, where the left hand side specifies the state before and the right hand side the state afterwards. Whenever the guard method is called in the code, the state expression is checked and, if the left hand side of the expression matches the current state, the right hand side expression is applied to update it. In the event that the state fails to match any of the left hand side expressions, the contract is considered to be violated. The following example states that, once the file *Secret.dat* has been opened, plaintext socket connections can no longer be used. Note that the skip keyword is used to represent no state change.

```
BEFORE File.Open(String path) PERFORM
     path == "Secret.dat" -> {private = true;}

BEFORE Socket.Send(String sd, String data)
PERFORM
     private == false -> {skip;}
```

Figure 1. Syntax of ConSpec

One of the attractive features of this approach is that the use of a finite state machine coupled with guards defined against explicit methods means that the ConSpec script defines not just the policy but also the means to identify it. However, we can also see from the above example that ConSpec was originally developed for use with single isolated pieces of software, written in a specific language (in the case of the Aniketos project, this is Java). This impacts the cross-composite requirements.

The language has therefore been extended to support composed services [1][2]. This can be achieved in one of two ways. First, a single ConSpec file can be defined to apply across a set of composed services. This requires there to be a single centrally managed finite automata state machine that all guard events refer back to. In this case, rather than specifying methods for the guards, a service identifier must also be specified. Service identifiers can also be passed as a parameter to the reaction, so that the state change can be predicated on service properties as well. In this case, earlier guards that identify particular functionality in a particular service can be used to correlate with guards identifying different functionality at a later time. It also allows more flexibility in defining contracts, since ideally the contract should be independent of the service composition that its applied to. Second, each service can be given its own ConSpec file. In this case there’s effectively an automaton applied to each service. However, there needs to be correlation between the services, so a further central automaton is needed at the composition level. State changes at the service level generate events, which are then matched against guards at the composition level which potentially update the central automaton. An attractive feature of using finite automata is that they are themselves compositional: this arrangement is equivalent to a finite automaton applied across all services. This allows cross-composition.

The policies written in ConSpec are easily interpretable by humans. It has a comparatively simple semantics, and is simple to learn. ConSpec is an automata-based language. Although this feature slightly reduces its expressiveness (in comparison with its predecessor PSLan [21], or other declarative languages as EventCalculus [10], XACML [22], PROTUNE [14], etc.), it allows automatic reasoning on it. For example, in the project we needed to check that requirements desired by a consumer could be fulfilled by a service provider. Furthermore, it is simple to define a policy decision point for monitoring purposes if an automaton is available. Finally, ConSpec defines different scopes of its application. Thus, we may define a policy for a single execution of a service or multiple executions. Overall, ConSpec provides an unambiguous, cross-composite and intelligible approach, which makes it a more suitable specification language for composite services.

IV. CONSPEC IN THE ANIKETOS PROJECT

Based on the above analysis, we selected the ConSpec language as a specification language for the Aniketos platform and extended it (as discussed above) to support the composite nature of services. In the scope of the Aniketos project, we have created a tool, which provides a graphical user interface for making and changing ConSpec policies. The tool is called a ConSpec Editor illustrated in Fig. 2.

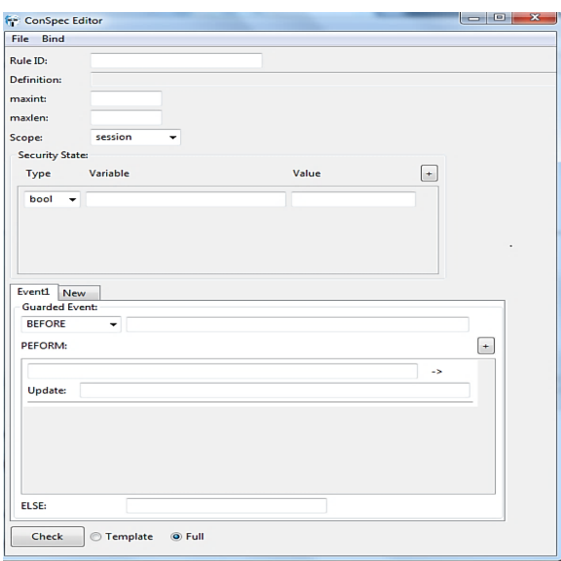

Figure 2. ConSpec editor

ConSpec policies can be created with the ConSpec Editor without knowing the ConSpec language. As an example, the ConSpec policies are used by a monitoring module developed as a part of the Aniketos project. The monitoring module is responsible for the runtime monitoring of a service to ensure that the service behaves in compliance with a pre-defined security policy. For more details about the monitoring framework, see [2].

Consider the following example where a service designer creates a travel booking composition that consists of several tasks, such as ordering, booking hotel, booking flight, payment and invoice, and each task is performed by a component service. The service designer might want that the payment service component should only be invoked when it has a trustworthiness value ≥ 90%. This requirement could easily be specified using the ConSpec language as shown in Fig. 3.

```
RULE ID Trustworthiness
SECURITY STATE
    String ServiceID=Payment;
   int trust_threshold = 90;
   /* assume trustworthiness is in [0%,..., 100%]*/

BEFORE invoke (serviceID)
PERFORM
   (eval_Trust(serviceID) >= trust_threshold) -> skip
                                 condition1 -> update
```

Figure 3. ConSpec policy example 1

The monitoring module in adherence to the policy monitors services to ensure that only a payment service with trustworthiness value ≥ 90% is used. In another example, where a service designer imposes the separation of duty constraint for a particular service composition, i.e., both service A and service B should be offered by different providers.

```
RULE ID SoD_Goal
SECURITY STATE
string serviceProvider = _;
string guardedTask1 = ServiceA;
string guardedTask2 = ServiceB;

BEFORE v#service.start
(string id, string type, int time, int date,
   string provider) PERFORM

(id == guardedTask1 || id == guardedTask2) &&
   serviceProvider == "_" -> {serviceProvider =
   provider; }

(id == guardedTask1 || id == guardedTask2) &&
   !(serviceProvider == "_") && !(provider ==
   serviceProvider) -> {skip}

   !(id == guardedTask1) && !(id == guardedTask2)
-> {skip}
```

Figure 4 . ConSpec policy example 2

The requirement for the above example can be specified in ConSpec as illustrated in Fig. 4.

V. CONCLUSION ON SPECIFICATION LANGUAGE CHOICE

The different languages discussed here exhibit interesting properties in relation to their suitability for composite service. However, comparing the requirements and needs that the Aniketos platform requires to express security policies and the previous descriptions of the different languages, we can conclude that ConSpec is the best solution for the main reasons summarized below:

- It is extended to offer unambiguous, cross-composite solutions with important elements of generalizability for composite services.
- It is developed as a language for representing security relevant behaviours of an application in terms of Java calls, which allows the rules to be generated and checked at runtime also by software or security automata.
- A policy written in the ConSpec language is easily interpretable by humans and the simplicity of the language allows a comparatively simple semantics and a reasonably fast learning curve.
- Although ConSpec does not allow any arbitrary type to represent the security state of a service, it includes tags for expressing security requirements in different stages of the application life cycle. It makes it possible to indicate constraints that can be applied to multiple executions of a service, as well as interactions with other services.

TABLE 2. MATCHING OF SPECIFICATION LANGUAGES TO REQUIREMENTS

Requirement	XACML	EC	WSDL/BPEL4WS	PROTUNE	OWL-based (Rei/KAoS)	ConSpec
Rec-01	X	X	X	X	X	X
Rec-01-01	X	X				X
Rec-02	X			X	X	X
Rec-02-01			X			X
Rec-02-02	X			X	X	X
Rec-03	X		X	X	X	X
Rec-03-01	X		X	X	X	X
Rec-03-02			X			X

Table 2 summarizes the requirements that are covered by each of the different languages presented above.

REFERENCES

[1] D. Llewellyn-Jones, M. Asim, Q. Shi, and M. Merabti, "Requirements for Composite Security Pattern Specification," Second International Workshop on Cyberpatterns 2013: Unifying Design Patterns with Security, Attack and Forensic Patterns, Abingdon, UK., 2013, pp. 70-77.

[2] M. Asim, D. Llewellyn-Jones, B. Lempereur, B. Zhou, Q. Shi, and M. Merabti, "Event Driven Monitoring of Composite Services,"The 5th ASE/IEEE International Conference on Information Privacy, Security, Risk, Washington D.C., USA, Sep 2013, pp. 550-557.

[3] Aniketos (Secure and Trustworthy Composite Services), http://www.aniketos.eu, retrieved: April, 2015.

[4] Aniketos Consortium, Deliverable D2.3: Models and methodologies for implementing Security-by-Contract for services, 2012,http://www.aniketos.eu/content/deliverables, retrieved: April, 2015.

[5] Aniketos Consortium, Deliverable D1.2: First Aniketos architecture and requirements specification, 2012, http://www.aniketos.eu/content/deliverables, retrieved: April, 2015.

[6] P. A. Bonatti et al. Rule-based Policy Specification: State of the Art and Future Work. Technical Report IST506779/Naples/I2-D1/D/PU/b1, Reasoning on the Web with Rules and Semantics, REWERSE, August 31st, 2004.

[7] REWERSE: Reasoning on the Web with Rules and Semantics, http://rewerse.net/, Retrieved: March, 2015.

[8] PrimeLife (Privacy and Identity Management in Europe for Life), Deliverable D5.1.1: Final requirements and state-of-the-art for next generation policies, August 2009, http://primelife.ercim.eu/, retrieved: March, 2015.

