BUSTECH 2012

The Second International Conference on Business Intelligence and Technology

ISBN: 978-1-61208-223-3

July 22-27, 2012

Nice, France

BUSTECH 2012 Editors

Malgorzata Pankowska, University of Economics – Katowice, Poland

Petre Dini, Concordia University - Montreal, Canada / China Space Agency Center - Beijing, China

BUSTECH 2012

Foreword

The Second International Conference on Business Intelligence and Technology (BUSTECH 2012), held between July 22 and 27, 2012 in Nice, France, continues a series of events covering topics related to business process management and intelligence, integration and interoperability of different approaches, technology-oriented business solutions and specific features to be considered in business/technology development.

The term Business Intelligence (BI) covers a large spectrum of applications and technologies used to collect, store, interpret and decide on the information about company operations with the aim of helping corporate entities with a comprehensive status and knowledge on their business. BI is integrating with Warehouses (DWs), on-line analytic (OLAPS), corporate performance management (CPM), business process management (BPM), and other technology-oriented business solutions. Web technologies, semantics and ontology mechanisms are now used to mine, integrate and interpret distribute corporate data, either real-time or intermittent, by filtering noisy data, interpreting business data in context, enforcing trust and security in handling corporate data, and providing access to data from anywhere, at anytime, and via any media. The complexity, volume and intrinsic semantic of data needed for conducting business require a tailored IT infrastructure and advanced methodologies and technologies for timely building competitiveness by intelligent business decisions. With the large spectrum of emerging technologies, such as cloud computing, sensors environments, and mobility there is a need for specialized supporting tools and business/technology decisions to optimize the business process and business performance.

We take here the opportunity to warmly thank all the members of the BUSTECH 2012 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 BUSTECH 2012. 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 BUSTECH 2012 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that BUSTECH 2012 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the area of business intelligence and technology.

We are convinced that the participants found the event useful and communications very open. We hope Côte d'Azur provided a pleasant environment during the conference and everyone saved some time for exploring the Mediterranean Coast.

BUSTECH 2012 Chairs:

BUSTECH Advisory Chairs
Lars Taxén, Linköping University, Sweden
Dieter Fensel, STI Innsbruck / University of Innsbruck, Austria
Malgorzata Pankowska, University of Economics – Katowice, Poland
Hannes Werthner, Vienna University of Technology, Austria

BUSTECH 2012 Research/Industry Chair
Tyrone W. A. Grandison, IBM Research - Hawthorne, USA

BUSTECH 2012 Special Area Chairs

Semantic/Ontology
Oscar Ferrandez Escamez, University of Utah, USA

Business-driven IT
Maribel Yasmina Santos, University of Minho, Portugal

Security
Pedro Soria-Rodriguez, Atos Origin - Madrid, Spain

Business continuity
Sokratis K. Katsikas, University of Piraeus, Greece

Services
Michael Parkin, European Research Institute in Service Science (ERISS), The Netherlands

BUSTECH 2012

Committee

BUSTECH Advisory Chairs

Lars Taxén, Linköping University, Sweden
Dieter Fensel, STI Innsbruck / University of Innsbruck, Austria
Malgorzata Pankowska, University of Economics – Katowice, Poland
Hannes Werthner, Vienna University of Technology, Austria

BUSTECH 2012 Research/Industry Chair

Tyrone W. A. Grandison, IBM Research - Hawthorne, USA

BUSTECH 2012 Special Area Chairs

Semantic/Ontology
Oscar Ferrandez Escamez, University of Utah, USA

Business-driven IT
Maribel Yasmina Santos, University of Minho, Portugal

Security
Pedro Soria-Rodriguez, Atos Origin - Madrid, Spain

Business continuity
Sokratis K. Katsikas, University of Piraeus, Greece

Services
Michael Parkin, European Research Institute in Service Science (ERISS), The Netherlands

BUSTECH 2012 Technical Progam Committee

Hassan Aït-Kaci, IBM, Canada
Antonia Albani, University of St. Gallen, Switzerland
Dumitru Dan Burdescu, University of Craiova, Romania
Joseph Barjis, Delft University of Technology, The Netherlands
Seyed M.R. Beheshti, University of New South Wales - Sydney, Australia
Kawtar Benghazi, Universidad de Granada, Spain
Peter Bollen, Maastricht University, The Netherlands
Jan Bosch, Chalmers University of Technology- Gothenburg, Sweden
Marko Bošković, Simon Fraser University Surrey, UK / Athabasca University, Canada
Mahmoud Boufaida, Mentouri University - Constantine, Algeria
Albertas Caplinskas, Vilnius University, Lithuania

Anis Charfi, SAP AG - Darmstadt, Germany
David Chen, University Bordeaux 1 - Talence, France
Claudia d'Amato, University of Bari, Italy
Zhi-Hong Deng (Denny), Peking University, China
Giuseppe A. Di Lucca, University of Sannio - Benevento, Italy
Johannes Edler, University of Applied Sciences Upper Austria - Hagenberg, Austria
Marcelo Fantinato, Universidade de São Paulo, Brazil
Dieter Fensel, STI Innsbruck/University of Innsbruck, Austria
M. Carmen Fernandez Gago University of Malaga, Spain
Oscar Ferrandez Escamez, University of Utah, USA
George Feuerlicht, University of Technology - Sydney, Australia
Agata Filipowska, Poznan University of Economics, Poland
Tyrone W A Grandison, IBM Research, Yorktown Heights, USA
Andrina Granić, University of Split, Croatia
Stewart Green, University of the West of England - Bristol, UK
Sven Graupner, Hewlett-Packard Laboratories - Palo Alto, USA
Pierre Hadaya, UQAM, Canada
Ioannis Hatzilygeroudis, University of Patras, Greece
Lucinéia Heloisa Thom, Federal University of Rio Grande do Sul, Brazil
Jingwei Huang, University of Illinois at Urbana-Champaign, USA
Edward Hung, Hong Kong Polytechnic University - Kowloon, Hong Kong
Stefan Jablonski, University of Bayreuth, Germany
Hoyoung Jeung, EPFL, Switzerland
Sokratis K. Katsikas, University of Piraeus, Greece
Marite Kirikova, Riga Technical University, Latvia
Agnes Koschmider, Karlsruhe Institute of Technology (KIT) , Germany
Robert Lagerström, Royal Institute of Technology - Stockholm, Sweden
Daniel Lemire, LICEF/ University of Quebec at Montreal, Canada
Olga Levina, Berlin Institute of Technology, Germany
Mario Lezoche, University of Lorraine, France
Niels Lohmann, University of Rostock, Germany
Yannis Manolopoulos, Aristotle University of Thessaloniki, Greece
Martin Molhanec, Czech Technical University in Prague, Czech Republic
Bela Mutschler, University of Applied Sciences-Weingarten, Germany
Lina Nemuraite, Kaunas University of Technology, Lithuania
Gustaf Neumann, WU Vienna, Austria
Andrzej Niesler, Wroclaw University of Economics, Poland
Manuel Noguera García, Universidad de Granada, Spain
Alexander Norta, University of Helsinki, Finland
Enn Ounapuu, Tallinn University of Technology, Estonia
Malgorzata Pankowska, University of Economics-Katowice, Poland
Eric Paquet, National Research Council - Ottawa, Canada
Andreas Pashalidis, Katholieke Universiteit Leuven, Belgium
Wolter Pieters, Delft University of Technology, The Netherlands
Alain Pirotte, University of Louvain, Louvain-la-Neuve, Belgium
Elke Pulvermueller, University of Osnabrueck, Germany
Manjeet Rege, Rochester Institute of Technology - NY, USA
Stefanie Rinderle-Ma, University of Vienna, Austria

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

Discovering Cross-Perspective Semantic Definitions from Process Execution Logs

Stefan Schönig, Christoph Günther, Michael Zeising, and Stefan Jablonski
University of Bayreuth
Chair of Applied Computer Science IV
Bayreuth, Germany
{stefan.schoenig, christoph.guenther, michael.zeising, stefan.jablonski}@uni-bayreuth.de

Abstract - **In this paper, we suggest a two-phase declarative process mining approach discovering explicit, cross-perspective semantic definitions. Cross-perspective semantic definitions are interesting and important for the analysis of business processes, because they reveal dependencies that are not obvious on the first look. They allow for a comprehensive examination of the recorded process execution information and enable the discovery of coherency between different process-involved entities and perspectives. Using the described cross-perspective semantic definitions, we additionally introduce an approach for simplifying less-structured process models.**

Keywords - Process mining; semantic definitions; business rules; ontology; semantic process modelling.

I. INTRODUCTION

Process modelling is an expensive and cumbersome task. Using process mining techniques, it is possible to discover process models automatically [10]. Moreover, event logs can be checked to assess conformance and compliance with respect to already defined processes [10]. Process mining has been applied in various domains ranging from healthcare and e-business to high-tech systems and auditing [5, 6]. However, many process mining techniques produce "spaghetti-like" models that tend to be large and complex, especially in flexible environments where process executions involve multiple alternatives [1, 2]. This "overload" of information is caused by the fact that traditional mining techniques construct imperative models explicitly encoding all possible behaviours [1, 10]. This sort of complexity arises when a huge number of execution paths exists (path complexity) [23]. When process models are becoming too complex, people cannot interpret them anymore and therefore cannot improve them. In order to face this problem, we leave the imperative world and focus on the generation of declarative process models. Declarative process modelling techniques reduce path complexity such that complex applications can be described by comprehendible process models.

In contrast to imperative modelling, declarative models concentrate on describing what has to be done and the exact step-by-step execution order is not directly prescribed. There are several process mining approaches like [1, 19, 20] that are discovering declarative process models. Here, the meaning of model elements is defined by explicit semantic definitions. Furthermore, several approaches [2, 3, 4, 12] filter the information contained in a log and to simplify less-structured imperative process models by discovering common execution patterns. However, the approaches named above have a common drawback: they are mainly examining the behavioural perspective, i.e., the control-flow. These methods are discovering semantic definitions considering the execution order of process steps without facing possible coherency with other perspectives. We think that especially this hidden coherency between perspectives should be outlined by discovery algorithms. That is why we suggest a mining approach, based upon user-defined cross-perspective semantic definitions. That means these semantic definitions spread over different entities and perspectives, e.g., a process execution order (behavioural perspective) could depend on the performing agents' position (organisational perspective). The semantic definitions are constituted through the analysis of the different perspectives recommended by the *perspective-oriented process modelling* (POPM) approach [15]. The user-defined assembly of semantic definitions allows analysts to shape the discovery process to extract the semantic definitions that are most important and interesting for them [1].

Cross-perspective semantics is especially interesting and important for the analysis of business processes, because it reveals dependencies that are not obvious on the first look. It allows for a comprehensive examination of the recorded process execution information and enables the discovery of coherency between different process-involved entities and perspectives. Beyond, our approach discovers semantic definitions that can be based upon properties of an ontology, containing further information about process-involved entities or participants. That means that the semantic definitions to be searched can partly consist of properties that have to be extracted from an underlying ontology. This functionality is similar to [13, 14], however we apply this possibility in the context of process discovery instead of conformance checking. Using the described cross-perspective semantic definitions, we additionally introduce an approach for simplifying less-structured process models.

This paper is organized as follows: Section II introduces the fundamental assembly and the two phases of the approach. In Sections III and IV, these two phases are described in detail. In Section V, related work is discussed. The paper is finally concluded in Section VI.

II. DISCOVERING CROSS-PERSPECTIVE SEMANTIC DEFINITIONS FROM PROCESS EXECUTION LOGS

Information systems typically log various kinds of information about process execution. The starting point for

 ISBN: 978-1-61208-223-3

process mining is an event log. An event log consists of a set of traces whereat each trace is a sequence of events corresponding to a particular case, i.e., one process instance [7]. Each events record refers to a single process step and typically has a timestamp. These facts also form the preliminaries for our approach. We assume an existing event log recording different perspectives of process execution. Table I shows a fragment of such a process execution log. While the PID-column assigns each event to unique process identifier, the case-column assigns each event to a single process case, i.e., a single process instance. Furthermore, the action type of each event is recorded. The following columns are an example for further information that should be logged during process execution in order to be able to discover cross-perspective business rules. We recommend to record data based upon the different aspects of the *perspective-oriented process modelling* (POPM) [15]:

Functional perspective: the functional perspective identifies a process step and defines its purpose. Also the composition of a process is determined by this perspective. Hence, the log should contain a common process identifier the corresponding event can be linked to.

Data perspective: the data (flow) perspective defines data used in a process and the flow of data between process steps. Therefore, the log should record documents or generally information that was used by the current process step as well as the data that was produced.

Operational perspective: the operational perspective specifies which operation (service) is invoked in order to execute a process step. It relates processes to services stemming from (external) service libraries. Here, the log should contain tools, applications or services that were used during performing the currently executed process step.

Organisational perspective: the organisational perspective defines agents (for instance users, roles) who are eligible and/or responsible to perform a process step. Therefore, the log contains information about the process executor. The personal information is enriched by group and role memberships.

Behavioural perspective: the behavioural perspective is used to define causal dependencies between process steps (e.g. step B may only be executed after step A). Often these dependencies are called control flow. The information in the log concerning this perspective is formed by the recorded timestamp of each event.

TABLE I. A FRAGMENT OF A PROCESS EXECUTION LOG.

Event	PID	Case	Action	Agents	Data	Tools	Time
1	A	**1**	Start	Head	Doc 1	Word	…
2	A	**1**	Finish	Head	Doc 1	Word	
3	D	*2*	Start	Agent 3	Doc 3	Excel	
4	D	*2*	Finish	Agent 3	Doc 3	Excel	
5	B	**1**	Start	Agent 2	Doc 2	Word	
6	C	*2*	Start	Agent 3	Doc 2	Excel	
7	B	**1**	Finish	Agent 2	Doc 2	Word	
8	C	2	Finish	Agent 3	Doc 2	Excel	
9	C	**1**	Start	Trainee	Doc 3	Word	
10	A	2	Start	Head	Doc 1	Word	
11	C	**1**	Finish	Trainee	Doc 3	Word	
12	D	**1**	Start	Trainee	Doc 4	Word	
13	D	**1**	Finish	Trainee	Doc 4	Word	
…							

The existence of an event log of such a shape allows for the comprehensive examination of various perspectives within one approach. Therefore, we propose a two phase approach to analyse a process execution log.

Phase-1 (Pre-processing the log to instance graphs; Section IV). Here, the event log is generally analysed and transformed into various graph data structures that allow for the flexible search of user-defined semantic definitions.

Phase-2 (Discovery of cross-perspective semantic definitions; Section V). This phase discovers cross-perspective semantic definitions concerning only one process as well as relations between two processes. Semantic definitions are used to encapsulate processes.

III. PRE-PROCESSING THE LOG TO INSTANCE GRAPHS

In this section, we focus on the construction of so-called *instance graphs*. An instance graph describes the execution order of process steps of a process (case), i.e. one single execution path of a process. For our particular cross-perspective purpose, we feature the graphs of [24, 25] with context data of the organisational, data and the operational perspective. Instance graphs also show parallelism if parallel (independent) branches have been executed. An instance graph consists of a set of nodes N and a set of edges E. Every node $n \in N$ has the following fields: process name, performing agent, used document and used tool support. Every edge $e \in E$ has two fields describing how two processes are connected: execution type (parallel or sequence) and distance (direct or transitive). Every instance graph is a complete graph. First, we separate the recorded events according to their corresponding case/instance id.

Therefore, we assemble a list for each case represented in the log and assign the events according to their case ids. With the help of these lists, we can now classify the relation between two (sub-)processes within one process case. The classification is based upon the event types of two succeeding events. Here, we make the same assumptions as [24, 25]. As already mentioned, we distinguish between parallel execution and direct sequential execution. Consider two processes *A* and *B*. We deduce that two processes are executed in parallel if process *A* is started before process *B* is started and completed before *B* is completed but after the start of *B*. This would result in the event sequence: *Start A*, *Start B*, *Finish A*, *Finish B* (Fig.1 ①). Furthermore, the two processes are also executed in parallel if process *A* is started before process *B* is started and completed after process *B* is completed. The resulting event sequence would look like this: *Start A*, *Start B*, *Finish B*, *Finish A* (Fig.1 ②). In addition to parallel execution, we mark direct sequential execution. Two processes *A* and *B* are executed in a direct sequence if process *B* is started directly after process *A* has been completed (we say "*B* is started after *A* finished"). The

resulting event sequence therefore is: *Start A*, *Finish A*, *Start B*, *Finish B* (Fig. 1 ③).

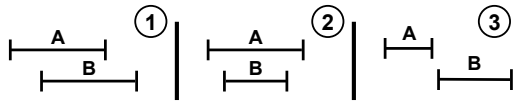


Figure 1. Classification of behavioural relations of processes.

On the basis of the classification above, an instance graph of an event list is created as follows. We generate a graph by running through the case-specific event list. For every newly occurring process *A* within the list, we create a new node *A* within a graph and assign the corresponding process context to the fields of the node, i.e., recorded agent, documents as well as tools. For every direct sequential-relation of two processes *A* and *B*, we add an edge of execution type "sequence" and distance "direct" between the node representations of *A* and *B* within the graph (formal: $A \rightarrow B$). For every parallel execution between two processes *A* and *B*, we add an edge of execution type "parallel" and distance "direct" between the node representations of *A* and *B* within the graph ($A \leftrightarrow B$). Finally, the existing graph is extended by edges that have been generated by the transitive closure of the graph. If the graph already contains two edges $A \rightarrow B$ and $B \rightarrow C$, we add an edge of execution type "sequence" and distance "transitive" between the node representations of the processes A and C ($A \gg C$). Note that in general, it is not possible to infer $A \leftrightarrow C$ from $A \leftrightarrow B$ and $B \leftrightarrow C$. Fig. 2 shows three different instance graphs of a process based upon the log fragment of Table 1 (for space reasons, the table just shows the activities of two instances). Considering graph 1, case 1 had the execution trace *A*, *B*, *C*, *D*, containing only direct sequential-relations and no parallelism. Exemplarily, graph 1 additionally contains the information that the agent "Head of Department" executed process *A* by using *Document 1* supported by *MS Word*. The two other graphs can be interpreted in the same way.

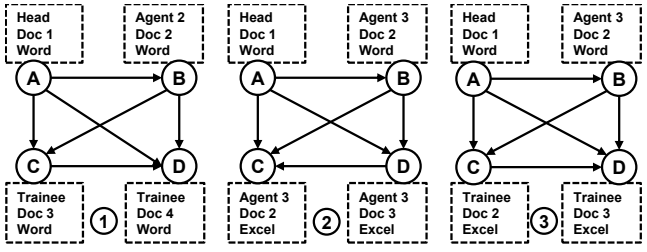


Figure 2. Instance graphs based on three different process cases.

IV. Discovery of Cross-Perspective Semantics

A. Classification of Semantics

Meta-modelling frameworks like the *Open-Meta-Modelling Environment* [16] offer the possibility to feature and adapt process modelling languages with a variety of user-defined (domain-specific) modelling elements. This requires a clear specification of the meaning of modelling constructs in order to avoid misunderstanding between modeller and programmer. This can be achieved by explicit semantic definitions of model elements. Semantic definitions are already used to validate executed processes in the context of conformance checking [10]. Hence, logs can be validated with the help of logical reasoners. In this paper, we scan event logs in order to discover semantic definitions. First, we introduce the assembly and representation of semantic definitions. Here, we use the *Semantic Web Rule Language* (SWRL) [21] to define semantics of modelling elements. Every semantic definition, e.g., SWRL rule, has a left and a right side. The left side contains the conditions that have to be satisfied so as to infer the consequences on the right side. Furthermore, the left side contains an indicator for the corresponding modelling construct whose semantics is defined. Conditions and consequences consist of atoms and assigned variables. The formalisation of such rules looks as follows:

$$Modeling\_Element\_Indicator \wedge Conditions \Rightarrow Consequences$$

The running example for this section consists of a domain-specific modelling element called *TraineeConnector*. This semantics expresses that any two processes *P1* and *P2* connected by this semantics have a strict execution order in case that a trainee employee performs the processes. Here, the cross-perspective nature becomes obvious: the behavioural perspective depends on the organisational perspective. The representation of this semantics as an SWRL definition looks as follows:

$$\begin{aligned} &\text{TraineeConnector}(?C) \wedge \text{from}(?C, ?P_1) \wedge \text{to}(?C, ?P_2) \wedge \\ &\text{hasAgent}(?P_1, ?A) \wedge \text{hasAgent}(?P_2, ?A) \wedge \text{Trainee}(?A) \\ &\Rightarrow \text{startedAfterFinished}(?P_2, ?P_1) \end{aligned}$$

Note, that the algorithm also is able to discover established (single perspective) declarative semantic definitions. Exemplarily, the semantic definition of a "standard" sequential execution order in any case (a name for such semantics could be *StrictOrderConnector*) would look like this:

$$\begin{aligned} &\text{StrictOrderConnector}(?C) \wedge \text{from}(?C, ?P_1) \wedge \text{to}(?C, ?P_2) \\ &\Rightarrow \text{startedAfterFinished}(?P_2, ?P_1) \end{aligned}$$

The assembly of the core algorithm demands for the definition of a few terms describing the type of semantic definitions.

Attribute: Attributes describe modelling elements that represent semantics concerning a single process. An attribute does not express a relation to another process. An example would be the fact that a specific process has always to be executed by the head of department.

Connector: Connectors describe modelling elements that represent relations, e.g., connectors, between two processes. An example for a connector is the *TraineeConnector* from above.

Group: Group semantic defintions typically assign functional entities (e.g., atomic processes) to other functional entities (e.g., complex processes or pools).

 ISBN: 978-1-61208-223-3

There may also exist other semantic definitions, however, in this paper we focus on discovery of semantic definitions defined above with respect to comprehension reasons.

B. Discovery of Semantic Definitions

The validation of given semantic definitions represents the core functionality of the approach. The general approach is as follows:

1. Search for explicit semantic definitions (attributes as well as connector semantics) using the instance graphs.
2. Generation of the semantic graph, that contains attributes of single processes as well as connectors between processes.
3. Encapsulation of processes and generation of process hierarchies examining group semantic definitions.

The main principle of the validation procedure is described by means of connector semantics. Therefore, every instance graph has to be analysed. There can be a relation between every combination of two processes (x, y) for every semantic definition. The core algorithm is build-up as follows: At first, we assume that a relation exists between the currently observed processes x and y. Now, the left side of the semantic definition is examined. Therefore, the processes and the corresponding process context information have to satisfy the conditions. If the information within the instance graphs is not sufficient, additional data is extracted from an ontology. If all conditions are satisfied, the consequences are examined. We make use of the principal proof by contradiction. All elements of the right side of the definitions are negated. If the negation of the consequence could be found within an instance graph, it means the algorithm disproved the current semantics for the processes x and y. If we could not find any counter-example within the instance graphs, the semantics is valid for the processes x and y. Note, that our assumption only holds, if the log is scanned completely every time. The extracted semantics is only valid for the currently analysed knowledge base. The processes x and y are connected with the corresponding model element, whose semantics is defined by the rule. The fact that we use a proof by contradiction is based upon the closed-world assumption. If we cannot find a counter-example within any instance graph (these in combination with an underlying ontology reflects our knowledge base), the semantics is declared as valid for the currently considered processes. The formalisation of the described proof is as follows:

Discovery of Semantics s between processes P_1, P_2:

$proof = \text{true}$
∀ Instance Graphs
{
  IF $s.Conditions$ **THEN**
  {
    IF $\neg s.Consequences$ **THEN**
      $proof = \text{false}$
  }
}

IF $proof$ **THEN**
  Semantic s is discovered between P_1 and P_2

The assembly of a validation algorithm for attribute-semantic definitions is similar to the one above. However, the algorithm takes only one process x and in case of a valid semantics, an attribute is assigned to x instead of a connector. *Example:* Consider the three instance graphs of Fig. 2 as the knowledge base for this example. We now apply the algorithm to discover the *TraineeConnector* between the processes *C* and *D*. At first, graph 1 (Fig. 2 ①) is examined. In this case all conditions are satisfied as the processes C and *D* have assigned agents and the performing agents are obviously members of a class "Trainee". Therefore, the algorithm demands for the examination of the consequences. This is, in the case of the *TraineeConnector,* the fact that process *D* has to be started after the completion of process *C*. As the algorithm follows the principle proof of contradiction, we have to examine if the graph contains the contrary. That is obviously not the case as *D* is performed after *C* and therefore the *proof* variable stays true. In the next step, graph 2 (Fig. 2 ②) is examined. Here, the conditions are not satisfied, as the performing agent "Agent 3" is not member of a class "Trainee". Hence, the consequence must not be examined and the *proof* variable stays true. Note, that in this case *C* was performed after *D* (wrong order). However, this is not relevant as the performing agent is not a trainee. In the last instance graph (Fig. 2 ③) the situation is identical to graph 1. This is why the *proof* variable stayed true during the whole examination and the algorithm discovered the *TraineeConnector* between the processes *C* and *D*.

C. Generation of the Semantic Graph

During the validation of semantics, we assemble a new graph (called *semantic graph*) containing all the extracted attributes and relations between, processes. The graph consists of a set of nodes N and a set of edges E. The nodes represent the processes, whereas the edges represent relations between these processes. Every node $n \in N$ has the following fields: process name, performing agents, used documents, used tools and furthermore a list containing the discovered attributes for this specific process. As described before, connector semantic definitions represent relations between two processes. Therefore, connector semantic definitions are depicted as edges $e \in E$ within the semantic graph. An edge e has two fields describing the connector between two processes: connector-type (i.e., *TraineeConnector*) and distance (direct, transitive). The proof of an attribute semantics adds an entry to the attributes-list of the node representation of the corresponding process. Furthermore, the proof of a connector semantics concerning the processes x and y adds an edge between the node representations of these two processes with the corresponding connector type. The semantic graph of Fig. 3 highlights three exemplary visualised semantic definitions extracted on the basis of the three instance graphs of Fig. 2. The algorithm discovered that process *C* and *D* have a strict execution order only in case that a trainee employee performs these processes. Note, that we assume a closed-world and cannot found a counter-example in any instance graph. That is why the semantic graph shows the

TraineeConnector (visualised as a dashed arrow) between *C* and *D*. Moreover, the graph shows various other semantic definitions, discovered by analysing the instance graphs of Fig. 2. Consider the processes *A* and *B*. Here, the algorithm discovered a strict execution order in any case (i.e., the *StrictOrderConnector* visualised by a continuous arrow), as B has always been executed after *A*. Furthermore, the algorithm discovered an attribute of process *A* revealing that *A* has always been executed by the head of department (that is why *A* is visualised with black filling).

