COLLA 2013

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

ISBN: 978-1-61208-287-5

July 21 - 26, 2013

Nice, France

COLLA 2013 Editors

Lasse Berntzen, Vestfold University College - Tønsberg, Norway

Petre Dini, Concordia University, Canada / China Space Agency Center, China

COLLA 2013

Foreword

The Third International Conference on Advanced Collaborative Networks, Systems and Applications (COLLA 2013), held between July 21 and July 26, 2013 in Nice, France, 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.

We take here the opportunity to warmly thank all the members of the COLLA 2013 Technical Program Committee, as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to COLLA 2013. 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 are grateful to the members of the COLLA 2013 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that COLLA 2013 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of advanced collaborative networks, systems and applications.

We are convinced that the participants found the event useful and communications very open. We hope that Nice, France provided a pleasant environment during the conference and everyone saved some time to enjoy the charm of this city.

COLLA 2013 Chairs:

COLLA Advisory Committee

Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Mohamed Eltoweissy, Pacific Northwest National Laboratory / Virginia Tech, USA
Jaime Lloret Mauri, Universidad Politécnica de Valencia, Spain
Abdulrahman Yarali, Murray State University, USA

COLLA 2013

Committee

COLLA Advisory Committee

Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Mohamed Eltoweissy, Pacific Northwest National Laboratory / Virginia Tech, USA
Jaime Lloret Mauri, Universidad Politécnica de Valencia, Spain
Abdulrahman Yarali, Murray State University, USA

COLLA 2013 Technical Program Committee

Imad Abugessaisa, Karolinska Institutet - Stockholm, Sweden
Leila Alem, CSIRO, Australia
Yasmine Arafa, University of East London, UK
Marcelo G. Armentano, ISISTAN Research Institute (CONICET/UNICEN), Argentina
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
Cornelia Boldyreff, University of East London, UK
Stainam Brandao, COPPE/Federal University of Rio de Janeiro, Brazil
Patrick Brezillon, LIP6, University Pierre and Marie Curie (UPMC) - Paris, France
Shu-Ching Chen, Florida International University - Miami, USA
Shiping Chen, CSIRO, Australia
Xueqi Cheng, Institute of Computing Technology / Chinese Academy of Sciences, China
Jin-Hee Cho, U.S. Army Research Laboratory (USARL) - Adelphi, USA
Kevin Curran, University of Ulster - Derry, Northern Ireland, UK
Claudia d'Amato, University of Bari, Italy
Fulvio D'Antonio, Polytechnic University of Marche, Italy
Arianna D'Ulizia, IRPPS-CNR, Italy
Antonio De Nicola, ENEA, Italy
Jerome Dinet, Univeristé Paul Verlaine - Metz, France
Anna Divoli, Pingar Research, New Zealand
Wael M El-Medany, University Of Bahrain, Bahrain
Mohamed Eltoweissy, Pacific Northwest National Laboratory / Virginia Tech, USA
Jens Eschenbächer, Priv. Fachhochschule für Wirtschaft & Technik (FHWT), Germany
Viviana Gaballo, University of Macerata, Italy
Carlos Granell, European Commission - Joint research Centre, Italy
Lei Guo, Ohio State University - Columbus, USA
Samer Hassan, Universidad Complutense de Madrid, Spain
Atsuo Hazeyama, Tokyo Gakugei University, Japan
Celso M. Hirata, Instituto Tecnologico de Aeronautica, Brazil
Chi Hung, Tsinghua University, China

Sergio Ilarri, University of Zaragoza, Spain
Kashif Javed, Abo Akademi University, Finland
Guillermo Jimenez, Tecnologico de Monterrey, Mexico
Xin Jin, University of Texas at San Antonio, USA
Henric Johnson, Blekinge Institute of Technology, Sweden
Rajkumar Kannan, Bishop Heber College (Autonomous) - Tiruchirappalli, India
Hassan Karimi, University of Pittsburgh, USA
Holger Kett, Fraunhofer IAO - Stuttgart, Germany
Byoung Uk Kim, Ridgetop Group Inc. - Tucson, USA
Dimitris Kotzinos, TEI of Serres, Greece
George Kousiouris, National Technical University of Athens, Greece
Roland Kübert, High Performance Computing Center Stuttgart / Universität Stuttgart, Germany
Julie Le Cardinal, Laboratoire Genie Industriel, Ecole Centrale Paris, France
Xiaoqing (Frank) Liu, Missouri University of Science and Technology, USA
Ying Liu, KAIST, Republic of Korea
Raymond A Liuzzi, Raymond Technologies - Whitesboro, USA
Wendy Lucas, Bentley University - Waltham, USA
M. Antonia Martínez-Carreras, Universidad de Murcia, Spain
Robert Matzinger, University of Applied Science Burgenland, Austria
Christoph Meinel, Hasso-Plattner-Institut GmbH - Potsdam, Germany
Klaus Moessner, University of Surrey, UK
Martin Molhanec, Czech Technical University in Prague, Czech Republic
Surya Nepal, CSIRO, Australia
Eugénio Oliveira, University of Porto, Portugal
Mourad Chabane Oussalah, LINA - CNRS Faculté des Sciences de Nantes, France
Philippe Palanque, IRIT, France
Willy Picard, Poznan University of Economics, Poland
Agostino Poggi, University of Parma, Italy
Yann Pollet, Conservatoire National des Arts et Métiers, France
Elaheh Pourabbas, National Research Council of Italy, Rome
Claudia Raibulet, University of Milano-Bicocca, Italy
Stuart Rubin, SSC-PAC, USA
Francesc Saiqi Rubió, Open University of Catalonia (UOC), Spain
Abdel-Badeeh M. Salem, Ain Shams University - Cairo, Egypt
Pascal Salembier, University of Technology of Troyes, France
Corrado Santoro, University of Catania, Italy
Marcello Sarini, University of Milano-Bicocca, Italy
Silvia Schiaffino, ISISTAN (CONICET - UNCPBA), Argentina
Thomas Schuster, FZI Forschungszentrum Informatik, Germany
Wolfgang Seiringer, Vienna University of Technology, Austria
Kewei Sha, Oklahoma City University, USA
Haifeng Shen, Flinders University, Australia
Dongwan Shin, New Mexico Tech, USA
Mohamed Shehab, University of North Carolina at Charlotte, USA
Mei-Ling Shyu, University of Miami - Coral Gables, USA
Alexander Smirnov, St. Petersburg Institute for Informatics & Automation of the Russian Academy of Sciences (SPIIRAS), Russia
Howard Spoelstra, Open University in the Netherlands, Netherlands

Julian Szymanski, Gdansk University of Technology, Poland
Yehia Taher, Université de Versailles Saint-Quentin-en-Yvelines – Versaille, France
Christophe Thovex, CNRS 6241 / University of Nantes, France
Katya Toneva, Middlesex University - London, UK
Thrasyvoulos Tsiatsos, Aristotle University of Thessaloniki, Greece
Manolis Tzagarakis, University of Patras, Greece
Michael Vassilakopoulos, University of Central Greece, Greece
Zhengping Wu, University of Bridgeport, USA
Mudasser F. Wyne, National University, USA
Kevin X. Yang, Wolfson College / University of Cambridge, UK
Yi Yang, The Catholic University of America - Washington DC, USA
Ting Yu, University of Sydney, Australia
Weihai Yu, University of Tromsø, Norway
Xiaokun Zhang, Athabasca University, Canada

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Collaborative Management Design Model for Career Planning

Weibao Chang
Academy of Mathematics and systems science
Chinese Academy of Sciences
Beijing, P.R. China
cwb@amss.ac.cn

Liqiang Wang
Department of Enterprise Platform Services
Microsoft (China) Co., Ltd
Beijing, P.R. China
bewang@microsoft.com

***Abstract* — This paper proposes the collaborative management design for career planning, which is an advanced management function in the modern human resource management system. Its implementation is quite complicated. Especially for a large enterprise, its huge size, complicated positions and diversified staffs make HR with career planning even more difficult and face two main challenges: one is coordination with other human resource modules (e.g., recruitment, training, performance appraisal, etc.), another is collaboration in the career planning process. It is needed to address issues such as: the problem of evolution of appraisal standards, the problem of satisfying corporation and employees in appraisal, the problem to involve every employee in career path planning. This article proposes a "data-driven, function-coordination" concept, and provides a collaborative design framework, based on its applications in Sinopec.**

Keywords – HR; position; career planning; Collaboration; CBR

I. INTRODUCTION

The intense competition among companies ultimately is the competition for talents, therefore a first-rate enterprise must have first-class human resource management (HR) [6], and furthermore it must have the first-class HR management software system. HR software includes many functional modules: recruitment management, staff training management, salary management, performance management, human resource planning, career planning, employee self-service, and so on. Issue that career planning [1] must resolve is how to make the promoted plan of next several positions and ultimate objective based the current position, certainly it does not exclude the possibility of being transferred to another position at the same level. The career planning process generally includes: self-assessment, present status examination, career path planning, and procedure management.

Career path planning covers the whole lifecycle of HR management and has interoperability requirements with almost all other modules. In a traditional career path planning software system, if the appraisal standards changed, the data for the old standards becomes obsolete and cannot be used anymore. Upstream systems and downstream systems, which are based on the old standards, must be modified before working with the new standards. In this article, a mapping model for the evolution appraisal standards is introduced to address the problem of change of standards. The article also describes how to improve the interoperability through providing Web Services and messaging mechanism.

In a modern enterprise, talents are the most valuable asset. A traditional appraisal is driven by the organization and managers. The employees are not actively involved and their ideas and opinions are not well considered. As a result, the employees do not have their career paths well aligned with the organization's goals. The organization must provide a way to enable every employee to be actively involved and to align their personal goals with the organization's goals [2]. The article describes collaboration solutions in the appraisal process and self-service career path planning, which will improve the satisfaction and loyalty of the employees. The growth of employee users supports the prosperity of the corporation.

The core function of career planning is to give the different career planning according to the different person. A large company's HR is more complex. It contains numerous subsidiaries, complex positions, and a huge number of employees. In Sinopec (China Petrochemical Corporation), for example, the number of employees is measured by millions and there are almost 50,000 branches. The positions in a large enterprise are complicated and so is the personnel career path planning. On the side, it has a large historical data as its advantage. CBR (case-based reasoning) technique in artificial intelligence is used to plan career paths of employees [4] [5]. By searching same or similar cases in the existing system, the next career paths can be designed. At the same time, the self-learning mechanism is used to adopt expert's advices and to add new cases to optimize, and improve the system.

In the following, the fundamental model of career planning is introduced in Section 2. Section 3 studies the coordination issue between the HR modules. Section 4 addresses the network coordination issue. Section 5 applies CBR to plan the career paths. Section 6 summarizes with the conclusion and future work.

II. GENERAL MODEL FOR CAREER PLANNING

The career planning serves each employee of whole companies. Our goal is to give different career paths (position paths) for each employee. A large enterprise contains giant subsidiaries, complex positions and a huge number of employees. In Sinopec, for example, the number of employees is measured in millions and there are hundreds of thousands of positions, which have a series of indexes to evaluate [9].

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Suppose there are ℓ employees and m positions in the enterprise, and set n indexes for evaluation.

The employee set is:

$$E = \{ E_1, E_2, \dots, E_\ell \}. \quad (1)$$

The position set is:

$$P = \{ P_1, P_2, \dots, P_m \}. \quad (2)$$

Each position includes two important attributes at least: *parent position* and *salary level*. The *parent position* means the superior position being report to. If the *parent position* is NULL, the position is at the top-level. A position tree based on the *parent position* attributes can be drawn as below.

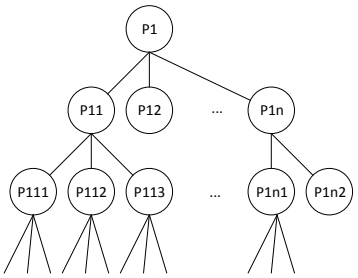

Figure 1. Example of a position tree.

The *salary level* is a measurement of the position's value. In general, if a position reports to another position, the latter has a higher *salary level*. $SL(P_i)$ is used to describe the *salary level* of the position P_i.

There are some specific indexes to evaluate a position. All the indexes make up the full index set to evaluate all the positions:

$$X = \{ X_1, X_2, \dots, X_m \}. \quad (3)$$

In the set, X_i can be the appraisal, professional interest, professional value or personality standard. One index can be applied to multiple positions.

Every index can have a serial scores and $\max(X_i)$ means the maximum score of the index X_i.

There is the essential attribute of index: *parent index*. Same as positions, an index tree based on the *parent index* attributes can be drawn as below.

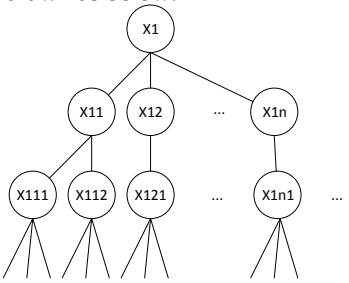

Figure 2. Example of an index tree.

In Sinopec, for example, the index tree is set up with 5 levels.

Some indexes in the index set can be used to evaluate a specific position, and the related thresholds and weights can be specified at same time.

For a specific position P_j, the index subset X^j, related threshold set T^j and weight set W^j can be defined as follows.

$$X^j = \{X_1^j, X_2^j, \dots, X_{nj}^j\} \subset \{ X_1, X_2, \dots, X_n \} \mid nj \le n; \quad (4)$$

$$T^j = \{T_1^j, T_2^j, \dots, T_{nj}^j\} \mid nj \le n; \quad (5)$$

$$W^j = \{W_1^j, W_2^j, \dots, W_{nj}^j\} \mid nj \le n. \quad (6)$$

After completing the definitions, all the employees with the positions can be evaluated. For the employee E^i at the position P_j, the evaluation result set V^{ij} can be calculated.

$$V^{ij} = \left\{V_1^{ij}, V_2^{ij}, \dots, V_{nj}^{ij}\right\} \mid i = 1,2,\dots,\ell;\ j = 1,2,\dots m; \quad (7)$$

$$nj \le n.$$

Theoretically, the employee E^i qualifies for the position P_j, if:

$$V_k^{ij} \ge T_k^j \mid k = 1,2,\dots,nj. \quad (8)$$

As a result, the entire position path chain for the employee E^i may be gotten.

$$P^i = \{P_1^i, P_2^i, \dots, P_{mi}^i\} \subset \{ P_1, P_2, \dots, P_m \} \mid mi \le m. \quad (9)$$

In general, the *salary level* of the position is monotone rise:

$$SL(P_k^i) \le SL(P_{k+1}^i) \mid k = 1,2,\dots,mi+1. \quad (10)$$

The goal of the career planning is to make the different position path chains for all the employees.

III. COLLABORATION WITH OTHER MODULES IN HR

In a modern corporation, there are many information systems to support daily operations and business growth. Career path planning system has close collaboration with other systems in recruiting, training, performance appraisal and promotion process [1]. The interoperability among these supporting systems is the key to make functional groups to coordinate smoothly together.

In order to support the collaboration of different modules in HR, the "data-driven, function-coordination" concept is proposed, which uses a generic data model for the appraisal process and a unified way to coordinate with other functions. To implement this concept, it needed to address two basic

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

problems: 1. the evolution of index standards; 2. the collaboration with other modules in HR.

A. Evolution of Standards

The lifecycle of appraisal is usually aligned with physical or fiscal year. At the beginning of the new cycle, the organization will define the standards based on the strategies and priorities. No matter if the change of standards is minor or significant, it will affect historical records and any upstream/downstream systems, which reference the changed indexes. The change can be categorized as adding, removing, updating an index or changing the calculation formulas.

The evolution mapping model is used to address the above problems. Given a sub-tree of index tree as example, it can be changed twice since the first year. The changes are illustrated in below diagram:

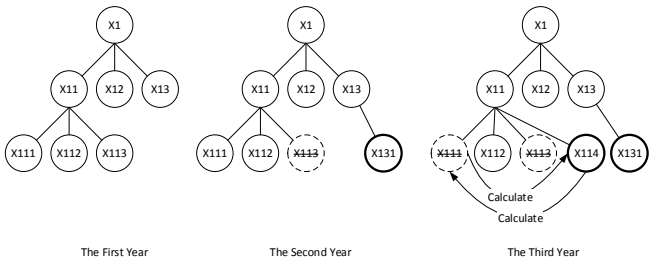

Figure 3. Illustration of the evolution mapping model.

In the first round changes, the index "X113" is removed and the index "X131" is added. In the second round changes, the definition of index "X111" is changed.

For deleting an index (e.g., "X113"), the index is marked as obsolete but will not be dropped physically. A user cannot set value for an obsolete index, but other systems can still use the value of this index as before. This will not break upstream or downstream systems. The data for obsolete indexes will be kept as historical data and can be used as fact data for business intelligence.

For adding an index (e.g., "X131"), it will change the calculation rule for the *parent index*. If the parent one was a leaf before, it may have inputting values before but will only have calculated value after the change.

For updating an index (e.g., "X111"), it will add a new index and mark the old index as obsolete. At the same time, it will allow the user to define the function from old to new index:

$$V_{X114} = f(V_{X111}). \quad (11)$$

The value of "F114" can be calculated based on this function f( ).

In order to be compatible with other systems, a reverse function can be defined as:

$$V_{X111} = g(V_{X114}). \quad (12)$$

With this function, it can calculate the value of "X111" when the value of "X114" is inputted or changed. Any system that reference "X111" can work as before and use the scores after the index changed.

Table 1 shows the example of the inputted and calculated values for an employee:

TABLE 1. EXAMPLE OF INPUTTED AND CALCULATED VALUES

Cycle	X1	X11	X111	X112	X113	X12	X13	X131	X114
1	C	C	F	F	F	F	F		C
2	C	C	F	F		F	C	F	C
3	C	C	C	F		F	C	F	F

In above table, "F" means inputted fact value and "C" means calculated value.

From the table, it can tell that values of the index "X113" are kept. When index "X111" was changed to "X114" on the third year, index "X114" has calculated values for first and second year by calling f(). On the third year, "X111" has calculated value by calling g(). This makes it possible to compare all the indexes over the three years.

To sum up, the mapping model of changed indexes makes it possible to change standards while keeping the history data and the interoperability with other systems.

For a large enterprise, the value of the history data will increase year by year. In Sinopec, it will turn into a huge data repository. A distributed database system can be built for long term data storage.

B. Collaboration with Mudules in HR

The key model for career path planning is appraisal standard of the indexes. Some values of the indexes are from other modules in HR, e.g., the degree data result from training module and attendance rate in check-in module. The appraisal module also provides information to other modules, such as open positions and position requirements for recruiting module, training requirements to training management module and projected employee growth for strategy analysis.

In order to provide data to downstream system and allow upstream systems to update information in career planning system, it provides Web Services around appraisal standards [7] [8]. The Web Services provide operations to query definitions of appraisal standards, query requirements of positions, update scores for indexes, make calculation and query appraisal results. All branches of Sinopec across all the regions can integrate with the central Web Services.

The systems for a corporation are loose coupled. In order to orchestrate them to work together, it provides a messaging mechanism. The messaging system provides ways to send message to other modules or persons. For example, when the appraisal cycle starts, it sends calendar items to all participants. The events will show up in calendar on a PC or a mobile. The receivers can access system by clicking the link in the event.

Based on the projected changes for next year, it can send out projected open positions to recruiting module. This makes it possible to start the talent recruiting work in advance.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

IV. COLLABORATION IN CAREER PLANNING PROCESS

Career planning is an employee driven process. In the context of the corporation and society, with the help of the corporation, experts, mentors and friends, the employee improves himself/herself in one or several appraisal index and moves from one position to another. The corporation establishes the index appraisal standards, which set the expectation for an employee's career development. At the same time, the corporation needs to provide a collaboration platform for self-serviced career planning [2].

A. *Appraisal with Collaboration*

A traditional appraisal process is mainly driven by a manager. The manager defines expectations based on the appraisal standards, and then communicates the expectations to the employee. The manager measures the actual performance, compares with expectations and provides feedbacks to the employee. The employee takes actions based on feedbacks and the manager will measure and compare again. This is an on-going measure and feedback process. At the end of a cycle, the manager records the results. Figure 4 shows the whole process of a traditional appraisal.

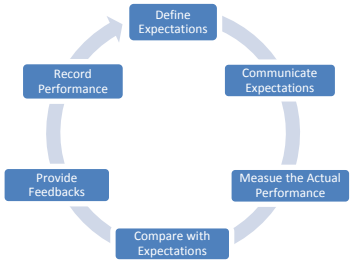

Figure 4. The appraisal process.

A limitation of the process is that employees are not actively involved. 360-degree performance appraisal was introduced to address the problems of the traditional appraisal [3].

Figure 5. The 360-degree performance appraisal.

The 360-degree performance appraisal is a process to get feedbacks from members of an employee's immediate work circle. The challenge is how to make the process efficient as it involves much collaboration across the teams. With the career planning system, all steps can be done online.

When the appraisal starts, it will send reminders to employees. An employee will see events on his/her calendar on a PC or a mobile. Following the guide in the reminders, the employee can nominate other employees to provide feedbacks. The nomination will be sent to the manager who then can choose from the list, add other employees if necessary and send out the invitations. The employees who are invited will receive reminders in their calendars. Opening the events, they can follow the guide to provide feedbacks. The feedbacks are for indexes defined by the organization.

The manager will grade an employee based on the feedbacks and other defined indexes, and then the manager can communicate the results to the employee. As the feedbacks are collected from a wide audience, the manager can grade more subjectively and the employee can get more helpful advices. The collaboration process improves the precision of the appraisal and the satisfaction of the employee.

B. *Self-Serviced Career Planning*

In additional to the appraisal, an organization can provide a self-serviced career planning system to support the growth of employees. Figure 6 shows four steps in self-serviced career planning.

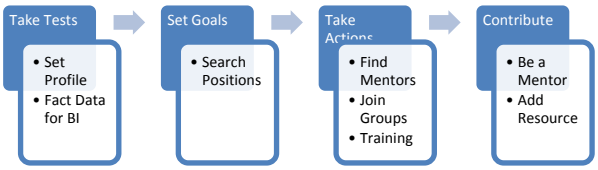

Figure 6. The Self-Serviced Career Planning Process.

Step 1: take tests. The system provides a career interest test, a career value test and a personality test. When an employee finishes those tests, he/she will have a profile in the system. The profile can be fact data for business intelligence analysis. The employee can search other employees based on his/her test results. Other employees can find theirs based on their profiles. It will build up a social system from the career perspective.

Step 2: set goals. Career planning is to set goals for career growth. The employee can search positions based on his/her test and appraisal results. The system can also recommend open positions based on an analysis of the employee and positions' requirements.

Step 3: take actions. The employee may find mentors for some fields, join in some groups and attend some training.

Step 4: contribute. As an employee grows up, he/she can provide help to other employees. He can be a mentor and add some resources. The system can also provide ranking or endorsement to mentors or experts. This will encourage volunteers to contribute to the system.

The four steps are iterative growth steps for all employees. They enable every employee to create a high-quality career path plan. With this information, the

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

organization can have the dynamic information of talents and win in competitions.

V. CAREER PATH PLANNING

Career path is a series of positions in the corporation. Section 4 describes how to get evaluation result from the collaboration process. This section will describe how to get the career path that fits an employee. In other words, how can an employee find a positive higher-level position? This section will solve the problem. Based on "Suitable is the best", the "*Qualifying Rate*" concept is proposed at first.

A. Defining Qualifying Rate

The position requirements and the evaluation are described in Section 2. For a specific position P_j, the index subset X^j, related threshold set T^j and weight set W^j can be specified. For the employee E^i at the position P_j, the evaluation result set V^{ij} can be calculated in Section 4.

Theoretically, the employee E^i qualifies for the position P_j, if $V_k^{ij} \geq T_k^j \mid k = 1,2,\dots,nj$. In general, that is impossible for the next higher-level position, because he has not worked at the position yet.

The *Qualifying Rate* Q^{ij} of the employee E^i for the position P_j is defined as follows.

$$Q^{ij} = \frac{\sum_{k=1}^{nj}\{\left(V_k^{ij} - T_k^j\right) * \frac{W_k^j}{\max\left(X_k^j\right)}\}}{nj}; \qquad (13)$$

$$i = 1,2,\dots,\ell;\ j = 1,2,\dots m.$$

Here, $\max(X_k^j)$ is the maximum value of the index X_k^j

It is obvious that:

$$-1 \leq Q^{ij} \leq 1 \mid i = 1,2,\dots,\ell;\ j = 1,2,\dots m. \qquad (14)$$

In general, positive Q^{ij} means that employee E^i qualifies for the position P_j, and negative Q^{ij} means the opposite. For example, in the following diagram, employee 1 qualifies for the position, and employee 2 does not qualify for the position.

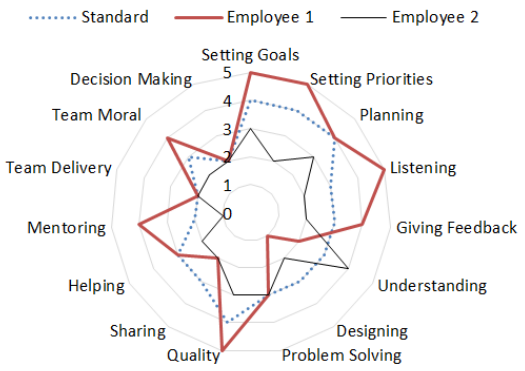

Figure 7. The capability model.

B. Finding the next higher-level position

From the above definition of *Qualifying Rate*, all the Q^{ij}s for an employee E^i to a position P_j can be calculated, which may be different from the current position. The next problem is how to select the next suitable higher-level position for the employee.

Actually, the maximum Q^{ij} is just the decision. The position P_k is the most suitable higher-level position where:

$$Q^{ik} = \max(Q^{ij}) \mid j = 1,2,\dots m, \qquad (15)$$

and the *salary level* (defined in Section 2) of P_j is higher than the one of the current position.

C. Finding the career path

The way to find the next higher-level position is proposed as above, but it is impossible to use the way to get the entire position path for the significant distance in position hierarchy. Career planning is to give the different career planning way according to the different person. This is complicated in a large enterprise, which normally has multiple levels and numerous regions. For instance, Sinopec Group has five levels of branches, three types of companies along the industry chain, millions of employees and almost 50,000 branches. Rich historical career path data may be used to find similar career path cases.

CBR technique in an artificial intelligence may be applied to career path planning.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

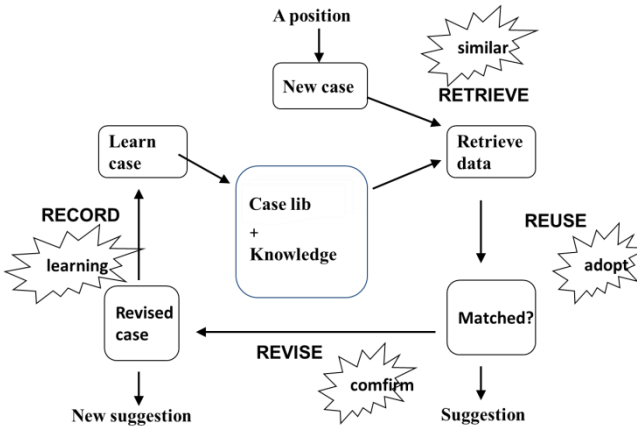

Figure 8. CBR proess diagram.

From the above diagram, the following "4-R" steps support the career path planning.

Step 1: Retrieving. For an employee with an existing position, it looks for the same cases in the case library.

Step 2: Reusing. If the same case is found, it recalls the historical data to plan the employee's position path.

Step 3: Revising. If no such case, the computer finds the most similar case, and saves it to a new case for the employee and HR manager to revise

Step 4: Returning. Add the new case to the case library, if there is not any problem.

The core technology is the matching algorithm. Based on current position, it can find the matching case, which has the minimum difference between the two persons. The distance definition of two persons E_1 and E_2 at position P_j may be given by:

$$D^{12} = \frac{\sum_{k=1}^{nj}\{abs\left(V_k^{1j} - V_k^{2j}\right) * \frac{w_k^j}{max\left(x_k^j\right)}\}}{nj}. \quad (16)$$

By constantly learning, the case library gradually becomes bigger and bigger. This model was considered effective when tested in Sinopec.

VI. Conclusion and Future Work

Career planning is a collaborative effort of the employees, their teammates, managers and the corporation. This paper described a framework of information system to support the collaborations in career planning process. It defined a model for appraisal standards, and then it elaborated how to manage the evolution of standards and interactions with other modules in HR system. After that, it proposed the appraisal process and collaborations in the process. Based on the analysis of the rich and well organized historical data, career path planning supporting information system can be a valuable assistant to the employees and the corporation.

The traditional career path selection is a manual process. The paper introduced an innovative solution to make automatic decision of the career path planning, which is based on the proposed *Qualifying Rate* to find the next-higher level positions and to get the whole career path by CBR.

As this is a general framework to provide support of collaborations in career planning, only skeleton was defined. The future work will elaborate the data model and define more details of CBR. The model will be adjusted and improved based its applications in the corporations.

Acknowledgment

This paper is supported by the National Center for Mathematics and Interdisciplinary Sciences, CAS。Thank Prof. Changshui Zhang from Tsinghua University for his instructions to this paper.

References

[1] Gary W. Carter, Kevin W. Cook, and David W. Dorsey, "Career paths: charting courses to success for organizations and their employees", Chichester, U.K., Wiley-Blackwell, 2009, pp. 69-90.

[2] Jessica Keyes, "Enterprise 2.0: Social Networking Tools to Transform Your Organization", Boca Raton, Fla. CRC Press, c2013

[3] Richard Lepsinger and Anntoinette D. Lucia, "The Art and Science of 360-degree Feedback, Second Edition", San Francisco, c2009.

[4] Qijun Xia and Ming Rao, Dynamic case-based reasonin for process operation support systems, Engineering Applications of Artifacial Intelligence, Vol. 12, Issue 3, June 1999, pp. 343-361.

[5] Agnar Aamodt and Enric Plaza, Case-based resoning: foundational issues, methodological variations, and system approaches, Artifical Intelligence Communications, Vol. 7, Number 1 / 1994, pp. 39-59.

[6] H. John Bernardin and Joyce E. A. Russell, Human Resource Management: An Experiential Approach, New York, McGraw Hill, 1998.

[7] Mohammad Mehedi Hassan and Eui-Nam Huh, Dynamic cloud collaboration platform: a market-oriented approach, New York, NY, Springer, 2013.

[8] Manlio Del Giudice, Elias G. Carayannis, and Maria Rosaria Della Peruta, Cross-Cultural Knowledge Management: Fostering Innovation and Collaboration Inside the Multicultural Enterprise, New York, NY, Springer, 2012, pp. 49-62.

[9] Weibao Chang and Changshui Zhang, Model for Evaluation and Optimal Distribution of the Service Branches of Large Business Enterprises, Proceedings of the 2012 International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government, WORLDCOMP'12, July16-19, 2012, Las Vegas Nevada, USA, pp. 273-279.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Effective Team Learning in the Cloud

Forming Teams for Motivating Productive, Creative or Learning Projects

Howard Spoelstra, Peter van Rosmalen, Peter Sloep
Centre for Learning Sciences and Technologies
Open University in the Netherlands
Heerlen, the Netherlands
howard.spoelstra@ou.nl, peter.vanrosmalen@ou.nl,
peter.sloep@ou.nl

Evert van de Vrie
School of Computer Science
Open University in the Netherlands
Heerlen, the Netherlands
evert.vandevrie@ou.nl

***Abstract*—Learning in the cloud can be a lonely activity for self-directing and self-organizing learners. Lack of sustained learner motivation can lead to less effective, less bond-creating learning experiences. By providing collaborative project-based learning opportunities these shortcomings can be overcome. A service design is introduced for the onset of collaborative project-based learning and team formation in the cloud, based on learning materials in the cloud, project definitions and characteristics, and learner 'knowledge', 'personality' and 'preferences'. The article specifies how the data required by the design can be gathered. Team formations rules are deduced from existing team formation research. They steer the team formation process towards facilitating learning, creative problem solving or increased productivity outcomes. The rules are implemented in three team formation equations. Deployment of the equations on a set of test data demonstrates the effectiveness of the team formation service.**

Keywords-Cloud learning; project-based learning; project team formation; self-directed learning; team formation rules

I. INTRODUCTION

Nowadays, everyone with a connection to the Internet can learn from the cloud of knowledge it provides. The individual learners freely using the resources available are considered to be self-directing and self-organising. But individual learners can find it difficult to remain motivated [1]. The introduction of collaborative project-based learning can help overcome the drawbacks of individual learning. Project-based learning is considered to be motivating, bond-creating and effective [2] [3]. However, how are such teams of cloud learners formed in the absence of human agents such as teachers? To address this need, we present a design for a service that can support these learners to set up project-based activities and form teams.

Prior research indicates that for project-based learning to deliver optimal results, experts should form the teams [19] [20] [21], using their knowledge about the learners. As these experts are not readily available in the cloud, our design aims to mimic expert behaviour, but also takes into account learner self-direction and self-organisation. The concept of *Fit* [16] [22] is used to indicate whether a team of learners (in its unique combination of knowledge, personality and preferences) is suitable to perform a specific project.

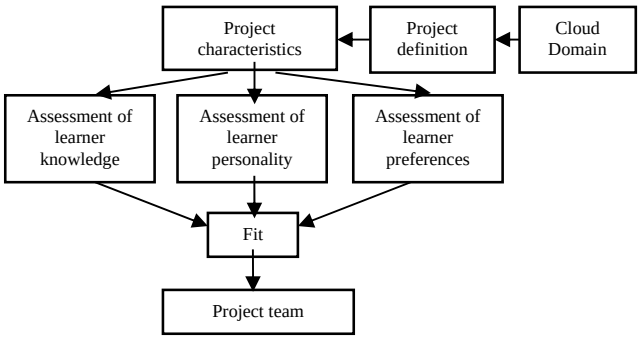

Figure 1. The team formation model

Following the team formation model depicted in Fig. 1 [4], our service is designed to work based on a knowledge representation of learning materials in some knowledge domain in the cloud. Such a representation can be created with language technologies such as Latent Sematic Analysis (LSA) [18]. In [23] the authors demonstrated that an LSA-based software tool is capable of fitting job descriptions to people's knowledge and learning materials.

The project-based learning and team formation process can be started by a learner or other project initiator by submitting a *project definition* (which details the project' aims), and the *project characteristics* (such as the preferred team size, duration, etc.) to the service.

The following is an example from the perspective of a prospective team member of how we envision the service can work: "*May 2013: Emma recently started her new job at the microelectronics department. For the first two months her main task was to strengthen her knowledge in this domain. She decided to follow a highly recommended MOOC course. On top of the regular lectures and other materials, the MOOC also offered a project wall with the possibility to apply for a project assignment. The assignments varied from assisting peer students to participating in small and large projects proposed by peer students, companies and research institutes. The larger projects followed an automated, open procedure to select the best applicants. Emma selected a project on bio-chip design. The project was an interdisciplinary project to be performed*

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

by at least 4 persons. She could apply by sending in a brief summary of around 100 words on her knowledge and skills with regard to a pre-defined list of topics, filling out her preferences (on language, availability, etc.) and taking a personality test. Emma decided to give it a try and sent in the required information and did the personality test. A few days later she received an invitation to participate in the project and contacted her fellow project members to make arrangements."

The team formation model defines that in order to assess whether prospective team members are suitable for a project, data is needed in three categories: i) *knowledge*, ii) *personality*, and iii) *preferences*. However, as the main focus of this paper is on how teams can be formed, we assume that the required data has already been gathered. Therefore, in Section II, we only briefly introduce how the assessments are designed to work. For the remainder of this article, the data are then assumed to be available.

As project-based activities can have different purposes, in Section III we define three general team formation rules, which enable the team formation service to form teams that are aimed at *increased productivity*, *creative problems solving* or *facilitating learning* as outcomes. The rules are based on prior research findings about team formation. Here, we translated them into three team formation equations. In Section IV, we report on the results of the application of the equations to a set of test data. Finally, in Section V, the results are discussed, conclusions are drawn and future work is indicated.

II. THE KNOWLEDGE, PERSONALITY AND PREFERENCES ASSESSMENTS

As introduced above, the team formation for project-based learning starts with the definition of a project related to a knowledge domain in the cloud. The service then deduces to how many and which specific topics in the domain the proposed project primarily refers, by using the project description as a query into the domain knowledge representation.

Next, the fit of the prospective team members is assessed with respect to their knowledge, personality and preferences, related to other members and the knowledge requirements of the project. We will briefly describe these assessments:

The *knowledge assessment* determines how much knowledge, if any, learners have available on the topics the project refers to. For this assessment the learners submit knowledge evidences, which are used as queries into the domain knowledge representation.

The *personality assessment* uses data on learner personality, which are gathered through a personality test. The resulting learner personality profile is made up from a person's *conscientiousness* [5] [6]. This personality construct consists of the personality characteristics carefulness, thoroughness, sense of responsibility, level of organization, preparedness, inclination to work hard, orientation on achievement, and perseverance. 'Conscientiousness' is chosen specifically because it predicts a person's future performance in a team. The learner' conscientiousness score is established by using the Big Five personality test [7]. We consider the inclusion of personality as a factor in the team formation process to be of particular importance as team formation literature shows a strong tendency to focus merely on knowledge as a general indicator for successful participation in a project, while other studies indicate that other factors better predict success [8].

The *preferences assessment* is performed on learner data entered on such variables as availability for the duration of the project, time zone, possible collaboration languages and preferred tools into a learner preferences profile. The assessment then determines the overlap between the project characteristics and the learners' project work related preferences. When preferences do not overlap, they constitute 'condiciones sine qua non' for inclusion in a team. (E.g., when one learner indicates to be available on Mondays, while another learner indicates to never be available on Mondays, their calendars are mutually exclusive and thus these two learners will never be matched in a team).

It is, however, important to notice that the data gathered from learners is *not of a static nature*, but can be refreshed every time a learner re-enters knowledge evidence for a project, retakes the personality test, or updates preferences.

As a first step in the team formation process, the preferences assessment can limit the number of learners to be considered for inclusion in a team. Learners might, however, negotiate preferences and re-enter them in their preferences profile. After this step is completed, the team formation process continues with the knowledge and personality data.

III. DEFINITION OF THE TEAM FORMATION RULES AND EQUATIONS FOR TARGETING SPECIFIC OUTCOMES

Assuming the data from the knowledge and personality assessments and the project characteristic "preferred team size" are available, the team formation service combines the two separate sets of data by following particular team formation rules. In the design of these rules we take into account prior research findings on team formation. We sort the findings into three possible teamwork outcomes (productive problems solving, coming up with creative solutions, and facilitating learning) and present the general rules we deduced for forming teams suited to achieve these outcomes. Based on these general rules, we present three team formation equations.

A. Teams fit for increased productivity

We considered the follow research outcomes for the creation of teams aimed at increased productivity:

a) Differences in conscientiousness scores impede task negotiations [9];

b) Members of productive teams should be capable and conscientious and must have domain knowledge [10].

The general team formation rule we deduce from these findings is: *Productivity is fostered when team members show high levels of conscientiousness and have supplementary high knowledge on the project topics*.

This rule is translated into the team formation equation for productive teams (1). When applied, it determines which teams have the highest average knowledge scores and the highest average conscientiousness scores.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

$$FitP_i = W_K * \frac{Avg\_K_i}{Max\_K} + W_C * \frac{Avg\_C_i}{Max\_C} \quad (1)$$

Explanation of the terms used in (1): $FitP_i$: Fitness of a team *i* for productive outcomes; Avg_K_i: Average knowledge of all members of a team *i* on all topics; *Max_K*: Maximum possible score on knowledge on a topic; Avg_C_i: Average conscientiousness score of all members of a team *i*; *Max_K*: Maximum possible score on conscientiousness; W_K, W_C: Weights.

B. Teams fit for creative problem solving

For the formation of creative problem solving teams, we considered the follow research outcomes:

a) Too much complementary fit in knowledge can lead to a loss of creativity and group thinking [11];

b) People with high conscientiousness scores tend to be less creative [12] [13];

c) Groups with members that possess different knowledge backgrounds will be more innovative because they contribute from different perspectives [14];

d) Successful research teams are heterogeneous [15].

The general team formation rule we deduce from these findings is: *Team creativity is fostered when team members have low scores on conscientiousness, while showing highly differentiated scores on knowledge of the project topics*.

This rule is translated into the team formation equation for creative teams (2). It reaches a maximum when team members have a maximum difference in knowledge between their best score and their second-best score over their own topic scores, and when there is a maximum difference in knowledge between the best score and the second best score on a topic. It minimises the average conscientiousness score in the team.

$$FitC_i = W_K * \frac{\sum_j DifK_j}{TeamSize * Max\_K} + W_E * \frac{\sum_t DifK_t}{NumTop * Max\_K} + W_C * \frac{Max\_C - Avg\_C_i}{Max\_C} \quad (3)$$

Explanation of the terms used in (3): $FitC_i$: Fitness of a team *i* for creative outcomes; $DifK_j$: difference between the highest and next highest topic knowledge score of member *j*; *TeamSize*: Number of team members; $DifK_t$: Difference between the highest and next highest knowledge score on topic *t*; *NumTop*: Number of topics; W_K, W_E, W_C: Weights

C. Teams fit for facilitating learning

For the formation of teams in which learning is facilitated, we considered that:

a) Learning is fostered when team members provide a complementary fit in knowledge backgrounds and show a supplementary fit in personalities [16];

b) Mutual teaching and learning are among the most important activities in defining and solving problems [14];

c) There is a maximum distance in knowledge (the zone of proximal development, or '*zpd*') that can be bridged when learning with more capable peers [17].

From these findings we deduce as general team formation rule: *Learning in a team is fostered when knowledge on the project topics is distributed over the members (allowing each member to learn and teach), that differences in the levels of topic knowledge should not be too high, and that the members' conscientiousness scores should all be high.*

This rule is translated into the team formation equation for learning teams (3). It reaches a maximum for teams whose members can teach and learn to and from each other inside each topic, while having a high score on Conscientiousness. It optimizes the match between possible teachers and learners in the team by using Vygotsky's "zone of proximal development" (*zpd*) to calculate teaching and learning effectiveness.

$$FitL_i = W_K * \frac{\sum_t \sum_l \sum_j \left| score_{t,j} - score_{t,l} \right|}{d_{jt} \cdot zpd \cdot n \cdot k} + W_C * \frac{Avg\_C_i}{Max\_C} \quad (2)$$

Explanation of the terms used in (2): $FitL_i$: Fitness of a team *i* for learning outcomes; $score_{t,j}$: Score on topic *t* of member *j* ; $score_{t,l}$: Score on topic *t* of member *l*; d_{jt}: Number of team members with a different score on topic *t* for student *j*; *zpd*: Zone of proximal development; *n*) Number of team members, *k*) number of topics; W_K, W_C: Weights.

These equations were deployed on a set of test data. The results of this experiment are presented in Section IV.

IV. RESULTS OF THE APPLICATION OF THE TEAM FORMATION EQUATIONS ON A SET OF TEST DATA

For the experiment, due to space limitations in this article, we used only a small set of test data (See Table 1). In the data set, the possible score for a learner on a topic (Topic 1 through Topic 3) ranges from 1 to 10. The knowledge scores are of the type ratio. The conscientiousness scores (Cons) range from 1 to 5. Following common practise, the conscientiousness scores are treated as type interval, even though they do relate back to the Likert scales with which the underlying personality characteristics were measured.

For each possible team, the team fitness values FitP, FitC and FitL are represented with a value between "0" and "1", with "1" indicating the highest possible fitness for that outcome. This allows comparing teams with respect to fitness over their different target outcomes. If a learner or other project initiator wishes to do so, weights can be used to prioritise the importance of e.g., knowledge over conscientiousness in the team formation process. However, in the equations below all weights sum up to 1, with weights set to 1 / the number of weights used in the equation. Other weight distributions are currently not considered. For this experiment, the team size was set to 4 learners per team, and the number of topics in the project was set to 3.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

TABLE I. THE TEST DATA SET.

Learner	Topic 1	Topic 2	Topic 3	Cons
L01	9	3	2	3
L02	4	8	3	2
L03	3	2	9	3
L04	6	10	8	4
L05	6	7	5	2
L06	9	8	6	4

When the equations above are applied to the test data set, all 15 unique combinations of 4 learners are calculated for their fit values. The number of unique combinations is calculated with [*n*! / ((*n* – *team_size*)! * *team_size*!)], where *n* is the total number of learners in the data set and *team_size* is the desired number of learners in a team.

The output in Table II lists all 15 possible teams and their scores on FitP, FitC and FitL. The scores in the three columns FitP, FitC and FitL are sorted from high to low. The results are truncated to three decimals.

TABLE II. TEAM FORMATION SUGGESTIONS FOR 15 TEAMS OF 4 LEARNERS, SORTED BY FITP, FITC OR FITL.

Team members	FitP	Team members	FitC	Team members	FitL
L02,L03, L04,L05	0.606	L01,L02, L03,L05	0.358	L02,L04, L05,L06	0.488
L02,L03, L04,L06	0.603	L01,L03, L05,L06	0.358	L02,L03, L04,L05	0.475
L01,L02, L03,L04	0.597	L01,L02, L03,L06	0.342	L03,L04, L05,L06	0.466
L02,L04, L05,L06	0.587	L01,L03, L04,L05	0.325	L01,L02, L04,L05	0.437
L01,L02, L04,L05	0.581	L02,L03, L05,L06	0.325	L02,L03, L04,L06	0.436
L01,L02, L04,L06	0.578	L01,L02, L03,L04	0.308	L01,L03, L04,L05	0.429
L03,L04, L05,L06	0.563	L01,L02, L04,L05	0.300	L01,L04, L05,L06	0.427
L01,L03, L04,L05	0.556	L01,L03, L04,L06	0.300	L01,L02, L03,L04	0.402
L01,L03, L04,L06	0.553	L01,L02, L05,L06	0.300	L01,L02, L04,L06	0.383
L01,L04, L05,L06	0.538	L03,L04, L05,L06	0.283	L01,L03, L04,L06	0.362
L02,L03, L05,L06	0.503	L01,L04, L05,L06	0.283	L02,L03, L05,L06	0.337
L01,L02, L03,L05	0.497	L01,L02, L04,L06	0.275	L01,L02, L03,L05	0.324
L01,L02, L03,L06	0.494	L02,L03, L04,L06	0.275	L01,L02, L05,L06	0.299
L01,L02, L05,L06	0.478	L02,L04, L05,L06	0.258	L01,L02, L03,L06	0.279
L01,L03, L05,L06	0.453	L02,L03, L04,L05	0.242	L01,L03, L05,L06	0.274

The highest scoring teams for FitP, FitC and FitL show fitness scores of 0.606, 0.358 and 0.488 respectively. This indicates that a team of 4 (consisting of learners L02, L04, L05, and L06), created from the set of learners best fits the outcome *increased productivity*, but still only with a value of 0.606. The best possible creative team from the set of learners would only receive a FitC value of 0.358, indicating a low probability of successfully achieving the outcome *creative problem solving* for that team. The best possible combination of learners for *facilitating learning* outcomes (a team with learners L02, L04, L05, and L06) scores a FitL of 0.488, which indicates the members can only enjoy approximately half of the maximum learner and teaching effectiveness possible.

When the results are sorted on FitP, the highest scoring team on FitC is found on position 12. The highest scoring team on FitL is found on position 4. When sorted on FitC, the results show the highest scoring team on FitP is found on position 15, while the highest scoring team on FitL is found on position 14. Sorting on FitL reveals that the highest scoring team on FitC is to be found on position 12, while the highest scoring team on FitP is to be found on position 2.

The calculations' results show how the three equations, through their different handling of learner knowledge and conscientiousness, produce teams of different compositions. The results reveal the best team for a particular outcome, but also how well a particular team fits to any of the outcomes.

V. CONCLUSIONS AND FUTURE WORK

Our take on learning in the cloud is that cloud-based learners are not necessarily provided, nor can easily provide themselves, with effective, bond-creating and motivating learning settings. We argued that providing these learners with the possibility to start project-based activities affords motivating collaboration opportunities [2] [3]. We therefore suggested a design for setting up project-based learning and team formation services in the domain the learners are interested in, based on our team formation model [4]. The design puts learners in control over the process of defining and staffing projects, thus honouring these learner's self-directing and self-organising behaviour, while at the same time being firmly rooted in team formation theory. The design uses the data categories 'knowledge', 'personality', and 'preferences' from the team formation model and describes the ways in which the data can be gathered and processed to arrive at team formations suggestions. A benefit of the design is that it is also based on personality characteristics, which is rarely the case in existing tools, but which – according to literature [8] – are highly relevant in the team formation process.

In order to determine how learners should be teamed up based on knowledge and personality, we analysed existing research on team formation principles and team outcome criteria. From the team formation principles and outcomes we deduced three general team formation rules for forming productive, creative, or learning teams. These rules were formalised in team formation equations. The application of the equations to a set of test data demonstrated their ability to form teams and to suggest different teams based on the desired team work outcomes. It also showed the ability of the equations to determine for which of the three outcomes a team would be most suited.

We acknowledge that knowledge might also be contained in other forms of evidence currently not taken into account. There might also be personality aspects besides the ones underlying the personality construct 'Conscientiousness'

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

that are important predictors of a learner' success in project work or that facilitate learning and working collaboratively. The current equation for the formation of creative teams favours low conscientiousness scores for all learners, based on [12] [13]. We do, however, plan to compare the current choice with other distribution methods.

Our work for the immediate future focusses on a large scale experiment with the team formation service, using real data on knowledge from learner self-assessments and real data from learners on personality and preferences.

ACKNOWLEDGMENT

The authors wish to thank their colleague Matija Obreza for his highly valued work on programming the team formation equations.

REFERENCES

[1] K. J. Kim, "Motivational challenges of adult learners in self-directed e-learning," Journal of Interactive Learning Research, 20(3), 2009, pp. 317–335.

[2] D. W. Johnson, R. Johnson, M. B. Stanne, and A. Garibaldi, "Impact of group processing on achievement in cooperative groups," The Journal of Social Psychology, 130(4), 1990, pp. 507–516.

[3] J. A. Marin-Garcia and J. Lloret, "Improving teamwork with university engineering students: the effect of an assessment method to prevent shrinking," WSEAS Transactions on Advances in Engineering Education, 5(1), 2008, pp. 1–11.

[4] H. Spoelstra, P. van Rosmalen, and P.B. Sloep, "Project team formation support for self-directed learners in social learning networks," Proc. IADIS International Conference on Web Based Communities and Social Media (ICWBC & SM 2012), July 2012, pp. 89-96.

[5] L. R. Goldberg, "An alternative description of personality: The big-five factor structure," Journal of Personality and Social Psychology, 59(6), 1990, pp. 1216–1229.

[6] J. Jackson, D. Wood, T. Bogg, K. Walton, P. Harms, and B. Roberts, "What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC)," Journal of Research in Personality, 44(4), 2010, pp. 501–511.

[7] M. R. Barrick and M. K. Mount, "The big five personality dimensions and job performance: A meta-analysis," Personnel Psychology, 1991. 44, pp. 1–26.

[8] B. W. Roberts, N. Kuncel, R. Shiner, A. Caspi, and L. Goldberg, "The power of personality: A comparative analysis of the predictive validity of personality traits, SES, and IQ," Perspectives in Psychological Science, 2, 2007, pp. 313–345.

[9] J. M. P. Gevers and M. Peeters, "A pleasure working together? The effects of dissimilarity in team member conscientiousness on team temporal processes and individual satisfaction," Journal of Organizational Behavior, 30, 2009, pp. 379–400.

[10] S. G. Isaksen and K. J. Lauer, "The Climate for Creativity and Change in Teams," Creativity and Innovation Management, 11, 2002, pp. 74–86.

[11] M. West, "Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups," Applied Psychology: An International Review, 51, 2002, pp. 355–424.

[12] J. M. George and J. Zhou, "When openness to experience and conscientiousness are related to creative behavior: An interactional approach," Journal of Applied Psychology, 86(3), 2001, pp. 513–524.

[13] U. Wolfradt and J. E. Pretz, "Individual differences in creativity: personality, story writing, and hobbies," European Journal of Personality, 15, 2001, pp. 297–310.

[14] P. Paulus, "Groups, teams, and creativity: The creative potential of idea-generating groups," Applied Psychology, 49, 2000, pp. 237–262.

[15] K. Dunbar, "How scientists think: On-line creativity and conceptual change in science," In Creative thought: An investigation of conceptual structures and processes, T. B. Ward, S. M. Smith, and J. Vaid, Eds. Washington, DC: American Psychological Association, 1997, pp. 461-493.

[16] J. D. Werbel and D. J. Johnson, "The use of person-group fit for employment selection: A missing link in person-environment fit," Human Resource Management, 40, 2001, pp. 227–240.

[17] L. S. Vygotsky, "Mind in society: The development of higher psychological processes," M. Cole [et al.], Eds. Cambridge, Harvard University Press, 1978.

[18] T. K. Landauer, P. W. Foltz, and D. Laham, "Introduction to Latent Semantic Analysis," Discourse Processes, 25, 1998, pp. 259–284.

[19] B. Oakley, R. M. Felder, R. Brent, and I. Elhajj, "Turning student groups into effective teams," Journal of Student Centered Learning, 2(1), 2004, pp. 10–34.

[20] S. B. Fiechtner and E. A. Davis, "Why some groups fail: A survey of students' experiences with learning groups," The Organizational Behavior Teaching Review, 9(4), 1985, pp. 75–88.

[21] A. Obaya, "Getting cooperative learning," Science Education International, 10(2), 1999, pp. 25–27.

[22] A. L. Kristof, "Person-organization fit: An integrative review of its conceptualizations, measurements, and implications," Personnel Psychology, 49, 1996, pp. 1–49.