[9] I. Aktug and K. Naliuka, "ConSpec: A Formal Language for Policy Specification.", In Proceedings of the First International Workshop on Run Time Enforcement for Mobile and Distributed Systems , 2007, pp. 2-12.

[10] M. P. Shanahan, The Event Calculus Explained, in Artificial Intelligence Today, eds. M. J. Wooldridge and M. Veloso, Springer-Verlag Lecture Notes in Artificial Intelligence no. 1600, Springer-Verlag, pages 409-430, 1999.

[11] A. D.H. Farrell, M. J Sergot, M. Salle, and C. Bartolini, "Using the Event Calculus for the Performance Monitoring of Service-Level Agreements for Utility Computing" First IEEE International Workshop on Electronic Contracting (WEC'04), 2004

[12] Web Service Description Language (WSDL) 1.1, http://www.w3.org/TR/wsdl, retrieved: March, 2015.

[13] Business Process Execution Language for Web Services Version 1.1, http://public.dhe.ibm.com/software/dw/specs/ws-bpel/ws-bpel.pdf, 5 may 2003, retrieved: March, 2015.

[14] P.A. Bonatti, J.L. De Coi, D. Olmedilla, and L.Sauro, "PROTUNE: A Rule-based PROvisional TrUst Negotiation Framework", 2010.

[15] Resource Description Framework (RDF), http://www.w3.org/2001/sw/wiki/RDF, retrieved: Feb, 2015.

[16] OWL Web Ontology Language Overview, http://www.w3.org/TR/owl-features/, WC3 Recommendation February 2004, retrieved: Feb, 2015.

[17] G. Tonti, J. M. Bradshaw, R. Jeffers, R. Montanari, N. Suri and A. Uszok, "Semantic Web Languages for Policy Representation and Reasoning: A Comparison of KAoS, Rei and Ponder," In Proceedings of the 2nd International Semantic Web Conference (ISWC2003). Springer-Verlag, 2003.

[18] L. Kagal, T. Finin and A. Joshi, "Declarative Policies for Describing Web Service Capabilities and Constraints", Proceedings of the 6th international conference on E-Commerce and Web Technologies, 2005

[19] Platform for Privacy Preferences Project (P3P), http://www.w3.org/P3P, retrieved: Feb, 2015.

[20] Electronic Privacy Information Center and Junkbusters, "Pretty Poor Privacy: An Assesment of P3P and Internet Privacy" (June 2000), http://epic.org/reports/prettypoorprivacy.html, retrieved: Feb, 2015.

[21] U. Erlingsson. The inlined reference monitor approach to security policy enforcement. PhD thesis, Department of Computer Science, Cornell University, 2004.

[22] eXtensible Access Control Markup Language (XACML) Version 3.0 (http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.pdf), retrieved: Feb, 2015.

SilentMeet - A Prototype Mobile Application for Real-Time Automated Group-Based Collaboration

Remus-Alexandru Dobrican, Gilles I. F. Neyens and Denis Zampunieris
University of Luxembourg
Luxembourg, Grand-Duchy of Luxembourg
Email: remus.dobrican@uni.lu, gilles.neyens.001@student.uni.lu, denis.zampunieris@uni.lu

***Abstract*—Today's growing world of mobile devices offers all the necessary elements for developing collaborative mobile applications. However, this brings new challenges like how to handle the high complexity of efficient collaborative mechanisms or automatize part of the user's interaction with the applications, as too many actions are required from the users in order to perform even the most basic operations. This paper describes an experimental mobile application, i.e., SilentMeet, that uses a rule-based middleware architecture for mobile devices and a new technique for exchanging information, for coordinating and for taking distributed decisions. More precisely, the application is designed to detect, based on collaboration, possible meetings or events with more than 2 participants and automatically switch the smartphone into silent mode. The goal of SilentMeet can be divided into 2 two main parts: 1) to develop a collaborative application with the help of rule-based systems; and 2) implement and evaluate Global Proactive Scenarios (GPaSs) in a real-case example.**

Keywords–mobile devices; collaborative applications; proactive computing; distributed group collaboration.

I. Introduction

Communication and collaboration, more precisely interactive collaboration, are two key aspects in today's mobile world. Basic mobile applications that are able to perform only local tasks do not address the increasing needs of the users anymore. The demand for services and applications which support communication and collaboration of mobile devices has raised significantly in the past years [1]. The latest interest in mobile collaboration can be explained by the large number of mobile devices around the world, which is continuing to grow from one year to another [2]. However, this mobile environment capable of performing distributed operations brings new challenges, such as intermittent connectivity, data heterogeneity, limited computational capabilities and users' mobility. Also important, is the fact that mobile networks, due to the high mobility of their users [3], differ a lot from static systems, where the users are always connected. This leads to the issues like determining the context information needed to trigger the collaboration process or like users being temporarily unavailable while they are still engaged in the collaborative operations.

Another important aspect to be addressed, when designing collaborative applications, is to establish up to which level will the users interact with the system. Because users may have basic skills or only limited experience when interacting with complex applications or because they do not want to spend a lot of their time giving instructions to the system, the applications can automatize a lot of their processes. One of the solutions for doing this is Proactive Systems, which are able to act on their own initiative and to take decisions on behalf of their users [4]. Recently, the possibility of implementing a Proactive Engine for mobile devices was investigated [5]. The added value is that, with the help of a mobile Proactive System, which is essentially an advanced rule-based system, developers can directly add the functionality they want to their applications by using Proactive Rules. From the developer's point of view, a Proactive Rule represents a tool for writing a set of instructions, while from the system's point of view, a Proactive Rule is a piece of code which has to be executed. More about Proactive Rules and examples with the rules used for this study will be shown in Section III.

In order to have a rule-based system capable of executing Proactive Rules on mobile devices, a middleware model was created for Android-based devices [6]. This represents an important achievement as until now only lightweight basic rule-based engines like [7] and [8] were developed for mobile platforms. These engines would allow applications to use simple conditional rules. The middleware model is also providing an information sharing method between the devices called Global Proactive Scenario (GPaS) [9]. This method was implemented to give the possibility to the applications to perform collaborative tasks.

Numerous studies [3][10][11] have been conducted that provide middleware architectures as tools for developing collaborative applications. One important difference is that these studies look at collaboration from a different angle. More precisely, they concentrate on user-centred collaboration, where the focus is to get the users to interact more and more with their applications on the mobile devices. The issue is that these applications would depend too much on the actions of their users and, if the users do not engage properly in each step of their interaction with their devices, the applications may remain at the same step. Opposite to this, Proactive Computing, which was defined by Tennenhouse as a new way of computing, for and on behalf of the user [12], tries to reduce the users' involvement by automatizing some processes. By doing so, the users can concentrate more on the most important parts of the collaboration.

Many mobile applications exist on the market, like Silence[13], Go Silent [14] or Advanced Silent Mode [15], which automatically switch off the sounds of mobile devices based on the user's preferences. These simple applications perform only local tasks like checking the user's predefined preferences or detecting calendar events. They do not use any kind of collaboration with other devices to make the application smarter.

For example, SilentTime [16] searches for weekly events in the local schedule and automatically silences the user's phone if a future event is detected. It offers the user the possibility to add exceptions, in case he/she is waiting for an important phone call. However, the application has a couple of downsides. First, it is exclusively based on the users input, i.e., a calendar event or an exceptions of a special situations will only be detected if the user creates them before, and second, it does not use any kind of communication with other devices to check if the events will take place or not. Another example is AutoSilent [17], which is slightly different from SilentTime because it adds an extra step of verification before muting the users phone, i.e., it will verify if the users location corresponds with the events location at a certain time. This extra feature is again just a simple check because it does not use any kind of collaboration, like, for example, checking also the location of the other participants.

In this study, we investigate how a mobile application, i.e., SilentMeet, which uses a proactive rule-based middleware system to communicate and collaborate, is automatically turning the devices into silent mode if a meeting is detected and confirmed between a predefined group of users.

The rest of the paper is structured as follows. Section II introduces the problem statement and a motivating scenario that points out the need for automatizing certain tasks and processes inside applications in order to reduce the user's involvement in unnecessary situations. Section III contains explanations about SilentMeet's architecture and about its way of reaching a global decision based on distributed reasoning, including how and which Proactive Rules were used in this case. Tests on real devices are discussed in Section IV and their results in Section IV-B. And finally, Section V contains the main conclusions and future work.

II. Motivating Scenario

There are quite a few mobile users who went through embarrassing situations when their phones rang during important meetings, lectures, exams, presentations, concerts, interviews or key talks offered at international conferences. Imagine, for example, that during a viola recital of a famous musician, the mobile phone of a person start ringing, like it did during a recital in Slovakia [18]. The musician is not only interrupted but he/she could also loose focus and find it difficult to continue. The are many more other examples when muting the phone is a mandatory requirement. The main problem is that each user has to manually configure his/her phone to be silent during important events. And often, they forget. A general common strategy or approach which performs collaborative actions is missing.

Let's imagine the following scenario: an important event is about to begin. The mobile devices of the participants, located in their pockets, go automatically into silent mode. The participants do not have to worry they forgot to silence their mobile phones, they can focus more on their important tasks. The meeting can continue without any interruptions or embarrasing situations.