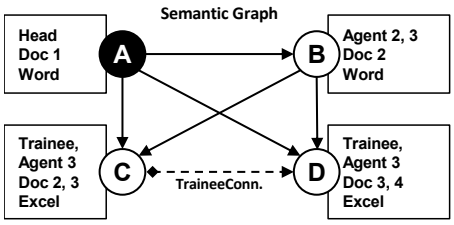


Figure 3. Semantic graph containing cross-perspective semantics.

Finally, the resulting graph contains all processes, the process context information and all attributes and connectors between processes that could have been extracted from a set of instance graphs.

D. Encapsulation of Processes

We propose to simplify and structure process models by examining group semantic definitions that additionally allow for encapsulating processes based on other perspectives. The approach allows for assigning processes to sets, i.e., groups, whose members offer common behaviour or characteristics with respect to a specific perspective. These groups of processes can be highlighted by assigning modelling elements to each predefined group semantics.

1) Encapsulation based on the behavioural perspective

In the first part of this section, we focus on encapsulation of processes on the basis of the behavioural perspective. Note that the execution order of processes could be based on various other perspectives. The execution order is for example mostly determined by the data flow. The encapsulation of processes based on the behavioural perspective has the aim to simplify process models by assigning atomic processes to complex processes as sub-processes. While an atomic process is associated with an executable activity, a complex process contains sub-processes [18]. This allows for the hierarchical composition of processes. Therefore, a complex process serves as a "*capsule*" for its sub-processes. The algorithm to encapsulate processes with respect to the behavioural perspective is based upon the following two rules:

$$\text{sameGroup}(?P_1, ?P_2) \wedge \text{startedAfterFinished}(?P, ?P_1) \Rightarrow \text{startedAfterFinished}(?P, ?P_2)$$

$$\text{sameGroup}(?P_1, ?P_2) \wedge \text{startedAfterFinished}(?P_1, ?P) \Rightarrow \text{startedAfterFinished}(?P_2, ?P)$$

The rules say: two processes P_1 and P_2 are in the same group (capsule) if every process P that is started (completed) after (before) P_1, is also started (completed) after (before) P_2. That means, two processes are in the same capsule, if they always have the same predecessors as well as the same successors with respect to all instance graphs. In order to discover this semantics, we need two rules as the control variable process P cannot be a predecessor and a successor at the same time. Therefore, we first have to check if P_1 and P_2 have the same predecessors and subsequently if they have the same successors. A formal definition of the encapsulation based on the behavioural perspective is as follows:

```
∀ process combinations (P1, P2)
∈ Set of processes recorded in the log
{
   proof = true
   ∀ Instance Graphs
   {
      P ∈ Set of processes recorded in the log
         ∀ p ϵ P \ {P1, P2}:
               IF s1.Conditions THEN
               {
                     IF ¬s1.Consequences THEN
                           proof = false
               }

               IF s2.Conditions THEN
               {
                     IF ¬s2.Consequences THEN
                           proof = false
               }
   }

   IF proof THEN
         P1, P2 ∈ Group x
}
```

It is obvious that the assembly of the algorithm is similar to the algorithm to discover connector semantics. However, here we have to examine two rules at the same time and a control variable *P* influences the examination.

Example: Consider again the three instance graphs of Fig. 2. We focus on the process combination *A*, *B*. In the first part of the formalism the focus is on common predecessors. Obviously, no instance graph contains predecessors for *A* and *B*, so the *proof* variable stays true. The second part of the algorithm examines the successors. Every process *P* that is started after the completion of *A*, has also to be started after the completion of *B*. Given that $P = \{C, D\}$ we have to examine which $p \in P$ are successors of *A* and check if p is also a successor of *B*. Focusing on the three instance graphs of Fig. 2 it is obvious that process *A* has the successors *C* and *D* in every case. Moreover, *C* and *D* are also successors of process *B* in every instance. Hence, the algorithm could not find a counter-example within the knowledge base and the *proof* variable stayed true during the examination. Finally, the algorithm assigned *A* and *B* to a set *Group x*. By applying this procedure recursively to the resulting sets, process hierarchies are discovered. In Fig. 4, we visualise

 ISBN: 978-1-61208-223-3

the group membership property by a box surrounding the contained processes.

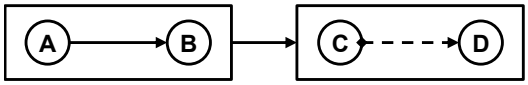


Figure 4. Exemplary visualisation of group semantics.

2) Encapsulation based on other perspectives

Besides, our semantic definitions allow for the encapsulation on the basis of other perspectives. In this paper, we exemplarily focus the encapsulation with respect to the organisational perspective. However, encapsulation based on any other perspective is possible. Exemplarily, we encapsulate processes based on the organizational perspective. If the agents' organization is not recorded in the log explicitly, we can extract it from ontologies by querying with the corresponding agent id. In our example, we visualise the group membership property again by a box element surrounding the contained processes. Note, that the resulting visualisation is related to pools in BPMN [17]. The group semantics to search for looks as follows:

$$\text{sameGroup}(?P_1, ?P_2) \wedge \text{hasAgent}(?P_1, ?A_1) \wedge \text{hasAgent}(?P_2, ?A_2) \wedge \text{employedAt}(?A_1, ?B1) \wedge \text{employedAt}(?A_2, ?B2) \Rightarrow sameAs(?B_1, ?B_2)$$

Here, we do not have to examine a common behaviour with respect to other processes and therefore there is no need for a control variable. In the case of the organisational perspective, we propose a solution to the challenge of cross-organisational mining. Consider a cross-organisational process, e.g., supply chains of manufacturing processes. Each involved organization records its activities to its specific event log. We propose to merge the different logs to one integrated log based on the recorded timestamps. The prerequisite in order to mine for instance graphs is that the involved organizations have a common instance ID. Note that especially in the case of a supply chain, the instance id could be based on a common product id. In the context of cross-organisational mining it could be interesting how the different parties interact with each other. Fig. 5 shows how cross-organisational interaction as for a product P_1 could be highlighted based on merging different logs and grouping different processes with respect to the performing agents' organization.

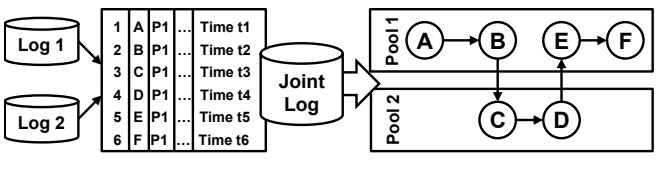


Figure 5. Merging of logs and encapsulation based on organizations.

V. Related Work

Starting point for process mining is a process execution log. The basis for analysis is pre-processing the available log. In this paper, the pre-processing is carried out by converting the log information to so-called instance graphs. These graphs are similar to the graphs of [11, 24, 25], however we feature them with further context data of process execution. There are already several algorithms and even complete tools, like the *ProM Framework* [8], that aim at discovering and generating process models automatically.

During the last decade, several algorithms have been developed, focusing different perspectives of process execution data. Van der Aalst et al. give a detailed introduction to the topic process mining and a recapitulation of research achievements in [5, 6, 10]. Many of these traditional process mining algorithms are imperative approaches. These methods construct imperative models explicitly showing all possible behaviours. Other ways to mine for process models are declarative approaches. Instead of explicitly specifying all the allowed sequences of events, declarative process models specify the possible ordering of events implicitly by constraints. Alongside with declarative approaches that are used for conformance checking of already existing models [10, 22], there are several declarative discovery algorithms like [1, 19, 20]. The major difference to these methods, that are mainly examining the behavioural perspective, is that our approach is based upon semantic definitions considering multiple perspectives of process data and includes the possibility to query ontologies. This enhances the declarative discovery process by functionality provided by semantic process mining [13, 14] and allows for revealing dependencies that are not obvious on the first look. In literature, several approaches [2, 3, 4, 12] are described to filter the information contained in a log and to simplify less-structured process models. However, these methods are limited to encapsulate processes based on common behaviour with respect to the control-flow. We propose to simplify and structure process models by examining cross-perspective semantic definitions.

VI. Conclusion and Outlook

In this paper, we suggested a declarative mining approach, based upon explicit, cross-perspective semantic definitions. Semantic definitions are constituted through the analysis of the different perspectives recommended by the *perspective-oriented process modelling* (POPM) approach and spread over different process-involved entities and perspectives. Cross-perspective semantic definitions are important for the analysis of business processes. Additionally, we proposed an approach to simplify and structure process models by examining cross-perspective semantics that allows for encapsulating processes based on various perspectives. By grouping processes with respect to the organisational perspective, we proposed a contribution to the challenge of cross-organisational mining. The *IEEE Task Force on Process Mining* stated the improvement of usability and understandability for non-experts as a challenge for future mining approaches [8, 9]. We strive for improving these issues with respect to the application environment as well as the presentation of results. The latter

 ISBN: 978-1-61208-223-3

is amended by the use of domain specific modelling elements, whereas the prospective integration in the *Open-Meta-Modelling Environment* [16] will allow for the simple definition of modelling elements with underlying semantics. The user should have the possibility to specify the semantics to search for and therefore to extend or reduce the scope of functions flexibly. A kind of instruction manual for process managers could look as follows: at first, analysts define the semantic definitions in SWRL notation they are most interested in. Next, a process modeller assigns user-defined model elements to these semantic definitions. The resulting model elements represent the domain specific process language that is used in the current case. After that, a log is selected as an input for the algorithm. Subsequently, the semantic process mining algorithm discovers the given semantic definitions from the language-catalogue. By traversing the resulting graph, it is possible to transform the information into a process model. Considering the assigned model elements, the process model based on the previously defined process meta-model, i.e., the language catalogue, is visualised. The discovered model can be discussed, possibly remodelled and finally be executed by a workflow engine.

We strive for an extensive application of our approach, including a detailed evaluation. Our future research activity in the field of process mining will face the problem that process events are typically not recorded in a unified manner. Therefore, we are developing an approach to align process logs recorded with different granularities.

REFERENCES

[1] Maggi, F.M., Mooij, A.J., and van der Aalst, W.M.P.: "User-Guided Discovery of Declarative Process Models". In: Proceedings of IEEE Symposium on Computational Intelligence and Data Mining (CIDM), April 2011, Paris, France, pp. 192–199.

[2] Li, J., Bose, R.P.J.C. and van der Aalst, W.M.P.: "Mining Context-Dependent and Interactive Business Process Maps using Execution Patterns". In: Business Process Management Workshops, Lecture Notes in Business Information Processing, Vol. 66, 2011, pp.109-121.

[3] Günther, C. and van der Aalst, W.M.P.: "Fuzzy Mining – Adaptive Process Simplification Based on Multi-Perspective Metrics". In: Business Process Management, LNCS, vol. 4714, 2007, pp. 328-343.

[4] Bose, R.P.J.C. and van der Aalst, W.M.P.: "Abstractions in Process Mining: A Taxonomy of Patterns". In: Business Process Management, LNCS, vol. 5701, 2009, pp. 159-175.

[5] Van der Aalst, W., Reijers,H., Weijters, A., Van Dongen, B., De Medeiros, A., Songa, M., and Verbeek, H.: "Business process mining: An industrial application". In: Information Systems, vol. 32 (5), 2007, pp. 713-732.

[6] Van der Aalst, W., Weijters, T., and Maruster, L.: "Workflow mining: Discovering process models from event logs". In: IEEE Transactions on Knowledge and Data Engineering, vol.16 (9), 2004, pp. 1128-1142.

[7] Van der Aalst, W.M.P., Pesic, M., and Song, M.: "Beyond Process Mining: From the Past to Present and Future". In: Advanced Information Systems Engineering, LNCS, vol. 6051, 2010, pp. 38-52.

[8] Van der Aalst, W.: "Challenges in Business Process Mining". In: Applied Stochastic Models in Business and Industry, 2010

[9] Van der Aalst, W.M.P., Dustdar, S.: "Process Mining Put into Context". In: IEEE Internet Computing, vol. 16, no. 1 (2012) 82-86.

[10] Van der Aalst, W.M.P.: "Process Mining: Discovery, Conformance, and Enhancement of Business Processes", Springer-Verlag, 2011, Berlin-Heidelberg, ISBN 978-3-642-19344-6.

[11] Van Dongen, B.F. and van der Aalst, W.M.P.: "Multi-phase Process Mining: Building Instance Graphs". In: Conceptual Modeling (ER), LNCS, vol. 3288, 2004, pp. 362-376.

[12] Bose, R.P.J.C, Verbeek, E.H.M.W., and van der Aalst, W.M.P.: "Discovering Hierarchical Process Models Using ProM". In: Proceedings of the CAiSE Forum, 2011, London, UK, CEUR Workshop Proceedings, vol. 734, pp. 33-40.

[13] De Medeiros, A.K.A, Pedrinaci, C, van der Aalst, W.M.P., Domingue, J., Song, M., Rozinat, A., Norton, B., and Cabral, L.: "An Outlook on Semantic Business Process Mining and Monitoring".In: On the Move to Meaningful Internet Systems (OTM Workshops), LNCS, vol. 4806, 2007, pp. 1244-1255.

[14] De Medeiros, A.K.A., van der Aalst, W.M.P., and Pedrinaci, C.: "Semantic process mining tools: core building blocks". 16th European Conference on Information Systems, June 2008, Galway, Ireland.

[15] Jablonski, S. and Goetz, M.: "Perspective Oriented Business Process Visualization". In: Business Process Management Workshops, LNCS, vol. 4928, 2008, pp. 144-155.

[16] Volz, B., Zeising, M., and Jablonski, S.: "The Open Meta Modeling Environment". In: Workshop on Flexible Modeling Tools (FlexiTools), May 2011, Waikiki, Honolulu, USA.

[17] Object Management Group (OMG): "Business Process Modeling Notation Version 2.0 – OMG Standard", last access: march 2012.

[18] Weber, B., Sadiq, S., and Reichert, M.: "Beyond rigidity – dynamic process lifecycle support". In: Computer Science - Research and Development, vol. 23, no. 2, 2009, pp. 47-65

[19] Chesani, F., Lamma, E., Mello, P., Montalo, M., Riguzzi, F., and Storari, S.: "Exploiting Inductive Logic Programming Techniques for Declarative Process Mining". In: Transactions on Petri Nets and Other Models of Concurrency II, LNCS, vol. 5460, 2009, pp. 278-295.

[20] Bellodi, E., Riguzzi, F., and Lamma, E.: "Probabilistic Declarative Process Mining". In: Knowledge Science, Engineering and Management, LNCS, vol. 6291, 2010, pp. 292-303.

[21] Horrocks, J., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., and Dean, M.: "SWRL: A Semantic Web Rule Language Combining OWL and RuleML". W3C Member Submission, last access: March 2012.

[22] Van der Aalst, W. M. P., de Beer, H. T., and van Dongen, B.F.: "Process Mining and Verification of Properties: An Approach Based on Temporal Logic". In: On the Move to Meaningful Internet Systems (OTM Workshops), LNCS, vol. 3760, 2005, pp. 130-147.

[23] Igler, M., Jablonski, S., Günther, C.: "Semantic Process Modeling and Planning". In: SEMAPRO 2010, The 4th International Conference on Advances in Semantic Processing, pp. 199-204, October 2010, Florence, Italy.

[24] Pinter, S., Golani, M.: "Discovering workflow models from activities' lifespans". In: Computers in Industry, vol. 53 (3), 2004, pp. 283-296

[25] Hwang, S., Yang, W.: "On the discovery of process models from their instances". In: Decision Support Systems, vol. 34 (1), 2002, pp. 41-57.

 ISBN: 978-1-61208-223-3

Learning Business Rules for Adaptive Process Models

Hans Friedrich Witschel, Tuan Q. Nguyen, Knut Hinkelmann
Fachhochschule Nordwestschweiz FHNW
Olten, Switzerland
hansfriedrich.witschel@fhnw.ch, nguyen.quoctuan@students.fhnw.ch, knut.hinkelmann@fhnw.ch

***Abstract*—This work presents a new approach to handling knowledge-intensive business processes in an adaptive, flexible and accurate way. We propose to support processes by executing a *process skeleton*, consisting of the most important recurring activities of the process, through a workflow engine. This skeleton should be kept simple. The corresponding workflow is complemented by two features: firstly, a task management tool through which workflow tasks are delivered and that give human executors flexibility and freedom to adapt tasks by adding subtasks and resources as required by the context. And secondly, a component that learns business rules from the log files of this task management and that will predict subtasks and resources on the basis of knowledge from previous executions. We present supervised and unsupervised approaches for rule learning and evaluate both on a real business process with 61 instances. Results are promising, showing that meaningful rules can be learned even from this comparatively small data set.**

Keywords – business process intelligence; process mining; knowledge work; workflow management

I. INTRODUCTION

Offering the best possible support for the execution of core business processes is a mission-critical requirement for many companies. Such support can be offered, e.g., through workflow management systems (WfMS), which help to automate process executions. The advantages of automating business processes with WfMS consist in the increased consistency, traceability and hence quality of process executions as well as gains in efficiency.

WfMS [1] require a model of the business process; such business process models are traditionally created in a top-down way by modeling experts, an approach which brings with it various problems:

- Agility: process models need to adapt changing conditions. This is hard to achieve if modeling requires an expert.
- Cost: modeling business processes is time-consuming, especially if a great accuracy should be achieved.
- Accuracy: process models should be consistent with reality in the sense that they either describe accurately how a process is executed or – in the case of automation through WfMS – that the execution really follows the model (i.e., no "by-passes" occur). Such consistency is almost never achieved because the cost would be too high in many situations.

These problems are especially serious in processes which we call *knowledge-intensive*, i.e., ones that involve complex decisions at runtime (e.g. exceptional situations, highly variable situations) and can include a wide variety of resources depending on the context.

In this paper, we present an integrated approach to modeling and managing business processes which addresses the above problems through a combination of top-down modeling – resulting in a simple *process skeleton* – with bottom-up techniques: end-user contributions and automatic *learning of business rules.*

The process skeleton includes what is stable – in most business processes, even if they are very loosely structured and highly knowledge-intensive, there exists a kernel of tasks that always has to be executed. This skeleton is usually easy to identify and can thus be modeled quickly and at low cost.

Then, the skeleton is combined with business rules that predict subtasks and resources based on the context (e.g. the attributes of previous activities in the process) and on historical user behavior. By observing user behavior, we ensure both accuracy and agility: because they are based on usage history, the rules reflect how the process is executed in reality – and new rules can be learned at any time.

Thus, taken together, the process skeleton and the business rules form an *agile process model* that is adaptive and maintainable at low cost. It is the aim of this paper to show that it can also be accurate, i.e. that useful business rules can be learned (semi-)automatically to complement the process skeleton. These rules are meant to *support* human knowledge workers who execute a process – they can still choose to ignore the resulting recommendations and act according to their own experience.

In the rest of this paper, we will first examine related work in section II and outline our specific contribution. After giving some background regarding an existing approach that we build on in section III, we will outline our method for learning business rules in section IV, describe our experimental setup for evaluating in section V and present the results of that evaluation in section VI. Finally, section VII concludes the paper.

 ISBN: 978-1-61208-223-3

II. RELATED WORK

In [2], three dimensions of change for business processes are introduced:

- Dynamism: adapting the model (at design time) to evolutions in the execution of a business process
- Adaptability: reacting to exceptional circumstances at runtime
- Flexibility: being able to execute on a partially specified model (where the full specification is made at runtime)

The need for dynamism has been recognized by researchers a long time ago and has resulted in the emergence of the new research area of business process reengineering, see, e.g., [3].

The accuracy of process models which are created by human experts is a related problem: the complexity of most business processes often leads to discrepancies between the model and reality [4]. Therefore, process mining has been proposed as a means to discover process models from usage data: traces of real process executions, e.g. recorded actions in event logs, are the basis for learning process models automatically [3][5]. Besides the discovery of process models, checking the conformance of a model with traces of real executions is another discipline in process mining [6]. Finally, van der Aalst [7] has proposed various ways of enhancing process models, e.g. with information about execution times and relationships between human executors. Learning (decision tree) classifiers to predict decisions in the case of choices is another kind of process model enhancement [8]. It is related to our approach since a decision tree can be represented as a set of business rules that predict subsequent activities. In [9], the history of process executions is used to learn a model that recommends activities in a running process instance.

However, decision mining and process mining in general focus exclusively on the control flow, i.e., activities of a process model are treated as atomic units, their internal structure, e.g., resources or subtasks, are not considered or predicted. In [10], we have investigated how possible subtasks of process activities can be recommended based on the analysis of informal traces of work (email). However, in that approach we did not build on an existing WfMS to collect traces of process execution.

The challenges regarding adaptability and flexibility result from the fact that the traditional distinction between design time and run time in workflow management [11] can be less strictly followed in cases where the flow of action (traditionally modeled at design time) is largely determined by input that is received only at runtime.

Modeling languages that are based on business rules (e.g. [12]) have been claimed to be more flexible and expressive than graphical models – in terms of flexibility, it is possible to create *partial* process descriptions with business rules, and the expressiveness of rules is higher because they can take into consideration the run-time process context. Although these advantages are confirmed by [13] in a comparative evaluation of the two approaches, the authors of that study also find that specifying workflows solely through business rules has the drawback of requiring more technical knowledge and being more difficult to understand and maintain by humans than a graphical model.

Various other solutions have been put forward to achieve flexibility: as discussed in [1] and [10], trying to model all possible choices at design time is rarely possible and results in complex and unwieldy models.

Therefore, combinations of process models and business rules have been put forward. By separating the business logic from the process logic and representing it as rules, the complexity of process models can be reduced. The KISS approach [15] combines semantically enhanced process models with business rules to increase their flexibility. A similar approach is taken in [1][14] where a combination of a core process model (similar to our proposed skeleton) with *pockets of flexibility* is proposed – in which predefined or new process fragments are inserted into the core model at run time via rules (contained in so-called *build activities*). But again, this approach concentrates on the control flow, and does not consider e.g. resources used in process activities.

Other approaches for ensuring flexibility in processes consist in giving more freedom and responsibility to the human executors of processes. Examples of such endeavours are case handling [16] and task management approaches based on task patterns [17][18]. The task pattern approach relies on collaborative development of process knowledge. A combination of process skeletons with pattern-based task management is proposed in [19].

However, the idea of having at most a very coarse process model ("flexibility by granularity" [2]) and leaving details of process execution to humans is problematic since it fails to adequately support these persons in their work. In general, as pointed out in [9], as the flexibility of execution increases, the support offered by process-aware information systems usually decreases.

Finally, there are approaches that also propose to model families of business processes on a coarse level and then make them configurable such that they can be adapted to the concrete needs of a situation or company [19][20]. The focus of these approaches is, however, not so much on flexibility for individual processes, but on the *re-use* of models of recurring processes across companies.

A. Contribution

Our approach builds on and extends previous work as follows: We start from the notion of a *process skeleton*, similar to the core process model in [2]. That model is deployed as a workflow, but its activities can be handled flexibly by the executors in a *task management* framework (see [17][19]) by adding resources and subtasks at their own discretion. That handling is captured in an event log, which serves as the basis for *learning business rules* that recommend resources and subtasks in later process executions. This learning approach is similar to decision mining [8], but goes beyond it since it predicts task features, not only the control flow.

All in all, our proposed approach combines the advantages of previously introduced ones to respond to the challenges mentioned earlier: It reduces the *cost* of

 ISBN: 978-1-61208-223-3

modeling since it only requires a rough process skeleton to be deployed. Furthermore, it ensures *flexibility* and *adaptability* by giving responsibility to humans (task management) and triggering rules based on the context. On the other hand, it offers adequate *support* to knowledge workers through recommendations. By learning the rules from real process executions, this approach also ensures the *accuracy* of the resulting model.

III. BACKGROUND: THE KISSMIR APPROACH

As outlined in the previous section, this work builds on the KISSmir system, as described in [19]. KISSmir is based on a process skeleton that contains those activities of a business process that are always executed. The process skeleton is deployed as a workflow – human executors are assigned tasks that they need to complete.

The tasks are loaded into a task management application where they can be modified by adding resources (persons or documents), notes, subtasks or statements of problems that occur during execution.

Task executors are supported by recommendations of resources, subtasks and problem/solution statements that have been contributed by other persons during execution of the same activity (see [19] for details).

Any modifications of tasks, e.g., addition of resources, subtasks or problem statements, be it by copying from the task pattern or manual creation, are logged by the system.

IV. LEARNING BUSINESS RULES

The learning of business rules that we propose in this work is based on the KISSmir system logs. It consists in an initial preprocessing, followed by either supervised or unsupervised learning, as described in the following subsections. Both rely on co-occurrence of attributes, e.g. the fact that the presence of a certain problem (exceptional situation) triggers a certain subtask to be executed. We thus understand a business rule as an implication of the form A → B, meaning that the presence of the element A in a case implies that B should also be present.

A. Preprocessing

The data from the KISSmir system logs is transformed into feature vectors, each of which represents exactly one process instance. Formally, a feature vector for a case c is given by

$$\vec{c} = (c_1, \dots, c_k, s_1, \dots s_l, p_1, \dots, p_m, r_1, \dots, r_n)$$

The attributes are derived as follows:

- $c_1, \dots, c_k$ are attributes that describe the case as a whole (process variables)
- $s_1, \dots s_l$ describe the subtasks that have been added to *any* of the activities of which the case consists. Each attribute $s_i \in \{0,1\}$ denotes whether a given subtask (from the set of all subtasks in the whole log) is present in the given case ($s_i = 1$) or not ($s_i = 0$).
- In the same way, $p_1, \dots, p_m$ and $r_1, \dots, r_n$ describe the presence of problem statements (p_i) and resources (r_i) in a case.