[23] D. Laham, W. Bennett Jr., and T. K. Landauer, "An LSA-based software tool for matching jobs, people, and instruction," Interactive Learning Environments, 8(3), 2000, pp. 171–185.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Powwow: A tool for collaborative software jam sessions

Paul Muntean, Damir Ismailović, Sebastian Paetzold and Bernd Bruegge
Department of Computer Science
Technische Universitaet Muenchen
Muenchen, Germany
{ paul.muntean | damir.ismailovic | sebastian.paetzold | bernd.bruegge }@cs.tum.edu

***Abstract*—The increasing time, complexity and cost of today video game development projects demand for new software tools capable to support the development of fast runnable video game prototypes. This paper presents our conceptual solution for building runnable video game prototypes. For editing video game levels we added editing capabilities in a mobile software tool. We present a mobile software tool which supports editing video game levels. Furthermore, the tool supports testing of previously designed levels with the help of design recognition mechanisms that facilitate loading and simulating of games. The tool supports distributed, collaborative design sessions as it is based on a client server software architecture. Game designers collaborate by sharing their designs on a file server from where the files can be pulled by other users for further editing. This tool is used for software jam sessions in order to support local collaborative game level design. In addition to the tool we present a tool evaluation in this paper. We introduce software jam sessions in order to be able to support local collaborative game level design. In the last part of the paper, we evaluated our tool. The goal is to find out if the working efficiency increases when developing video game levels from scratch with our tool.**

Keywords-Collaborative tool; map based video game; fast prototyping; software jam sessions

I. INTRODUCTION

The usage of the paper and pencil to capture requirements in the video game industry is causing delays and engineering overhead during the development process [1]. Thus, shifting deadlines and increasing costs negatively influence the software development process.

When conceiving new game ideas, game designers should not stick to traditional methods of capturing requirements and then develop costly prototypes. We foresee that game designers should have an easy to use tool that they can carry easily around and that helps to test ideas in a matter of minutes. Such kind of tool can link the requirements elicitation phase to the implementation of fast runnable video games. Users with no programming experience should be able to use the tool right away.

When developing mobile applications and addressing user experience there is a great need for creating many user interfaces [11]. Every video frame of a game can be abstracted to a single user interface. We think that a software tool should have game editing features and an integrated simulator that help to build fast runnable games at a fair cost of time and effort.

Nowadays, it does not exist a specialized process model for mobile applications but it can be observed that large development projects have moved away from a process-intensive approach toward a more agile approach, with the Scrum approach and other agile techniques, e.g., test driven development, finding widespread acceptance [11].

With the emergence of new mobile devices and the begining of the post-PC era [12] new posibilities arrise that can help to speed up the development of runnable video game prototypes. During a project requirements meeting with the customer a software tool should facilitate not only requirements elicitation support but also the development of a runnable game prototypes. With such a tool we want to close the gap between requirements elicitation and the implementation by being able to develop fast runnable game prototypes that help to get fast feedback from the customer. In this regard, our solution is thought to be focused on a specialized tool deployed on a mobile device accompanied with the corresponding development technique.

We have identified three main aspects regarding our conceptual solution:

Editing: The user should be able to place different objects and elements on a grid map and immediate feedback should be provided by the user interface.

We think that a grid is needed to help the user place the game elements on predefined places on the grid. This represents in our opinion the framework for other game genres because a plane is common to: FPS, Soccer, car races games, etc.

Collaborative design: Firstly it should be possible to work collaboratively in near real time and in an asynchron manner where designers collaborate and share ideas over a file server. We call this approach distributed collaborative design.

Secondly it should be possible for designers to work collaboratively on the same prototype in real time. The users should get instant feedback from others. By tailoring tasks and profiting from shared team knowledge. We think that the process of collaborative video game editing can be performed more efficiently. We call this local collaborative design.

Thirdly a development technique should be defined that contains a set of rules and guidelines which do not constrain the participants but rather help them to profit from the setting type in which they are working. This technique should help to achieve real time collaborative work. The second and third

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

aspects remain currently work in progress and we will present only the achievements that our research has produced until the moment of writing this paper.

Design recognition: The collaboratively designed levels can be interpreted in different ways. We want to be able to load them into the integrated game simulator and recognize all previously designed game elements consistently. The goal is to test the game playability.

The paper is structured as follows. The section 1 contains the problems related to development of video games on mobile devices and the description of our solution used to achieve fast collaborative video game development. Related work is presented in section 2. In section 3 we analyze the development method and tools used in the Battlefield Wars case study and in section 4 we present our developed tool. Section 5 contains work in progress about software jam sessions. In section 6 we evaluate the Powwow tool. Section 7 contains our conclusions and future work.

II. RELATED WORK

At the moment of writing this paper there were only a few apps in the App Store that support fast video game prototype development.

The *Codea* app, [2] offers the possibility to modify code of the already deployed video game app. This approach provides deep level control on how the app behaves during user interaction with it. The *PGC* app, [3] is based on predefined prototypes of video games. Haladjian et al. [5] present a quick prototyping tool that is based on code generation that can be used by users with no programming experience to develop physics based game levels. The *Battle Map 2*, [4] app is designed for building map based games levels. This approach is interesting since it wants to be a replacement for pencil and pen when conceiving new game ideas.

These tools lack collaborative work capabilities and demand programming experience. Furthermore, they restrict the user by providing a fixed number of game templates.

Game designers that have no programming experience should have the possibility to use a tool right away. Limiting game designers with predefined game templates constrains game creativity in our opinion. In order to develop complex games, collaboration mechanisms should be incorporated also in an app.

The overall video game development is not suited for typical software life cycle methods, such as the waterfall model [6]. The stages of development are done in a serial manner linking the project phases rigidly together maintaining a high project risk during the whole project. Thus, requirements updates are difficult to be performed. To close the gap between user model and design model [9] specialized tools that support informal communication are needed [7]. We believe that a video game development technique supported by tools can speed up the development process of video games.

The design at run time concept described in [8] is used in the context of reconfigurable ubiquitous software systems. The design at run time concept can be extended for developing video games using an iterative developing technique.

In our approach we have assumed that the designers have no programming experience so we choose not to expose code-editing features as in the previously mentioned example.

III. CASE STUDY: BATTLEFIELD WARS PROJECT

The *Battlefield Wars* project had the goal of producing a framework that allows light interaction between users and map-based games and supporting development of multiuser map-based video games.

The goal of our *Battlefield Wars* case study was to find out how real video game projects are developed by an experienced team of video game developers from the point of view of software tools used and development methods. The first research question was, *RQ1: Do developers use specialized tools for developing map based video games? If yes which ones?* The second research question is the following, *RQ2: Which development technique or process model do developers use? Is this adapted to the special needs of video game projects?*

We have observed the team of developers during their four weeks development work. We were for two days per week with the developers and wrote down every detail regarding tools and development methods used. Firstly we have focused on how project tasks were addressed and solved. Secondly we were concerned with inter-team communication during the project and how this has influenced the project outcome.

The results of the case study revealed that at the beginning of the project until the end the developers have worked independently and without using specialized tools.

Figure 1. Battlefield Wars game

Co-located Collaborative work was not possible since the developers did not use any specialized tools that support collaboration. The development team has used the waterfall development model being forced to stick to a sequential development style. The developers reported that it was

difficult to add new requirements to the product backlog and that creativity was "damped" because of the incapacity of team members to efficiently communicate and test their ideas. The developers agreed to use in the future agile development methods and they suggested that they need an iterative and adaptive development technique tailored to their needs.

IV. THE POWWOW TOOL

The Powwow tool prototype that we have developed is based on our conceptual solution. It is available for iOS and can be deployed on the iPad.

The editing features are available in hidden pop up menus that appear by tapping on the buttons placed on the two tab bars located in the upper and lower part of the screen. The main aspects of the tool are highlighted with numbers (1-7). Not to clutter the UI we have chosen to have two fixed tool bars in the upper and lower part of the screen.

The editing, persisting, sharing and simulation features are available by tapping on the buttons present on the two toolbars. The number (1) indicates an initial map where every tile of the map represents a second freely editable map. In Fig. 2 we have a red dot, near number (1), representing a house. Tapping the red house tile the user opens a second "endless map view" where all the previously saved map elements can be further on edited. The user has the possibility to zoom in and out when editing so that the tiles do not appear too small as in Fig 2. After pressing one of the buttons located above the number (2) the user has the possibility to select different layers of the map, to save, delete, position, undo/redo and to center the map on the screen.

After the saving process is done all previously added elements are saved in a TMX meta-format file which can be easily parsed and shared with other designers.

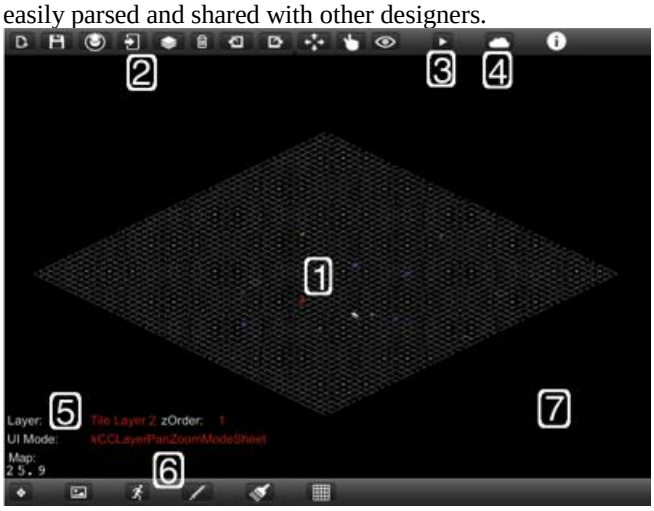

Figure 2. Powwow tool prototype

During the saving process a second file format is saved representing the same game level. This file contains all game elements, which are objects, in a serialized form. We used this second format because of performance reasons, mainly because it can be loaded and saved faster then the TMX file format, which has to be parsed. We also use the second object files for presentation reasons on the first map view indicated with number (1).

After successfully loading the game level we observe that all game elements from the TMX file are present on the level. The level can be further on edited on other iPads that run the Powwow app or with the help of the desktop program Tiled [10]. Successful design recognition consists in the TMX file parsing and the game objects instantiation.

Number (3) indicates the play button which triggers a sub view when pressed, where we can select a previously saved game level and play on it in order to find out if the game logic fits our needs and decide if further editing is necessary.

The Powwow users indicated in Fig. 3 can push all locally available game levels on the distributed server and pull all the remote available levels on their iPad. The user also has the possibility to erase every locally and remote available file. The users can be located in different locations and can collaborate by sharing these files. The files can be further on edited and pushed on the distributed file server. Every user has the possibility to play on the level that he is currently editing. The only restriction is that the level has to be previously stored on the iPad.

Number (4) represents a button with a cloud. When pressing this button, a sub view appears which asks the user if he wants to connect the Powwow tool with a distributed file server, which offers file services. After accepting this option another sub view appears offering the options of pushing, pulling, local deletion, remote deletion and disconnecting from the distributed server. At this stage of development we have added all our options for collaboration in this submenu.

The *File Server*, Fig. 3, files can be synchronized with local file folders distributed on desktop PCs. This offers the possibility to edit the prototype files on the PC by using the program called Tiled [10]. In order to be able to collaborate locally in real time without having to use a distributed file server we want to add real time capabilities to the Powwow tool. This issue will be addressed in section 5.

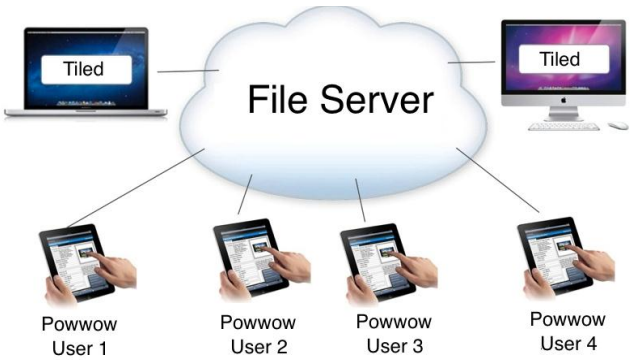

Figure 3. Distributed collaboration infrastructure

Number (5) labels indicate the current editing layer; the current selected drawing mode and the number of FPS (Frames per Second) available are also indicated. These labels can by hidden if needed.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Number (6) indicates several buttons that contain pop up menus with objects that can be added in order to edit our game level. The assets are at this stage of development restricted to only several types of objects. We also have a brush with different brush sizes that can be selected during editing.

Number (7) indicates a black area where the user cannot add tiles. The initial map can be dragged around on the black area by performing sweep gestures. This is handy when positioning and zooming the initial map and the second "endless map".

Design recognition is achieved by parsing the TMX file and then instantiating game objects. These are used afterwards to populate the game level.

V. COLLABORATIVE SOFTWARE JAM SESSIONS

This section represents work in progress and addresses the local collaborative design aspect presented in the introduction. We want to address the collaborative design not only from the tool point of view but also from the process development technique perspective. Our collaborative software jam sessions concept aims at porting the concept of musical jam sessions to quick games prototyping. The whole concept relies on the idea of jamming together, Fig. 4, in a group when developing a video game level.

The video game level artifacts represent the components that compose the game level. We want to design software instruments capable to build these artifacts. In order to design these instruments we need to identify the type of relations between instruments and artifacts.

The Fig. 5 contains the *JAMInstrument* class with which we modeled a software tool capable to produce different types of game artifacts represented by the *Artefact* class. These artifacts represent game level logic and level design assets. The produced artifacts compose our video game level.

Figure 4. Local collaborative jam session concept

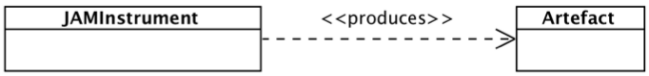

Figure 5. Jam session meta model

The Fig. 8 represents a high level view on the jam session technique. It illustrates our software jam session concept that we think it is superior in this context to the waterfall development process. Collaborative work increases communication and encourages knowledge sharing between participants that have different backgrounds. The jam session technique relates to the agile methods because it is an iterative and incremental activity that supposes that team members can organize themselves. As we perform research on this topic we think that this diagram will suffer further modifications.

The four swim lanes represent our main project stakeholders: user/player, customer, designer and tester. Looking at the Fig. 8 from left to right we have in the beginning of a jam session the requirements elicitation activity.

The first activity is the kick-off meeting where one user story is selected from the backlog and the working strategy is discussed with all stakeholders. After this phase the software jam session starts and we observe here work done in parallel and collaboratively.

In the second activity game designer designs game assets, the developer develops game logic code, the user plays/tests the current runnable prototype. Every stakeholder has the possibility to review the game level prototype at any instant in time. After the first iteration we have a wrap-up discussion where the participants analyze how the tasks were accomplished and further on feedback goes in the continuous development activity.

After several iterations we have the review product activity. At the end of each iteration we have a potential shippable product increment. Again feedback goes directly to each participant and to the continuous development activity. This helps to reduce the risk of ill-defined requirements and helps to update the product backlog. It provides a mechanism for collaborative knowledge sharing that helps the participants to improve themselves for the next jam sessions. In the end of the process the result is a potential shippable product.

We think that the software jam session technique can help to improve collaborative work by allowing 7+/-2 participants to design together in the same location a video game. We believe that software jam sessions will encourage knowledge sharing between participants and enforce creativity.

Currently, we are capable to create collaboratively a playable video game level in a matter of minutes with the help of the Powwow tool. Without having real time capabilities built in the tool yet.

We think that collaborative design with the support of the jam sessions technique can speed up the development of complex game levels where workload has to be tailored between participants.

VI. PRELIMINARY EVALUATION

We conducted a quasi-experimental study where we measured the time needed for developing a video game prototype with the Powwow tool.

First we describe how we tested and then show the results together with our framework current limitations. In Fig. 6 the *X* axis represent the five users and the *Y* axis the time measured in minutes. The blue, red and green color represents the three runs each user made. In Fig. 7 the *X* axis represents questions and the *Y* axis points.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

We wanted to find out if the working efficiency is increasing when using the Powwow tool instead of writing code and if the tool is usable for developing runnable video game levels.

The quasi-experimental study was performed with 5 testers. The testers had never used the Powwow tool previously. We used shadowing to observe the testers during the experiment. The description of the task was provided to the testers at the beginning of the experiment. The time needed to complete the task was measured.

The study contained two parts. First the tester was introduced to the Powwow tool, which took around 5 minutes, the task was given to the tester, after finishing reading the task the time keeping was started, the tester finished the task, the time was stopped, the results were evaluated. The total time for each user was around 20 minutes.

The testers got on a sheet of paper the following task. *Please design a prototype having: one player, one enemy, one friend, one live item, one house, one tree and a five by five squared plane. Save the prototype. Simulate the prototype. Share the prototype onto the distributed file server.*

After each tester finished the task, they got a questionnaire, Table I., with five qualitative questions. All the questions had to be answered by checking a checkbox associated to each question. The possible answers were presented on a scale: 1 point (unsatisfactory), 2 points (satisfactory), 3 points (fair), 4 points (good), 5 points (very good).

The results show that one Powwow tester needs around three minutes to complete the task and the other 4 testers need between 5 and 10 minutes. Afterwards four users wanted to try the tool again.

TABLE I. THE SAMPLE QUESTIONAIRE

Questions	1	2	3	4	5
Q1:How do you find the usability?					
Q2:Are the tools implemented usefull?					
Q3:Did you had dificulties during prototype design?					
Q4:Are the pictures used for the buttons appropiate?					
Q5: Would you recommand the tool to a friend?					

The colors in the Fig. 6 represent the runs for each tester. After the second run almost all users improved their times. The time values presented in Fig. 6, of 0 min, 3.30 min, 50 sec, 2.20 min and 1.34 min represent the time difference between the first run and the third run for each of the testers.

Also we can observe that a learning curve appears for each of the testers. This means that the time needed to accomplish the same task reduces after the first attempt. Fig. 7 indicates that only 7 answers from a total of 25 answers are under the 2.5 average values. This means only 28% of the answers have obtained under 2.5 points, the maximum value being 5. Thus, 72% of the answers lie between 3 and 5 points in the Fig. 7.

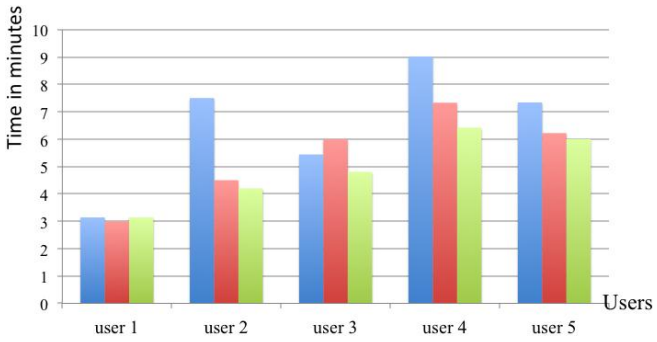

Figure 6. Time need for each tester

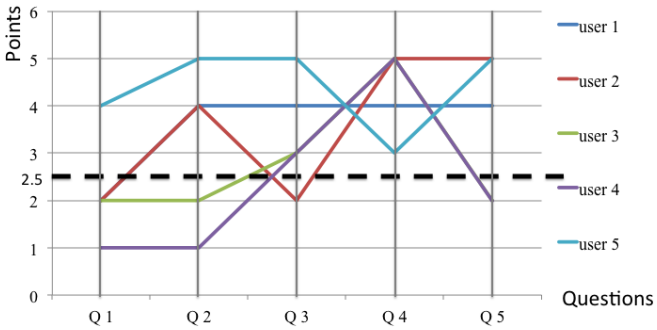

Figure 7. The results of the questionaire

The Powwow tool is not capable to support real time jam sessions yet. This issue has to be addressed in the future in order to achieve real time feedback during collaborative design of game levels. We did not test our collaborative jam session concept presented in section 5 because this represents work in progress. We plan to do this experiment also in the future when the Powwow tool is capable to support real time collaborative work.

VII. CONCLUSION AND FUTURE WORK

In this paper we described Powwow, a software tool that represents an alternative for developing fast runnable video game prototypes. Powwow can be used in the requirements elicitation phase and during the software jam session that we introduced in this paper too. The tool enforces communication and knowledge sharing through the interactive development work style.

We have made a case for rapid game prototyping as it can help to close the gap between the design model and user model. This does not mean that we not believe in the standard approach of firstly gather requirements and then develop incremental prototypes. We think that our tool is an alternative of gathering requirements and building fast runnable video game prototypes right away from the first

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

meeting with the customer.

By binding all the editing stages presented in section 4 together and implement in further releases of Powwow mechanisms that support real or near real time collaboration. We think that the tool should perform all the synchronization operations with the local file folder and the distributed file server independently.

In our case study we found out that developers need specialized apps for developing and testing RPG (role playing games) games. They would like to have a tool where they can right away test new game ideas without having to write necessarily code.

We also have introduced collaborative software jam sessions as an alternative development technique to take advantage of creative and ambitious game designers.

Finally, we have presented a preliminary study where we used our tool in order to develop runnable game prototypes. Until now we are able to work collaboratively with the Powwow tool and can develop a video game level in matter of minutes. We think that this result will motivate us to focus in the future on the development of real time capabilities in order to perform local collaborative jam sessions too.

ACKNOWLEDGMENT

Special thanks go to Damir Ismailović, the leader of the DRG (Dance Research Group). We would like to also thank the members of the DRG who provided key insights during the research phase of this paper.

REFERENCES

[1] D. Wuest, N. Seyff, M. Glinz, "Flexible, lightweight requirements modeling with flexisketch" Fourth International Conference on Mobile Computing, Applications and Services (MobiCASE), Springer-Verlag, Berlin, Heidelberg, October 2012, pp. 225-244, doi: 10.1007/978-3-642-36632-1_13.

[2] Codea app, http://twolivesleft.com/Codea/, retrieved: March, 2013.

[3] PGC app, https://itunes.apple.com/us/app/prototype-game-crafter-1.0/id419847669?mt=8, retrieved: March, 2013.

[4] Battle Map 2 app, https://itunes.apple.com/us/app/battle-map-2/id384800918?mt=8, retrieved: March, 2013

[5] J. Haladjian, D. Ismailović, B. Köhler and B. Bruegge, "A quick prototyping framework for adaptive serious games with 2D physics on mobile touch devices" IADIS International Conference Mobile Learning, March 2012, pp. 197-204, ISBN (Book): 978-972-8939-66-3.

[6] B. Bates, Game Design 2th edition. ISBN-10: 1592004938. Course Technology PTR, 2004, pp. 225.

[7] B. Bruegge, S. Krusche and M. Wagner, "Teaching tornado from communication models to releases" (EduSymp12), Proceedings of the 8th edition of the Educators' Symposium, Innsbruck, Austria, October 2012, pp. 5-12, doi: 10.1145/2425936.2425938.

[8] A. MacWilliams A Decentralized Adaptive Architecture for Ubiquitous Augmented Reality Systems Dissertation, Technische Universität München, June 2005, pp. 121.

[9] D. A. Norman and S. W. Draper, User centered system design; new perspectives on human-computer interaction. L. Erlbaum Associates Inc., 1986.

[10] Powwow tool, https://www.youtube.com/watch?v=njS2caLWFRI, retrieved: March, 2013.

[11] Tiled program, http://www.mapeditor.org/ , retrieved: March, 2013.

[12] A. I. Wasserman, "Software Engineering Issues for Mobile Application Development. " Proceeding FoSER '10 Proceedings of the FSE/SDP workshop on Future of software engineering research, 2010, pp. 397-400, doi: 10.1145/1882362.1882443.

[13] E. M. Maximilien and P. Campos, "Facts, trends and challenges in modern software, International Journal of Agile and Extreme Software Development, Vol. 1, No. 1, July 2012, pp. 1-5, doi: 10.1504/IJAESD.2012.048305.

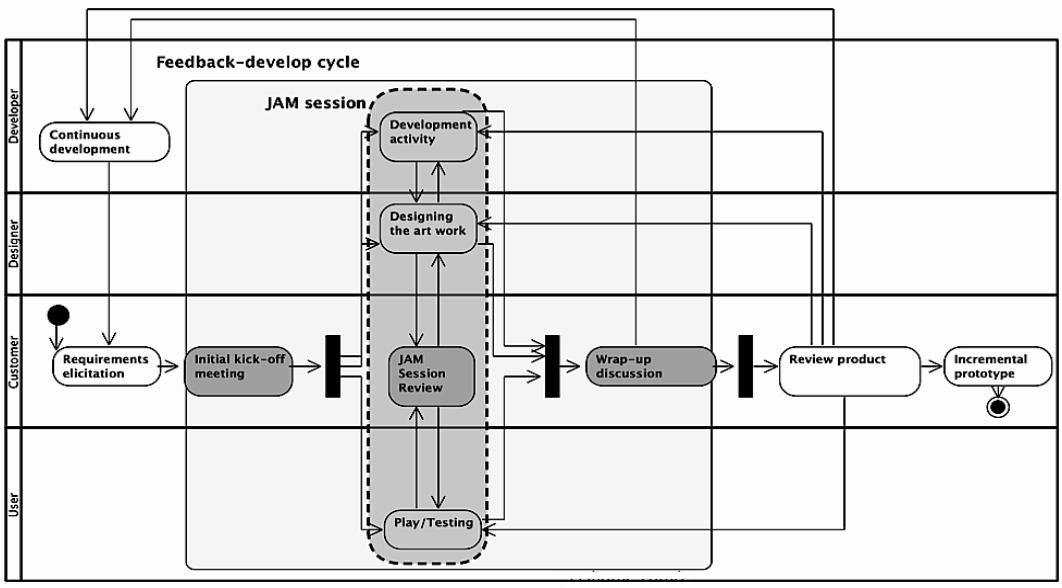

Figure 8. Collaborative software jam sessions

Exploring Collective Intelligence in Online Brainstorming

Yann Veilleroy
Faculté Libre de Médecine
Université Catholique de Lille
Lille, France
e-mail : yann.veilleroy@ed.univ-lille1.fr

Gabriel Eurin, Frédéric Hoogtoel, Luigi Lancieri
Laboratoire d'Informatique Fondamentale de Lille
Université Lille1 Sciences et Technologies
Villeneuve d'Ascq, France
e-mail : gabriel.eurin@polytech-lille.net,
{frederic.hoogstoel, luigi.lancieri}@univ-lille1.fr

***Abstract* — How can we know what has happened within a work group? How can we visualize the non-visual? How can we evaluate forms of interaction in the context of group work? These are some of the questions we tried to answer with Qlim, using online brainstorming activities. To achieve this goal, we collected data related to users' activities, then we used several visualization techniques to show the evolution of the interactions: zoomable graphs to get more detailed information and synthetic graphs to allow an overall perception of the activity using colorization and fill.**

Keyword-Interactions; collaboration; emergence; monitoring; visualization.

I. Introduction

Collective intelligence emerges from interactions between individuals when a group is involved in a creative task or in solving a problem [1]. The singularity of such a phenomenon is that the effect of the group action is greater than the sum of that of all individuals. Collective interaction is complex, involving direct and indirect causes that are difficult to observe and to link together.

Most of the time, the only indicator of collective efficiency is the final global result of the collaboration. On a large scale, open sources software development or Wikipedia are some examples of crowd sourcing results. Outside observation of these effects, the inner mechanisms of the collective phenomena are theorized, but few are measured in real situations. For this purpose, we developed a platform dedicated to supporting collective intelligence with a full statistical and graphical monitoring environment. This allows us to observe various causes and effects of the collective action (consensus evolutions, influence, changes of opinion, etc.) from a structural and temporal perspective.

The Qlim (in French: Questionnaire en Ligne Interactif et Malléable - Interactive, Tailorable, Online Questionnaire) platform was designed to support online group interaction on the basis of a shared feedback mechanism as occurs in a brainstorming session. In short, the process starts with few questions and answer choices (QA). Participants are invited by e-mail to come to answer these questions, add new questions and add new answer choices. They are notified by e-mail of all additions made during the day. They can return to modify their answers at any time or to make new propositions for question or answer choices, and so on. This system was tested in a dozen asynchronous brainstorming experiments, each involving an average of 20 people.

This paper is organized as follows. First, we will present Qlim and its role in supporting collective intelligence. Then, we will describe the monitoring environment, some examples of statistical and visualization results and an overview of the latest developments in the field of group dynamics monitoring. Finally, we will discuss these results and present our future work perspectives.

II. Collective phenomena in interactive questionnaires

As a tool to support collective intelligence, Qlim is a web-application (see Figure 1) that relates to the brainstorming technique. It is made to build questionnaires, but like a brainstorming session, Qlim is useful for collecting data on opinions or new ideas. These questionnaires, very easy to use, are spread out in time, and it is up to the creator to decide when the questionnaire ends. Qlim presents some special functions. It is an interactive questionnaire, which means the participants are not just passive respondents. With Qlim, respondents can add some new answer choices, and some questions, which comes down to extending the questionnaire to suit their desires and their trends. A questionnaire can thus take the form of an open debate, which can stimulate creativity and bring a lot of new thinking. Qlim has also several other features. It supports the respondents by sending a summary in the evening in case of additions made during the day, to encourage them to further fuel the debate. Qlim also has color user guidance to highlight the news, ensuring the respondent will focus on the news, if he wishes.

"To solve a problem, much of the solution lies in the wording of the question" [2]. Qlim allows expression and creativity, it helps participants formalize their subjects of interest into a structured frame (unlike an online open debate forum). It is based on a question and radio buttons system, designed to encourage the participants to make a precise choice or to ask some new questions or to add new answer choices, to react, to criticize or to develop; it is designed to be creative and to give an open direction to the discussion unlike traditional questionnaires. The goal of Qlim is not specially to lead to a consensus, it is a creative space: the final question can be off-topic with regard to the first question.

Qlim captures what is exchanged in a group. The problem in common brainstorming systems is the lack of structure; a regular brainstorming session is not especially structured (it returns a raw list of ideas). However, the Qlim structure allows easy exploitation of the results of any interaction. Qlim uses the frame (textbox, radio buttons) so the participants have to identify and break down what they want to say, to be synthetic, so they structure their thoughts. This way, all interaction can be logged and these traces can be exploited to understand the behavior, developments and influences.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

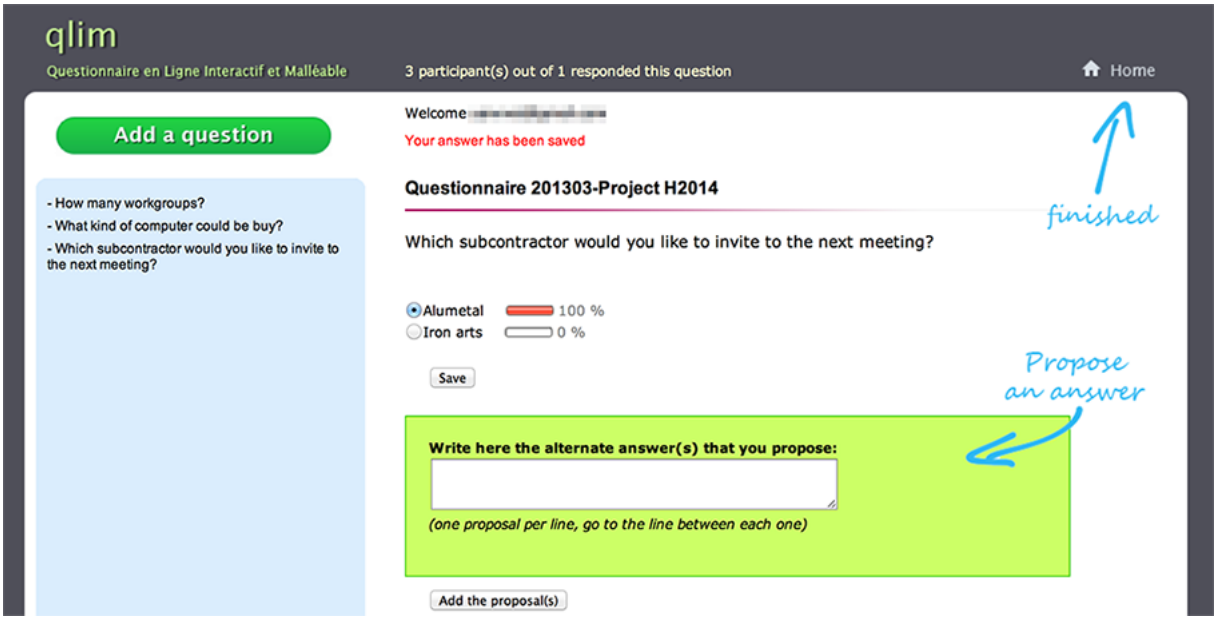

Figure 1. QLIM's page, to respond, add questions and answer choices, with guides.

Qlim's first aspect is the enhanced use of questionnaires, with two kinds of people, the creator who initiates the questionnaire and the participants. The second aspect is the use of the logged traces in order to monitor the interaction.

A. Qlim usage, from the creator's point of view

Figure 2 shows the user interface involved in this process. The "New questionnaire" button is available from the creator home menu. Creating a new questionnaire is simple. There are only four textboxes to fill, namely:

- the heading of the first question (for more questions, enter the questionnaire and press the "Add a question" button)
- some answer choices (if desired),
- the participants' e-mail addresses
- a name for the questionnaire

Figure 2. Creation principle

Before pressing the Save button, one important thing remains to be done: to choose "With" or "Without" the scores. Figure 3 shows the distinctive elements displayed when the creator chooses with the scores, namely, the number of participants who answered the question at the top of the page and a percentage at the end of the line for each answer choice (the "score"). It represents the success rate of each answer choice among the participants who answered the question. If present, these data may influence the participants because it becomes possible to know the trend of the group.

An Administration menu gives access to private functions to manage participants and questionnaires. The creator is able to list, to add or delete a participant to/from one of his questionnaires; he is also able to delete, to disable or enable his own questionnaires.

B. Qlim usage, from the participant's point of view

When a questionnaire is created, each participant receives an e-mail that is an invitation to participate in a Qlim questionnaire. Inside the message, a link directly leads to the web page with the first question. The participant makes a choice selecting a radio button and saves it by clicking the Save button. All the questions are listed in a column on the left side; the participant clicks the question he wants to answer (see Figure 1), in the order he wishes. If he changes his mind, he can go back to modify his answer at any time. If the existing answers choices do not suit him, the participant can propose a new one simply by writing one or more answer choices in the textbox just below the question. It is also possible to add a new question by a click on the big button "Add a question"

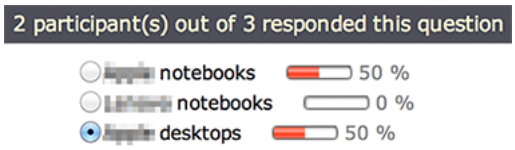

Figure 3. The distinctive elements of a question -with- scores

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

(see Figure 1). This can happen if a participant wants to ask a question to know the opinion of the others, to fuel the debate, or if he finds "it's not a good question!" or a relevant one, or if he wants to complete or to correct answers. He puts the heading and some answer choices if he wants, then clicks the Save button. It is also possible to add a question without any answer choices, in order to encourage questioning and to see the answers brought by the others. All new things, question and answer choices are available for all participants. Every evening, the new questions and the new answer choices are counted. If there were some creations, the summary of the day's activity is sent by e-mail to every participant.

Participants are helped to rapidly identify changes through the user interface colored marks, a little bit like unread messages in a mailbox that are highlighted in a bold font (see Figure 4). A question that did not exist the last time the participant logged on appears in a red font and a question with new answer choices appears in an orange font. This is to draw attention to what's new to encourage participation. The main menu gives access to the help function and enables every participant to know the results.

III. Monitoring tool in Qlim

Usually, traditional questionnaires provide statistics and graphs about their results. But the traditional questionnaires do not use the time or interactions links, so their results are like a snapshot. In comparison, Qlim brings new features that describe the complexity of the interaction more accurately. A questionnaire, like a brainstorming session, needs time. So the results cannot be just a mere snapshot, because it would be impossible to understand what happened without tracking the chronology of the interactions. A mere snapshot would be a great loss of valuable information. That's why, in addition to the usual standard numerical results, Qlim supplies trends over time, which represent, in fact, the evolution of the interaction. By building Qlim, our intention was to try to understand collective phenomena inside a group. So, from the beginning, Qlim was designed to record each interaction. Every time an interaction occurs, it is recorded. We determined five features that are stored in relation to the user id through the questionnaire's life duration: Visit a question, Answer a question, Add a new response choice, Add a new question, Modify an answer.

Used in Creative Problem Solving, the well-known 5 Ws (and one H) [3] (who, what, where, when, how, with what, and how), is an old formula used by police, journalists and researchers [4]. It inspired us, and in our database, the data structure that store the interactions matches the five Ws, and was even extended to store more parameters. The Who column stores the e-mail, the What column stores the type of interaction, the Where column stores the page name on which the interaction occurred, the When column stores the date (and hour), and the others store values like the questionnaire index, the question index, etc. By consulting this table, we generate reports, tables of numbers and graphs.

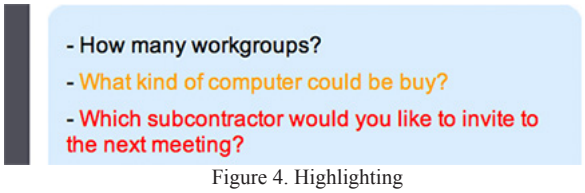

Figure 4. Highlighting

Exploring and analyzing vast amounts of data can be very difficult [5], but the way we stored the data is very helpful (storing using this method has already been used [6]). This storage method simplifies the queries a lot and allows a large number of graphs. We made general graphs for the whole questionnaire (see Figure 5) with a graph for each action, to show their evolution, graphs (histograms and pie charts) of the activity of each participant, and graphs of the actions for each question. It becomes possible to study both individual behavior and collective phenomena.

The first set of graphs uses the time as the abscissa axis and the different actions as the ordinate axis. So, we can produce the evolution in time of the five actions (Visit, Answer, Add a new response choice, Add a new question, Modify). Technically, when the creator asks for graphs, a data set is created using several loops, then sent to a graphic library for display (JPGraph). We get histograms, grouped bar charts, accumulated bar charts showing the quantity of interaction in one time slot, or histogram and line combined showing each interaction plus accumulated visits, either for the whole questionnaire or for one question at a time, for all participants or one participant at a time.

"We showed that our eyes can quickly evaluate a situation of cause and effect, without the help of our cognitive system" said Patrick Cavanagh in January 2013 [7]. Thus, we have visually arranged all our graphs in vertical alignment (see Figure 6). The time scale is the same for all graphs, so it is possible to read in a set. This enables us to find parallels between the different actions at the same time (e.g., the effect caused by the arrival of a particular question). It makes it possible to see in one look the collective phenomena, cause and effect relationships, influences that could explain a turnaround, anything that cannot be observed with traditional questionnaires.

A second set of graphs has come from the use of another recent graphic library, D3[1], which is made of JavaScript and Ajax and allows some interaction once the graph is plotted. For example, the abscissa axis of the graph is "zoomable" using a cursor. It does not bring new data, the data remain the same, but it allows a larger view, which can be useful at times. Another feature causes a resampling of the data table on display, to get this time a more detailed view. To get a more detailed view is a shared concern, it seems [8].

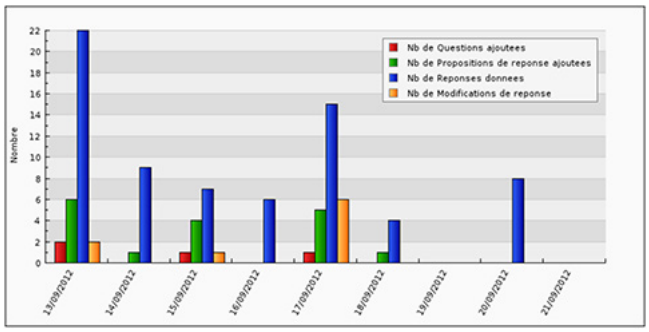

Figure 5. General graph for the whole questionnaire

[1] http://d3js.org (May 2013)

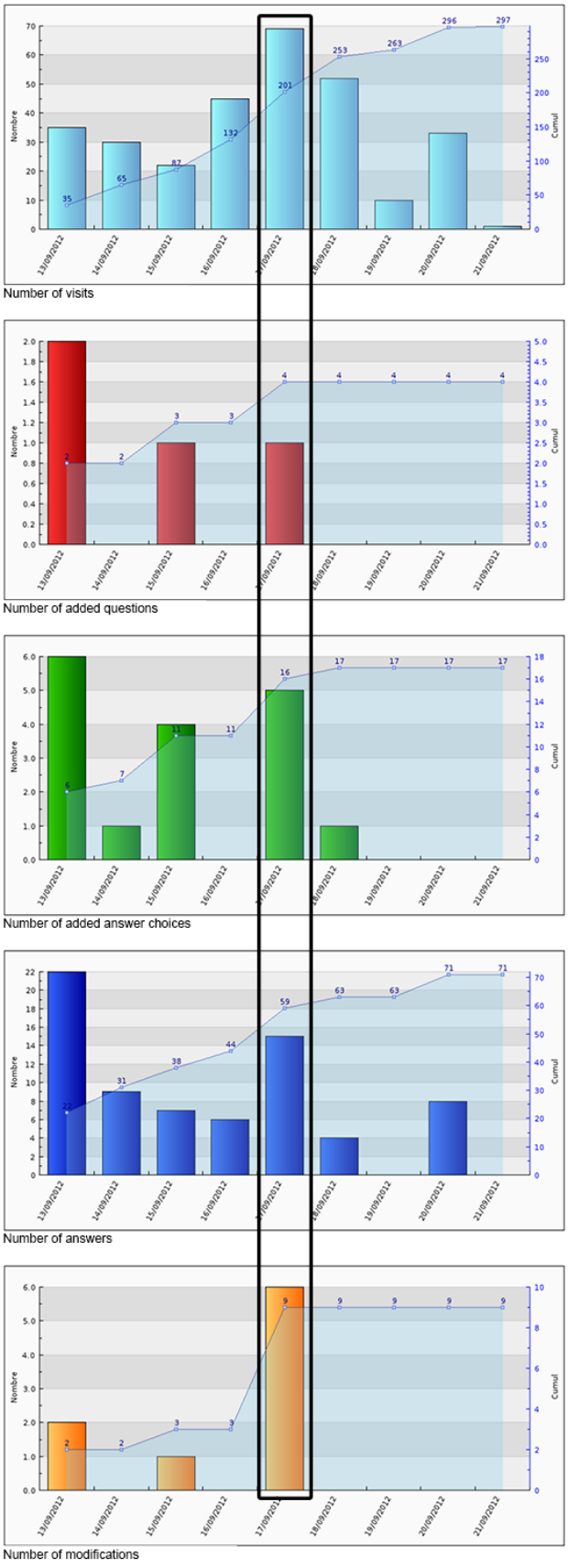

Figure 6. Each histogram represents an interaction

Besides the histograms, a third set of graphics is available. Collective intelligence stems from how well the group works together. Woolley [9] shows that a group where one person is dominant is less collectively intelligent and has poorer results (in solving problems - puzzles, games, etc.) than groups where conversational turns are more evenly distributed. The individual abilities of its members do not determine the performance of a group, and moreover a group whose members have higher social sensibility is more collectively intelligent. So, we made a view of the collective activity, of the distribution of the interaction and how many participants interacted. On a first graph (see Figure 7), we have drawn some donuts. The participant who made the highest number of interactions has a full donut, in proportion to his interactions. He is the maximum, the point of reference; all the others are calculated compared to him. We can see that even a lack of interaction has a specific meaning. We used the "Nothing" (transparent) interaction to fill the remainder of the donut to 100% when a participant made fewer interactions than the reference. There is one donut per participant on the graph and all the participants are on the graph, so this is another way to measure at a glance the collective activity, how it was distributed and how many participants have been interacting. A second graph represents a table (see Figure 8) with the time as the abscissa axis and the participants as the ordinate axis. In each cell, a matrix (2x2), contains four small color squares. Each one of the five interactions is represented by a color, for this participant, at this time. Calculated on the maximum number of interactions, the strength of the color (pale, medium, dark) shows how active the participant has been. This way, the group can take a dominant color or be a colored patchwork. The group colors make it possible to perceive at a glance the interactions that occurred. No red means that no questions were added, for example. A lot of orange squares means that the group changed its mind many times, etc. The more the group is monochrome, the less variety there has been in the interaction. On the contrary, a patchwork of colors means there have been many different interactions. This is a way to measure at a glance the collective activity, how it was distributed and how many participants have been interacting. We also included a wordcloud which is interesting to get a quick look at the most used words in a questionnaire. Data can be exported to a .csv file for further manipulation, for some graphics, this is a subject for our future work.

IV. State of the art

Traditional questionnaires are very regulated, framed, constrained (check boxes, radio buttons, textboxes for open-ended questions, etc.), impossible to go back (later) to modify answers. They are static: no "ping-pong" game between the participants, no real interaction is possible. Nevertheless, it was observed that the respondents respond willingly when they are given the opportunity to do so, and the web mode also appears to be easier for this: open-ended questions by e-mail collect longer answers and more information than a paper survey [10, 11, 12]. That's why an interactive questionnaire can be a place to collect lots of valuable information. We have noticed two other questionnaire-based collective intelligence tools: e-Brainstorming and Real-Time Delphi.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

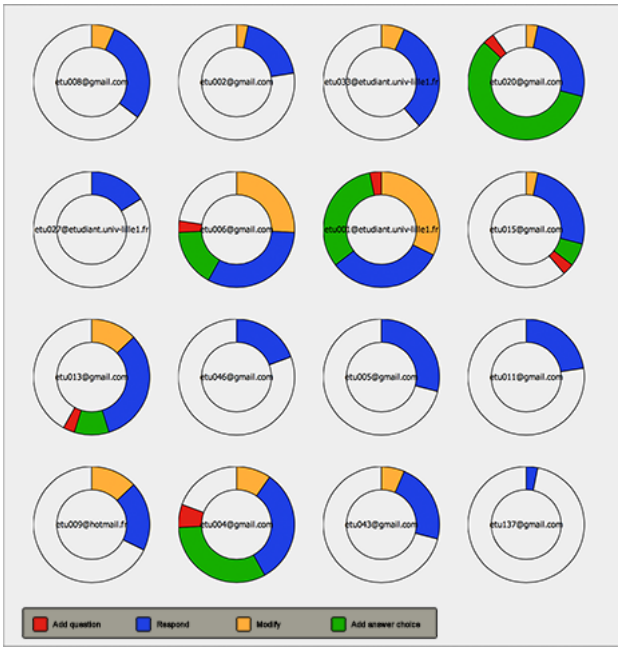

Figure 7. The collective activity by a donuts graph

A. e-Brainstorming

The eBrainstorming is an initiative from the Orange Labs in Caen [2]. It is a computerized system of close-ended questions, a multiple choice questionnaire system (MCQ), intended to simplify and synthesize the opinions of a group. The system allows the respondents to add new questions and possible answers, and they can leave a comment in a free comments zone inside a question. It works without a moderator; the group has to be self-moderated. The questionnaire is scripted using simple and easy tagged language. It is written in one form, on a mobile phone application for example [14], then sent to a web server. The server generates the questionnaire, then a return mail is sent to the author with the URL of the web questionnaire, to be distributed to the chosen participants. The participants can access the synthesis (statistics, graphics). The system uses traces to evaluate the collective intelligence phenomena, and data can be exported.

Face-to-face brainstorming was an idea of A. Osborn [15] in the nineteen-forties, introduced to make his company more creative. In a Brainstorming session, there must be no criticism in order not to hinder creativity, so any idea can be a starting point for creative development from other participants. It gives a raw list of ideas in a short time.

B. Delphi, Real-Time Delphi

The Delphi method is a structured communication technique, which tries to get opinions, judgments and justifications from the participants. It seeks a consensus (if possible), with a carefully prepared predefined set of questions, but here the creativity is controlled and contained. There are multiple rounds where questionnaires allow experts to provide their judgment, then to revise their answers (to be more accurate [16]). The process stops when a pre-defined criterion has been reached. The reasons and arguments collected can be highly valuable and useful. But Delphi takes

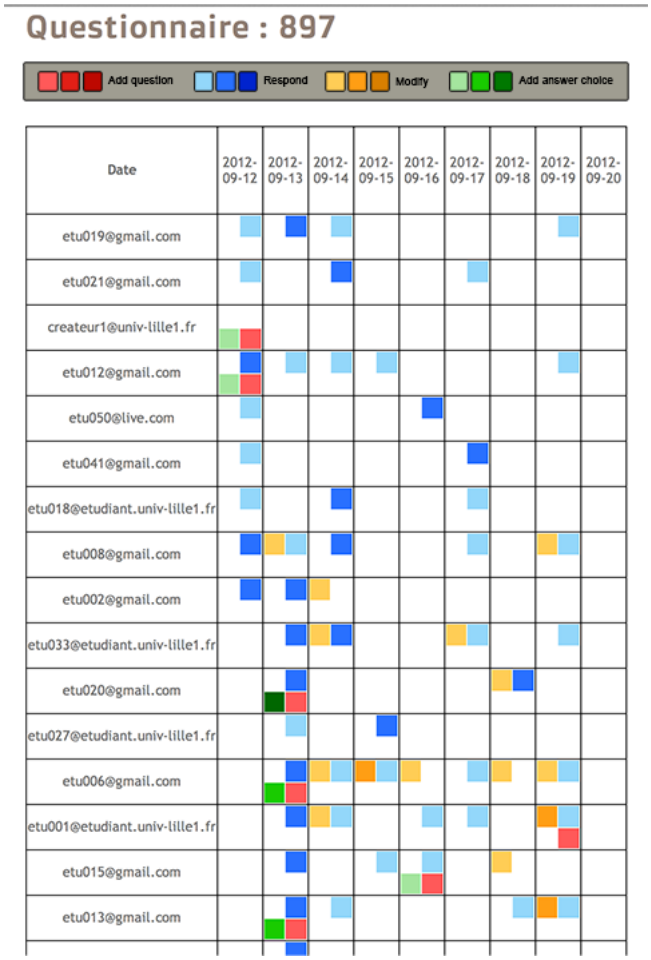

Figure 8. The collective activity by a table graph (anonymized)

time and requires good time management. It may be found long, expensive, tedious and requires a lot of effort [17].

Real-Time Delphi [13] is a computerization of the Delphi method, where Artificial Intelligence and Natural Language replace the human monitoring. It works “roundless”: every participant can come at anytime to update his input. Each question comes with some information (the average/median response of the group, the number of responses, the reasons).

C. Visualization of interactions

Wanting to know what happens during group work, and to read the interaction is a current need, at the origin of a large number of very diverse projects. In 2006, Calvani et al. [18] wanted to visualize effective interaction in online collaborative groups. They found several methods in the literature: Quantitative methods for content analysis, the most used (De Wever et al. [19] and Van Keer [20] made a review). This method, they say, consists roughly of coding single messages and statistically analyzing them to read the frequency and identify the relationships. Social Network analysis is another method used to analyze interaction among the members of a community, about this latter method, Calvani et al. [18] cite Cho et al. [21]; De Laat et al. [22]; Garton et al. [23]; Reffay and Chanier [24], but they found that the tools (very roughly: an analysis of individual messages between individuals) did

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

not fit their situation. They considered quantitative methods based on Conversation Analysis [25, 26] to be time-consumers. Calvini et al. [18] were working on a forum inside Moodle[2], and they improved this module in order to monitor interaction. However, to observe interaction, a forum has one serious drawback, that is, the different actions are not quite distinct. Indeed, it is possible to add a post, for example, and we agree that this is an interaction which is possible to trace. However, one can do a large number of things in a single post, and it is difficult to analyze every interaction in that case. A forum does not seem to be the best tool to achieve that goal, because the actions are not distinct enough.

V. Conclusion and future work

As we have seen, this paper presents a tool that has two strengths in its combined features. First, the malleability (tailorability) of Qlim simplifies and enhances collaboration within groups. This helps the emergence of creative new ideas and solutions. Secondly, Qlim provides monitoring of non-visual aspects of the collective phenomena. In this regard, we can say that Qlim is more a laboratory tool than an operational one. There remain several possible improvements, for example to compare various experiments. Another kind of representation is needed, based on new indicators of collective collaboration with more descriptive graphics.

References

[1] T. Malone, R. Laubacher, and C. Dellarocas, "Harnessing crowds: Mapping the genome of collective intelligence," 2009.

[2] L. Lancieri, A. Lavallar, and P. Manson, "E-Brainstorming: Optimization of collaborative learning thanks to online questionnaires," In Proceedings of CELDA IADIS International conference. 2005.

[3] D. Ruellan, "La routine de l'angle." Dans Questions de communication, 2006, no. 10, pp. 369-390.

[4] L. Lancieri and N. Durand, "Internet user behavior: compared study of the access traces and application to the discovery of communities," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2006, vol. 36, no. 1, pp. 208-219.

[5] D.A. Keim, "Information visualization and visual data mining," Visualization and Computer Graphics, IEEE Transactions on, 2002, vol. 8, no. 1, pp. 1-8.

[6] J. Ćosić, Z. Ćosić, and M. Baća, "An Ontological Approach to Study and Manage Digital Chain of Custody of Digital Evidence," Journal of Information and Organizational Sciences, 2011, vol. 35, no. 1, pp. 1-13.

[7] M. Rolfs, M. Dambacher, and P. Cavanagh, "Visual adaptation of the perception of causality," Current Biology, 2013, vol. 23, no. 3, pp. 250-254.

[8] A. Lavallard and L. Lancieri, "Observation de l'évolution des communautés d'intérêts," 2004.

[9] A.W. Woolley, C.F. Chabris, A. Pentland, N. Hashmi, and T.W. Malone, "Evidence for a collective intelligence factor in the performance of human groups," Science, 2010, vol. 330, no. 6004, pp. 686-688.

[2] http://moodle.org (May 2013)

[10] D. Bachmann, J. Elfrink, and G. Vazzana, "Tracking the progress of e-mail vs. snail-mail," Marketing Research, 1996, vol. 8, pp. 31-35.

[11] K.B. Sheehan, "E-mail survey response rates: A review. Journal of Computer-Mediated Communication," 2001, vol. 6, no. 2.

[12] A.M. Paolo, G.A. Bonaminio, C. Gibson, T. Partridge, and K. Kallail, "Response rate comparisons of e-mail-and mail-distributed student evaluations. Teaching and Learning in Medicine, 2000, vol. 12, no. 2, pp. 81-84.

[13] T. Gordon and A. Pease, "RT Delphi: an efficient,"round-less" almost real time Delphi method," Technological Forecasting and Social Change, 2006, vol. 73, no. 4, pp. 321-333.

[14] K. Hamadache, P. Manson, and L. Lancieri, "Pervasive services, brainstorming in situation of mobility," In : Pervasive Computing and Applications, 2008. ICPCA 2008. Third International Conference on. IEEE, 2008. pp. 709-714.

[15] A. Osborn, "Your creative power," 1950.

[16] T.J. Gordon, "The Delphi method," Futures research methodology, 1994, pp. 1-33.

[17] J.P. Ekionea, P. Bernard, and M. Plaisent, "Consensus par la méthode Delphi sur les concepts clés des capacités organisationnelles spécifiques de la gestion des connaissances. Recherches qualitatives, 2011, pp. 168.