III. A rule-based solution - SilentMeet

SilentMeet is a collaborative application which is developed in order to minimize the risk of interruptions and their distracting effects during an important event such as a meeting, interview or public event. Moreover, for having an efficient distributed algorithm, part of the user's actions are automatized with the help of Proactive Rules. We assume that groups of people are predefined when an event is created by each user. More precisely, when a calendar event is created, the user also adds the participants. Users can perform collaborative actions only if they are part of the same group of the same event. So, users first have to build their own groups or agree to be part of already created groups. For example, in a company, the secretary of a department creates a group for the employees of that department that have meetings regularly. By joining this group, the members agree that their mobiles phones can be silenced by the application of the other members, after multiple rounds of negotiation. More about the negotiation process is presented in Section III-D. Also, more conditions and checks are taken into account like the location of the event and the participants, the date and the hour of the event and the local preferences of each user.

A. Middleware model - Proactive Engine for Mobile Devices

The Proactive Engine is a middleware architecture developed to support the execution of Proactive Rules. It was designed to perform background operation and to interact with the user only when necessary. Moreover, it comes equipped with a Rules Engine to process rules, a data storage mechanism to store different parameters and with a communication layer, to be sure the Proactive Engine is able to share important information.

Proactive Engines communicate with each other by sending JavaScript Object Notation (JSON) messages. The messages can contain questions, answers or commands, depending on their purpose. For example, a Proactive Engine can send a question to another engine to ask for various context information. Based on the received answer, if some conditions are fulfilled, the engine can then send a command to the other engine to perform an action. Messages are forwarded to a server and to the cloud. The server and the cloud are in charge of assigning each device with a device ID and with forwarding the JSON messages to the targeted devices. They also handle special cases such as lost JSON messages or devices that are not temporary available on the network.

B. Global Proactive Scenarios (GPaSs)

The idea of SilentMeet is that the devices participating in a collaboration process can take global decision, not only local ones. Each device is able to make use of the global knowledge, which is created by all the devices. For example, a basic application would only be able to detect an event based on the local information provided by the calendar of a device. SilentMeet is able to query all the devices to obtain more precise information about that event with the help of GPaSs. A GPaS is a data exchange mechanism, which allows devices to dynamically acquire relevant context information by merging data from multiple sources. It works between all mobile devices with an integrated Proactive Engine. GPaSs, from a technical point of view, are composed of sets of Proactive Rules.

SilentMeet contains one GPaS, as it uses a distributed reasoning algorithm to reach a decision and to execute specific tasks. In this specif GPaS, each device needs additional information from the other devices before taking a decision. The idea is that if multiple devices, part of a collaboration group, have an event in their local calendar, with the same date, time and location, it is very probable that the event will take place. We presume that the same information about an event coming from 2 different devices part of the same group is enough for the application to decide what to do next. In this case, it will switch the corresponding devices into silent mode when the event will take place. The minimum number of 2 devices is motivated by the fact that a device should not be able to mute, by itself, other devices without any kind of agreement. Also, a decision can be taken without the confirmation of the event from all the participants, as this is very difficult to achieve in real-life situations, where each user is expected to manually add the event into the calendar.

C. Proactive Rules

Proactive Rules, as shown in Figures 1, 2, 3, 4 or 5, contain a set of instructions, which are written by the developer. These rules are to be executed by the Proactive Engine when different events are detected or when they are missing. The initial structure of a Proactive Rule [19] was used for creating the rules necessary for SilentMeet. It contains 5 main parts such as *data acquisition*, *activation guards*, *conditions*, *actions* and *rules generation*. These parts are important as they decide when a rule is executed, if the rule performs its actions, if the rule will generate other rules or will just simply clone itself. Proactive Rules can have different execution times because their activation depends on the local settings of each device and on the user's actions. For example, 2 users creating a new calendar event at different hours on their phones, trigger, at different time intervals, the rule which starts the negotiation process of SilentMeet.

For achieving its goal, SilentMeet only needs 5 Proactive Rules. SilentMeet will come together with the 5 Proactive Rules, when installed on each device. Initially, only one Proactive Rules, the first one, will be executed by the Proactive Engine. Then, all the rules can be activated, if their execution conditions are meet.

Illustrated in Figure 1, the first Proactive Rule, i.e., R001-DetectMeeting, is used to detect new meetings added in the calendar of each device. Adding the meeting in the calendar was either manually added by the user or automatically added by another application. If a new meeting is detected, all the participants of that meeting will be retrieved and will be contacted by a second rule. If the first rule is not activated, it will continues to clone itself for being executed at the next iteration of the Proactive Engine. When activated, rule R002 - ContactAttendees, shown in Figure 2, is in charge of sending a request to all the participants for more information about the meeting. The request is then forwarded with specific parameters like the sender ID, the destination ID and the full details of the event itself.

R001 - DetectMeeting

Description: This Rule is the initial rule in the proactive engine. It will check for new events in the calendar and then create a ContactAttendees rule for each event.

```
parameters
    None
data acquisition
    Event[] events= getNewEvents();
activation guards
    return !empty(events);
conditions
    return true;
actions
    None
rules generation
    if (!activationGuard());
        cloneRule(DetectMeeting);
    else
        foreach event in events:
            createRuleContactAttendees(event);
    end if
```

Figure 1. First Proactive Rule in pseudo-code

R002 - ContactAttendees

Description: This Rule sends an AskMeetingConfirmation Rule to every attendee of an event.

```
parameters
    Event event;
data acquisition
    String[] attendees= event.getAttendees();
activation guards
    return true
conditions
    return true;
actions
    foreach attendee in attendees:
        sendAskMeetingConfirmationRule(attendee,event,deviceID);
rules generation
    None
```

Figure 2. Second Proactive Rule in pseudo-code

The third rule, or rule R003 - AskMeetingConfirmation, presented in Figure 3, sends back a response about the event to the device which has previously sent the request for extra information. The answer is positive if an event with the same date, hour, location and participants is detected in the calendar and a negative answer otherwise. The fourth rule, R004 - ConfirmMeeting, checks for the answers of each participant and, if there is at least one confirmation, it validates the requirements of having at least 2 users that will attend the same meeting. If the previous condition is meet, it activates the last rule, which is in charge of muting the mobile phones during a selected event. And the last rule, R005 - MuteCommand, is the one that checks if the meeting is about to begin, and, if the device's location is close to the location of the event, it will activate the silent mode that that particular device. The local preferences of a the user are also checked because there are situations where the user is expecting an important phone call. For example, a man would like to be called if his pregnant wife is giving birth at the hospital, even if, at the time of the call, the man would be in a meeting.

R003 - AskMeetingConfirmation

Description: This Rule will check whether an event is present in the calendar of this user. If this is the case it will send a ConfirmMeeting Rule back to the device which started the negotiation.

parameters
 Event event;
 String senderDeviceID;
data acquisition
 boolean answer=meetingExists(event.startTime,event.endTime);
activation guards
 return answer;
conditions
 return true;
actions
 sendConfirmMeetingRule(senderDeviceID,event);
rules generation
 None

Figure 3. Third Proactive Rule in pseudo-code

R004 - ConfirmMeeting

Description: Upon receiving this Rule the application will issue a MuteCommand Rule to the device that confirmed the meeting.

parameters
 Event event;
data acquisition
 None
activation guards
 return true;
conditions
 return true;
actions
 sendMuteCommandRule(senderDeviceID,event);
rules generation
 None

Figure 4. Fourth Proactive Rule in pseudo-code

R005 - MuteCommand

Description: This Rule will silence the phone when the event starts.

parameters
 Event event;
data acquisition
 None
activation guards
 return eventStarted(event);
conditions
 return atMeetingLocation(getCurrentLocation(),
 event.getLocation()) and checkPreferences(event));
actions
 mutePhone();
rules generation
 if (!activationGuard()):
  cloneRule(MuteCommand);
 end if

Figure 5. Fifth Proactive Rule in pseudo-code

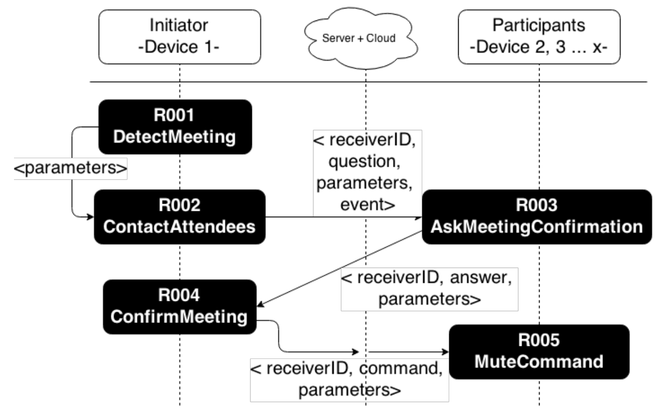

Figure 6. Sequence diagram with the rules activation and the communication process of SilentMeet

D. Negotiation Process

For muting the mobile devices of the participants of a group, after a calendar event is detected, SilentMeet passes through a couple of rounds of negotiation. As shown in Figure 6, when a device detects an event in its calendar, with the help of rule R001, it immediately checks for the participants of that event. Then, it starts, with the help of rule R002, contacting and asking each participants about the event. On the receiving devices, rule R003 is activated and start looking for an event in the calendar with the particular characteristics as the ones received in the list of parameters. An answer is then forwarded to the initiating device and rule R004 is activated. Rule R004 will check if the Initiator receives at least one positive answer, i.e., event detected on another device, it will send a command to the participants which confirmed the event to activate rule R005. The last rule will then be activated and it will wait for all the conditions to be satisfied in order to perform its actions, i.e., to put the device on silent mode.