Subtasks and problem statements are identified by their name. Obviously, the same kind of (sub)task (resp. problem) can be described by different names and it is thus sometimes difficult to match the subtask names that describe the same activity. On the other hand, subtasks are frequently accepted recommendations, which ensures consistent naming across process instances. The set of features used is given by all resources, subtasks and problem statements that actually occurred in the past, not limited to any fixed or predetermined set.

B. Supervised learning

Supervised learning through classification consists in predicting one value of a feature vector from some or all values of the other attributes, i.e., we want to learn a classifier that predicts the presence of a particular subtask s_i in a given case, based on the presence of other subtasks, resources and/or problems in the same case. This is promising since often the presence of an exceptional situation (as documented by problem statements) triggers the execution of a subtask, or one subtask triggers another etc.

1) Learning decision trees

Because of the temporal dependency between activities – as given by the workflow – we do not use the full feature vectors in each case: when we train a classifier for an attribute, say subtask s_i, we prune the feature vectors of all training cases such that they only include the case attributes c_i, plus those subtask, problem and resource elements that have been used in activities prior to or equal to the one to which the current class attribute (s_i in our example) belongs. In our experiments, we used the Weka machine learning library [22] and its implementation of the J48 decision tree learner [23] – we loaded the feature vectors into the tool, which then produced candidate decision trees out of them.

2) Selecting decision trees, deriving business rules

In a real-life scenario, most elements of feature vectors are 0, i.e., subtasks, resources etc. are used in only a few cases. For those attributes that do not strongly depend on the presence of other elements, this strong bias means that their corresponding decision tree will consists of only one node which is root and leaf at the same time and decides for a value of 0 – i.e., the classifier predicts the absence of the element in all cases. Obviously, these are not interesting business rules.

Only attributes whose presence depends significantly on others will result in a decision tree with more than one node. Therefore, choosing all learned decision trees that have more than one node is a simple, but – as we will see later – effective approach for filtering the initial set of decision trees.

Since each decision tree can be expressed as a set of rules, we can easily derive our targeted business rules from

 ISBN: 978-1-61208-223-3

the resulting set of trees. We will see later that in practice, most trees are not very deep (two or at most three levels), such that the resulting rules are simple and human-readable.

C. Unsupervised learning

Mining association rules is another approach to derive business rules which determine appropriate subtasks for certain problems, i.e., we are looking for co-occurences of problems and subtasks in process instances. For this, we transform feature vectors representing the process instances into *transactions*, where only the attributes with a value of 1 appear, say, e.g., $\{c_2, s_3, s_7, r_4\}$.

1) Mining associations

We then use the Apriori algorithm [24] from the Weka library to derive association rules. Because of the known weaknesses of approaches that use support and confidence – confidence does not accurately measure the statistical dependence between antecedent and consequence of a rule [25] – we chose to additionally analyse the co-occurrence patterns in the data using a likelihood ratio measure [26] that is particularly reliable in cases where frequency distributions are skewed (which is the case for our data set). For this purpose, we used the software tinyCC[1] which implements the likelihood ratio measure for computing the significance of co-occurrences of words in sentences of natural language text. We therefore represented the transactions as "sentences" where each sentence is formed of the subtasks, resources or problems that are present in a given case – each such element is a word of the sentence.

2) Selecting association rules

For the unsupervised variant of the rule learning, we need to rely on thresholds to filter rules. In case of the Apriori results, we can filter by confidence, in case of the tinyCC results, we can filter by the likelihood ratio values. The threshold needs to be set manually.

In both cases, we need to additionally discard rules that point into the past, i.e., ones where the consequence of a rule appears in an activity that occurs temporally before the activity to which the antecedent of the rule belongs.

V. EXPERIMENTAL SETUP

Within the School of Business at our university, there are two master programmes – "Business Information Systems" (BIS) and "International Management" (IM) – that have a very similar student selection process. The so-called "matriculation process" for the two master programmes, i.e., the process of checking and deciding on student applications and of communicating these decisions to applicants, forms the context of our experiments.

For each of the two master programmes, there is one secretary in the administration office (see the middle layer of the process model below) who performs the majority of activities within the student selection process.

The work of the two secretaries in the matriculation process is supported by an implementation of KISSmir. In this evaluation, we have concentrated on the sub-process "Check application" that consists of four sub-tasks as shown in Figure 1 and that is performed by the secretaries alone.

We logged the execution of 61 instances of this process with the KISSmir tool and then transformed the log data into feature vectors as described above. In this case, the case attributes were the previous degree of the applicant (DEGREE) and her nationality (COUNTRY).

Table 1 shows statistics about the other attributes that occurred in the 61 matriculation cases.

Regarding the number of resources, the statistics are somewhat misleading since 61 of the 64 distinct resources were attached to the tasks automatically by the workflow and contain information about the applicant.

This means that there were only 3 distinct resources being used. Since these were mutually exclusive, we created a single attribute (called WEBNAME below) that had the title of the resource chosen as value.

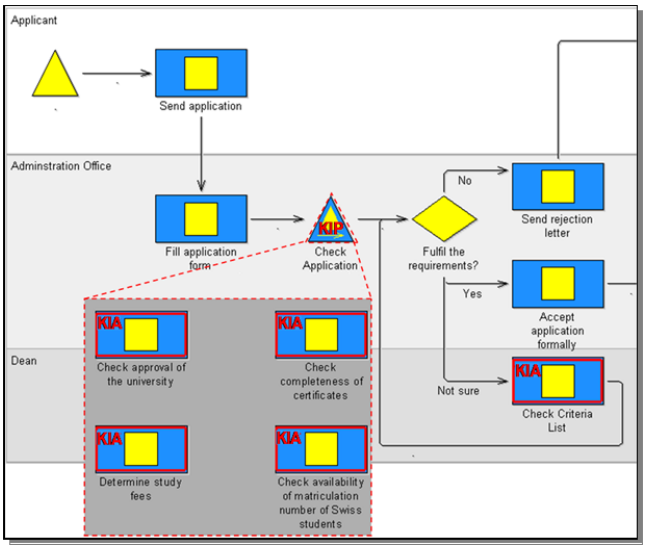


Figure 1. The matriculation process

TABLE 1. STATISTICS OF THE LOG DATA

Measure	Value
Number of cases (process instances)	61
Number of subtasks added by secretaries	34
Number of distinct subtasks	16
Number of problems used	19
Number of distinct problems used	7
Number of resource attachments	314
Number of distinct resource attachments	64

Thus, including the two case attributes, there are 28 attributes in each feature vector (2 case attributes, 16 subtask attributes, 7 problem attributes and 1 resource attribute).

[1] http://wortschatz.uni-leipzig.de/cbiemann/software/TinyCC2.html

 ISBN: 978-1-61208-223-3

VI. RESULTS

A. Supervised Learning

Classifiers are usually evaluated by measures such as accuracy of area under the (ROC) curve (AUC). In our case, however, we are not so much interested in the quality of the classification, but rather in the quality of the *model* that is learned – i.e., the decision trees. Because of the small scale of our experiment and since we know the process rather well, we were able to judge the degree to which the extracted trees made sense.

By applying the procedure for extracting business rules as described above in section 0, we found five non-trivial decision trees. Three of them have only two levels. We will thus describe them directly as rules where the antecedent of the rule corresponds to a value of the root node of the decision tree and the consequence of the rule describes the value of the attribute that should be predicted. For each rule, we add a short description that explains the rationale of the rule in terms of the matriculation process. In each case, we only present the "positive" variant, i.e., the variant that predicts the *presence*, not the absence of the class attribute.

- *Applicant has degree in a complete different area = 1 → Ask the dean = 1.* This rules predicts the presence of the subtask "Ask the dean" (consequence) in cases where an applicant had a degree that was definitely not compatible with the requirements of the master programme (e.g. the student had a Bachelor degree in French language and applied for the International Management master programme, antecedent). This accurately describes the current practice of the secretaries.
- *COUNTRY = Other → Check anabin = 1.* This rule predicts the presence of the subtask "Check anabin" in cases where the applicant is neither from an EU nor from an EFTA country. Anabin is an information platform where university degrees and accreditations can be compared for equivalence. Hence, this rule predicts that the platform should be consulted for students from non-European/non-EFTA countries.
- *Ask the student for a description of the financial situation = 1 → Forward scholarship documents to commission = 1.* This rule predicts the presence of a subtask – namely forwarding an applicant's scholarship application documents to a commission – from the presence of another subtask, as given in the antecedent. When a request for scholarship arrives, the secretaries will first ask the student for a description of his/her financial situation. When that description is received, the documents will be forwarded to the commission.

In addition to these simple rules, two three-level decision trees were discovered. One of them is shown in Figure 2: its leaf nodes (light grey boxes) contain the values of the attribute WEBNAME. That attribute describes which of three letter templates were chosen by the secretaries in their last task – "Accept application formally".

The tree should be read as follows: if a problem statement is present that indicates that the Bachelor degree of a student is still missing because the student is still studying, the student can be accepted, but one needs to use a special letter template in which the applicant is asked to hand in the certificate as soon as (s)he finishes his/her studies. If this is not the case, but the problem "Applicant has degree in complete different area" is present, a letter template should be used in which the student is informed that (s)he needs to take part in a so-called pre-master course to catch up with important foundations for their studies.

Only if none of the two problem statements is present, the standard letter template will be used. Although this tree is largely correct, the two attributes are actually independent, i.e., theoretically – yet very improbably – both could occur together.

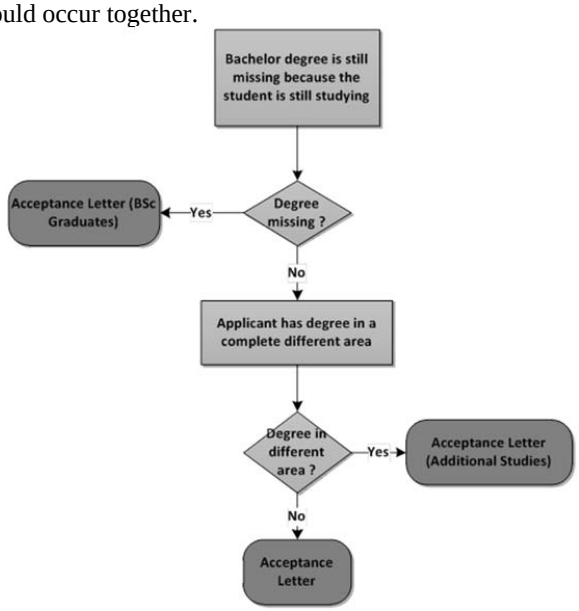


Figure 2. Decision tree for the attribute WEBNAME in the activity "Accept application formally".

Unsupervised learning

Table 2 shows the set of rules extracted with tinyCC using a likelihood ratio threshold of 10.0 and filtered by temporal consistency. They are to be read from left to right; the letters in brackets behind each attribute indicate whether it is a case attribute (c), a subtask (s), problem (p) or resource (r).

Some of these rules – numbers 1, 3 and 6 – are also extracted with supervised learning. The rest can be divided into useful additional rules (4 and 8, e.g., rule 4 will remind the secretaries to forward documents to the commission when a scholarship is requested) and over-generalisations (rules 2, 5 and 7, e.g., rule 2: checking the result of an interview with the student does not always result in using the acceptance letter that notifies about pre-master studies).

All in all, unsupervised learning produces slightly more rules, including more useful rules, at the cost of some noise.

 ISBN: 978-1-61208-223-3

VII. CONCLUSION AND FUTURE WORK

We have presented an approach to flexibly support business processes and to make process models more accurate by learning business rules. The approach starts from a simple process skeleton which is executed through a workflow engine. Human executors are given freedom and support to adapt tasks to their needs. Finally, the log files of this application are evaluated to learn business rules, which can be incorporated into the workflow using a rule engine to recommend subtasks in future process executions.

We have experimentally compared a supervised and an unsupervised approach on a real-life business process and found that both approaches yield meaningful rules even with a small training set. The supervised approach was more reliable, the unsupervised approach had higher recall.

In the future, we plan to implement our proposed approach in an environment with a larger number of participants and cases to test its scalability.

REFERENCES

[1] W.M.P. v. d. Aalst and K. van Hee, *Workflow Management: Models Methods, and Systems*. MIT Press, 2002.

[2] S. Sadiq, W. Sadiq, and M. Orlowska, "Pockets of Flexibility in Workflow Specifications," in *Proceedings of ER'01*, 2001, pp. 513 - 526.

[3] K. D. Swenson and K. Irwin, "Workflow technology: trade-offs for business process re-engineering," in *Proceedings of COCS 95*, 1995, pp. 22-29.

[4] W. M. P. van der Aalst, W. M. . van der Aalst, A. J. M. . Weijter, and L. Maruster, "Workflow Mining: Discovering process models from event logs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, 2003.

[5] J. E. Cook and A. L. Wolf, "Discovering models of software processes from event-based data," *ACM Trans. Softw. Eng. Methodol.*, vol. 7, no. 3, pp. 215-249, 1998.

[6] A. Rozinat and W. M. P. van der Aalst, "Conformance checking of processes based on monitoring real behavior," *Information Systems*, vol. 33, no. 1, pp. 64-95, Mar. 2008.

[7] W. M. P. van der Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer Verlag, 2011.

[8] A. Rozinat and W. M. P. van der Aalst, "Decision Mining in Business Processes," 2006.

[9] H. Schonenberg, B. Weber, B. Dongen, and W. Aalst, "Supporting Flexible Processes through Recommendations Based on History," in *Proc. of BPM'08*, 2008, pp. 51-66.

[10] S. Brander et al., "Refining Process Models through the Analysis of Informal Work Practice," in *9th International Conference on Business Process Management*, vol. 6896, S. Rinderle-Ma, F. Toumani, and K. Wolf, Eds. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2011, pp. 116-131.

[11] D. Hollingsworth, "The workflow reference model." The Workflow Coalition, 1993.

[12] J. Bae, H. Bae, S.-H. Kang, and Y. Kim, "Automatic control of workflow processes using ECA rules," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 8, pp. 1010-1023, Aug. 2004.

[13] R. Lu and S. Sadiq, "A survey of comparative business process modeling approaches," in *Proc. of BIS'07*, 2007.

[14] S. Sadiq, M. Orlowska, and W. Sadiq, "Specification and validation of process constraints for flexible workflows," *Information Systems*, vol. 30, no. 5, pp. 349-378, Jul. 2005.

[15] D. Feldkamp, K. Hinkelmann, and B. Thönssen, "KISS - Knowledge-Intensive Service Support: An Approach for Agile Process Management," in *RuleML*, 2007, pp. 25-38.

[16] W. M. P. van der Aalst, W. M. . van der Aalst, M. Weske, and D. Grünbauer, "Case Handling: A New Paradigm for Business Process Support," *Data and Knowledge Engineering*, vol. 53, 2005.

[17] E. Ong, O. Grebner, and U. V. Riss, "Pattern-based task management: Pattern lifecycle and knowledge management," in *Proc. of WM'07)*, 2007, pp. 357-367.

[18] B. Schmidt and U. V. Riss, "Task patterns as means to experience sharing," in *Proceedings of ICWL'09*, 2009, pp. 353-362.

[19] H. F. Witschel et al., "A Collaborative Approach to Maturing Process-Related Knowledge," in *Proc. of BPM'10*, 2010, vol. 6336, pp. 343-358.

[20] M. La Rosa, M. Dumas, A.H.M. ter Hofstede, and J. Mendling, "Configurable multi-perspective business process models," *Information Systems*, vol. 36, no. 2, pp. 313-340, 2011.

[21] I. Montero, J. Pena, and A. Ruiz-Cortes, "From Feature Models to Business Processes," in *SCC'08*, 2008, pp. 605-608.

[22] "Weka." [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka. [Accessed: 18-Jun-2012].

[23] R. Qinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.

[24] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proceedings of ACM SIGMOD*, 1993.

[25] S. Brin, R. Motwani, and C. Silverstein, "Beyond market baskets: generalizing association rules to correlations," in *Proc. of ACM SIGMOD*, 1997.

[26] T. Dunning, "Accurate Methods for the Statistics of Surprise and Coincidence," *Computational Linguistics*, vol. 19, no. 1, 1993.

TABLE 2. ASSOCIATIONS OF ATTRIBUTES EXTRACTED USING TINYCC

Nr	Attribute 1	Attribute 2	Sig
1	Applicant has degree in a complete different area (p)	Ask the dean (s)	26.55
2	Check interview result (s)	Acceptance Letter (Additional Studies) (r)	18.63
3	Bachelor degree is still missing because the student is still studying (p)	Acceptance Letter (BSc Graduates) (r)	17.74
4	How to handle Scholarship requests (p)	Forward Scholarship Documents To Commission (s)	17.39
5	Ask The Dean (s)	Acceptance Letter (Additional Studies) (r)	15.34
6	Applicant has degree in a complete different area (p)	Acceptance Letter (Additional Studies) (r)	10.8
7	Ask The Dean (s)	Check interview result (s)	10.78
8	iso3166#Other (c)	Anabin (r)	10.36

 ISBN: 978-1-61208-223-3

Business Process Simulation for Predictions

With focus on decision mining and execution time of tasks

Milan Pospisil
Department of Information Systems
BUT, Faculty of Information Technology
Brno, Czech Republic
ipospisil@fit.vutbr.cz

Tomas Hruska
Department of Information Systems
BUT, Faculty of Information Technology
Brno, Czech Republic
hruska@fit.vutbr.cz

***Abstract*—Simulation can be used for analysis, prediction and optimization of business processes. But models often differ from reality. Data mining techniques can be used for improving these models based on observations of process and resource behavior from detailed event logs. More accurate process models can be used not only for analysis and optimization, but for prediction and recommendation. This paper focuses on decision mining and the duration of tasks in conjunction with personal performance based on case data, workload, and other factors. Some existing ideas are an improvement and others are new. Part of the research was validated on real data.**

Keywords-business process simulation; business process intelligence; data mining; process mining; prediction; optimizatio; recomendation

I. INTRODUCTION

Classic simulation can be used for the analysis of business processes. We can try many scenarios, measure the effects, and then decide on the optimal process settings. For example, we can redesign the process, change resource allocation, and search for the most optimal configuration with respect to requirements (price, effectiveness, customer satisfaction, etc.). Or the current process can be tested for how many cases it can handle.

Nowadays, these models are often built manually, which is error-prone, and time consuming; the main drawback of this approach is that it cannot be used for operational decision support, but only for strategic decisions. This is because classic simulation models have several simplifications – probability the routing and statistical distribution of execution time of tasks. These models are sufficient in long-term simulation (usable for analysis), because simulation parameters are the result of long running processes. But, operational decision support needs short-term simulation. In this situation, we know the running and incoming cases, and the actual resource allocation. Therefore, actual running processes can differ from long measured processes. For example, task A needs to be done and there is a standard execution time of about 30 minutes; but, we have allocated a skilled resource and it is able to execute it under 20 minutes. These, and more problems need to be solved to obtain the simulation model for operational decision support.

Predictions, recommendations, and dynamic optimizations could be accomplished by operational simulation. The system can warn us, that some cases will be probably late. Then some different scenarios can be simulated and evaluated, then the system can recommend us actions and provide dynamic optimization of current running cases – for example; give extra resources from non-critical case to critical, or use a different sub-process – when we have a slower / cheaper version or faster but more expensive.

This work deals with the building of simulation models for operational decision support using data mining, because there is need to find deeper dependencies.

The paper is organized as follows. Related work is described in Section 2. Section 3 reveals the problem with classical simulation and answers the question, why it can not be used for operational decisions. Decision mining in Section 4 advances classic simulation. It is based on current research, but new improving ideas are sketched. Section 5 is about predicting execution time of cases and goes beyond current research in this area. Section 6 compares prediction using simulation and the standard approach by classification (or regression). At the end, Section 7 concludes whole paper.

Our work extends current research of simulation for operational decisions and new ideas are described, some new, some inspired by other works described below. Emphasis is placed on better decision mining and prediction of time execution of task with conjunction of personal performance based on case data, workload, and other factors.

II. RELATED WORK

Data mining techniques can be used in Business Process Management. This new area was called Process Mining [3, 6, 12, 13, 14]. It was based on analysis of information from event logs, that were produced by business processes. Process discovery is one of the methods and it is able to find a process model from an unknown process using many sequence examples of tasks.

 ISBN: 978-1-61208-223-3

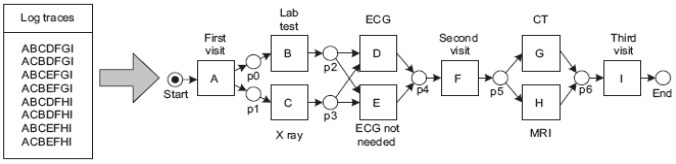


Figure 1. process discovery. We are able to discover a process model from log. The discovered process model must be able to replay most log traces.

A process log (figure 1) contains a sequence of tasks and we are able to discover what process model fits that log. Many algorithms are available for that nowadays and they were successfully used in practice. Discovered process models can be used for simulation, even if the model is not explicitly given or it is not usable (too low level detail, not all paths are described, and so on).

Different techniques are focused on performance analysis [4] (figure 2), where influential factors of Key Performance Indicators (KPI) are investigated.

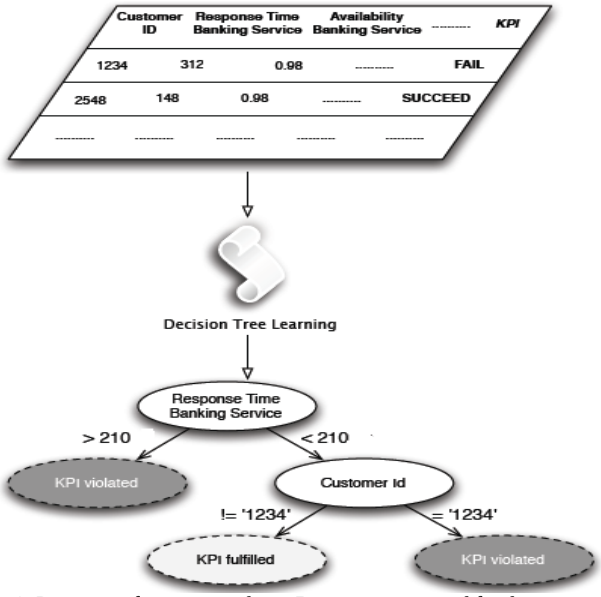


Figure 2. Process performance analysis. Decision tree is used for discovering factors that leads to KPI violation. We can see that KPI is violated when response time of banking service is larger than 210.

The table above (Figure 2) shows data that was collected by running cases (every row is one case). The target value is Key Performance Indicator (KPI) and it reports, if the case was executed right or not (it could be time, quality, or anything else, it depends on process manager). Based on these provided attributes, the target value (KPI) is predicted by a decision tree. Although we used the term 'prediction', this form of decision tree is usable only for analysis of historic data, not for real-time monitoring. We will discus it later. But the tree can tell us what combination of factors lead to KPI violation. For example, we can see, that if the response time of a banking service is higher than 210, the KPI is always violated. And also, cases with customer id 1234 has a problem with KPI too. Performance analysis can be valuable for managers because they can discover and focus on critical factors of processes.

Other work focused on the prediction of execution time [1, 2, 11] using classifiers [1, 2] or process discovery with time information [11]. Note that [1, 2] used similar techniques as [4], not for analysis, but for prediction. While [11] can be used only for time estimation, [1, 2] can predict other things like some events.

Work in papers [1, 2] is based on classifiers. Running cases can produce much usable data. For example, time execution of tasks (start, end) or some data passed from task to task. This information is written into the table - multiple execution of same task in loop is written only twice – first and last occurence. Then some classifier (neural network, decision tree, regression tree…) can be used for the prediction of the target value (total execution time of case or some other quality atribute). Authors test that method in some industrial applications and results were promising. This method will be discussed in Section 6.

Rozinat et al. [5] and Rozinat et al. [10] introduced the idea of building operational simulation models using Process Mining techniques as we described in the introduction. These methods were based mainly on process discovery and decision mining.

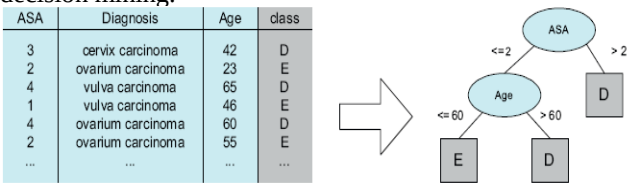


Figure 3. Decision mining. We can discover decision rules in routing points (OR-split nodes). Class attribute represents next task in running process. Decision tree is used for rule discovery.

Decision mining (Figure 3) enhances process discovery with decision rules. It is good to know the process model, but it is useless for simulation, when we do not know what task (in OR node) will be next and routing by percent (which is used in classic simulation models) is not sufficient for operational decisions.

Additional work deals mainly with resource modelling problems [7, 8, 9], which is now a topic of interest, because resources are one of the hardest things to simulate. They try to discover how to simulate resources and what factors influences their productivity – for example it was discovered (from some industrial experiments), that people tend to work faster when there is lot work to do – it is common knowledge, but we need some methods to compute it for every particular person (some people work more at a constant speed, some do not).

III. BUILDING SIMULATION MODELS

Imagine a typical example process (Figure 4) of handling warranties. The process model can be taken from a system, or discovered by Process Mining [5]. The first item is received and then checked for more information and the warranty. Then, a decision is made: the repair process is canceled (warranty not applicable) or send to repair. The

 ISBN: 978-1-61208-223-3

repair is either basic or advanced. At last, the item is returned back to customer.

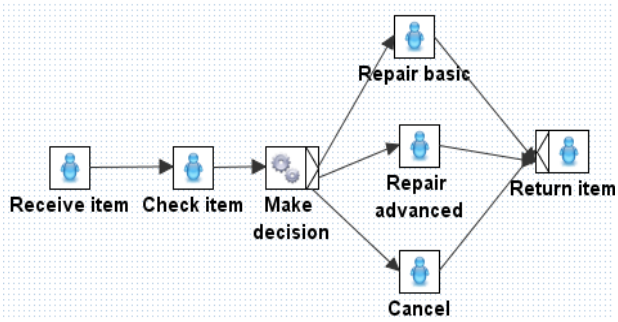


Figure 4. Example process model of handling warranties.