[18] A. Calvani, A. Fini, M. Molino, and M. Ranieri, "Visualizing and monitoring effective interactions in online collaborative groups," British Journal of Educational Technology, 2010, vol. 41, no. 2, pp. 213-226.

[19] B. De Wever, T. Schellens, M. Valcke, and H. Van Keer, "Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review," Computers & Education, 2006, vol. 46, no. 1, pp. 6-28.

[20] L. Rourke, T. Anderson, D.R. Garrison, and W. Archer, "Methodological issues in the content analysis of computer conference transcripts," International Journal of Artificial Intelligence in Education (IJAIED), 2001, vol. 12, pp. 8-22.

[21] H. Cho, M. Stefanone, and G. Gay, "Social information sharing in a CSCL community," In : Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community. International Society of the Learning Sciences, 2002. pp. 43-50.

[22] M. de Laat, V. Lally, L. Lipponen, and R.J. Simons, "Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis," International Journal of Computer-Supported Collaborative Learning, 2007, vol. 2, no. 1, pp. 87-103.

[23] L. Garton, C. Haythornthwaite, and B. Wellman, "Studying online social networks", Journal of Computer-Mediated Communication, 1997, vol. 3, no. 1.

[24] C. Reffay and T. Chanier, "How social network analysis can help to measure cohesion in collaborative distance-learning," In : proceeding of Computer Supported Collaborative Learning conference (CSCL'2003). 2003. pp. 343-352.

[25] P. Ten Have, "Doing conversation analysis," Sage Publications Limited, 2007.

[26] H. Sacks, "Lectures on conversation,", 1995.

Impact of LinkedIn on Boosting Students' Learning Motivation and Career Prospects

Bernard Montoneri
Department of English Language, Literature and Linguistics
Providence University
Taichung, Republic of China (Taiwan)
bmon@pu.edu.tw

Abstract—**This study aims at evaluating the impact of the professional network LinkedIn on students' motivation and learning performance and on boosting their career prospects. This paper's main contribution is to apply Data Envelopment Analysis to demonstrate that student's learning motivation, communication, and writing skills can be improved by using LinkedIn.**

Keywords-Data envelopment analysis; LinkedIn; Malmquist index; professional network.

I. INTRODUCTION

There are around 160 Higher Educational Institutions in Taiwan for a population of 23 million; however, the birthrate is the lowest in the world, which will inevitably cause damage to Taiwanese society, economy, and education system [1]. As a result, students will find it increasingly difficult to graduate (higher tuition fee, more competition) and to be hired (higher unemployment, notably in the sector of education). Professional networking might be a way for students to develop their skills, their motivation, and to boost their career prospects. However, a rapid search on the Internet shows that a surprisingly small number of Taiwanese teachers and students have their profile on LinkedIn, the world's largest professional network. We have begun a study aimed not only at analyzing the present situation, but also at encouraging students to connect in order to improve their writing and communication skills, to increase their learning motivation, and to boost their chances on the job market. In that matter, so far, academic research on LinkedIn is scarce. However, this field of study is very promising in Taiwan and in other countries facing lowering student enrollment, higher graduate unemployment, credential inflation, and even department or university closure.

This paper presents some publications related to networking and to the selected evaluation method (Data Envelopment Analysis [1, 11, 12, 13, 15] and Malmquist index [13, 19]) in the literature review. The next section presents the objectives, methodology, and data source. The discussion is preliminary and can only present the expected results as we are still in the process of collecting the data.

II. LITERATURE REVIEW

A. Professional networking

Rosenfeld defines a network as a place where "the members choose each other, for a variety of reasons; they agree explicitly to co-operate in some way and to depend on each other to some extent." [2].

In the first decade of the twenty-first century, hundreds of social and professional networks have been created online, with more or less success. MySpace, Facebook, Twitter, and LinkedIn are obviously among the most famous. LinkedIn is a business-oriented social networking site founded at the end of 2002 and launched officially on May 5, 2003, that is, more than 10 years ago. Data about how many professionals are connected vary widely; LinkedIn reports more than 200 million in more than 200 countries and territories (including 74 millions in the US); the largest industries were at the beginning of 2013 information technology and services, financial services, higher education, and computer software [3]. Forbes states that LinkedIn is nowadays the most advantageous social networking tool available to job seekers and business professionals [4].

Many articles in newspapers [5], books [6, 7, 8], and portals [9] focus on how to help students finding a job. But, very little has been published so far in academic journals about professional networking and how it can boost students' career plan. A study published in 2009 [10] shows that on the question "How beneficial would a graduate-level professional networking course be to helping you improve your skills?" 34 % of the students interviewed answered "somewhat beneficial", 32% "beneficial", 18% "very beneficial", and 8 % "invaluable". Moreover, thirty-one percent of elite professionals spend one to two hours each week networking. Clearly, it takes time to build connections and make friends. Professional networking is a long-term investment, but twenty-seven percent of respondents network only when there is an immediate need, such as when starting a job search [10].

B. Data Envelopment Analysis and Malmquist index

We are now in the process of collecting the data necessary for this study. He expects to apply Data Envelopment Analysis (DEA), which is a quantitative evaluation method that can evaluate the relative performance of units called Decision Making Units (DMUs). There is a large body of literature on application of DEA in various fields, including education [11, 12]. Various studies evaluate the performance of departments [13, 14], classes [15], or individual students [16]. We have demonstrated the reliability and relevance of applying DEA in various academics papers, such as [17][18].

Malmquist index, a bilateral index defined by Caves et al. [19], has notably been applied in the field of education [13].

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

This index can help measure and analyze the improvement in motivation and learning over a period of time, e.g., from the time students register on LinkedIn to the time they find a job. By combining DEA with Malmquist index, this study will show the evolution of students' motivation, of their scores, and of their participation to LinkedIn from February to June 2013. This study is a work in progress and more time is needed to collect, run, and analyze the data.

III. METHODOLOGY

A. Research Issue

In February 2013, Maclean's Magazine in Toronto published an article stating that Higher Educational Institutions do very little to help students plan for future careers [5]. Some students feel that teachers and universities don't do enough to help them connect to the market place. Perhaps, in times of economic crisis, HEIs should offer more concrete career advice and help students post their profile on professional networks (as to February 2013, on more than 11,000 students at our university, less than 10 had a profile on LinkedIn and few of the students interviewed so far had heard about the network). Nowadays, social and professional networks are a great and fast growing place where people can connect and boost their career prospects. Students are often clueless about future pay and job prospects. Surprisingly, they know little about how to build their network and how recruiting works.

B. Research objectives

This research explores a number of issues in relation to networking practices. The objectives of the research are as follows:

- Observe the state and evolution of students' motivation from the beginning to the end of the semester;
- Ascertain the perceived understanding of networking by Taiwanese students;
- Explore how students form, maintain and develop their network and
- Follow students from the moment they register on LinkedIn; see if they can ultimately find a job.

C. Research methodology

The research methodology for this study is both qualitative and quantitative. At the beginning of the semester, students answer a questionnaire we have designed. This questionnaire includes 20 questions concerning students' knowledge about and participation to social and professional networks. Some questions are related to indicators of good teaching. Students can rank which indicators they feel are the most important (such as fair grading, being supportive, communication skill, teaching contents, etc.). The research methodology is also quantitative and DEA combined with Malmquist index will assess the progress and motivation of students from the beginning (period 1) to the end of the semester (period 2). Some of the indicators include the number of connections and the number of groups students follow or create on LinkedIn, scores during the semester (notably evaluation of their resume, portfolio, and profile), and students' evaluation of teaching.

D. The data source

The study case is a private church-sponsored university in Taiwan. Preliminary data consist of four classes of students (sophomore to senior students from various departments). Later on, a larger number of classes will participate to the study. We will follow students from the moment they register to LinkedIn to their first job offers.

IV. PRELIMINARY RESULTS AND DISCUSSION

Out of 11,000 students at our university, less than 10 had a profile on LinkedIn. As to March 2013, there were 81 followers of the institution, mostly registered teachers, staff members, and students from the university. All included, only 73 people had registered, often with an empty profile; that is, no picture, no information about education, experience, interests, and almost no connections.

We propose to define three steps in order to help students boost their career on LinkedIn:

1. We register first and connect to a large number of universities, companies, and recruiting agencies around the world. We registered on LinkedIn on February 12, 2013. One month later, he had more than 350 connections (more than 50 in Taiwan, same in Japan, 27 in China, and many in the US, England, France, etc.), many recommendations and endorsements from colleagues. As to July 2013, we had 1370 connections and we were linked to 12 million people.
2. We help students register and create their profile (developing their writing and communication skills). We create a discussion group on LinkedIn including all the students registered and we invite teachers and professionals to join. The group has been created in March 2013. It is called "The Student Connection".
3. We help students build connections according to their skills, interests, and career plan. The teacher can here recommend students and endorse their skills. One of the students has already 90 connections and many skills endorsed. We have accepted to recommend him.

A. Some preliminary results

On 93 students who responded to a questionnaire at the beginning of March 2013:

- 23 had heard about LinkedIn.
- 80 thought about going online to look for a job.
- 91 had a Facebook account.
- Only 13 used Twitter.
- 23 had penpals abroad.
- 24 ran a website.
- 25 were blogging.
- 62 thought their teachers were helping students enough to connect to companies.

We additionally asked 26 students coming from various departments at a national university in Taiwan to fill the questionnaire. 100% of the respondents answered yes to the

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

question "Do you have a Facebook account?" Students enjoy social networking, but don't consider it might help them find a job. Only 6 students had heard about LinkedIn (including 2 who already had a profile) and only 5 were using Twitter.

B. Discussion

In order for this research to become reliable, we will improve the questionnaire, interview thousands of students, and analyze the data by using SPSS, DEA model, and Malmquist index. So far, the number of students who responded to the questionnaire is too small to draw reliable conclusions; however, some trends appear: even though almost all the students use Facebook, few of them are connected to LinkedIn. Most students seem to be unaware of the fact that CEOs and recruiting agencies are now using LinkedIn to hire graduates. Only 10 out of 93 students who answered the questionnaire said they have a very good idea about what job they want to do after graduation.

This study expects to show that:

- Student's learning motivation will increase if they built a professional network and connect with CEOs and recruiting agencies;
- Students' grades in English writing will improve by using the network and by communicating with people around the world in English;
- Students' connections will help them to have a better understanding of the current job market and of what employers are expecting from graduates (employability skills);
- Undergraduate students will increase their chances to enter a graduate school in Taiwan or abroad by connecting to professors in their field of research;
- Some students participating to this study will eventually find a job with the help of their connections.

This is obviously a preliminary research, and it appears that more time is needed to introduce LinkedIn to workers and students at the institution studied in this paper and to analyze the impact of having new members registered.

V. CONCLUSION AND FUTURE WORK

Even though many university educators share their well-designed profile on LinkedIn, they rarely invite their students to connect. Most students are clueless about how to make their profile and about what good can come out of joining the network. Out of 11,000 students at our university, less than 10 had a profile on LinkedIn and almost none of the students interviewed so far had heard about it. In March 2013, around 100 more students connected and learned how to use LinkedIn in order not only to improve their English, but to boost their career prospects. From March 2013 on, we will follow and guide them and many other students to observe their progress in writing and communication skills, hoping for the best, that is, job offers. Clearly, there is little academic research done on LinkedIn and its impact on students' motivation and career prospects. According to a recent article [20], 97 percent of recruiters are using LinkedIn to find new employees, while only 27 percent are using Twitter and just 22 percent turn to Facebook as a source for recruiting.

REFERENCES

[1] B. Montoneri, C. C., Lee, T. T., Lin, and. S. L., Huang, "A Learning Performance Evaluation with Benchmarking Concept for English Writing Courses," Expert Systems with Applications, vol. 38(12), Dec. 2011, pp. 14542-14549, doi: 10.1016/j.eswa.2011.05.029.

[2] S. Rosenfeld, Industrial Strength Strategies: Regional Business Clusters and Public Policy. Washington, DC: Aspen Institute, 1995.

[3] D. Nishar, 200 Million Members! Blog from January 9, 2013, [retrieved: July, 2013] http://blog.linkedin.com/2013/01/09/linkedin-200-million.

[4] Daily Muse, Your LinkedIn Intervention: 5 Changes You Must Make. Forbes, 7/06/2012, [retrieved: July, 2013] http://www.forbes.com/sites/dailymuse/2012/07/06/your-linkedin-intervention-5-changes-you-must-make.

[5] J. Dehaas, Finding work: the missing link for university students, Maclean's Magazine On Campus. February 9th, 2013. [retrieved: July, 2013] http://oncampus.macleans.ca/education/2013/02/09/the-missing-link/

[6] F. Agin and L. Howes, LinkedWorking: Generating Success on LinkedIn ... the World's Largest Professional Networking Website. Columbus, Ohio: Four Eighteen Enterprises, 2009.

[7] B. Schepp and D. Schepp, How to Find a Job on LinkedIn, Facebook, Twitter, MySpace, and Other Social Networks. New York: McGraw-Hill, 2009.

[8] T. Vickey, Social Capital and the Role of LinkedIn to Form, Develop and Maintain Irish Entrepreneurial Business Networks. Newcastle upon Tyne: Cambridge Scholars Publishing, 2011.

[9] C. Allen, LinkedIn Student Job Portal: Helping Recent Graduates Find Jobs. LinkedIn's Official Blog, 2011′s Top 10 LinkedIn Blog Posts, [retrieved: July, 2013] http://blog.linkedin.com/2011/03/21/linkedin-student-job-portal.

[10] R. Gregoire (editor; The Write Idea), Professional Networking and Its Impact on Career Advancement. Professional Networking and Career Advancement Report, Upwardly Mobile, Inc., 2009.

[11] F. R., Førsund and N., Sarafoglou, "On the Origins of Data Envelopment Analysis," Journal of Productivity Analysis, vol. 17(1-2), 2002, pp. 23–40.

[12] B. Montoneri, T. T., Lin, C. C., Lee, and. S. L., Huang, "Application of data envelopment analysis on the indicators contributing to learning and teaching performance," Teaching and Teacher Education, vol. 28(3), pp. 382-395, doi.org/10.1016/j.tate.2011.11.006.

[13] A., García-Aracil and D., Palomares-Montero, Changes in universities' efficiency over the time: Differentials according to the missions. First ISA Forum of Sociology: Sociological Research and Public Debate, Barcelona, 2008. [retrieved: July, 2013] http://digital.csic.es/handle/10261/10097.

[14] A., Colbert, R., Levary, and M., Shaner, "Determining the relative efficiency of MBA programs using DEA," European Journal of Operational Research, vol. 125, 2000, pp. 656-660.

[15] B. Montoneri, C. C., Lee, T. T., Lin, and. S. L., Huang, "Application of DEA on teaching resource inputs and learning performance," Proc. in S. N. Chin (Ed.),

Emerging trends in higher education learning and teaching. TARC International Conference on Learning and Teaching, 2010, pp. 31-38.

[16] B., Montoneri, M., Moslehpour, and H.-N. C., Chou, "Applying DEA to assess English writing progress of university students using MyAccess in Taiwan," Proc. ACLL 2012, The Asian Conference on Language Learning, Osaka, Japan. 2012, pp. 165-175. ISSN: 2186-4691. http://iafor.org/ACLL2012_proceedings_reduced.pdf.

[17] B. Montoneri, T. T., Lin, C. C., Lee, and. S. L., Huang, "Application of data envelopment analysis on the indicators contributing to learning and teaching performance," Teaching and Teacher Education, vol. 28(3), 2012, pp. 382-395.

[18] B. Montoneri, C. C., Lee, T. T., Lin, and. S. L., Huang, "A Learning Performance Evaluation with Benchmarking Concept for English Writing Courses," Expert Systems with Applications, vol. 38(12), 2011, pp. 14542-14549.

[19] D. W., Caves, L. R., Christensen, and W. E., Diewert, "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," Econometrica, vol. 50(6), 1982, pp. 1393–1414.

[20] D. Mielach, The Social Network Most Recruiters Use, BusinessNewsDaily, LiveScience.com. May 2, 2013. [retrieved: July, 2013] http://news.yahoo.com/social-network-most-recruiters-191750747.html.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Does the Acquaintance Relation Close up the Administrator Community of Polish Wikipedia?

Analysing Polish Wikipedia Administrator Community with use of Multidimensional Behavioural Social Network

Justyna Spychała
Faculty of Polish Studies
University of Warsaw
Warsaw, Poland
Email: jspychala@gmail.com

Mateusz Adamczyk
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw
Warsaw, Poland
Email: adamczyk@mimuw.edu.pl

Piotr Turek
independent
Warsaw, Poland
Email: piotrt@gmail.com

***Abstract*—The aim of this paper is to answer a question if a group of administrators of Polish-language Wikipedia is closing up because of relation of acquaintance. The administrators are usually nominated by the community and then elected by voting. This research is an attempt to find out whether the community is not becoming less open to new users and new potential administrators, because of their lack of chances to gain reputation. The key here is the understanding of social aspects driving the process of Request for Adminship (RfA) votings. Based on our previous work focused on social networks induced from collective activity of wikipedians, this paper extends it by the annual analysis of obtained statistics and examination of clustering coefficient as an approximation of social capital. We present the dynamics of relationships between voters and candidates across several years of Polish Wikipedia development.**

Keywords*—*Multidimensional Behavioral Social Network; Wikipedia; Request for Adminship; Clustering Coefficient.

I. INTRODUCTION

Administrators (or sysops) are very dedicated and trustworthy participants of the Wikipedia projects in all language versions. Thanks to community decision, they have received special privileges and—due to their entitlements they are able to use administrative tools—they exercise preventing and policing functions. Administrators have the right to edit all the Wikipedia articles as well as many other privileges—understood rather as duties. These powers are not to suggest editorial control over the project, but rather provide mentoring and technical assistance in other wikipedians' work. Administrators also serve by helping, especially to newcomers, in editing of Wikipedia—all newly registered users get their guides—the administrators to whom they can always turn for help and be sure they will receive it as soon as possible.

Due to increasing amount of management work at Wikipedia, such as content quality control, coordination, maintenance, that are caused by the increasing popularity and amount of content in Wikipedia [1], the importance of administrators is increasing. This creates a potential risk that administrators may become overwhelmed by the amount of work and their response time become longer.

However, the Polish community of Administrators is growing slower than expected; hence, the question whether and why this community closes up. Currently, there are 149 administrators on Polish Wikipedia—for comparison, 1,147 administrators work currently on the English version. Of course, the English version is much more developed, but sheer number of people with administrative privileges is impressive.

Administrators are elected during a special procedure the rules of which are clearly defined. This procedure is called Request for Adminship (RfA). As it was already mentioned, the privileges for administrators are granted by Wikipedia community. They are granted by voting of the Wikipedia users who are well-known and respected members of the community and know and respect the established rules on the website. Wikipedians who candidate for the administrator must "have a minimum of 1,000 not deleted edits, first of which has to be made at least 3 months prior to the date of filing the candidacy". Nominations for administrator candidates are adopted by a special form on the web page that also contains the regulations and the list of candidates. New administrators are elected during a voting that lasts a week (168 hours). Wikipedians who are allowed to vote must be registered for at least one month and must have a minimum of 500 not deleted edits.

Interestingly, in case of English version of Wikipedia, no formal conditions are required in order to declare a candidacy for an administrator. The only conditions are possession of an account and trust among other users. Despite this, the page with the declaration forms contains the information that in case of self-nominating, it is recommended to have at least 2,000 edits for a minimum period of 3 months. Another important difference is that in the case of English version of Wikipedia, new administrators are elected not by voting, but by discussion. Moreover, "the consensus in RFA is not achieved by exceeding a threshold, but by the strength of the justification of the candidacy".

This paper is a refinement of work done by Turek et al. [2]. It covers the range of years 2005-2011 and its aim is to answer the question if the reason for decreasing in successful RfA votings in Polish Wikipedia is choosing Administrators based

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

on acquaintance. We argue that it is not the case. Probably, it is due to growing expectations about new candidates.

The rest of this work is divided as follows: in Section II, the related work is presented. Section III contains base statistics, which show that growth of Polish-language Wikipedia Administrators group has slowed down. Data presented there are extended by year 2011, in comparison to [3]. In Section IV, Multidimensional Behavioural Social Network is used to analyse historical voting data. This analysis is the main contribution of this paper. Section IV also contains answer to question stated above. Section V presents conclusions and suggestions for future work.

II. Related Work

The problem of evaluation and recommendation of users requesting for adminship in Wikipedia has been addressed in several papers. In one of them, Burke et al. [4] try to indicate the features and qualities determinative for the user selection to the position of administrator. On the basis of publicly available tips for candidates [5], a set of attributes, that a future administrator should have has been developed. Behavioural data and comments, not page text, were used to evaluate candidates. Authors counted each candidate's edits in various namespaces (article, article talk, Wikipedia, Wikipedia talk, wiki projects, etc.) to calculate total contribution as well as contribution diversity. They also measured user interaction, mainly activity on talk pages, but also participation on arbitration or mediation committee pages and a few others. There are also several other statistics, but the ones mentioned seemed to be the most relevant to the candidate's success. Especially successful were candidates with strong edit diversity, mere edits in Wikipedia articles didn't add much more chance of success. In user interactions, article talk page edits were the best predictor of success, with other authors talk page edits being rather poor. Burke et al. also confirmed Kittur's [6] results that the percentage of indirect work (coordination, discussion, etc.) grows over time, the share of articles in all Wikipedia edits is decreasing.

It is noteworthy that in [4] only qualities of each user were evaluated. Leskovec et al. [7] have shown that the outcome of the voting depends on the candidate and his or her place in the community. They found out that the probability of one person's vote to be positive is correlated with the basic relative figures such as: who—voter or candidate has more edits, who has more barnstars (awards given by other Wikipedia users), the extent of collaboration of the two, etc. Authors strongly noted that the vote value (positive or negative) is not just a function of candidate, but both voter and candidate.

In [8], the impact of the similarity of users on their mutual assessment has been analysed. The examined data were collected from three websites: Wikipedia, Stack Overflow, and Epinions. The important feature of those websites is the possibility of mutual evaluation between their users. In case of Wikipedia it is the RfA voting. Two users were considered similar, when they have performed similar actions, which in case of Wikipedia were edits of articles. The authors concluded that, in case of Wikipedia, the possibility of casting a vote for a candidate increases with the increase of the similarity between the candidate and the voter. The voters, who are similar to the candidate, are less driven by the objective qualities (status in the community), such as experience in development of Wikipedia. Candidate's status determines casting a vote for that candidate when the voter and candidate are only slightly similar.

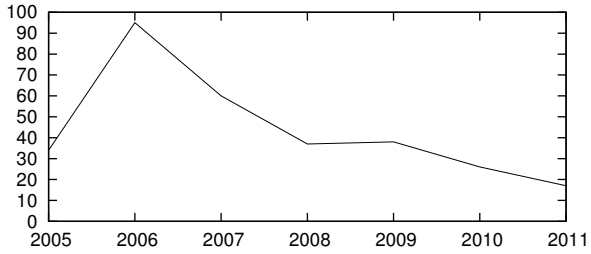

Fig. 1. Number of votings in each year

An interesting observation is that during voting, there are much more voters similar to the candidate in a group with higher status than in a group with relatively lower status. This may suggest that during RfA the voters do not constitute a representative sample of community. This allows for the outcome of the election to be predicted when profiles and similarity of a few first voters and the candidate are known. To effectively predict the result of the voting one does not even need to know the votes given by the first voters.

The quality of Wikipedia articles depends on the level of cooperation of the editors. Rad et al. [9] decided to examine the history of article edits, and on that basis determine the mutual attitude of the editors and how controversial is the given article. Casting a vote during new administrator election was adopted as an indicator of relation between two users. If the voter has a positive attitude toward the candidate, the vote will be positive. In case of a negative attitude, the vote will be negative. The authors decided that the co-edit of the article is a pair of changes of the same section of the article, which were set apart in time by less than a fixed number of revisions. A social network with nodes labelled with users' profiles and directed edges labelled with users' co-edits, was also considered. This graph was used to induce a decision system and train a classifier, which was highly effective in predicting votes. What is important, is that this approach is complementary to the ones described earlier. It is based on the analysis of Wikipedia articles and their edit history and not on the aggregated statistics of the community. What is interesting, it turned out that it is relatively easy to predict positive votes. It seems that they are influenced by the most recent history of cooperation. On the other hand, the high quality of prediction of the negative votes required appropriately bigger and richer history of cooperation. The authors risked the statement that the users can remember disagreements for a long time and during a voting they can be guided by hidden qualities, like for example, the votes already cast in a given voting.

III. Statistics for Request for Adminship Procedure

As of December 31, 2011 the Polish-language Wikipedia had 171 administrators. Since 2005, there were 307 votings on RfA. 177 of those ended with granting the administrator

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

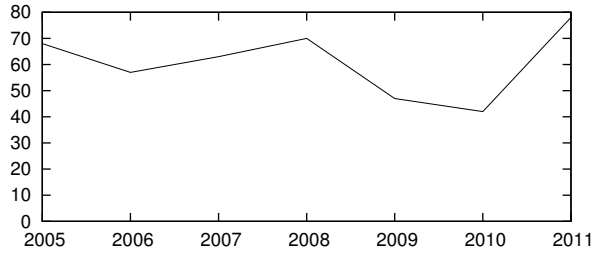

Fig. 2. The percentage of accepted RfA in each year

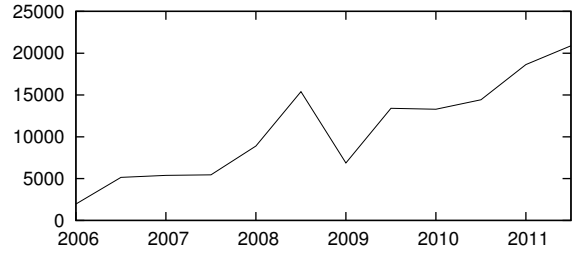

Fig. 3. The average number of edits made by user at the moment of receiving administrative privileges

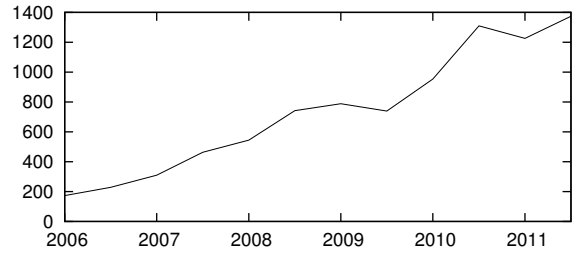

Fig. 4. The average number of days since creating an account at the moment of receiving administrative privileges

privileges to a candidate, in 110 of those, the candidates were rejected and in about 40 votings, the candidates resigned before the end of the voting and about 30 votings were cancelled (due to statutory requirements or lack of acceptance of the nomination by the candidate). About 38 administrators were chosen before the introduction of RfA procedures in March 2005. The data on RfAs do not sum up for several reasons. Among them are: verification votings and losing privileges by administrator either by giving them up or being revoked by the Arbitration Committee.

In the current version, the procedure states that a candidate for an administrator must have an account for at least three months and at least one thousand not deleted edits. In order to participate in the voting, user must have an account for at least 2 weeks and at least 500 article edits. Voting begins at the moment when the candidates confirm that they are willing to take the administrator position, as users can apply by themselves or be nominated by other users. In order to receive administrator privileges, the candidate need to receive at least 20 votes “for” and it must constitute of at least 80% of the sum of the votes “for” and “against”. If the candidate does not receive the required number of positive votes or do not meet the formal requirements, he or she can apply again after at least 60 days since the end of last voting. A similar rule applies to the administrators who resigned from their position but would like to receive the privileges again.

Figure 1 is presenting the number of votings in each year; the peak can be observed in the year 2006, when that number reaches 95, while a year before it reached only 34. One year after the peak, the number of votings dropped to 60. With the exception of the years 2006-2007, the number of votings never exceeded 38. In the years 2010-2011 that number declined below 34. The number of RfAs between year 2006 and 2011 decreased by nearly three quarters (form 95 to 26).

The percentage of the accepted nominations in each year (see Figure 2) can be divided into three periods. The first one consists of the years 2005-2008, when the percentage of the accepted candidates ranged from 57 to 70. The second period are the years 2009-2010, with the percentage below 50 (47% and 42% respectively). Between the years 2008 and 2010 the percentage of the positive RfAs fell by almost a half (from 70% to 42%). The third period, which accounts for the year 2011, is characterized by the relatively high number of positive RfAs. However, it should be noted that the number of the votings performed at that time was significantly lower than in the previous years.

The next study, related to the experience of candidates prior to granting them administrator privileges, has been conducted on 97 users, who recently received them. In case of those elected before, the gathering of data was impossible because of gaps in the logs of Polish Wikipedia.

One of the factors causing the most discussion during the voting is the number of edits performed by the candidate. RfA rules contain the following sentence: “Users who want to candidate for adminship (...) must have at least 1000 not deleted edits”. Often, however, this number is considered by the voters to be too small. Basing on the analysis of the number of edits, it can be seen (Figure 3) that the minimum falls on the first half of 2006 with an average of 1,957 edits. This value then grows up to 2011 when it slightly exceeds 20,000 edits. This indicates that year by year, candidates needed to have greater experience in order to be accepted as administrators. The difference between the level of experience required by the regulations and the level widely accepted has been increasing as well. A similar phenomenon can be observed in the German Wikipedia, where—according to the voters—in the second half of 2010, candidates were accepted only if they had over 10,000 edits.

Another factor that stirs up emotions at the time of voting is the seniority (understood as time since the first not deleted edit) on Wikipedia. The terms of voting set the following requirements: “Users who want to candidate for adminship (...) must have at least 1000 not deleted edits, the first of which took place at least 3 months before the date of candidacy proposition”. The seniority (in days) of candidates, before the date of registration and acquiring the administrator rights, had been analysed (see figure 4). This, however, is not exactly the same value as the required by the regulations. The measured seniority in the first half of the year 2006 was 173 days. This value has been gradually increasing: from 463 days in the second half of 2007, to 788 days in the first half of 2009, with a slight decline in the second half of 2009 (739 days). In the

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

second half of 2010 the value reached 1310 days. This result, however, may be unreliable due to the fact that during that period only two votings took place. In the second half of 2011 the measured value reached 1374 days. The overall analysis of the chart shows that in the year 2006 candidates had less than a year of seniority, however, since mid-2008 the seniority is at least two years. The last two candidates who had less than one year of experience were selected in February 2009 and November 2008.

Both Figure 1 and Figure 2 show the downward trend in the total number of appointed administrators between 2006 and 2011. This decrease gives reason for serious concern as the amount of administrative work on Wikipedia is constantly growing. This phenomenon may have several possible explanations. The first explanation is the declining number of candidates who accept their nominations for administrators (that would explain the decreasing number of RfA votings), but the confirmation of this hypothesis is beyond the scope of this paper. Nevertheless, related works have shown that in recent years Wikipedia has experienced a downtrend in the amount of user contributions, which reflects the general decline in motivation [10].

The second explanation states that the number of positive nominations decreases due to the changing criteria for selecting and accepting candidates. Those criteria can vary in many ways; however, our research shows that they are connected to the candidate's experience. This experience can be initially estimated on the basis of the edits performed, but the more accurate measurement (presented in [4]) represents the number of article edits in a specific category.

The more damaging prospect is the fact that the administrator community is chosen on the basis acquaintance between current administrators and candidates. The next section discusses, if that is the case.

IV. Acquaintance in the administrators society

A. Data description

Data and multidimensional behavioural social network used for this paper were gathered, aggregated and made available by the team led by dr. Adam Wierzbicki. Methodology, data and networks are described in greater detail in [3]. Examined period encompasses the years 2005-2011.

Basically, the network consists of four dimensions:

- Co-edits,
- Reverts,
- Discussion,
- Topics.

Weights in co-edits dimension are based on number of words written by one author next to the text written by some other one in the text of articles. The authorship information for a particular fragment of text was obtained by analysing its first occurrence in the whole edit history of examined page.

Edge strength in reverts dimension is based on the number of edits made by one author and reverted by other. It was obtained by searching identical revisions before the examined one. If it was found, each pair of examined revision and revisions after the other identical one was used to calculate number of reverts.

Similar to co-edits, edge strengths in discussion network were stated as number of words written by one author next to text created by other one. But in this dimension, the talk pages were considered.

The last dimension, topics, was a little different to other ones. It was a bipartite graph connecting authors with categories in which they have edited at least one article. The edge weight was exactly the number of article edits made by given author in the particular category.

One of the most important observations made in [3], is that discussion network can be interpreted as social relation of acquaintance. Jankowski-Lorek et al. conducted [3] another research, a survey among Polish Wikipedia users. However, interpretations of other dimensions have not been confirmed.

The data contained two more graphs: positive votes network and negative ones. If, during RfA procedure, user has cast positive vote for candidate, then an edge in the positive votes network has been created. Its weight was equal to number of positive votes cast by the user for the candidate. Weights of more than one were possible only if the user was a candidate more than once. Network of negative votes has been created in an analogous manner, but taking the negative votes instead.

In [2], [3], each dimension has been intersected with positive and negative votes networks, in order to examine correlation between social network dimensions and RfA votings. Both graphs were analysed separately and features distinguishing them have been found.

Research presented in this paper studies only on the discussion dimension. The reason for such decision is that discussion network can be interpreted as a real relation—acquaintance. For each year, graph of discussion network has been intersected with positive and negative votes networks. Some authors suggest using one, signed network [11], especially, when there is a strong correlation between both networks as shown in [12]. Two separate graphs were used for two reasons:

- To maintain consistency with analysis presented in [2], [3],
- To separately check positive and negative impact on RfA procedure of acquaintance relation.

B. Base statistics

In order to compare graphs resulting from intersecting the discussion dimension graph and votes nets for each year, base statistics were obtained for edges' weights. The used measures were: minimum, maximum, arithmetic mean, median, first and third quartile.

For each graph: discussion network, discussion intersected with positive and discussion intersected with negative votes in each year, empirical distribution functions were calculated. Distribution graphs for selected years are presented in Figure 5. Values of x-axis are logarithms of edge strengths. Since 2007, the distribution of data is analogous to that described in the article. Both arithmetic mean and median are significantly

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

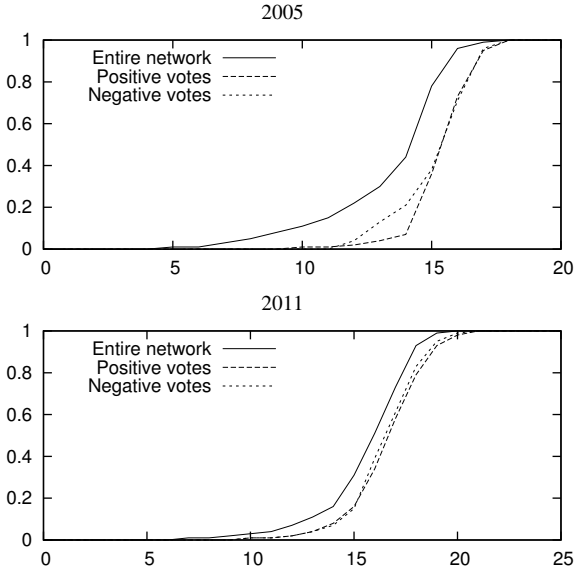

Fig. 5. Distributions of discussion network edge strengths

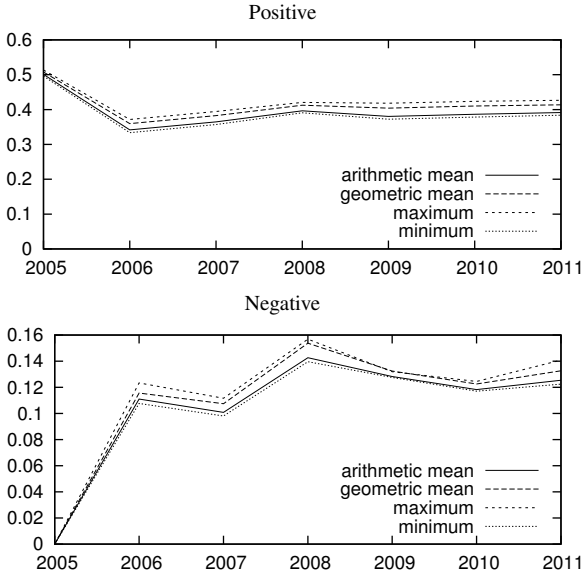

Fig. 6. Clustering coefficients

higher for positive votes. As a result, it can be concluded that the Wikipedia user community is developing steadily.

The stability of the development of the Wikipedia administrator community is also reflected by the empirical cumulative distribution charts. The shapes of the curves are similar, so it can be concluded that the probability distributions describing different parts of the Wikipedia user community originate in the same distribution family. This means that the behaviour of the voters is not subject to sudden changes, but at most it undergoes a calm evolution.

C. Clustering coefficients

Clustering coefficient is a measure of degree to which nodes in a graph tend to be clustered together. The global version, which is used in this article, was designed to give an overall indication of the clustering in the network. Basically, for undirected graphs, it is a ratio between number of closed triplets (three nodes connected by two links) to number of all triplets (three nodes connected by either two or three links).

For directed, weighted graphs a generalization was proposed, it is described in detail in [13]. Opsahl et al. proposed four measures to calculate triplet value:

- Minimum of edges' weights (mi),
- Maximum of edges' weights (ma),
- Arithmetic mean of edges' weights (am), and
- Geometric mean of edges' weights (gm).

The intuition is as follows: the minimum version is used to find the weakest group in graph, the maximum to find the strongest. Both means give an indication of the strength of ordinary clusters. Opsahl et al. [13] and [14] also created tnet library [15] for R software [16].

For each year from 2005 to 2011, the clustering coefficients were obtained for intersections of acquaintance networks with graphs of positive and negative votes. Those coefficients are presented in Figure 6. There are four values (calculated for each of the measures mentioned before) for both graphs.

A few facts can be observed. The first is that there are no very weak or strong groups in Polish Wikipedia society. There is no "elite", which governs RfA procedure or has taken over the administrator society and has power to rule Polish Wikipedia.

The second fact is that clustering coefficients are relatively low and their growth rate is low and negligible. We argue, that decrease in successful administrator elections is not a result of a building up acquaintance relation. Voters do not cast positive votes for their acquaintances or cast negative votes for strangers. The anomaly in year 2005, that clustering coefficients have abnormal values, is most likely caused by the fact, that data for year 2005 were not complete.

V. Conclusions and Future Work

This paper presented the analysis of the development dynamics of the community of administrators of Polish Wikipedia. We have used multidimensional behavioural social networks as a tool to model relationships between wikipedians. The aforementioned analysis included examination of the community in each year from 2005 to 2011 as well as the analysis of the social network corresponding to the final state of the community. The analysis was based on the data from public Wikipedia data dumps.

The fundamental question which we sought the answer to was: "Is the administrator community of the Polish Wikipedia closing up?" It turns out that the answer is not straightforward and it depends on what aspects of the problem one put the greater emphasis, or how to define the "closing up" society.

The conducted analysis of the social network allows us to draw conclusions about the impact of the social system on the nominations of the new administrators. The results of this analysis clearly show that this phenomenon does not exist in the Polish Wikipedia. This is one of the arguments

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

for the statement that the community of the administrators is not closing up. The administrator community is open to new members in the same way as it was in the beginning of the Polish-language Wikipedia.

However, the pace of growth of the administrator community is lower than it could be expected in case of a young and dynamically growing society. In the early years of development the number of votings was much higher than in the recent years and the number of new administrator appointments strongly declined. That could indicate, however, that the community is closing up after all.

Slower pace of growth and acceptance of new members can be caused by various factors. One such factor may be higher entrance requirements for candidates. Both administrators and regular editors of Wikipedia continue to develop and gain experience in new areas. At the same time, the history of their activity is freely available. For that reason, new users may have trouble with showing equally high achievements and contribution to Wikipedia development. This can be interpreted as closing up of the community by making prohibitive requirements for the new candidates, or as a kind of professionalization aiming to increase the substantive level of the Polish Wikipedia.

Our conclusion is that it cannot be claimed with certainty that the Polish Wikipedia community is closing up. We believe that the increase in the requirements of the current administrator community and users entitled to speak during RfA process toward administrator candidates stems from the community's desire to raise the quality and ensure maximum involvement of all the administrators in the development of Wikipedia.

The results presented in this study describe the community of Polish Wikipedia administrators only partially. Further research should focus on the detection of new relations between the users and social networks associated with them. It is important to find methods that will allow the development of community to be automatically analysed on the basis of widely available data. The multidimensional behavioural social networks seem to be an ideal tool for this purpose. Richer description of the community could help predict the direction of its development, which may result in the early identification of threats. This will give the opportunity to counteract those threats and ensure the correct development of Wikipedia.

User community of Polish Wikipedia—in contrast to other language versions—is relatively little known and researched, although, it is an ideal subject for researchers dealing with social informatics. It can be an interesting subject for two types of research: new research, previously not conducted on such a social group, and repeated research, taken from a different version of Wikipedia and performed on the Polish version in order to compare the results and draw conclusions on the development of the latter in comparison to other versions.

Tools used to create multidimensional behavioural social networks for Polish Wikipedia were unable to create such graph for larger instances, e.g., English one. In order to conduct comparative research, scalability problems should be addressed. There is also possibility, that more scalable algorithms can be made on-line. This can allow development of on-line recommendation algorithm for RfA votings.

ACKNOWLEDGEMENT

Authors of this paper thank the team led by dr. Adam Wierzbicki from Polish-Japanese Institute of Information Technology in Warsaw for sharing processed Polish Wikipedia data and multidimensional behavioural social network of Wikipedia users.

The second author was supported by the Polish National Science Centre grant 2011/01/B/ST6/03867.

REFERENCES

[1] A. Kittur, B. A. Pendleton, B. Suh, and T. Mytkowicz, "Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie," *World Wide Web*, vol. 1, no. 2, p. 19, 2007.

[2] P. Turek, J. Spychała, A. Wierzbicki, and P. Gackowski, "Social mechanism of granting trust basing on polish wikipedia requests for adminship," in *SocInfo*, ser. Lecture Notes in Computer Science, A. Datta, S. W. Shulman, B. Zheng, S.-D. Lin, A. Sun, and E.-P. Lim, Eds., vol. 6984. Springer, 2011, pp. 212–225.

[3] M. Jankowski-Lorek, L. Ostrowski, P. Turek, and A. Wierzbicki, "Modeling wikipedia admin elections using multidimensional behavioral social networks," *Social Network Analysis and Mining*, pp. 1–15, 2013. [Online]. Available: http://dx.doi.org/10.1007/s13278-012-0092-6 [Accessed: 15.05.2013]

[4] M. Burke and R. Kraut, "Taking up the mop: identifying future wikipedia administrators," in *CHI Extended Abstracts*, M. Czerwinski, A. M. Lund, and D. S. Tan, Eds. ACM, 2008, pp. 3441–3446.

[5] (2013, 05) Wikipedia: Guide to requests for adminship. [Online]. Available: http://en.wikipedia.org/wiki/Wikipedia:Guide_to_requests_for_adminship [Accessed: 15.05.2013]

[6] A. Kittur, B. Suh, B. A. Pendleton, and E. H. Chi, "He says, she says: conflict and coordination in wikipedia," in *CHI*, M. B. Rosson and D. J. Gilmore, Eds. ACM, 2007, pp. 453–462.

[7] J. Leskovec, D. P. Huttenlocher, and J. M. Kleinberg, "Governance in social media: A case study of the wikipedia promotion process," in *ICWSM*, W. W. Cohen and S. Gosling, Eds. The AAAI Press, 2010.

[8] A. Anderson, D. P. Huttenlocher, J. M. Kleinberg, and J. Leskovec, "Effects of user similarity in social media," in *WSDM*, E. Adar, J. Teevan, E. Agichtein, and Y. Maarek, Eds. ACM, 2012, pp. 703–712.

[9] H. S. Rad, A. Makazhanov, D. Rafiei, and D. Barbosa, "Leveraging editor collaboration patterns in wikipedia," in *HT*, E. V. Munson and M. Strohmaier, Eds. ACM, 2012, pp. 13–22.

[10] B. Suh, G. Convertino, E. H. Chi, and P. Pirolli, "The singularity is not near: slowing growth of wikipedia," in *Int. Sym. Wikis*, D. Riehle and A. Bruckman, Eds. ACM, 2009.

[11] S. Maniu, B. Cautis, and T. Abdessalem, "Building a signed network from interactions in wikipedia," in *DBSocial*, D. Barbosa, G. Miklau, and C. Yu, Eds. ACM, 2011, pp. 19–24.

[12] J. Leskovec, D. P. Huttenlocher, and J. M. Kleinberg, "Signed networks in social media," in *CHI*, E. D. Mynatt, D. Schoner, G. Fitzpatrick, S. E. Hudson, W. K. Edwards, and T. Rodden, Eds. ACM, 2010, pp. 1361–1370.

[13] T. Opsahl and P. Panzarasa, "Clustering in weighted networks," *Social Networks*, vol. 31, no. 2, pp. 155–163, May 2009. [Online]. Available: http://dx.doi.org/10.1016/j.socnet.2009.02.002 [Accessed: 15.05.2013]

[14] T. Opsahl, *Structure and Evolution of Weighted Networks*. University of London (Queen Mary College), London, UK, 2009. [Online]. Available: http://toreopsahl.com/publications/thesis/ [Accessed: 15.05.2013]

[15] (2013, 05) tnet: Analysis of weighted, two mode, and longitudinal networks. [Online]. Available: http://toreopsahl.com/tnet/ [Accessed: 15.05.2013]

[16] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org [Accessed: 15.05.2013]

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Developing an Online Academic Community of Practice

KatyaToneva
Information Services
St Mary's University College
London, UK
Katya147@yahoo.com

Iain Cross
School of Management and Social Science
St Mary's University College
London, UK
Iain.Cross@smuc.ac.uk

Martin Scarrott
Information Services
St Mary's University College
London, UK
Martin.Scarrott@smuc.ac.uk

***Abstract*—Constructivist approaches to learning suggest that learning within groups is more effective than individual learning. Communities of practice are an example of how individuals can unite around a common theme and collaboratively work together in order exploit the collective knowledge of the group members. Higher education professionals may find particular benefits of working within such communities due to the diverse nature of the profession and the wide knowledge and experience that can be found among their peers. At St Mary's University College, London, U.K., a strong tradition of excellence in teaching and learning guides institutional strategies to enhance the learning experience for students. An online academic community of practice is being developed in order to heighten the dissemination of good practice among colleagues and improve the sense of community among staff. This paper describes the stages of the development process. It is argued that the developing community of practice is already improving collaboration and dissemination of good practice among St Mary's staff. The key benefits are accessibility, interactivity and feeling of ownership. As the community evolves and matures, it is expected that knowledge sharing and collaboration will continue to increase, if the barriers that have been identified can be overcome successfully.**

Keywords–online Academic Community of Practice (ACoP); Microsoft SharePoint; Moodle; social constructivism; interaction; collaboration.

I. INTRODUCTION

Contemporary education often draws upon constructivist approaches, which suggest that learning is more effective when knowledge is constructed by individuals, rather than by the ordering and memorisation of facts [1, 2]. Constructivism originates in the work of Jean Piaget and Lev Vygotsky [3]. Whilst both emphasised the role of the individual in constructing models of understanding, Vygotsky argued that individual learning was mediated by social interaction and cultural setting. Therefore, knowledge is not isolated from the environment in which it is constructed but is instead copied from, or directly reflects, existing models of socially-produced understanding [3].

Vygotsky formulated the concept of the 'zone of proximal development': that is, the distinction between what an individual can achieve by independent thought and what is possible 'through problem solving…in collaboration with more capable peers' [4]. Other workers (including Peter Woods and Jermone Bruner) have suggested that the gradual withdrawal of this collaborative support leads to the development of 'scaffolding' [5, 6]. Scaffolding forms an architecture of understanding, providing a structure in which the learner can develop their knowledge whilst reducing the level of support required from 'more capable' peers. Learning within groups or communities in which the interactions between individuals are applied to resolve practical or real-world problems has gained widespread recognition among contemporary educationalists.

Communities may take different forms based on the aim of the community and the way that individuals interact. Communities can take different forms: knowledge-based communities, learning communities and collaborative communities of practice [7]. Knowledge-based communities are typically focussed on resolving problems based on the group's collective knowledge, and undergo a process of re-adaption in response to other group members. The aim of a learning community is to further understanding: knowledge acquisition is equally shared among group members and the sum outcome of group's endeavours is greater than that of any individual. A third type of community, communities of practice, have their origins in the model of professional development through apprenticeships (the original work of Etienne Wegner [8, 9] and Jean Lave).

Communities of practice (CoPs) focus on the development of professional practices, where 'like-minded' individuals are united by 'joint enterprise, mutual engagement and shared repertoire' [7]. CoPs can therefore be defined as 'groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly' [8].

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Numerous examples of CoPs exist. MirandaNet [10] provides an online community for professionals associated with information and communications technology (ICT) education. The Education Communities portal [11] provides access to numerous CoPs with a collective focus on education and pedagogic research, and offers the ability to easily establish new CoPs. One growing focus for the use of CoPs is in developing and strengthening the links between education and the business and enterprise communities through providing an online arena in which ideas can be exchanged and collaborations developed (i.e., [12]).

CoPs may be particularly beneficial for professional development as, for example, the nature of the 'problem' of interest, or focus of the community, is often difficult to define, and identifying exactly what it is may require extensive negotiation amongst individuals [7] potentially deepening their understanding of the 'problem'. Members of a CoP are strongly united by their shared interest in 'problems', which may be derived from the shared institutional identity and process, or from a common professional practice context [7, 9]. These principles may be particularly relevant to professional development in higher education. It is recognised that there is continued tension within the definition of the Higher Education Professional (HEP).

Pressure to be a skilled teacher and researcher has led to the identity of HEPs being increasingly diverse, complex and open to constant re-negotiation [13, 14, 15]. The teaching aspect of the dual professionalism of HEPs can be overlooked if HEPs feel defined by their subject specialism. Many HEPs have higher degrees and research degrees that are subject-specific rather than focussed on teaching.

Developing as an HE teacher is often seen as a 'bolt-on' to subject-specialisms, evidenced in the U.K. by newly appointed HEPs undertaking teaching qualifications, such as postgraduate certificates and achieving fellowships of the Higher Education Academy (an independent organisation in the U.K. dedicated to supporting professional recognition for HEPs and enhancing teaching and learning [16]). It has been suggested that in order for HE teaching to become a 'valid' profession, HEPs must engage in scholarship and research comparable to that dedicated to their subject specialism [17]. CoPs provide a constructivist environment in which HEPs can develop and share teaching expertise.

Drawing on this theoretical framework, St Mary's University College, U.K., is establishing an academic community of practice (ACoP) in order to support academic staff in disseminating good practice and developing the use of e-learning (Section 2). In particular, the ACoP will be used to support a parallel project investigating the use of audio feedback for students (Section 3). Section 4 outlines the technical platform that supports the newly developed online ACoP. Section 5 provides a framework for the development of the ACoP and an analysis of how potential difficulties and hindrances to the success of the ACoP project may be overcome. The last section outlines our intentions for the future development of the ACoP.

II. CONTEXT – ST MARY'S UNIVERSITY COLLEGE

St Mary's is a relatively small higher education institution with around 5,000 students and 150 academic staff. It has a Catholic foundation and although initial teacher education is at the heart of its mission it is engaged with a broad range of disciplines from English and Media Arts to Management Studies and Applied Physics. Learning technologies have been used for a number of years and online learning began to develop around 2003. At the current time almost all of the University College's programmes of study use online learning technology in some form although much of the practice is limited to content dissemination rather than using it to support more contemporary pedagogical approaches.

St Mary's has a Learning, Teaching and Assessment Strategy and the use of technologies to improve students' learning opportunities is embedded into that strategy. The strategy is enabled (in part) by a Learning Resources and Technology Strategy which focuses on the provision of various tools and supporting academic staff (and students) in their use. There is no "top down" managed approach for technology enhanced teaching and learning and so practice has evolved over a number of years. A number of staff clearly see ways in which technology can enhance students' learning and such staff have been leading the way in trying to encourage their colleagues to adopt new and innovative approaches. Recently, St Mary's has recognised the need for setting up institutional requirements for using the internal Virtual Learning Environment (VLE) and all academic staff will be expected to follow these guidelines from September 2013.

One of the barriers to improving pedagogical practice has been a lack of sharing. Staff work in academic schools and learning is delivered in the form of programmes (e.g., Sports Science or Geography). Staff tend to identify strongly with the programme of study with which they are connected but are perceived to have limited contact with staff in other academic disciplines or support roles.

A number of attempts have been made to increase cooperation and collaboration between staff including the establishment of a formal teaching qualification for staff, promotion of Higher Education Academy fellowships and regular teaching and learning 'away days'.

In relation to the use of learning technologies, a network of e-champions has been established. The e-champions have a key role not only in supporting their colleagues pedagogically and technically in the use of the available tools, but also in disseminating their own practice and experience. These efforts will now be supplemented by the establishment of the ACoP, coordinated by the University College's e-Learning Advisor.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

The e-Learning Advisor role is a supporting role designed to help motivate staff to embrace the use of technology to enhance learning and provide appropriate support. This role is located in the Information Services Department which is responsible for providing and supporting a broad range of learning support for students and staff.

The technical infrastructure is supported by the IT Department although a number of the online learning tools used are outsourced to third parties rather than being provided internally.

The ACoP will enhance a wide range of online interactions and will promote collaboration, ubiquitous access, increased flexibility, choice and support. Due to the rapid pace at which e-learning is developing within higher education, it is envisaged that the ACoP will provide a flexible working framework which will be adapted as necessary to ensure that the university college can respond pro-actively to such a changing environment.

The ACoP supports the aims of the Teaching, Learning and Assessment Strategy which outlines the specific objective for academic staff "to develop and implement an effective framework for sharing good practice". In the context of providing increased flexibility, the Information Services Department and the e-Learning Committee highlighted the implementation of the online ACoP as a high priority in its staff development strategy.

III. CONTEXT – JISC TRANSFORMATION PROGRAMME

St Mary's has received funding from Jisc (a U.K. charity which works throughout higher and further education to support and champion the use of digital technologies [18]) to develop the use of audio technologies for delivering formative feedback to students. This is part of Jisc's wider funding stream, the Transformations Programme. This programme is funding a number of higher education institutions with the aim of supporting 'large-scale organisational change' [19]. The audio feedback project at St Mary's focusses on the Enhanced Student Experience strand of the programme [20].