IV. Tests

Tests were conducted locally at our university on 3 different devices: a Samsung Galaxy Note 3, a Samsung Galaxy S3 and a Nexus 4, as shown in Figure 7. All 3 devices use an Android operating system and have the Proactive Engine middleware installed in order to be able to execute rules and collaborate with each other in GPaSs. The devices were part of a predefined group of 3 participants. During the tests, all 3 devices were connected via WiFi to the same network. Initially, all the devices had their sound turned on and had an event in their calendar with the same date, time and location. The first rule, R001 - DeetectMeeting, was activated on each device when the event started to take place. The Rule Engine was set to execute the rules present in its queue every 30 seconds.

The participating devices used in the tests were part of the same predifined group for the given event. For registering to the group, the user of each device had to use a unique email address, e.g., user1@uni.com, an email address provided by the user of the Galaxy Note 3.

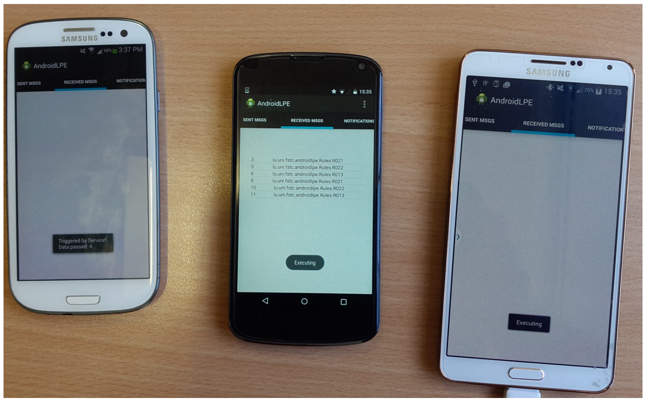

Figure 7. Devices used for testing SilentMeet, in the process of activating the silent mode

A. Measurements

The main goal was to check if the application behaved as expected and, based on the algorithm for distributed agreement, if all three devices were muted, after checking if there is at least one other person who is still attending the meeting. Another point of interest of the tests was the time the devices needed to reach a common agreement and to perform the actions of the last rule R005 - MuteCommand, which muted all the devices.

B. Results and discussions

The tests showed that the application behaves as expected and that all three devices were muted after the negotiation process. In the given settings, it took around 10 seconds to reach a common agreement that the meeting will take place and to mute all three devices. However, this time is highly dependent on the frequency parameter of the Rule Engine, meaning that setting a lower time interval between two iterations will also lead to a faster execution of the GPaS.

V. CONCLUSION AND FUTURE WORK

In this paper, we show that it is possible to develop a collaborative application on top of a rule-based middleware engine and with the help of Proactive Computing, more precisely by using Global Proactive Scenarios. The application is able to detect and acquire relevant context-information, use a distributed reasoning algorithm and take global decision. At the same time, several parts of the collaboration process were automatised and the user's involvement reduced only to the most important operations.

Future work includes developing more complex collaborative applications and other Global Proactive Scenarios on top of the Proactive middleware engine. One possible direction is to develop another GPaS for turning off the sounds of each smartphone, based on the location of the participants. If the location of the participants would correspond to the location of the event on the given date the Proactive Engines would start to alert each other and mute each smartphone. This would show how more than one GPaS can be used in the same application.

REFERENCES

[1] Forrester. Latest IT Trends For Secure Mobile Collaboration. Forrester Consulting. [Online]. Available: http://www.connectedfuturesmag.com/docs/byod_forrester_tap_latest_it_trends_wp_en.pdf [retrieved: May, 2015]

[2] CISCO. VNI Mobile Forecast Highlights. CISCO Systems. [Online]. Available: http://www.cisco.com/c/dam/assets/sol/sp/vni/forecast_highlights_mobile/index.html [retrieved: May, 2015]

[3] V. Sacramento and et al., "MoCA: A Middleware for Developing Collaborative Applications for Mobile Users," Distributed Systems Online, IEEE, vol. 5, no. 10, Oct 2004, pp. 2–2.

[4] A. Salovaara and A. Oulasvirta, "Six modes of proactive resource management: a user-centric typology for proactive behaviors," in Proceedings of the third Nordic conference on Human-computer interaction. ACM, 2004, pp. 57–60.

[5] R.-A. Dobrican and D. Zampunieris, "Moving Towards Distributed Networks of Proactive, Self-Adaptive and Context-Aware Systems: a New Research Direction?" The International Journal on Advances in Networks and Services, vol. 7, 2014, pp. 262–272, ISSN: 1942-2644.

[6] G. I. F. Neyens, R.-A. Dobrican, and D. Zampunieris, "Enhancing Mobile Devices with Cooperative Proactive Computing," COLLA - The Fifth International Conference on Advanced Collaborative Networks, Systems and Applications, 2015, to be published.

[7] M. Slazynski, S. Bobek, and G. J. Nalepa, "Migration of Rule Inference Engine to Mobile Platform. Challenges and Case Study," in Proceedings of 10th Workshop on Knowledge Engineering and Software Engineering (KESE10) co-located with 21st European Conference on Artificial Intelligence (ECAI 2014), Prague, Czech Republic, August 19 2014., 2014. [Online]. Available: http://ceur-ws.org/Vol-1289/kese10-08_submission_4.pdf

[8] C. Choi, I. Park, S. J. Hyun, D. Lee, and D. H. Sim, "MiRE: A minimal rule engine for context-aware mobile devices," in Third IEEE International Conference on Digital Information Management (ICDIM), November 13-16, 2008, London, UK, Proceedings, 2008, pp. 172–177.

[9] R. Dobrican and D. Zampunieris, "A Proactive Approach for Information Sharing Strategies in an Environment of Multiple Connected Ubiquitous Devices," in Proceedings of the International Symposium on Ubiquitous Systems and Data Engineering (USDE 2014) in conjunction with 11th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2014). IEEE, 2014, pp. 763–771.

[10] J. Gabler, R. Klauck, M. Pink, and H. Konig, "uBeeMe - A platform to enable mobile collaborative applications," in Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on, Oct 2013, pp. 188–196.

[11] P. Coutinho and T. Rodden, "The FUSE Platform: Supporting Ubiquitous Collaboration Within Diverse Mobile Environments," Autom. Softw. Eng, vol. 9, 2002, pp. 167–186.

[12] D. Tennenhouse, "Proactive Computing," Communications of the ACM, vol. 43, no. 5, 2000, pp. 43–50.

[13] "Silence App," 2015, URL: https://play.google.com/store/apps/details?id=net.epsilonlabs.silence.ads [accessed: 2015-05-13].

[14] "Go Silent App," 2015, URL: https://play.google.com/store/apps/details?id=com.eventscheduler [accessed: 2015-05-13].

[15] "Advanced Silent Mode," 2015, URL: https://play.google.com/store/apps/details?id=com.joe.advancedsilentmode [accessed: 2015-05-13].

[16] "Silent Time," 2015, URL: https://play.google.com/store/apps/details?id=com.QuiteHypnotic.SilentTime&hl=en [accessed: 2015-05-13].

[17] "Auto Silent," 2015, URL: https://itunes.apple.com/us/app/autosilent/id474777148?mt=8 [accessed: 2015-05-13].

[18] Alastair Plumb. Slovakian Violist Lukas Kmit Interrupted By Nokia Ringtone, Incorporates It Into Recital. Huffington Post. [Online]. Available: http://www.huffingtonpost.co.uk/2012/01/23/slovakian-violinist-lukas-kmit-nokia-ringtone_n_1223086.html [retrieved: May, 2015]

[19] D. Zampunieris, "Implementation of a proactive learning management system," in Proceedings of" E-Learn-World Conference on E-Learning in Corporate, Government, Healthcare & Higher Education", 2006, pp. 3145–3151.

An Experience on Leadership Identification in Social Cognocracy Network

Alberto Turón
e-mail: turon@unizar.es

José María Moreno-Jiménez
e-mail: moreno@unizar.es

Juan Aguarón
e-mail: aguaron@unizar.es

María Teresa Escobar
e-mail: mescobar@unizar.es

Grupo Decisión Multicriterio Zaragoza (GDMZ)
Facultad de Economía y Empresa, Universidad de Zaragoza
Zaragoza, Spain

***Abstract*—In this work we show the results of a decision making experience conducted through the Social Cognocracy Network. This collaborative social network, developed by the Zaragoza Multicriteria Decision Making Group (GDMZ), is based on the principles that support the cognitive democracy known as e-cognocracy. The network considers three levels of interaction: information, content creation and decision making. E-cognocracy uses two rounds in order to incorporate preferences through an e-voting module and an intermediate round of discussion in which the arguments that support the individual positions and decisions are added by means of a forum. In addition, the forum provides quantitative measures that reflect the reputation of the actors and the relevance of topics and comments. This quantitative information is used to propose a procedure for the identification of the social leaders, the persons whose opinions influence the preferences of others.**

Keywords—Policy Making, Social Cognocracy Network, Reputation, Influence, Leadership Identification.

I. Introduction

E-cognocracy [1][2] is a cognitive democracy oriented to the extraction and sharing of knowledge associated with the scientific resolution of public decision making problems related with the governance of society. It uses two rounds in order to incorporate preferences through an e-voting module and an intermediate round of discussion in which the arguments that support the individual positions are added by means of a forum.

In the voting round the priorities associated to the alternatives are compared on two separate occasions; in the discussion step, the arguments for and against these alternatives, which are defended by the decision makers, are incorporated by posting messages and comments to the messages. In addition to the text that contains the decision makers' arguments, each post includes the information about three quantitative measures that reflect the importance given by the author and the reader to the post and the extent of their agreement with it. All this quantitative information and the relationships and influence indicators within the discussion network are used to propose a procedure for the identification of the social leaders –the persons whose opinions influence the preferences of others.