The typical simulation model would provide us a probability of routing (for example 20% of cases are cancelled, 60% sent to basic repair and 20% sent to advanced repair) and an average execution time of tasks with a standard deviation of some distribution (Repair basic takes 1 hour with 0.5 hour deviation, Repair advanced takes 4 hours with 2 hours deviation). We can use this simulation model for long-term simulation over several weeks, but not for short-term simulation over several days. Lets say, we have some running process instance. We are now behind the Check item and we have filled in information about the item (item type, damage type, etc). We want to predict the total execution time. Then routing probabilities are not sufficient, because we now know what type of repair it is and the differences between basic and advanced repair is significant.

IV. DECISION MINING

Decision mining can be used to discover what influences the decision of routing. Of course, we can also take that decision from the system (if available). But there is a catch. Decision expression would be probably simple and based on a few attributes known at the time of the decision (filled by a human based on the previous several attributes). In our process, decision rule could look like that – if RepairType is 'Cancelled' – go to Cancel, or if RepairType is 'Basic' – go to Repair basic, if 'Advanced – go to Repair advanced. That rule can be useless if one has not yet filled the attribute RepairType by the time of the decision. But we could have filled some important attributes – for example at the middle of the task Check item. Based on the provided information, one can make a better prediction of the next steps than the basic probability described above.

This can be solved as a classification problem. We have a table of attributes needed for a decision (Figure 5) and we want to predict the next step in the process – if the item will be repaired as basic, advanced, or cancelled.

Item age	Item type	Item type name	Damage	Repair type
3,1	Notebook	Notebook Acer	Broken monitor	Canceled
2,1	Notebook	Notebook HP	Broken HD	Basic
1,2	Mouse	Genius	Unknown	Basic
1,6	Notebook	Notebook Acer	Broken matherboard	Advanced
2,3	Mouse	Logitech	Broken glass	Basic
2,7	Notebook	Notebook Acer	Broken monitor	Canceled
1,3	Mobile	Errickson	Damaged keyboard	Basic
..	..	..	..	..

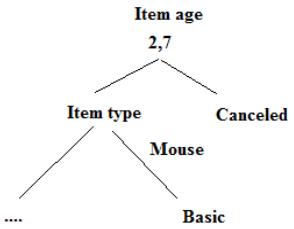


Figure 5. Decision mining. The data in the table is used to predict the next step in the process represented by the attribute Repair Type. The decision tree classifier is used here. Decision tree can be used for prediction and also for analysis.

A. *Extension*

Similar approach was described by Rozinat et al. [5], but there were several unsolved problems. First, the decision tree with these provided attributes will find this rule – if Repair type is 'Basic' – go to "Basic", etc. (we have discussed it above). It is not a mistake of the algorithm, the decision tree will simply find the attribute that decides everything with 100% precision. Of course, we can delete this attribute from the list. But still, that type of decision mining is good for analysis of historical data, but not for prediction, where not all attributes are provided at runtime. The most important attribute Repair type is filled last, so our prediction will work very poorly.

But, note that the Repair type attribute is filled by a human based on previous provided attributes, so we can still predict the next step even if we do not have all the attributes. The more important attributes we have, the Berger the prediction will work. Most classificators do not work so well on sparse data, so several classificators have to be used at different milestones of the process – a similar problem was solved in [1, 2]. That means we suppose some order of attributes and we build several classifiers using more and more attributes according to that supposed order; from Figure 5, we can see, that if the item type is mouse, it will be probably sent to basic repair. The item type will be one of the first filled attributes.

Another important thing is that many times, we cannot determine one precise decision (mostly when not all attributes are available). In that situation, more decisions have to be provided with some order of probability. A classic decision tree could give us only one final decision, but this can be solved quite easy by providing probabilities of classes in every node (mostly in leaves). Or, we can use a more advanced classifier like a neural network. A neural network can have output neurons corresponding to the next following available task – in our example, we can have three output

 ISBN: 978-1-61208-223-3

neurons, first will give us fitness (0..1) of the decision of Repair basic, second about Repair advanced and third about Cancel. A neural network also has some disadvantages. It works as a black box (opposite as a decision tree) and it needs more data to train. Also Bayesian classification could be used, if the attributes are not so dependent. That type of classification will work better than a neural network when not as much data is provided.

V. TIME EXECUTION OF TASKS

In [5], task execution times were modeled classically by a distribution with mean and deviation. This is not sufficient for short-term simulation. Task execution times can depend on several things. We will describe a new and better approach to this problem.

Some people work faster, some slower. Some are good at one task, some at another task. So resource information influences execution time. We have an example in Table I. From this table, we can deduct, that John is faster than Karl (he is also able to repair advanced items, but that information can be also found in process definition). More dependencies could be found (but not at Table I., we do not have so much space), for example mice are repaired faster than notebooks, etc. We can predict task execution time also by classification (Table I.).

These techniques and attributes can be used for the execution time of tasks and for decision mining. In fact, decisions can be influenced by resource, who is responsible for them, weather, time of day, etc. It is the reason, why these approaches have to be done semi-automatically. The process designer has to decide what attributes are needed for what decision. Also, there is an option to automatically find important data. But still, a human has to provide all the important data to system.

Additional works about resource modeling in simulation are in [7, 8, 9]. Human productivity can be influenced by many factors – by weather, day of week, time of day (especially after lunch or dinner). Another important factor is workload [9]. People tend to work faster, when there is full work queue, but not so long. After some time (it depends on the individual) productivity fails. There is a place for future research – resources have several attributes and those attributes can be measured. It could be productivity variance, ability to increase performance, when there is too much work, endurance to illness (people are usually weaker when the weather changes or during a flu epidemic). The question is how to deal with that information, because some tasks can be more influenced by resource productivity, and some tasks not. Mainly stereotyped work or machine operations – These tasks will probably be influenced more by data parameters of case (in our example - item type, damage type, etc.) than personal productivity.

We can put that information into the table and provide the same data mining techniques as in decision mining. It could work, but it requires more data to learn than the classifier. So, there is a space for research how to accomplish that with measuring resource parameters and without the need to put all this information into the classifier.

VI. SIMULATION AND CLASSIFICATION

Grigori et al. [1][2] uses classification to predict total execution time of cases (and potentially some events) using all important case parameters. It works similar to our approach, with distinction, that process model is not used. All process attributes are in one table and the final predicted value is the total execution time of cases.

In our simple example, this method does not differ from the simulation model, because we also used classification for decisions and predictions of execution time. It is because our model is too simple and we are using almost all attributes to predict both decision and time. In a more complicated model, not all attributes will be needed for classification.

What are advantages and disadvantages of these two approaches? Clearly, when there is no predictable process model behavior, or even no model available, classification based on all attributes will be better.

First, the simulation model will be better, when something changes – for instance, a faster machine, a change in the process model will be much worse for the classifier. We do not have to learn the whole classifier again, but we have to deal only with one change – the data needed to predict the execution task of the new machine can be provided from an expert, for example. Second, short-term simulation can give us what-if analysis. We can simulate several situations (with different resource allocation) and the system can choose the best solution.

Third, and maybe the most important advantage, is that this method is contextual. Prediction based on classification of all attributes without a model is bad for resource modeling. If case attributes show that the case will be in time that does not have to be true, because we can be short of workers. It is hard to give information to the classifier about the resource workload. But in the simulation model, we know what resources are available, what tasks can be accomplished by what resources, etc.

VII. CONCLUSION AND FUTURE WORK

We proposed an improvement of the existing simulation model for operational decisions. Improvement was based on better decision mining and mainly on the execution time of tasks. This work is now in progress, so ideas are described and compared with some other approaches.

We did some industry experiments and results were quite good. We were able to predict execution time 40% better than methods that do not také into account parameters of cases and were based only on global mean and deviation

 ISBN: 978-1-61208-223-3

of all execution time of cases. So we belive, that methods are able to improve prediction in some industrial companies.

Next research could be focused on industrial experiments and dealing with resources – we need to meause resource productivity at particular tasks.

We believe, this type of simulation will be able to support operational decisions and predict execution times of cases. But more work need to be done, mainly at the field of resource modeling.

ACKNOWLEDGMENT

This research was supported by the grants of MPO Czech Republic MPO TIP FR-TI3 039, by the grant FIT-S-10-2 and by the research plan no. MSM0021630528.

REFERENCES

[1] D. Grigori, F. Casati, M. Castellanos, U. Dayal, M. Sayal, and M.C. Shan, Business Process Intelligence, Computers in Industry, Volume 53, Issue 3, Process / Workflow Mining, April 2004, Pages 321-343, ISSN 0166-3615, DOI: 10.1016/j.compind.2003.10.007.

[2] D. Grigori, F. Casati, U. Dayal, and M.C. Shan, Improving Business Process Quality through Exception Understanding, Prediction, and Prevention Proceedings of the 27th VLDB Conference,Roma, Italy, 2001, 1-55860-804-4

[3] W.M.P. van der Aalst, H.A. Reijers, A.J.M.M. Weijters, B.F. van Dongen, A.K. Alves de Medeiros, M. Song, and H.M.W. Verbeek, Business process mining: An industrial application, Information Systems, Volume 32, Issue 5, July 2007, Pages 713-732, ISSN 0306-4379, DOI: 10.1016/j.is.2006.05.003.

[4] B. Wetzstein, P. Leitner, F. Rosenberg, I. Brandic, S. Dustdar, F. Leymann, "Monitoring and Analyzing Influential Factors of Business Process Performance," *Enterprise Distributed Object Computing Conference, 2009. EDOC '09. IEEE International, pp. 141-150, 1-4 Sept. 2009* doi: 10.1109/EDOC.2009.18

[5] Anne Rozinat, R.S. Mans, M. Song, and W.M.P. van der Aalst, Discovering simulation models, Information Systems, Volume 34, Issue 3, May 2009, Pages 305-327, ISSN 030

[6] Minseok Song and W.M.P. van der Aalst, Towards comprehensive support for organizational mining, Decision Support Systems, Volume 46, Issue 1, December 2008, Pages 300-317, ISSN 0167-9236, DOI: 10.1016/j.dss.2008.07.002.

[7] W.M.P. Van der Aalst, Business Process Simulation Revisited, 2010, ISSN: 1865-1348

[8] J. Nakatumba, A. Rozinat, and N. Russell, Business Process Simulation: How to get it right, 2010,Springer-Verlag,doi=10.1.1.151.834

[9] J. Nakatumba and W.M.P.V.D. Aalst, "Analyzing Resource Behavior Using Process Mining", in Proc. Business Process Management Workshops, 2009, pp. 69-80.

[10] Anne Rozinat, M.T. Wynn, W.M.P. van der Aalst, A.H.M. ter Hofstede, and C.J. Fidge, Workflow simulation for operational decision support, Data & Knowledge Engineering, Volume 68, Issue 9, Sixth International Conference on Business Process Management (BPM 2008) - Five selected and extended papers, September 2009, Pages 834-850, ISSN 0169-023X, DOI: 10.1016/j.datak.2009.02.014.

[11] W.M.P. van der Aalst, M.H. Schonenberg, and M. Song, Time prediction based on process mining, Information Systems, Volume 36, Issue 2, Special Issue: Semantic Integration of Data, Multimedia, and Services, April 2011, Pages 450-475, ISSN 0306-4379, DOI: 10.1016/j.is.2010.09.001.

[12] W. M. P. van der Aalst, and A. J. M. M. Weijters, Process mining: a research agenda, Computers in Industry, Volume 53, Issue 3, Process / Workflow Mining, April 2004, Pages 231-244, ISSN 0166-3615, DOI: 10.1016/j.compind.2003.10.001.

[13] W. M. P. van der Aalst, B. F. van Dongen, J. Herbst, L. Maruster, G. Schimm, and A. J. M. M. Weijters, Workflow mining: A survey of issues and approaches, Data & Knowledge Engineering, Volume 47, Issue 2, November 2003, Pages 237-267, ISSN 0169-023X, DOI: 10.1016/S0169-023X(03)00066-1.

[14] W. M. P. van der Aalst, Process Mining, Berlin, Heidelberg 2011, ISBN 978-3-642-19344-6

TABLE I.

Item age	Item type	Item type name	Damage	Repair type	Resource	Time
3,1	Notebook	Notebook Acer	Broken monitor	Canceled	Mary	1:00
2,1	Notebook	Notebook HP	Broken HD	Basic	Karl	2:00
1,2	Mouse	Genius	Unknown	Basic	John	1:30
1,6	Notebook	Notebook Acer	Broken matherboard	Advanced	John	10:50
2,3	Mouse	Logitech	Broken glass	Basic	Karl	3:30
2,7	Notebook	Notebook Acer	Broken monitor	Canceled	Mary	1:10
1,3	Mobile	Errickson	Damaged keyboard	Basic	Karl	1:50
..	..	..	..	..		

Execution times of tasks. Six attributes are used to predict execution time.

 ISBN: 978-1-61208-223-3

Transformation Framework for SBVR based Semantic Queries in Business Information Systems

Algirdas Sukys, Lina Nemuraite, Bronius Paradauskas, Edvinas Sinkevicius
Kaunas University of Technology
Kaunas, Lithuania
sukys.algirdas@gmail.com, lina.nemuraite@ktu.lt, bronius.paradauskas@ktu.lt, edvis.s@gmail.com

***Abstract*—The paper presents transformation framework from questions in structured language based on Semantics of Business Vocabulary and Rules (SBVR) into SPARQL queries over ontologies defined in Web Ontology Language OWL 2 and, possibly, supplemented with Semantic Web rules SWRL. Such transformation depends on OWL 2 ontology related with corresponding SBVR vocabulary and rules. The current work considers a family of transformations and metamodels required for relating ontologies, rules, SPARQL queries and real business data supported by computerised information systems, as well as establishes requirements for harmonizing the coexistence and preserving semantics of these different representations.**

Keywords-SBVR; SBVR question; business vocabulary; business rule; ontology; SPARQL

I. INTRODUCTION

Semantics of Business Vocabulary and Business Rules (SBVR) is the OMG metamodel [1], which has gained a great attention both in commercial applications and academic world. It is streamlined towards raising an abstraction level one more step upwards in the field of modelling business software systems as makes it possible to conceptualize software artefacts in some kind of structured language understandable for business people and computing systems. SBVR could not serve as the last step in the vision of communicating with computers in natural languages but it can be seen as intermediate specification between linguistically processed text and computer supported knowledge models as e.g. Web Ontology Language OWL 2 [2] and Semantic Web Rules (SWRL). Currently, there are many SBVR related applications, for example, generation of software models and code [3], creation of terminological vocabularies for various fields [4], auditing compliance between governmental regulations and business actualities [5], etc. In such context, our work is related with SBVR support for analyzing business information – i.e., formulating SBVR questions and transforming them into ontology query language SPARQL [6] for immediate access to business data from various sources (databases, business process management and enterprise resource planning (ERP) systems, documents), supported with ontological descriptions.

For implementing such SBVR into SPARQL transformations in real world enterprises, the complex infrastructure is needed for keeping connections between business vocabularies, vocabularies of business rules, ontology storages, databases and software service systems. The growing complexity of dealing with business information demonstrates need for such infrastructures, especially for large organizations or public institutions. As a little attention is given for SBVR questions in research literature and SBVR specification itself [1], we have investigated the peculiarities of modelling SBVR questions and representing them as SPARQL queries in our previous works [7], [8]. In the current paper, we present a transformation framework among SBVR questions and OWL 2, SWRL, SPARQL representations of business data.

The rest of the paper is organized as follows. Section 2 presents related works. Section 3 presents the framework of metamodels and transformations for transforming SBVR questions into OWL 2 and SPARQL; section 4 – examples of SBVR question patterns in SPARQL. Section 5 draws conclusions and outlines the future work.

II. RELATED WORKS

The SPARQL is a Semantic Web query language for RDF and OWL. SPARQL, as well as RDF and OWL, are knowledge representation centric languages oriented towards computer processing, therefore, they are not suitable for supporting immediate access to business data for business people. SBVR metamodel, concentrated on specifying the meaning of knowledge in terms of concepts, propositions and questions, is devoted for filling this gap by proposing the abstract syntax, expressible in structured languages, understandable by human, and, contemporaneously, having formal ontological representation. SBVR may be treated as multilingual representation of meaning in structured language intermediate between natural languages and executable software specifications [9]. SBVR and OWL 2 were created to be compliant with Common Logics [10]; therefore, SBVR concepts can be mapped to OWL 2 concepts. Rough correspondences between SBVR and first OWL version were declared in SBVR specification, but OWL was overridden by more expressive OWL 2, and explicit SBVR into OWL 2 transformation is still under investigation [11]. Though some implementations of SBVR into OWL 2 transformation exist, they are embedded in commercial products as Collibra tool suite [12] or ONTORule project components [13] in the restricted manner directly unusable for other research projects or applications. As SPARQL queries are executed over RDF and OWL 2 ontologies, transformation of SBVR vocabularies and rules

 ISBN: 978-1-61208-223-3

into OWL 2 (SBVR2OWL2) is important for possibility of expressing SBVR questions in SPARQL.

OWL 2, the latest version of Web ontology language [2], is provided with two kinds of semantics – direct semantics [14], based on Functional style syntax, and RDF based semantics [15]. As OWL 2 metamodel has changed from the previous OWL version, standard RDF based metamodel of OWL 2 still does not exist though OMG is making efforts for finalizing a new version of ontology definition metamodel [9], which should harmonize OWL 2 Functional and RDF syntax. For practical purposes, the Manchester University has created converter between various syntaxes of OWL 2, so it is possible to convert OWL 2 Functional style ontologies into RDF documents; but the reverse is not always true. OWL 2 Functional style ontologies are limited to Description Logics whereas RDF format is capable representing OWL 2 FULL, the most expressive language specified by the W3C OWL 2 standard.

Similarly, in Twouse project [16], SPARQL Abstract Syntax (SPARQLAS) metamodel is proposed, which is based on OWL 2 Functional Syntax and directed towards querying and two-way transforming of UML and OWL 2 models for software design, code generation and OWL 2 ontology engineering. SPARQLAS directly includes part of OWL 2 metamodel, so the advantage of using the SPARQLAS is a conceptual clarity perceived by developers, and decidability of reasoning and querying tasks; its shortcoming – the limitation to OWL 2 Description Logics (DL). There are other interesting works related with SPARQL, e.g. [17], which authors present solutions for transforming object oriented queries into SPARQL RDF query language and, vice versa, for translating SPARQL queries into object oriented ones. The approach allows to access RDF data from object models as well as to implement SPARQL endpoints for object oriented applications.

The mentioned works raise the question what SPARQL metamodel to choose for SBVR2SPARQL transformation. SPARQLAS has advantages for writing OWL 2 queries as users can do it faster; SPARQLAS queries can be automatically translated into SPARQL and executed using an OWL 2 reasoner and a SPARQL engine. SBVR to SPARQLAS transformations also would be much simpler but such transformations may be not suitable for all cases because the newest version, SPARQL 1.1, has dependencies on OWL 2 Full. Current ontology reasoners as Pellet or HermiT are based on OWL 2 DL, but there already are proposals as [18] for implementing OWL 2 Full reasoners on the base of first order logic formulas. As SBVR semantics also is not limited to OWL 2 DL, we have given the preference to RDF based SPARQL metamodel against OWL 2 Functional Syntax based one regarding the fact that the latter solution could be not applicable in a general case for SBVR questions that are considerably more expressive and should not be limited to OWL 2 DL capabilities.

RDF data model is based on subject-predicate-object expressions (triples) and is considered to be the most relevant standard for data representation and exchange on the Semantic Web. It gives a basis for other standards such as RDF Schema (RDFS) and OWL. OWL 2 ontologies can be mapped into RDF graphs and vice versa using special mapping rules [19].

Our transformation framework uses SPARQL RDF based metamodel capable to represent OWL 2 Full. Contemporaneously, we map SBVR concepts to ontology concepts with the requirement that SBVR transformation into OWL 2 (SBVR2OWL2) should preserve the same semantics as SBVR to SPARQL. For providing querying capabilities equal to SBVR questions, the SBVR transformation into to OWL 2 should represent in ontology SBVR related concepts as preferred, not-preferred and prohibited designations, synonyms and synonymous forms, predefined fact types etc. However, these terminological concepts should not make impact on behaviour of reasoners and other ontology processing means. Therefore, terminology and linguistics related SBVR elements should be separated from conceptual knowledge for efficient querying using synonyms or synonymous forms and getting correct answers.

III. SBVR BASED QUERYING FRAMEWORK

The process for accessing data and services of Business Information Systems by the means of SBVR questions over business vocabularies is presented in Figure 1. Business data and service software models could be generated from SBVR vocabularies and rules, and queried using SBVR questions.

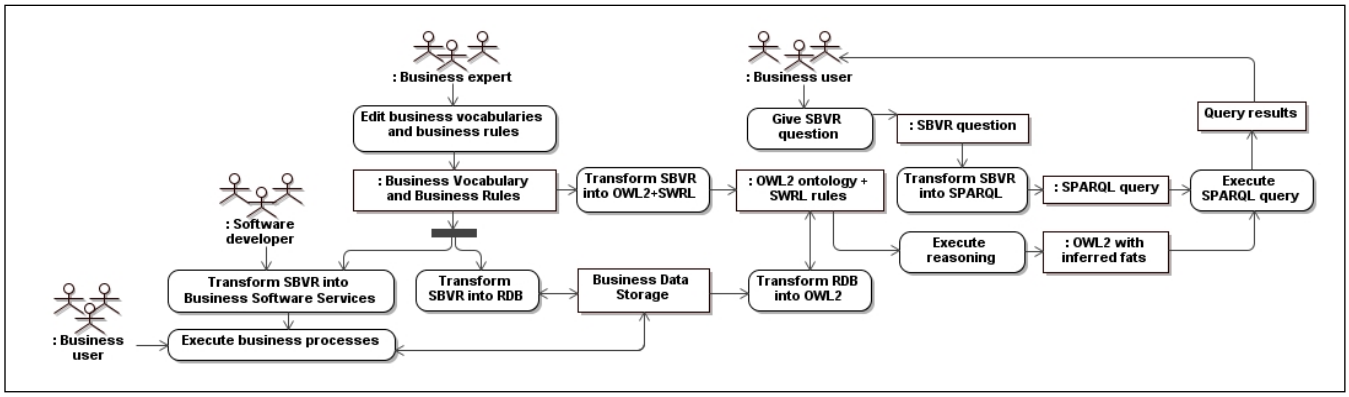


Figure 1. SBVR as a intermediate representation between data and services of Business Information Systems and user-understandable language

For the wholeness of the picture, we have included transformations among SBVR business vocabularies&rules and software services&databases. The latters are beyond the scope of the current paper but they are essential by giving a possibility for implementing business supporting software systems and querying about business state using the same business vocabulary and rules used for creating new (or conceptualizing existing) information technology assets. We suppose the usage of reasoners acknowledging OWL 2 axioms and SWRL rules that allow inferring additional facts and making SPARQL queries simpler – what should simplify translation from real natural languages into SBVR structured languages as well.

Metamodels and transformations required for implementing the querying process are represented in Figure 2. Two kinds of transformations are used: model–to–model transformations in ATL language for transforming SBVR models into OWL 2+SWRL (SBVR2OWLSWRL) and SBVR into SPARQL models (SBVR2SPARQL); and model–to–text transformations for transforming OWL 2, SWRL, SPARQL models into OWL 2, SWRL, SPARQL textual representations.

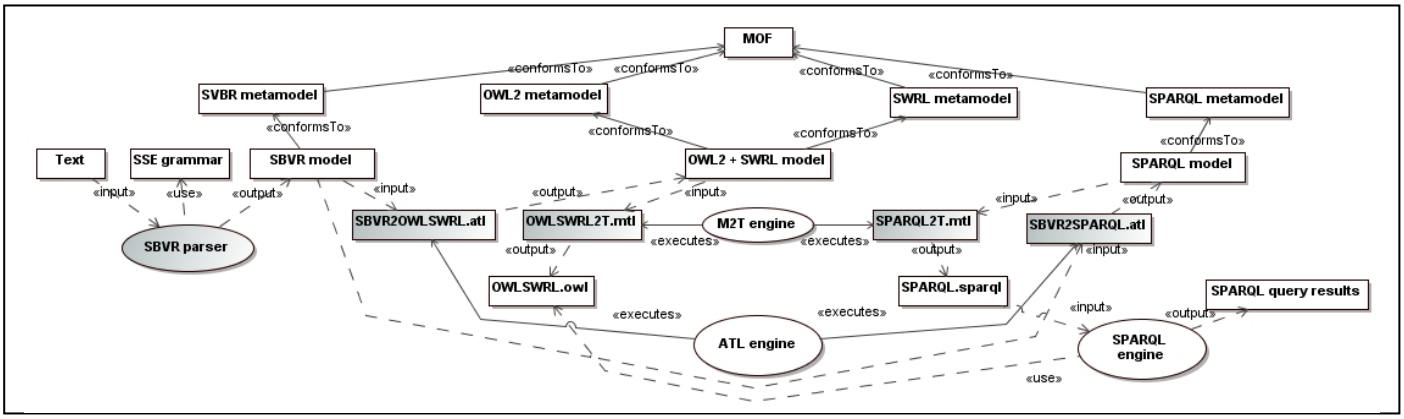


Figure 2. Metamodels and transformations of SBVR-based querying framework (shadowed components are implemented by authors and [11])

For alignment of SBVR, SPARQL, OWL 2, SWRL and RDF metamodels, we should establish mappings between relevant SBVR, OWL 2, SWRL, RDF, and SPARQL concepts.

Main concepts of SBVR used in formulating questions are object types, individual concepts, fact types, facts, variables and various kinds of logical formulations. Following recommendations for correspondence SBVR and OWL concepts [1], we suppose mapping SBVR object types to OWL 2 classes; SBVR binary fact types to OWL 2 object properties; SBVR *is_property_of* fact types to OWL 2 data properties; SBVR concept generalization to OWL 2 `SubclassOf`, `SubObjectProperty` or `SubDataProperty` with corresponding disjointness axioms; SBVR individual concepts to OWL 2 individuals, and some of SBVR logical formulations to OWL 2 axioms and restrictions. Additionally, we consider mapping of n-ary relations, roles and categorization schemes according [11], and the wider set of logical formulations. Only conceptual elements, i.e. SBVR meanings represented by preferred designations are transformed into OWL 2. SBVR synonyms and synonymous forms, non-preferred or prohibited designations should not be lost but may be included in terminological ontologies or annotations in such a manner that they would not make impact or require additional processing from reasoners acting on domain ontology.