The work being undertaken at St Mary's involves a group of academics from several diverse programmes using audio feedback. The e-learning team has supported academics in using of a number of alternative approaches to delivering audio feedback, such as delivering collective feedback through one large audio file or uploading individual commentaries on pieces of work. The team also have provided advice on the technical dimensions to the project.

The ACoP will be used to disseminate examples of good practice to academics throughout St Mary's in order to ensure that the project meets Jisc's requirement for 'organisation-wide impact' [19]. A rich and diverse range of feedback has been gathered by the e-learning team (including audio and video interviews with staff and students, and quantitative and qualitative questionnaire data) regarding the experiences of students and academics in using audio feedback.

To support academics who are interested in using audio feedback, these materials will be developed into case studies and disseminated on the ACoP. This will enable the whole academic community at St Mary's to engage in the project outcomes and to seek advice, examples and support from those who have experience of using audio feedback. Ultimately, this will enhance the student experience at St Mary's and increase student satisfaction and engagement with feedback activities.

IV. THE TECHNICAL PLATFORM

St Mary's University College uses Microsoft SharePoint 2010 to provide internal staff and student intranet services (branded as StaffNet website for service information and collaboration for staff, and as SIMMSpace website for advice and service information for current students).

The StaffNet website is used for publishing of web-based news, service information and internal documentation for staff members, for collaboration websites for staff groups, for personal storage areas and profile pages for individual staff. StaffNet is the web area that is being used to develop the online ACoP.

SharePoint has delivered significant benefits to the institution by providing a reliable and stable platform on which to host internal communication and information services. In addition, SharePoint is serving as the gateway to other online services such as the Moodle VLE (branded simmsCAPital). The simmsCAPital VLE (integrated with the TurnitIn plagiarism detection service) is used for developing the courses of study, for online submission of coursework and providing feedback to students. Feedback is delivered through both traditional formats (e.g., written comments and numerical grades) and e-learning can be instrumental for establishing CoPs.

This technical platform offers a significant opportunity for the development of a CoP. Technology can facilitate collaborative and community-driven learning, for example by overcoming geographic barriers, facilitating sharing of resources and removing the need for scheduled times for face-to-face meetings [21, 22].

CoPs can be supported by blogs, wikis and other resource-sharing facilities (such as Google Docs and Microsoft SharePoint) which enable collaborative learning. These technologies tend to offer a relatively informal approach to structuring ideas and interactions which is well aligned with the social constructivism inherent within functioning CoPs [7, 8, 9].

This project will establish an ACoP implemented within Microsoft SharePoint that supports professional development and the sharing of good practice in the context of teaching in higher education.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

V. RATIONALE AND DEVELOPMENT FRAMEWORK

The rationale for the design of the St Mary's ACoP was informed by conducting interviews with the e-champions. They identified the following themes as important aspects in ensuring that the online ACoP would be successful at meeting the aim of sharing best practice to enhance learning:

- Sharing of information to promote best practice;
- Avoiding duplication and effort (e.g., in the production of guidelines);
- Disseminating high-quality information;
- Facilitating the retrieval of appropriate information at the point of need;
- Supporting the teaching and practice within the University College programmes;
- Encouraging collaborative working;
- Improving access to knowledge and information;
- Bringing different streams of information together;
- Promoting communities of practice by bringing together disparate groups with a shared interest (tutors, administrators, senior managers);
- Providing on-going support, feedback and shared knowledge for the ACoP stakeholders;
- Raising the profile of academic groups and identifying ways in which they can contribute;
- A potential for forming links with other organisations to provide a coordinated approach to information-provision.

These themes are therefore considered to be the desired 'core characteristics' of the St Mary's ACoP. In order to embed these characteristics into the design of the ACoP, the following sequence of stages of development is being utilised: a) the potential stage; b) the building stage; c) the engaged stage; d) the active stage and e) the adaptive stage [23]. The development of St Mary's ACoP is evaluated in the following sections based on the descriptions given in [20], based on feedback collected from members of St Mary's staff.

a) The potential stage: This involves facilitating connections among individuals who are potential members of the community. Communities at this stage are a group of individuals with some occasional interaction concerning a shared interest [24].

During this stage a conceptual framework of the ACoP integrating the relevant pedagogical theories and technological solutions was developed, in order to respond to the requirement for a more systematic way of interacting.

The theoretical and practical appreciation of management of change (for example, discussing suitable approaches for involving e-champions and senior managers in the ACoP development) was also considered. Web-spaces and meetings, such as staff development workshops, have been used to promote the ACoP to staff throughout the University College and to engage their interest in the project. At this stage, face-to-face meetings and discussions have been held with academic leaders, e-learning strategy developers and other academic staff and tutors to identify existing expertise and the needs of staff.

b) The building stage: During this stage, core members of CoPs develop a shared understanding of the purpose of the community and the ways in which it will function. The community begins to develop the necessary structures to allow expertise to be shared within the community [23].

It is also the stage at which the community is acknowledged to exist and potential members of the community are reached out to. Since this is considered a 'fragile' part of the community development [24] specific action was taken to develop the St Mary's online ACoP to encourage contribution and interest from members.

A case study template was established within the ACoP in order to support a consistent way for presenting show cases by academic staff. Standard activities, such as writing collaborative documents and articles were introduced.

This has resulted in the production of a structured ACoP (Figure 1) with a number of features to facilitate community interactions and the sharing of good practice, including personal webpages, interactive feedback features and details of other projects within the University College (e.g., the Jisc Transformation Project).

Figure 1. The front page of the ACoP on SharePoint

c) The engaged stage: A community at this stage has gained enough momentum to survive. The community has a common purpose as it increases in size. Community members continue to build trust, loyalty and commitment and provide outreach to new members. Members also collaborate to share their knowledge [23].

This was evidence in the St Mary's ACoP as teaching and learning materials, links to social and digital media and embedded objects such as videos and audio files have been posted online.

In order to continue expanding the size of the ACoP, training (including online tutorials and face-to-face sessions) is being provided by the e-learning advisor to enable staff to make contributions which showcases their best practice in teaching and learning.

d) The active stage is when communities continue to expand their membership. At this point it is clear to the community what it is able to contribute to the wider organisation. Members of the group collaborate to exploit their combined knowledge base. The community is publicised throughout the organisation to highlight how it can benefits that organisation. Members may also begin to work with other external groups [23].

The St Mary's ACoP has not fully reached this stage. It is expected that the full transition into this stage will be evident from staff being actively and sustainably engaged with the ACoP and developing the use of a wider range of digital media to do so (e.g., blogs, wikis, discussion forums and collaborative working). Some evidence that the St Mary's ACoP is beginning to reach this stage can be seen in the use of the ACoP to write collaborative articles.

e) The adaptive stage: At this stage a community is able to respond to changing external boundaries. It does this by continually renewing and evaluating its purpose and functionality to effectively use knowledge and resources. The community may move towards setting the agenda of its collective shared expertise and may demonstrate influence beyond its immediate area of expertise [23].

This is an advanced stage of community formation and one that the St Mary's ACoP is expected to achieve in the future through links with national and international CoPs, demonstrating a leading role in developing teaching and learning strategies at the University College and providing a focus for professional development at St Mary's.

VI. POTENTIAL BENEFITS, BARRIERS AND CRITICAL SUCCESS FACTORS OF THE ACoP.

Our observations and the collected feedback from staff show that there are potential benefits to be derived from sharing and learning within the online ACoP. There is a sense of connectedness and a deepening of knowledge to be derived from the interactions and peer support via the ACoP.

This will also help instil a sense of community among St Mary's staff. Peer support is particularly valuable in the context of professional development, as members of the community who are new to HE are able to draw on the experiences of other staff.

One significant potential benefit of the St Mary's ACoP is in the dissemination of other projects running across the institution, such as the Jisc audio feedback project.

The ACoP is already facilitating engagement from across St Mary's as a range of resources showcasing innovative and successful use of audio feedback has been made available via the ACoP.

However, academic staff who have been involved in the ACoP commented on some potential barriers to these benefits being realised. For example, some members of the community may have basic technical literacy skills and may therefore feel intimidated or not confident in participating in the ACoP.

Another potential barrier to the success of the ACoP is the commitment of individuals to participate. Whilst there has been a good initial response to the ACoP and individuals have begun to demonstrate a willingness to engage in with the community, as the ACoP matures there is the risk that the enthusiasm to participate will decline. Engagement may only extend as far as viewing the material rather than contributing expertise (i.e. being a 'read only' participant).

One mechanism to help develop a sustainable community is to ensure that there is leadership [25]. In the case of a CoP, this may be a moderator or facilitator.

A facilitator has been established for St Mary's ACoP, and is encouraging tutors to engage with this newly developed online environment for professional development. They will be working to promote the ACoP and encourage its development through activities such as marketing on the StaffNet intranet, embedding the ACoP into professional development and promoting the ACoP and its advantages through the relevant College Committees.

However, it is acknowledged that frameworks for developing online CoP in other institutions may vary in terms of technological and pedagogical approaches, and may focus on different activities.

For example, one of the authors has had experience in developing an online CoP for Middlesex University, London, where the technological framework was based on Google Apps and the main collaborative activities were around project and research work [26].

Critical success factors (CSFs) for the online ACoP have been identified. These include usability of the available technology, appreciation of the cross-institutional dimension of the ACoP, a sense of belonging and a common sense of purpose among members and shared understandings and ownership.

The potential benefits, potential barriers and critical success factors are shown in Table I (a new adaptation from [26]). Our initial exploration of these potentials and success factors may help to both address the difficulties of establishing an ACoP at St Mary's and also provide a basic framework for other HE institutions developing their own ACoPs.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

TABLE I. POTENTIAL BENEFITS AND BARRIERS, AND CRITICAL SUCCESS FACTORS (CSFs) OF THE ACoP

Potential Benefits	**Potential Barriers**	**CSFs**
Enhanced learning environment	Perpetuation vs. variety and flexibility	Good use of interactive tools based on Microsoft SharePoint and Moodle (demonstrated by the online activities and engagement by staff and students, analysed in review reports).
Knowledge sharing	Culture of independence	Institutional promotion of the online ACoP as a communication media (demonstrated via workshops and other institutional events).
Building up knowledge and expertise	Maintaining the interaction and information flow	Common values, shared understanding, varied communications (demonstrated by the online activities).
Feeling of connection	Read-only participants	Sense of purpose, sense of belonging, sense of ownership (demonstrated by active online communications and sharing).

To ensure the longevity and sustainability of the online ACoP, engaging and collaborating through the ACoP needs to become part of standard practice rather than seen as a novel standalone project. In an effort to achieve this, there should be clear institutional aims and time frames.

For example, the ACoP could become a primary repository for material relating to workshops, induction events and guest speaker seminars. We will work towards this by involving representatives from each department to collect or design relevant resources and publish these at the online ACoP.

The longevity of the ACoP can also be ensured by effectively incorporating future developments in social networking technologies. Allowing the ACoP to adapt to these will help ensure its continued relevance to new members and facilitate new forms of interaction and collaboration.

Links to other professional online communities of practice should also be established in future to exploit of other sources of expertise and enhance the depth and breadth of discussion within St Mary's ACoP.

VII. CONCLUSION AND FUTURE WORK

The initial experience of developing an online ACoP at St Mary's University College has been successful. The process has demonstrated that Microsoft SharePoint can be used effectively to establish an online ACoP. The St Mary's ACoP has been designed to support the development of collaborative resources, to gain access to e-resources and disseminate good academic practice. It is anticipated that the online ACoP will help to continuously enhance the professional practices of St. Mary's staff and facilitate collaboration with other stakeholders as it develops in the future. The use of the ACoP will be promoted throughout the institution and become increasingly embedded within the culture of academic best practice at St Mary's.

The initial assessment of the use of the ACoP suggests that it is being successfully used as a focal point for the dissemination of good practice in teaching and learning. This includes communicating the outcomes of other projects that are enhancing teaching and learning at St Mary's. It is evident that there are a series of benefits to the staff and institution which arise through the use of the ACoP, including accessibility, flexibility, interactivity and the sense of ownership and community that the users have begun to develop. This provides the rationale for on-going development and promotion of the ACoP.

The ACoP has developed through a series of stages, each associated with specific activities. This sequence and set of activities may serve as a framework for other institutions who wish to establish an ACoP, although it is recognised that the potential benefits and barriers to success may vary in other organisations depending on the institutional context and aims of individual ACoPs. The ACoP is expected to expand in size, develop a wider range of multimedia resources and become central to professional development at St Mary's.

Having established an ACoP, future research will focus on the long term development of the St Mary's ACoP. A three-year longitudinal study is being initiated which will provide a valuable insight into the development of online CoPs. In particular, the relationship between the cultural characteristics of the organisation and knowledge sharing will be investigated. Furthermore, future work will address the potential for critical success factors of St Mary's ACoP to vary in different stages of the development of communities. This will contribute to the evolution of future communities by facilitating approaches that can overcome the barriers to their use, particularly in the context of CoPs for professional development in higher education.

REFERENCES

[1] T. Ward, K. Monaghan, and R. Villing, "MyVLE: A case study in building a universal telematic education environment for a small university," EURODL,

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

http://eprints.nuim.ie/1267/1/Ward_Monaghan.htm., October 2010. (Accessed: 16/05/2013)
[2] G. M. Kundi and A. Nawaz, "From objectivism to social constructivism: The impacts of information and communication technologies (ICTs) on higher education," Journal of Science and Technology Education Research, vol. 1, pp. 30-36.
[3] B. J. Wadsworth, Piaget's Theory of Cognitive and Affective Development, Boston: Pearson, 2004.
[4] L. S. Vygotsky, Mind in Society: The development of higher psychological processes, Massachusetts: Harvard, 1978.
[5] N. Selwyn Education and Technology: key issues and debates, London: Continuum, 2011.
[6] R. D. Pea, "The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity," The Journal of the Learning Sciences, vol. 13, pp. 423-451, 2004.
[7] D. McConnell, E-learning Groups and Communities, Maidenhead: Open University Press, 2006.
[8] E. Wenger, Communities of practice: a brief introduction, 2006, http://www.ewenger.com/theory (Accessed: 18/05/2013)
[9] E. Wegner, Communities of practice: learning, meaning and identity, Cambridge: Cambridge University Press, 1998.
[10] MirandaNet http://www.mirandanet.ac.uk/, 2013. (Accessed: 15/05/2013)
[11]EducationCommunities, www.educationcommunities.org/, 2013 (Accessed: 15/05/2013)
[12] Jisc Business & Community Engagement, http://bce.jiscinvolve.org, 2013. (Accessed: 15/05/2013)
[13] B. Macfarlane, Teaching with integrity, Abingdon: RoutledgeFalmer, 2004.
[14] J. Nixon, "Professional identity and the restricting of higher education," Studies in Higher Education, vol. 21, pp. 5-16, 1996.
[15] G. Nicholls, Professional development in Higher Education, London: Kogan-Page, 2001.
[16] Higher Education Academy, "About Us", http://www.heacademy.ac.uk/about, 2013. (Accessed: 18/05/2013)
[17] G. R. Luedeke, "Professionalising teaching practice in higher education: a study of disciplinary variation and 'teaching-scholarship'," Studies in Higher Education, vol. 28, pp. 213-228.
[18] Jisc, "About Us", http://www.jisc.ac.uk/aboutus.aspx, 2013. (Accessed: 15/05/2013)
[19] Jisc, "Transformations Programme", http://www.jisc.ac.uk/whatwedo/programmes/transformations.aspx, 2012. (Accessed: 15/05/2013)
[20] Jisc, "Transformations Programme, Programme Strand A–Enhanced Student Experience", http://www.jisc.ac.uk/whatwedo/programmes/transformations/stranda.aspx, 2012. (Accessed: 15/05/2013)
[21] E. Wenger "Communities of practice – a brief introduction", www.ewenger.com/theory, June, 2006. (Accessed: 18/05/2013)
[22] S. Ziovas, M. Grigoriadon and M. Samarakou, "Supporting learning in online communities with social software: An overview of community driven technologies", http://cdn.intechopen.com/pdfs/8604/InTech-Supporting_learning_in_online_communities_with_social_software_an_overview_of_community_driven_technologies.pdf (Accessed: 18/05/2013)
[23] P. Gongla and C. R. Rizzuto, "Evolving communities of practice: IBM Global Services experience," IBM Systems Journal, vol. 40, pp. 842-862, 2001.
[24] E. Wenger, R. McDermott, and W. M. Snyder, "Cultivating communities of practice", Harvard Business Press: Harvard, 2002.
[25] P. Gannon-Leary and E. Fontainha "Communities of Practice and virtual learning communities: benefits, barriers and success factors", http://www.elearningeuropa.info/files/media/media13563.pdf (Accessed: 18/05/2013)
[26] K. Toneva, K. Doncaster and D. Bravenboer, "Using Virtual Spaces for Online Communities to Facilitate Project Development and Collaborative Learning", The Fourth International Conference on Mobile, Hybrid, and On-line Learning, January, 2012, http://www.thinkmind.org/index.php?view=article&articleid=elml_2012_4_60_50105 (Accessed: 18/05/2013)

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Lebanese Industrial Center for Assessment and Development

A Collaboration Platform

Ramy Harik
Department of Industrial and Mechanical Engineering
Lebanese American University
Byblos, Lebanon
ramy.harik@gmail.com

Wissam El Hachem
Department of Industrial and Mechanical Engineering
Lebanese American University
Byblos, Lebanon
elhachemw@gmail.com

Alexandre Durupt
Département Génie des Systèmes Mécaniques
Université de technologie de Compiègne
Compiègne, France
alexandre.durupt@utc.fr

Benoit Eynard
Département Génie des Systèmes Mécaniques
Université de technologie de Compiègne
Compiègne, France
benoit.eynard@utc.fr

***Abstract*— Lebanese manufacturing industries face both severe external and internal competitions. External competition from low wage production countries and adoption of free market strategies force SMEs (Small and Medium Enterprises) to tap into productivity enhancement strategies. Internal competition comes from the lack of collaboration tools and infrastructure. This research aims at establishing a positive collaboration platform between industries and enabling them to share knowledge to reduce external fierce competition. A first step towards this effort would be to establish a relevant framework that would assess the current situation of an industry through the sustainability diamond matrix. Next stages would include grouping industries according to common deficiencies and establishing a cooperation framework.**

Keywords-Collaboration; Sustainability; Diamond Index; Center

I. INTRODUCTION

This paper presents the first step of establishing a center of industrial collaboration (LICAD)[1]: assessing the current industrial situation through a balanced quantitative sustainability approach. A drastic reshaping of how modern manufacturing is expanding its capabilities is needed for it to become sustainable in the long term. Reduction of emissions, control of waste generation and disposal, monitored energy consumption and responsible material selection (renewable ones) are the main pillars supporting every company wishing to be labeled as an environmentally-conscious manufacturing firm. However, for manufacturing-environmental sustainability to gain momentum of its own it has to be socially-economically viable. So, a sustainability framework index with 4 main pillars (Manufacturing, Environmental, Economical and Societal) will be presented later on, built specifically after an extensive literature review for the Lebanese manufacturing industry with an initial collection of multi-disciplinary indicators. This will serve to classify industries based on their primary needs as a first step towards the establishment of the collaboration platform.

Following, we will give a brief topic overview that will eventually lead to the identification of sustainability scales. In section three, we justify using casual loop diagrams, in systems dynamics, to represent the various components of our variables. In section four, we will detail the application of analytic hierarchy process to aggregate the indicators and group deficiencies together. The paper ends with a conclusion on the importance of assessing further industries to refine the diamond index.

II. TOPIC OVERVIEW

Sustainability is an elusive concept hard to delimit. We will adopt a mixed quantitative/qualitative estimation of sustainability. The approach will attempt to integrate unrelated indicators. For example, the comparison of water consumption [2] with Brand management [3] is tricky, to say the least, and to valorize mostly intangible notions is prone to subjectivity which might compromise the validity of the assessment [4]. However, using the Analytic Hierarchy Process introduced in the next section, the need for sustainability performance measurement is crucial, and by capturing the main indices and inducing the learning process, a leap towards local industrial sustainability is the ultimate goal. It goes without saying that the sustainability indices collected and aggregated are by no means absolute nor completely reliable, and they serve as the first step in the LICAD initiative to holistically enhance Lebanese industry. As it is called Lebanese sustainability diamond index, it is natural to deduce that it is an iterative process and it needs to be monitored for several years to set the baseline expectations and form a clearer picture of trend-indicators to periodically update indices.

Four main scales can be identified into which sustainability can be applied [5]:

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

1-Global systems: global warming, ozone depletion, biodiversity, etc. Treaties and joint ventures are most effective on this scale.

2-Bounded systems: urban planning, transportation, etc. Economics, law, ecology among others are most effective on this scale.

3-Business systems: through energy efficiency, cleaner technology, recycling policies, leaner supply chains, etc. business sustainability is possible.

4-Technological systems: in providing clear economic value while operating in a cleaner manner can technology sustainability be achieved.

This section enabled us to situate the complexity behind our work: integrating quantitative and qualitative data to generate a compilation of indexes. The following sections will explore factors influence on one another and the normalization of this influence.

III. SYSTEM DYNAMICS

Many of the factors involved in sustainable manufacturing interact, influencing and depending on each other. This is why we propose using system dynamics and, in particular, causal loop diagramming in this paper to model the different direct depending relationships and, through these, the many chains of cascading interactions between often very odd groups of variables. In order to integrate a multitude of variables and indicators, we will need to distribute them into pillars, categories and sub-categories. Their representation through a casual loop diagram will allow for a better more holistic view of the various components of sustainable manufacturing and enable decision makers to better assess companies. Their chain reaction (domino) effect will thus be best represented: a variable under a certain pillar might as well affect a variable theoretically not related. We identify four separate domains for analysis, i.e., manufacturing, environmental, financial, and social domains. The manufacturing occupies the central part of the diagram, and thus it evaluates policies intended to promote sustainable manufacturing practices [6].

IV. GENERATING THE SUSTAINABILITY INDEX

Since a hybrid qualitative/quantitative approach will be adopted, a comprehensive index is preferred for the decision making purposes. Through numerical aggregation, a single index will emerge that will assess the company's sustainability performance across the several domains [7]. There has been numerous methods proposed to compare indicators of different nature in a systematic manner, while capturing and controlling the inherent judgment subjectivity. Among the most adopted methods are AHP (Analytic Hierarchy Process) introduced by Saaty [7] and the Multi-attribute Utility Theory proposed by Keeney [8].

In this paper, the AHP method will be adopted with a few proven modifications in literature [9]. By now, there are many proposed sustainability assessment frameworks with a great variety of indicators and units. It is better, yet challenging, to conduct comparisons among sectors and companies relying on many performance indicators. AHP makes it possible to prioritize alternatives based on different criteria, thus enabling judging and comparing companies based on multi-disciplinary indicators. Criteria are split between quantitative and qualitative scales. This makes direct integration into a single composite index not possible. For example, measuring the percentage of female managers is straight forward, while assessing deployed processes requires a certain expertise and human judgement. It is the prioritization derived by pair-wise comparison that solves the multi-scale judgments problem and enables calculation, through weighing and adding, of the overall priorities of alternatives which mark their impact on the primary goal of sustainability. AHP has several clear benefits; it can handle in a relatively easy manner data of qualitative and quantitative nature, and perhaps one of its most appealing feature it that it allows compensatory rules in multi-criteria decision making process. What this infers is that by adopting the AHP, poor performance of certain criteria can be offset by good performance of others, i.e., bad score in the social dimension can be offset by a good one in the environmental dimension leading to an overall passing performance with respect to company sustainability. We will present the procedure, based on AHP, of calculating a composite index aggregating indicators ranging over environmental, societal, economical and manufacturing levels which will permit quick and efficient ranking of companies within a sector with respect to sustainability performance.

A. Selection of Indicators

After extensive literature review [10][11], suitable performance indicators which fall under Societal, Economical, Environmental and Manufacturing dimensions are selected with the goal of preliminary assessment of a company sustainability performance and benchmarking industries within a given sector. Indicators, whenever possible, ought to be of a quantitative nature for the obvious reason of less risk of subjectivity. However, for some indicators, a qualitative judgment is more appropriate. An important requirement for the continued accuracy of this methodology is the periodical review of indicators to track any changes in status and possibly add a few if the need arises. An indicator needs to have clear unit of measurement if quantitative and clear judgment scale if qualitative.

Let us set the main sustainability dimensions (or group of indicators) as the j's (economic j =1, environmental j = 2, and societal j=3, manufacturing j=4). The manufacturing dimension has been separated as a standalone dimension since the field of application is restricted to primary and secondary industries. The respective indicators serve as an initial assessment enablers of how much a company is seeking to modernize and achieve the much sought over status of sustainable development.

- The economic group of indicators serves as an overview of the company's economic impact on its stakeholders as well as the local and national economic system. It uses basic economic interactions of expenses versus revenues.
- The environmental group of indicators serves as an overview of the company's impact on its

surroundings (water, atmosphere, land). It uses basic environmental indicators such as pollution and recycling.

- The societal group of indicators serves as an overview of the company's endeavors to properly treat its employees, suppliers, customers. It uses basic societal interactions such as employee turnover rate and safety.
- The manufacturing group of indicators serves as an overview of the company's technological advancement so as to minimize production related waste and emissions. It uses basic manufacturing indicators such as inventory tracking and machine scheduling.

Some of these dimensions will be split into sub-categories for concentrated judgment (to be explained in the next section) and for easier calculation (also to be explained in later section, mainly to keep down the number of pair-wise comparison). The environmental indices are presented in table 1 with three different sub-categories (1= atmosphere, 2= policies, 3= standards) labeled between parentheses from 1→3.

B. Judging the Indicators

The indicators have to be sorted between those who positively and negatively influence the company's goal towards achieving sustainability. So, let us mark the indicators with I's so to have I_ji (Indicator i from group j) and with + or – signs like I_(ji+) or I_(ji-) so to indicate respectively whether or not if they increase in value they contribute towards enhancing overall sustainability score. For example increased energy consumption per unit of output has a negative effect on the overall score.

C. Weighing the Indicators

It is a very critical step the weighing of these indicators. As with most cases, lack of data hinders absolute certainty when judging and placing weights. However, the pair-wise comparison technique derives relative weights in a practical and highly accurate manner relying on a selected group of experts. Pair-wise comparison enables the estimation of weights for each indicator i relative to other indicators within the same group j and it is done by comparing each pair of indicators. By answering the question of which of the two indicators i and j is more influential in the overall sustainability of the company. By indicating a preference for one over the other on a 1-9 (this range was chosen because it offers limited yet sensitive judgment enabling distinction) scale, the intensity of relevance is detected. For example, if indicator i is given a score of 4 when compared to indicator j, this means that i is 4 times more important than j in determining the overall score of the company.

After comparing each pair of indicators, a NxN positive reciprocal matrix is formed (1) with the reciprocal property (if i is 5 times more important than j, then j is 1/5 times more important than i).

$$a_{ij} = \frac{1}{a_{ji}} \quad (1)$$

Equation 1: Reciprocal property of a reciprocal matrix

A 3 step approximation of the normalized weights of the indicators:

1- Sum entries each column of the reciprocal matrix (2)

$$\sum_{i=1}^{N} a_{lk} \quad (2)$$

Equation 2: Sum of each column in reciprocal matrix

2- Divide each entry by the sum of its column to obtain the normalized pair-wise comparison N x N matrix (3)

$$\bar{a}_{lk} = \frac{a_{lk}}{\sum a_{lk}} \quad (3)$$

Equation 3: Normalized pair-wise comparison

3- The criteria weight or priority vector w is obtained by averaging the sum of each row (4)

$$W_k = \frac{\sum \bar{a}_{lk}}{N} \quad (4)$$

Equation 4: Priority column vector W

D. Checking Consistency

For an N x N matrix, (N-1) comparisons are enough to derive weights for the N indicators. However in AHP, N(N-1)/2 comparisons are made, which is more computationally exhaustive for the purpose of rendering the weights less sensitive to our inherent inconsistency of judgments and its redundancy permits the calculation of a ratio to measure the data's degree of inconsistency and determine the validity of the results called the consistency ratio. The smaller the ratio the better, and, as a general rule of thumb it should be ideally less than 0.1, however acceptable up to the limit of 0.2. A perfectly consistent matrix has a CI=0 as per condition listed in the equation (5) below.

$$lambda_{max} = lambda = n \quad (5)$$

Equation 5: Principal Eigen value for a consistent matrix

Two people will most definitely differ in their belief, at least as for the intensity of the degree of difference between two indicators, the same person will most likely report inconsistent judgments (i.e., A much more important than B, B slightly more important than C and C is slightly less important than A), which is part of our nature. The transitive property of logic of preference is what determines the consistency of the judgments, however too much consistency is impossible even undesirable in this case since it makes the AHP looses one of its more desired features

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

which is to systematically combine experts opinions in a structured manner without losing their subjectivity enabling the determination of real experience weighted indicators. The rank of the indicators has to be transitive; however, the degree of intensity preference does not.

$$I_{ij} \times I_{jk} \neq I_{ik} \quad (6)$$

Equation 6: Relaxed (non-transitive) intensity preference

Three steps are required to compute the consistency ratio:

1. Compute the principal eigen value which the summation of product of summation of each column in the reciprocal matrix and the priority vector.

$$lambda_{max} = sum\ product[\sum_{l=1}^{N} a_{lk}, W_k] = sum\ product[\sum_{l=1}^{N} a_{lk}, \frac{\sum_{k=1}^{N} \bar{a}lk}{N}] \quad (7)$$

Equation 7: Principal Eigen value

2. Compute the consistency index

$$CI = \frac{lambda_{max} - n}{n-1} \quad (8)$$

Equation 8: Consistency index

3. Compute the consistency ration CR relying on prof Saaty's random consistency index RI

$$CR = \frac{CI}{RI} \quad (9)$$

Equation 9: Consistency ratio

- If CR<= 0.1, the judgments are said to be consistent enough
- If CR<=0.2, the judgments are said to be acceptable and need to be revisited if possible
- If CR>0.2, the judgments are said to be inconsistent and a re-evaluation of the alternatives is in order.

Random consistency index RI is given in the following table 1:

TABLE I. RI VALUES

	n values									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

E. Normalizing Indicators

In order to aggregate different indicators of different natures and behaviors from every group j across several years t into a single index, normalization is necessary. The fitting time-dependent Min-Max normalization will be used, and the normalized indicators will range between 0 and 1. The following two equations are used for more is better and less is better indicators, respectively.

$$I^{+}_{N,ijt} = \frac{I^{+}_{ijt} - I^{+}_{min,jt}}{I^{+}_{max,jt} - I^{+}_{min,jt}} \quad (10)$$

Equation 10: Min-Max normalization for positive indicators

This normalization allows for different indicators to become compatible and therefore possible to aggregate. However, for this normalization to be possible and in order to benefit from the more holistic approach of mixed quantitative/qualitative assessment there remains the problem of assessing the qualitative indicators values in such a way to still be able to use the more accurate relevant model and not having to switch entirely to the rating model. The following proposed steps allow valorizing of qualitative indicators to fit alongside the quantitative ones in the relevant model scoring methodology:

1. Using the qualitative rating scales used in the rating model (for example High, Medium, Low), a pair wise comparison will derive values for each of the scale components (ex: High=0.5, Medium=0.3, Low=0.2).
2. Now, based on the expert's opinions, a company will be deemed to rank on each one of the rating components given a certain percentage of certitude (ex: 30% high, 60% medium, 10% low).
3. By multiplying the scale components value with their respective probability we get an approximate value of the indicator (ex: high=0.5*0.3=0.15, medium=0.18, low=0.02)
4. We idealize these number by dividing each one by the largest one (ex: high=0.833, medium=1, low=0.111)

F. Calculating Sub-Indices

Before calculating the global sustainability index, sub-indexes $I_{N,ijt}$ for the four domains need to be calculated.

$$I_{jt} = = \sum_{jit}^{n} W_{ji} I^{+}_{N,ijt} + \sum_{jit}^{n} W_{ji} I^{-}_{N,ijt} \quad (11)$$

Equation 11: Sub-index

$$\sum_{ji}^{n} W_{ji} = 1, W_{ji} \geq 0 \quad (12)$$

Equation 12: Weight of indicator i in group j

I_{jt} is the sustainability sub-index for group j in time t and W_{ji} is the weight of indicator i in group j.

G. *Combining sub-indicesinto the global index*

The global sustainability index I_S is calculated using the weighted sum of the sub-indexes (13):

$$I_{St} = \sum_{jt}^{n} W_{jt} \times I_{jt} \quad (13)$$

Equation 13: Sustainability index

V. CONCLUSION AND FUTURE WORK

This paper presented the index that will be used to assess around 337 local industries. The results will serve to categorize companies based on common deficiencies and support the collaboration process. The deficiencies will be categorized according to the preliminary four categories as well as to the sub-categories. Currently, the 337 industries to be investigated were selected and a collaboration database is being established on [1] .

Future steps will include data gathering from the selected industries and surveys sent for technical managers. Following a thorough analysis of the gathered information, the selected indicators will be re-investigated and confirmed. An online tool will then be made available for industries to test their sustainability and to identify which norms they need to investigate to enhance their index.

ACKNOWLEDGMENT

The authors would like to thank a NCSR Grant in Lebanon that allowed the completion of this work.

REFERENCES

[1] www.licad.net, last access on June 25, 2013

[2] L. Bettoni and S. Zanoni, "Energy Implications of Production Planning Decisions," IFIP Advances in Information and Communication Technology, vol. 384, 2012, pp. 9-17, http://dx.doi.org/10.1007/978-3-642-33980-6_2, last access on 4 July 2013

[3] S. Gupta and V. Kumar, "Sustainability as corporate culture of a brand for superior performance," Journal of World Business, vol. 48 (3), July 2013, pp. 311-320, http://dx.doi.org/10.1016/j.jwb.2012.07.015, last access on 4 July 2013

[4] R. K. Singh, H. R. Murty, S. K. Gupta, and A.K. Dikshit, "An overview of sustainability assessment methodologies" Ecological Indicators, vol. 9, May 2008, pp. 189-212, http://dx.doi.org/10.1016/j.ecolind.2008.05.011, last access on June 25, 2013

[5] S. Sikdar, "Sustainable Development and Sustainability Metrics," Institute of Chemical Engineers Journal, vol. 49 (8), August 2003, pp. 1928-1932, http://dx.doi.org/10.1002/aic.690490802, last access on June 25, 2013

[6] S. Jain and D. Kibira, "A framework for multi-resolution modeling of sustainable manufacturing", Proceedings of the 2010 Winter Simulation Conference, pp. 3423-3434, 5-8 Dec. 2010, Baltimore, USA, http://dx.doi.org/10.1109/WSC.2010.5679032, last access on June 25, 2013

[7] T. L. Saaty, "Decision making with the analytic hierarchy process,", International Journal of Services Sciences, vol. 1 (1), 2008, pp. 83 – 98, http://dx.doi.org/10.1504/IJSSCI.2008.017590, last access on June 25, 2013

[8] R. L. Keeney, "Multiplicative Utility Functions," Operations Research, vol. 22, 1974, pp. 22–34, http://dx.doi.org/10.1287/opre.22.1.22, last access on June 25, 2013

[9] D. Krajnc and P. Glavic, "How to compare companies on relevant dimensions of sustainability," Ecological Economics, vol. 55 (4), 2005, pp. 551-563, http://dx.doi.org/10.1016/j.ecolecon.2004.12.011, last access on June 25, 2013

[10] J. Uddin, S. Parvin and L. Rahman, "Factors influencing importance of country of brand and country of manufacturing in consumer product evaluation," International Journal of Business and Management, vol. 8, 2013, pp. 65-74, http://dx.doi.org/10.5539/ijbm.v8n4p65, last access on June 25, 2013

[11] C.M. Stracke, "Process-oriented quality management," Handbook on Quality and Standardization in E-Learning, 2006, pp.79-96, http://dx.doi.org/10.1007/3-540-32788-6_6, last access on June 25, 2013

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Experiments in Collaborative Cloud-based Distance Learning

Aleksandar Karadimce, Danco Davcev
Faculty of Computer Science and Engineering, University Ss Cyril and Methodius
Skopje, R. Macedonia
akaradimce@ieee.org, danco.davcev@finki.ukim.mk

***Abstract*—Mobile devices have revolutionary changed the process of multimedia consumption in distance learning systems and provide flexible user-oriented experience. M-learning systems typically include different kinds of multimedia resources helping learners to be more interactive and interested in collaboration. The proposed multi-tenancy group collaboration within the cloud-based learning platform (as a main contribution of this paper) provides easy and simple access to the cloud-based Platform as a Service (PaaS) model and delivers improved interactive tools for distance learning. This collaborative and interactive cloud-based learning environment that students are using to develop and deploy intelligent multimedia content resources has been subjected to quality of experience (QoE) estimation. Using the Questionnaire for User Interaction Satisfaction (QUIS) as estimation tool, the system has provided an effective collaboration environment between the students and the professor, by continuous interaction in the multi-tenant cloud environment.**

Keywords - distance learning; cloud computing; quality of experience; collaborative group learning.

I. INTRODUCTION

Traditional university applied teaching requires the professor to be present in the classroom to provide scaffolding and collaboration with students aimed to complete all assigned tasks that they are facing. The process of dissemination of the multimedia content in distance learning systems is a challenging task and requires more than just simple static HTML-based websites. With the introduction of the Web 2.0 technologies in distance learning context, we are facing a new era in the practice of e-learning [1]. L. Wan [1] has proposed a framework that consists of Web 2.0 tools, e-learning 2.0 application and e-learning 2.0 learning modes. Recent distance learning systems are directed to provide on-line collaborative tools, which are using Web-based e-learning architecture [1].

The multimedia material that is used in collaborative distance learning environment should be adjusted according to the mobile device capabilities, student domain knowledge level, interaction style and skills [2]. There have been several distance learning systems, like in [3], that propose a decentralized and distributed adaptive architecture, which would bring the possibility to support collaborative activities for sharing resources between students with their mobile devices in constantly changing conditions. Challenges in distance learning systems exist in the process of delivering individualized content according to student's particular learning needs, by using several technologies to ensure maximum collaborative learning tasks [4]. Taking into consideration the results from research by I. Jung [4], we have concluded that *interaction* is an important dimension in evaluating the quality of e-learning.

Mobile devices provide an interactive and user friendly interface that was investigated in mobile video retargeting and location-based systems [5], as well as in mobile health care applications [6]. Especially interesting is mobile multimedia computing because it is a computationally intensive task on mobile devices limited by battery and computing power [5]. Modern m-learning [7] applications are complex because they deliver learning content usually as multimedia material, which combines texts, pictures and video/audio files. Because of the limited mobile resources, algorithms designed for mobile multimedia systems need to be lightweight or when Internet access is available, part of the computing tasks to be offloaded in the cloud where more resources (such as computing, memory, and relevant data) are available. The existing research by C. Ciurea and P. Pocatilu [7], proposes m-learning applications and mobile learning processes inside a virtual university campus that have the great potential to enable innovative and effective distance learning techniques. Some of the positive benefits of using cloud computing architectures in distance learning systems development for m-learning systems are presented in [7], [8] and [9]. M. S. Fathi et al. [9] presented and developed the concepts and potential of innovative collaborative tools, such as Context-Aware Cloud Computing Information Systems, in order to ensure the delivery of applicable and reliable information to enhance the collaboration of the parties within the construction supply chain and project stakeholders.

The main contribution of this paper is the advantage of using the multi-tenancy group collaboration in cloud-based learning platform (CCLP) during distance learning processes in order to provide increased quality of e-learning. It will be estimated by the quality of experience (QoE) analysis conducted with a group of university students. We propose our collaborative cloud-based learning platform to provide easy and simple access to the cloud-based Platform as a Service (PaaS) model, in order to provide improved interactive tools for distance learning application.

This paper is organized as follows: Section II presents our proposed architecture of collaborative cloud-based architecture. Section III presents a case study for distance m-learning of a database course. Section IV gives overview on the results from the QoE estimation model. Finally, Section V concludes the paper.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

II. COLLABORATIVE CLOUD-BASED ARCHITECTURE

M-learning systems typically include different kinds of multimedia resources helping learners to be more interactive and interested in collaboration. The proposed collaborative cloud-based architecture for distance learning system uses Platform as a Service (PaaS) [11] cloud model that comes with integrated developer tools, database management system and Web server. This platform provides multi-tenant environment where users can work on different operating systems from mobile devices, and that way students and professors can access the cloud-based platform simultaneously from any location, at any time. The proposed collaborative cloud-based architecture is given in Figure 1, where different layered components in distance learning environment are illustrated.

Collaborative cloud-based architecture consists of three main parts: user interface layer, network layer and cloud-based multimedia delivery system. Mobile clients (smart phone, tablet and laptop) are members of the *user interface layer* that provide various access points for accessing our cloud-based system. The *network layer* is responsible for daily operation of the platform and monitoring Quality of Service (QoS). Networks are the main means to deliver the multimedia content through a variety of protocols (LTE, WiMax, WiFi, GSM, EDGE, etc.). The third component of this architecture is based on the cloud computing model and is composed of hardware (infrastructure) layer, virtualization layer, platform layer and multimedia service. The cloud management server uses the request broker agent that collects the users' requests for multimedia content and, depending on the estimated user profile and context-aware conditions, the adaptation agent is delivering the appropriate multimedia content to the user. The proposed cloud-based technologies provide high interactive environment (on a single or a group of users' level) as a basis for delivery of personalized and adapted multimedia content.

Using this cloud-based architecture, the student can work on their application independently and efficiently without additional problems with software installation and compatibility issues.

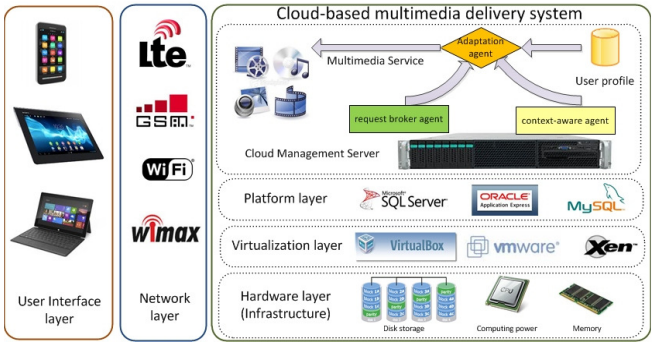

Figure 1. Collaborative cloud-based architecture.

The cloud provides ease of transfer of multimedia content between the students, using mobile devices, and the university distance learning portal. Students have the main developer role, they are creating database objects, and they write SQL request queries and develop Web-based applications. The professor has a supervision role, but also very important interactive possibilities based on PaaS cloud model. The learning process starts when mobile users enroll to the Web 2.0 university distance learning portal, then locate the link to PaaS cloud model learning platform. This way, the student has direct access from his personal mobile device to the remote Database Management System (DBMS) and he/she is able to send the SQL query to the cloud and receive results on his mobile device. Using the proposed cloud-based architecture, professors can monitor the students' progress and can provide scaffolding and propose more efficient solutions or can interactively support the error debugging in the application development progress.

III. CASE STUDY – DISTANCE M-LEARNING OF A DATABASE COURSE

The proposed collaborative cloud-based learning platform has been demonstrated to the students using the Web 2.0 university distance learning portal, see Figure 2. We have created a distance learning course for learning the database course by using the ORACLE database. The course covers the Entity Relations (ER) – real world modeling techniques that provide basic knowledge for designing relational database systems, see Figure 3. Our case study provides interactive communication with real database problems that students need to handle.

In order to conduct QoE estimation in our case study using the university distance learning system, the students were split into two teams: control and experimental groups. The first control group was subjected to regular distance learning course for SQL query data manipulation, and on the other hand, the second experimental group of students has been provided with collaborative cloud-based learning platform - APEX.

Figure 2. University Web 2.0 distance learning portal.

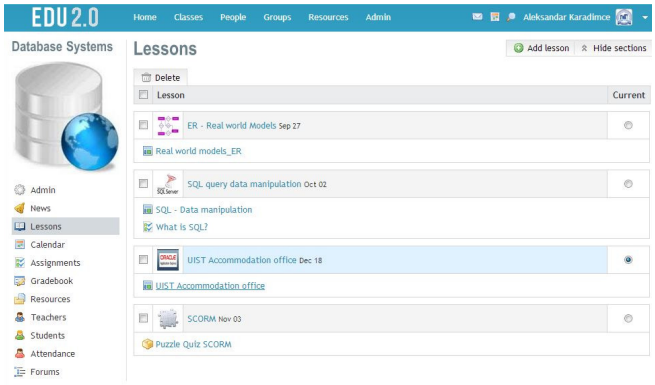

Figure 3. List of lessons for the course Database systems.

For the purpose of this case study, both lessons where available for learning on the Web 2.0 university distance learning portal under the database course in the section Lessons, see Figure 3.

A. *Control group of students*

The control group consisted of 15 students who have been offered a guided instructors' video and audio course that is explaining the process of constructing correct SQL data manipulation queries, see Figure 4. This course presented an existing database to the control group of students, where using video and audio guidance have been given scaffolding for SQL query development process. For this part of the research we have used only multimedia streaming services, without any help from cloud computing technology in order to deliver the instructions to the control group of students that were using mobile devices.

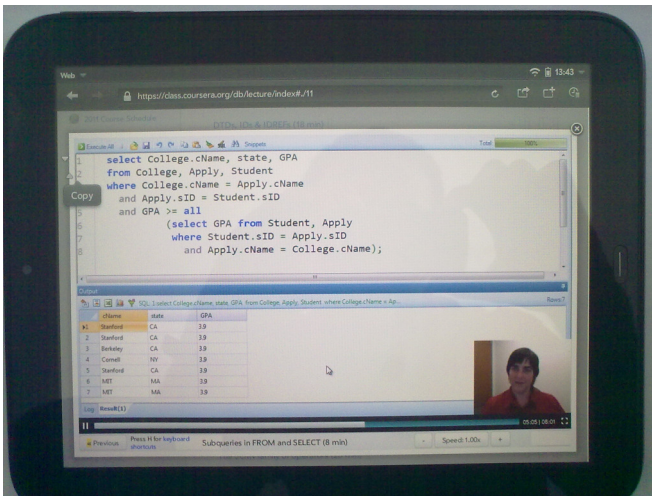

Figure 4. Course presented to the control group of students.

B. *Experimental group of students*

Using the proposed collaborative cloud-based learning platform, we have delivered customized interactive learning course for the experimental group of 15 students. Oracle APEX [10] is a collaborative learning environment for processing Web-based SQL requests, and offers a platform for fast, reliable development and running Web applications. Oracle Application Express (APEX) is a rapid Web application development tool for the Oracle database, which relies on Web browser in order to gain hardware independence. Using mobile devices to access this platform, students are able to learn how to build SQL queries and incorporate different types of items and shared components in the Web-based application [10]. Their learning effort was completed with developing a Web-based application "UIST accommodation office" (see the mobile screenshot in Figure 5) that was created from the proposed cloud-based platform. This system has provided an effective collaboration environment between students and professor, by continuous interaction in the multi-tenant cloud environment.

The database Web application was developed for the needs of the University Accommodation Office. It contains information about students, accommodation room types, lease agreement types, invoices and student flats inspections. Using this Web-based application, all the users can easily see their status of acceptance in the University Accommodation Campus. To estimate the benefits of this innovative collaborative cloud-based PaaS model for distance learning, we conducted a quality of experience (QoE) analysis for both university student groups.

IV. RESULTS FROM THE QOE ESTIMATION MODEL

Mobile devices have revolutionary changed the process of multimedia consumption in distance learning systems and provide flexible user-oriented experience. Today, there is a growing trend to estimate the user satisfaction of video, audio and interaction level with the collaborative distance learning systems. These collaborative and interactive environments that students are using to develop and deploy intelligent multimedia content resources have to be subjected to estimation using a QoE model. The QoE estimation, according to the International Telecommunications Union (ITU) has been defined as "the overall acceptability of an application or service, as perceived subjectively by the end-user" [12]. This requires subjective perception of the end-users and indicates the importance of the context in which a system is being used, as well as the expectations that users have about the system [12].

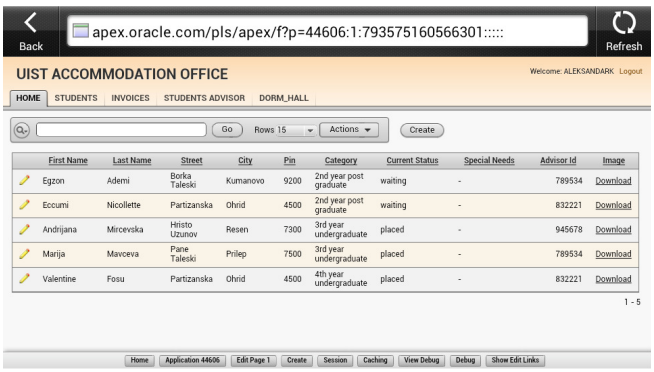

Figure 5. Course presented to the experimental group of students.

In our case study, of distance M-learning of database course, we have used rating scale for subjective evaluation of the student's opinion, as required. The aim of Questionnaire for User Interaction Satisfaction (QUIS) is to measure users' satisfaction with the human-computer interface and involves users' satisfaction to examine the relationship with the users' knowledge [13]. Students from both groups, control and experimental, were therefore required to respond to 22 questions, using a 9-point response rating scale. Scores were recorded so that a high score denoted a factor which was highly satisfactory, and a low score indicated a more negative response. This questionnaire provided a qualitative approach of the QoE that is applied to obtain the details of the participants' interaction with mobile device interface based on the individual experience. The QUIS questionnaire presented observations from the students overall reaction to the software, mobile screen, software terminology and system information, process of learning and system capabilities [13].

We have conducted our survey for QoE evaluation model in order to measure the quality of achieved learning. Distance learning systems should be supported with increased interaction using multi-tenant mobile cloud computing environment. The QUIS survey questions were answered by both of the groups consisted of 15 students, control and experimental groups, which participated in the database course.

The results from the control group of students that participated in regular distance learning course for SQL query data manipulation are given in Figure 6. We can conclude that control group of students have demonstrated better results from 12% to 14% percentage response for the values 6-7 on the 9-point rating scale.

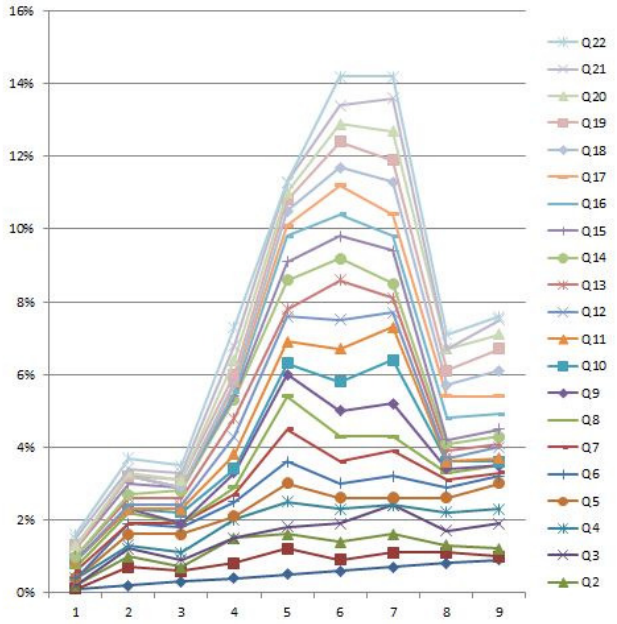

Figure 6. Results from the control group of students.

On the other hand, the experimental group of students participated in the course with *collaborative cloud-based learning platform* and created the mobile application UIST Accommodation Office. The analysis of the QUIS questionnaire completed by the students that participated in this group is given in Figure 7. From that chart we concluded that experimental group of students have demonstrated better results from 15% to 20% percentage response for the value of 8 on the 9-point rating scale.

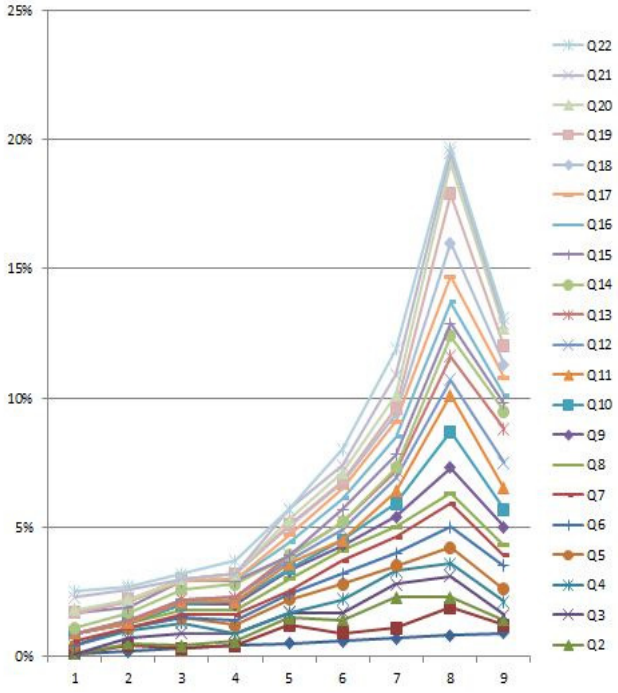

Figure 7. Results from the experimental group of students.

This research examines individuals' performance with mobile device interface according to QoE estimation model in order to confirm the consistency of this research. Comparison of the results from the QUIS survey completed by the two groups of students has allowed us to conclude that the group of students that was using collaborative cloud-based learning platform had increased overall attention and provided educational benefit in the process of learning. It is obvious that the cloud-based learning platform, using the benefit of the multi-tenancy collaboration, had provided higher interaction between professor and students.

V. CONCLUSION AND FUTURE WORK

The proposed collaborative cloud-based architecture enhances the importance and the benefits of collaborative and interactive knowledge in distance learning systems. This proposed system is based on the PaaS cloud model, which comes with integrated developer tools, database

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

management system and Web server. That way, students and professors can access the cloud-based multi-tenant platform simultaneously from any location in any time. We have understood the advantage of using multi-tenancy group collaboration in cloud-based learning platform through the quality of experience (QoE) analysis conducted with a group of university students. The QoE model for evaluation was done using the QUIS in order to measure users' satisfaction with the human-computer interface and student's satisfaction that leads to improved distance learning.

In addition to the benefits from the proposed cloud-based collaborative environment, in our future work, we plan to provide an intelligent decision rules algorithm for adaptation of the multimedia content based on the user request, profile and context.