To do so, it is necessary to measure how an actor's vote is influenced by the opinions of the other actors.

II. Background

The experience involved a citizen's participation process based on the use of a social network, Social Cognocracy Network, that integrates the voting and discussion processes and provides tools to analyze the resulting preference structures.

A. *The Social Cognocracy Network*

Social Cognocracy Network (SCN) is a social network designed by the Zaragoza Multicriteria Decision Making Group (GDMZ). Based on the e-cognocracy, SCN promotes the citizen participation. Three participation levels are possible: (i) information (ii) content creation (iii) decision making. The citizens can achieve one of those participation levels depending on their involvement and their responsibility.

B. *Reputation, Relevance and Influence*

From the point of view of Social Cognocracy Network, three different types of nodes can be defined, according to their role in the process:

- Persons: individuals that take part in the voting and discussion processes, either decision makers or guests.
- Topics: discussion threads proposed, grouped into categories defined during the problem-design stage.
- Comments: any of the messages, arguments… expressed by the participants as messages published in the forum.

Each comment comes with the identification (nickname) of its author. All comments about a topic are grouped in a page under a header containing the topic title.

Two basic indicators of influence are defined:

- Reputation of an actor: measurement of the prestige the actor has among the participants.

 ISBN: 978-1-61208-436-7

- Relevance of a topic or a comment: measurement of the capability of the topic or comment to shape the actors' opinions.

A person P_j can give his/her opinion about other persons, topics and comments, by giving values from 0 to 10 to three quantitative indices:

Index	Rates
Confidence C_{ij}	The author P_i of a comment
T-Importance I_{ij}^{T}	A topic T_i
C-Importance I_{ij}^{C}	A comment C_i

Persons can also assess the importance of their own topics and comments and even to themselves (*self-trust*).

From these indices, the reputation R_i^{P} of persons and the relevance of topics (R_i^{T}) and comments (R_i^{C}) are obtained using the expressions:

$$R_i^P = \frac{\sum_{j=1}^{n} R_j^P C_{ij}}{\sum_{j=1}^{n} R_j^P} \quad (1)$$

$$R_i^T = \frac{\sum_{j=1}^{n} R_j^P \mathrm{I}_{ij}^T}{\sum_{j=1}^{n} R_j^P} \quad (2)$$

$$R_i^C = \left(1 + \frac{n_c}{N}\right) \frac{\sum_{j=1}^{n} R_j^P \mathrm{I}_{ij}^C}{\sum_{j=1}^{n} R_j^P} \quad (3)$$

being N the total number of comments that a topic has received and n_c the number of answers to a specific comment in that topic.

A person's reputation, as well as the importance the community gives to a topic or a message, can influence the preferences of other persons. Under the e-cognocracy, this influence will result in changes in the preference structures of these other persons from one round to another.

III. CASE STUDY

Social Cognocracy Network was used in an experience related to the design of the metropolitan public transport network of the city of Zaragoza (Spain). Four alternatives were proposed. Representatives of political parties defended their proposals in a classroom with students of E-government and Public Decision Making (4th course of the Degree in Economy at University of Zaragoza). After a first voting round, in which only the students were allowed to participate, a discussion was developed in the forum, with the participation of the students and the political representatives. In the forum, the participants could valuate the reputation of the others, as well as the importance of the comments posted (Figure 1). Then, a second voting round was performed. The voters' preferences were expressed by using Analytic Hierarchy Process [3].

IV. CONCLUSIONS

Numerical and visual analyses show the influence of the opinions in the change of the actors' preference structures. The quantitative valuations of the posts allow calculating the influence of the different participants and their comments. An interactive 3D visualization tool is used to explore the results (Figure 2). Individual preference structures are represented on the simplex $x_1+x_2+x_3+x_4=1$, being x_i the preferences given by each voter to each alternative A_i, $i=1,\ldots,4$. From the visual analysis some relevant facts stand out: 24 zones define the different positions of the voters with respect to the four alternatives; although the results of the two rounds offer few significant differences, only an inversion in the preferences of the alternatives chosen in second and third places, it seems clear that the participants with greater reputation persuaded to take their opinion (influenced) voters who, in the first round, had favored other alternatives. Analyzing specifically the behavior of these voters, it can be observed that the change in their opinion is in line with the opinions expressed by the four best valued voters. The importance of the comments in which these opinions have been exposed is also high.

ACKNOWLEDGMENT

This research was partially funded by the Spanish Ministry of Economy and Innovation Research Project "Social Cognocracy Network" (Ref. ECO2011-24181).

REFERENCES

[1] J. M. Moreno-Jiménez, "E-cognocracia: Nueva Sociedad, Nueva Democracia," ("*E-cognocracy: New Society, New Democracy*"), Estudios de Economía Aplicada, vol. 24(1–2), 2006, pp. 559–581.

[2] J. M. Moreno-Jiménez, and W. Polasek, "E-democracy and Knowledge. A Multicriteria Framework for the New Democratic Era," Journal Multicriteria Decision Analysis, vol. 12, 2003, pp. 163–176.

[3] T. Saaty, "The Analytic Hierarchy Process," McGraw-Hill, New York, 1980.

Figure 1. Discussion window with the thread of comments of a topic, showing the input fields used for assessing the reputation of the actors (persons) and the relevance of the topics and comments.

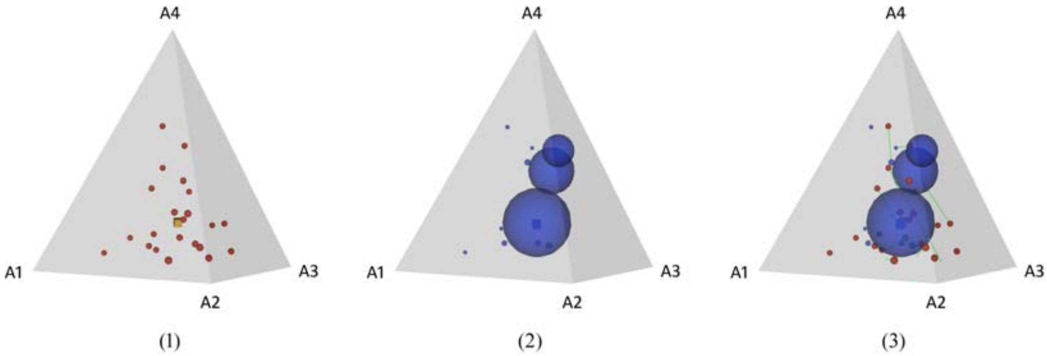

Figure 2. Several 3D views of the simplex showing the voters' preference structures: (1) after the first round (2); after the second round, being the size of each point proportional to the reputation of the voter; (3) changes in the preference structures of each voter. The box represents the overall (group) preference structure.

An Attempt to Evaluate Chances and Limitations of Social Information Retrieval

Christoph Fuchs
Department of Informatics
TU München
Garching bei München, Germany
Email: fuchsc@in.tum.de

Georg Groh
Department of Informatics
TU München
Garching bei München, Germany
Email: grohg@in.tum.de

***Abstract*—We present an approach to evaluate a novel concept for distributed social information retrieval. The concept is based on the idea that users can query private information spaces of socially close people ("friends"), facilitating social interactions that correspond with typical human information sharing behavior to a higher extent. Thereby, we hope to establish more efficient and socially compatible information sharing among peers in social and collaborative networks. We give a short overview on related research from information science, psychology, and economics, explain and motivate our research questions, summarize our early findings, and sketch the setup of our upcoming empirical experiment.**

Keywords–Social search; distributed social information retrieval; information seeking; social networking

I. INTRODUCTION

Relying on our social network to satisfy information needs is a strategy that is deeply linked to human social behavior [1]. Social media heavily builds upon the users' willingness to participate and share information. While current social networking sites like Facebook, Google+, or Twitter offer features to target information items to a specific audience, they don't facilitate social information retrieval in a way that optimally corresponds with human behavior. Users' readiness to share previously unshared information is impacted by a set of social mechanisms. We would like to leverage these concepts in order to provide a more efficient way of distributed social information retrieval, allowing users to benefit from collaboration with their contacts. Thereby, we focus on a scenario, where information seekers can query other users' information spaces (related to asking questions in normal human-human interaction). An information space constitutes a repository of private information items, generated by (but not necessarily limited to) analyzing the user's actions (e.g., web browsing, online transactions, communication) and contextual data (e.g., location, type of activity, other people present). Information seekers can query the information spaces of others (referred to as information providers) using their agent (e.g., implemented in a mobile device). An information provider's agent which received a query would analyze the information provider's private information space and – subject to the concrete configuration – recommend matching information items to the information provider as potential items for sharing. Upon confirmation, the results and additional comments given from the information provider would get shared with the information seeker.

The paper's main contribution is the documentation of the social search concept and the method to evaluate it. Other approaches for social search [2], [3], distributed social networks [4], or P2P file sharing [5] focus on questions related to the technical implementation, whereas our work aims to offer a different perspective on the topic, namely the social mechanisms that influence information dissemination in social networks fostering collaboration among users and the potential benefits from integrating social context.

The paper is structured as follows: Section II explains and motivates the research questions, Section III details relevant areas of research, which either build a basis for our concept or describe an approach that goes a similar track (in the latter case, the differences to our concept are explained). In Section IV, we introduce the planned empirical experiment, state how we will answer the research questions, and present first results obtained from pre-studies.