In our chosen RDF based SPARQL metamodel (based on Eclipse EMFTEXT project), OWL 2 ontologies are represented by RDF triples. It means that all complexity of OWL 2 axioms is expressed using a few concepts as *VarOrTerm*, *Verb*, *Object*, representing OWL 2 `Class`, `DataProperty`, `ObjectProperty`, `Individual`. It hides conceptual clarity of OWL 2 but gives a great flexibility e.g. by allowing to access OWL 2 classes and axioms in the same manner as individuals and their property assertions.

SPARQL query language uses the power and flexibility of RDF. The conditions of queries are composed using sets of graph patterns, which are written in a same form as RDF triples, but can have variables in position of subject, predicate or object. SPARQL is not very user friendly especially when one needs to express complex OWL 2 statements in RDF to compose query condition. As our approach is based on generating SPARQL automatically from SBVR questions, written in structured natural language, the complexity of SPARQL makes no impact to end users.

SPARQL query consists of two main parts – at least one variable and conditional element – `WHERE` clause, containing graph pattern elements – *TriplesSameSublectLeft* (Figure 3). Graph pattern element of SPARQL consists of two *VarOrTerm* elements, representing subject and object, and *VarOrIRIRef* element, representing predicate of graph pattern.

OWL 2 ontologies are comprised of axioms that can be expressed with RDF. Having OWL 2 ontology in RDF format one can query it using SPARQL. World Wide Web Consortium defines mappings between structural specification of OWL 2 and RDF graphs. Using this mapping any OWL 2 ontology can be transformed into RDF graph without changing formal meaning of the ontology [19]. Some OWL axioms (for instance, `SubClassOf`, `SubObjectProperty`, `SubDataPropertyOf`) that can be represented using binary relation are transformed into single

triples. Other axioms (e.g., `EquivalentClasses`, `EquivalentObjectProperties`) are transformed into several triples. Annotations of axioms are also transformed into sets of triples.

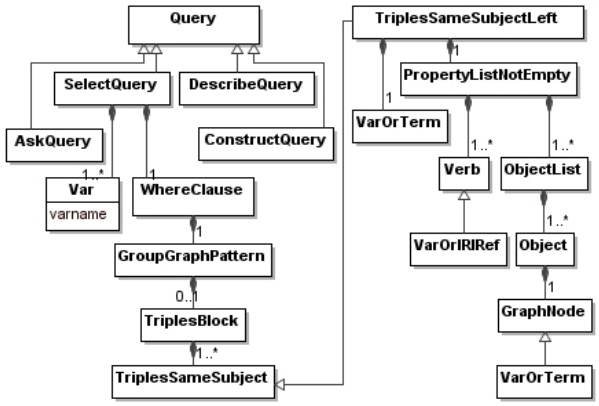


Figure 3. SPARQL metamodel fragment

IV. EXAMPLES OF APPLYING SBVR, OWL 2 AND SPARQL CONCEPTS FOR FORMULATING QUERIES

Transforming SBVR questions into SPARQL, it is essential to consider SBVR mappings to OWL 2 concepts for preserving the intended semantics during executing queries. It must be known how SBVR concepts are represented in OWL 2 to compose proper SPARQL queries. For example, SBVR uses instantiation formulations to classify things, while in OWL 2 there is an `rdf:type` property for this purpose. Only preferred designations of SBVR concepts are transformed for representing domain ontology in OWL 2; consequently, SBVR question concepts also are transformed into SPARQL using preferred designations. In the following we will show how SBVR concepts and formulations should be treated to compose SPARQL queries for OWL 2 ontology corresponding to SBVR questions in SBVR Structured English (SSE).

A. SBVR Facts and Binary Fact Types

SBVR binary fact types are expressed using OWL 2 object properties or RDF properties. Instances of fact types (facts) are expressed using triples with individuals in positions of subject or object and a property in a position of predicate.

For searching facts by SPARQL, graph pattern with variables in positions of subject or object (depending on what individuals we are searching for) should be used. Example of representing SBVR fact type and fact in SBVR SSE and RDF Turtle format is presented in TABLE I; query example – in TABLE II.

TABLE I. EXAMPLE OF REPRESENTING SBVR FACT TYPES AND FACTS

Language	Representation expressions
SBVR SSE	person1 *has_parent* person2 Ina_Griniute *has_parent* Jurgis_Grinius
RDF	:*has_parent* rdf:type owl:ObjectProperty rdfs:range :Person rdfs:domain :Person :Ina_Griniute :*has_parent* :Jurgis_Grinius

TABLE II. EXAMPLE OF QUERYING SBVR FACTS

Language	Query expressions
SBVR SSE	Who *is_parent_of* person Ina_Griniute?
SPARQL	SELECT ?person { :Ina_Griniute :has_parent ?person }

B. SBVR assortment fact type

SBVR assortment fact type is expressed using `rdf:type` property. Example of assortment fact type is presented in TABLE III, querying instances of a particular class – in TABLE IV.

TABLE III. EXAMPLE OF SBVR ASSORTMENT FACT TYPE

Language	Representation expressions
SBVR SSE	Ina_Griniute *is_a* person
RDF	:Ina_Griniute rdf:type :person

TABLE IV. EXAMPLE OF QUERYING ASSORTMENT FACT TYPE

Language	Query expressions
SBVR SSE	What subjects *are_of_type* person?
SPARQL	SELECT ?subjects { ?subjects rdf:type :person }

C. SBVR Specializations

For expressing that a particular concept is a category of more general concept, SBVR uses concept specialization, which corresponds to OWL 2 classes related by property `rdfs:subClassOf` stating that one class is a subclass of another class. We can use specializations for querying more general concepts. Also, we can access OWL 2 metaconcepts (classes, object and data properties, etc.) in the same manner as individuals and properties of individuals. For example, we can find subclasses of a particular class (TABLES V and VI). Similarly, specialization may be defined for fact types what corresponds to OWL 2 `SubObjectPropertyOf` or `SubDataPropertyOf`.

TABLE V. SBVR SPECIALIZATION FACT TYPE IN RDF

Language	Representation expressions
SBVR SSE	book *is_a* publication
RDF	:book rdfs:subClassOf :publication

TABLE VI. EXAMPLE OF QUERYING SBVR FACTS WITH SPARQL

Language	Query expressions
SBVR SSE	What class *is_sublass_of* publication?
SPARQL	SELECT ?class { ?class rdfs:subClassOf :publication }

D. SBVR Implication Formulations

SBVR implication formulations can be expressed using SWRL rule language. SWRL rules can infer new facts in OWL ontology. SPARQL queries can access all inferred

 ISBN: 978-1-61208-223-3

facts in a same way as asserted ones, using OWL 2 object properties or RDF properties.

E. SBVR Conjunction

Conjunction is a logical operation, which is true if each of its logical operands is true. It can be used to link SBVR fact types, propositions or logical formulations. In SBVR SSE language keyword "and" is used to indicate conjunction. In OWL 2, conjunction is represented by intersection of triples expressing axioms or restrictions for defining classes. Such definitions are used for inferring classes that could be directly accessed by SPARQL. In SPARQL queries, conjunction is used for finding facts satisfying conjunction condition. E.g. facts presented in TABLE VII are used in SPARQL query conjunction expression (TABLE VIII).

TABLE VII. EXAMPLE OF SBVR FACTS USED IN CONJUNCTION

Language	Representation expressions
SBVR SSE	Ina_Griniute *has_friend* Tadas_Gudas Ina_Griniute *has_kin* Jurgis_Titas
RDF	:Ina_Griniute :has_friend :Tadas_Gudas :Ina_Griniute :has_kin :Jurgis_Titas

TABLE VIII. EXAMPLE OF QUERYING SBVR FACTS USING CONJUNCTION

Language	Query expressions
SBVR SSE	What person *is_friend_of* Tadas_Gudas and *is_kin_of* Jurgis_Titas?
SPARQL	SELECT ?person { ?person :is_friend_of :Tadas_Gudas ; :is_kin_of :Jurgis_Titas }

F. Disjunction of Facts

Disjunction in SBVR is a logical operation, which is true if at least one of its logical operands is true. In SBVR SSE, it is indicated using keyword "or". Disjunction is useful when we want to find individuals having at least one of certain properties. For expressing disjunction in SPARQL, alternative graph patterns specified by UNION keyword can be used. As a result, all matching solutions are returned. Example of facts, used in alternative graph pattern, is presented in TABLE IX, example of query – in Table X.

TABLE IX. EXAMPLE OF SBVR FACTS, USED IN DISJUNCTION

Language	Representation expressions
SBVR SSE	Publication1 *has_author* Tomas_Gudas Publication2 *has_editor* Tomas_Gudas
RDF	:Publication1 :has_author :Tomas_Gudas :Publication2 :has_editor :Tomas_Gudas

TABLE X. EXAMPLE OF QUERYING SBVR FACTS USING DISJUNCTION

Language	Query expressions
SBVR SSE	What publications *has_author* Tomas_Gudas or *has_editor* Tomas_Gudas?
SPARQL	SELECT ?publications { {?publications :has_author :Tomas_Gudas} UNION {?publications :has_editor :Tomas_Gudas } }

G. Negation

Logical negation is a logical operation in SBVR having one logical operand. Negation, introduced in SPARQL 1.1, filters query solution by checking if query graph pattern does not match ontology. It is specified using keyword NOT EXISTS (TABLE XI).

TABLE XI. EXAMPLE OF QUERYING USING NEGATION

Language	Query expressions
SBVR SSE	What persons *are_author_of* publications that *are_not_cited_by* any publication?
SPARQL	SELECT ?persons { ?persons :is_author_of ?publication . FILTER NOT EXISTS {?publication :is_cited_by ?o}}

H. Synonyms and Synonymous Forms

SBVR synonyms and synonymous forms have the same meaning but different representation. They allow more flexible formulation of questions as synonyms are treated as equivalent classes or data properties, and synonymous forms are treated as equivalent object properties. Synonyms and synonymous forms in SBVR questions are transformed into SPARQL using preferred designations (i.e. in the same way as in SBVR2OWL2 transformation). Example of synonyms is presented in TABLE XII, example of querying using synonyms – in TABLE XIII (the query will find individuals of type "dictionary").

TABLE XII. EXAMPLE OF SBVR SYNONYM

Language	Representation expressions
SBVR SSE	vocabulary Synonym: dictionary Book1 *is_a* dictionary
RDF	:Book1 rdf:type :vocabulary

TABLE XIII. EXAMPLE OF QUERYING USING SYNONYM

Language	Query expressions
SBVR SSE	What books *are_of_type* dictionary?
SPARQL	SELECT ?books { ?books rdf:type :vocabulary }

Similarly, SBVR synonymous forms allow saying "person *is_author_of* publication" or "publication *is_published* by person". Example of using synonymous forms is presented in TABLE XIV and TABLE XV.

TABLE XIV. EXAMPLE OF SBVR SYNONYMOUS FORM

Language	Representation expressions
SBVR SSE	publication *is_published* by person Synonymous_form: person *is_author_of* publication Tadas_Grinius *is_author_of* Book1
RDF	:Book1 is_published_by :Tadas_Grinius

TABLE XV. EXAMPLE OF QUERYING USING SYNONYMOUS FORM

Language	Query expressions
SBVR SE	What person *is_author_of* Book1?
SPARQL	SELECT ?person { :Book1 :is_published_by ?person }

I. Predefined SBVR fact types

Some concepts and fact types having dedicated meaning are predefined in SBVR metamodel [1] (e.g. number1 *is_greater_than* number2, number1 *is_less_than* number2), SBVR extensions [4] (e.g. time interval1 *is_before* time interval2) or may be introduced into business domain vocabularies by users. Extensibility of SBVR allows conceptualizing terminology of specific domains for the way people think and communicate using natural language, and support computer understandable reasoning. E.g., we can introduce a predefined meaning for SBVR fact type *is_number_of* to express counting operation in SPARQL (Table XVI).

TABLE XVI. EXAMPLE OF QUERYING USING COUNT AGGREGATION

Language	Query expressions
SBVR SE	What *is_number_of* persons who is_child_of person Tadas_Grinius?
SPARQL	SELECT count(?persons) { ?persons :is_child_of :Tadas_Grinius}

V. CONCLUSIONS AND FURTHER WORKS

Transforming SBVR questions into SPARQL queries has shown that it is very important to synchronize SBVR to OWL 2 and SBVR to SPARQL mappings because it makes impact on correctness of SPARQL queries. Transforming SBVR vocabularies and rules into ontologies brings specific features into OWL 2 ontologies that could be rationally managed for obtaining more semantics from domain concepts.

Currently, main transformations from SBVR into SPARQL were realized and analysed; however, the efficiency of such transformations could be reached only by implementing the overall framework relating business vocabularies, rules, questions and executable business models. Our future work is streamlined towards elaborating the framework components, applying them in the real business environment, and, possibly, relating with linguistic models for connecting a really natural language and business software systems.

REFERENCES

[1] OMG. Semantics of Business Vocabulary and Business Rules (SBVR). Version 1.0. December, 2008, OMG Document Number: formal/2008-01-02.

[2] B. Motik, P. F. Patel-Schneider, B. Parsia, OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax. W3C Proposed Recommendation 22 September 2009.

[3] A. Marinos, P. Krause, "An SBVR Framework for RESTful Web Applications", Rule Interchange and Applications, Springer Berlin/Heidelberg, Germany, LNCS, Vol. 5858, 2009, pp. 144–158.

[4] OMG. Date-Time Vocabulary (Date-Time). OMG Document Number: bmi/2011-08-01.

[5] K. R. Bouzidi, C. Faron-Zucker, B. Fies, and N. L. Thanh, "An Ontological Approach for Modeling Technical Standards for Compliance Checking", Web Reasoning and Rule Systems, LNCS, Vol. 6902, Springer-Verlag Berlin, Heidelberg, 2011, pp. 244–249.

[6] W3C. SPARQL 1.1 Query Language. W3C Working Draft 14 October 2010.

[7] A. Sukys, L. Nemuraite, B. Paradauskas, and E. Sinkevicius, "Querying ontologies on the base of semantics of business vocabularies and business rules", Information Technologies' 2011: proceedings of the 17th international conference on Information and Software Technologies, IT 2011, Kaunas, Lithuania, April 27-29, 2011, pp. 247–254.

[8] A. Sukys, L. Nemuraite, B. Paradauskas, and E. Sinkevičius, "SBVR based representation of SPARQL queries and SWRL rules for analyzing semantic relations", the First International Conference on Business Intelligence and Technology, (Bustech 2011), Sep 25–30, Rome, Italy, IARIA, 2011, pp. 1–6.

[9] M. Kleiner, P. Albert, J. Bézivin, "Parsing SBVR-Based Controlled Languages", Model Driven Engineering Languages and Systems, LNCS, Vol. 5795, 2009, pp. 122–136.

[10] OMG. Ontology definition metamodel. OMG Document Number: ptc/2008-09-07.

[11] J. Karpovic, L. Nemuraite, "Transforming SBVR business semantics into Web ontology language OWL2: main concepts", Information Technologies' 2011 : proceedings of the 17th international conference on Information and Software Technologies, IT 2011, Kaunas, Lithuania, April 27–29, 2011 pp. 231–238.

[12] Collibra. Data governance software. http://www.collibra.com/ [Accessed 10 Mar 2012]

[13] Ontorule project. Ontologies meet business rules. http://ontorule-project.eu/ [Accessed 10 Mar 2012]

[14] I. Horrocks, B. Parsia, U. Sattle, OWL 2 Web Ontology Language Direct Semantics. W3C Recommendation 27 October 2009.

[15] J. Carrol, I. Herman, P. F. Patel-Schneider, OWL 2 Web Ontology Language RDF-Based Semantics. W3C Recommendation 27 October 2009.

[16] M. Schneider, "SPARQLAS – Implementing SPARQL Queries with OWL Syntax", Proceedings of the 3rd Workshop on Transforming and Weaving Ontologies in Model Driven Engineering, Málaga, Spain, Jun 30, 2010, pp. 1–7.

[17] G. Hillairet, F. Bertrand, and J. Y. Lafaye, "Rewriting Queries by Means of Model Transformations from SPARQL to OQL and Vice–Versa", ICMT '09 Proceedings of the 2nd International Conference on Theory and Practice of Model Transformations", Springer-Verlag Berlin, Heidelberg, 2009, pp. 116–131.

[18] M. Schneider, "Reasoning in expressive extensions of the RDF semantics", ESWC'11 Proceedings of the 8th Extended Semantic Web Conference on the Semanic Web: Research and Applications", Springer-Verlag Berlin, Heidelberg, LNCS, Vol. 6644, 2011, pp. 487–491.

[19] W3C. OWL 2 Web Ontology Language Mapping to RDF Graphs. W3C Recommendation 27 October 2009.

 ISBN: 978-1-61208-223-3

Context-Based Adaptation of Process Definition

Vojtěch Mates
Department of Information Systems
BUT, Faculty of Information Technology
Brno, Czech Republic
e-mail: imates@fit.vutbr.cz

Tomáš Hruška
IT4Innovations Centre of Excellence
BUT, Faculty of Information Technology
Brno, Czech Republic
e-mail: hruska@fit.vutbr.cz

***Abstract*—There are several ways to reach a goal. The right way often depends on the context of the process that produces the desired product. The context can be understood as a set of conditions including changing requirements and resource properties. The method described in this paper is a possible way of dealing with a changing context of a process. Our solution uses a product dependency tree and a database of processes. The dynamic profiles of the process are maintained by a workflow management system that produces a valuable data stream that records past behaviour.**

Keywords – business process management; business process intelligence; workflow management system; business process optimizing

I. INTRODUCTION

There are many events that can significantly influence the result of a process. Unexpected events are caused by changes in the environment of the process, which are not covered by business rules. Thus, it is also necessary to deal with the context of a process. Since the environment of a process is variable, it is usually necessary to change the logic of the process. This is usually done by replacing one part of the process by another one that produces adequate results. For example, if the input resources of a subprocess are temporarily unavailable, a need will arise to change the logic structure, e.g., by outsourcing the particular part of the process. The process execution is also limited by certain constraints such as costs, time, or success rate. Thus, the logic of the process should also be adapted for possible changes in these attributes that have an influence on the choice of variants of the process.

Our solution uses a workflow management system to collect data from previous process performances. At the same time, dynamic profiles of the processes are created from the collected data. The workflow management system thus works as a measurement tool.

There are more approaches for adaptive managing systems. The challenge is to provide flexibility and offer process support at the same time [19]. Another presented solution is the combination rule-based, case-based, and agent-based approach [20]. Our approach generates a structured process definition which can adapt to process context. Furthermore, the internal context variables are updated in runtime due to an existing workflow management system. Our approach is a hybrid between a declarative approach, rule-based approach, and a goal-based approach with the ability to monitor itself and use the information to optimize the process performance by the user's preferences, as defined by the users query.

The approach is also a combination of data-driven [14][15] and control flow approaches. It merges the subprocess definitions by using interfaces previously defined based on product compatibility. The algorithm uses dynamic profiles which describe the historical behaviour of processes (e.g., time, success rate, costs) to choose the optimal way to reach the goal of a process. To create the dynamic profiles of processes would be almost impossible without automatic tracking of processes by the workflow management system.

The next important step is to build a product dependency tree. Its purpose is to map all possible ways to create the desired product. A route in the product dependency tree is used to build the final structure of a modified process. The product dependency tree represents a product hierarchy. In other words, it reflects which income products are needed to produce the desired product.

Our method has several advantages. It allows switching over to another way of producing the product according to changing conditions. This can be done due to the declarative approach. With this approach, it is necessary to define the final product and a set of conditions, which are related to the context of the process. By using databases of processes that include process definitions and dynamic characteristics, it is possible to build new processes according to specific conditions. Thus, the process can be modified according to the current context, e.g., costs, time, success rate, resource problems, etc. This is a great advantage because information about the context of a process is often not available before runtime.

Product substitutability is what makes our solution so flexible. The substitutability means that within a group of products needed to create a final product, one product can replace another product so that the target product characteristics are not changed.

The approach is a hybrid of a rule-based and a goal-based approach. It deals with contexts using a top-down declarative approach using context variables. The process definition uses a structured definition of subprocesses and combines them according to the current context. The basic idea can be compared to a real-time decisions process because managers compare the possible ways of reaching certain goals using the currently available process. If something important happens, then the process has to be rescheduled.

 ISBN: 978-1-61208-223-3

A. Example

The desired final product is lunch. There are several ways to get lunch, e.g., going to a restaurant, making lunch at home, or hiring a cook. Each of these options (like each product) is desired to perform a certain subprocess. For example, to have lunch in a restaurant requires making a reservation and going to the restaurant. To make lunch at home requires buying the ingredients at a shop, and preparing the meal. Under certain circumstances, lunch in a restaurant and lunch at home, including the subprocesses connected with both possibilities of getting lunch, can be substitutable. The suitability of substitution depends on the context of the process of creating the final product (lunch) because the cheapest or the fastest solution is not always the best for each case. Current location should be considered as one of the factors in the choice of the optimal solution because it influences the price and time depending on transport conditions. The process of shopping can also involve some subprocesses. For example, the ingredients can be bought at different shops that have different prices. If the optimal way according to the current context is not applicable, the algorithm chooses the second best option. The behaviour can be compared to computer networks routing algorithms, especially OSPF (Open Shortest Path First) algorithm. If only the structure based approach is used then all the possible combinations of variables have to be defined. If a variant is added (e.g. buying ingredients to deliver to home), it is necessary to redefine all related processes which use the subprocess and make an evaluation of the processes using the changed subprocesser. The process can have a different priority, thus it is not possible to simply change the subprocess in each related process.

In case management, it is possible to adapt a process according to the current context, but each kind of change has to be defined, for example for price changes, each supplier can use a different order process. Overloading due to continuous change can eliminate the advantage of case-based decisions.

This paper aims to show how to optimize the process definition by using data captured by the workflow management system for managing dynamic profiles of processes, and how to use the declarative skill for context-based online adaptation.

The rest of the paper is organized as follows. Section II provides an introduction to the Top-Down Declarative approach, Section III overviews the workflow management system used for creating dynamic profiles. In Section IV, the merging of processes using Petri nets is briefly presented. Section V introduces a query-based approach using the declaration of final products. Section VI covers prerequisites of the optimizing method based on The Product Dependency Tree. Section VII describes the database of dynamic profiles of processes used for optimal decision in The Product Dependency Tree. Section VIII describes the approach to managing dynamic profiles. Section IX introduces The Product Dependency Tree, section X explains searching in The Product Dependency Tree. Section XI overviews experiments related to rescheduling decisions comparing an algorithm to human decision. Section XII summarizes the paper.

II. Top-down Declararitve Approach to Business Process Management

The declarative approach to process definition is quite similar to real-life planning. It is necessary to set the goal of a project first and then to find a way to reach the goal. Thus, the final goal generates other goals. The approach is also frequently used in the field of production. The process of production of the final product is decomposed into several subprocesses with the ordering of material being the basic subprocess. The approach is usually used for the calculation of costs. Different strategies can be chosen to create a product – all the processes needed to produce the product can be performed within a company or certain processes or their parts can be outsourced. The declarative approach is useful for runtime adaptation because a process can be defined as a set of conditions related to a changing environment and the final structure of the process is created according to the conditions in a particular context. The new process is feasible as its subprocesses generating interchangeable products can be modified.

The common declarative approach is described in [12][13][16][17]. However, our approach is focused more on hierarchy (hierarchy of goals). The approach is similar to database queries. The generated process can be launched in a rule-based workflow management system. The query approach is useful as an analytic tool.

III. Architecture of the Worklow Management System

The workflow management system can manage the cooperation between the workflow management system participants (including the interaction between people and software) in order to perform a business process.

The main purpose of the workflow management system lies in managing resources according to the process definition, which is usually set by the company management. Making changes in a business process is easier because the process definition is not hard-coded into the system and the company is therefore more adaptable to changing conditions. The workflow management system also allows easier business process re-engineering, as well as provides valuable audit data (in addition to managing and monitoring of the processes) which can later be used for analysis. Information extracted from the audit data can be also used for better managing as will be illustrated further in the paper. The architecture of the workflow management system will be described very briefly using the reference model created by the Workflow Management Coalition.

The workflow management system should consist of the Process Definition Tool, Workflow management engine, Work List Handler, and User Interface [5]. Creating a process definition is the first part of the process deployment in the workflow management system. This is usually done by the Process Definition Tool. A process described by the WFDL (Workflow Definition Language)

 ISBN: 978-1-61208-223-3

[10] is an output of this tool. A process definition should include information about tasks such as who can perform individual tasks (mostly the role-based approach), and information about routing between the tasks. The business process execution logic can be described as routing that will be dealt with in more detail in the following text.

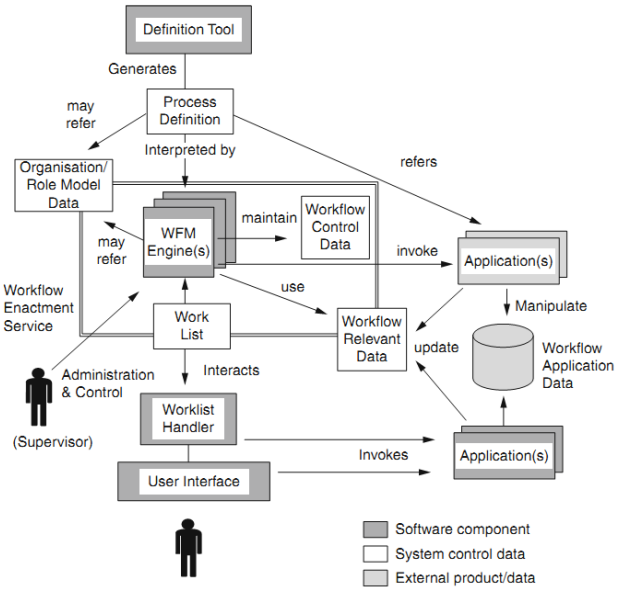


Figure 1. Architecture of the workflow management system [4]

Instances, i.e. particular processes, originate from the process definition. The process definition is transported to the workflow management engine where it can be launched as a process instance. Work lists are created by running particular instances according to the process definitions with specific data. For example, during the execution, applications or people can update a workflow's relevant data. Work list Handlers then assign tasks to resources, e.g., people or software (mostly web services) after launching a process instance. A workflow management system participant is an object in the workflow management system taking part in a process instance (usually people or web service). The user communicates with the system via several user interfaces like email, application interfaces, etc. All information about the performances of particular processes can be stored for monitoring and analysis purposes.