REFERENCES

[1] L. Wan, "Application of Web 2.0 Technologies in E-Iearning Context," Proceedings of the 2nd International Conference on Networking and Digital Society (ICNDS) 2010. 30-31 May 2010.Wenzhou. vol.1. pp. 437-440, DOI: 10.1109/ICNDS.2010.5479229.

[2] V. Glavinic, S. Ljubic, and M. Kukec, "A Holistic Approach to Enhance Universal Usability in m-Learning," Proceedings of the 2008 Second International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM '08). 29 Sept.- 4 Oct. 2008. Valencia, pp. 305-310, DOI:10.1109/UBICOMM.2008.54.

[3] D. Gil, J. Andersson, M. Milrad, and H. Sollervall, "Towards a Decentralized and Self-Adaptive System for M-Learning Applications," in Proc. 2012 Seventh IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education (WMUTE 2012), pp. 162-166.

[4] I. Jung, "The dimensions of e-learning quality: from the learner's perspective," Educational Technology Research and Development, Aug 2011, vol.59, n4, pp. 445-464, ISSN-1042-1629.

[5] W. Zeng, "Mobile Media in Action: Remote Target Localization and Tracking," Multimedia, IEEE. Vol.19, Issue: 3 July - Sept. 2012, pp. 74-80, DOI:ieeecomputersociety.org/10.1109/MMUL.2012.37.

[6] S. Sultan and P. Mohan, "Designing a Peer-Facilitated Self-Management Mobile Application: A User-Centred Approach," Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012. 21-24 May 2012. San Diego, pp. 228-231, DOI:org/10.4108/icst.pervasivehealth.2012.248707.

[7] C. Ciurea and P. Pocatilu, "Designing M-Learning Applications for Collaborative Virtual Environments", International Journal of Education and Information Technologies, Issue 1, vol. 6, 2012, pp. 150-156. ISSN 2074-1316.

[8] P. Pocatilu, F. Alecu, and M. Vetrici, "Using Cloud Computing for e-Learning Systems," in Recent Advances on Data Neworks, Communications, Computers, Academy of Economic Studies, Bucharest, Romania, pp. 54 -59.

[9] M. S. Fathi, M. Abedi, S. Rambat, S. Rawai, and M. Z. Zakiyudin, "Context-aware cloud computing for construction collaboration," Journal of Cloud Computing, 2012, pp. 1-11, ISSN 2330-7889.

[10] D. Baker and T. Jennings, Oracle Database 2 Day + Application Express Developer's Guide, Release 4.0. E15516-04. Oracle, 2010. http://docs.oracle.com/cd/E17556_01/doc/appdev.40/e15516.pdf [retrieved: 29/12/2012]

[11] M. Wang and J. W.P. Ng, "Intelligent Mobile Cloud Education. Smart anytime-anywhere learning for the next generation campus environment," In proceedings of Eighth International Conference on Intelligent Environments, 2012, pp. 149-156, DOI: 10.1109/IE.2012.8.

[12] D. Geerts, et al., "Linking an integrated framework with appropriate methods for measuring QoE," In Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on Quality of Multimedia Experience, pp. 158-163.

[13] QUIS website, [Online] Available at: http://lap.umd.edu/quis [retrieved: 18/01/2013]

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Argument Schemes and Provenance to Support Collaborative Intelligence Analysis

Alice Toniolo, Federico Cerutti, Nir Oren and Timothy J. Norman
Department of Computing Science, University of Aberdeen,
Scotland, UK
e-mail: {a.toniolo, f.cerutti, n.oren, t.j.norman}@abdn.ac.uk

***Abstract*—Intelligence analysis is the process of interpreting scattered information to form hypotheses and testing those against evidence. Collaboration enhances this process reducing effort and bias of an individual, while permitting more criticism and different perspectives to be considered. Existing analytical tools support an analyst in collecting information and evaluating hypotheses. However, for effective collaboration, analysts must work together in forming hypotheses from information. We propose an evidential reasoning service that aims at improving collaborative sense-making. It exploits argumentation schemes for structuring annotation and analysis of information, and records provenance to track data and reviews. This service aims to support the analysis by introducing automated reasoning about competing evidence for identifying plausible hypotheses. It provides a uniform reasoning structure that permits integration and facilitates sharing of analyses from different contributors.**

***Keywords*—intelligence analysis; argument schemes; provenance; collaborative applications.**

I. INTRODUCTION

Intelligence analysis is an iterative process of interpreting information and elaborating evidence about situations and events [4]. The product of the analysis is a report discussing hypotheses supported by evidence. These reports can be used by decision makers in informing strategies, tactical operations, non-kinetic activities, additional collection requirements, trend analysis, and so on. Collaboration is common among analysts of an intelligence agency [6], however analysts may form coalitions across different organisations in order to address complex tasks. Collaborative analysis is a difficult task within, and more so across, agencies because analysts have access to diverse sources of information that may report conflicting data and their analysis may have different purposes. Further, analysts have different expertise, part of the analysis may be dependent upon other contributors and the analytical approach may differ from agency to agency. A system that supports collaborative annotation and sense-making of information is crucial to deliver timely and accurate intelligence reports.

In order to reduce analyst workload, software tools, which logically organise information, have been developed. TRELLIS [5], for example, focuses on annotating information received from different sources, highlighting contradictions and trustworthiness of sources. XIP-Cohere [2] supports mixed automatic and human annotation. However, very few tools permit collaboration in the analytical process. Entity Workspace [1] supports collaboration in comparing and deciding upon the most likely hypothesis. POLESTAR [10] instead allows an individual portfolio of analysis to be shared across different users that can make suggestions. Although these tools have been positively adopted for collaborative analysis, they focus on *sharing* and *functional* collaboration [6], those being the activities of managing information and expertise and editing reports for completing the analysis. In our research we address collaboration at the *content* level, whereby analysts work together to reason about information and evidence. This type of collaboration permits the elaboration of more hypotheses and enables greater criticism in reasoning. However, problems with integrating and maintaining partial intelligence analyses made by analysts that differ in capabilities, access to information and analytical approaches make this a difficult challenge.

Our core research question is: *How can we support analysts within a coalition to collaborate in sense-making at the content level of analysis?* We propose an evidential reasoning service to enhance content analysis. This service follows the iterative approach of analysts to intelligence analysis and uses *argumentation schemes* for structuring annotation of information and reasoning about competing evidence to build hypotheses. Within argumentation theory, argumentation schemes are used for structuring critical thinking about a controversial statement [3]. Intuitively, these schemes provide structures for making inferences, and for exploring evidence for or against a claim [12]. Assume that an analyst investigating criminal activities asserts that "Jill collaborates with Bob and Bob is a smuggler, then Jill is a smuggler". This inference may be seen as an argument. The fact that Jill is a smuggler may be tentatively accepted unless other evidence dismissing this claim is collected. The structure of this logical inference is extracted and represented as a reasoning pattern that constitutes an argumentation scheme. In addition, the service records *provenance* (i.e., origin of the information) to track data and reviews for assessing the quality of the analytical workflow.

Using argument schemes, our evidential reasoning service guides analysts to uniformly structure links and inferences amongst information. Such structure can be used for autonomous reasoning about competing evidence to assist users in the construction and validation of hypotheses. The service facilitates the integration of partial analyses from different contributors improving collaboration at the content level of analysis. Moreover, it considers provenance that will enhance sharing and maintenance of these analyses. In this paper we discuss: challenges of the reasoning service in Section II; argumentation schemes and provenance for collaborative analysis in Section III; and future work in Section IV.

II. CHALLENGES OF THE REASONING SERVICE

Our evidential reasoning service aims at assisting analysts to collaborate throughout the reasoning process in order to create timely and accurate intelligence reports. We envisage our service to be adopted within a platform that enables

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

analysts to work together providing data-stream maintenance and managing the contributions to the analysis. The initial information is extracted from soft or hard sources and expressed via a human/machine readable language permitting both user and automatic elaboration (such as in [2]). Although there is no established computational representation of intelligence elements, we can broadly identify *entities* such as people and places, *events* including actions or activities, and *facts* about situations. The analysis focuses on making sense of scattered information by exploring relations between these elements. Important relations are *temporal* relations representing correlation, causality and chronological distribution of events and activities, and *association* relations that connect people and events. The mental process of intelligence analysis is identified by two main loops [11]: the *Foraging Loop* intended as the process of collecting information and extracting evidence; and the *Sense-Making Loop* that aims to structure and annotate evidence to form plausible hypotheses and prepare documents to be presented to the decision makers. In this research, we focus on the formation of evidence ("Ev") from information ("Inf") and the formation of hypotheses ("Hyp") from evidence (see Figure 1). Following this model, our service provides a virtual space for each level of elaboration (for a topic) that can be accessed (as appropriate) by a team of contributors and it must proactively assist analysts in these phases.

Within a coalition, experts in different fields have access to different sources, and to a vast amount of domain specific information. The identification of connections is complex because analysts may only be allowed to view or share partial information, they may exploit different evidence, and draw different or even incompatible conclusions. Biases such as confirmation bias, whereby an analyst considers only information that confirms one's beliefs, may also prevent an analyst to draw accurate conclusions. Moreover, teams of analysts may iteratively or concurrently review the analysis and modify hypotheses according to new information. The workflow of partial analyses must be considered to better assess their quality. In the development of the evidential reasoning service we aim to address the following questions:

- How do we support analysts in structuring information and making links by exploring the relations among information?
- How do we support analysts in summarising and annotating evidence for building hypotheses?
- How do we support analysts in identifying acceptable hypotheses?
- How does the history of analysis and manipulation affect the acceptability of claims for an analyst?

III. ARGUMENTATION SCHEMES AND PROVENANCE

Our approach to the above questions is to employ argumentation schemes for structuring the analysis and provenance for recording the workflow of the analysis.

Argumentation schemes. An argumentation scheme is a structured way of making presumptive inferences, stating explicitly what the premises are and what conclusions can be drawn from these premises. Associated with an argumentation scheme are critical questions (CQs), which can be used to challenge the validity of arguments. Argumentation schemes represent a method for formulating arguments in argumentation

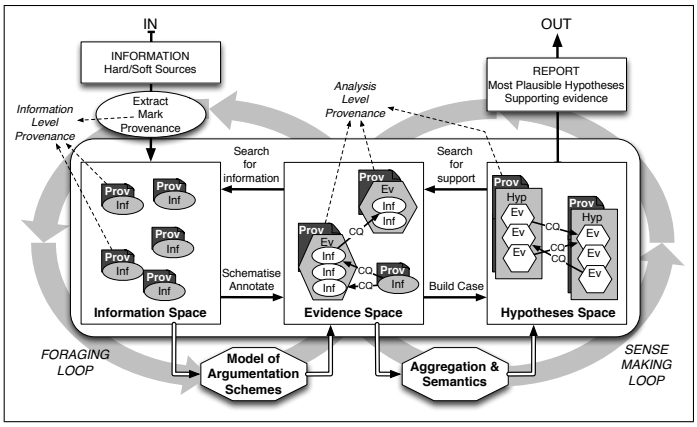

Fig. 1: Evidential Reasoning Service, adapted from [11]

theory, whose computational aspects revealed to be of great interest since the seminal work of [3]. Argumentation theory has increasingly received attention in artificial intelligence as a mechanism for representing autonomous reasoning with uncertain and incomplete information [8], by providing methods for deriving the acceptability status of arguments. In [3] an argument is rationally acceptable if it is defended against attacking arguments. For example, consider argument A_1, "Jill collaborates with Bob, Bob is a smuggler, thus Jill is a smuggler too". An attack is an argument A_2, "Jill had no contact with Bob, thus Jill does not collaborate Bob". The fact that Jill is a smuggler cannot be rationally accepted since A_2 attacks A_1. However, if A_2 is attacked by a new argument A_3, "Jill and Bob has been introduced by Mark, therefore Jill and Bob had a direct contact", claim A_1, defended by A_3, may then be reinstated. In an argumentation framework, where arguments and attack and support relationships between arguments are specified, a criterion, called semantics, is defined to establish the acceptability of arguments considering relationships between arguments.

In computational systems the use of argumentation schemes has been introduced to formulate arguments since they provide structures that can be applied to diverse information and permit the representation of reasoning in complex domains [8]. Empirical studies in domains such as law provide a variety of argumentation schemes derived from patterns of common human reasoning and dialogue [12]. Argumentation schemes are defeasible in the sense that the premises warranting the conclusion are tentatively accepted according to existing evidence. In the light of new received information the conclusions may be discarded if this invalidates the claim. In our example, argument A_1 can be structured using the scheme (from [12]):

- **Premise 1** - M member of group G has quality Q,
- **Premise 2** - if M has property Q, G will have Q,

$\Rightarrow$ **Conclusion** - Therefore, every member of G have Q.

where quality Q is being a smuggler, M is Bob and the group G is formed by Bob and Jill. This scheme can be applied to form a different argument with similar structure for example "A batch of goods M in the warehouse G is smuggled, thus all the goods in the warehouse G are smuggled".

We believe that argumentation schemes can be used in

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

collaborative intelligence analysis for structuring information to form evidence in terms of relations between entities, events and facts (highlighted in Figure 1). The service should assist users in selecting schemes that may fit the information available. Furthermore, we will investigate methods to aggregate and combine schemes to summarise interpretations of evidence for assisting analysts in building hypotheses. In this process critical questions play an important role. CQs may identify supporting relations among information or may highlight missing information such as premises that are necessary for a specific inference to be supported, but for which there is insufficient evidence. CQs make explicit links to alternative explanations, and may prevent confirmation biases by highlighting evidence against the most favourable hypotheses, ensuring a more objective evaluation. Similar to argumentation schemes, other structures have already been tested with existing tools for intelligence analysis. For example, TRELLIS [5] presents predefined constructs for asserting statements and associated reasons. However, argumentation schemes provide structure to those reasons that can be analysed by automated reasoning mechanisms to assist analysts in drawing conclusions. The defeasible of arguments is suitable to handle the dynamic intelligence analysis process, where new information is continuously being collected. An argumentation semantics can be used to assist analysts in deciding the status of the hypotheses.

Provenance. For effective collaborative analysis, analysts examining different information must consider *how, when, where* this information has been gathered and *by whom* it has been manipulated. Existing toolkits provide support for automatic or manual provenance annotation following data-model recommendations (e.g., PROV-DM [7]), and provide repositories for storing such metadata and querying services to extract relevant information [9]. We will employ such models to annotate the provenance of incoming data and the process of analysis. In particular, as shown in Figure 1, we will record provenance at different levels: *a)* the information level, where provenance of documents and information is stored, including sources and context of collection; and *b)* the analysis level, where provenance is a record of each phase of the analysis for a topic, including in particular sources and contributors, the creation of new schemes, data used, timestamps and updates.

This contextual information is used to retrieve the history of analysis and it must be integrated with the reasoning service in order for an analyst to better assess claims. In fact, understanding the workflow of analysis may affect conclusions and may lead to discard some hypotheses or evidence because they come from flawed reasoning processes. Argumentation schemes will be designed for relating inferences to, for example, the expertise of the source, when the claim was made, the temporal consistency of claims, expert reliability, trustworthiness of a source, and so on. Such schemes can be used as bridge between the reasoning process and the analysis of provenance records. To date, however, provenance within argumentation is yet to be explored. Furthermore, provenance has important applications in maintaining large volumes of heterogeneous data, but it has not been integrated within intelligence analysis tools. An introductory work is proposed in [13] for visual analytics. Although this work proposes a layered method to record analysis, here, we focus on the introduction of provenance into the reasoning process, not only as an external record that an analyst may consult. The analysis of contextual information contributes to the overall sense-making process in collaborative analysis.

Example. We introduce, here, an example to illustrate relations amongst information, how a scheme will support the generation of hypotheses, and how the introduction of contextual information affects the analysis. The goal is to establish the presence of any criminal activities across the border in a named area of interest, between locations L1 and L2 (L1-L2). The first scheme discussed is an *abductive argument from effects to cause* [12], which explores the causal relation between a set of facts F, and its plausible causal explanation C. Statements C and F may be collected from different sources, added by the analyst or suggested by the system. Here, the scheme is used by an expert of video recordings to state that some aerial images show that there is a gang G suspected of smuggling forbidden products P1 across the border L1-L2, and this is a possible explanation of why product P1 arrived in L2.

- **Premise 1** - The set of events $F = \{$"A: there is a forbidden product P1 in location L2", "B: a gang G smuggles products P1 in location L1"$\}$ has been recorded,
- **Premise 2** - "C: Smugglers G crossed the border L1-L2" is the best satisfactory causal explanation of F so far,

$\Rightarrow$ **Conclusion** - Therefore, C is plausible as the cause of F.

This scheme can be challenged with the question "How strong is the explanation C?" or "Is there any better explanation?". This partial analysis is shared with an analyst specialising in trafficking within the region. The new scheme is an *abductive argument from evidence to a hypothesis* [12]. Here, the hypothesis is that, since the smugglers crossed L1-L2, criminal activities have occurred at the border.

- **Premise 1** - If hypothesis "H: There are criminal activities at the border checkpoint L1-L2" is true then "C: Smugglers G crossed the border L1-L2" will be observed to be true,
- **Premise 2** - C has been observed to be true,

$\Rightarrow$ **Conclusion** - Therefore, hypothesis H is true.

A question that links to the previous analysis is "Has C been observed to be true?". Other questions may exploit alternative reasons for C being true; e.g., that the smugglers bypassed the checkpoint, thus the suspected activities are not associated with the checkpoint. Moreover, introducing information from provenance records about the reliability of the sources of F can invalidate the hypothesis if not supported by stronger evidence. The *argumentation scheme from expert opinion* may be used to relate a statement to its source [12]. Assume that the fact that product P1 was found in L2 has been asserted by E, an expert in identifying illegal products.

- **Premise 1** - Source E is an expert in domain "identification of products P1,
- **Premise 2** - E asserts that proposition "A: there is a forbidden product P1 in location L2" is true,

$\Rightarrow$ **Conclusion** - Therefore, A may be taken to be true.

In contrast with the others, this scheme permits challenges against both claim A and the association of A with source E. If there is evidence that leads the analyst to think that E is unreliable (e.g., through previous interactions), the claim can be rendered invalid. In fact, the new evidence attacks A and challenges the inferences C and H made on the basis of A.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

IV. Conclusion and Future Work

Effective collaborative intelligence analysis must be supported at the content level, in reasoning about information and formulating consistent hypotheses. We proposed an evidential reasoning service that guides analysts in structuring the process of analysis employing argumentation schemes, where partial elaborations can be shared more easily and an autonomous reasoning mechanism can be applied to support the sense-making. Recording provenance data permits one to assess the quality of information and analyses. The first step to develop this service is to identify an appropriate set of argumentation schemes and related critical questions to use for the analysis. These will be based on a computational language and a collection of ground facts for representing intelligence. We use existing schemes as a starting point [12], however such schemes may not be sufficiently expressive. We will design new schemes for representing relations amongst information by employing diagramming methods [12] for extracting schemes from corpora of intelligence reports. More importantly we aim to engage with experts in the field for validating and defining relevant schemes. A further question is how to include provenance data within the schemes, such as the temporal persistence of information, the workflow of analysis, the trustworthiness of data, and so on. We will identify argumentation semantics able to reflect how analysts consider some hypotheses acceptable, based for example on their expertise or on the reliability of the supporting evidence in order to ensure an objective evaluation.

We will explore methods for dealing with uncertain information that may come from untrusted sources or sources that have reported unreliable data in the past. We will study how this affects the construction of evidence and the acceptability of hypotheses. Furthermore, the information accessible by an analyst may be incomplete because many reports have limited distribution, clearance levels may dictate what information is available, and so on. Critical questions may help in identifying gaps and prompt the user to seek more information. Provenance may also help by disclosing some contextual information to ensure that if only limited analysis is shared an analyst may at least assess its reliability. However, these options may not always be allowed. Thus, we must study methods for addressing information gaps when sharing is limited. An important issue to be addressed is how users perceive the employment of argumentation schemes. Most of the work on argumentation is for autonomous reasoning, although such schemes have been used to analyse real world arguments [12]. Here, however, argumentation schemes are used in a system that interacts with users to assemble scattered information, and many issues may arise. For example, the support is not adequate for the inferences that the analyst intends to report. In deploying the reasoning service, we must consider what schemes should be suggested to the users, and which ones would best fit the information for providing valid support. We previously claim that our evidential reasoning service aims at improving the sense-making in collaboration at the content level, in order to assess such a claim we must perform experiments in order to test how effective our system is in identifying plausible explanations in comparison to the use of other tools (e.g., [1,10]) according to both the quality of the hypotheses identified and the ease of use of the system.

In conclusion, we believe that our evidential reasoning service based upon argumentation schemes and provenance is suitable to support collaborative intelligence analysis. This service will facilitate the integration of analyses from different contributors permitting a more objective evaluation of hypotheses and reducing the workload of analysts. Furthermore, tracking the use of data within the analytical process may help analysts in understanding the utility of information for planning more focussed collection activities.

Acknowledgment

The authors would like to thank Susan Toth at US Army Research Laboratory for the valuable feedback. Research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

[1] E. Bier, S. Card, and J. Bodnar, "Entity-based collaboration tools for intelligence analysis," Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, 2008, pp. 99–106.

[2] A. De Liddo, Á. Sándor, and S. Buckingham-Shum, "Contested collective intelligence: Rationale, technologies, and a human-machine annotation study," Computer Supported Cooperative Work, vol. 21, 2012, pp. 417–448.

[3] P. M. Dung, "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games," Artificial Intelligence, vol. 77, 1995, pp. 321–357.

[4] B. Fischhoff and C. Chauvin, Intelligence Analysis: Behavioral and Social Scientific Foundations. The National Academies Press, 2011.

[5] Y. Gil and V. Ratnakar, "TRELLIS: An interactive tool for capturing information analysis and decision making," Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, vol. 2473, Springer Berlin Heidelberg, 2002, pp. 37–42.

[6] Y.-a. Kang and J. Stasko, "Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study," Proceedings of the IEEE Conference on Visual Analytics Science and Technology, 2011, pp. 21–30.

[7] L. Moreau and P. Missier, "PROV-DM: The PROV Data Model," 2013, Available at: http://www.w3.org/TR/prov-dm/ [retrieved: May, 2013].

[8] N. Oren, T. J. Norman, and A. Preece, "Subjective logic and arguing with evidence," Artificial Intelligence, vol. 171, 2007, pp. 838–854.

[9] E. Pignotti, P. Edwards, A. Preece, N. Gotts, and G. Polhill, "Enhancing workflow with a semantic description of scientific intent," The Semantic Web: Research and Applications, vol. 5021, Springer Berlin Heidelberg, 2008, pp. 644–658.

[10] N. Pioch and J. Everett, "POLESTAR: collaborative knowledge management and sensemaking tools for intelligence analysts," Proceedings of the 15th International Conference on Information and Knowledge Management, 2006, pp. 513–521.

[11] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," Proceedings of the International Conference on Intelligence Analysis, 2005.

[12] D. Walton, C. Reed, and F. Macagno, Argumentation schemes. Cambridge University Press, 2008.

[13] B. Wong, K. Xu, and S. Attfield, "Provenance for intelligence analysis using visual analytics," Proceedings of the CHI Workshop on Analytic Provenance, 2011.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Privacy Policy Negotiation Architecture for Pervasive Computing Environments

Nidal Qwasmi, Khalil El-Khatib, Ramiro Liscano, Julie Thorpe
University of Ontario Institute of Technology
Oshawa, Ontario, Canada

nidal.qwasmi@uoit.ca, khalil.el-khatib@uoit.ca, ramiro.liscano@uoit.ca, julie.thorpe@uoit.ca

Abstract— **In a highly connected and collaborated world, privacy is always a pressing issue. New technologies, such as cloud computing and service-oriented architecture, made system collaboration possible, which, in return, made privacy and security a challenge. This research paper focuses on designing a comprehensive technical architecture to negotiate and preserve privacy policy in pervasive computing environments. The architecture addresses the problem of user's privacy by developing dynamically controlled mechanisms that would allow the inhabitants of the environment to control what information is collected about them, how it is being processed, and under what circumstances it can be shared. The architecture allows users to have more control and negotiate their privacy policy as well as resolve any conflict that may occur during normal operations of the system.**

Keywords- APPEL; privacy; architecture; conflict; P3P; resolution; exchange; negotiation.

I. INTRODUCTION

Establishing and maintaining individual privacy policies for Internet users remains unpopular mainly due to the technology limitations such as the complexity of Platform for Privacy Preferences (P3P), Preference Exchange Language (APPEL) and the "take it or leave it" approach for the web which make any privacy solution deployment a challenge. However, the Platform for Privacy Preferences (P3P) and preference language called APPEL [1] are considered the most recognizable efforts to enable users to have control over their privacy while online. P3P became a W3C recommendation in April 2002 [2] while APPEL became a working draft on April 20th, 2000 [2]. P3P and APPEL interact to enable a machine to programmatically check and apply user's privacy preferences and policy. Unfortunately, P3P and APPEL do not interact very well due to fundamental design choices that result in serious problems when using APPEL to define user's preferences [3]. Deficiencies with APPEL are mainly because users can only define what is unacceptable, not what is acceptable. In addition, APPEL rules are hard to express which makes it inefficient and tedious [3].

Negotiation mechanism was a working item for W3C in their early drafts of the P3P specification, but that initiative had been abandoned in support of easy and quick implementation. Negotiation mechanisms are still not part of the latest version of the P3P 1.1 specification [2]; however, it is planned for future versions of P3P [4]. Policy conflicts may cause serious privacy breaches and vulnerabilities such as blocking legitimate operations or permitting unwanted operations, hence privacy negotiation becomes a very important requirement for any privacy architecture in practice. Many real life examples demonstrate the importance and usefulness of negotiation features to any privacy architecture, for example:

- In medical environments [5], where a patient negotiates his own privacy policy with his nursing service provider.
- In distance education [6], where, for example, a trainee negotiates the access limit to his education record with an education center.
- In online retailing [7], where a customer negotiates how his address information would be handled with an on-line bookstore.

Our contribution in this research is presenting a new design of a privacy architecture for pervasive environments that supports negotiation, utilizing available technologies (P3P and APPEL). Adapting P3P and APPEL for a pervasive computing environment requires fundamental changes to P3P XML policy structure as well as its reference language APPEL. Some of these changes are necessary to accommodate the special needs of the environment; other changes are dictated by the basic system requirement of enabling conflict negotiation and resolution. Our suggested changes to P3P and APPEL are discussed in Sections III and IV. The rest of the paper is organized as follows. Section II presents related work. In Section III, we describe privacy policy for pervasive computing environment. In Section IV, we discuss reference language for pervasive computing environment. In Section V, we present our proposed architecture for privacy policy negotiation. Section VI presents our research conclusion and future work.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

II. RELATED WORK

There are many previous research papers related to the topic of privacy protection in pervasive computing environments. Here, we describe these papers and how they differ from our present paper.

Shankar et al. [8] presented the necessity of having a policy-based management architecture, which is capable of handling conflict between the policies within the system. The proposed architecture is based on the concept of Event-Condition-PreCondition-Action-Post Condition (ECPAP). To accomplish this task, the authors proposed providing an axiomatic specification of rule actions to define the status of the impacted system before and after the execution of a policy. The authors in this research acknowledged the issue of policy conflict and proposed an approach to detect the conflict; however, they did not propose any viable way to dynamically resolve this conflict.

Langheinrich [9] provided comprehensive information about ubiquitous computing environments, which covers background analysis, privacy mechanisms and principles, privacy awareness systems, and related work. The authors also proposed a conceptual privacy system architected for ubiquitous computing environments, which is based on the following three main concepts:

1. An upfront announcement of a system's privacy policies.
2. An automatic configuration of available services based on the user's privacy policy.
3. An enforcement of privacy policy usage on the stored and collected information.

Langheinrich's work inspired us in designing our proposed system as it provides a set of privacy principles. However, this research did not address the issue of conflict between privacy policies and preferences.

Babbitt et al. [10] present the necessity of having a privacy management system (PMS) in ubiquitous computing environments where inhabitants can have control over their privacy. To accomplish this task, the authors proposed to create a conceptual model for a privacy management system, which can be used in the implementation process of any smart home environments. This research highlights many privacy issues within the ubiquitous environment.

Lahlou et al. [11] explained the gravity of infringing upon an individual's privacy. They also touched on the reasons behind taking privacy issues lightly by technical professionals. The authors also present the European Union's approach to enhance individual's privacy within ubiquitous computing systems. This research presents real life cases of privacy violation as well as the importance of having a privacy negotiation mechanism, which can help avoiding any unwanted consequences of privacy violation.

El-Khatib [7] has discussed a conceptual privacy negotiation protocol (PnP) protocol for web services. The author explained the benefits of having a privacy negotiation protocol, and presented an architecture for a negotiation protocol that can be used with web services. To implement the suggested protocol, the author envisions using P3P and APPEL with modifications. Our present paper differs from [7] in the targeted environment, as we envision our proposed architecture to be implemented in client-server rather than B2B (Business-To-Business) environments.

Kapadia et al. [12] presented a model for protecting user's privacy by giving users control over the dissemination of their digital footprint such as location or other private habits. Their model is based on building virtual walls, which simulate the actual live physical walls and follow the same concept of real life privacy protection mechanisms. The proposed virtual walls model is based on three levels of transparency, which are: transparent, translucent, and opaque. Users in this model can choose any level of transparency, which translates into a level of their own privacy protection. Authors in this research designed pre-determined levels of privacy and assumed that it will be applicable to all users. We believe that user privacy preferences are varied and it is too complex to be summarized in a limited number of choices.

Schmandt et al. [13] listed many real life examples and issues to support the importance of protecting an individual's privacy. Some of these examples are surveillance, grocery shopping in an intelligent environment, exchange of personal information over the Internet. The authors anticipated a controversial issue in the case of users having an interface to control their own privacy that is "what is the level of granularity at which users wish to control what is revealed about them. And how those choices are expressed" [13].

Babbitt et al. [14] proposed a privacy system model based on an authorization model, the administration model, the domain model, and the performance model. The authorization model is responsible for controlling how users, services, devices, and policies act and interact. The administration model is responsible for monitoring the changes that users need to apply on the authorization model. The domain model is responsible for controlling the context of where the user is trying to access the system, as different domains would need different privacy policies. The performance model is in charge of overseeing the efficiency and scalability of the privacy systems. This model does not deal with policy conflict or negotiation.

III. PRIVACY POLICY FOR PERVASIVE COMPUTING ENVIRONMENT

In order to adapt P3P specifications into a pervasive environment, some modifications are necessary on the privacy policy file structure (XML); these modifications are mainly to accommodate the idea of communicating simultaneously to multiple devices with different policies while providing the capability of automated or semi-automated negotiation between a system and its clients.

Fig. 1 shows the existing P3P policy file structure, while Fig. 2 shows our modified version of the P3P policy file structure. The difference is that we have added an

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

additional segment for negotiation and changed the existing "data segment". The proposed changes are explained here for each segment:

A. Negotiation-group Segment:

This new segment is mandatory. The purpose of this new segment is to allow for automated or semi-automated negotiation between a system and the system's clients. In the normal state of the system, this segment should be free of any conflict elements but if conflict exists, then either side (system or client) may populate the resolution part of the segment, depending on which side had initiated the conflict. In this case, one of the two sides (system or client) either accepts the resolution or rejects it. When any of the two sides (system or client) rejects the resolution(s) then the receiver side may propose another resolution for the conflict. The negotiation process would continue until mutual agreement has successfully concluded or connection ended (time-out).

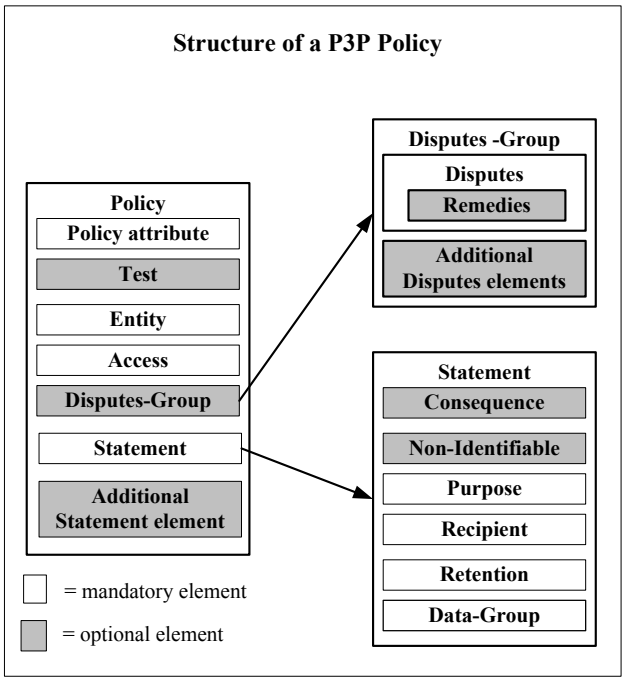

Figure 1. P3P XML policy structure [15]

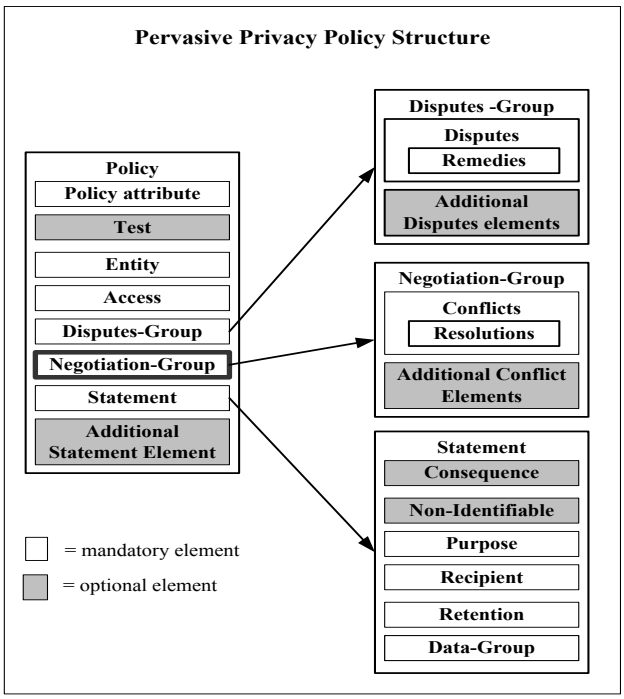

Figure 2. Privacy policy structure for pervasive computing

For our proposed model to work, we need a simple rule-set language which can be used to set and control agent's rules, conflict and resolution. The Extensible Stylesheet Language family (XSL) [18] would be an ideal tool to serve our model so the system can automatically negotiate privacy policies for the client's device(s). We discussed available language options in more detail in Section IV. This segment will also show the importance of resolving the policy's conflicts as well as providing a means to inform the client of any privacy policy violation and available paths to resolution. At the same time, this change would serve the purpose of increasing the level of client privacy by informing the user right from the beginning about any privacy policy conflict or violation and engaging the system's automated negotiation process in order to resolve the conflict according to the client's rules without much involvement from the client or disruption to the device's normal activities. If applicable, the negotiation segment would also contain data about the conflict resolution process which has not been successfully resolved by the negotiation process. Example of such data could be about other alternatives to resolve a conflict, which still exists after the end of system automated negotiation.

B. Data-group Segment:

This group is expanded into three sub-segments: the first is for mandatory device information, the second is for personal information, and finally the third is for optional additional data that needs to be collected.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

IV. REFERENCE LANGUAGE FOR PERVASIVE COMPUTING ENVIRONMENT

APPEL is privacy reference execution language which is currently in use for privacy over the Internet; however it has been retired [16]. Another alternative is to use the XPref [3] and the Enterprise Privacy Authorization Language (EPAL) [17]. EPAL is a language for writing enterprise privacy policy and exchanging it in a standard format between applications or enterprises [17]. The main purpose of EPAL is to control data usage in the enterprise system according to fine-grained positive and negative authorization rights [17]. EPAL is very useful for enterprise IT systems but provide little or no benefit for client side interface where policy negotiation should take place, therefore EPAL with its features and capabilities is not suitable for client side interface.

XPref [3] is a script language based on XPath [19] and borrows some of its syntax and semantics from APPEL. The main goal of designing XPref is to overcome the deficiencies in APPEL, which are explained in [3]:

- Allows specifying unacceptable rules, but not acceptable rules.
- Rules are constrained with limited number of logical expressions that could be provided in a single rule.

Despite the good features and great flexibility that comes with the XPref language, it is still not capable of handling the negotiation part of the system. Therefore, our search extended to other alternatives that may satisfy our system requirements. The first step in selecting any programming language is to define the requirements of the system we intend to use it for. Accordingly, our system requirements are: XML based language, mature with international standard, easy to use, simple syntax, can accept XML file as an input and generate an XML file as an output, and lightweight so it can be installed on devices with low computation power such as PDA and cell phones.

To implement the system, our search concluded by selecting The Extensible Style-sheet Language Family (XSL) [18], more specifically XPath [19] and XSLT [18]. The reasons behind our selections are:

1. XSL is mainly designed for XML based applications.
2. It is a mature language as it is adopted by W3C.
3. It has a simple and easy syntax.
4. It can transform any XML file into another XML file (same or different structure), HTML, or XHTML
5. XSL client can be a standalone or can be add-on to a web browser. Most of the Internet browsers are supporting XLS. Firefox from version 1.0.2, Mozilla, Netscape from version 8 (uses the Mozilla engine); Opera from version 9 and Internet Explorer from version 6 all have supported XML and XSLT (and CSS). If a web browser add-on is being used then XSLT and XML files can be in separated files and implemented directly or "javascript-ECMAScript-262" [20] can be used to process the document.

The Extensible style-sheet language family consists of the following three main components.

- XSLT: language for transforming XML files [18].
- XPath: an expression language used by XSLT to navigate through an XML document [19].
- XSL-FO: an XML vocabulary for specifying formatting semantics [18].

The main objective of using XSL is to analyze the XML file, which carries the privacy policy, apply system's or client's privacy preferences, and to resolve any privacy policy conflict. If a conflict occurred then a new XML file should be generated with its conflict segment been populated. Both side (client and system) should be able to receive, process and generate an XML file.

As shown in Fig. 3, XML standards will be used to create privacy policies (system and client) while XSLT will be used to express privacy preferences and resolve any conflict that may have occurred. As we mentioned previously, XSL can transform any XML file into another XML file, which may be (or may not) the same structure or it can transform it into HTML/XHTML file. Also, XSL uses XPath to access any part of the XML file and it uses XSL-FO to format/control the presentation layer. These features become very valuable when a user interface is needed such as when a user needs to be prompted for a decision.

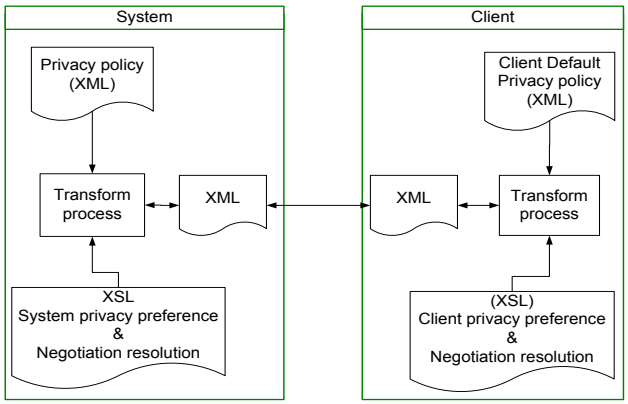

Figure 3. Privacy policy processes

V. PRIVACY POLICY NEGOTIATION

Steven L. McShane defines negotiations as a "process occuring whenever two or more conflicting parties attempt to resolve their divergent goals by redefining the terms of their interdependence" [21]. Although this definition was in the context of human behavior, it is also applicable to privacy policy negotiation. If we take a closer look at the definition, then we find that it has three keywords that can summarize the whole process of negotiation; they are conflicting, redefining, and resolve. Indeed, it required a conflict to occur between two or more parties about something in order to start a negotiation process, which is

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

using redefining techniques of terms and conditions to achieve the ultimate outcome of the process that is resolving the conflict. Based on this definition, a basic pervasive policy negotiation process might proceed as follows:

1. The server system tries to automatically access the client's privacy policy. The client can store the privacy policy in a pre-determined location (public access) on the device where the system can access it automatically. In the absence of the client's privacy preferences policy, the system will send its own privacy policy file (XML file format).
2. If the client's device has a publicly accessible privacy policy preference file, then the system tries to apply the client's privacy policy preferences. In the case of any conflict, the system populates the conflict and resolution segments (if applicable) in an XML file and sends it back to the client. In the absence of privacy preferences file from the client device, the server system will send its own privacy policy file. Upon receipt, the client's device using the reference language (XSL) will try to apply the client's privacy rules, fill the conflict and resolution segments (if applicable), and then send the XML file back to the server.
3. Normal operation status required that no conflict should exist.
4. If conflict existed, then either side would automatically evaluate the resolution(s). In this case either side has three options:

a) Accept the resolution and change the privacy policy accordingly.

b) Add a new resolution. Here it implies the other side had rejected the suggested resolution and would like the client to suggest another one.

c) Reject the resolution. In this case it means either side doesn't agree to the service agreement and would like to walk away, which also require prompting the user about the conflict and asking for consent to leave the system service area.

5. After a pre-determined number of automatic negotiation cycles, the client will be prompted for a decision.

Fig. 4 shows the message sequence for exchanging the privacy policy between the client and the server system. Our assumption in this scenario is that the client doesn't allow public access to the privacy policy file.

Fig. 5 shows a graphical representation of our previously described process of privacy negotiation. The left side of the graph shows the activities taking place in the system side of the process (service provider). In the right side of the graph, it shows the activities taking place in the client side of the process (service consumer). Our graph clearly shows that links between the two sides are the transactions of exchanging privacy policy file (XML file). The process of applying the privacy preference and negotiation is completely dependent on the Extensible style-sheet language family; mainly XSLT and XPath.

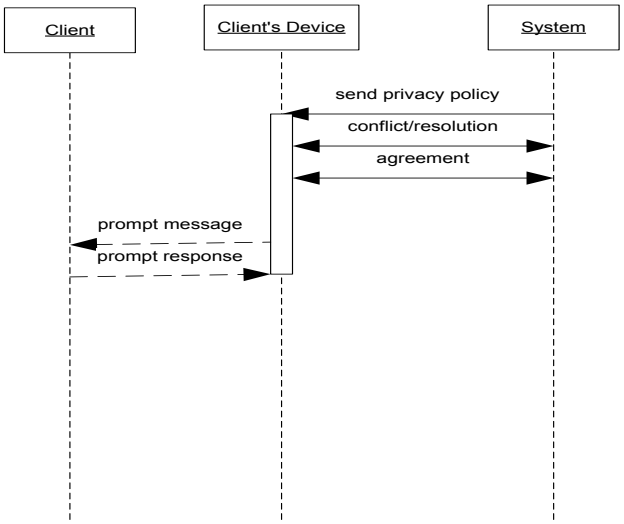

Figure 4. System sequence for accessing privacy policy

VI. CONCLUSION AND FUTURE WORK

Our new architecture adapts P3P and APPEL for pervasive computing environments, which requires some fundamental changes to the P3P policy file structure as well as its reference execution language APPEL. Some of these changes are necessary to accommodate the environment special information needs. Other changes are dictated by the basic system requirements of enabling conflict negotiation and resolution.

Our proposed architecture requires changes to the P3P structure as well as its script reference language. Fig. 2 shows our modified version of the P3P policy file structure. As we have seen from previous sections of this research, APPEL is not even sufficient for static privacy policy needs. Our other alternatives we looked at into were XPref [3] and EPAL [17].

Our research concluded by selecting The Extensible Style-sheet Language Family (XSL) [18] more specifically XPath [19] and XSLT [18]. The reasons behind our selections are, first it is an XML based implementation, second it is a mature language, third its simplicity, fourth it can transform any XML file into another XML file (same or different structure), HTML, or XHTML, and fifth is XSL agent is flexible to implement.

With all the efforts that have been put into this research but we think that it is still coming short to cover some important areas of the privacy system for pervasive computing environment; especially the following two topics:

- Interface Agent: This research topic focuses on communicating between users and XLS. It covers the issues of converting user's privacy preferences to XLS rules. Although XLS syntax is easy to understand by technical professional but it is still very inconvenient for non-technical person to compose his preference file. Our visionary solution for this requirement is to be a Java

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

standalone application utilizing standard web browser as presentation engine.

- Communication Agent: As we previously described in this research, the privacy system depends on exchanging privacy policies, which are an XML files. The need for communication agent is to facilitate, organize, and control the process of exchanging the privacy policy files between negotiated parties.

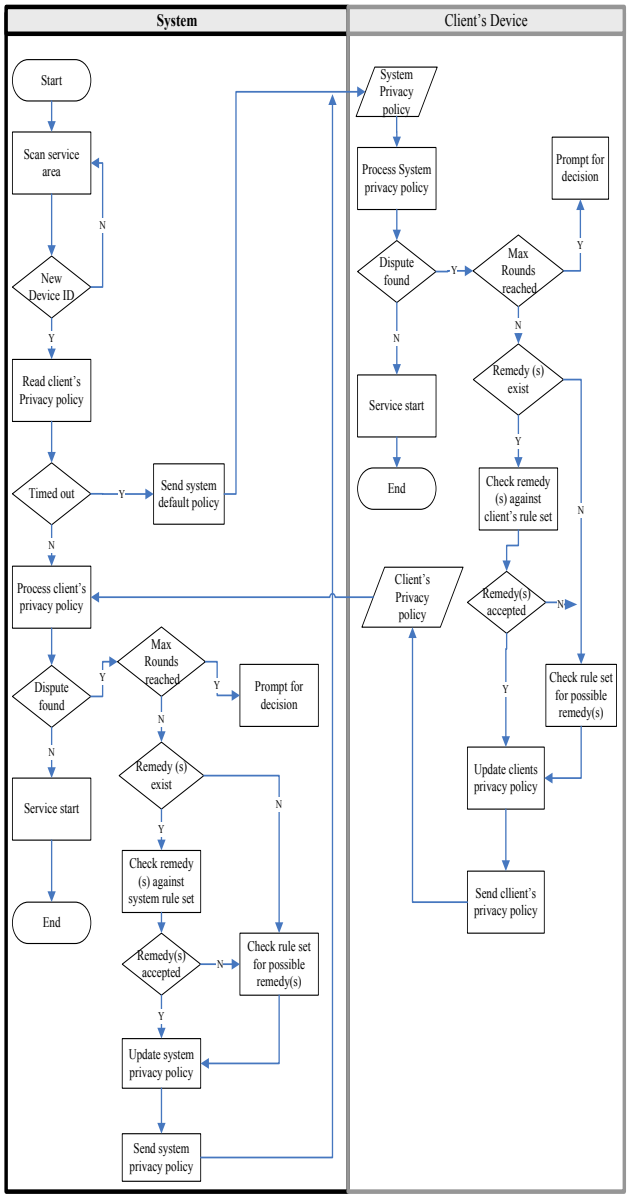

Figure 5. System negotiation flowcharts

REFERENCES

[1] A P3P Preference Exchange Language 1.0 (APPEL1.0) http://www.w3.org/TR/P3P-preferences/ [retrieved: 04, 2013]

[2] The Platform for Privacy Preferences 1.0 (P3P1.0) Specification http://www.w3.org/TR/P3P/ [retrieved: 04, 2013]

[3] Agrawal, R., Kiernan, J., Srikant, R., Xu, Y., "XPref: a preference language for P3P", science direct-Computer Networks 48 (2005) 809–827, 16 February 2005

[4] W3C, P3P – Future Versions of P3P (Presentation, slide 10), http://www.w3.org/Consortium/Offices/Presentations/P3P/Overview.html#(10) [retrieved: 04, 2013]

[5] Yee, G., Korba, L., "Feature Interactions in Policy-Driven Privacy Management", Proceedings, Seventh International Workshop on Feature Interactions in Telecommunications and Software Systems, Ottawa, Canada, 2003

[6] Yee, G., Korba, L., "the Negotiation of Privacy Policies in Distance Education", Proceedings of the 4th International IRMA Conference. Philadelphia, Pennsylvania, USA. May 18-21, 2003.

[7] El-Khatib, K, "a Privacy Negotiation Protocol for Web Services", Workshop on Collaboration Agents: Autonomous Agents for Collaborative Environments Halifax, Nova Scotia, Canada. October 13, 2003

[8] Shankar, C., Campbell, R., "Policy-based Management Framework for Pervasive Systems using Axiomatized Rule-Actions", Fourth IEEE International Symposium on Network Computing and Applications, July 27-July 29 2005

[9] Langheinrich, M., "Personal Privacy in Ubiquitous Computing - Tools and System Support." Zurich, Switzerland: ETH Zurich, 2005.

[10] Babbitt, R., Information privacy management in smart home environments: Modeling, verification, and implementation. Proc. of 30th Annual Int. Computer Software and Application Conf. COMPSAC 2006, vol. 2, 2006, pp. 344-346.

[11] Lahlou, S., Langheinrich, M., and Ro¨cker, C., "Privacy and Trust Issues with Invisible Computers," Comm. ACM, vol. 48, no. 3, pp. 59-60, 2005.

[12] Kapadia, A., Henderson, T., Fielding, J., Kotz, D., "VirtualWalls: Protecting Digital Privacy in Pervasive Environments", In Proc. of the 5th Int'l Conf. on Pervasive Computing, 2007

[13] Schmandt, C., Ackerman, M., "Personal and Ubiquitous Computing Issue on privacy and security", Journal Personal and Ubiquitous Computing, Volume 8 Issue 6, November 2004

[14] Babbitt, R.; Wong, J.; Chang, C.; , "Towards the Modeling of Personal Privacy in Ubiquitous Computing Environments," Computer Software and Applications Conference, 2007. COMPSAC 2007. 31st Annual International , vol.2, no., pp.695-699, 24-27 July 2007

[15] Cranor, L. F., "Privacy Policy, Law, and Technology", http://lorrie.cranor.org/courses/sp04/p3p-slides.pdf , p 19, [retrieved: 04, 2013]

[16] W3C http://www.w3.org/TR/#tr_P3P [retrieved: 04, 2013]

[17] Enterprise Privacy Authorization Language (EPAL 1.2), http://www.w3.org/Submission/2003/SUBM-EPAL-20031110/ [retrieved: 04, 2013]

[18] The Extensible Stylesheet Language Family (XSL) http://www.w3.org/Style/XSL/ [retrieved: 04, 2013]

[19] XPath http://www.w3.org/TR/xpath20/ [retrieved: 04, 2013]

[20] ECMAScript, http://www.ecmascript.org/docs.php [retrieved: 04, 2013]

[21] McShane, S., "Canadian Organization Behaviour", fifth edition, McGraw-Hill, 2004

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

MobWEL - Mobile Context-Aware Content-Centric Workflow Execution Language

Anna Kocurova, Samia Oussena, Peter Komisarczuk
School of Computing and Technology
University of West London
London, UK
Anna.Kocurova, Samia.Oussena, Peter.Komisarczuk@uwl.ac.uk

Tony Clark
School of Engineering and Information Sciences
Middlesex University
London, UK
Email: T.N.Clark@mdx.ac.uk

***Abstract*—Collaboration among people has been empowered by using mobile devices in daily life. However, increasing user demands for a better mobile collaboration experience require constant evolution and adaptation of existing mobile collaborative technologies. In this paper, the collaborative workflow technology is adapted to enhance mobile peer-to-peer collaboration. With context awareness integrated, workflows are adapted to collaborators' needs and circumstances. Extra management support for content behaviour is incorporated in order to increase content awareness in workflows and enable communicating progress among collaborators. This paper introduces MobWEL, a context-aware content-centric workflow language designed for mobile peer-to-peer collaboration. MobWEL extends BPEL, using constructs from existing workflow approaches, Context4BPEL and BPEL^{light}, and adopting elements from the BALSA workflow model.**

***Index Terms*—mobile; peer-to-peer; context-aware; content; workflows.**

I. INTRODUCTION

Mobile phones have become ubiquitous, opening new possibilities for collaboration, communication and sharing ideas in real-time. Using smartphones allows people to remain productive regardless of their geographical locations and collaborate while on the move. However, with the spread of mobile technologies, collaborators expect to have tools that support collaboration, content sharing and management. The raising user expectations for mobile collaboration include the demand for an intelligent mobile collaborative environment which responds to people's needs.

Collaborative workflow is a widely accepted technology which offers rich support for multi-party collaboration, coordination of distributed teamwork and content sharing. With the arrival of mobile computing on the scene, mobile devices have been integrated into workflow management scenarios and the technology has been adapted for various sets of requirements. Smanchat et al. [1] states that when workflows are utilised in ubiquitous environments, adaptability and context-awareness are the features that should be included in the workflow mechanism. Earlier workflow management approaches based on a server-client model required centralised workflow management systems deployed on servers. However, this centralised approach is impractical in various situations in which people prefer to collaborate solely by using mobile devices. A typical example of such a situation can be a meeting between business partners who share private information. Moreover, in the centralised architectural style, control and management of workflows is driven by a central management unit with mobile devices behaving only as thin clients. The central management unit makes all important decisions and drives the whole workflow execution without considering the current situation of users, and the context the devices reside in. In this work, Dey's definition of context is used: 'context is any information that can be used to characterize the situation of an entity' [2].

The need for a mobile device-centric workflow process and systems that manage the workflow processes in a completely distributed manner has been recognised in the work of Pajunen and Chande [3]. Mobile collaboration in the distributed manner, also labelled as mobile peer-to-peer collaboration, enables data and content distribution by direct exchange. Additionally, all decisions can be made by mobile devices and based on local, context information, thereby adapted to current collaborators' circumstances and needs. The work presented in this paper tackles the problem associated with representing context information available on a single mobile device, and using context information to shape peer-to-peer workflow execution. Context describes the current situation of a user, a device or an environment regarding a specific purpose. By expressing the context model explicitly for each workflow scenario, each mobile device can monitor, acquire and process only required context information. By adapting workflow logic to consume context information, the workflow execution can become more dynamic and responsive to an individual collaborator's situation.