II. RESEARCH QUESTIONS

To evaluate the corner stones of a social information retrieval system as outlined above, we analyzed literature in related fields (computer science, psychology, and economics). For some assumptions, documented research does either not provide an unambiguous answer or does not take the specific circumstances into account. Therefore, we planned to design a large-scale experiment which should help to answer the following questions:

A. RQ1: How do social context and interaction archetypes influence users' data sharing sensitivity?

One of the basic ideas of social media is that people proactively share information with a wider audience (e.g., a group of other users considered as "friends"). One possible reason why social media can't harness its full potential in reality is that a large number of positive use cases rely on users who share the information – what they don't do to the required extent. We would like to analyze to which degree social context, i.e., the social closeness of information seeker and provider, and the type of interaction (e.g., directed/broadcasted request/reply, anonymous/not anonymous request/reply) influence the users' willingness to ask for or share information. It seems to be reasonable that the type of interaction may influence the users' willingness to share information with other users. This hypothesis is going to be evaluated with the dataset obtained from the described experiment. By considering the natural information sharing preferences of users in the design of a distributed social information retrieval system we hope to increase the amount of individually available information for social search.

 ISBN: 978-1-61208-436-7

B. RQ2a: How relevant are information items taken from non-public information spaces of socially close people when satisfying information needs?

When solving information needs, having access to originally private information does not necessarily result in higher satisfaction levels for the information seeker. Unpublished information held in private information spaces might, e.g., not have undergone the same degree of rigorous review as published work, might be common everyday knowledge or irrelevant and therefore possibly not be suited to satisfy information needs. The main objective of this research question is to verify whether friends' private information spaces contain relevant information.

C. RQ2b: Does social context imply a valuable contribution to retrieving information from the unconscious information need (serendipitous information)?

Theories on homophily [6] suggest that socially close people have similar preferences and therefore keep information items that are of potential (mutual) interest in their information spaces. Referring to Mizzaro's model of relevance [7] we would like to investigate whether information spaces of socially close users could foster finding information items that are considered as serendipitous by the information seeking users.

D. RQ3: Which social concepts impact the users' routing decisions?

In a majority of existing approaches for P2P document retrieval systems, routing decisions for queries are based mainly on content characteristics (i.e., does a specific node store a certain document?). In these settings, the social relationship between information seeker and provider is not important – the document is standardized and not linked to the specific social context: It is not important for the seeker where the document comes from – as long as it contains the expected content. In scenarios where social search is expected to perform best [8], the source of an information can be of importance. Selecting a user as an information provider is not only based on the availability of content, but on other (more social) criteria as well (referred to as "social concepts"). Given a system which allows routing of queries to potential information providers within ones own social network to satisfy information needs, we would like to understand why some people are chosen as information providers and others not. Borgatti and Cross [9] already published some theories on social interaction in information retrieval, but focus only on professional settings (in the work environment of large organizations). We assume that a more general situation might reveal different results since the workplace implies a specific set of rules, which do not apply in broader environments.

E. RQ4: Which category of information needs could benefit from social information retrieval?

Traditional web search engines excel at finding published factual information, like Mozart's date of birth. The information items in the information spaces of others might be much more subjective than publicly available information – therefore they might be relevant for certain types of information needs, which would profit from various subjective recommendations and/or opinions, as Oeldorf-Hirsch et al. [8] already suggested for SMQA (it is important to note that Oeldorf-Hirsch et al.'s study only allowed public broadcasting of questions and therefore investigated a different setting).

III. RELEATED WORK

A. Information Retrieval

1) Models & Approaches: Elementary concepts of information retrieval approaches include traditional vector space models based on *term frequency-inverse document frequency* (TF-IDF, [10, p. 118]), where documents are represented as vectors in a multi-dimensional vector space. The dimensions are defined by the terms which are derived from the words occurring in the documents of the collection. The position of a document within this vector space is defined by calculating the *term frequency* (how often does a specific word occur in a document) and the *inverse document frequency* (how many documents of the collection contain the word). By combining those two factors, it is possible to identify words describing the prevalent content of the document and at the same time differentiating the document from the other documents in the collection [10]. One of the main assumptions in most TF-IDF based approaches is that the order of words does not matter (*bag-of-words assumption*). BM25F, an improved version of BM25 (Best Matching), also relies on the bag-of-word assumption but distinguishes between different fields of a document and adjusts the weighting according to the importance of the respective field [11]. More recent approaches, like *term weights-IDF* (TW-IDF) [12], use a graph-based representation which outperforms classical approaches, like BM25, by considering the relations between terms using a unweighted directed graph.

Our work builds upon these concepts: each user's information space has to be indexed using those classical techniques in combination with probabilistic elements like topic models/Latent Dirichlet Allocation (LDA) [13] to suggest personal information items to be shared with an information seeker.

2) Relevance & Serendipity: In naive TF-IDF approaches, the relevance of a document for a query is often calculated using a metric like cosine similarity or Euclidian distance to compare document vectors and query vector. Beyond this mechanical way of calculating relevance, Mizzaro [7] distinguished between *real ("objective") information need* and *perceived ("subjective") information need* which have some overlap, but are not equal. The information seeking user is only aware of the perceived information need and uses this as a starting point when formulating the query. During the search process, the user's mental model about the topic of interest evolves and the user's subjective information need iteratively changes while consuming more information (and ideally would cover more of the real information need). As a consequence, information items might be relevant according to the user's *real* information need, but might not be considered as relevant by an algorithm which is designed to maximize the relevance for the user's query (because the latter is derived from the *perceived* information need). This allows *serendipity*, where items do not necessarily fit the entered query but are considered as relevant by the user. Previous literature covers measurement [14], exploration [15], [16], and formalization [17], [18], [19] of serendipity.

In our work, we would like to evaluate which social relations make occurrences of serendipity happen more likely to allow the user to consciously manipulate the answer's degree of serendipity by selecting information providers from different social communities. Thereby, we define serendipity as combination of relevance and unexpectedness of the result.

3) Distributed & Personalized Search: Several approaches for *Distributed Search* have been proposed in the past, examples relying on agents are (among others) DS4 [2], [3], Blogracy [20] (where the authors also provide a comprehensive overview of distributed social networking approaches), RAIS [21], or DIAMS [22]. YaCy [23] forms a distributed index for web search and Callan [24] provides an early overview of distributed information retrieval, coining the term "federated search" where an information seeker queries several search engines in parallel. Shokouhi and Si [25] profoundly summarize approaches for the sub-steps in federated search. To include the information seeker's individual context in the evaluation of potential search results, several attempts have been made to *personalize* search results. Micarelli et al. [26], Steichen et al. [27], and Ghorab et al. [28] published comprehensive surveys, clustering the existing approaches. Some approaches also personalize results based on the information seeker's social network [29] or use the social network to rank information [30]. Carmel et al. [31] compare different strategies to use information obtained from different social networks (familiarity-based, similarity-based, overall network with both types of edges) to personalize search results. Their comparison with personalization based on topics suggests that all three personalization strategies relying on social networks outperform the topic-based approach (e.g., indexing and ranking).

Our intended concept combines several characteristics of other approaches: The idea of selecting different information providers (and thus repositories) is comparable to federated search [24] or DS4 [2], [3], where the social component is not taken into account to the same extent. Other approaches like YaCy [23], Carmel et al.'s [31], or SNDocRank [30] cover parts of the search process.

B. Social Search

1) Definition: McDonnell and Shiri [32] list a variety of definitions for social search; for the remaining part of this chapter, we define social search broadly as integrating others in the search process and therefore are very close to Evan et al.'s definition [33].

2) Social Context: By analyzing usage patterns of mobile search, Teevan et al. [34] and Church and Oliver [35] show that social context highly influences the search process, either by searching collaboratively or by discussing the search results with others. While Teevan et al. and Church/Oliver only consider the short-term social context during the search process, Kramr et al. [36] use clusters of users with similar interests to disambiguate queries (and thus rely on the long-term social context). After having conducted an online survey with 150 participants, Evans and Chi [37] conclude that social interactions "play a key role throughout the search process". Their findings suggest that existing tools do not fully meet the users' requirements.

Our objective is to understand the users' social behavior and to build a concept to improve information dissemination among users.

3) Social Media Question Asking: While many studies cover the social aspects of search performed using traditional web search engines, some also investigate social interaction when users try to get information from their social contacts. In *Social Media Question Asking* (SMQA), information seekers satisfy information needs by asking other people via social network platforms like Twitter, Facebook, or Google+. Efron and Winget [38] propose a taxonomy of questions asked in a microblogging environment, Paul et al. [39], and Teevan et al. [40] identify patterns for question asking and answering, Lampe et al. [41] investigate Facebook's value as an information service while Oeldorf-Hirsch et al. [8] compare SMQA with searching on traditional web search engines.

Those findings show that there are specific types of information needs which people prefer to solve by leveraging the knowledge of their peers. Our concept aims to improve this process by reducing involved social costs and increasing the efficiency of the process (e.g., by recommending information items, which might be suitable answers, to the information provider – in later versions, an automatic reply could also be possible).