The workflow management system produces data that can be used as input data to dynamic profiles of certain parts of the process definition.

IV. MODELLING A WORKFLOW USING PETRI NETS

The Petri nets [3] formalism is widely used for modelling the dynamic aspects in systems. Petri nets consist of places, transitions, and arcs. Places represent conditions in a workflow. Transitions represent the performances of tasks. Oriented arcs connect places with transitions. A process performance depending on certain conditions can be simulated by moving tokens from one place to another place within a Petri net. Specific patterns in Petri nets are used to model a business process such as: AND split, AND-join, OR-split, OR-join, and other patterns [11].

The High-level Petri net Formalism also supports colour (passing several attributes in a token), hierarchy, and time (time stamps). Colour extension enables tokens to hold several attributes of a certain type. Hierarchy allows the building of a process from a subprocesses. Time extension can be used for time related conditions [3]. Modelling business processes can be a very complex issue [1][6][9]. The mathematical basis for Petri nets is suitable to demonstrate the algorithm that makes several Petri nets merge into one Petri net which will be the final process definition. The dynamic change of Petri nets [2], the correctness criteria [18] has to be considered in merging Petri nets. Our approach uses the previously mentioned interface-based process merging which eliminates many problems by correcting the merged process.

V. CUSTOM-MADE PROCESS BASED ON THE FINAL PRODUCTS

The aim of this task is to generate a complete process structure by specifying the final products of the process, which are desired to be created. It is also possible to set some specific additional constraints that reduce the number of the feasible solutions. When a query about the final product is made, the tool tries to find the optimal solution to the process definition according to the launched query.

A. Main Advantages

- It is easy to use. The tool only needs to have a set of desired products and constraints
- As the process definition can be built during runtime, it is possible to get a better solution because more up-to-date information is available
- The process can be custom-made so that it fits a particular purpose
- It is possible to modify the process structure according to temporary constraints which are only known in runtime (e.g., it is not possible to launch a particular task when the process is running, which may be caused by missing resources or a limited amount of time)

The problem can be defined as how to create the optimal process structure while keeping the process logic (the process has to produce the same result after its transformation).

Let us analyse what can be changed within the process definition. Is it possible to make changes in the business logic of a process that will lead to automatic product creation? The answer depends on the particular definition of the process. If a part of the process is changed randomly, the logic of the process will be destroyed. What can be changed automatically to avoid destroying the general logic?

Which two processes can be substituted with each other? It is possible to replace the original process with a process that produces results, which can substitute for the results of the original process. A product does not necessarily have to be a physical object, but it can also be information.

 ISBN: 978-1-61208-223-3

A process can be described as a stream of changing products, which leads to the final product. An input product can be a result of some agreement process, too.

In order to make any queries, it is necessary to have a database containing sufficient information.

The approach of following the product value chain is typically used in value based management.

VI. PREREQUISITES OF THE OPTIMIZING METHOD

In order to keep the logic of a whole process, it is necessary to keep the transformation of the process structure according to certain constraints. A requirement of the optimizing method is that each process has to produce at least one product. This is not an unusual requirement because processes are primarily optimized to produce a specific product. The product created by the process can also be information, e.g., where to find a shop offering the lowest price. A process usually contains several subprocesses. The process often needs some other products to run. There are several ways to obtain a certain kind of a product because in reality, there is a possibility of obtaining a substitute for the desired product. Nevertheless, each way of obtaining the substitute of the desired product can be different in certain properties. Examples of such properties are the time needed to get the product, costs paid to get the product, need for other resources, and special requirements for the product.

It also has to be mentioned that there are several ways to make the same product in a process. It is possible to demonstrate the situation by means of a simple example. A new table is desired. There are several options to get a table. It can be bought at a shop, or someone can be hired to get it. Another option is to make a table according to a table construction scheme. The first option will probably be more cost-effective than the second, but it will also be more time-consuming as some time has to be spent on choosing a table and comparing prices before the ideal table is found. The second option is the least time-consuming but someone (e.g., an interior decorator) has to be paid to find the best table according to the requirements. The third option is favourable in that no concrete product has to be found because it can be made according to the requirements. However, some wood is needed as well as tools to work with wood, some nails, etc.

There is no general answer to the question of which option is best. It depends on the context of performing the process called *get the table*.

The first option is probably the most cost-effective provided there is nobody specialized in making tables to produce one. A disadvantage of this option is that some time has to be spent in order to save some money. The second option is probably the fastest if the product has been specified but it is necessary to pay for searching, table transfer, etc. The third option is the most cost-effective if the process is launched in a company specialized in making tables because it already has equipment for making tables. Otherwise, buying all equipment and material for making one single table would cost more than a joiner would be paid to make a table. The process *get a table* can be a part of a more complicated process, for example the process of arranging a new house or business office. Therefore, one option is preferably chosen according to the context of the process performance.

The question arises which constraints are supposed to be the most suitable for the iterative composition of the parts of the process? Let us suppose that there is a database of all possible ways to make some products. The database also contains information about the properties of each of the ways.

A. Examples of the Properties

- time to accomplish the process
- process costs
- success rate of the process

These properties are changing according to the actual performance of the process. The time to accomplish the process can be changed as a result of a redesign of the process or changing resources. The cost of a product is calculated as the sum of fixed and variable costs while the part of the variable costs also depends on the working time of a particular worker. Reliability can also be captured as the success rate of the process. It is almost impossible for the staff to manage every single part of the database of processes but it is possible to use the workflow management system to do all the work.

VII. DATABASE OF THE DYNAMIC PROFILES OF PROCESSES

The database of processes has to be provided with some specific additional information that is required for an automatic change in the composition of the original process definition. As the process definition is continuously updated, the optimal solution may be changed from the previous version during runtime. The costs and time variables can be changed due to better training or equipment. There are many possible changes in the process factors that consequently influence the optimal solution, such as changes in the inner structure of the process, constant changes in the process characteristics (time, costs, success rate, etc.), choice of a different alternative how to make the same product, updated requirements on the final products, etc. It is also very useful to have several scenarios prepared that can be used for a simulation of potential evolution of the system, for example an increasing amount of products dependent on an increasing demand.

A. Content of the Database of Processes

- The process definition, which was designed for producing the desired product
- A set of places in the process definition is used for the purpose of merging several processes into one
- A set of input products/events is important for finding other dependent processes
- A set of output products is necessary for finding processes, which are using products as an input
- Runtime details such as the performance time, costs (calculated from fixed and variable costs depending on time), success rate, etc.

 ISBN: 978-1-61208-223-3

The process definition can be described by means of high-level Petri nets. The process definition contains a plan for performing tasks. The hierarchy of sub processes can be modelled by hierarchy extension in high-level Petri nets.

It is necessary to specify a set of places, which enable the process to be connected with other processes modelled by Petri nets. The set of places will be used for the synchronization of the output and input products.

A set of input products represents the prerequisites, which have to be satisfied as a condition for successful finishing of the process. For example, it is necessary to buy a table and then to take it home. *Buying the table* is an input event for *the transfer of the table*. An input product for the process called *transfer of the table* is the *purchased table*. An output product of the process *purchase of the table* is the *table*. The product *table* has to be used for the purposes of synchronization.

The runtime details represent a set of general properties which can be used for optimizing purposes, e.g., costs, time, success/failure rate. The workflow management system can be used for managing such kinds of information. It will be described in more detail in the next section of the text.

VIII. Managing the Dynamic Properties of Activities Using the Workflow Management System

To support the dynamic updating properties of the database of processes, it is possible to use the workflow management system, which can collect input data for processing to obtain the up-to-date profiles (sets of properties) of processes.

A. Static and Dynamic Properties of the Process Profile

In the case of the static properties, it is possible to manage them manually. The information can be updated by filling in some forms and can be managed by administrators responsible for the process performances. These properties do not rely on process performances. Examples of such kinds of properties are costs per time unit or fixed costs related to the process. In the case of dynamic properties, it is very difficult to keep the correct values up-to-date. Examples of dynamic properties are the time to accomplish the process, process costs, or reliability of the process. To manage the dynamic properties, updating of the specific values is triggered by events, which are created by the workflow management system. The events are happening when an instance of a particular process is being performed. Therefore, the dynamic profile of the process is updated continuously. However, the profiles could be extracted from execution logs. On the other hand, using triggers is more suitable from the performance point of view, because redesigning processes in runtime requires fast responses.

The structure holds a collection of items related to a particular process. Every item holds a set of monitoring indicators (time to accomplish the process, process costs, reliability of the process, etc.) for each process.

As can be understood, the dynamic properties describe the performance of a particular process. They can reflect changes in the processes performance. A combination of the static and dynamic properties can provide the calculation of the current process costs. Another important property is the time needed to perform the process. It can be directly calculated from the historical data captured by the workflow management system. However, to update the dynamic properties manually would be almost impossible without the automatic support of the workflow management system. The process of dealing with the dynamic properties is similar to [7]. Dynamic properties can be used in the decision process see [8].

The process profile should contain at least the time to accomplish the process, process costs, and reliability of the process. The time to accomplish the process can be calculated from the historical performance of the process.

The costs depend on the time and value of the fixed and variable costs for each task in the process. The reliability of the process can be found out by counting positive and negative results of the test in the checkpoint. Results closer to the present time certainly have higher significance than a result that was retrieved a longer time ago.

IX. Product Dependency Tree

Each activity has some inputs that are necessary for producing some products that can be used as input for other processes.

A. Phases of the Optimizing Process Structure

- It is necessary to build and keep an up-to-date database of processes
- The dependency tree based on the database of processes has to be built
- When the previous steps are accomplished, it is possible to find the optimal solution respecting the selected constraints
- The optimal solution is then to find all processes and their subprocesses within the dependency tree that are needed to create the final product. These are then merged into one final process that brings the same results as the original process.

The main idea of the algorithm is to connect the input and output events. Each event has to belong to a particular group. The events, which are placed in the same group, can be substituted with each other.

B. Types of the Processes

- Leaf process – the process does not depend on any other process, it has no input event that has to be produced by another process. The process only uses resources directly
- Root process – the process is derived from the final desired product
- Regular process – the process that has sub processes but it is not the root process at the same time

C. Steps of the Algorithm

- Select the desired products
- Find all processes that can be used to produce the products selected

 ISBN: 978-1-61208-223-3

- You will get a set of processes as the result. Search for all activities that can produce inputs for the previous set of processes. Exclude the processes that were used in the previous steps (prevention of cycling in the tree)
- Repeat until the activity is not a leaf activity

X. SEARCHING THE PROCESS DEFINITION IN THE DEPENDENCY TREE

The dependency tree covers all possible ways of getting the desired product. Each node of the tree represents one process. The relationship between nodes represents the relationship between the input and output products of processes. The nodes in the tree are synchronization points and contain additional data about the historical process performances. The algorithm follows the route from the root to leaf process. The route is marked according to the additional properties of the process. For example, there are two routes representing two options. One option is more cost-effective but slower, another is faster but more expensive. If the slow way meets the constraints, both the routes will be followed to lower levels until the leaf processes are reached. The higher level of the tree also covers related sub processes so that there are already aggregated values in the higher level of the tree due to the workflow management system. The aggregated values can be used in the decision process. If several solutions are available fitting the query, the solution has to be chosen that has the priority (e.g., costs, time, success rate, GPS of resources or priority set by the user, etc.) given by the query has to be chosen. The user query defines what the process should look like in order to be optimal, e.g., cost, time, quality.

A. Example of a query

The query can look as follows – select a process in which the desired product is *1 bread or 10 rolls*, the time to perform the process is less than 10 days, and the maximum costs are less than 2 euro. Choose the cheapest process. To accomplish this aim, these steps are to be followed:

- Focus on the root process, which is derived from the product (*1 bread or 10 rolls)*
- Select the routes that satisfy the constraints of time shorter than 10 days and costs not higher than 2
- The process must contain the subprocess with product "wheat"
- Order all the processes by costs.
- Choose the first process from the list
- Make the dynamic substitution of the substitutable parts in the process definition
- Launch the modified process by the workflow management engine

The query is similar to regular database queries, thus a SQL-like language will illustrate how it looks.

```
SELECT p /* a process variable name */
FROM processDatabase /* the source for the
query */
/* final product definition includes the
amount */
WHERE p.finalProducts IN (1 bread, 10 rolls)
/* constraits for context variable based on
historical performance of process p */
  AND (p.time <= 10 days)
  AND (p.cost <= 2)
/* defining constraints for subProduct which
the desired product should use */
  AND p.subProduct IN (wheat)
/* order the results which meet the
previously specified conditions according to
their costs */
ORDER BY p.cost
LIMIT 1./* limiting the number of results */.
```

The result of a query is a ready for launching process for getting the desired products. There are many context variables which could be used in a query, the variable can be user-defined and depend on the specialization of the process (e.g., data from sensors - temperature, humidity, etc.), but time, costs, location, and the role of person starting the process are general variables. If it is necessary to change the priority of a process then it should be done simply by changing the query. If the variant of making a substitutable product is chosen, the algorithm has to re-evaluate the process just before launching. The final process will be run-time generated according to the context values. Even if the number of variants of making a desired product is the same and the query remains the same, the historical performance data is still changing. Each change means checking the process if it is optimal according to the current situation.

XI. EXPERIMENTS

The currently implemented workflow management system will also support the adaptation of processes for changing conditions and is currently being implemented. Research has been made based on data describing changes in project scheduling. The retrieved data covers only 21 projects, however, many causes of rescheduling in the projects have been proven similar, and the resulting solutions have been similar, too. The outcome of the research is that almost all projects (over 80%) are rescheduled due to missing the deadlines of the sub processes of the projects or due to temporary unavailability of the resources. After detecting these kinds of problems, the solution is often to outsource some parts of the projects. The supplier is usually selected according to their success rate of meeting the deadlines, quality level, experience, and prices depending on the context and the priority of the projects. The integration of the adaptation into the workflow management system is therefore expected to reflect and react to natural changes in processes that occur in practice. The concept of the product dependency tree is currently being tested for the dynamic selection of the most suitable supplier. The concept of the dynamic profile is being tested for scheduling purposes and the data used for the testing include over one million records. The automatic selection behaves correctly (compatible with manager decision) in almost all projects (over 90%). In some projects (less than 10%), the decisions made by the algorithm were even better than decisions made by managers. The main reason was identified as the lack of time to collect and evaluate all information for decision. In some cases, the

 ISBN: 978-1-61208-223-3

algorithm chose the solution that was not optimal because of a changing variable that was not taken into account. In order to fix the problem, the general process has to be confirmed by the manager. However, the tool enabled fast analysis based on user queries, using historical performance data.

XII. CONCLUSION AND FUTURE WORK

The context of a process should be taken into consideration in process management. This article aims to describe the idea of the runtime context-based process definition adaptation based on replacing the substitutable parts of the process with its other parts. This paper shows how the original workflow management system is used for maintaining the dynamic process profiles which are used for making decisions in the product dependency tree. The extension can solve the problem of the automatic redesigning process according to constant changes in the context, which enables using advantages in rule-based and case-based approach. The context-based approach is very helpful in dynamic rescheduling in case of changing user priorities. For example, time is suddenly preferred more than costs or due to a problem with resources; the most optimal variant of the process cannot be performed. The solution also supports fault tolerant processes management and includes risk management into the common process management. The solution is suitable for business process reengineering because it can simulate and compare the old and new versions of a process. The changes in the process definition are briefly illustrated via high-level Petri nets that have good support for varying regions in the process definition. The dynamic process structure allows higher flexibility of the process management during runtime. An advantage of this approach is performing a process, which is accomplishable, and it is able to adapt to changing context. It can also manage flexibility at the same time. The case-based management system has the advantage of flexibility. The rule-based management system has the advantage of standardization of a process. The contribution at present is complex approach which provides the ability to make a dynamic change which is driven on by user queries and historical process performances.

ACKNOWLEDGMENT

This research was supported by the grants of MPO Czech Republic MPO TIP FR-TI3 039, by the grant FIT-S-10-2, by the research plan no. MSM0021630528 and by the IT4Innovations Centre of Excellence project CZ.1.05/1.1.00/02.0070.

REFERENCES

[1] W. M. P. van der Aalst and K. van Hee, “Workflow management Models, Methods, and Systems”, The MIT Press. Massachusetts 2004, ISBN 0-262-72046-9.

[2] C. Ellis, K. Keddara, and G. Rozenberg, “Change Within Workflow System”, http://portal.acm.org/citation.cfm?id=224021, retrieved: April, 2012.

[3] C. Girault and R. Valk: Petri “Nets for System Enginnering”, Springer; 1 edition 2002, ISBN-13: 978-354041217.

[4] D. Hollingsworth, “Workflow Management Coalition The Workflow Reference Model”, UnitedKingdom 1995, http://www.wfmc.org, retrieved: March, 2012.

[5] D. Hollingsworth, “WfMC The Workflow Reference Model 10 Years On”, United Kingdom2003, http://www.wfnc.org, retrieved: March, 2012.

[6] D. Marlon, W.M.P van der Aalst, and A. Hoftstede: “Process-aware information systems”, JOHN WILEY & SONS, INC, New Jersey 2005, ISBN: 978-0-471-66306-5.

[7] V. Mates, “Using Workflow Management System pro Analysis Based on Properties of Resources”, DATAKON 2010 Proceedings (Ed. Petr Šaloun), Mikulov, CZ, Ostravsk univerzita, Ostrava, 2010, pp. 161-168,

[8] V. Mates, “Modification of Workflow Management System Focused on Human”, Information Systems Development, , New York, US, Springer, 2011, ISBN 978-1-4419-9645-9, pp. 479-490.

[9] A. Sharp and P. McDermott: “Workflow Modeling: Tools for Process Improvement and Application Development”, Second Edition, ARTECH HOUSE, INC., MA 02062 USA 2009, ISBN-13: 978-1596931923.

[10] Workflow Management Coalition, “Workflow Management Coalition Terminology & Glossary”, United Kingdom 1999, http://www.wfmc.org, retrieved: March, 2012.

[11] Workflow Control-Flow Patterns : “A Revised View, BPM Center Report BPM-06-22”, http://www.workflowpatterns.com/documentation/documents/BPM-06-22.pdf, retrieved: April, 2012.

[12] I. Barba, B. Weber, and C. Valle, “Supporting the Optimized Execution of Business Processes through Recommendations“, Business Process Management Workshops, Bd. 99, F. Daniel, K. Barkaoui, S. Dustdar, W. M. P. van der Aalst, J. Mylopoulos, M. Rosemann, M. J. Shaw, and C. Szyperski, Hrsg. Springer Berlin Heidelberg, 2012, pp. 135–140.

[13] R. Hull, F. Llirbat, E. Siman, J. Su, G. Dong, B. Kumar, and G. Zhou, “Declarative workflows that support easy modification and dynamic browsing“, Int’l Joint Conference on Work activities coordination and collaboration, New York, NY, USA, 1999, pp. 69–78.

[14] V. Künzle and M. Reichert, “Towards Object-aware Process Management Systems: Issues, Challenges, Benefits“, Proc. 10th Int’l Workshop on Business Process Modeling, Development, and Support (BPMDS’09), Amsterdam, The Netherlands, 2009.

[15] V. Künzle and M. Reichert, “PHILharmonicFlows: towards a framework for object-aware process management“, Journal of Software Maintenance and Evolution: Research and Practice, Bd. 23, Nr. 4, 2011, pp. 205–244.

[16] M. Pesic, H. Schonenberg, N. Sidorova, and W. M. P. van der Aalst, “Constraint-Based Workflow Models: Change Made Easy.“, in OTM Conferences (1), 2007, pp. 77–94.

[17] M. Pesic, and W. M. P. van der Aalst, “A Declarative Approach for Flexible Business Processes Management“, in Business Process Management Workshops, Bd. 4103, Springer Berlin / Heidelberg, 2006, pp. 169–180.

[18] S. Rinderle, M. Reichert, and P. Dadam, “Correctness Criteria for Dynamic Changes in Workflow Systems – A Survey.“, Data & Knowl. Engineering, Bd. 50, Nr. 1, 2004, pp. 9–34.

[19] H. Schonenberg, R. Mans, N. Russell, N. Mulyar, and W. M. P. van der Aalst, “Process Flexibility: A Survey of Contemporary Approaches“, in Advances in Enterprise Engineering I, 2008, pp. 16–30.

[20] B. Weber, W. Werner, and R. Breu, “CCBR–Enabled Adaptive Workflow Management.“, in Proc. European Conf. on Case-Based Reasoning (ECCBR’04), Madrid, 2004, pp. 434–448

 ISBN: 978-1-61208-223-3

User as Innovator in IT development process

Malgorzata Pankowska
Information Systems Department
University of Economics
Katowice, Poland
e-mail: pank@ue.katowice.pl

***Abstract*—In this paper, the benefits of user participation in value creation in the information technology (IT) development process are discussed. The literature studies allow for the conclusion that user role is very important and new methods are developed to increase user participation. The research results indicate that the user participation, although requested, is limited, therefore in the paper some changes are suggested.**

Keywords-innovation; lead user; user centered design; user experience; user participation

I. INTRODUCTION

Encouraging users to be value creators in IT development processes is an important step in the competitive effectiveness increase activity. Many authors argue the user involvement in IT process is critical for IT implementation success because under the service-dominant logic users are contributing to the process of IT exploitation.

According to Claycomb et al. [12] users can actively participate in creating a solution, when IT service failure occurs by applying specialised skills and knowledge. For example a user can diagnose their laptop problems based on the product's user manual. Research on user participation so far has focused on how to employ users to increase productivity in the service delivery context [18]. The purpose of this research is to bridge the gap in the literature by investigating user participation in IT development process and to present some suggestions on users' future co-creation behaviours. The paper is to show that nowadays users are no longer passive audience, but under certain circumstances they are active co-producers. Williford [27] has noticed that for many years, the computer-literate users developed simple applications to increase their personal productivity. User applications development has evolved to include complex application development by groups of users and shared across departmental boundaries. However, there still prevails a common mistake that user applications are not significant, transient, disposable as not production-oriented.

In 2009, Anderson [2] had noticed that in the markets, there are a lot of niche goods, the costs of reaching them are going down because of the IT development. Anderson [2] argues that for long-term demand development is the democratizing the tools of production. So, individuals can do now what just a few years ago only professionals could do. He argues that now people have the tools and the methods to become amateur producers. Some of them have talent and vision, so they are able to produce by and for themselves.

The main problem of the paper is to explain the role of users as innovators in the business information system process. The paper consists of two main parts. The first covers analysis of the works from other researchers on user innovativeness and on user participation methods. The second includes considerations connected with empirical research on the user different activities, strong points and weaknesses of their behaviours. The further research could cover explanation of users attitudes differentiation according to the industry.

II. CUSTOMERS AS INNOVATORS

Innovation can be seen as a management activity, which involves focusing on the organization's mission, searching for unique opportunities, determining whether they fit the organization's strategic directions. Innovators look to the future with knowledge of the past. They are expected to deliver something new - new products, new processes and new designs, which create change in goods as well as in lifestyles. Roberts and Frohman [21] generalize that innovation is an invention plus the exploitation. The invention process covers all efforts aimed at creating new ideas and getting them to work. The exploitation process includes all stages of commercial development, application and transfer, including the focusing of ideas or inventions toward specific objectives, evaluating these objectives, downstream transfer of research and development results, and eventually a broad-based utilization, dissemination and diffusion of technology-based outcomes.

Tidd and Bessant [25] argue that innovation is driven by the ability to see connections and opportunities and to take advantage of them. The creativity in innovations adds value to the individual and the community and is based upon perceiving and capturing an opportunity. They consider four types of innovations:

- Product innovation - changes in the products or services that a business organization offers.
- Process innovation - changes in the ways in which they are created and delivered.
- Position innovation - changes in the context in which the products and services are introduced.

 ISBN: 978-1-61208-223-3

- Paradigm innovation - changes in the underlying mental models which frame what the business organization does.

The innovation process consists of an idea that comes from some recognized need that is developed into a concept, followed by invention and then taken through development, production, diffusion and adoption by end users. The innovation process will be guided by the type of innovation, the importance of innovation, the elapsed time and expected time of introduction, sources of innovation, the character of the organizational infrastructure, and the number of technology and market unknowns. There are many models of innovation process [19], [24], [25]. All of them emphasize the necessity to recognize the opportunity and through the research and design transform it into a market product.