Additionally, workflow processing often involves content manipulation such as reviewing and approving proposals or pictures. Mobile content such as picture, document or video/audio is usually user-generated or adapted for use on mobile devices. Mobile content is described by metadata and can be semantically enriched by context metadata such as location where it was created. Creating, processing and disposition of content are the basic stages of its lifecycle. However, content can flow through a more complex management process involving stages such as content editing, reviewing or approving. Workflow process activities modify content or

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

its metadata. By way of illustration, if a proposal has been approved, its metadata specifying the state of approval has been changed, the proposal is considered as being in the 'Approved' state and often shortly referred as 'approved'. So at any given instant of time, content is in a specific state that is defined by values of its metadata. There are several benefits to increasing content awareness in mobile workflows and integrating advanced content management capabilities in the workflow management system. Firstly, content behaviour is visible and its lifecycle is expressed as explicitly as the process logic is. Secondly, especially in peer-to-peer collaboration, content state-related information can be used significantly in communicating progress among collaborators. Dissemination of such information enables easier content synchronisation over a number of devices. Furthermore, there can be relationships and dependencies between various content items. For example, the proposal can have more sections edited by co-workers on their devices. Receiving regular information about each section state would enable faster processing of the proposal.

This paper develops the ideas presented in our previous work [4] and presents our solution for increasing context and content awareness in the collaborative workflow technology. In particular, a mobile context-aware content-centric workflow execution language (MobWEL) is introduced and an architecture of a workflow management system that carries out MobWEL workflows is described. The remainder of the paper is structured as follows. Section 2 presents the usage scenario. An overview of related work is given in Section 3. Domain analysis and the constructs of the MobWEL language are described in Section 4. The MobWEL workflow language is presented in Section 5. Section 6 outlines the logical architecture of the MobWEL workflow management system. This paper is concluded in Section 7.

II. Mobile Peer-to-Peer Collaboration Scenario

The concept introduced in this paper is illustrated by using the following usage scenario. A team of ten designers work on interior design of buildings. Designers often work out in the field using smart phones to communicate and directly share pictures. Although each design project is assigned to a particular designer, design decisions are never done by a single person. The following work pattern is used to complete projects:

1) Designer Jane takes a picture of a new room design by using her smart phone.
2) A simple rating system is used to quickly assess design ideas. Jane adds her own rating to the picture.
3) The picture is sent to her fellow co-workers. They work out in the field so each of them can review the picture by using own mobile phone.
4) Reviewer's subjective opinion can be captured by adding a comment.
5) Reviews and comments are sent back to Jane. She finally reassesses her idea according to opinions of other designers.
6) If the idea is good, the picture is sent for final approval to customer.
7) Approved picture is added to Jane's completed work.

The real work pattern can be abstracted into a collaborative workflow. However, the workflow is not adapted to individual collaborator's needs. Context and content awareness could enrich the workflow, as shown in Fig. 1.

The simplified workflow process is illustrated from a user's point of view in the middle of the figure. It shows three roles that participate in this collaboration: Interior Designer, Reviewer and Customer. The lifecycle of picture is outlined on the left side. The picture can go through a number of states such as *Initial, Reviewed, Assessed, Approved or Final*. Tasks

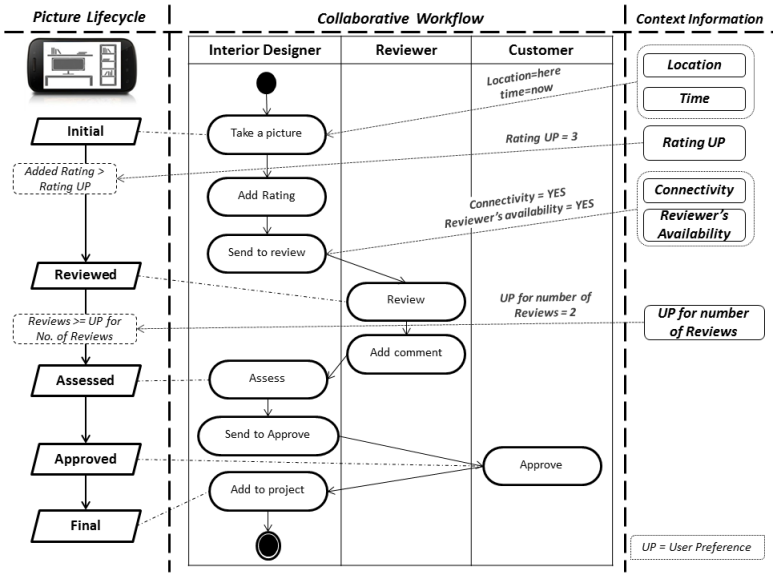

Fig. 1. Collaborative Workflow

or activities that can influence a change of the picture state are marked. For instance, if the *Review* task is performed, the picture can move to the *Reviewed* state or if the *Approve* task is performed, the picture may go to the *Approved* state. Relevant context data are shown on the right side of the figure. Context information can be used in several ways. Firstly, it can enrich picture such as *location* can specify the place where the picture is taken. Secondly, Jane can specify that only pictures with rating larger than her *Rating User Preference (RatingUP)* can move from the *Initial* state to the *Reviewed* state. So if the *RatingUP* is set to 3 and the rating added to the picture is 4, then the picture can be sent to be reviewed. Finally, knowing work context and current availability of fellow co-workers would enhance the mobile collaboration. For instance, the picture can be sent only to those co-workers who are currently available to review it (*Reviewer's Availability = YES*).

III. Related Work

A. Workflow Contextualisation

Workflow contextualisation has been addressed in a number of works [5][6][7], however, there are many research challenges to make context-aware workflow systems ready for practical use. Three main research challenges related to a context-aware process design have been recognised by Rosemann et al. [8]: context description, design for context and process adaptation. *Context description* refers to the identification and description of context variables relevant for a business process, respectively to a relevant context model. This includes the integration of the context model with an existing metamodel of a process modelling language. *Design for context* refers to the incorporation of contextual elements in the design of business process. In particular, how the knowledge can be embedded and utilised in the process design for context. Finally, *Process adaptation* considers the support for context-aware business processes.

Although the need for an explicit context description and identification of context variables that influence process design and execution has been highlighted, using context in applications is difficult because of the nature of context information. Numerous general context frameworks have been already developed to facilitate context modelling, recognition, reasoning and management. One of the earlier widely acknowledged works, the Context Toolkit introduced by Salber et al. [9], provides a framework for context modelling and management based on context widgets. Focus on mobile platforms aiming to simplify the development of context-aware mobile applications has been addressed in another context management framework proposed by Korpipaa et al. [10]. The framework employs the context manager to provide control between the acquisition and use of context information by applications. One of the context management approaches introduces a preference context model for representing context-dependent application requirements [11]. Based on the existing frameworks, a summary of the requirements, which should be met in the development of context modelling and reasoning techniques, has been elaborated by Bettini et al. [12]. An original and holistic view of the existing approaches for context data distribution for mobile ubiquitous systems has been undertaken by Bellavista et al. [13].

There have been various context management frameworks developed. However, the frameworks are either too general and designed to be used with a wide range of context-aware applications, or the frameworks are designed for a specific class of scenarios and a particular set of requirements. However, we believe that the context modelling and management approach need to be adapted for the problem it is applied to. Thereby, to provide different types of context information described in the usage scenario, a layered, self-contained approach to context acquisition has been designed and described in our previous work of Kramer [14]. In this paper, the use of context information is also considered, thereby, the approach is extended and the extended approach to a workflow-specific context representation and management is described.

Following the same vision that workflow meta-models should support context modelling and its use in workflows, a context-aware workflow process model has been developed by Wieland et al. [15]. The implemented process model based on Business Process Execution Language (BPEL) has been named Context4BPEL. Context4BPEL enables modelling of context-aware workflows directly without hiding the context usage in the invoked web services. The approach is based on the context-related concepts such as context event, context query and context decision. Context4BPEL workflow management system has been coupled with the Nexus context-provisioning platform, therefore, the use of the concepts in Context4BPEL follows the Nexus access pattern. The concepts have been adopted in our workflow language. However, the Nexus provisioning platform is not suitable for our scenario as it has been developed as a framework that supports a global context model, as described by Grossman et al. [16]. In our work, only local context situations of mobile devices need to be monitored, thus the context provider should reside on each device. It means that the Context4BPEL context-related concepts have been adapted in order to be used in alignment with our context provisioning platform and proposed workflow-specific context management approach.

B. Object Awareness in Workflows

So far, however, too little attention has been paid to content processed in process-centric workflows. Many workflow approaches are organised around process-centric and activity-centric workflow models, which are relatively flat with limited focus on artifacts processed in workflows. Some data-centric and artifact-centric workflow models have been developed, however, Hull et al. [17] states that the field of artifact-centric workflows is still in infancy. The need for more data-driven and object-aware workflow processes has been recognised in the research community and various approaches, presented in this section, have been developed to outline or address the associated challenges.

The belief that traditional process modelling approaches which focus on activities fail to capture informational structure

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

relevant to the business contexture has been supported by Liu et al. [18]. Business artifacts, such as Purchase Order or Insurance Claim, are seen as an additional dimension with which business analyst can model their business. A business artifact has been described as *'identifiable, self-describing unit-of-information through which business stakeholders add value to the business'*. Therefore, the artifact has an id to be uniquely identified within the business and its attributes are so named that their use within the domain is apparent. This approach highlights the need for discovering and modelling of the artifact behaviour.

The issue that entity, such as content or business artifact, is integrated in workflows as an input or output of an activity and the effects of how performed activities influence entity's behaviour are not visible, motivated the development of the Business Artifacts with Lifecycles, Services and Associations (BALSA) workflow model [19]. The artifact-centric workflow model comprises four key elements or dimensions: business artifact information model, business artifact (macro-level) lifecycle, services, and associations. By varying the paradigms used to specify the information model, lifecycle, services and associations, numerous BALSA models can be obtained. For example, the information model might be specified as attributes with scalar values or XML; the lifecycle might be specified by using flowcharts or finite state machines; and the services might be specified in black box or BPEL activities. The choice among the various paradigms depends on the intended area of application. Therefore, the BALSA workflow model is flexible in its use and provides a guided framework how to develop a customised artifact-centric workflow approach. This motivated us to use some aspects of the BALSA model in building the MobWEL workflow approach for mobile platform.

Another framework for integrated process and object life cycle modelling has been developed by Wahler [20]. Business objects processed by business processes can be associated with distinct states abstracted from the details of the performed tasks. The states mark the milestones of the overall processing, and are useful in communicating progress to stakeholders who are unaware of the exact process logic. Unlike other approaches which base entity lifecycles on variants of finite state machines, Hull et al. [21] introduces the guard-stage-milestone lifecycle model. The model is an evolutionary work on previous approaches based on business entities with lifecycles, but is more declarative than finite machines and supports hierarchy within a single entity instance.

Object-awareness in process-centric workflows is still very limited, and a holistic approach to integrate data, processes and users is undertaken in the development of the PHILharmonicFlows framework by Künzle et al. [22].

Most of the presented studies support the use of a process model and an object model as two complementary assets in object-aware workflows. However, research to date has tended to focus on the lifecycle of the artifact without considering its adaptation to context changes and its evolution in peer-to-peer workflow execution. This motivates the need for MobWEL workflows to support context-awareness in both models and be designed for peer-to-peer collaboration.

IV. Domain Analysis

The objective of this work is to enhance mobile peer-to-peer collaboration by developing context-aware content-centric workflows. Such workflows can be described by a workflow language that possesses the following main characteristics: a) is executable in a mobile peer-to-peer environment; b) is context-aware; and c) is content-centric. Building the workflow language from scratch would have been inefficient and impractical. Moreover, there have been many suitable existing workflow constructs with well-defined semantics that could have been reused. This section describes the MobWEL constructs that have been produced by domain analysis.

A. Peer-to-peer collaboration

Peer-to-peer collaboration imposes requirements on functional aspects of workflows. Workflow partitions are executed on mobile devices, therefore, the workflow language must provide the support to describe control flows in the partitions. The second requirement is that messages and information need to be exchanged directly between peers. To satisfy this requirement, asynchronous conversation needs to be supported and coordinated.

BPEL has been used as a base for MobWEL because of its characteristics. BPEL is platform-agnostic; expressed entirely in XML; extensible and adaptable; a widely accepted and adopted workflow standard language with well-defined semantics; and provides robust interaction model that enables peer-to-peer conversation. Although BPEL has been designed for processes in a Web Service world, it can be adapted for mobile peer-to-peer workflow execution. Strong coupling of process logic with the web services technology in BPEL is not very convenient and flexible approach to be used for mobile peer-to-peer interactions. Workflow participants and their roles are known beforehand so a much lighter interaction model can be used.

BPEL^{light} is a WSDL-less BPEL extension which decouples process logic from interface definition [23]. The BPEL^{light} approach allows modelling of the workflow process independently of Web service technology by introducing a single type of interaction activity. The interaction activities ($<receive>, <reply>, <invoke>, <pick>$) defined in BPEL 2.0 are resumed by single $<interactionActivity>$. In addition, BPEL^{light} processes can be modelled without specifying interface definitions (port types), hence they can be used in non-WS environment. By discarding the static specification of port types, this approach enables direct message exchange between workflow partners. The BPEL^{light} approach offers better flexibility than BPEL and is more suitable for the description of peer-to-peer mobile workflows in which all participating and interacting parties are known beforehand. Hence some concepts defined in BPEL^{light} have been adopted in MobWEL. The $<interactionActivity>$ has been further

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

adapted for the use of collaboration-related context information. An optional *collaborationContext* attribute has been added to determine whether the availability of collaborator(s) should be consider when a message is sent.

B. Context Management

To make workflows context-aware, workflows need to have the ability to react to context changes and adapt their behaviour to the changed environment. To achieve this, workflows need to be adapted for the use of context information in two ways:

- context information can be queried when needed,
- context changes and events need to be captured and handled at any time of the workflow execution.

The Context4BPEL context-related concepts have been adopted in the MobWEL workflow approach to support both ways of the use of context information.

In addition, to ensure that only relevant and meaningful context information is consumed, the abstraction level of the delivered context data should be raised. There are numerous reasons why the context raw data should be processed before its consumption in workflows. Firstly, the raw context data is heterogeneous and inconsistent. Secondly, if a context provisioning platform distributes the raw context data, workflow management system would need to deal with irrelevant context information or make context-related decisions very often. To prevent these situations, high-level context information is derived. For example, knowing when battery level is low might influence decisions whether certain operations should be performed or not. If a consuming workflow requires being informed only about the change when the battery level drops down to 5 %, constant notifications about every battery level change would be completely inefficient. By specifying a high-level context information such as *LOW* for battery level in a range between 0% and 5% and *HIGH* for a range between 5% and 100% would ensure that the context-aware workflow is notified only when the context value is changed from *HIGH* to *LOW*.

Further, relationships and dependencies between various context information can be indicated through the accommodation of context aggregation. An example of context aggregation is illustrated in Fig. 2.

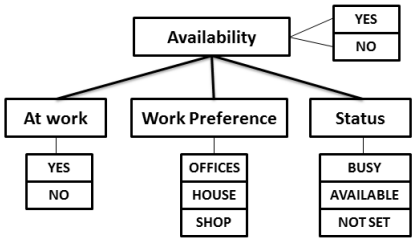

Fig. 2. Example of context aggregation

Work context of interior designers is expressed as a context aggregation of three user preferences: *At work, Work Preference* and *Status*. Context values for the *At work* context are *YES* and *NO* representing the designer's work status. Each designer can specify own work preference, for illustration purposes only *OFFICES*, *HOUSES* and *SHOPS* are shown. The *Status* is used to show whether collaborator is currently busy. *Availability* of the designer has an informative character to indicate whether the task can be taken by the person. The context value of *Availability* is determined from children context values by using associated rules. The aggregated context value can be: *(Availability-YES)* or *(Availability-NO)*.

C. Content Behaviour

Adding the content management support means that workflows should be adapted to use content state-related information similarly as they use context information:

- content state can be queried,
- content state-related changes and events are captured and handled at any time of the workflow execution.

To achieve this, some ideas developed in the BALSA workflow model seem to be practical and applicable for modelling of content-centric mobile workflows, despite the fact that context awareness has not been considered in this model. Two elements of the model, namely the Business Artifact Information Model and the Business Artifact Lifecycle have been adopted in our workflow approach. Firstly, the substance of the Business Artifact Information Model has inspired us to consider the content-related data. To gain control over processed content and its behaviour, a set of certain content-related metadata that is accessible and manageable by the workflow management system needs to be identified beforehand. This can be briefly illustrated by an example from usage scenario. If there is a rating system created for pictures, an attribute called *'ratingScore'* should be associated with each picture to hold an information about added rating. By adding the metadata to a picture, the information can be accessed at any time and can be used to make certain management decisions.

Secondly, the concept of the BALSA Business Artifact Lifecycle element is adopted in the construction process of our context-aware content lifecycle. Similarly, the context-aware content lifecycle is represented by a variant of finite state machines. Generally, a state machine contains a number of states, each state corresponding to a stage in the content lifecycle. Therefore, a *content state* is an essential construct in the content lifecycle. Content might move from one stage to another. The connections between two states are called *transitions*. When content is created, it is in an initial state with no incoming transition. At the end of its evolution, there is a final state that indicates the end of its lifecycle and has no outgoing transition. Between the initial and the final state, there are states with incoming and outgoing transitions. *Conditions* may be attached to these transitions. A condition can depend on a certain value of content attribute or external event.

A context change is an external event that can trigger a transition between two content states. Because of this, conditions placed on transitions need to be modified to deal

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

with context. With context considered, two possible situations can happen.

- In the first situation, the further evolution of content depends on the current context of the execution environment. For example, if a rating has been added to a picture, the picture moves to the *Rated* state.
  After that, based on available context information, the picture can go to either the *Ready to Review* or *Archived* state, see Fig. 3a. So if the added picture *ratingScore* is greater than or equal to the value of the user preference, the picture goes to the *Ready to Review* state. Otherwise, it is moved to the *Archived* state. In this situation, context information has to be acquired at real time. When context information is obtained, the conditions are evaluated. Context information presented in this example is simplified, considering only *RatingUP* as a representant of the current context.
- In the second situation, the transition between two content states occurs only when specific context emerges. The situation is illustrated in Fig. 3b. Rating added to the picture is stored in the *ratingScore* attribute. For example, it can have a value of '2'. Current *RatingUP* is set to '3', therefore the picture cannot move to the *Ready to Review* state at this time. Instead, it waits whether any context change in the *RatingUP* happens. The picture remains in the *Rated* state until the *RatingUP* is changed and becomes less than or equal to '2'. In contrast to the first situation, this situation requires an awaiting, monitoring and filtering mechanism for context events.

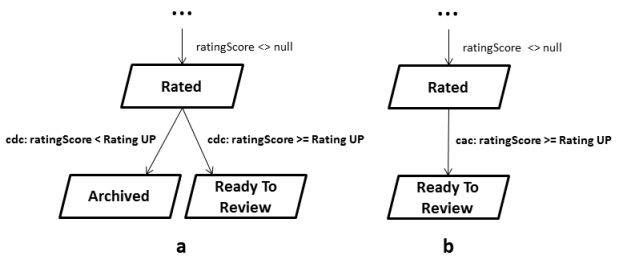

Fig. 3. Context conditions

Therefore, to distinguish between these situations, two types of context conditions are created: *context-driven condition (CDC)* and *context-aware condition (CAC)*.

D. Support for context and content awareness

Based on the requirements outlined for context and content awareness, the control flow has to be adapted for information querying and events handling. Again, BPEL has the potential to be extended to fulfil the requests. Firstly, BPEL provides event handlers to deal with occurring events. Event Handlers can be used if a context change or a content state-related change occurs. Secondly, using BPEL constructs for conditional branching, such as *switch* or *while*, enables to query information at decisions points. To query context information or content state-related information, two extension functions need to be supported:

(Object contextValue) mobwel:getContextValue(String contextName)

(String contentState) mobwel:getContentState(String contentVariable)

The use of the functions is illustrated in Fig. 4.

```
<switch>
  <case condition="mobwel:getContextValue(RatingUP)=3">
    ...
  </case>
  <case condition="mobwel:getContentState(picture)='Reviewed'">
    ...
  </case>
  <otherwise>
    ...
  </otherwise>
</switch>
```

Fig. 4. Example of the use of the extension functions

Constructs presented in this section have been accommodated in the design of the MobWEL workflow language. The language is defined next.

V. MobWEL Language Definition

This section introduces the MobWEL workflow language. Firstly, the anatomy of a MobWEL workflow is outlined (Fig. 5).

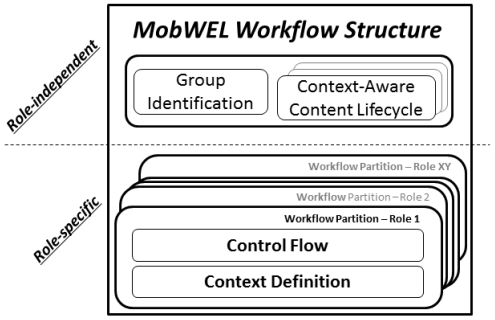

Fig. 5. Anatomy of MobWEL workflow

MobWEL workflows are designated for mobile peer-to-peer collaboration, thereby a number of roles and workflow participants can be involved in the workflow execution. Workflow participants have assigned roles, each role performs only activities defined in an allocated workflow partition.

MobWEL workflows are composed of various constituents, which are either *role-specific*, explicitly defined for each participating role, or *role-independent*, defined same for all participating roles. The workflow constituents are described next.

A. Group Identification

In collaborative workflows, tasks are performed by collaborators. In this concept, workflow collaborator is a person who uses a mobile device to collaborate, share content and communicate with other team members in order to achieve a common goal. To execute workflows in peer-to-peer collaboration, each mobile device needs to be given the relevant details about fellow collaborators. The *Group Identification* workflow constituent specifies all *Collaborators* participating in a workflow and their associated *Roles* (Fig. 6).

Each *Collaborator* is described by personal name and the attributes which identify collaborator's mobile device such as

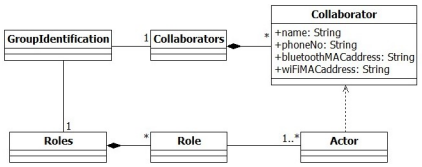

Fig. 6. Metamodel of Group Identification

phone number, Wi-Fi and Bluetooth MAC addresses (assuming that a collaborator uses only one device to collaborate). A collaborator can play more than one role in a workflow (not in same workflow instance). A good example is given in our usage scenario where Jane can play the role of designer but also the role of reviewer if asked to review a picture created by her co-worker. Therefore, the *Roles* element describes roles involved in workflow. Each *Role* can be played by a number of actors, each *Actor* refers to a *Collaborator* and is identified only by the *phoneNo* attribute of the collaborator. This enables collaborators to be associated with more roles.

B. Workflow-Specific Context Definition

Context Definition Metamodel, see Fig. 7, expresses all the constructs, concepts and relationships between the constructs needed to build workflow-specific context definition models.

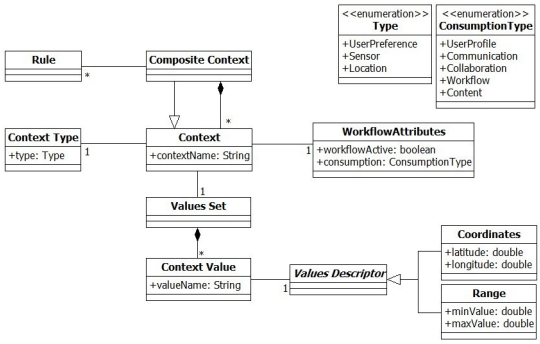

Fig. 7. Metamodel of Context Definition

The workflow specific context definition modelling approach allows building context models that are specific for each workflow partition. It enables to express context aggregations, sets of values and other attributes related to a workflow partition.

Context is described by its contextName, *Context Type* and *Values Set*. For example, context can have the name: *Status*, the context type of *User Preference*, and the context values set: {*Busy, Available, Not Set*}. A high-level *Context Value* can be derived from raw context data. We use the *Values Descriptor* construct to associate the high-level context values with raw data. Two examples: *Range* and *Coordinates* are outlined. For instance, *LOW* as a high-level context value for battery can be defined for a range with 'minValue set to 0' and 'maxValue set to 5' (%). Coordinates are used to specify a location or a place of interest.

Composite context, dedicated for context aggregation, is designed as a context container. It is a subtype of *Context* and inherits its attributes and behaviour. A composite context can be built as an aggregation of other composite and atomic contexts. Based on the context values of child contexts, *Rule* is used to determine a context value of the composite context. Therefore, the composite context values set can be fully defined by workflow designers. For example, the context value of *Availability* can be defined as *YES* if the current values of *At work* is *YES*, *Work Preference* is *HOUSE* and *Status* is *AVAILABLE*.

Workflow Attributes are the attributes required for better functioning of the workflow management system. Firstly, the workflow management system does not need to obtain context information of all defined contexts. For example, in order to execute a workflow step, it might need to obtain only context information: *Availability-YES* without knowing that the context value of *At work* is set to *YES* or the value of *Status* is set to *AVAILABLE*. Therefore, each context is associated with the *workflowActive* attribute which can be set to *true* if a particular workflow is interested in listening to changes of this context, respectively to *false* if the context is only auxiliary and its values do not directly influence the workflow execution. Secondly, the *consumption* attribute is used to identify the consumption of context information within the workflow management system.

C. Context-Aware Content Lifecycle Definition

A number of content items can be processed across multiple mobile devices in a MobWEL workflow case, each content item following its own content lifecycle. The evolution of content items is distributed across all collaborators, therefore a set of content lifecycles is defined once and remain same for all workflow partitions. Based on the design specifications presented in the previous section, the metamodel for a context-aware content lifecycle has been constructed. All constructs related to a content lifecycle and their relationships are shown on the metamodel in Fig. 8.

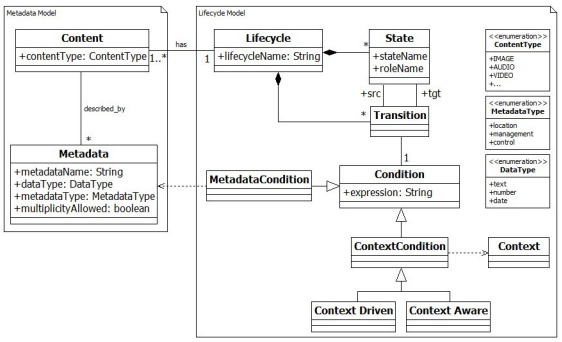

Fig. 8. Metamodel of Content Lifecycle

There are two parts presented, one representing content information model and another representing content lifecycle. *Lifecycle* is described by its lifecycleName. The lifecycle can describe the behaviour of several content items, however, in a particular definition it can be associated only with one *Content* item. The *Content* item is characterised by its contentType

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

such as image, audio or video and a set of workflow-specific *Metadata*, each metadata of a specific *metadataType*. Furthermore, *Lifecycle* is composed of a number of *State*s and *Transition*s. *State* is characterised by stateName and roleName. The latter attribute determines the role associated with the ownership of the state. Each *Transition* has a source and target *State* and can be associated with a *Condition*. To trigger the transition of the content item between the source and target state, the associated condition must evaluate to true. There are two subtypes of conditions, *MetadataCondition* and *ContextCondition*. Fulfillment of *MetadataCondition* depends on the value of particular metadata. *ContextCondition* depends on certain context and can be *ContextAware* or *ContextDriven*.

D. Process Control Flow

The main part that integrates other constituents is the control flow of the MobWEL workflow. MobWEL extends BPEL, therefore, the major part of its syntax and semantics have been inherited from BPEL. The focus is put on constructs that have been introduced or existing constructs that have been conceptually modified. These constructs depicted in Fig. 9 are as follows:

- *MobWEL Workflow Process:* A workflow process identified by its unique *targetNamespace*. The *groupIdentification, contentLifecycleDescription and contextDefinition* attributes specify namespaces of definitions of group identification, content lifecycle and context.
- *Variable:* A variable is used in a standard BPEL process. The variables used for content items have the default value for the variable type set to 'content'.
- *Lifecycle:* This element refers to the description of an explicitly defined content lifecycle. Each lifecycle is specified by its *name* and *resource*, a path to the explicit definition. A *variable* of a content type must be also associated with the lifecycle. More lifecycles can be declared within one workflow process.
- *ContextDefinition:* This element refers to the explicit context model. The *contextDefinitionSource* attribute specifies the path to the context definition model. A workflow process can have zero or one context definition declared.
- *CollaboratorsGroup:* A group of collaborators is explicitly defined and imported in the workflow definition through this element. It is specified by its *groupName* and *groupSource*, a path to the group description. A workflow process has one group of collaborators declared.
- *Partner and Conversation:* A declaration of partners and conversations involved in collaboration. The constructs are adopted from the BPEL^{light} workflow definition. A workflow process can have one or more partners.
- *MobWEL Activity:* A base type for MobWEL activity.
- *InteractionActivity:* An interaction activity adopted from BPEL^{light} used for all interaction activities between partners and adapted for collaboration-related context.
- *ContentActivity:* An activity with an informative character. The activity is used to access content-related data stored in process variables and communicate them out to external parties.

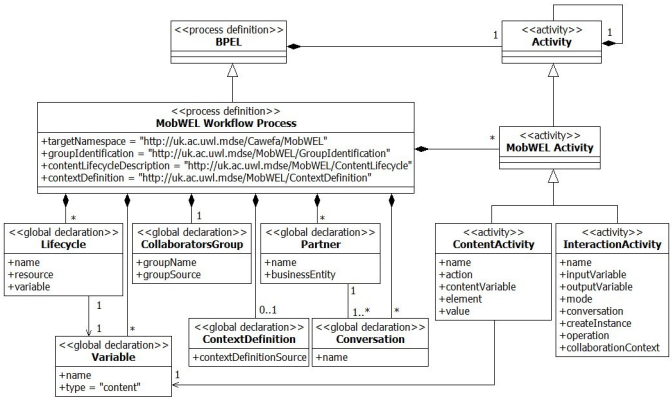

Fig. 9. MobWEL constructs

Most of the introduced constructs are global declarations in the process. Only *interactionActivity* and *contentActivity* are specific activities used in the description of the control flow.

VI. MobWEL Workflow Management System

MobWEL workflows are carried out on mobile devices. Management and execution of MobWEL workflows in a distributed manner across a number of devices require an adapted workflow management system that is deployed on each device. In this section, the architecture of such mobile workflow management system is described.

The high-level architecture of the MobWEL workflow management system is shown in Fig. 10.

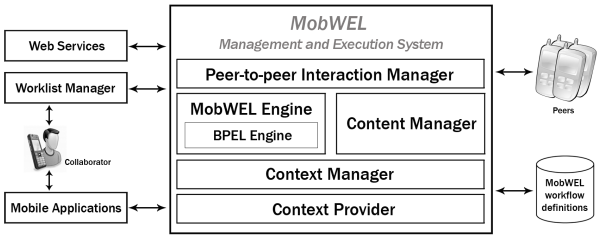

Fig. 10. High-Level Logical Architecture

The system is composed of the following components: Context Provider, Context Manager, Content Manager, MobWEL Engine, and Peer-to-peer Interaction Manager.

A. Context Provider

Context Provider is a component responsible for context monitoring, acquiring, processing, aggregating and disseminating. Context Provider interprets context models and creates corresponding internal data structures. Each particular context concern is managed through a self-contained context component. When a context change occurs, raw context data is processed, and context information in the *AppKey-ContextName-ContextValue-ContextDate* format is broadcasted to interested parties.

B. Context Manager

Context Manager acts as intermediator between Context Provider and other internal workflow management components. It manages the use and consumption of context information and drives context routing to other components of the MobWEL workflow management system at run-time.

It contains mechanisms to synchronously and asynchronously communicate with Context Provider. To communicate effectively, both components need to be aware about the same contexts. When the definition for a new workflow is deployed to Context Manager, Context Manager also passes it to Context Provider. Because of this, the same workflow-specific context definition is deployed to both components, but they extract different information from it. While Context Provider is more interested in extracting information related to context hierarchies and context reasoning, Context Manager focuses only on active contexts. Context Manager extracts *contextName, contextValuesSet, workflowID* and *workflowUse* for all contexts labelled as workflow-active. This ensures that Context Provider broadcasts context messages that are related to the workflows deployed in the workflow management system, and Context Manager can process them further. Context Manager has implemented a mechanism that monitors and filters messages broadcasted by Context Provider. As opposed to this asynchronous way of communication, Context Manager contains also a mechanism that enables synchronous binding to Context Provider and supports querying of context information at real time.

Context information is consumed within the workflow management system in several ways. Based on the workflow-specific context *consumption* value extracted from the context definition, Context Provider knows which component is interested in obtaining the context information. For example, if it is content-related, Content Manager is informed, if it is workflow-related, MobWEL engine gets the context notification. Context Manager also persists last context values.

C. Content Manager

Content Manager provides advanced content management functionalities and manages content lifecycles. This component parses the *Context-aware content lifecycle* elements of the MobWEL workflow definition. Each lifecycle element contains description of content-related metadata and description of content lifecycle. Content-related data is persisted and managed by using *Content Provider*. When workflow case is instantiated, multiple content items are created that follow the same lifecycle. *Content State Transition System* is used to manage and control content lifecycles.

D. MobWEL Engine

MobWEL Engine is the execution engine and the heart of the MobWEL workflow management system that parses and instantiates workflows, and provides an overall control over the management and execution of MobWEL workflow instances. The MobWEL engine is based on a BPEL engine. Sliver is an open source lightweight workflow engine that supports the execution of SOAP services and BPEL processes on a wide range of devices, and communication protocols [24]. This BPEL engine has been used as a base for the MobWEL engine.

The MobWEL engine contains a main parser and an execution unit. The parser interprets the whole MobWEL workflow definition and invokes corresponding parsers in other components for parsing of all workflow constituents.

The execution unit extends the BPEL execution unit, therefore, only the support for the execution of constructs introduced in the MobWEL language, namely *interactionActivity* and *contentActivity*, has to be added. The semantics for *interactionActivity* is inherited from BPEL^{light}. The *contentActivity* has only an informative character with a simple role to inform Content Manager and its execution does not change the data flow, thereby it is handled in the same manner as other BPEL activities and its internal execution depends on the action specified in the activity.

E. Peer-To-Peer Interaction Manager

Peer-to-peer Interaction Manager manages communication and messages transmissions with other mobile devices. *Event Handlers* handle incoming messages from other devices and requests coming from mobile applications or services. An event handler can be designed for each form of communication such as incoming binary data message, multimedia message, or messages coming via Bluetooth connection. Structured information and messages are sent between devices as a sequence of bytes, *Message Parser* is used to convert workflow-related data and objects into such message format that can be transmitted across the network. The parser also extracts workflow objects from incoming messages. *Identity Manager* is used to store and manage the details about participating collaborators. The details include collaborators' names, their roles, and other data needed for device-to-device communication.

F. Validation

To prove that the theory is functional, concrete research artefacts have been constructed. Firstly, the metamodel of MobWEL workflow language has been mapped into an XML schema. The MobWEL XML Schema is a tool that can be used by workflow designers who would be able to model and define the concrete workflows in an XML format. Secondly, formal MobWEL semantics has been defined. Finally, software prototypes have been implemented for the Android platform. Context provider operates on a mobile device and can be built as an internal component of the workflow management system, which provides services solely to the MobWEL workflow management system. However, there might be numerous context-aware applications running on the same mobile device which need to use the same context provisioning services. Thereby, Context Provider has been implemented as an external component, and named *ContextEngine*. The remaining components of the MobWEL workflow management system have been implemented as a system named *CAWEFA* (Context-Aware Workflow Engine For Android). A MobWEL workflow definition has been constructed for the

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

usage scenario. The workflow definition has been deployed to CAWEFA and ContextEngine, and instantiated several times. The expected behaviour of each workflow instance expressed by using the MobWEL semantics has been compared with the actual behaviour of running workflow instances. The conducted experimentation and obtained results have shown that the MobWEL workflow language can be used to describe executable mobile context-aware content-centric workflows.

VII. Conclusion and Future Work

In this paper, we have presented MobWEL, a context-aware content-centric workflow language designed for mobile peer-to-peer collaboration, and the architecture of a MobWEL workflow management system. One of our main goals has been to build a language which is context-aware, visualises and integrates content behaviour, and supports peer-to-peer interaction. We have extended BPEL, considered the dimensions of the data-centric approach proposed in the BALSA model, and adopted some existing constructs from Context4BPEL and BPEL^{light}. In this paper, only a first milestone in the definition, management and execution of MobWEL workflows has been presented.

Integration of context and content awareness into the workflow technology offers numerous benefits. Firstly, by adding context definitions to the workflow description, the execution of workflow partition can be adapted to individual collaborator's needs and situation in the current run-time environment. Secondly, additional content management functionalities enable to monitor content behaviour and disseminate progress to other peers. However, there have been also some limitations discovered that should be addressed in future work. For example, the development of the MobWEL language has been tailor-made for a specific class of workflows. It would be interesting to use and validate the MobWEL workflow approach in other scenarios. Further, although integration of context awareness enables workflow adaptation to individual user needs and preferences, the proposed approach requires a prediction of possible context situations beforehand. However, context emerges in the moment and cannot be fully predicted. In addition, content sharing between devices might be a time consuming and costly operation, especially when one collaborative task may be accomplished by a number of actors with the same role but only few of them might be able to perform the task. So another valuable extension of this work would be in the development of an appropriate content sharing strategy.

A major limitation of this work is that it is up to workflow designers to construct MobWEL workflows in a way that ensures the consistency between the process-based workflow models and associated content lifecycle(s). Thereby, further work is required to establish a MobWEL workflow design methodology.

References

[1] S. Smanchat, S. Ling, and M. Indrawan, "A survey on context-aware workflow adaptations," in Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia, 2008, pp. 414–417.

[2] A. K. Dey, "Understanding and using context," in Personal and Ubiquitous Computing, vol. 5, 2001, pp. 4–7.

[3] L. Pajunen and S. Chande, "Developing workflow engine for mobile devices," in Enterprise Distributed Object Computing Conference, 2007, pp. 279–286.

[4] A. Kocurova, S. Oussena, P. Komisarczuk, and T. Clark, "Context-aware content-centric collaborative workflow management for mobile devices," in Second International Conference on Advanced Collaborative Networks, Systems and Applications, 2012, pp. 54–57.

[5] M. Rosemann, J. Recker, and C. Flender, "Contextualisation of business processes," in International Journal of Business Process Integration and Management, vol. 3(1), 2008, pp. 47–60.

[6] O. Saidani and S. Nurcan, "Context-awareness for adequate business process modelling," in Third International Conference on Research Challenges in Information Science, 2009, pp. 177–186.

[7] M. Adams, A. Ter Hofstede, N. Russell, and W. Van der Aalst, "Dynamic and context-aware process adaptation," in Handbook of Research on Complex Dynamic Process Management: Techniques for Adaptability in Turbulent Environments/Ed. MM Wang, Z. Sun, 2009, pp. 104–140.

[8] M. Rosemann and J. Recker, "Context-aware process design: Exploring the extrinsic drivers for process flexibility," in The 18th International Conference on Advanced Information Systems Engineering, Proceedings of Workshops and Doctoral Consortium, 2006, pp. 149–158.

[9] D. Salber, A. Dey, and G. Abowd, "The context toolkit," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI is the limit, 1999, pp. 434–441.

[10] P. Korpipaa, J. Mantyjarvi, J. Kela, H. Keranen, and E. Malm, "Managing context information in mobile devices," in Pervasive Computing, vol. 2(3), 2003, pp. 42–51.

[11] K. Henricksen and J. Indulska, "Developing context-aware pervasive computing applications: Models and approach," in Pervasive and Mobile Computing, vol. 2(1), 2006, pp. 37–64.

[12] C. Bettini et al., "A survey of context modelling and reasoning techniques," in Pervasive and Mobile Computing, 2010, pp. 161–180.

[13] P. Bellavista, A. Corradi, M. Fanelli, and L. Foschini, "A Survey of Context Data Distribution for Mobile Ubiquitous Systems," in ACM Computing Surveys, vol. 45(1), 2013, pp. 1–49.

[14] D. Kramer, A. Kocurova, S. Oussena, T. Clark, and P. Komisarczuk, "An Extensible, Self Contained, Layered Approach to Context Acquisition" in Proceedings of the Third International Workshop on Middleware for Pervasive Mobile and Embedded Computing, 2011, pp. 1–6.

[15] M. Wieland, O. Kopp, D. Nicklas, and F. Leymann, "Towards Context-aware Workflows," in CAISE 07 Proceedings of the Workshops and Doctoral Consortium, 2007, pp. 11–15.

[16] M. Grossmann et al., "Efficiently managing context information for large-scale scenarios," in Pervasive Computing and Communications, 2005, pp. 331–340.

[17] R. Hull, "Artifact-centric business process models: Brief survey of research results and challenges," On the Move to Meaningful Internet Systems: OTM 2008, pp. 1152–1163.

[18] R. Liu, K. Bhattacharya, and F. Wu, "Modeling business contexture and behavior using business artifacts," in Advanced Information Systems Engineering, 2007, pp. 324–339.

[19] K. Bhattacharya, R. Hull, and J. Su, "A data-centric design methodology for business processes," in Handbook of Research on Business Process Management, 2009, pp. 503–531.

[20] K. Wahler, "A Framework for Integrated Process and Object Life Cycle Modeling," PhD Thesis, 2009.

[21] R. Hull et al., "Introducing the guard-stage-milestone approach for specifying business entity lifecycles," in Web Services and Formal Methods, 2011, pp. 1–24.

[22] V. Künzle and M. Reichert, "PHILharmonicFlows: towards a framework for object-aware process management," in Journal of Software Maintenance and Evolution: Research and Practice, vol. 23, 2011, pp. 205–244.

[23] J. Nitzsche, T. Van Lessen, D. Karastoyanova, and F. Leymann, "BPEL^{light}," in Proceedings of the 5th International Conference on Business Process Management, 2007, pp. 214–229.

[24] G. Hackmann, M. Haitjema, C. Gill, and G. Roman, "Sliver: A bpel workflow process execution engine for mobile devices," in Service-Oriented Computing, 2006, pp. 503–508.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Search Computing for E-government

Alessandro Campi, Matteo Aletti, Yao Cheng, Edoardo Colombo,

Piero Franceschini, Mariangela Rizzo, Davide Sanapo

Politecnico di Milano
Dipartimento di Elettronica e Informazione
Piazza Leonardo da Vinci 32,
20133, Milano, Italy
campi@elet.polimi.it,
[matteo.aletti|yao.cheng|edoardo.colombo|piero.franceschini|mariangela.rizzo|davide.sanapo]@asp-poli.it

***Abstract*—People shop online, compare online, book hotels and flights online. This happens because the data needed to complete these tasks are easily accessible, and a lot of Web sites allow users to query the Web to obtain enough information to be confident. The aim of this work is to propose a framework tailored to extend the internet revolution to public administration. This work is the first step towards an infrastructure allowing people to know in a very easy way the information they need. This paper exploits the Search Computing paradigm. It is a new way for composing data. While state-of-art search systems answer generic or domain-specific queries, Search Computing enables answering questions via a constellation of cooperating data sources, called search services, which are correlated by means of join operations. Search Computing aims at responding to queries over multiple semantic fields of interest; thus, Search Computing fills the gap between generalized search systems, which are unable to find information spanning multiple topics, and domain-specific search systems, which cannot go beyond their domain limits.**

Keywords-search; integration; query; e-government

I. INTRODUCTION

One hundred years ago, sending 10 words cost 50 dollars. Today gigabytes of information can be sent for a fraction of that cost, and huge quantities of information can be sent and received without it costing anything. This has changed the way people shop, the way people travel, the way people do business. The internet revolution has actually gone all the way through societies in many different ways, but it hasn't yet touched the way states are governed [1].

People shop online, compare online, book hotels and flights online. This happens because the data needed to complete these tasks are easily accessible and a lot of Web sites allow users to query the Web to obtain enough information to be confident. The aim of this work is to propose a framework tailored to extend the internet revolution to public administration. This work is the first step towards an infrastructure allowing people to know in a very easy way the information they need. People could search what operations work out properly, what records doctors have, the cleanliness of hospitals, who does best at infection control, etc.

Moreover, our work is a step forward to the transparency of public administration, and toward a real accountability of public services, and real awareness of citizens about their governments politics. The Missouri Accountability Portal [2] is an example of a Web portal that made available online all the data of one state in America. Every single dollar spent by that government is searchable, is analyzable, is checkable. Any business that wants to bid for a government contract can see what currently is being spent and possibly can offer to deliver it in a cheaper way.

Finally, our work allows one to relate objective data available online or offline with news articles, blog posts or other comments available online enriching objective data with information about the mood of people (as shown in Figure 1). In today's politics the need of a fast reaction to discontent is paramount and the possibility to relate numbers, facts, and sentiment analysis is very important.

Our proposal exploits the Search Computing paradigm [3]. It is a new way for composing data. While state-of-art search systems answer generic or domain-specific queries, Search Computing enables answering questions via a constellation of cooperating data sources, called search services, which are correlated by means of join operations. Search Computing aims at responding to queries over multiple semantic fields of interest; thus, Search Computing fills the gap between generalized search systems, which are unable to find information spanning multiple topics, and domain-specific search systems, which cannot go beyond their domain limits. Paradigmatic examples of Search Computing queries are: "Where is the school closest to my home, offering a high teaching quality and a good food service?", "Who is the best doctor who can cure insomnia in a nearby public hospital?", "Which are the highest risk factors associated with the most prevalent diseases among the young population?". These queries cannot be answered without capturing some of their semantics, which at minimum consists in understanding their underlying domains, in routing appropriate query subsets to each domain-expert search engine, and in combining answers from each engine to build a complete answer that is meaningful for the user.

A prerequisite for setting such goal is the availability of a large number of valuable search services. We could just

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

wait for SOA (Service Oriented Architecture) to become widespread. However, in the public administration scenario, very few data are offered by services designed to support search, and, moreover, a huge number of valuable data sources are not provided with a service interface. In this paper, we tackle the important issue of publishing service interfaces suitable for Search Computing so as to facilitate the widespread use of data sources on the Web and to simplify their integration in Search Computing applications.

The goals of our proposal is to create a virtuous system where the Public administration shares its data with citizens (increasing the transparency of the government and the trust in the public administration). Citizens analyze the available data to learn something but also to provide hints to the public administration helping to offer better services.

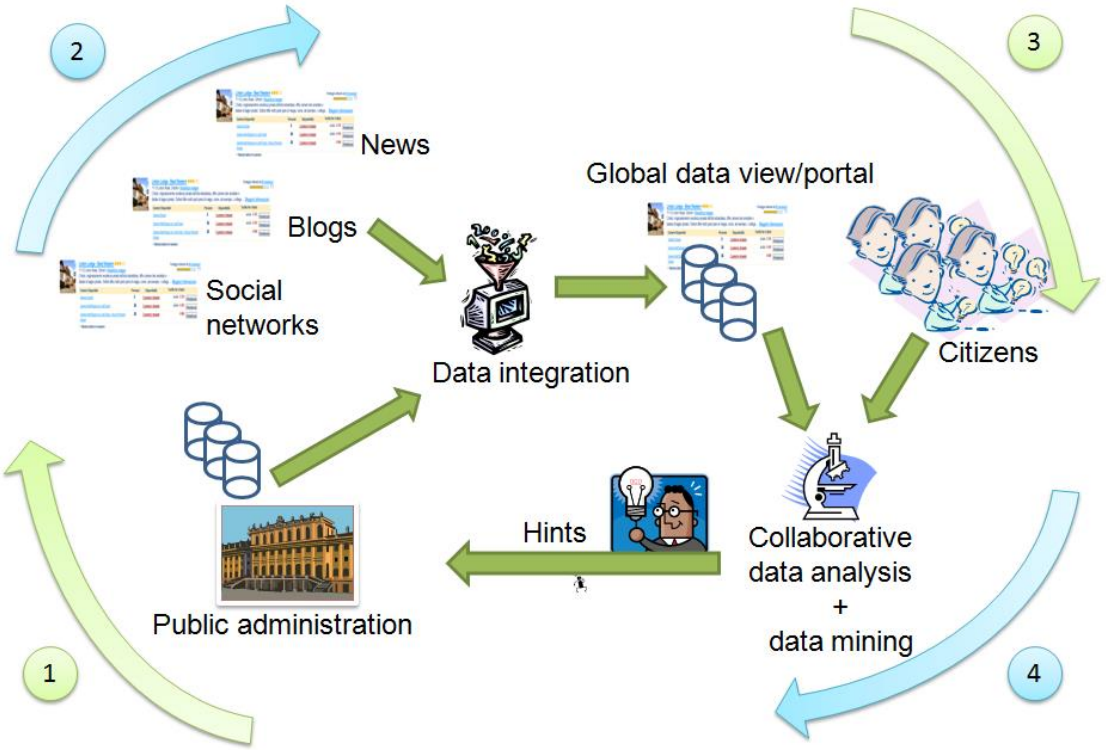

Figure 1: The data life cycle

II. THE FRAMEWORK

Figure 2 shows an overview of our framework, composed of several sub-frameworks. As shown in the figure, the data framework provides the scaffolding for wrapping and registering data sources. The core concept of the data framework are Service Marts. The user framework provides functionality and storage for registering users, with different roles and capabilities. The query framework supports the management and storage of queries: a query can be executed, saved, modified, and published for other users to see. The execution framework is responsible of an efficient execution of the previously defined query plan. Finally, the invocation framework masks the technical issues involved in the interaction with the data sources, e.g., the Web service protocol and data caching issues.

The core of the framework aims at executing multi-domain queries. The query manager takes care of splitting the query into sub-queries (e.g., "Who is the best doctor that cures insomnia?" ; "Who are the doctors of a certain hospital?"; "Which hospital is close to my house?") and bounding them to the respective relevant data sources registered in the data framework repository (in this case, hospitals could be retrieved using GeoPlaces, places close to my home using Google Maps, …); starting from this mapping, the query framework produces an optimized query execution plan, which dictates the sequence of steps for executing the query. Finally, the execution framework actually executes the query plan, by submitting the service calls to designated services through the service invocation framework, building the query results by combining the outputs produced by service calls, computing the global ranking of query results, and producing the query result outputs in an order that reflects their global relevance.

To obtain a specific application, this general-purpose architecture is customized by users, supported by appropriate design tools. The development process involves users with different roles and expertise (see Figure 3).

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

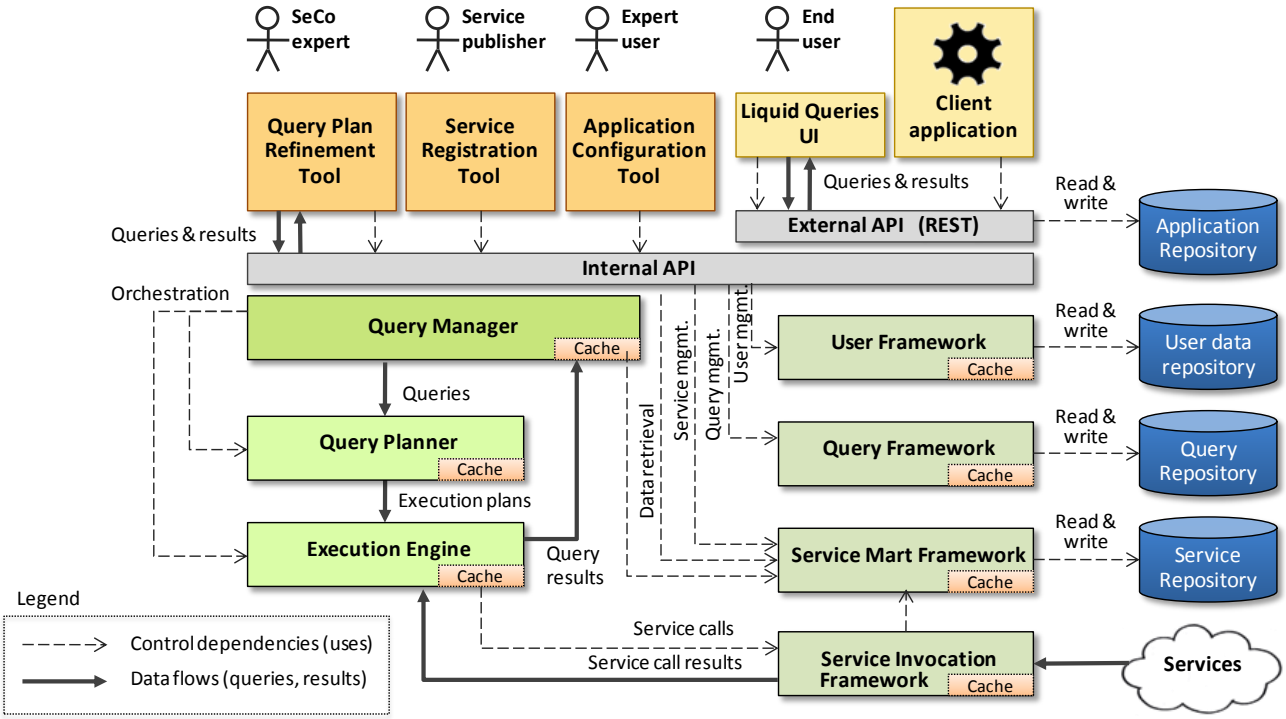

Figure 2. The framework (more details in [3])

Data Publishers: they are in charge of implementing mediators, wrappers, or data materialization components, so as to make data sources compatible with the service mart standard interface and expected behavior; they register service mart definitions within the service repository, and declare the connection patterns usable to join them.

Expert Users: they configure Search Computing applications, by selecting the service marts of interest, by choosing a data source supporting the service mart, and by connecting them through connection patterns. They also configure the user interface, in terms of controls and configurability choices for the end user.

End Users: they use Search Computing applications configured by expert users. They interact by submitting queries, inspecting results, and refining/evolving their information need according to an exploratory information seeking approach, which we call Liquid Query [4].

The development process steps lead to the final application accessed by the end user. The Liquid Query interface, instantiated during the application configuration phase, supports the "search as a process" paradigm, based on the continuous evolution, manipulation, and extension of queries and results; the query lifecycle consists of iterations of the steps of query submission, when the end user submits an initial liquid query; query execution, producing a result set that is displayed in the user interface; and result browsing, when the result can be inspected and manipulated through appropriate interaction primitives, which update either the result set (e.g., re-ranking or clustering the results) or the query (e.g., by expanding it with additional service marts or requesting for more results). This approach to development takes into account the trend towards empowerment of the user, as witnessed in the field of Web mash-ups [5]. Indeed, all the design activities from service registration on do not ask to perform low-level programming.

III. CONCLUSIONS AND FUTURE WORKS

This proposal is a step forward towards the possibility to exploit all the information we have about government and society for a better understanding of the actual situation. This work exploits the Search Computing paradigm in order to allow people to relate objective data available online or offline with news articles, blog posts or other comments available online enriching objective data with information about the mood of people.