C. Motivation to share information

Alan P. Fiske suggests in [42] that human social life could be explained by combining four psychological models, namely *communal sharing* (CS), *authority ranking* (AR), *equality matching* (EM), and *market pricing* (MP). Following this approach, information sharing could be considered as a social act, allowing to express the underlying motivation as a combination of Fiske's models. In CS, people treat members of their specific group as equivalents. People within the group behave altruistic and are sometimes linked by kinship. In AR, people are ordered linearly according to some social hierarchical dimension. People with higher ranks typically have privileges, prestige, and prerogatives, which people with lower ranks don't have. EM describes a relation between two people who try to keep the balance of their relationship even. This is the standard behavior between people who meet multiple times and follow a tit-for-tat strategy or some other reciprocal granting of favors. In contrast, in MP relationships all relevant features are reduced to a lower dimensional value or utility metric (e.g., price) that is used to compare different factors. This is the default relationship for people who only meet once and don't plan any further encounters. Applying these concepts to distributed social search, the motivation to provide information to the information seeker highly depends on the type of relationship: Following a CS regime, people would be much more interested in sharing information while offering information to socially more distant friends or even strangers would follow a more strict EM or even MP regime. AR regimes could be characteristic for certain professional settings.

Manski [43] models social interactions based on the concepts *expectation* (agents choose actions based on the experience of others who had the same problem), *constraint* (the respective good is limited and therefore needs to be shared/allocated wisely), *preference* (own choice depends on others' choice), and *equilibrium* (occurs when all agents' actions are mutually consistent). Jackson and Rogers [44] analyze theoretic games in social networks and distinguish between *strategic complements* and *strategic substitutes*. Examples for strategic complements are, e.g., the majority game (a user's payoff is higher when she/he does the same as her/his neighbors), which could model for example the adoption of new technologies. Strategic substitutes describe situations where a user's payoff is lower when she/he does the same as her/his neighbors (e.g., best shot public goods game; it doesn't make any sense for a user to buy a book when she/he can borrow it from her/his neighbors). With chronobot [45], Li and Chang implemented a bidding system for tasks with the required time as currency, including a proposal to determine exchange rates according to preference or expertise levels. Using Fiske's model above, this system would follow a clear MP approach, offering standardized prices calculated based on objective input parameters. Social relationships and human traits in sharing information are not explicitly considered.

By conducting our planned experiment, we would like to understand users' social behavior when sharing or asking for information in the social information retrieval scenario described in in the introduction section. We will build upon Fiske's work [42] and use it as a framework to distinguish different types of social interactions. Assuming that users are acting rationally, Jackson and Roger's theories [44] might show parallels when analyzing the underlying market mechanisms of our scenario.

IV. First Results & Proposed Research Outline

We plan to conduct a larger experiment on distributed information retrieval with $150 - 200$ participants within the coming months, allowing us to generate empirical data to answer the research questions outlined above. In some cases, we already did pre-studies to obtain first insights. The participants will disclose their individual social networks, assign weights to their social edges (tie strength, knowledge similarity, social context similarity, sympathy) and provide an index to something we consider as their private information space (visited URLs on the web, extracted from their browser history, crawled and indexed using LDA [13]). In addition, we will receive information about individually viewed and bought products from a leading online store. The experiment consists of three parts (manual query mode, automatic query mode, and semantic product search):

Manual query mode – Participants define three queries they would like to solve by asking people within their social network. During query definition and assignment of potential information providers to the queries, users are asked to justify their decisions. In addition to the self-defined queries, each participant will also be asked to satisfy three predefined information needs taken from domains which are suitable for social information retrieval (based on [8]). Information providers will be asked to answer the query and to fill out an online survey, information seekers are expected to review the received results and rate them.

Automatic query mode – Participants define queries, a background task uses a randomly selected strategy to choose potential information providers (possible strategies include tie strength, knowledge similarity, social context similarity, and sympathy). Since all participants of the experiment uploaded an index to their private information space in advance, it is possible to query the index of the identified group of information providers in the background. In case of any matches, the respective information provider is informed about the incoming query and the identified result within her/his information space and is asked to provide the information (i.e., the URL to the site corresponding to the index position) and fill out a survey. After the information seeker received the response, she/he is also asked to evaluate the results. To allow further comparisons, one of the answering information providers for each query will be an (undisclosed) technical user account, querying a traditional web search engine and returning the first five resulting URLs to the information seeker.

Semantic product search – The participants will be asked to search for items using a customized user interface to a well-known online store, where items which have been viewed at and/or bought by friends are highlighted. In addition to the results obtained from the normal search functionality of the online store, products viewed/bought by close friends (identified by various strategies) will be added to the result list (without revealing the friends' identities). Those additional products do not necessarily match the search query exactly, but might be considered as relevant due to the social relationship between the product owner/viewer and the information seeker. Participants will be asked to evaluate usefulness and degree of predictability/novelty for each item in the result list. In addition, click behavior will be recorded.

A. RQ1: How do social context and interaction archetypes influence users' data sharing sensitivity?

In a small pre-study (online survey with 112 participants) [46], we re-run a modified and reduced version of Oeldorf-Hirsch et al.'s [8] experiment. Oeldorf-Hirsch's observation was that people are quite hesitant to ask others for help in a social media environment. Our hypothesis was that this might be caused by the fact that users had to post questions visible to a majority or all of their friends ("broadcast") on the social network platform. Our findings suggest that people's willingness to ask for information highly depends on the audience – if it is possible to target a single recipient or a limited audience, social means for information retrieval are considered much more often.

In the upcoming experiment, we plan to ask the user who provides an information item whether she/he has already shared the information item on any public media channel (like Facebook, Twitter, Google+, etc.), whether she/he would share it on a social media channel, and whether she/he would share it with a friend who asks for it. In addition, we ask the potential information provider to assess the information item's degree of privacy (using a slider on a scale from 0 to 100, with expressive descriptions for minimum/maximum values). One possible outcome is that the degree of information sharing (i.e., how much "privacy" does someone share?) highly depends on the social context (audience) and the type of interaction (reactive, proactive).

B. RQ2a: How relevant are information items taken from non-public information spaces of socially close people when satisfying information needs?

Analyzing two datasets obtained from Twitter and Facebook [47], our early findings suggest that content created by socially close people is of higher interest for us than content from strangers. During the social information retrieval experiment we will ask the information seekers to assess the quality of responses given by the information providers in order to obtain a measureable value for relevance, novelty, and personalization. In addition, we plan to correlate the results with social metrics like tie strength, knowledge similarity, social context similarity, and sympathy.

C. RQ2b: Does social context imply a valuable contribution to retrieving information from the unconscious information need (serendipitous information)?

Focusing on the use case to buy a product, we will compare the relevance of product items returned from a well-known online store with those items taken from the list of viewed and bought items from people within the own, individual social network (defined using the max/min values of tie strength, knowledge similarity, social context similarity, and sympathy). We will compare the degree of relevance and novelty between the different groups of origin.

D. RQ3: Which social concepts impact the users' routing decisions?

When adding a user as a recipient of a search query in the manual query mode of the experiment, we will ask the user to justify her/his choice. During the assessment of the result quality, we ask the user which other contacts could have also been potential information providers (and why the user hesitated to nominate them). We also gather information about the motives for sharing information to be able to describe the relationship between the information seeker and provider using Fiske's [42] model.

E. RQ4: Which category of information needs could benefit from social information retrieval?

In a different study, we used highly specialized websites as proxies for specific topics to derive the degree of socialness for the topic. After a first initial pre-study we plan to elaborate on this, scaling the experiment with the help of crowdsourcing platforms and a larger URL database. In the upcoming study, we will ask the participants to provide queries which are considered to be suited for social information retrieval. We plan to validate these expectations using the quality assessments of the results and will compare these findings with previous literature.

V. Summary & Conclusion

The objective of this paper is to propose a concept for "distributed social information retrieval" and a possible evaluation approach. We explained and motivated the research questions, described first results and outlined the agenda for the following experiment. By understanding the underlying psychological details, we would like to create a system that facilitates information retrieval among social peers and allows to efficiently incorporate the knowledge that is available within one's own social network.

References

[1] C. A. Johnson, "Choosing People: The Role of Social Capital in Information Seeking Behaviour," Information Research, vol. 10, no. 1, 2004. [Online]. Available: http://InformationR.net/ir/10-1/paper201.html [retrieved: 2015.07.21]

[2] D. Kontominas, P. Raftopoulou, C. Tryfonopoulos, and E. G. Petrakis, "DS4: A Distributed Social and Semantic Search System," in Advances in Information Retrieval – 35th European Conference on IR Research, ser. ECIR '13. Springer, 2013, pp. 832–836.

[3] P. Raftopoulou, C. Tryfonopoulos, E. G. Petrakis, and N. Zevlis, "DS4: Introducing Semantic Friendship in Distributed Social Networks," in Proceedings of the 21st International Conference on Cooperative Information Systems, ser. CoopIS '13, 2013.

[4] H. Li, K. Bok, and J. Yoo, "An Efficient Mobile Social Network for Enhancing Contents Sharing over Mobile Ad-hoc Networks," in Proceedings of the International Conference on Parallel and Distributed Computing, Applications and Technologies, ser. PDCAT, 2012, pp. 111 – 116.

[5] K. Chen, H. Shen, and H. Zhang, "Leveraging Social Networks for P2P Content-Based File Sharing in Mobile Ad Hoc Networks," in Proceedings of the International Conference on Mobile Adhoc and Sensor Systems, ser. MASS, 2011.

[6] D. Kempe, J. Kleinberg, S. Oren, and A. Slivkins, "Selection and Influence in Cultural Dynamics," arXiv:1304.7468 [cs.GT], 2013. [Online]. Available: http://arxiv.org/abs/1304.7468 [retrieved: 2015.07.21]

[7] S. Mizzaro, "How Many Relevances in Information Retrieval?" Interacting with Computers, vol. 10, no. 3, 1998, pp. 303–320.