IT innovation as any innovation demands the co-evolution and co-existence of the information technology invention and the business process of any organization. Innovations are realized in the context of socio-economic institutions. The object of IT innovation does not stand alone, but is set in the economy, cultural and business practices, social values and interests. Innovation research has emphasized the importance of understanding user needs in the process of new product development. However, it is not sufficient to understand or even satisfy existing customers, but rather it is necessary to lead existing customers and create new customer segments. Therefore, companies could be interested in customer involvement in production or service delivery process. Customers are expected to support companies to reduce research and development costs, to differentiate services, reduce the time to market, facilitate user education, improve market acceptance, or even to provide more original and valuable proposals than professional developers. General classification of customers in IT sector covers lead user, normal user and user community [3], [15]. Lead users are critical to the development and adoption of complex products. They demand innovations ahead of the general market of other users. They face needs that will be the future trends in the market place and they are actively engaged in the innovative process. Normal users could help to provide superior and differentiated services, reduce cycle time, increase acceptance of the new service. Community users are interested in technological products and they play important role in advance of technologies [23]. Communities are treated as a decentralized virtual design teams, as designers and as a source of innovations. Websites such as Twitter, Facebook, MetaCafe, Wikipedia, Flickr among others, have all been introduced within the last decade and rapidly grew in user communities. Organizations are beginning to invest time and effort in developing a social media presence e.g. on Facebook, to capitalize on a growing user population that is interested in creating, retrieving, and exploring the Websites. Organizations are beginning to realize the potential benefits that can be captured when users and organizations co-create value. Users benefit from their positive experience that fulfils personal needs and interests. Experience is defined as an intense individually involved event. Designing for use and testing by use are the essential characteristics of user-innovators; they may subcontract production and parts supply. The user-innovators are motivated by the users' own desires for a better product. However, there is a risk that the user involvement in the design process will never be properly rewarded and an organization fails to provide a positive user experience, so the negative consequences can occur such as negative publicity and loss of user engagement. Therefore, only through the social interactions, perceived dialogue online and social accessibility and transparency the value can be co-created. Some initiatives have institutional backing involving professional staff, others build on communities of practitioners and rely on their voluntary work. Repositories can be organised as a place to share and exchange resources, which means that people are either users or producers, or they can promote the collaborative production of common resources. There are a couple of models of open collaboration:

- User-producer model: centralized model, although real costs can be met with resources other than money, most initiatives need to raise some capital.
- Co-production model: equal participation,
- The replacement model, open content replaces other uses and benefits from cost savings.
- The foundation, donation or endowment model in which funding for the project is provided by an external actor.
- The segmentation model, in which the provider offers value-added services to user segments and charges them for these services.
- The conversion model, in which "you give something away for free and then convert the consumer to a paying customer".
- Membership model, voluntary support model, based on fund-raising campaigns or paying members [13].

In innovation process, customers are perceived as having an integral pro-active role in collaboration to innovate. Bhalla [8] discussed the changed profile of the new customer (see Table I).

TABLE I. CUSTOMER PROFILE CHANGES

	Classic Management Organization	**New Economics Organization**
Identity	***Consumers, Recipients***	***Creative Partners***
Role	Passive	Active Collaborators, Co-producers of Value
Source of Insights	Surveys, observation of customers	Conversations, online comments
Relations with company	Transaction-based	Experience-based
Location	Fixed and visible customers	Virtual customers
Influence media	Advertising, expert opinions	Social media, peer-to-peer
Source of value	Company offers	Customer preferences and experiences

Source: [8].

 ISBN: 978-1-61208-223-3

For many years mass production and product customization were business strategies that aimed at fulfilling individual user needs quickly and efficiently. However, concepts such as customer orientation, user centered design, close to the customer, customer segmentation, customer relationship management reveal the importance of the user involvement.

III. USER PARTICIPATION METHODS

Barki and Hartwick [4] proposed a distinction of user involvement from user participation. They define user participation as the assignments, activities and behaviours that users and their representatives perform during the system development process. User involvement refers to the subjective psychological state reflecting the importance and personal relevance that a user attaches to a given system. User participation is defined as the degree to which the customer is involved in producing and delivering the service. System designers have promoted techniques requiring user participation, such as prototyping, rapid application development and joint application design.

A user is a person using services on a day-to-day basis within the business. This means that IT staff need to be involved in the development and production of appropriate and relevant Service Level Requirements (SLRs) and Agreements (SLAs) that detail the business quality targets, together with required business functionality.

According to Cartlidge [11], an Informed Customer (IC) is a term that came into use in the late 1990s. to describe a customer with some information system (IS) perspective. Typically the areas of involvement of ICs are:

- The alignment of business and IT plans and strategies.
- The development of Business Unit objectives and requirements for IS.
- The establishment and co-ordination of user groups.
- The development, negotiation and agreement of SLRs and SLAs.
- Managing the provision of the IS services on behalf of their Business Unit.
- Shared risk and reward, e.g., agreeing how investment costs and resultant efficiency benefits are shared.

For example, in PRINCE2 project management methodology [6], [16], the Senior User represents the interests of the users, who will use the final products of the project, those for whom the product will achieve an objective, or those who will use the product to deliver benefits.

A. User Experience

The concept of user experience is understood as the subjective relationship between user and application. It goes beyond the usability of the application, focusing on the personal outcome that the user gets from interacting with the application while performing a task [10]. In Human Computer Interaction (HCI), the term of experience design is about considering the user, the task and the context when designing a computer application [9]. Usually, the projects have a large context that the users should understand and IT people should integrate into their planning. This context is the project's ecosystem and it includes the environment they are working within the company culture, the general type of work they all will be engaged and the people with whom they interact within their roles and responsibilities. According to Beccari and Oliveira [5] the User Experience orientation points to project goals, but not just to attain effectiveness, efficiency and satisfaction, but it aims to enhance the entire experience resulting from the use of a product, system or service.

B. User Centered Development Process

The purpose of User Centered Development (UCD) is to develop products with a high degree of usability. The user becomes the centre of focus in the product development process. UCD is defined as a user interface design process that focuses on usability goals, user characteristics, environment, tasks and workflows in the design of an interface [20], [7]. According to Goncalves and Santos [14] the UCD is a philosophy that is based on the needs and interests of user and requires an investigation of user activities, profile, environment and goals. The key common concepts in UCD are as follows:

- Focus early on users and tasks to understand user cognitive, behavioural and attitudinal characteristics.
- First design the user interface.
- Involving the users in design and design reviews.
- Insisting on iterative prototyping and evaluation [17].

Design methods include prototyping and participatory design. Among the evaluation methods, there are usability inspection methods, and user testing methods such as laboratory and field tests. Participatory design concerns the direct participation of the beneficiaries of the introduction of a computer application. Major issues considered in this approach cover the expertise regarding users' own work, sustainable innovation opportunities, multiple viewpoints and taking differences seriously as facts and resources, the linking of the work practices, technology, and work environment context [26].

IV. EMPIRICAL RESEARCH ON USER INVOLVEMENT

The literature studies created the need to empirically verify that very optimistic attitude of academic publications' authors towards user involvement. Therefore, the empirical research was done in 2011. The research covered the interviews with Chief Information Officers (CIOs) from 270 firms in Poland. Characteristics of the surveyed firms are presented in Table II. User involvement in this paper is considered as participation in the business information system development process measured as a set of activities that users have performed. In this research, CIOs answered the questions concerning the activities of users at their

 ISBN: 978-1-61208-223-3

companies. Historically, there are several kinds of adaptive methods of information system development that build a model of users' knowledge and their involvement in that process. Active participation of a person in a community is a powerful indicator of the person's interests, preferences, beliefs and social and demographic context. Community members are a part of users' model and can contribute to tasks like personalized services, assistance and recommendations.

TABLE II. SURVEYED COMPANIES FEATURES

Feature	N=270
Number of employees	
Micro Enterprises (1-9 employees)	44,4%
Small Enterprises (10-49 employees)	29,3%
Medium Enterprises (50-250 employees)	15,2 %
Big Companies (more than 250 employees)	11,1 %
Dominating Activities	
Production	9,3%
Commerce	22,6%
Services	50,4%
Mixture of above activities	17,8%
Main Clients	
Individual	61,1%
Institutional	38,9%
Scope of Activities	
Local market	27,8%
Regional market	23,7%
National market	35,6%
International market	7,4%
Global market	5,6%

Involvement of the end users in IT projects covering IS development is presented in Table III. In Table III, the following activities of users have been specified: goal specification and project concepts (GSPC), business logic analysis and business process modelling (BLA BPM), requirements engineering (RE), information system design (ISD), information system implementation (ISI), information system testing (IST), information system installation and migration to a new IT environment (ISE), information system maintenance (ISM), security of information system (SIS), information system usage (ISU).

TABLE III. PARTICIPATION OF USERS IN IT PROJECTS

	User					
	Passive	***Evaluator***	***Co-creator***	***Partner***	***Producer***	***Prosumer***
GS PC	15%	17%	**33%**	24%	10%	1%
BLA BPM	**32%**	22%	19%	20%	5%	1%
RE	**37%**	19%	17%	16%	10%	1%
ISD	**34%**	20%	19%	15%	10%	1%
ISI	**39%**	16%	19%	15%	9%	1%
IST	18%	20%	**24%**	21%	15%	2%
ISE	22%	**27%**	24%	14%	10%	2%
ISM	20%	23%	**24%**	19%	12%	3%
SIS	**33%**	18%	20%	14%	13%	1%
ISU	7%	23%	20%	**30%**	15%	4%

In Table III, six different profiles of users has been included. Passive users and users-evaluators are oriented towards the observation and acceptance of other people efforts. Co-creator supports IT staff in business information system development works. User as the partner plays equally important role as IT professional in the system development process. User as the producer is self-dependent and has got sufficient competencies to utilize IT independently of the IT staff help. The last, i.e., prosumers are able to utilize IT by themselves and for their work purposes. In this paper the definition of prosumption was adapted from the work of Xie et al. [28]. According to Xie et al., prosumption consists of individual and social acts by users in an exchange relationship that help to co-produce and gives rise to sociopsychological experiences for the buyers (i.e., users) in cooperation with the sellers (i.e., IT people) [28]. Taking into account the survey results, it should be noticed that only some percent of users were evaluated as prosumers. Generally, they are from SMEs sector and they are working on their Web portals and e-business system development. Prosumption implies that buyers produce products for their own consumption. Although most contribution in marketing today have been constructed with the view of consumers as passive buyers of what others produce, there is a perspective that it is rather limited approach and in e-business users have tools to produce Website for their own usage. Taking into account the results included in Table III you can notice that users are rather inactive. CIOs evaluate users as inactive at business analysis and business process modelling stages as well as at requirements engineering, system design and implementation. IT people do not demand the technical expertise from users, they should be helpful at the initial stages of business information system development process. Users were evaluated as co-creator in project concepts specification, information system testing and maintenance. Security of IS is the domain of IT professionals, and of course the strong activity of users is revealed at the business information system exploitation stage.

In the survey, CIOs were asked about the attitudes of the users towards changes of IT/IS. Eventually, CIOs admitted that users have got innovative ideas concerning information systems functionalities, but they do not solve the problems independently. Radical as well as evolutional changes of information systems functionalities are implemented with involvement of user as well as IT professionals. Users are not able to find the solution, solve the problem and create the software product independently. They prefer the product developed and maintained by IT professionals or in co-operation with them. CIOs were asked about the user support in IT product promotion process (see Table IV). According to CIOs, users are not involved in the distribution of information about IT product. Mostly, they distribute information about the product usage (see Table IV).

CIOs provided opinions on users' involvement in the IT product promotion process. The process was divided into three stages:

- Stage 1 (S1), before entrance of the product to the market.
- Stage 2 (S2), when the IT product is on the market.
- Stage 3 (S3), when the IT product is removed from the market.

 ISBN: 978-1-61208-223-3

TABLE IV. IT PRODUCT DISTRIBUTION SUPPORT

User supports distribution of the IT product information		
Support of IT product information on:	***Yes***	***No***
IT product design	39%	**61%**
IT product installation methods	41%	**59%**
IT product maintenance methods	41%	**59%**
IT product usage	**59%**	41%
The implemented product in comparison with other similar products	50%	50%

As it is presented in Table V users are not interested in promotion of IT products on the market. CIOs explain that the IT products are for business usage, and not for the meeting of the private needs, therefore the users are not involved and they do not identify themselves with the enterprise and its IT resources utilization strategies. CIOs admit that lead users are able to define the IT trends, but preferences of lead users and normal users are not the same. The normal users focus not only on new functionalities of IT products, but also take into account non-functional requirements, i.e., cost, processing time, scalability, interoperability, compatibility, reliability and security of IT products, and technical support.

TABLE V. USER INVOLVEMENT IN IT PRODUCT PROMOTION

	User involvement in IT product promotion				
	Very frequently	***Frequently***	***Rarely***	***Very rarely***	***Never***
S1	11%	17%	24%	17%	**30%**
S2	18%	**24%**	21%	13%	23%
S3	4%	11%	18%	12%	**55%**

CIOs have noticed the development of user communities, but that process is still at the beginning stage and the final results will be available in a few years (see Table VI).

TABLE VI. USER COMMUNITY DEVELOPMENT

User Communities Development		
Type of online community	***Percent of reviewing companies***	
	Yes	***No***
Community created by the enterprise management	35%	**65%**
Community created by the enterprise users	31%	**69%**
Community created by the similar product proponents	22%	**78%**

The user involvement in the IT product development was evaluated through the questions concerning the user participation in IT research (see Table VII).

CIOs argue that users are involved in the IT research. They are less engaged in the basic research, but their efforts are highly appreciated in the developmental works, implementation and exploitation processes. For example, users are encouraged to take part in the Operational Programme Innovative Economy projects for SMEs, where they have the opportunities to use the knowledge and expertise for new IT products implementation and exploitation.

TABLE VII. USER PARTICIPATION IN RESEARCH

User Participation in Research		
Type of research work	***Percent of reviewing companies***	
	Yes	***No***
Technology development research	25%	**75**%
Application development research	39%	**61**%
Developmental works	49%	**51**%
Pilot implementations	**66**%	34%
Commercial product exploitation and maintenance	**69**%	31%

V. DISCUSSION AND FURTHER RESEARCH

The approach concerning the user as an innovator has led to a new innovation paradigm, in which customers play a central and very active role. Rather than seeing users as the market, the new paradigms and IT product development methods focus on users involvement. According to Xie et al. the realization that users are actively involved in creating value and benefits for their own consumption is aligned with the post-modernist view that the user is a participant in the IT development process. The user involvement is important because of the necessity to develop end user computing (EUC) and reduce the risks of EUC [22]. The risks associated with the end user computing include: weak security, inefficient use of resources, inadequate training, inadequate support, incompatible and redundant systems, ineffective implementations, copyright violations, the destruction of information by computer viruses, unauthorized access or changes to data and programs, unauthorized remote access, reliance on inaccurate information [1].

In the aspect of end users, their involvement in the information system development process results from the opportunities to openly criticize insufficient ICTs solutions provided to the customers i.e., users. Users have got their own personal satisfaction and have possibilities to learn new products. They want to be followers of their competitors and business partners, if only the IT solutions implementations have provided financial benefits. Users can learn new products in the production process and through this "learning by doing" can reduce the cost of trainings. The interviewed CIOs share the opinions that in the co-development process, users benefit because:

- Their own ideas enrich the software products functionalities as well as non-functional characteristics of the products.
- Software firms encourage users to be involved in the production process and pay them.
- The software implementation costs are reduced.

 ISBN: 978-1-61208-223-3

Therefore, the further research of users attitudes seems to be important.

VI. CONCLUSION

Perhaps the research results are not very impressive in comparison with the literature reviews and other authors' studies, which include a more optimistic vision of end user involvement in IT implementation and exploitation processes. The references cover works where qualitative research and theoretical considerations are included. This statistical survey presents that users are not strongly involved and only spectacular cases describe an innovative approach of users. According to this research results, users are afraid of novelty and risks and they prefer to rely on IT people knowledge and preferences.

However, the research results reveal generally positive opinions about user participation in business information system development projects. Probably, there is a need to create the culture of end user involvement and participation in the social and business organizations to encourage users to IT development activities. Within an organization there are various groups of interests that have their own cultures. Across the entire company there is often a corporate culture that is the collected set of values, traditions or other elements that characterize the company. Therefore, the cultures should develop generally positive attitudes towards users' innovativeness.

ACKNOWLEDGMENT

The work is funded by the National Science Centre in Poland. The grant number is 4100/B/H03/2011/40.

REFERENCES

[1] S.D. Allen-Senft and F. Gallegos, "Managing Risks in User Computing," in IS Management Handbook, 8th Edition, C.V. Brown, H.Topi, Eds. Auerbach Publications, London, 2003, pp. 771-780

[2] Ch. Anderson, The longer Long Tail, How Endless Choice is Creating Unlimited Demand. Business Books, London, 2009.

[3] C. Baldwin, Ch. Hienerth and E. von Hippel, "How user innovations become commercial products: a theoretical investigation and case study," Research Policy, 35(9), 2006, pp.1291-1313.

[4] H. Barki and J. Hartwick, "Rethinking the concept of user involvement, and user attitude," MIS Quarterly, 18(1), 1994, pp. 59-79.

[5] M.N. Beccari and T.L. Oliveira, "A Philosophical Approach about User Experience Methodology," in Design, User Experience and Usability, Theory, Methods, Tools and Practice, A.Marcus, Ed. Springer, Heidelberg, 2011, pp. 13-22.

[6] C. Bentley, PRINCE2, A Practical Handbook. Butterworth-Heinemann, Amsterdam, 2010.

[7] H. Beyer, User-Centered Agile Methods. Morgan & Claypool Publishers, Boston, 2010.

[8] G. Bhalla, Collaboration and Co-creation, New Platforms for Marketing and Innovation. Springer, Berlin, 2011.

[9] B. Buxton, Sketching User Experiences. Morgan Kaufmann, San Francisco, CA, 2007.

[10] E.H. Calvillo-Gamez, P. Cairns and A.L. Cox, "Assessing the Core Elements of the Gaming Experience, " in Evaluating User Experince in Games, Concepts and Methods, R. Bernhaupt, Ed. Springer, London, 2010, pp. 47-72.

[11] A. Cartlidge, Best Practice for Business Perspective: The IS View on Delivering Services to the Business, ITIL, the key to Managing IT services. TSO, Norwich, 2004.

[12] C. Claycomb, C.A. Lengnick-Hall and L.W. Inks "The customer as a productive resource: A pilot study and strategic implications," Journal of Business Strategies, 18(1), 2001, pp. 47-69.

[13] Giving Knowledge for Free, The emergence of open education resources, Centre for educational reserach and innovation, Organisation for economic co-operation and development, OECD, Paris, 2007. http://www.oecd.org/document/41/0,3343,en_2649_35845581_38659497_1_1_1_1,00.html, retrieved: May 2011.

[14] J. Goncalves and C. Santos, "POLVO- Software for Prototyping of Low-Fidelity Interfaces in Agile Development," in Human-Computer Interaction, Design and Development Approaches, J.A.Jacko, Ed. Springer-Verlag Berlin, Heildeberg, 2011, pp. 63-71.

[15] J. Matthing, B. Sandem and B. Edvardsson, "New service development: learning from and with customers," International Journal of Service Industry Management, 15(5), 2004, pp. 479-498.

[16] M. Pieper and J. van Bon, Project Management Based on PRINCE2. Van Haren Publishing, Zaltbommel, 2005

[17] H. Pradeep, User-Centered Information Design for Improved Software Usability. Artech House, Inc. London, 2004.

[18] C.K. Prahalad and V. Ramaswamy, "Co-opting customer competence", Harvard Business Review, 78(1), 2000, pp. 79-87.

[19] J.B. Quinn "Managing Innovation: Controlled Chaos," Harvard Business Review, May-June 1995, pp. 73-84.

[20] C. Righi and J.James, User-Centered Design Stories. Elsevier, Amsterdam, 2007.

[21] E.B. Roberts and A.L. Frohman "Strategies for Improving Research Utilization," Technology Review, March-April, 1997, pp. 33-36.

[22] S. Senft and F. Gallegos, Information Technology Control and Audit, CRC Press, Taylor and Francis Group, London, 2009

[23] D. Tapscott and A.D. Williams, Wikinomics: How Mass Collaboration Changes Everything. Penguin, New York 2007.

[24] J. Tidd, J. Bessant and K.Pavitt, Managing innovation, Integrating technological, market and organizational change, 3rd Edition. J.Wiley & Sons, Chichester, 2005.

[25] J. Tidd and J. Bessant, Managing innovation, integrating technological, market and organizational change. J.Wiley and Sons, Ltd, Chichester, 2009.

[26] B. Torpel, A.Voss, M. Hartswood and R. Procter, "Participatory Design: Issues and Approaches in Dynamic Constallations of Use, Design and Research," in Configuring User-Designer Relations, A. Voss, M. Hartswood, R. Procter, M. Rouncefield, R.S. Slack, M. Buscher, Eds. Springer Verlag London, 2009, pp. 13-30.

[27] S.M. Williford "Reviewing User-Developed Applications," in IS Management Handbook, 8th Edition, C.V.Brown and H. Topi, Eds. Auerbach Publications, London, 2003, pp. 781-798.

[28] Ch. Xie, R.P. Bagozzi and S.V. Troye, "Trying to prosume: toward a theory of consumers as co-creators of value," Journal of the Acad.Mark.Sci (2008) 36, 2008, pp. 109-122.

 ISBN: 978-1-61208-223-3

Mobile Context-Aware Product Assistant

Antti Nummiaho, Timo Laakko
VTT Technical Research Centre of Finland
Vuorimiehentie 5, P.O.Box 2000, FIN-02044 VTT, Finland
antti.nummiaho@vtt.fi, timo.laakko@vtt.fi

***Abstract*—In this paper we present an application that manufacturers and retailers can use to provide contextual product information for their customers. The application also provides a direct interactive feedback channel that the customers can use in asking questions about the products from the manufacturers and retailers. In addition to that the application also offers functionality for the customers to discuss about the products amongst themselves. The application consists of a web interface for the manufacturers and retailers and both a web interface and a mobile Java application for the customers. We conclude that this kind of an application has clear benefits for both the manufacturers and retailers as well as for the customers. The customers get guidance and assistance in using the products and the manufacturers and retailers get valuable information that they can use in their product development. We also discuss some potential difficulties that the concept might face.**

***Keywords*-context-aware messaging; context-aware application; mobile application; product guidance; product assistance; mCRM; mobile customer relationship management**

I. INTRODUCTION

It is all too often the case that the customer gets all sorts of guidance when buying a product, but after the purchase is made, getting help and assistance in using the product turns out to be a lot more difficult. To tackle this problem we have designed and implemented a tool that enables manufacturers and retailers to provide product information and guidance for the customers in the situation where the customers actually need it. The tool provides an interactive feedback channel that the customers can utilize in communicating with the manufacturers and retailers. This not only improves the customers' satisfaction but also generates valuable information on the problems and issues that the customers have in using the product, which can then be used in further product development.

The importance of context-aware user guidance was acknowledged already in the 1990's. One of the first applications in the area was Cyberguide which used tourists' location and time to provide them information services such as suggestions for places of interest and background information about their current location [1]. Also, many similar systems were developed for example for museum visitors and exhibition tourists [2].

In our work we have implemented an application that lets manufacturers/retailers provide contextual product information about their products to their customers. Another important aspect is the ability for the customers to ask questions about the products directly from the manufacturers/retailers. In addition to that, the customers can also discuss about the products with other customers.

II. RELATED WORK

Mobile customer relationship management (mCRM) has been the focus of many recent studies. In his PhD thesis [3] Jaakko Sinisalo discusses the topic thoroughly. He concludes that although mCRM has an inherent potential to benefit both company and customer, it has been utilized quite cautiously to date. Reinhold et al. [4] come to the same conclusion in their study of utilizing mobile devices in collaborative CRM processes. They state that only if the offered mobile services show benefits in the perception of customers they will make use of them and engage into collaborative processes. Some ideas for the future of mCRM are also presented by Steimer et al. [5]. They focus on the possibilities of mCRM in the automotive industry and tourism, and present different location based and personalized services that could take advantage of the enhancing mobile infrastructure. Based on the current research, one can therefore say that mCRM has lots of unutilized potential. To address this, our application aims at providing practical ideas on how mCRM could be utilized.

Other recent work has focused on recommending items based on users' context and preferences. Shin et al. [6] propose a general framework for taking into account not only user preferences, but also context information for generating recommendations. Hong et al. [7] propose an agent-based framework that utilizes users' context history for automatically predicting users' preferences and providing personalized services and products for the users. Our application is more focused on supporting the users with their current products, but can also be used in recommending new products that the users might be interested in based on the existing products that the users have.

Our work also involves utilizing context information in providing the most relevant information. In addition to raw context information such as GPS coordinates, time and accelerometer data, it is often useful to create more abstract contexts such as 'driving' or 'at work'. The raw context information is utilized in deducing the more abstract contexts

 ISBN: 978-1-61208-223-3

thus creating a context hierarchy. One recent suggestion for a general context hierarchy is presented by Reichle et al. [9].

Another aspect in our work is community-based product information sharing as users can discuss about the products amongst themselves. This topic has been studied before by von Reischach et al. [10] who in their paper present a concept called APriori that enables consumers to access and share product recommendations using their mobile phone. They discuss issues related to community-based product recommendations such as how willing the users are to trust other users' opinions and how motivated they are in sharing their own.

III. The Application

In the application the manufacturers/retailers provide product information that can be assembled from different data types such as text, images, pdf and video. The product information can also be tied to a context such as location. The users can view the product information according to their preferences and/or context. They can also leave questions for the manufacturers/retailers to answer and leave comments for other users related to certain product information.

A. Web Interface

For manufacturers/retailers the application offers a web interface and for the customers the application offers both a web interface and a mobile Java ME application. Figure 1 shows the manufacturers/retailers' view of the web interface. The customers' web interface share a similar look and feel, but offer different functionality as manufacturers/retailers can add, modify and delete the product information, set contexts in which the product information is to be available in the mobile application and answer customers' questions while the customers can only view the product information, ask questions and discuss about the products with other customers.

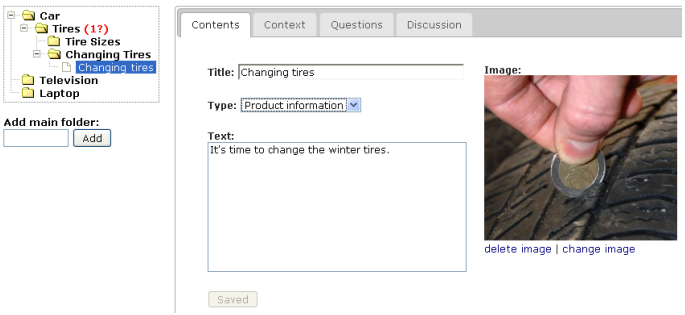


Figure 1. The manufacturers/retailers' web interface. The red numbers in brackets in the tree menu indicate that the corresponding folders / items have that many unanswered questions.

The product information items can be tied to a context so that the customer can browse only the most relevant items in the mobile application. For example, a reminder to change to winter tyres is only relevant in October, November and December when the temperature drops below five degrees Celsius. Another example would be a car repair shop providing its services in a certain location. However, it is not necessary to define a context for a product information item as in all cases the information may not be context specific.