In the future, we envision the spreading of software components and methodologies for e-government. In particular, we would like to study how to make the search of data related to e-government as easy as a Google (or Yahoo! or Bing) query.

IV. ACKNOWLEDGMENTS

This work was partially supported by Search Computing ERC Grant.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

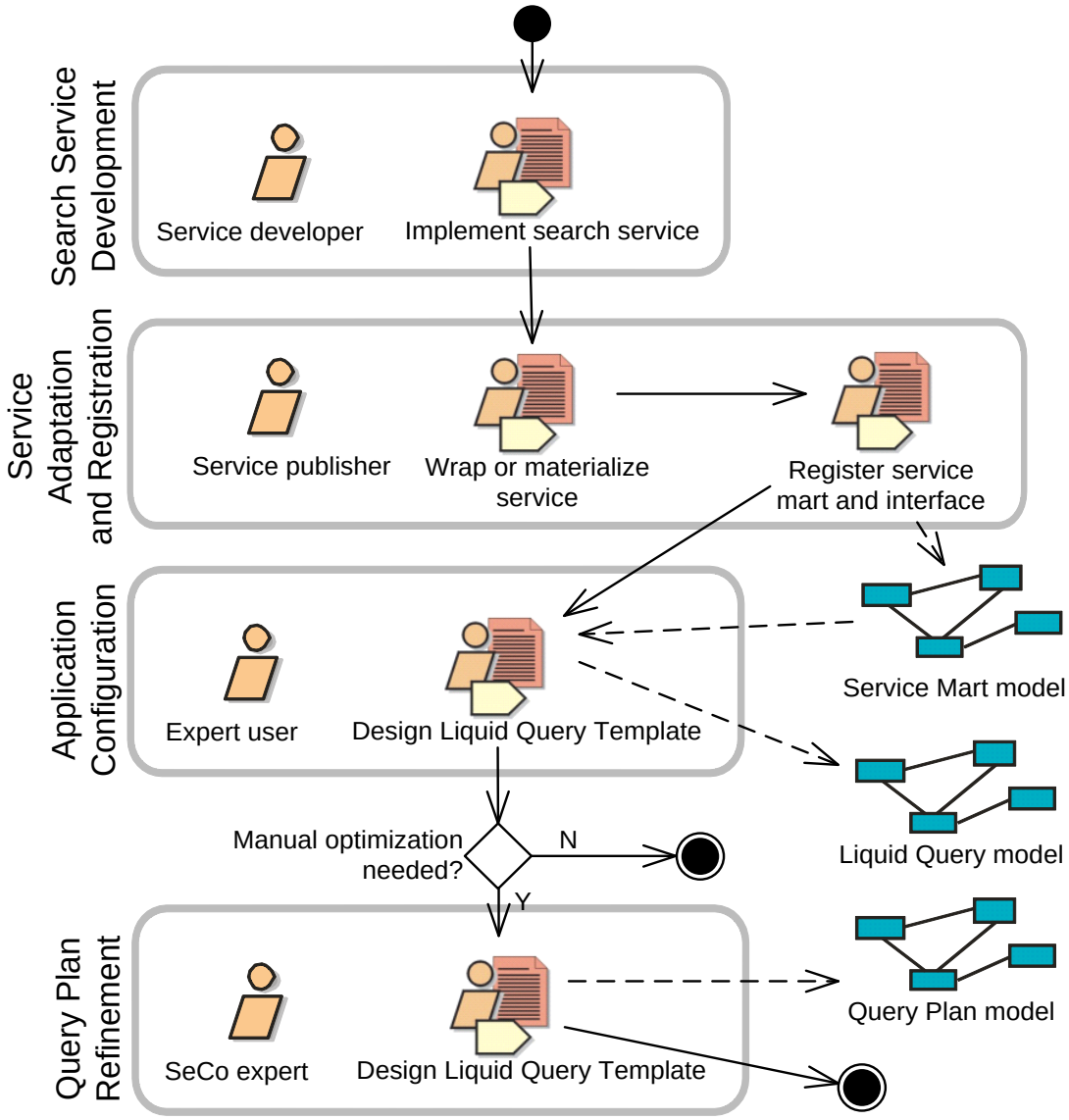

Figure 3. Users (more details in [4])

V. REFERENCES

[1] David Cameron. The next age of government. http://blog.ted.com/2010/02/16/the_next_age_of/ (accessed 1/2/2011)

[2] Map your taxes. http://mapyourtaxes.mo.gov/MAP/Portal/Default.aspx (accessed 1/2/2011)

[3] Ceri, S., Brambilla, M. (eds.). Search Computing Challenges and Directions. Springer LNCS vol. 5950, March 2010.

[4] A. Bozzon, M. Brambilla, S. Ceri, P. Fraternali. Liquid query: multi-domain exploratory search on the web. In *Proc. of the 19th int. conf. on World wide web* (WWW '10). ACM, New York, USA, 161-170.

[5] Braga, D., Ceri, S., Daniel, F., Martinenghi, D. Mashing Up Search Services. IEEE Internet Comp. 12(5) (2008), 16-23.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

A Proposal of New Autonomous Decentralized Structure Formation Based on Huygens' Principle and Renormalization

Kenji Takagi, Masaki Aida
Graduate School of System Design,
Tokyo Metropolitan University
Tokyo, Japan
Email:{takagi-kenji,aida}@sd.tmu.ac.jp

Chisa Takano
Graduate School of Information Sciences,
Hiroshima City University
Hiroshima, Japan
Email:takano@hiroshima-cu.ac.jp

Makoto Naruse
National Institute of Information
and Communications Technology (NICT)
Tokyo, Japan
Email:naruse@nict.go.jp

***Abstract*—This paper proposes an autonomous distributed algorithm that can construct spatial structures for clustering in MANETs. Since the topology of a MANET changes frequently, a fast, light-weight, and autonomous clustering mechanism is required. However, existing autonomous clustering mechanisms are based on differential equations and thus demand a lot of calculations for generating the spatial structures that yield clustering. This paper proposes an autonomous clustering algorithm that is based on Huygens' principle. The most remarkable characteristics of our proposal are calculation simplicity and fast convergence on the cluster structures. We verify the basic characteristics of the proposed algorithm.**

Keywords-autonomous decentralized control; structure formation; asymptotic stability; Huygens' principle; renormalization transformation.

I. INTRODUCTION

In large-scale communication networks, hierarchical architectures are effective for scalable network control. Let us consider how to introduce a hierarchical structure into a network. In a network having fixed topology, we can consider the desired hierarchical structure when designing the network. Unfortunately, this is not possible in a network with a dynamic topology. A typical example of such a network is the mobile ad hoc network (MANET) [1]. A MANET consists of mobile terminals that offer routing functions and data forwarding. Two terminals can directly communicate if their coverage areas overlap. If the areas do not overlap, the terminals do not directly communicate, but by relaying data through terminals between the two terminals, they can establish multihop communication. To achieve multihop communication, routing is one of most important issues in MANETs.

The primitive approach to route finding is called flooding [2]. In flooding, the sender terminal sends route finding packets to all adjacent terminals, which resends them to all their adjacent terminals until at least one copy of the packet reaches the destination terminal. The total amount of route search packets sent in MANETs increases exponentially with network size (the number of terminals). One of the challenges in MANETs, realizing scalable routing control [3], [4], is best addressed by setting a hierarchical structure through clustering [5]–[7].

Hereafter, we call a MANET terminal a node. An autonomous clustering mechanism for generating a hierarchical structure requires several characteristics, as follows:

- Each node acts autonomously based on local information about its neighboring nodes.
- The generated cluster structure should reflect the state information of the network (e.g., battery power of nodes).
- The generated cluster structure should be flexible so that it can adapt to the dynamic environment.
- The convergence time of clustering should be sufficiently shorter than the timescale of topology change enforced by node movement. This is because clustering should dynamically adapt to the network topology.
- Action rules of each node should be as simple as possible in order to reduce the battery power consumed by computation or processing at the node.

Since ad hoc networks are expected as an effective communication tool after serious disaster, the above requirements are essential for realizing clustering in ad hoc networks.

Takano *et al.* has proposed a clustering mechanism based on the Fokker-Planck equation and includes the drift motion given by back-diffusion [8], [9]. Let us call it the back-diffusion based approach. This mechanism satisfies the first two requirements listed above. Hamamoto et al. [10] recently proposes a guaranteeing mechanism of the asymptotic stability, and clarifies that, by using the guaranteeing mechanism, the back-diffusion based approach can satisfy the third requirements listed above. In addition, it also implies that we might be able to make the clustering algorithm that satisfies all the requirements listed above by replacing back-diffusion algorithm with other simple and fast-converging rule. This is because the guaranteeing mechanism of the asymptotic stability does not depend on details of the clustering mechanism. In this paper, we use Huygens' principle [11] as the simple and fast-converging rule, and propose a new clustering mechanism that satisfies all the requirements listed above.

The paper is organized as follows: Section II explains the guaranteeing mechanism of the asymptotic stability, which is the foundation of this research. Section III proposes an autonomous clustering mechanism based on Huygens' principle. Section IV discusses the initial condition of the proposed mechanism with the goal of ensuring robust controllability of cluster size. In addition, it also shows cluster structures generated by the proposed mechanism using numerical examples and verifies that they reflect the network condition. The

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

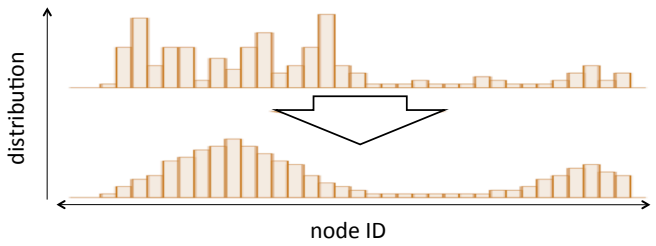

Fig. 1. Concept of cluster forming.

conclusion is discussed in Section V.

II. PRELIMINARY

A. Concept of Cluster Formation

In our clustering model, each node has certain value and cluster formation is conducted by the distribution of the values of nodes. The initial value is determined by considering the network condition (e.g., battery power of each node). The clustering algorithm is to extract a coarse grained spatial structure from the initial distribution of the values and this procedure corresponds to clustering. Fig. 1 shows an example cluster formation in a simple 1-dimensional network. The horizontal axis represents node ID, and the vertical axis represents the value of the distribution for each node. The top part of Fig. 1 represents the initial initial distribution, which is reflecting network state (e.g., battery power of each node). The bottom part of Fig. 1 represents the generated coarse grained spatial structure. The peaks of the coarse grained distribution correspond to clusters and they are reflecting the initial condition.

The back-diffusion based approach is an example of this mechanism, and has a relatively faster convergence rate than conventional bio-inspired approach [12]. However, this clustering mechanism does not consider the change of the initial condition, and therefore it cannot adapt to the dynamic environment. That is why this mechanism does not satisfy the third requirement listed in the previous section.

B. Guaranteeing Mechanism of Asymptotic Stability

To adapt the spatial structure to the dynamic environment, the guaranteeing mechanism of the asymptotic stability of cluster structures has been proposed by Hamamoto *et al.* [10]. In this mechanism, generated cluster structure can adapt dynamic environment, and it also can generate stable spacial structure under the fixed initial condition.

Let us consider a one-dimensional network model for simplicity, and let $q(i,t)$ be the value of distribution at node ID i at time t. The distribution $q(i,t)$ determines the cluster structure. Examples of the initial condition $q(i,0)$ and cluster structure $q(i,t)$ obtained at time t are shown in Fig. 1. The conventional back-diffusion based approach described in Takano *et al.* [9] presents a rule governing the temporal evolution of the distribution $q(i,t)$. However, as shown in the previous section, it is difficult to guarantee the stability of $q(i,t)$ for large t. In other words, the cluster structure is not stable in a dynamic environment.

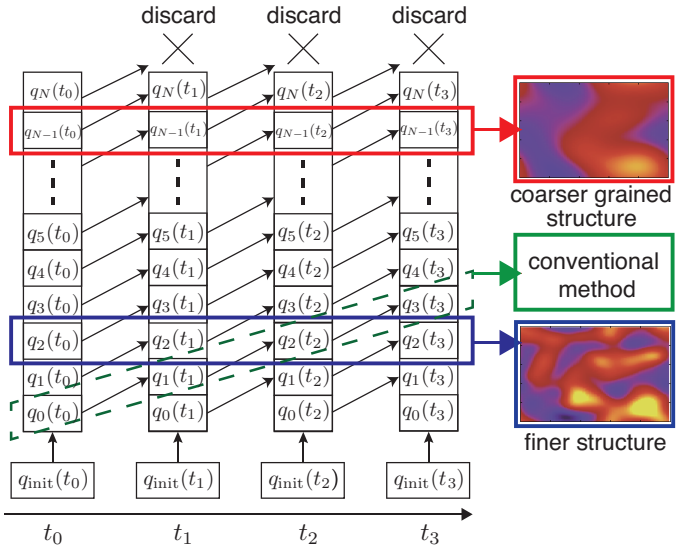

Fig. 2. Outline of the guaranteeing mechanism of asymptotic stability.

Let us consider discrete time t_k $(k = 1, 2, \dots)$, and distribution $q(i,t_k)$. Since we need an autonomous decentralized algorithm, the temporal evolution of distribution $q(i,t_k)$ is determined by its local information. By introducing the temporal evaluation operator of $\mathcal{T}$, the temporal evolution is formally described as

$$q(i,t_{k+1}) = \mathcal{T}(q(i-1,t_k),\, q(i,t_k),\, q(i+1,t_k)) \qquad (1)$$

This rule states that the distribution of node i at the next time is completely determined by the values of the present distribution at node i and its adjacent nodes.

To guarantee the mechanism of asymptotic stability, we consider a vector of the distribution. Each node i has the following $N+1$ dimensional vector

$$\boldsymbol{q}(i,t_k) = \{q_0(i,t_k),\, q_1(i,t_k),\, \dots,\, q_N(i,t_k)\} \qquad (2)$$

Here, we define the rule for the temporal evolution of the vector $\boldsymbol{q}(i,t_k)$. Let $q_{\mathrm{init}}(i,t_k)$ be the distribution describing the network state (e.g., battery power of a node) at time t_k. Then we set

$$q_0(i,t_{k+1}) = q_{\mathrm{init}}(i,t_{k+1}) \qquad (3)$$

If $q_{\mathrm{init}}(i,t_k)$ is independent of time, $q_0(i,t_{k+1}) = q(i,0)$, that is, the initial condition of the conventional mechanism. Note that, in general, we allow the time-dependence of $q_{\mathrm{init}}(i,t_k)$. Next, for $q_{n+1}(i,t_{k+1})$ $(n = 0,\, 1,\, \dots,\, N-1)$, we set

$$q_{n+1}(i,t_{k+1}) = \mathcal{T}(q_n(i-1,t_k),\, q_n(i,t_k),\, q_n(i+1,t_k)) \qquad (4)$$

Although the above rule may look complicated, we can easily understand it through graphical representation. Fig. 2 explains the temporal evolution of vector (2) at node i. The horizontal axis represents discrete time as t_0, t_1, $\dots$, and $q_{\mathrm{init}}(t_k)$ expresses a certain metric of network state of a node at time t_k. Each component of the vector is a distribution value and (4) is the temporal evolution rule for the nth component $(n = 1, 2, \dots, N-1)$. The temporal evolution of each component will be updated to the upper-right component in

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

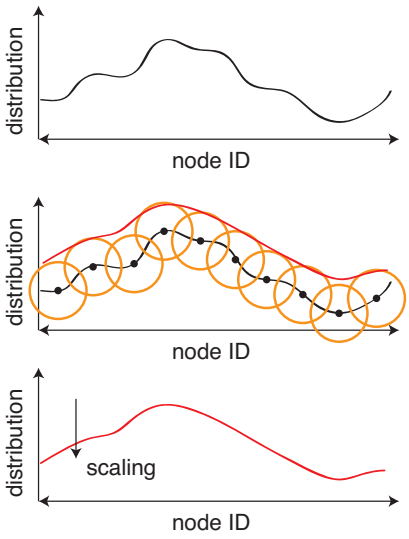

Fig. 3. Renormalization transformation as per Huygens' principle.

Fig. 2. The component at the bottom, $q_0(i, t_{k+1})$, is overwritten by the network condition $q_{\rm init}(i, t_{k+1})$ at the present time. The component at the top will be discarded.

The temporal evolution of the conventional mechanism corresponds to the sequence indicated by the green broken line in Fig. 2. In the guaranteeing mechanism of asymptotic stability, we focus on the sequence of the same vector components. If we choose small n for the nth component, we obtain a finer-grained spatial structure as indicated by the blue line. A large n yields a coarse grained spatial structure as indicated by the red line.

III. Design of the Autonomous Structure Formation Technology based on Huygens' Principle

A. Huygens' principle and Renormalization

Huygens' principle [11] describes the temporal evolution of the wavefront and can explain the laws of reflection and refraction. Let us consider spherical waves originating at each point on a wavefront. The envelope of these spherical waves gives the temporal evolution of the wavefront. This is called Huygens' principle or the Huygens-Fresnel principle. Renormalization is a way to extract simple and important macroscopic characteristics from a large-scale and complex system, and its procedure is defined as the renormalization transformation. This procedure is suitable for generating a simple cluster structure extracted from the spatial structure of the network state. The renormalization transformation is defined as the combination of coarse-grained transformation and scaling. In this paper, we adopt the renormalization transformation based on Huygens' principle as temporal evolution operation $\mathcal{T}$. Concrete procedures of the renormalization transformation are shown below.

Let us consider a one-dimensional network and a distribution on the network. The panel at the top of Fig. 3 shows an example of the distribution at the present time. We consider the shape of the distribution as the wavefront. The panel at the middle of Fig. 3 shows the temporal evolution of the wavefront as given by Huygens' principle. This procedure has smoothing effect such that the fine-grained structure that takes the shape of the distribution becomes smooth. The temporal evolution of the distribution causes an increase in the value of the distribution, that is, the wavefront proceeds upward. In order to compensate for this increase, we introduce scaling as shown in the panel at the bottom of Fig. 3. We define the renormalization transformation as the combination of such temporal evolution and scaling.

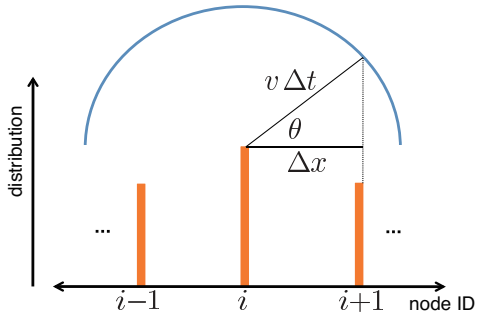

Fig. 4. The wavefront of the spherical wave reaches an adjacent node.

Let the value of distribution at node i at time t_k be $q(i, t_k)$, and let the set of nodes that are adjacent to node i at time t_k be $M(i, t_k)$. In addition, $\tilde{q}(i, j, t_{k+1})$ is the wavefront of sphere wave at node i at time t_{k+1} that originated from node j at time t_k. Our renormalization transformation is expressed as

$$q(i, t_{k+1}) = \frac{1}{b} \max_{j \in M(i,t_k)} \tilde{q}(i, j, t_{k+1}), \tag{5}$$

where, the maximizing operation in (5) means Huygens' principle; it determines the most advanced wavefront of spherical waves that originated from the node itself and its neighborhood, and $b > 1$ is the scaling parameter.

Next, we consider the concrete form of $\tilde{q}(i, j, t_{k+1})$. Let the propagation speed of spherical wave be v, the distance between two adjacent nodes be Δx, and the interval of the temporal evolution (renormalization transformation (5)) be Δt (i.e., $t_{k+1} - t_k = \Delta t$). Here, Δx is not physical distance but is a kind of hop count, so we can choose $\Delta x = 1$. We consider the situation that the temporal evolution (5) is determined only by adjacent nodes, v is chosen as $1 \le v\,\Delta t < 2$. As shown in Fig. 4, the wave front of spherical wave originated from node i influences both node i and its adjacent nodes. It is expressed as

$$\tilde{q}(i \pm 1, i, t_{k+1}) = q(i, t_k) + v\,\Delta t\,\sin\theta, \tag{6}$$

$$\tilde{q}(i, i, t_{k+1}) = q(i, t_k) + v\,\Delta t, \tag{7}$$

where θ is a constant and, from $v\,\Delta t\,\cos\theta = \Delta x$,

$$\theta = \arccos\left(\frac{\Delta x}{v\,\Delta t}\right) \tag{8}$$

Since v, Δt, and $\sin\theta$ are constants and we can know them in advance, the temporal evolution (5) is a simple operation. In addition, with iterations of the temporal evolution (5), the maximum value of distribution $q(i, t_k)$ converges to $p^* := v\,\Delta t/(b-1)$, regardless of initial condition $q(i, 0)$.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

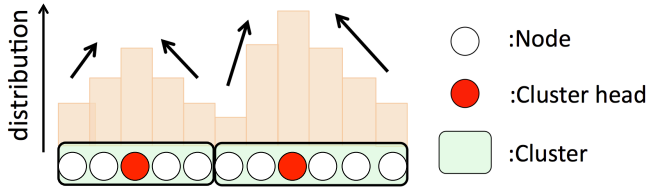

Fig. 5. Determination of the cluster and cluster head.

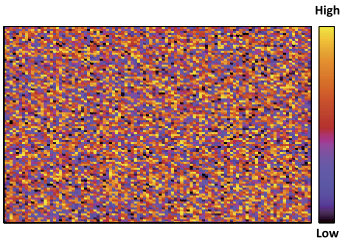

Fig. 6. An example of randomized initial condition

B. Amplification of Amplitude of the Distribution

Our renormalization transformation (5) makes the distribution flat and we can obtain coarse-grained spatial structure. However, differ from physical phenomena, there are situations that the distribution does not change when the difference of distribution values are small. This is because the positions of nodes in the network are discrete. If the value of the distribution at a node can affect that of the adjacent node, the following relation is required,

$$|q(i \pm 1, t_k) - q(i, t_k)| > v\,\Delta t\,(1 - \sin\theta) \qquad (9)$$

When the smoothing proceeds and the condition (9) is no longer met, two adjacent nodes do not interact and the distribution is unchanged. To avoid this phenomenon, we introduce amplification of the amplitude of the distribution in addition to the renormalization transformation (5). The additional operation is

$$q(i, t_{k+1}) \leftarrow p^* + a\,\left(q(i, t_{k+1}) - p^*\right), \qquad (10)$$

after the renormalization transformation (5). This operation means that the difference between the value of distribution and p^* is amplified by a factor of a times. Here, aforementioned $p^* = v\,\Delta t/(b-1)$ is the fixed point of the renormalization transformation, and also is the convergence point. The value of the parameter a should be chosen as $a > b$.

Finally, we explain how to determine clusters and cluster heads from the generated spatial structure (Fig. 5). By following the direction of the steepest gradient of the distribution, we can find a node with local maximum value. We define it as a cluster head, and the nodes belonging to the same cluster head belong to the same cluster.

IV. Performance Evaluation

This section investigates the convergence speed with respect to the range of the initial distributions, and proposes a way to guarantee the controllability of our proposed mechanism. In addition, we verify the initial distribution dependency of our clustering.

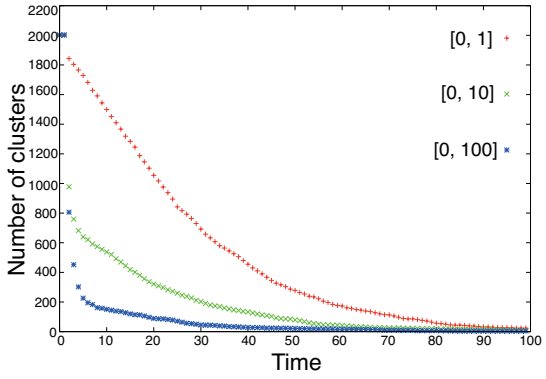

Fig. 7. The number of generated clusters for randomized initial conditions w.r.t. time.

A. Dependence Characteristics on the Range of Initial Distribution

First, we investigate the convergence speed with respect to the range of the initial distributions. The network model in this evaluation is a two-dimensional lattice network with torus boundary and it has 100×100 nodes. The reason of torus boundary condition is to eliminate the effect of the network edge, and to concentrate our attention on the characteristics of clustering mechanism itself. The initial distribution of $q(i, 0)$ for all the node position, i, are given by a uniform distribution; three kinds of uniform distributions are examined: their ranges are $[0,\ 1]$, $[0,\ 10]$, and $[0,\ 100]$. An example of an initial condition is shown in Fig. 6.

We calculate the temporal evolution of the distribution by using (5) and (10), every discrete time. The interval of discrete time is set to be $\Delta t = 1$. Here, we investigate the change in the number of generated clusters. Fig. 7 shows the temporal evolutions of cluster number, from three different initial conditions. The parameters were set as $v = 1.5$, $a = 1.2$, $b = 1.1$. From this result, we can recognize that the number of clusters strongly depends on the range of the initial condition. Huygens' principle or the maximizing operation in (5) has strong impact when the adjacent nodes have very different values.

This characteristic triggers a loss of control over cluster size (or the number of clusters). The number of clusters in the initial state (the number of local maximums in Fig. 6) is about $2,000$ as shown in Fig. 7. The horizontal axis denotes time but also corresponds to iteration number of temporal evolution or the component of the vector (2). If we need $1,000$ clusters, we can choose about 20 iterations or the 20th component of vector $\boldsymbol{q}(i, t_k)$, for the initial condition of $[0,\ 1]$. However, we cannot choose the appropriate value for $[0,\ 100]$. Since we cannot know the network condition in advance, we cannot control the number of clusters.

Let us consider how this situation corresponds to difficulty in controllability. If we describe the initial condition of nodes by their battery power, we should express the battery power in numerical value. There are many ways to express the battery power in numerical value: ampere-hour [Ah], mili-ampere hour [mAh], coulomb [C], etc. Incidentally, 1 Ah = 1,000 mAh = 3,600 C. The above difficulty in controllability implies that

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

the initial distribution having different range gives different clustering structure even if the distributions come from the same physical situation (only difference is in way to express it).

B. Robust Controllability of Cluster Size

The cause of the above problem is the excessive sensitivity of cluster forming with respect to the range of the distribution. The mechanism of the excessive sensitivity with respect to the range of the distribution can be recognized by Fig. 8. Fig. 8 shows the behaviors of the proposed clustering mechanism for the distributions having wide and narrow ranges. First, each node performs temporal evolution obeying Huygens' principle, and next, scaling. Small-valued node, which is next to the large-valued node, is greatly influenced by the large-valued node, and the difference of the values between them is rapidly decreasing. So, if the range of the initial distribution is wide, the distribution is rapidly uniformized.

In order to avoid the above problem, we redefine the initial condition. The details are as follows. We do not use the network condition directly as the initial condition, but we use

$$q(i,0) = \log(1 + q_{\text{init}}(i,0)) \tag{11}$$

In the vector formulation, we replace (3) with

$$q_0(i,t_{k+1}) = \log(1 + q_{\text{init}}(i,t_{k+1})) \tag{12}$$

The reasons for introducing a logarithmic function are as follows:

- It is possible to maintain the magnitude relation of the value of the original initial distribution.
- As a value of the original initial distribution is large, new value is smaller in the sense of the ratio.

Fig. 9 shows similar evaluations by using the redefined initial condition (11). We recognize that the impact of the initial condition is sufficiently weakened. In particular, the initial conditions of $[0, 10]$ and $[0, 100]$ yield almost the same result. This means we have a robust clustering mechanism that can control the number of clusters by appropriately choosing the number of iterations or the component of the vector (2).

C. State Dependent Characteristics of Clusters

Since the initial condition reflects the network state (e.g., battery power of each node), the generated cluster must be influenced by the initial condition. This subsection introduces a numerical example that shows that the cluster structures generated by the proposed mechanism do reflect the network condition.

We use the same network model and parameter setting as the previous subsection. To verify the dependence on initial conditions, we use two types of initial conditions. One is a randomized condition, that was used in the previous subsection. This initial condition is determined by logarithm of random values, which obey a uniform distribution with range of $[0, 10]$, as shown in Fig. 6. The other initial condition has spatial patterns as shown in Fig. 11. Three areas have relatively high values, and the value of these areas is determined by random values which obey a uniform distribution with range $[0, 10]$. The value of other area is determined by random values which obeys the uniform distribution with range $[0, 1]$. The logarithm of these values is used as the initial condition.

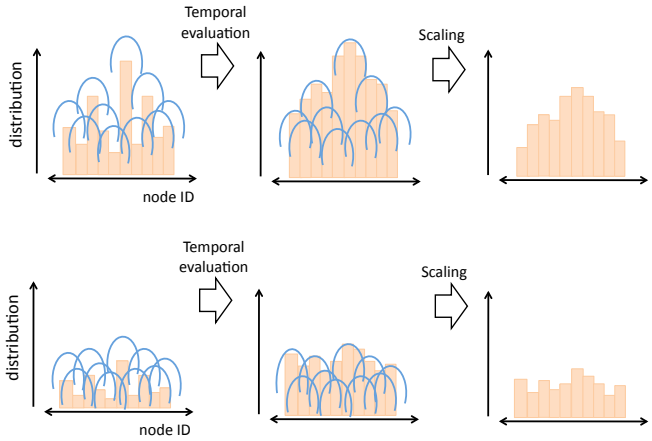

Fig. 8. An example of behavior of the our proposed mechanism on two different distribution ranges

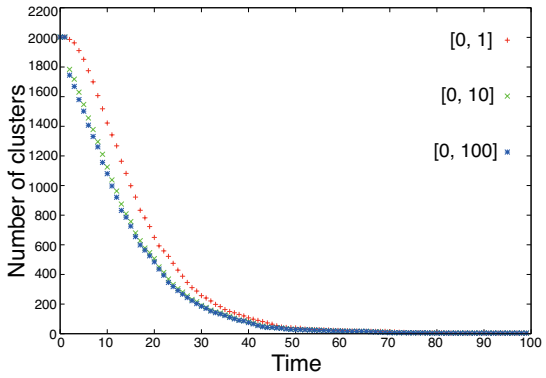

Fig. 9. The number of generated clusters for logarithmic initial conditions, w.r.t. time.

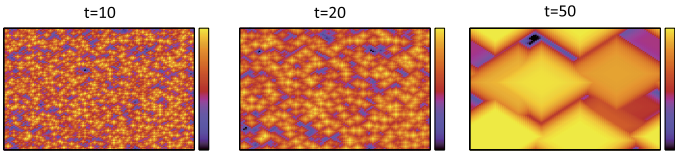

Fig. 10. The cluster structures generated from randomized initial condition.

Figs. 10 and 12 show the cluster structures generated from the initial conditions of Figs. 6 and 11, respectively. The three panels of each figure show the number of iteration or, equivalently, the vector component. If we choose few iterations, we obtain a finer-grained cluster structure, and if we choose more iterations, we get a coarse grained cluster structure. We can also recognize that the cluster structures reflect the spatial structures of the initial conditions.

V. Conclusion and Future Work

In this paper, we proposed an autonomous clustering mechanism based on Huygens' principle and renormalization. In

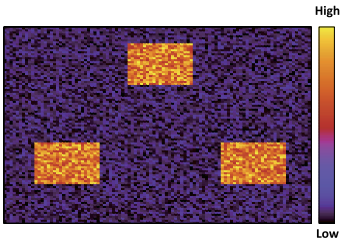

Fig. 11. An example of initial conditions with the spatial structure

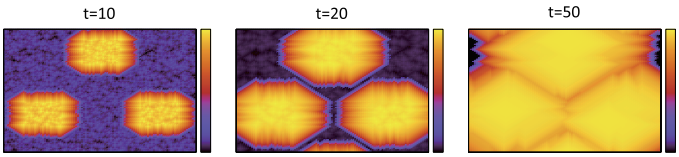

Fig. 12. The cluster structures generated from the initial conditions with spatial structure.

verification, we used a two-dimensional lattice network to evaluate the characteristics of the proposed algorithm. The pros of the proposed algorithm are in its simplicity and in the ability to keep the spatial structure of the initial condition in the configuration of clusters. However, unfortunately, the convergence speed of cluster configuration strongly depends on the value of the initial distribution. Since we can not know the value of distribution for each node in advance, the difference of convergence speed causes the situation that we cannot control the number of clusters. To avoid this problem, we introduced new distribution defined by the logarithm of the original distribution. Consequently, the difference of the convergence speed is significantly reduced, and the number of clusters becomes controllable. The above obtained characteristics are suitable for clustering in MANET. As future work, we will consider the adaptability of our mechanism in dynamic environment.

ACKNOWLEDGMENT

This work was supported by a Grant-in-Aid for Scientific Research (B) No. 23300031 (2011-2013) from the Japan Society for the Promotion of Science, and by National Institute of Information and Communications Technology (NiCT).

REFERENCES

[1] C.E. Perkins, Ad Hoc Networking, Addison Wesley, December 2000.

[2] C.E. Perkins, E.M. Royer, and S. Das, “Ad hoc on-demand distance vector (AODV) routing,” Request for Comments 3561, July 2003.

[3] J. Broch, D.A. Maltz, D.B. Johnson, Y.C. Hu, and J. Jetcheva, “A performance comparison of multi-hop wireless ad hoc network routing protocols,” ACM International Conference on Mobile Computing and Networking (MobiCom ’98), 1988, pp. 85–97.

[4] S. Seys, and B. Preneel, “ARM: Anonymous routing protocol for mobile ad hoc networks,” International Journal of Wireless and Mobile Computing (IJWMC), 2009, pp. 145–155.

[5] S. Basagni, “Distributed clustering for ad hoc networks,” Proceeding of International Symposium on Parallel Architectures, Algorithms and Networks, 1999, pp. 310–315.

[6] C.E. Perkins, E.M. Royer, and S. Das, “Ad hoc on-demand distance vector (AODV) routing,” Request for Comments 3561, July 2003.

[7] T. Ohta, S. Inoue, Y. Kakuda, and K. Isida, “An adaptive multihop clustering scheme for ad hoc networks with high mobility,” IEICE Transactions on Fundamentals, 2003, vol. E86-A, no. 7, pp. 1689–1697.

[8] C. Takano, M. Aida, M. Murata, and M. Imase, “New framework of back diffusion-based autonomous decentralized control and its application to clustering scheme,” IEEE Globecom 2010 Workshop on the Network of the Future (FutureNet III), Miami,USA, December, 2010, pp. 362–367.

[9] C. Takano, M. Aida, M. Murata, and M. Imase, “Proposal for autonomous decentralized structure formation based on local interaction and back-diffusion potential,” IEICE Transactions on Communications, May, 2012, vol. E95-B, no. 5, pp. 1529–1538.

[10] R. Hamamoto, C. Takano, K. Ishida, and M. Aida, “Guaranteeing asymptotic stability of clustering by autonomous decentralized structure formation,” The 9th IEEE International Conference on Autonomic and Trusted Computing (ATC 2012), Fukuoka, Japan, September 2012.

[11] David A. B. Miller, ”Huygens’s wave propagation principle corrected,” Optics Letters, September, 1991, Vol. 16, p.p. 1370-1372.

[12] K. Takagi, Y. Sakumoto, C. Takano and M. Aida, “On convergence rate of autonomous decentralized structure formation technology for clustering in ad hoc networks,” IEEE ICDCS 2012 Workshops (ADSN 2012), Macau, P.R.China, June 18-21, 2012.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Context Awareness Monitoring Model for Smart Office Environment

Nor Azlina Aziz Fadzillah, Nasiroh Omar, Siti Zaleha Zainal Abidin, and Nor Azylia Ahmad Azam
Faculty of Computer And Mathematical Sciences
University Technology MARA,
Shah Alam, Malaysia
e-mail: norazlina@perlis.uitm.edu.my; nasiroh@tmsk.uitm.edu.my;
sitizaleha533@salam.uitm.edu.my; azylia@tmsk.uitm.edu.my

***Abstract*—Administrators conventionally write information on paper or on a notice board to monitor subordinates' activities within an organization. Unfortunately, the information cannot be stored and retrieved for future use. This paper presents a case study for the implementation of a smart notice board. The analysis focuses on the presence and transition of staff work activities. The staff activities are evaluated to allow administrators to be aware of their work status. Based on the smart office board design, we propose a context-aware monitoring model in a smart office environment. This model provides a platform to gather useful information of staffs during their working hours. Therefore, top-level administrators can identify staff status (present, absent or busy) and perform statistical analyses on staffs' availability. Thus, they can plan, schedule, make decisions and prepare reports instantly.**

Keywords-collaborative; context-aware; monitoring user activities; workspaces; smart office; smart notice board.

I. INTRODUCTION

Generally, an organization uses a notification board to inform their staff of theirs' daily activities. The activities depend on the current work mode, which deals with different tasks, tools, and information [1]. Nevertheless, networked collaborative virtual environments allow workers to work in a shared virtual space at real-time. However, virtual space lacks contextual elements as in a real life-working situation. The virtual workspace should make optimum use of human computer interaction. It should provide relevant information or services to adapt the system behaviours for a specific user [2]. In addition, context-awareness is one of the vital features, which helps users to map their tasks by considering the parallel interaction in networked virtual environments [3].

A context-aware service should provide accurate information to assist people in setting goals, making plans and performing the necessary actions in their daily work activities. Recognizing people activities in a virtual work space will help users to react automatically to a person's identity, location, activity, and facial expression [4]. They should be informed about others' work to obtain the information at the right time, right place and communicate in a meaningful way.

Today, most people spend a lot of time working in spaces, like offices and committee rooms. Most of the researchers are focused on the development of smart and intelligent working environments [5]. A smart environment is defined as a situation in which a user interacts with another user in a flexible, personalized and domain-specific manner [6],[7]. The goals of a smart environment are to support and provide task-relevant information services to sense context in a variety of user environments [8]. The smart environment can provide a unique service in designing, researching and addressing issues in a multi-agent system [9]. There are some examples of smart environments such as the smart home [10], [11], [12], the smart hospital [13], the smart freeway [14] and the smart office [5].

The smart office is defined as an environment that enables the user to adapt user needs, routine tasks, preferences, and access on-time services [15]. It represents a user's work, as in a normal office. In addition, it allows the user to acquire and apply knowledge to improve their experience and synchronous communication in a particular environment [7],[16]. The synchronous communication and data are integrated with other information to enable users to gain knowledge of current and past activities in the working environment [1].

Furthermore, the smart office handles several devices that supports everyday tasks, anticipates user intentions, and reduces the conflict in decision making [5]. There are several examples of smart office systems such as intelligent meeting rooms [5], the virtual secretary [17], and the smart doorplate project [18]. This paper focuses on designing a smart notice board as part of the smart office environment. Designing a context-aware system in a smart office environment is a complex and challenging issue. Most of the studies do not fully address specific characteristics of context management [19]. Therefore, in Section II, we will analyze and categorize the most common characteristics used in the smart office environment.

The rest of this paper is organized as follows: Section III observes smart office scenarios and related activities for a smart office in an academician environment. Next, Section IV presents a context-aware monitoring model for the smart office environment. Finally, Section V discusses the conclusion in relation to the research work.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

II. CONTEXT CATEGORIZATION

Human activities can be classified into personal, community, and social activity [20]. Each of the activities will involve different types of context depending on the location, time, resources, task, and situation. The context has been used as a tool to select the correct action, improve the quality and efficiency of case-based reasoning within the diagnostic domain virtual workspace [21]. When dealing with a different state or situation, context is the key element used to infer possible action and support the user with contextual information [22]. It is crucial to understand context and its relation with adaptability (i.e. device, environmental setting, and time) to construct context-aware software development that is beneficial to the users [23].

Many researchers define and classify context to support and develop context-aware systems. Dey and Abowd [24] identified four primary context types (i.e., location, identity, time, and activity). It provides task-relevant information and interaction between the user and the context-aware application. Prado [2] defines three types of context (i.e., device-related, environmental and user context). This context deals with the impact of application on different user interfaces, information and functionality of context. In order to model the relationship of activities and situations in a context-aware system, Kofod and Cassens [22] suggest five categories of user context (i.e., personal, task, social, spatio-temporal and environmental context).

Soylu et al. [23] propose eight categories to develop context-aware systems (i.e., user, device, application, information, environmental, time, historical and relational context). In addition, Villegas and Muller [25] classify five categories of context (i.e., individual, human, artificial, activity and compositional relations context) to control and govern context information within a smart environment. Furthermore, Kapitsaki et al. [26] apply three context categories within the tourist service (i.e., user, context information and services). It is found that user, system, physical, time and history are the most important context for mobile map applications [27]. Generally, it is concluded that most of the context-aware systems involve three main contextual entities: *user*, *activity* and *events*.

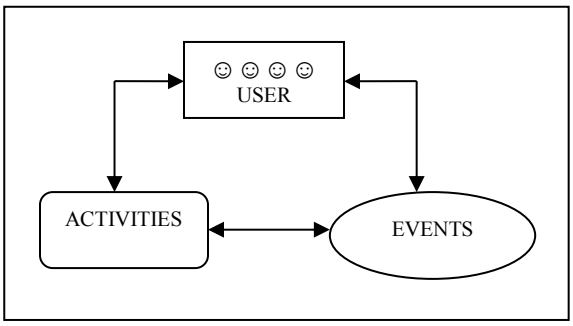

Figure 1. The contextual entities.

Fig. 1 depicts the most common context entities within the smart environment. All of these entities are dependent on the application, environment and goal, which the context-aware system is used. Based on this contextual entities, we observed the academicians' work modes to illustrate the context–aware activities within the smart office environment.

III. OBSERVING SMART OFICE ACTIVITIES

This study is conducted as an observation case study on staff activities during office hours. It is built on monthly activity that requires users to fill in information on the notice board. The aim of this case study is to extract important features to be used in designing a smart notice board. Contextual elements in the case study are analyzed in order to provide accurate information about staff work activities. The information is useful for the management to plan, schedule, and take appropriate action. Through the smart notice board, top-level management can identify whether staff are present, absent or busy. In addition, staff data can be used for further statistical analyses when top level management need to make decisions based on the system's report.

Scenario: Assume academicians are users in the smart notice board working environment. Let U be the set of lecturers = $\{U_1,\ldots..,U_n\}$. Lecturers can insert their monthly activities in a smart notice board as shown in Table I. The activities are based on daily activity, which involves contextual ontology. Contextual ontology is referred to as the shared understanding of some domains such as entities, relations, functions, and instances [28]. It provides a shared vocabulary and matches specific user queries, service descriptions and information in a context-aware system [29]. Each member of staff will input all monthly activities for the duration of the office hours (8.00 am – 5.00 pm).

TABLE I SMART NOTICE BOARD REPRESENTATION

Day / Person	1	2	3	……	……	……	31
U_1							
U_2							
U_3							
:							
:							
:							
U_n							

Users are categorized into two levels; academicians and administrators. The users' contextual activities can be in the form of teaching, research and development (RnD Context), administration and personal context. Contextual activity relates to an event based on three contextual elements,

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

which are sensed-data, time-line and location. Table 2 presents the details of the user classification, their contextual activities, and events as portrayed in Fig. 1.

TABLE II CONTEXTUAL ONTOLOGY EXAMPLE

USER	CONTEXTUAL ACTIVITY	EVENTS

• Administration
• Academicians
  • Dean (U_1)
  • Assistant Dean (U_2)
  • SeniorLecturer (U_3)
  • Lecturer (U_4)
	• Teaching Context
	  • Teaching(C_1)
	  • Evaluation (C_2)
	  • Invigilation (C_3)
	• RnDContext
	  • Conference (R_1)
	  • Consultation (R_2)
	• AdminContext
	  • Clerical Work (A_1)
	  • Meeting (A_2)
	  • Personal Context
	  • On Leave (P_1)
	  • Undefined (P_2)
		• SensorData
		• DateNTime
		  • Morning(T_1)
		  • Breakfast(T_2)
		  • Afternon(T_3)
		  • Lunch (T_4)
		  • Evening(T_5)
		• Location
		  • Home (L_1)
		  • Office (L_2)
		  • Classroom (L_3)
		  • Meeting room (L_4)
		  • Conference room(L_5)

The activities can be divided into two states: present and not present as described in Table 3. The data will be input based on individual activity. The contextual activities are classified as present (i.e., teaching, consultation and meeting) and not present (i.e., conference, on leave, undefined). For example, a teaching schedule will be extracted from a class schedule, and the faculty provides these schedules for each semester. Consultation and on leave will be input as individual activities and conferences are based on the application and calls for meeting. Each of the users will input their activities for the whole semester based on the academic calendar, except for on call activities.

TABLE III CONTEXTUAL ACTIVITY OF STAFF PRESENCE

User	Contextual Activity		Event
	Present	Not Present	
Dean Assistant Dean SeniorLecturer Lecturer	Class Consultation Meeting	Conference On Leave Undefined	Breakfast Lunch Classroom Meeting room Conference room

When dealing with a context-aware system, it should select the most appropriate content and features depending upon the usage context, such as context views [30]. Fig. 2 summarizes the context view for a smart notice board application.

User:	Contextual activities	Events:
Lecturer (U_4)	Teaching (C_1), Evaluation (C_2), Meeting(A_2), Consultation(R_2)	• Timeline : Morning(T_1),; Afternoon(T_3);Evening(T_5) • Location-: Office(L_2), Classroom(L_3), Meeting room(L_4)

Figure 2. Context entities for a smart notice board application

The context view is used to view shifts from one-node to other in the state transition activities. The graphical notation simplifies and helps the software developer to analyze, redesign and define the software requirements between users and environmental services [31],[32]. Fig. 3 shows a general context view of the state of transition for academician activities during office hours. This scenario shows how one user (U_4) represents normal daily activities, which involve the *time-line* (i.e., morning (T_1), afternoon (T_3), and evening (T_5)), *location* (i.e., office (L_2), classroom (L_3), meeting room (L_4)) and *contextual activities* (i.e., teaching (C_1), evaluation (C_2), consultation (R_2) and meeting (A_2)).

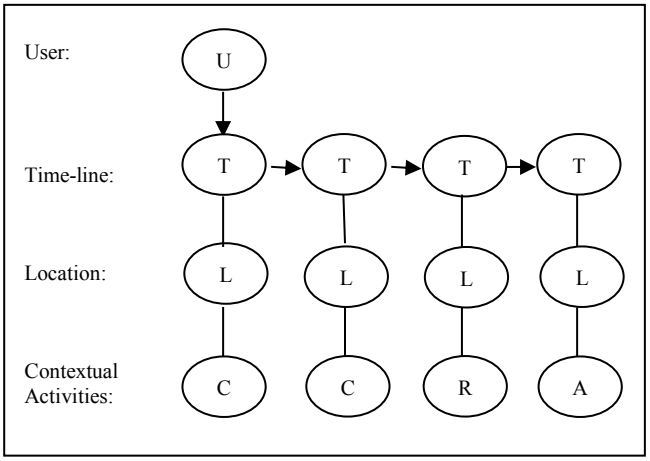

Figure 3. General context view of state transition activities

The activities are considered as dynamic since it might change from one-node to another node. All of the activities will be analyzed and monitored by top-level administrators. They should be aware of each activity in order to identify the work status whether the staffs are present, absent or busy. The provision of awareness is a key factor for keeping users up-to-date with what happens around them [33]. Therefore, a notification system is important to deliver current information effectively without causing unwanted distraction to ongoing tasks [34]. Through the notification system, the top-level administrators can use the information to make necessary actions and achieve the task objectives.

We illustrate four scenarios in academic work activities to determine the most available time for all academicians' staff. The scenarios show common activities for academicians, which involve three contextual entities (regardless of their location):

i. User - Dean (U_1), Assistant Dean (U_1), Senior Lecturer (U_1), Lecturer (U_4)
ii. Contextual activities - Available (√), Not available (X)
iii. Events - D ay (1, 2, 3, and 4)
Time slots: Morning (T_1), Breakfast (T_2), Afternoon (T_3), Lunch (T_4) Evening (T_5)

The four scenarios represent the daily activities from day one until day four. In each scenario, it will involve the four types of user (i.e., dean, assistant dean, senior lecturer, and lecturer). The contextual activities are considered as available (√) and not available (X). The available activities refer to free time, whereas not available means the users are on duty (i.e., teaching, conference, meeting, or on leave). The events are based on five standard time slots (morning, breakfast, afternoon, lunch and evening). The time slot for breakfast and lunch are blocked out as break-time. Each user will input data into the smart notice board system.

Scenario 1: In this scenario, it shows that all users are busy and not available.

TABLE IV SCENARIO 1 OF ACADEMICIANS ACTIVITIES

DAY	1				
TIME	T_1	T_2	T_3	T_4	T_5
USER	CONTEXTUAL ACTIVITIES				
U_1	X		X		X
U_2	X		X		X
U_3	X		X		X
U_4	X		X		X

Scenario 2: In scenario 2, it shows only 50% of the users available during the morning slot and most users are not available during the afternoon and evening slots.

TABLE V SCENARIO 2 OF ACADEMICIANS ACTIVITIES

DAY	3				
TIME	T_1	T_2	T_3	T_4	T_5
USER	CONTEXTUAL ACTIVITIES				
U_1	X		X		√
U_2	X		X		X
U_3	√		X		X
U_4	√		√		X

Scenario 3: In scenario 3, it shows 80% of the users are available in the evening slot.

TABLE VI SCENARIO 3 OF ACADEMICIANS ACTIVITIES

DAY	2				
TIME	T_1	T_2	T_3	T_4	T_5
USER	CONTEXTUAL ACTIVITIES				
U_1	X		√		X
U_2	√		X		√
U_3	X		X		√
U_4	X		X		√

Scenario 4: In scenario 4, it shows all users (100%) are available in the morning slot.

TABLE VII SCENARIO 4 OF ACADEMICIAN ACTIVITIES

DAY	4				
TIME	T_1	T_2	T_3	T_4	T_5
USER	CONTEXTUAL ACTIVITIES				
U_1	√		X		X
U_2	√		√		X
U_3	√		X		X
U_4	√		√		X

The scenarios show, that the most available time for all staff is during scenario 3 (80%) and scenario 4 (100%). Therefore, the administrator can monitor and make a decision to do any activity such as call a staff meeting, a conference or an evaluation of staff performance. As data is stored within the computer system, statistical reports along with general reports can be produced at real time. Based on the context awareness-monitoring model, a smart notice board system will be developed.

IV. CONTEXT AWARENESS MONITORING MODEL

Adapted from a previous study (Fig. 4), the monitoring manager will interpret, filter and select rules to analyze and manage the interactive data.

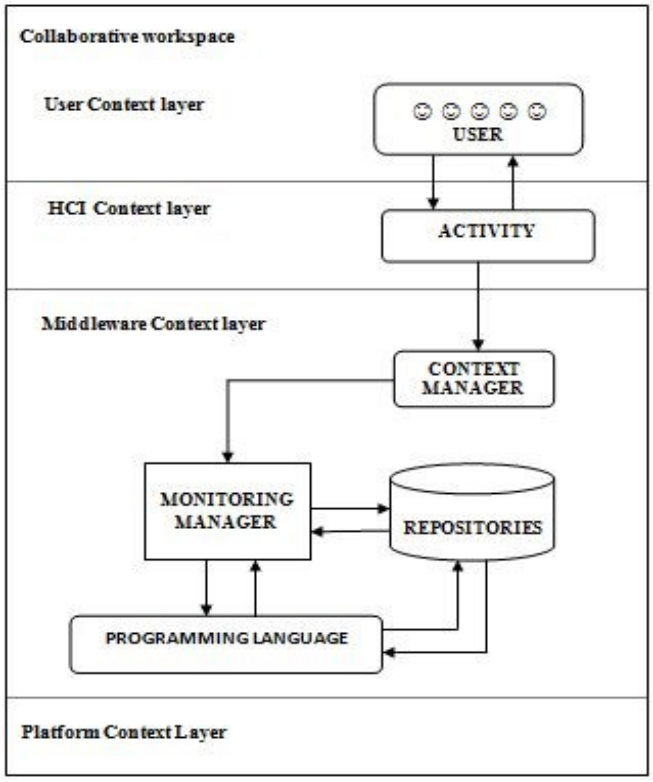

Figure 4. A general context awareness monitoring model [35]

Based on the four scenarios, the rules represent the logic and reasoning mechanism to choose the most available time for all staff.

Scenario 1-4 : General Rules

IF (user({U_1,.....,U_n})
AND (day(({D_1,.....,D_n})
AND (time(({T_1,.....,T_n})
THEN
isAvailability ({yes, no})

The monitoring process deals with the communication between at least two of the users who should observe their enviroment and activities (shown in Fig. 5). The awareness provides information about the activity and the availability of all users. Besides the rules, the *contextual media* (e.g., text, graph, colour-coded, keyword, image, and messages), will be considered as an important elements to monitor user activities within the collaborative workspace [35],[36]. For example; the change of colour in text (e.g., available, or not available), shows different action has been taken by a user. The person who monitored the system, should also be aware of the signal given by the other user.

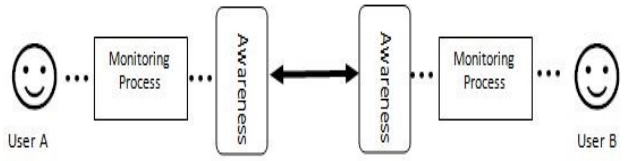

Figure 5. The user interaction in contextual awareness activities [36]

The implementation of the smart notice board system, involves the communication between an object (i.e., user, task, and resources) and the monitoring system [37]. This communication is governed by a monitoring process, which consists of the contextual monitoring components (i.e., contextual activity, event, time, and location) as depicted in Fig. 6. The sensor-data detects the presence of any action during the execution time in the monitoring process system.

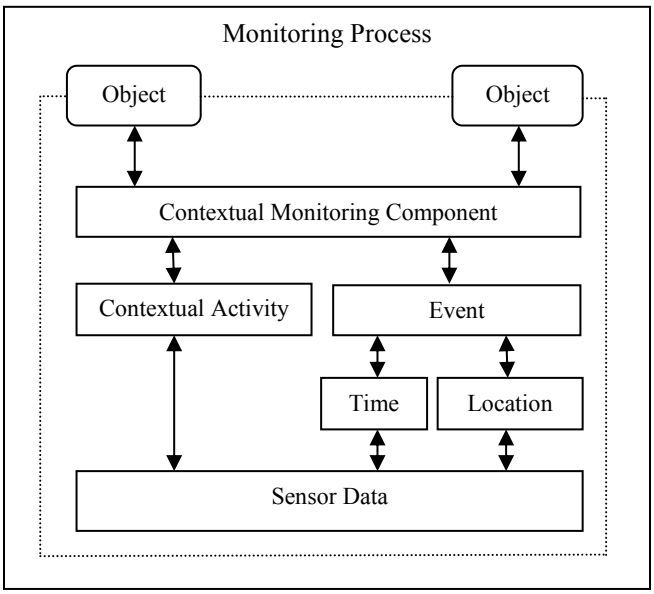

Figure 6. The monitoring process components

In order to design the smart office notice board system, all of the monitoring process components are mapped into the context-monitoring model (shown in Fig. 7).