[8] A. Oeldorf-Hirsch, B. Hecht, M. R. Morris, J. Teevan, and D. Gergle, "To Search or to Ask: The Routing of Information Needs Between Traditional Search Engines and Social Networks," in Proceedings of the 17th Conference on Computer Supported Cooperative Work & Social Computing, ser. CSCW '14. New York, NY, USA: ACM, 2014, pp. 16–27.

[9] S. P. Borgatti and R. Cross, "A Relational View of Information Seeking and Learning in Social Networks," Management Science, vol. 49, no. 4, 2003, pp. 432–445.

[10] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval. Cambridge University Press, 2008.

[11] H. Zaragoza, N. Craswell, M. Taylor, S. Saria, and S. Robertson, "Microsoft Cambridge at TREC-13: Web and HARD Tracks," in Proceedings of TREC-2004, 2004. [Online]. Available: http://trec.nist.gov/pubs/trec13/papers/microsoft-cambridge.web.hard.pdf [retrieved: 2015.07.21]

[12] F. Rousseau and M. Vazirgiannis, "Graph-of-Word and TW-IDF: New Approach to Ad Hoc IR," in Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, ser. CIKM '13. New York, NY, USA: ACM, 2013, pp. 59–68.

[13] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," The Journal of Machine Learning Research, vol. 3, 2003, pp. 993–1022.

[14] I. Bordino, Y. Mejova, and M. Lalmas, "Penguins in Sweaters, or Serendipitous Entity Search on User-Generated Content," in Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, ser. CIKM '13. New York, NY, USA: ACM, 2013, pp. 109–118.

[15] M. Dörk, S. Carpendale, and C. Williamson, "The Information Flaneur: A Fresh Look at Information Seeking," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '11, no. 10. New York, NY, USA: ACM, 2011, pp. 1215–1224.

[16] M. Dörk, N. H. Riche, G. Ramos, and S. Dumais, "PivotPaths: Strolling Through Faceted Information Spaces," in IEEE Transactions On Visualization and Computer Graphics, ser. InfoVis '12, vol. 18, no. 12. IEEE, 2012, pp. 2709–2718.

[17] T. E. Workman, M. Fiszman, T. C. Rindflesch, and D. Nahl, "Framing Serendipitous Information-Seeking Behavior for Facilitating Literature-Based Discovery: A Proposed Model," Journal of the Association for Information Science and Technology, vol. 65, no. 3, 2014, pp. 501–512.

[18] A. Thudt, U. Hinrichs, and S. Carpendale, "The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries Through Information Visualization," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '12. New York, NY, USA: ACM, 2012, pp. 1461–1470.

[19] M. Schedl, D. Hauger, and D. Schnitzer, "A Model for Serendipitous Music Retrieval," in Proceedings of the 2nd Workshop on Context-awareness in Retrieval and Recommendation, ser. CaRR '12. New York, NY, USA: ACM, 2012, pp. 10–13.

[20] E. Franchi, A. Poggi, and M. Tomaiuolo, "Supporting Social Networks With Agent-Based Services," International Journal of Virtual Communities and Social Networking, vol. 5, no. 1, 2013, pp. 62–74.

 ISBN: 978-1-61208-436-7

[21] M. Mari, A. Poggi, M. Tomaiuolo, and P. Turci, "Enhancing Information Sharing Through Agents," in Proceedings of the 8th International BI Conference on Agent-oriented Information Systems IV, 2006, pp. 202–211.

[22] J. R. Chen, S. R. Wolfe, and S. D. Wragg, "A Distributed Multi-Agent System for Collaborative Information Management and Sharing," in Proceedings of the Ninth International Conference on Information and Knowledge Management, ser. CIKM '00. ACM, 2000, pp. 382–388.

[23] M. Christen, "YaCy Decentralized Web Search," 2015. [Online]. Available: http://yacy.net/en/index.html [retrieved: 2015.07.21]

[24] J. Callan, "Distributed Information Retrieval," The Information Retrieval Series, vol. 7, 2000, pp. 127–150.

[25] M. Shokouhi and L. Si, "Federated Search," Foundations and Trends in Information Retrieval, vol. 5, no. 1, 2011, pp. 1–102.

[26] A. Micarelli, F. Gasparetti, F. Sciarrone, and S. Gauch, "Personalized Search on the World Wide Web," Lecture Notes in Computer Science, vol. 4321, 2007, pp. 195–230.

[27] B. Steichen, H. Ashman, and V. Wade, "A Comparative Survey of Personalised Information Retrieval and Adaptive Hypermedia Techniques," Information Processing & Management, vol. 48, no. 4, 2012, pp. 698–724.

[28] M. R. Ghorab, D. Zhou, A. O'Connor, and V. Wade, "Personalised Information Retrieval: Survey and Classification," User Modeling and User-Adapted Interaction, vol. 23, no. 4, 2013, pp. 381–443.

[29] D. Lu and Q. Li, "Personalized Search on Flickr Based on Searcher's Preference Prediction," in Proceedings of the 20th International Conference Companion on World Wide Web, ser. WWW '11. New York, NY, USA: ACM, 2011, pp. 81–82.

[30] L. Gou, X. L. Zhang, H.-H. Chen, J.-H. Kim, and C. L. Giles, "Social Network Document Ranking," in Proceedings of the 10th Annual Joint Conference on Digital Libraries, ser. JCDL '10. New York, NY, USA: ACM, 2010, pp. 313–322.

[31] D. Carmel, N. Zwerdling, I. Guy, S. Ofek-Koifman, N. Har'el, I. Ronen, E. Uziel, S. Yogev, and S. Chernov, "Personalized Social Search Based on the User's Social Network," in Proceedings of the 18th ACM Conference on Information and Knowledge Management, ser. CIKM '09. New York, NY, USA: ACM, 2009, pp. 1227–1236.

[32] M. McDonnell and A. Shiri, "Social Search: A Taxonomy of, and a User-Centred Approach to, Social Web Search," Program: Electronic Library and Information Systems, vol. 45, no. 1, 2011, pp. 6–28.

[33] B. M. Evans, S. Kairam, and P. Pirolli, "Exploring the Cognitive Consequences of Social Search," in CHI Extended Abstracts on Human Factors in Computing Systems, ser. CHI EA '09. New York, NY, USA: ACM, 2009, pp. 3377–3382.

[34] J. Teevan, A. Karlson, S. Amini, A. J. B. Brush, and J. Krumm, "Understanding the Importance of Location, Time, and People in Mobile Local Search Behavior," in Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, ser. MobileHCI '11. New York, NY, USA: ACM, 2011, pp. 77–80.

[35] K. Church and N. Oliver, "Understanding Mobile Web and Mobile Search Use in Today's Dynamic Mobile Landscape," in Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, ser. MobileHCI '11. New York, NY, USA: ACM, 2011, pp. 67–76.

[36] T. Kramár, M. Barla, and M. Bieliková, "Disambiguating Search by Leveraging a Social Context Based on the Stream of User's Activity," Lecture Notes in Computer Science, vol. 6075, 2010, pp. 387–392.

[37] B. M. Evans and E. H. Chi, "Towards a Model of Understanding Social Search," in Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, ser. CSCW '08. New York, NY, USA: ACM, 2008, pp. 485–494.

[38] M. Efron and M. Winget, "Questions are Content: A Taxonomy of Questions in a Microblogging Environment," in Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem, ser. ASIS&T '10. American Society for Information Science, 2010, pp. 27:1–27:10.

[39] S. A. Paul, L. Hong, and E. H. Chi, "Is Twitter a Good Place for Asking Questions? A Characterization Study," in Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, 2011.

[40] J. Teevan, M. R. Morris, and K. Panovich, "Factors Affecting Response Quantity, Quality, and Speed for Questions Asked via Social Network Status Messages," in Proceedings of the 5th International Conference on Weblogs and Social Media, ser. ICWSM '11. Association for the Advancement of Artificial Intelligence (AAAI), 2011, pp. 630–633.

[41] C. Lampe, J. Vitak, R. Gray, and N. Ellison, "Perceptions of Facebook's Value as an Information Source," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '12. New York, NY, USA: ACM, 2012, pp. 3195–3204.

[42] A. P. Fiske, "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations," Psychological Review, vol. 99, no. 4, 1992, pp. 689–723.

[43] C. F. Manski, "Economic Analysis of Social Interactions," NBER Working Paper Series, no. 7580, 2000. [Online]. Available: http://www.nber.org/papers/w7580 [retrieved: 2015.07.21]

[44] M. O. Jackson and Y. Zenou, "Games on Networks," Handbook of Game Theory, vol. 4, 2014.

[45] X. Li and S.-K. Chang, "User Profiling in the Chronobot/Virtual Classroom System," International Journal of Software Engineering and Knowledge Engineering, vol. 17, no. 2, 2007, pp. 191–206.

[46] C. Fuchs and G. Groh, "Appropriateness of Search Engines, Social Networks, and Directly Approaching Friends to Satisfy Information Needs," in Proceedings of the 5th Workshop on Social Network Analysis in Applications, ser. SNAA '15, Paris, France, August 2015, forthcoming.

[47] C. Fuchs, J. Hauffa, and G. Groh, "Does Friendship Matter? An Analysis of Social Ties and Content Relevance in Twitter and Facebook," in Proceedings of the Service Summit Workshop and Service Summit 2015. Karlsruhe Institute of Technology, 2015.