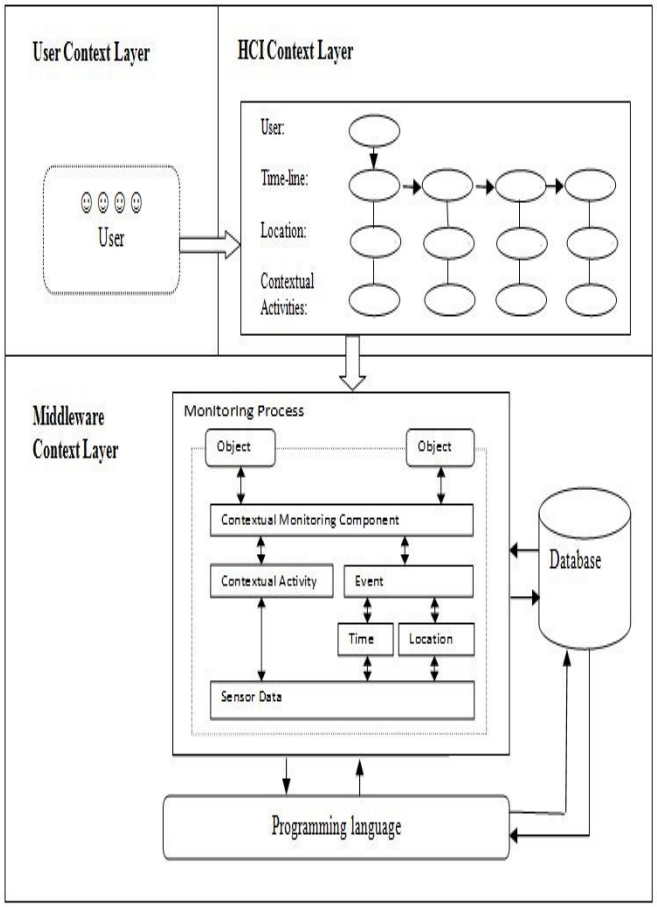

Figure 7. Context monitoring model for smart office environment

The context monitoring model design focuses on the development for a smart office notice board system. This system will be developed and manipulated by programmers in commonly available languages (e.g., Java, C, or C++). The data analysis is based on the user activities and the assessment on the availability of free time for each user. Basically, the monitoring process deals with the contextual media elements, which automatically senses the contextual activities. This will allow sensors to configure the appropriate action in response to user requirements. The sensors will recognize and respond to the performance of the activity. The result of tracking activities will be displayed in an assessment console. Through the data analysis, this information will be utilized and helps the top-level management in decision making.

V. CONCLUSION AND FUTURE WORK

This paper has analyzed user activities in a smart office environment. Basically, a context-aware system involves three common entities namely: user, activity and events. We

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

illustrated a scenario of contextual elements for academicians in their working environment by using a smart notice board prototype. The contextual awareness elements provide information about the activities and availability of all staff. Finally, a context awareness-monitoring model for the smart office environment has been proposed. The context-aware system should offer easier and more intuitive ways for the service provider to add and to subscribe new services through the historical context, log in service and feedback systems so as to enhance and better deliver a more appropriate quality of services [38]. Therefore, our future work will analyze and construct a general platform for monitoring a user activities system, which can be applied to any smart office environment.

ACKNOWLEDGMENT

The authors would like to thank Ministry of Higher Education and University Technology MARA Malaysia (RAGS Grant) for the financial support.

REFERENCES

[1] W. Geyer, H. Richter, L. Fuchs, T. Frauenhofer, S. Daijavad, S. Poltrock, and I. B. M. T. J. Watson, "A team collaboration space supporting capture and access of virtual meetings," 2001, pp. 188–196.

[2] A. G. De Prado, "Context-Aware Services : A Survey on Current Proposals," *Service Computation 2011, The Third International Conferences on Advanced Service Computing*, pp. 104–109, 2011.

[3] G. Niezen, B. J. J. van der Vlist, J. Hu, and L. M. G. Feijs, "From events to goals: Supporting semantic interaction in smart environments," *The IEEE symposium on Computers and Communications*, pp. 1029–1034, Jun. 2010.

[4] J. Krumm, S. Shafer, and A. Wison, "How a Smart Environment Can Use Perception," in *Workshop on Sensing and Perception for Ubiquitous Computing (part of UbiComp 2001)*.

[5] C. Ramos, G. Marreiros, R. Santos, and C. F. Freitas, "Smart Offices and Intelligent Decision Rooms," *Handbook of Ambient Intelligence and Smart Environments*, pp. 851–880, 2010.

[6] J. Weissman and Z. Zhang, "Smart Environments : Middleware Building Blocks for Pervasive Network Computing (A Position Paper)," *Developing an Infrastructure for Mobile and Wireless Systems*, pp. 90–105, 2002.

[7] M. Youngblood, D. J. Cook, and L. B. Holder, "Seamlessly Engineering a Smart Environment," *2005 IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 548–553.

[8] A. K. Dey, G. D. Abowd, and D. Salber, "A Context-Based Infrastructure for Smart Environments," pp. 1–15, 1999.

[9] D. J. Cook, "Multi-agent smart environments," *Journal of Ambient Intelligence and Smart Environments*, vol. 1, pp. 51–55, 2009.

[10] M. Lyons, P. Lyons, A. T. Cong, H. J. Steinhauer, S. Marsland, J. Dietrich, and H. W. Guesgen, "Exploring The Responsibilities Of Single- Inhabitant Smart Homes With Use Cases," *Journal of Ambient Intelligence and Smart Environments*, vol. 2, pp. 211–232, 2010.

[11] M. Jahn, M. Jentsch, C. R. Prause, F. Pramudianto, A. Al-akkad, S. Birlinghoven, and S. Augustin, "The Energy Aware Smart Home," in *Future Information Technology (Future Tech)*, 2010, pp. 1–8.

[12] A. Helal, D. J. Cook, and M. Schmalz, "Smart home-based health platform for behavioral monitoring and alteration of diabetes patients.," *Journal of diabetes science and technology*, vol. 3, no. 1, p. 141, Jan. 2009.

[13] A. Coronato, M. Esposito, and G. Pietro, "A multimodal semantic location service for intelligent environments: an application for Smart Hospitals," *Personal and Ubiquitous Computing*, vol. 13, no. 7, pp. 527–538, Apr. 2009.

[14] S. K. Lo, "A collaborative multi-agent message transmission mechanism in intelligent transportation system–A smart freeway example," vol. 184, pp. 246–265, 2012.

[15] I. Marsá-maestre, E. De Hoz, B. Alarcos, and J. R. Velasco, "A Hierarchical , Agent-based Approach to Security in Smart Offices," 2006.

[16] C. Röcker, "Services and Applications for Smart Office Environments - A Survey of State-of-the-Art Usage Scenarios," *International Conference on Computer and Information Technology (ICCIT 2010)*, pp. 385–401, 2010.

[17] M. Danninger and R. Stiefelhagen, "A context-aware virtual secretary in a smart office environment," in *Proceeding of the 16th ACM international conference on Multimedia, 2008*, 2008, pp. 529–538.

[18] J. Petzold, F. Bagci, W. Trumler, and T. Ungerer, "Next Location Prediction Within a Smart Office Building," *Cognitive Science Research Paper Unversity of Sussex CSRP*, vol. 577, p. 69, 2005.

[19] H. Vahdat-nejad, K. Zamanifar, and N. Nematbakhsh, "Context-Aware Middleware Architecture for Smart Home Environment," vol. 7, no. 1, pp. 77–86, 2013.

[20] Y. Lee, S. Oh, C. Shin, and W. Woo, "Ubiquitous Virtual Reality and Its Key Dimension" *Proc. of IWUVR*, vol. 9, pp. 5–8, 2009.

[21] M. Vladoiu and Z. Constantinescu, "U - Learning Within A Context-Aware," vol. 3, no. 1, 2011.

[22] A. Kofod-petersen, "Using Activity Theory to Model Context Awareness : a Qualitative Case Study," *Modeling and Retrieval of Context*, pp. 1–17, 2006.

[23] A. Soylu, P. De Causmaecker, and P. Desmet, "Context and Adaptivity in Context-Aware Pervasive Computing Environments," *2009 Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing*, pp. 94–101, 2009.

[24] A. K. Dey and G. D. Abowd, "Towards a Better Understanding of Context and," 1999, pp. 304–307.

[25] N. Villegas and H. Muller, "Managing dynamic context to optimize smart interactions and services," *The smart internet*, pp. 289–318, 2010.

[26] G. M. Kapitsaki, G. N. Prezerakos, N. D. Tselikas, and I. S. Venieris, "Context-aware service engineering: A survey," *Journal of Systems and Software*, vol. 82, no. 8, pp. 1285–1297, Aug. 2009.

[27] A. Nivala and L. T. Sarjakoski, "Need for Context-Aware Topographic Maps in Mobile Devices," *Proceeding of ScanGIS 2003*, pp. 15–29, 2003.

[28] T. Gu, X. H. Wang, H. K. Pung, and D. Q. Zhang, "An Ontology-based Context Model in Intelligent Environments," *Proceeding of communication networks*

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

and distributed systems modeling and simulation conference, pp. 270–275, 2004.

[29] T. Broens, S. Pokraev, M. Van Sinderen, J. Koolwaaij, and D. Costa, "Context-aware, ontology-based, service discovery," *Ambient Intelligent*, pp. 72–83, 2004.

[30] T. Hussein, W. Gaulke, and T. Linder, "Improving collaboration by using context views ¨," 2010.

[31] R. Paiano, A. Pandurino, L. Mangia, P. Marra, and M. Russo, "Collaborative Business Applications: A New Methodological Approach," *IBIMA Business Review Journal*, vol. 2011, pp. 1–18, Jun. 2011.

[32] A. R. Tripathi, D. Kulkarni, and T. Ahmed, "A specification model for context-based collaborative applications," *Pervasive and Mobile Computing*, vol. 1, no. 1, pp. 21–42, Mar. 2005.

[33] L. Ardissono and G. Bosio, "Context-dependent awareness support in open collaboration environments," *User Modeling and User-Adapted Interaction*, vol. 22, no. 3, pp. 223–254, Aug. 2011.

[34] L. McCrickard, D. S., Czerwinski, M., & Bartram, "Introduction: design and evaluation of notification user interfaces.," *Journal of Human-Computer Studies*, vol. 58(5), pp. 509–514., 2003.

[35] N. A. A. Fadzillah, N. Omar, and S. Z. Z. Abidin, "A Framework for Monitoring User Activities in Networked Collaborative Workspace." Unpublished.

[36] N. A. A. Fadzillah, N. Omar, and S. Z. Z. Abidin, "Activity-Based Entities Relationship in Monitoring Context Awareness," in *Context-Aware Systems and Applications*, Springer Berlin Heidelberg, 2013, pp. 1–9.

[37] N. A. A. Fadzillah, N. Omar, and S. Z. Z. Abidin, "Application-based context awareness in collaborative workspaces: A Review," 2012, pp. 411–415.

[38] S. Lee, J. Chang, and S. Lee, "Survey and Trend Analysis of Context-Aware Systems," *Information An International Interdisciplinary Journal*, vol. 14, no. 2, pp. 527–548, 2011.

Copyright (c) IARIA, 2013. **ISBN: 978-1-61208-287-5**

Intelligent Product States

Exploring the Synergy of Intelligent Products and State Characteristics in Collaborative Manufacturing

Benjamin Knoke, Thorsten Wuest, Karl Hribernik, Klaus-Dieter Thoben
BIBA - Bremer Institut für Produktion und Logistik GmbH
Bremen, Germany
{kno, wue, hri, tho}@biba.uni-bremen.de

***Abstract*—Along its lifecycle, a product passes through several states, which can be described by their characteristics. This concept of product state characteristics can be utilized in manufacturing process chains to improve the process quality by e.g., increasing transparency. The basic approach of describing a product state by its state-characteristics can be transferred into a collaborative manufacturing environment, however, in that case, a performing communication structure is indispensable. The goal of this paper is to explore if this concept is applicable for Intelligent Products and to what extent they can contribute to facilitate this endeavor. Accordingly, the idea of Intelligent Product States is introduced, including Intelligent Products sensing their own current state, interacting with the process and / or other products in order to achieve better results (e.g., quality) for themselves or following generations of products. Concluding, a classification model is presented to introduce a first structure of characteristics describing the different levels of Intelligent Product States (IPS).**

Keywords-Intelligent Products; Intelligent Product State; Product State Characteristics

I. INTRODUCTION

Effectivity and efficiency are terms that are repetitively taught in management seminars throughout the world. They imply the aim of maximizing the effect and minimize the use of resources in the actions of organizations. When transferred to the business objectives of a manufacturer, these principles manifest in the task of achieving the maximal product quality possible, with minimal use of resources. Both factors can be influenced by the actions of the manufacturer, which are taken within its manufacturing processes. The quality of these manufacturing processes is a key factor to determine the product quality and necessary effort of resources [1].

In a collaborative manufacturing network, qualitative manufacturing processes are not only important to evade costly manufacturing errors, but also to maintain a good reputation within the network. To achieve this, the improvement of process quality necessitates different types of knowledge to determine the current state, the targeted state, and the methodology to get there [2], [3]. From a deterministic point of view, these states can be described with certain characteristics that require knowledge about the current state of the product and the parameters of the process [4].

In order to gain and implement such viable knowledge within a collaborating manufacturing chain, certain capabilities are required. Upon sensing, communication, and interpretation, inter- and intra-organizational barriers may arise towards knowledge exchange [5].

This paper explores to what extent Intelligent Products can facilitate the crossing of the above mentioned barriers and support the measuring, storage, and communication of knowledge based on state-characteristics and process parameters within a collaborative manufacturing environment. Therefore, first, a brief introduction in product state and state-characteristics along with an application scenario in a collaborative environment is given. Followed by an introduction of Intelligent Products as the main pillars this paper is based on. Finally, these concepts are merged into a combined concept, named Intelligent Product State. Concluding, the impact of this new concept on collaborations and collaboration performance is explained.

II. PRODUCT STATE IN COLLABORATIVE MANUFACTURING

To provide an overview on the possibilities of state-characteristics to support collaborative manufacturing process chains, this section provides a definition, depicts application scenarios, and introduces intelligent products as a possible approach to realize a knowledge transfer based on state-characteristics.

A manufacturing process chain's purpose is to transform raw materials to final products through different value adding processes in order to satisfy the customer requirements. Consequently, the overarching goal of every manufacturing process chain is to add value to a work piece, component or product [6], (hereinafter the term *product* will be used comprehensively for a physical item) with each process step. Adding value in manufacturing implies physical transformation of the physical product (e.g., transformation of form, hardness, chemical composition, etc.). The specific purpose of every process step is to execute a part of the physical transformation of the product. Thus, the state of the product is changed at least with every (value adding) process. Looking at a product by its state has the advantage of being able to describe or record the transformation. Therefore, looking at the product state along the whole manufacturing process chain accumulates a complete picture of realized measures and transforming processes.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

A. Product State-Characteristics

The concept of the product state based view describes a product at certain times during a manufacturing process chain or after, through a combination of relevant state characteristics. As definable and ascertainable measures, state characteristics can be described in a quantitative or qualitative way. The state of a product changes due to external influence from one point in time to another when at least one descriptive state characteristic changes [4]. In other words, every product has a deterministic state that can be described at every time by its characteristics. These state-characteristics changes can appear during the whole product lifecycle, but are especially present during the beginning of life phase, the manufacturing stage.

The various characteristics can include measures of physical, chemical, or biological nature (e.g., location, dimensions, chemical composition, and internal tension).

However it is a challenge to identify a set of relevant state characteristics, which are sufficient to describe the product completely. Furthermore, the complexity is increasing when taking the influencing factors, most of all the manufacturing processes, triggering the state change into account.

The state characteristic change can, among others, be characterized as *intentional (planed)*, an example for an intentional state change is through machining, or *unintentional (unplanned)*, e.g., state change through corrosion when a metal part is stored out in the open. A slightly different characterization of state change is if the observer is *aware*, e.g., diameter change after machining or *unaware*, e.g., inflicting residual stress through clamping, of the state change. This aspect is very closely related to the following knowledge reasons.

Further difficulties occur on the attempt to derive an information model on the basis of the product state. The state characteristics and the influencing factors on state change can be *known*, e.g., influence of heat treatment of the harness of steel or *unknown*, e.g., effects of residual stress allocation on distortion *(knowledge reasons)*.

Another aspect, which is more of a technical *nature*, is the state characteristic change *measureable*, e.g., weight or *unmeasureable*, e.g., in an economical sensible way during the manufacturing process (*technical and/or economic reasons*).

The above reasons may make it difficult to derive certain state characteristics; they might though be relevant by having a significant impact on the quality of a product and thus the success of a manufacturing process chain. This impact is described by an exemplary application scenario within the next section, against a collaborative background.

B. Application Scenario for State-Characteristics in a Collaborative Manufacturing Environment

In the past decades a trend in manufacturing led from single integrated companies over static supply-chains to dynamic collaborations [7]. It has therefore to be taken into consideration, that manufacturing process chains are not executed at a single location, but spread geographically as well as organizationally instead.

An application scenario for the application of state-characteristics has been described by [8], as in Figure 1. It describes the manufacturing of a steel disc through the different process groups of casting, forging, machining, and heat treatment. In this level of detail, the steel disc passes through five different states, which can be described by certain characteristics. Due to the final heat treatment, internal tensions may become visible by resulting in a change of the disc's geometry. If these geometries exceed acceptable limits, the disc is rendered a defective good. As no practical models to incorporate these changes are available, the geometry is usually adjusted approximately before the heat treatment.

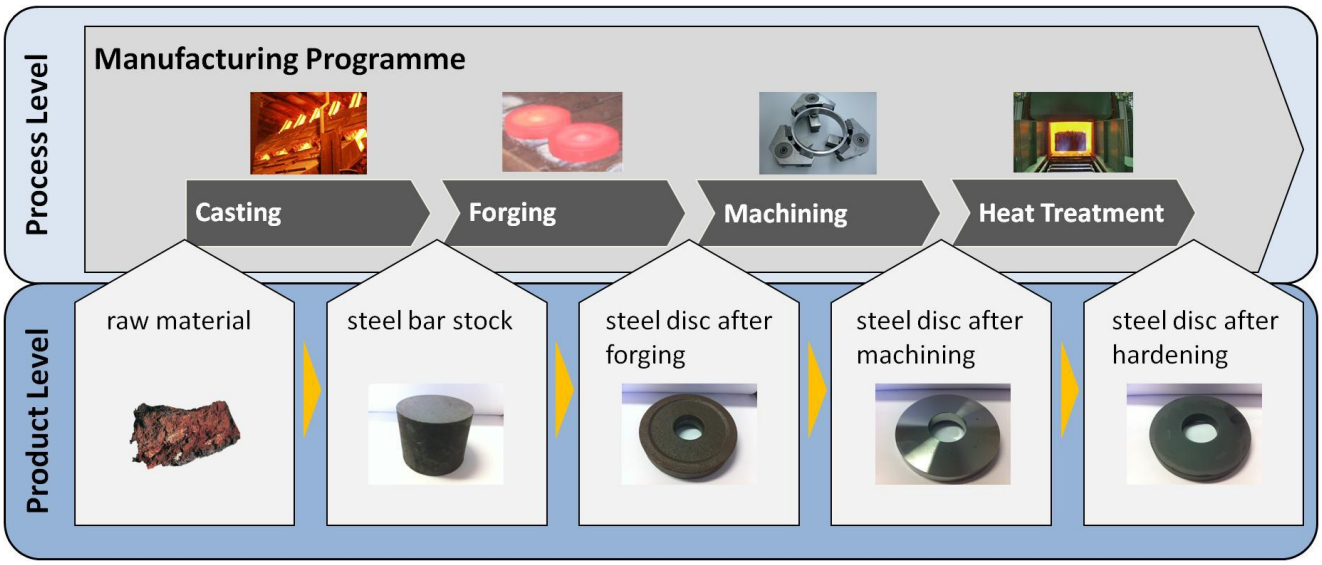

Figure 1. Transforming of a product during a multi-stage manufacturing programme [8]

One simple approach to manage this manufacturing process chain with state-characteristics is to establish a feedback loop about the geometrics of the steel disc after hardening to adjust the parameters of the machining process. This feedback adds transparency to the process chain and can thus increase its quality, assuming that the steel discs after forging have a sufficient homogeneous quality.

In a collaborative manufacturing environment, each process group of this manufacturing program can be executed by a different company. Such a feedback loop would therefore have to be established between organizations. This adds further complexity to the manufacturing, as additional logistics and information exchange needs to be implemented. This approach necessitates a clear structure within the organization of the collaboration to identify, share/distribute and use product and process information [9].

The shift towards collaboration in manufacturing paradigms is accompanied by an increasing development of ICT applications, which enable new approaches of operations management. Intelligent products are one outcome of this development, which can be combined with production states, to create the concept of Intelligent Product State (IPS), as described in the following section.

III. INTELLIGENT PRODUCTS TO SUPPORT THE USAGE OF STATE-CHARATERISTICS IN COLLABORATIVE MANUFACTURING ENVIRONMENT

The management of a manufacturing process chain from a product state based view can be supported by the capabilities of Intelligent Products. To explore this approach, this section provides a definition of Intelligent Products, and describes the concept of the Intelligent Product State (IPS) along a classification model.

A. Intelligent Products

Intelligent Products, also known as Smart Products, are physical items, which may be transported, processed or used and which comprise the ability to act in an intelligent ("smart") manner. McFarlane et al. (2003) defines the Intelligent Product as "[...] a physical and information based representation of an item [...] which possesses a unique identification, is capable of communicating effectively with its environment, can retain or store data about itself, deploys a language to display its features, production requirements, etc., and is capable of participating in or making decisions relevant to its own destiny" [10].

[11] characterizes intelligent products by describing related attributes:

- **Situatedness:** recognition of situational and community contexts.
- **Personalization**: in terms of tailoring the product according to buyer's and consumer's needs and affects
- **Adaptiveness**: the ability to change product behavior according to buyer's and user's responses and tasks
- **Pro-activity**: anticipation of user's plans and intentions
- **Business-awareness**: consideration of business and legal constraints
- **Network capability**: the ability to communicate and bundle with other products

These attributes describe the abilities of Intelligent Products, but are somehow qualitatively and leave a certain variance in their degree of fulfillment. Another dimension to determine the degree of intelligence of Intelligent Products can be added by addressing the characteristics of their tasks. They may exhibit varies from simple data processing to complex pro-active behavior. This is the focus of the definitions in [10] and [12]. Three dimensions of charactering Intelligent Products are suggested by [13]: Level of Intelligence, Location of Intelligence and Aggregation Level of Intelligence, as shown in Figure 2.

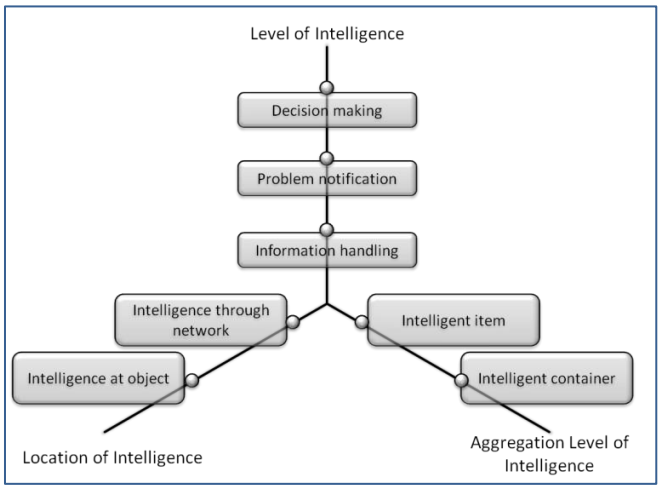

Figure 2. Classification model of Intelligent Products [13]

The first dimension describes whether the Intelligent Product exhibits information handling, problem notification, or decision making capabilities. The second shows whether the intelligence is built into the object, or whether it is located in the network. Finally, the aggregation level describes whether the item itself is intelligent, or whether intelligence is aggregated at container level.

Intelligent Products have been shown to be applicable to various scenarios and business models. For instance, [12] describe the application of the concept to supply network information management problems. Other examples are the application of the Intelligent Products to supply chain [14], manufacturing control [10], and production, distribution and warehouse management logistics [15]. A comprehensive overview of fields of application for Intelligent Products can be found in survey paper by [13].

Thus, an Intelligent Product is more than just the physical product – it also includes the enabling information infrastructure. Intelligent Products can make use of available advanced information infrastructures enhancing the quality of information and accessibility for humans who interact with them. Furthermore, Intelligent Products can make use,

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

e.g., of RFID, sensors, and embedded computing throughout their lifecycles in order to collect data for example about their usage, service, maintenance, upgrading, decommissioning and disposal. They thus can contribute significantly to closing the information loops throughout the product lifecycle and are fundamental to a holistic implementation of Closed-loop PLM in many products.

B. Intelligent Product State (IPS)

The gist of the Intelligent Product State (IPS) is the combination of Intelligent Products with the concept of the product state based view. This approach leads to Intelligent Products that support a product state based process management. Against this background, a product state can be intelligent, if its characteristics are monitored by an intelligent product. The extent of this support strongly depends on the product's capabilities. Their contribution can span from passive communication to corrective actions. The different levels of intelligence in managing with state-characteristic can be drawn according to the classification model of Intelligent Products, as shown in Figure 3.

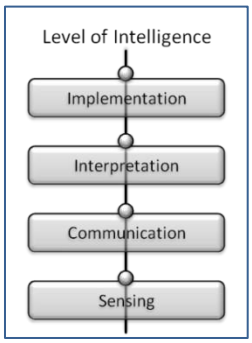

Figure 3. Levels of Intelligence of the Intelligent Product State

These levels of intelligence determine the ability of the intelligent products to process information about their state-characteristics:

- **Sensing capabilities** are the basic requirement to gather any information about state-characteristics. These capabilities require the knowledge about the relevant characteristics, and the methodology to measure them.
- **Communication** capabilities build upon the sensing state-characteristics. The ability of Intelligent Products to actively communicate the sensed information to an internal or external interpreter.
- **Interpretation** is the task of understanding the information by semantic analysis and contextualization.
- The **Implementation** of measures to change its own state is based on the previous interpretation and identified correction potential.

The latter determines a level of intelligence, at which products pro-actively participate in their own management. Based on one or multiple characteristics, they are capable of recognizing their own state, can identify an improved state, and possess the methodology to change accordingly.

Some exemplary applications of these capabilities could be:

- To enable an adjustment of functionalities during usage-phase (e.g., car-seat with settings management)
- To enable a complex customization of the product on instance level (e.g., a car-seat with memory functions for different users)
- To enable more complex functionalities (e.g., a car-seat providing settings in context of a driving situation)

When transferred to collaborative manufacturing, the IPS can be viewed from a different perspective that is described in the following section.

C. IPS in Collaborations

By equipping products with the ability to participate in the control of these processes, collaborators could facilitate the structure of their networks by reducing their necessary management efforts. This especially impacts on the early phase of collaboration. Nevertheless, a new layer of complexity occurs through the distributed process structure, but is accompanied by synergistic factors as well.

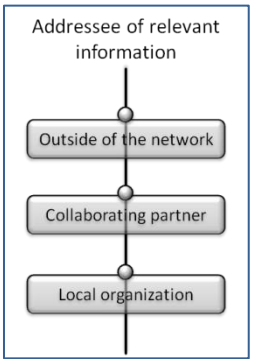

Figure 4. Addressee of Relevant Information of the IPS

If information concerning state-characteristics is obtained within a collaborative manufacturing environment, it might not be directly applicable, but instead address a process at another location. Thus, organizational barriers might be crossed from the sensing of a state-characteristic until the implementation of corrective measures. As visualized in Figure 4, the addressee of relevant information can be located:

- Within the local organization
- Aim at the processes of a collaborating partner
- Outside of the network

While information, addressing the local organization, might be easier to process, the interaction with partners and outsiders becomes more difficult, as not only intra- but inter-organizational barriers apply [5]. Although an Intelligent Product does not face barriers due to personal differences or a lack in encouragement, another layer of organizational barriers arises. Information has to be treated confidential according to the contractual basis and the level of trust in the collaboration [16]. When using IPS in collaborations, data

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

protection should though be considered. Especially, if IPS are active outside of the network, e.g., an Intelligent Product interacts with a user.

One approach to manage communication between Intelligent Products and users is the Product Avatar since its initial introduction as a technical concept, it has evolved into the concept of a customisable "digital representation" of product-related information [17].

This product-user interaction is influenced by the location of intelligence. While an independent intelligent product could also be used offline, a connection to a network is necessary, if the product is depending on the network's intelligence. In collaborative manufacturing process chains, a network connection is mandatory to not only just take correcting or improving actions, but also to allow IPS to learn upstream the manufacturing process chain. The location of intelligence can therefore be considered another dimension to describe the IPS.

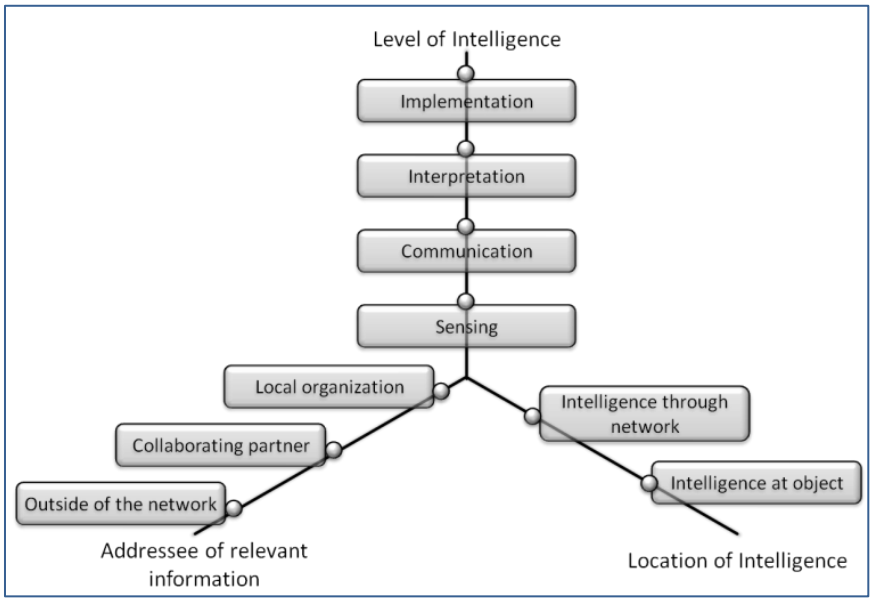

Figure 5. Classification Model of IPS

The Classification Model of IPS is comprised of the Level of Intelligence, the addressee of relevant information, and the location of intelligence. It has been summarized in Figure 5.

IV. CONCLUSION

The capabilities of Intelligent Products can contribute to increase the transparency and thus quality of a manufacturing process chain, based on the concept of product states. This combination has been described as IPS. In a collaborative manufacturing environment, these IPS can facilitate and improve a distributed management, but at the same time requires an open-minded culture throughout the process chain and faces concerns in sort of data protection.

The research of this paper was somehow limited, as only the application of IPS on the manufacturing of products has been considered. But also "virtual" state-characteristics, such as successful quality control, or customer satisfaction can be defined. Accordingly, an intelligent product-service combination, which is aware of its own state and is able to adjust itself to act in order to achieve a better result, could be applied to realize an Intelligent Service State.

This paper describes the concept of IPS, which is only the theoretical product of the combination of product states with Intelligent Products. This concept will be elaborated, and supported by a case study, which will be conducted for its application.

Upon this, future research will explore the characteristics of intra- and inter- organizational barriers to exchange information and knowledge against this background, as they significantly differ, as soon as the exchange is supported by the application of Intelligent Products.

ACKNOWLEDGMENT

This work has been funded by the European Commission through ICT Project BIVEE: Business Innovation and Virtual Enterprise Environment (grant agreement: 285746). The authors wish to acknowledge the Commission for their support. We also wish to acknowledge our gratitude and appreciation to all the BIVEE project partners for their contribution during the development of various ideas and concepts presented in this paper.

REFERENCES

[1] E. Brinksmeier, "Prozeß- und Werkstückqualität in der Feinbearbeitung," Fortschritt-Berichte VDI Reihe2: Fertigungstechnik Nr. 234. Düsseldorf: VDI Verlag. 1991.

[2] B. Knoke, T. Wuest, and K.-D. Thoben, "Characteristics of knowledge and barriers towards innovation and improvement in collaborative manufacturing process chains," accepted for

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

publication at the 25th International Conference on Advanced Information Systems Engineering, 2013.

[3] K. Kyriakopoulos and K. de Ruyter, "Knowledge stocks and information flows in new product development," Journal of Management Studies, vol. 41:8, 2004, pp. 1469–1498.

[4] T. Wuest, D. Klein, and K.-D. Thoben, "State of steel products in industrial production processes," Procedia Engineering, vol. 10, 2011, pp. 2220-2225.

[5] B. Knoke, T. Wuest, and K.-D. Thoben, "Fragmented knowledge in collaborative manufacturing process chains," accepted for publication at the 2013 International Conference on Collaboration Technologies and Systems (CTS 2013), San Diego, California, 2013.

[6] S. Kalpakjian and S. R. Schmid, "Manufacturing engineering and technology," New Jersey: Prentice Hall, 2009.

[7] M. Seifert, "Collaboration formation in virtual organisations by applying prospective performance measurement," Verlag Mainz. Aachen, 2009.

[8] T. Wuest, C. Irgens, and K.-D. Thoben, "An approach to quality monitoring in manufacturing using supervised machine learning on product state based data," Journal of Intelligent Manufacturing, March 2013, pp. 1-14.

[9] Y. Merali and Z. Bennett, "Web 2.0 and network intelligence, context and semantics for Knowledge Management," Technologies for Personal Productivity, vol. 11, 2011, pp.11-26.

[10] D. McFarlane, S. Sarma, J. L. Chirn, C. Y. Wong, and K. Ashton, "Auto ID systems and in-telligent manufacturing control," Eng Appl of Artif Intell 16, 2003, pp. 365-376.

[11] W. Maass and U. Varshney, "Preface to the focus theme section: Smart Products," Electronic Markets, vol. 18, 2008, pp. 211–215.

[12] M. Kärkkäinen, J. Holmström, K. Främling, and K. Artto, "Intelligent products - a step towards a more effective project delivery chain," Computers in Industry, vol. 50, 2003, pp. 141-151.

[13] G. G. Meyer, K. Främling, and J. Holmström, "Intelligent Products: a survey," Computers in Industry, vol. 60, 2009, pp. 137-148.

[14] O. Ventä, "Intelligent Products and systems. Technology theme - final report," VTT, Espoo: VTT Publications, 304, 2007.

[15] C. Wong, D. McFarlane, A. Zaharudin, and V. Agarwal, "The Intelligent Product driven Supply Chain," in: IEEE Int. Conference on Systems, Man and Cybernetics, vol. 4, 2002.

[16] D. Malhotra and F. Lumineau, "Trust and collaboration in the aftermath of conflict: the effects of contract structure," Academy of Management, vol. 54, 2011, pp. 981-998.

[17] T. Wuest, K. Hribernik, and K.-D. Thoben, "Digital representation of Intelligent Products: Product Avatar 2.0," In: Abramovici, M. et al. (Eds.) (2013). Smart Product Engineering. Proceedings of the 23rd CIRP Design Conference, March 11th - March 13th, 2013, Bochum, Germany, 2013.

[18] K. Hribernik, L. Rabe, K.-D. Thoben, and J. Schumacher, "The product avatar as a product-instance-centric information management concept," International Journal of Product Lifecycle Management, vol. 1, 2006, pp. 367-379.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Structured Analysis of Interactions in Collaborative Environments

Elena Troubitsyna
Åbo Akademi University, Dept. of IT
Joukhaisenkatu 3-5A, 20520,
Turku, Finland
e-mail: Elena.Troubitsyna@abo.fi

Abstract— **Collaborative computing environments are dynamic compositions of communicating components that interact with each other to achieve a common goals. The collaborations continuously reconfigure to achieve the required goals. Ensuring correctness of complex component interactions is cumbersome and requires structuring techniques that allow us to model and analyse component interactions in a systematic way. In this paper, we propose a set of modelling abstractions that allow us to define component interactions in dynamic collaborative environments. We propose a structured approach to analysing possible deviations in the component interactions based on HAZOP – Hazard and Operability Study -- and formally define the impact of deviations in component interactions on achieving the required goals.**

Keywords-*dynamic collaborations; interactions; goals; deviation analysis; formal modelling.*

I. INTRODUCTION

Over the recent years, collaboration has became a one of the primarily engines to create new services, achieve higher productivity or enable creating novel applications. Increasing openness of software and advances in networking has led to a proliferation of collaborative computing environments in different domains. Among the most remarkable examples of collaborative environments is the Internet of Things [10]. The term is introduced to stress the growing outreach of connectivity towards sensors, machines and variety of appliances. The wide-spreading use of the collaborative approach amplifies the need for novel communication paradigm that enables dynamic flexible collaboration creation and function.

The inherently dynamic mode of collaborations requires novel approaches that allow the designers systematically analyse the dynamics of collaborative environments and in particular, predict how deviations in the component behaviour and interactions impact objectives that a collaborative environment should achieve.

It has been recognized that it is convenient to formalize objectives that a system should achieve by a notion of *goals* [4]. The collaborations are formed to achieve certain goals. The components forming collaboration provide certain individual functionality that contributes to achieving overall goal. When a component fails or components communicate inappropriately, collaboration might fail to achieve the required goal. Therefore, we should analyse the possible deviations in the component behaviour and formally define the impact of these deviations on achieving overall goals.

To systematically study possible deviations in the component interactions we propose to use Hazard and Operability Studies -- the HAZOP method [1,2]. We define the main types of deviations in the components interactions and define their impact on achieving system goals.

We believe that the main contribution of this paper, i.e., a formal link between goals and possible deviations in component interactions, can potentially facilitate design of complex collaborative environments

The paper is structured as follows. In Section II, we define collaborative environments in terms of the goals that should be achieved. In Section III, we describe generic scenarios of component interactions. Section IV shows how to systematically analyse deviations in the component interactions using HAZOP. Finally, in Section V, we discuss the proposed approach and overview the related work.

II. GOALS IN COLLABORATIVE ENVIRONMENTS

In this paper, we define a *collaborative environment* as a set of collaborations, i.e.,

$$ColENV = \{C_1, C_2, ..., C_N\}$$

where *Ci* is an id of collaboration.

Collaboration is a dynamic composition of components. The components join and leave collaboration depending on the current goal and the states of the components. In general, any collaboration is formed to fulfil a certain goal [4]. The set of goals, which an entire collaborative environment can achieve, is denoted as *GOALS:*

$$GOALS = \{G_1, G_2, ..., G_M\}$$

The set consists of the constants defining the names of the goals. We assume that each particular collaboration is formed to perform a certain goal from the set *GOALS*.

A *goal* is an objective that collaboration should achieve. A goal can be decomposed into a set of subgoals, and furthermore, into a set of sub-subgoals that each component in the collaboration should perform. Each component carries a special attribute describing the functionality that it implements. Often these attributes are called roles. Usually a component implements a set of roles chosen according to the tasks that it should perform in each particular collaboration.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

The goal that collaboration should implement defines how many components executing each role collaboration should have to achieve a goal. Therefore, a configuration can be defined as follows:

Config $\in$ *CONFIG*, where *CONFIG : ROLES -> N*

where *ROLES* is a set of roles and *N* is a set of natural numbers.

Often a configuration of collaboration is defined not only by a goal but also non-functional parameters, e.g., performance. We assume that goals are distinct if their non-functional parameters are different. Therefore, we can unambiguously map a set of goals on the set of configurations.

For each goal *Gi, Gi* $\in$ *GOALS,* we can define the minimal sufficient configuration as a function

MINCONF : GOALS -> CONFIG

The function defines how many components in each role collaboration should have to be able to achieve a certain goal. The function *MINCONF* defines the minimal necessary conditions. Obviously, collaboration can have more components that might be inactive while achieving a certain goal or used as standby to implement fault tolerance in case some components fail.

In practice, at each particular moment of time, a collaborative environment *ColENV* does not try to achieve all the goals defined by the set *GOALS* at once. Therefore, we can distinguish between a set of the active (trigged) goals, i.e., the goals that a collaborative environment tries to achieve at a certain moment of time and the goals that are not trigged. This defines a partitioning of the set of goals into two non-intersecting subsets:

GOALS = ACT_G $\cup$ *PAS_G,*
where *ACT_G* $\cap$ *PAS_G = Ø*

In our modelling, we assume that the components are not kept idle but rather are getting engaged in different collaborations (as soon as their roles match the roles required in the collaboration). Therefore, when a goal is trigged, it might be the case that the conditions defined by *MINCONF* are not satisfied because the required components are still engaged in some other collaboration. If the required configuration is established then the collaboration executes the required actions to achieve the goal. We introduce a set

C_STATE : {Active, Activated, Dormant}

to designate the status of the collaboration and introduce the function C_STATUS that maps the id of the collaboration to its status:

C_STATUS : CNAME -> C_STATE

The function *CUR_CONFIG* is defined as follows:

CUR_CONFIG : CNAME-> CONFIG

It designates the current configuration of the collaboration.

Next, we formally define the relationships between the status of the collaboration, goals and configurations.

The collaboration *Ci* is active, i.e.,

C_STATUS (Ci) = Active

if

Gj $\in$*GOALS* $\wedge$
Gj $\in$*Act_G* $\wedge$
MINCONF(Gi)$\leq$*CUR_CONFIG(Ci)*

where the ordering relation $\leq$ is defined over the configurations as follows:

For $Conf_k$ and $Conf_l$, such that $Conf_k$, $Conf_l$ $\in$*CONFIG,* $Conf_k \leq Conf_k$ if

$$\forall r_n.r_n \in dom(\mathrm{Conf_k}) \Rightarrow r_n \in dom(Conf_l)$$
$$\forall r_n.r_n \in dom(\mathrm{Conf_k}) \Rightarrow Conf_k(r_n) \leq Conf_l(r_n)$$

When a collaboration *Ci* is set to achieve a certain goal but has not established the required configuration or an execution of a scenario required to achieve a goal is suspended due to failures, its status is *Activated, i.e.,*

C_STATUS (Ci) = Activated

if

Gj $\in$*GOALS* $\wedge$
Gj $\in$*Act_G* $\wedge$
$\neg$ *(MINCONF(Gi)*$\leq$*CUR_CONFIG(Ci))*

Finally, collaboration can be inactive, i.e.,

C_STATUS (Ci) = Activated

if

Gj $\in$*GOALS* $\wedge$
Gj $\in$*Pas_G*

We assume that components are involved in the collaboration with the status *ACTIVE* communicate with each other by exchanging messages. To achieve a certain goal, collaboration should perform a predefined scenario. In the next section, we define generic scenarios performed by the components in collaboration.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

III. MODELLING COMPONENT INTERACTION

We can describe a scenario by a UML [5] use case model and define the details of the communication by the sequence diagram. A high-level generic scenario is defined as follows:

Description of use case
Collaboration Ci achieves goal Gj

Precondition *Goal is eligible for execution and trigged*
$Gj \in GOALS \wedge$
$Gj \in Act\_G \wedge$

Postcondition *Collaboration achieves goal or Collaboration reports failure*

Includes: *Recover_Scenario_Ci_Gj*

Normal sequence of events:

1. The coordinator of *Ci* receives a notification that a goal is activated and changes the status of the collaboration, i.e.,

   C_STATUS (Ci) := Activated

2. The coordinator broadcasts an invitation to join a collaboration to the components of *ColENV* and monitors that the required configuration is established
3. When a configuration is established, i.e.,

   $MINCONF(Gi) \leq CUR\_CONFIG(Ci)$

   it broadcasts the message engaged to the involved components and changes the status of the collaboration, i.e.,

   C_STATUS (Ci) := Active

4. Components communicate to each other to perform the tasks required to achieve goal and the coordinator monitors the status of the components. If it discovers a component failure then go to step 8.
5. When goal is achieved the components report to the coordinator about completion of scenario.
6. Coordinator hands over the control to the collaborative environment manager and changes the status of the collaboration, i.e.,

   C_STATUS (Ci) := Dormant

7. The coordinator broadcasts disengage message to all components.
8. The collaboration coordinator re-evaluates the status of the collaboration. If the condition of the sufficient configuration is not satisfied then it changes the status of the collaboration to *Activated* and activates timer.
9. If the components recover within the timeout then the status is changed to *Active* and the normal execution is resumed.

If the components fail to recover within timeout then switch to executing failure recovery scenario *Recover_Scenario_Ci_Gj.*

Description of use case *Recover_Scenario_Ci_Gj*

Precondition
Normal execution of scenario to achieve goal Gj by collaboration Ci failed.
Status of Ci is Activated

Postcondition Reconfiguration and resuming normal execution or permanent failure

Extends: *Collaboration Ci achieves goal Gj*

Sequence of events:

1. The coordination of *Ci* broadcasts a new invitation to join a collaboration and activates a timer
2. If within the timeout the coordinator receives a respond from components whose roles match the roles of failed components then continue. Otherwise the scenario terminates, i.e., go to 4.
3. The coordinator sends engagement message to the newly joining components and changes the status of the collaboration to Active. Normal execution resumes, i.e., the use case *Collaboration Ci achieves goal Gj* resumes.
4. The collaboration sends the failure message to the collaborative environment manager and changes the status of the collaboration *Ci* to *Dormant.*
5. The coordinator broadcasts disengage message to all components.

Let us now depict the described scenario as a sequence diagram. An excerpt from the sequence diagram is shown in Fig. 1. We use the sequence diagram as an input for conducting analysis of deviations in the component interactions. Next, we present our analysis method, HAZOP, adapted for analysis of dynamic behaviour.

IV. GOALS IN COLLABORATIVE ENVIRONMENTS

HAZOP was originally developed in chemical industry [1,2]. Essentially, HAZOP provides a structured basis for brainstorming by a group of experts about possible deviations in the behaviour of the system. As a result of conducting HAZOP, experts indentify hazards and propose means to mitigate them.

HAZOP is conducted by applying the list of guidewords to certain system parameters. The list of the guidewords is presented in Table I.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

TABLE I. LIST OF GENERIC HAZOP GUIDE WORDS

Guideword	Interpretation
No/None	Complete negation of the design intention. No part of the intention is achieved and nothing else happens
More	Quantitative increase
Less	Qualitative increase
As Well As	All the design intentions is achieved together with additions
Part of	Only some of the design intention is achieved
Reverse	The logical opposite to design intention is achieved but something quite different happens
Early	Something happens earlier than expected relative to clock time
Late	Something happens later than expected relative to clock time
Before	Something happens before it is expected, relating to order of sequence
After	Something happens after it is expected, relating to order or sequence

In Table I, we present the generic guideword list from the Defence Standard 00-58 [1] and IEC-61882 [2]. The HAZOP methods has been adapted to various domains and received several interpretations that allows the designers to focus on a wide spectrum of aspects – from human errors to software.

For models of dynamic system behaviour, e.g., such as sequence diagrams, many guidewords interpretations can be used for exploring deviations during the component interactions. In this paper, we adopt the reinterpretation of the guidewords for HAZOP proposed in [3] for the UML sequence diagrams. The adopted interpretation of the HAZOP guidewords [3] to sequence diagram is given in Table II.

Let us now demonstrate an application of the guidewords to the basic scenario of component interactions. We consider only a few examples that have resulted in identifying deviations that hinder achieving the required goal.

Messages outgoing from the coordinator:

Invite message:
No: Execution of scenario is not trigged
Before: Message sent when the goal is not trigged
Earlier: Message sent before the goal is trigged
Later: Message sent with the delay

Messages from the components:
Confirm participation
No: Message might block execution of the goal if no other component confirm
After: Message delays execution of scenario

Inter-component Communication Message:
No: No message is sent after completing execution: Deadlocks goal execution
More than: several messages sent after completing execution: scenario is executed in wrong order
Before /Early: message is sent before task completes and trigs earlier than required execution of tasks in another components
Later: execution of the goal is delayed.

TABLE II. INTERPRETATION OF HAZOP GUIDE WORDS

Attribute	**Guideword**	**Interpretation**
Predecessor/ successors during interactions	No	Message is not sent
	Other than	Unexpected message sent
	As well as	Message is sent as well as another message
	More than	Message sent more often than intended
	Less than	Message sent is often as intended
	Before	Message sent before intended
	After	Message sent after intended
	Part of	Only a part of a set of messages is sent
	Reverse	Reverse order of expected messages
Message timing	As well as	Message sent at correct time and also incorrect time
	Early	Message sent earlier than intended time
	Later	Message sent later than intended time
Sender/ receiver objects	No	Message sent but never received by intended object
	Other than	Message sent to wrong object
	As well as	Message sent to correct object and also an incorrect object
	Reverse	Source and destination objects are reversed
	More	Message sent to more objects than intended
	Less	Message sent to fewer objects than intended
	No/none	The conditions is

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

Message guard conditions		not evaluated and can have any value (omission)
	Other than	The condition is evaluated true whereas it is false, or vice versa (commission)
	As well as	The condition is well evaluated but other unexpected conditions are true
	Part of	Only a part of conditions is correctly evaluated
Message guard conditions (cont.)	Late	The conditions is evaluated later than required (other dependant conditions have been tested before)
		The conditions is evaluated later than correct synchronisation with environment
Message parameters/ return parameters	No/None	Expected parameters are never set/returned
	More	Parameters values are higher than intended
	Less	Parameter values are lower than intended
	As Well As	Parameters are also transmitted with unexpected ones
	Part of	Only some parameters are transmitted Some parameters are missing
	Other than	Parameter type/number are different from those expected by receiver

Our analysis allows us to derive recommendation how to mitigate the impact of deviations. For instance, it clearly demonstrates that a message omission leads to the system deadlock. Therefore, a time out mechanism should be implemented to ensure that the goal execution progresses despite possible message omissions.

If a component sends a confirmation of a task completion then the consequent task might start in an incorrect state. To mitigate this hazard, a coordinator might additionally send a check to ensure that the required task was indeed completed.

V. CONCLUSION AND RELATED WORK

Our analysis is based on formal definition of relations between the goals that collaboration should achieve and states of the components. A formalization of a goal-oriented development was proposed in [6]. In this paper, the focus was not only on formal representation of relationships between the agents and goals but also on the systematic analysis of deviations. An approach to integration with other techniques for safety analysis was proposed in [8]. This work is relevant to a high-level analysis of collaboration. An approach to analysis of collaborative behaviour in the context of mode-rich systems was proposed in [9]. The focus of this work was on reasoning about modes of collaborating components.

A formalization of agent collaboration has been performed in [7]. The focus of this work was on tolerating temporal agent failures, while in our work we focused on systematic analysis of deviations in component interactions.

HAZOP analysis has been adapted to analyse human computer-interactions as well as process deviations. Our use of HAZOP is similar to the former and allows us to reason about interactions of components participating in collaboration.

In this paper, we proposed a systematic approach to analyse component interactions in collaborative environments. We formally defined relationships between the state of components and ability of collaboration to achieve the required goals. We have demonstrated that the HAZOP method allows us systematically study deviations in the component interactions and establish a link between errors in interactions and goal achieving.

As a future work, it would be interesting to apply the proposed approach to complex collaborative environment from the Internet of Things domain.

REFERENCES

[1] DefStan00-58. HAZOP studies on systems containing programmable electronics. Defence Standard, Ministry of Defence, UK, 2007.

[2] IEC61882. Hazard and operability studies (HAZOP studies). Application guide. International Electrotechnical Comission, 2001.

[3] J. Guiochet, D. Martin-Guillerez, and D. Powell. “Experience with Model-Based User-Centered Risk Assessment for Service Robots”. HASE 2010: pp. 104-113.

[4] A. van Lamsweerde “Goal-Oriented Requirements Engineering: A Roundtrip from Research to Practice”. *Proceedings of RE’04*, 12th IEEE Joint International Requirements Engineering Conference, Kyoto, Sept. 2004, pp. 4-8.

[5] OMG-UML2.OMG unified modeling language (UML). Superstructure, v2.1.2. Onject Management Group.

[6] I. Pereverzeva, E. Troubitsyna, and L. Laibinis. “Formal Goal-Oriented Development of Resilient MAS in Event-B”. In Proc. of *Ada-Europe 2012 --17th International Conference on Reliable Software Technologies.* Lecture Notes in Computer Science 7308, pp. 147–161, Springer, June 2012.

Copyright (c) IARIA, 2013. ISBN: 978-1-61208-287-5

[7] L. Laibinis, E. Troubitsyna, A. Iliasov, and A. Romanovsky. “Rigorous Development of Fault-Tolerant Agent Systems”. In M. Butler, C. Jones, A. Romanovsky and E. Troubitsyna (Eds.), *Rigorous Development of Complex Fault-Tolerant Systems*. Lecture Notes in Computer Science, vol. 4157, pp. 241-260, Springer, Berlin, November 2006.

[8] K. Sere and E. Troubitsyna. “Safety Analysis in Formal Specification”. In J.M. Wing, J. Woodcock, J. Davies (Eds.) *Proc. of FM'99 - Formal Methods: World Congress on Formal Methods in the Development of Computing Systems,* Lecture Notes in Computer Science 1709, pp. 1564 – 1583, Springer, France, September 1999.

[9] A. Iliasov, E. Troubitsyna, L. Laibinis, A. Romanovsky, K. Varpaaniemi, D. Ilic, and T. Latvala. “Developing Mode-Rich Satellite Software by Refinement in Event-B”. *Science of Computer Programming,* 78(7), pp. 884-905, 2013.

[10] Internet of Thigs. www.**internet-of-things**.eu/. Accessed July, 2013.

Copyright (c) IARIA, 2013. **ISBN: 978-1-61208-287-5**