Currently, only location names (using GeoNames[1]), maximum and minimum temperature and months can be used to define the context, but we have also defined a larger context hierarchy that we intend to support in the future. The context hierarchy is illustrated in Figure 2. It builds on lower level contexts that can be automatically recognized in the mobile device (calendar information, GPS, Cell ID, accelerometer data), and lower level contexts that can be automatically recognized in the server (time, weather). These lower level contexts are used in deducing higher level contexts (location, period of time, temperature) and user level contexts (at a hobby, at home, at work, in a car). Higher level contexts typically have distinct values while user level contexts are either active or inactive. Higher level contexts (location) as well as user level contexts can also be inputted manually by the user.

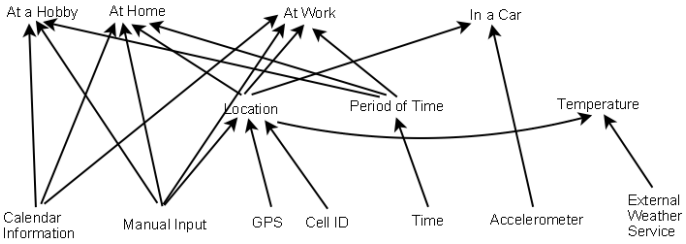


Figure 2. A suggestion of a context hierarchy for the application.

The customers can ask questions about the products from the manufacturers/retailers, and they can then answer the questions using their web interface. The customers can also discuss about the products with other customers, and the manufacturers/retailers are only allowed to view the discussion in their web interface. In the current implementation, the questions, answers and discussions can only contain text, but allowing more complex multimedia content could be useful as, e.g., the customers could take a photo of some problem in the product, which could be more informational for the manufacturers/retailers as plain text.

B. Customer's Mobile Application

Before the customers can start using the mobile application, they must login using their username and password.

[1] http://www.geonames.org/

 ISBN: 978-1-61208-223-3

After they have logged in once, the username and password are stored in the mobile device's memory and they are not required to login again unless they want to change the user. The main screen of the application offers options for browsing the product information, updating current location, changing preferences, changing user and exiting the application. The customer's mobile application is illustrated in Figure 3.

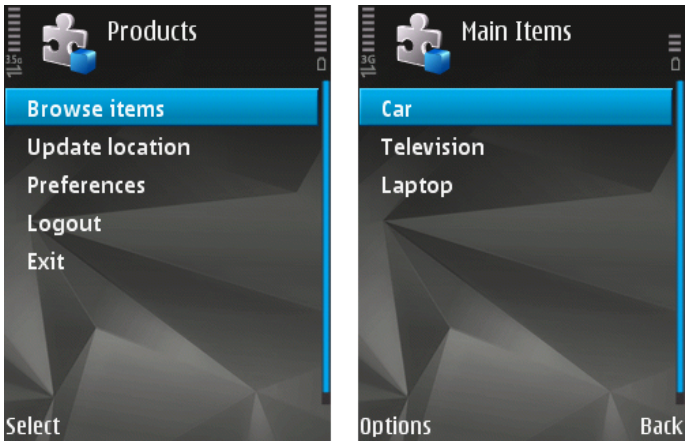


Figure 3. Customer's mobile application.

If the customers wish to restrict the information to what is most relevant in a certain location, they must input their location. This could be achieved automatically using GPS, but currently we have only implemented a way to input the information manually. The customers can input a location name and the application searches for matching location names from GeoNames. If matching location names are found, they are shown to the customers in a list of which they can choose their current location. Previously used location names are provided in a separate list, so that customers do not need to search for locations unless they go to a new location.

IV. Discussion

In summary, the features of the application that we have presented are:

1) the application lets manufacturers and retailers provide contextual product information to their customers
2) the application lets customers ask questions about the products directly from the manufacturers and retailers
3) the application lets customers discuss about the products with other customers

The benefits of this kind of an application are clear. The customers get a direct interactive feedback channel to the manufacturers/retailers that they can use to get guidance and assistance in using the product therefore increasing customer satisfaction. On the other hand, manufacturers and retailers get to make their customers happier and at the same time get valuable information that they can use in their product development. However, there are also some possible difficulties that this kind of an application can have. These issues are discussed next.

First of all, the customers need to install and use yet another application in their mobile device. Are the customers willing to do that? To start with, the installation process should be made as easy as possible. Perhaps the application could be installed by scanning a QR (Quick Response) code with the mobile phone's camera or by touching a RFID (Radio Frequency IDentification) tag with the mobile phone when buying the product, and the application could then provide a link to the web interface also. This way the application could be used without a username/password combination. On the other hand, if every product came with its own application, the number of applications in the customer's mobile device could become so big that the customer would not bother searching for the right application for a certain product anymore and would stop using the application. Therefore, having an option to use the same application for different products should also be possible. In addition to that, it is also important to have the customer's web interface as typing large amounts of text using a mobile device is often more difficult than using a regular keyboard. However, as current mobile devices are getting better and better (keyboards, touch screens, larger displays, etc.), one possible solution could be to only have the web interface and just make a more mobile browser friendly version of it.

Privacy issues are also always relevant when talking about context-aware applications. Would the customers be willing to reveal to the manufacturers/retailers when and where they are using the products? There should always be an option to restrict the context information that is exposed.

All in all, even though a lot of information can be searched from the Internet, this kind of an application with its direct interactive feedback channel between customers and manufacturers/retailers would provide the customers with more extensive and trustworthy answers to their questions. Many applications, like YouTube, Flickr, Twitter, etc. really become interesting only after they get a substantial amount of users. However, this kind of an application would be beneficial for the users directly after its launch as interacting with the manufacturers/retailers is the application's main functionality. Only the functionality of interacting with other customers would require a larger number of users for it to really become useful. For the manufacturers/retailers it would easily be worthwile to dedicate personnel for answering the questions that the customers are making using the application as the benefits of doing so would be clear.

The main focus in the future is to make better use of context and user preferences in providing the users the most relevant information about the products. In the current implementation we have only set up some simple preferences and basic contexts. We intend to support a more extensive

 ISBN: 978-1-61208-223-3

set of context information and user preferences, and also to implement the automatic recognition of context for the mobile application. Also, in the current implementation the users can only discuss about the products with all other users. It would be interesting to provide a possibility to create separate user groups also. For example, users who share the same hobby could discuss about products related to the hobby amongst themselves. In the future we also intend to verify the concept by running some experiments with actual users.

V. CONCLUSION

In this paper we presented an application that manufacturers and retailers can use to provide contextual product information for their customers. While the application is still in its preliminary stages, we focused on discussing the possibilities and challenges of the concept behind it. The main benefit that we discussed was that this kind of an application would provide guidance and assistance to the customers and a direct feedback channel to the manufacturers and retailers. Some of the potential difficulties that we highlighted were privacy issues (what kind of context information the customers would be willing to reveal) and whether the application should be a stand-alone mobile application or a web application. In the future we intend to develop the prototype further and explore these issues with actual users.

REFERENCES

[1] Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R. and Pinkerton, M.: Cyberguide: A Mobile Context-Aware Tour Guide. In: Wireless Networks Vol. 3, No. 5, pp. 421–433. Kluwer Academic Publishers, Hingham, MA, USA (1997)

[2] Chen, G., Kotz, D.: A Survey of Context-Aware Mobile Computing Research. Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College (2000)

[3] Sinisalo, J.: Mobile Customer Relationship Management: a Communication Perspective. PhD Thesis. University of Oulu (2010).

[4] Reinhold, O., Alt, R.: Enhancing Collaborative CRM with Mobile Technologies. In: Proceedings of the 22nd Bled eConference, pp. 97-116 (2009)

[5] Steimer, F.L., Steimer, T.: Future Mobile CRM in the Automotive and Tourist Area. In: Proceedings of the 3rd IEEE International Conference on Digital Ecosystems and Technologies, pp. 177–180 (2009)

[6] Shin, D., Lee, J., Yeon, J., Lee, S.: Context-Aware Recommendation by Aggregating User Context. In: Proceedings of the 2009 IEEE Conference on Commerce and Enterprise Computing, pp. 423–430. IEEE Computer Society, Washington, DC, USA (2009)

[7] Hong, J., Suh, E-H, Kim. J., Kim, S.: Context-Aware System for Proactive Personalized Service Based on Context History. In: Expert Syst. Appl. Vol. 36, No. 4, pp. 7448–7457. Pergamon Press, Inc., Tarrytown, NY, USA (2009)

[8] Han, B.-J., Rho, S., Jun, S., Hwang, E.: Music Emotion Classification and Context-Based Music Recommendation. In: Multimedia Tools Appl., Vol. 47, No. 3, pp. 433–460. Kluwer Academic Publishers, Hingham, MA, USA (2010)

[9] Reichle, R., Wagner, M., Khan M. U., Geihs, K., Lorenzo, J., Valla, M., Fra., C., Paspallis, N., Papadopoulos, G. A.: A Comprehensive Context Modeling Framework for Pervasive Computing Systems. In: Proceedings of the 8th IFIP WG 6.1 International Conference on Distributed Applications and Interoperable Systems, pp. 281–295. Springer-Verlag, Berlin, Heidelberg (2008)

[10] von Reischach, F., Guinard, D., Michahelles, F., Fleisch, E.: A Mobile Product Recommendation System Interacting with Tagged Products. In: Proceedings of the 2009 IEEE International Conference on Pervasive Computing

 ISBN: 978-1-61208-223-3

Adaptive OLAP Caching

Towards a better quality of service in analytical systems

Pedro Marques
ALGORITI R&D Centre
University of Minho
PORTUGAL
pcmarkes@gmail.com

Orlando Belo
ALGORITI R&D Centre
University of Minho
PORTUGAL
obelo@di.uminho.pt

Abstract — **Nowadays, the use of Multidimensional Data Systems has become a part of everyday actions in medium and large companies. These systems, which concern mainly in aiding their users in the process of decision-making, have a large flexibility in data exploration and high performance response levels to queries. Despite all the existing techniques, it is sometimes very hard to maintain such levels of performance that the user demands. With the purpose of tackling eventual performance losses, other techniques where developed trying to reduce data servers load. One of such mechanisms is the creation of OLAP caches maintaining previous queries and serving them upon subsequent requests without having to ask to the server. Due to OLAP systems organization, it is possible to identify the characteristics of its users and its exploration patterns – what queries will a user submit during a session, their frequency and resources involved. However, it is possible to go one step further, and to predict exactly what data will be requested by a specific user and, especially, the sequence of those requests. This is called the prediction phase and is followed by the pre-materialization of views that correspond to the user's requests in the future. These views are then stored in the cache and served to the user in the appropriate time. The technique we propose here consists in maintaining a positive ratio between the time spent to predict and materialize the most relevant views to users, and the time that would be spent if no prediction had been done.**

Keywords – on-line analytical processing; analytical servers; caching; association rules mining; cache content prediction.

I. INTRODUCTION

Due to the amazing increase of companies' data repositories in the last decade, attentions are now turned to the implementation of more powerful ways of analyzing data. As a consequence, Decision Support Systems, and more specifically, *Online Analytical Processing* (OLAP) [2] [18] systems are being implemented in a large scale when compared to few years ago scenarios. As we know, one of the greater advantages of OLAP Systems is the fact that they can cope with large volumes of data and execute *ad-hoc* queries within various analysis perspectives giving to decision makers an exceptional way to get more structured insights into company's data. OLAP systems were so well accepted by decision makers that soon they started loading more and more data into them and issuing more complex queries, which quickly surfaced some critical performance issues. As fast as an OLAP Server could be, there is always some space to apply new optimization strategies, trying to improve OLAP servers' performance and OLAP users' satisfaction. Thus, the usage of caching mechanisms in OLAP platforms is a natural (and viable) technological choice when one is concerned to improve the quality of service of an OLAP platform.

Despite being widely implemented and tested, conventional caching mechanisms were not prepared to handle OLAP data. One of the reasons why this type of information was not ideal for caching was due to its dynamic nature (i.e. versus the static nature of HTML information where caching techniques have a particularly good fit). Other aspect to be considered when deal with OLAP data is the dimension of the data to be kept in cache, both in terms of volume of data as well as in terms of data structure complexity. Comparing again with HTML data, which represents a little effort in terms of space needed to keep it in cache, OLAP data requires a great amount of space simply due to the fact that any response to a typical MDX (Multidimensional Expression) query involves a lot of data, usually materialized in a multidimensional data view (a data cube). Even with the diversity of the data to be maintained, several techniques were developed to apply caching mechanisms to OLAP data [10] [11] [16], revealing benefits good enough to keep the focus on improving caching techniques in order to integrate them effectively on OLAP server systems.

The work developed was based on an analysis of today's caching mechanisms and their application in the OLAP field, and based on selected information about user's querying patterns. In order to obtain these patterns, OLAP server logs were fetched, analysed and mined in order to obtain a set of association rules that represent the actions (and consequences) of user's queries (usage profiles), providing us the means to predict future user's querying tendencies. Such predictions unlock the possibility of issuing a query even before the user, put in cache the results that support responses to a specific user query, providing it faster than if no cache was available in the OLAP platform.

In the next sections, a more in-depth analysis to this process will be conducted, explaining the various stages reached along the evolution of the work, as well as discussing

 ISBN: 978-1-61208-223-3

some of most relevant considerations needed to understand the complexity of predicting the multidimensional content of a cache for a specific OLAP platform. This paper is organized into more five sections, namely: Section II shows a detailed overview about OLAP caching, its advantages and disadvantages; Section III presents some related work and discuss major characteristics of the problems we could face when dealing with high specialized caches, as are OLAP caches; Section IV, it's where we present our approach for a new model of OLAP caching; Section V reveals and discusses the results of the performed tests to validate the OLAP caching model proposed; and finally, section VI, presents some final remarks and conclusions, as well as some future research lines.

II. TO CACHE OR NOT TO CACHE

Whatever the specific area of implementation could be, when implementing caching mechanisms one has to remember that the space available for storing the cache is not unlimited. As a direct consequence we need to choose (and evaluate) what data should be kept (or not) in a cache and what data should be removed giving space for new (and hopefully more relevant) data to the users' needs. Keeping this in mind, researchers started to test quite well known algorithms – frequently referred as cache management algorithms – that up to that time had only been used in other types of environments such as for caching HTML pages with great success. As results became known, there was a clear notion that there should be promoted some additional efforts to develop new breads of algorithms that focused OLAP scenario in particular. One of the most common ways of evaluating the value of any cache algorithm is using a metric called *Hit Ratio*. This metric is the ratio between the number of requests that were in cache and the total number of requests that were made. However, *Hit Ratio* is not a perfect metric. For instance, even with a higher *Hit Ratio*, the number of bytes served directly by the cache could be smaller than a cache with a lower *Hit Ratio*, which lead to the creation of another metric: *Byte Hit Ratio*. The latter metric has been vastly used to evaluate how a cache can satisfy its clients' requests.

As a user of any OLAP (or other) system launches his queries, the cache management algorithm has to check if the necessary information is stored in the cache or, if it is not, to decide whether it should or shouldn't be added to the cache. If the request cannot be satisfied directly from the cache, there are two possible outcomes:

1) the cache still has space to accommodate the new data, and so it is added without further due, or
2) the cache hasn't enough space to store the new data.

In the former case, the content is added, and after that time, when it is requested, it will be served from cache instead of being satisfied directly by the OLAP Server. If there is no space available in the cache management system, the algorithm can either discard this information or free some space in cache in order to add this new data. This is the main decision that cache algorithms have to make. As we know, this decision will affect the way a cache behaves in the presence of new information to be added. One of the most basic ways to do this selection is to use a FIFO approach, which means that the oldest record to have been added to cache will be removed in order to create space for a new entry. If this is not enough, the second (the third, and so on) oldest record will be removed as necessary, record by record. The main problem with this technique is the fact that it doesn't take into account the nature of the data, despite of its size or actuality, data has an intrinsic value that cannot be measured as simplistically as these approaches propose. Other (more sophisticated) decision metrics were developed taking into account the timestamp of last access to a specific piece of data [15], the frequency of access to the data [6], or other more complex information such as the ones used by the Greedy Dual algorithm [5], for instance. All these metrics, in one way or another, take into account the intrinsic value of data and the relevance each piece of data has to the users and, therefore, they are much more suited to do the (caching) job correctly than others that simply look to the characteristics of the data neglecting its nature and its relevance to users.

III. OLAP CACHING

One of the most common operations performed when querying an OLAP Server are the well known drill-down and roll-up operations. The first of these two operations consists of lowering the grain at which the data is being analysed. For instance, we can go down in a hierarchy, detailing systematically, level by level, the grain of the data, from a country-level view to a district-level one. The roll-up operation is its direct counterpart, allowing viewing data at a higher level following a determined hierarchy. In an OLAP Server, the data is stored at the lowest level of granularity and then aggregated to a level required by a specific multidimensional request. Kalnis and Papadias [10] proposed a solution where this characteristic is explored, mainly by sharing the cache over several cache servers, specifically *OLAP Cache Servers* (OCS). In this approach, each OCS has the capability to apply transformations (aggregations and other operations) to multidimensional structures and, thus, combine them to satisfy, at least, part of a request that has been launched by a user. This way, whenever a user issues a query, the various OCS are asked if they have the needed information and, even if they don't, they are asked again if they can compute it from the data they have at a lower grain than the user requested. This means that an OCS can satisfy not only requests that have been issued before (and cached) but also other issues that involve computations over the data that exists in the OCS.

Kalnis et al. [11] proposed another alternative where individual caches of users are shared through a Peer-to-Peer network created between users of a same OLAP System – PeerOLAP. Essentially, this approach was based on the Piazza System [8], and intended to allow a very high level of

 ISBN: 978-1-61208-223-3

autonomy in the cache network due to the dynamic nature of Peer-to-Peer networks, where users can connect and disconnect without significantly affecting the overall usability and performance of the system.

As mentioned before, OLAP data is quite dynamic by nature, which means that it is very difficult to predict when the cached data will become out-dated. To deal with this problem, an active caching technique was created [4]. It consists of keeping in the cache server a Java applet that is invoked every time a cache hit occurs. This applet has the role to check with the OLAP Server if the cache information stills valid or if it has changed since the last time it was requested by one or more users. If data stills valid, it will be returned to the user who requested it. If not, the full request will be redirected to the OLAP Server.

One other question that was placed often by researchers, it was focused on what would be the optimal level of granularity to store data in a cache, in order to not only be able to aggregate it as needed but also to be able to do that in a timely fashion manner. Deshpande et al. [7] used a new indivisible unit called chunk. This data unit, with a low granularity level was mapped in the cache in order to be aggregated to satisfy user's requests. The mapping occurs in the server and denotes the relationship between a chunk and the basic units stored in the OLAP Server, allowing for the complementary fetching of data from a central server. When a cache server receives a request from a user, it calculates the parts of that request that it can be satisfied accessing directly the cache, and the information that it need to be requested to the OLAP Server (at a low level of granularity). When all the required data is located in the cache server, it combines it and sends the results to the user, without him ever knowing if the information he received came from the central server or the cache server.

As a last reference we selected the work presented by Sapia [16], an approach particularly interesting to us. In that work, the author proposed a predictive system for user behaviour in multidimensional information system environments that explore characteristic patterns users use to show when explore multidimensional data structures. It is an OLAP caching approach that complements other techniques, such as the ones presented by Albrecht et al. [1] or Deshpande et al. [7].

IV. A New OLAP Caching Approach

This work was based on the assumption that OLAP System's users have predictable patterns of data that they use to consult on their regular OLAP sessions. The nature of most OLAP users in a company – decision makers – use to induce them to be focus on a relatively small subset of data stored in a data warehouse. The day-to-day activity of a decision maker may begin with an analysis of a pre-defined dashboard or an interactive report, and based on the information he gathers from the analysis of the data, will continue his exploration in a lower level view of the same data – probably appealing to a typical drill-down operation. This shows us that for any given user his behaviour will be repeated during a certain period of time, revealing then a regular usage pattern.

One possible way of extracting these patterns is by analyzing de OLAP Server's logs that contain information about what multidimensional queries users had submitted and when they happened. It is also possible to know, for a given user, the sequence of queries he launched between his login and his logout in a specific OLAP session. From the analysis of this kind of information, giving a certain period of an OLAP system exploration, another problem arose: how far back in the logs should we go to make sure that the retrieved rules are truly representative of the user's exploration patterns?

On one hand, if we analyse the OLAP exploration habits (and tendencies) for a short period of time, we may get rules that represent the most recent patterns and not what the user usually does in the "long run". However, on the other hand, if we analyse a larger period, we may extract rules that represent older OLAP exploration patterns that do not represent what users are doing currently (users may change their exploration habits due to a large variety of reasons, demanding that the algorithm should be able to adapt to such changes).

Taking these constraints into consideration, we began our approach by retrieving the OLAP Server's log files, preparing them to be analysed by a specific data mining algorithm with the ability to generate a set of association rules that represent the most relevant exploration user patterns – we designate a set of usage patterns by an OLAP profile. We used the well-known *Apriori* algorithm [3], which is one of the most used algorithm for mining frequent itemsets, having prove its effectiveness so many times analysing a set of transactions and surfaces the relationships between them, given a minimum value for support and confidence.

As it is well known, association rules are usually represented in the format: $A \rightarrow B$ *(sup=*α*; conf=*β*)*, where *sup* and *conf* represent, respectively, the support and the confidence values of a rule. From an association rule (and from its support and confidence values) we can retrieve two important things, namely the:

- *support* (*sup*), that represents the ratio between the number of times that a sequence of queries A followed by a sequence of queries B was found in the dataset and the total number of queries in that dataset:

$$sup(A \rightarrow B) = \frac{\#(A \rightarrow B \text{ in the dataset})}{\#(\text{queries in the dataset})}$$

- *confidence* (*conf*), that represents the number of times a sequence of queries *A* is followed by a sequence of queries *B* in the dataset, divided by the number of times a query *A* (independently of what query followed it) was found in the same dataset:

$$\text{conf}(A \rightarrow B) = \frac{\#(A \text{ followed by } B \text{ in the dataset})}{\#(A \text{ in the dataset})}$$

If we take $A \rightarrow B$ *(sup=0.3; conf=0.8)*, as an example, we can say that for every time a user issues the query *A* he will, in

 ISBN: 978-1-61208-223-3

80% of the cases, issue the query *B*, right after that. On the other hand, generally speaking, we can say that for the analysed dataset, a sequence of queries *A* followed by a sequence of queries *B* occurred in 30% of all cases. This technique allows us to establish probabilities for the sequence of queries that a user will issue between the beginning and the end of an OLAP session. With this information some actions may be taken to improve the OLAP Server's response time to queries. We could then simulate the user's interaction and place in cache the views he used mostly. The main problem with this is the high value of rules that are going to be generated. This could easily produce untreatable results.

Keeping this problem in mind, our work focused on reducing the number of queries that should be taken into account in the prediction phase, without affecting significantly the results. To do this, we choose to map all the sequences of queries predicted by the mining algorithm, representing them in a Markov chain [9] as a way to provide a better visual insight of the entire set of generated rules. Next, we defined the minimum value for the confidence associated with the rules that should be taken into account in the prediction phase (*minconf*). Soon, we discovered that this action would not be enough if we wanted to reduce effectively the number of predicted queries. We needed to optimize a little bit more the process. When removing the rules with a confidence value smaller than *minconf*, we realized that some rules remained without the possibility to be predicted as a sequence of any other query. If we think of the sequence of queries as a graph, and we start removing some of the nodes, there are some of them that lose their entrance arches. Those "nodes" represent the queries that were removed in this second optimization step. This way we also risk an increased number of cache misses, but provide us an alternative way of reducing the number of views to be pre-materialized in the cache.

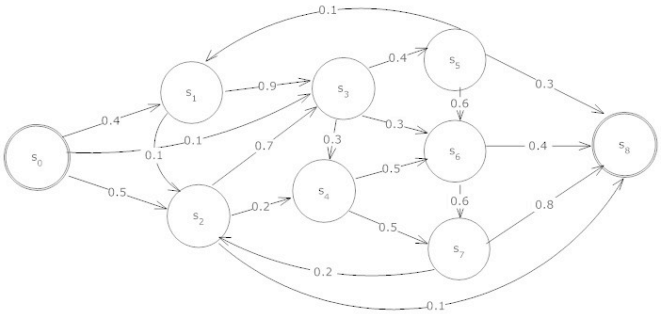


Figure 1. A query sequence prediction for the first dataset

V. VALIDATING THE PROPOSED TECHNIQUE

In order to test the technique proposed here, we decided to promote two different test cases, considering the number of query hits achieved before and after the proposed optimization scenarios for a given set of artificial queries (generated by artificial processing algorithms, not representing the actual usage of an OLAP Server). In Figure 1, we can see the sequence of queries (in the form of a Markov chain) that were predicted by the mining algorithm that was used – S_0 and S_8 represent, respectively, the begin of the session provoked by the user's login and the end of that session. The vertices' values represent the transition probabilities between two different states (or queries in this case).

All the tests conducted over this dataset basically used various values for *minconf* simplifying the rules accordingly. The chosen values for *minconf* were, respectively, 0.3, 0.4 and 0.5 (Table I). One other simplification that was introduced, and named as "main route", simplistically put in the cache the sequence of queries that an user most likely follow in a future data exploration process. In Figure 1, easily we can identify such candidate sequence of queries. It will be the sequence represented by the path $S_0 \rightarrow S_2 \rightarrow S_3 \rightarrow S_5 \rightarrow S_6 \rightarrow S_7 \rightarrow S_8$. It can be easily found by following the higher transition probabilities between the S_0 and S_8 nodes. The results of the tests, for the different values of *minconf* and for the "main route" simplification models, can be found in Figure 2.

TABLE I. TEST RESULTS FOR THE FIRST DATASET

Minconf	**0.3**	**0.4**	**0.5**	***"main route"***
Pre-materialized views (%)	100	86	28	71
Cache Hits (%)	100	89.8	38.3	79.78

As a comparison value, if we add 50% of all queries to the cache, intuitively we think we would achieve almost 50% cache hits for any given user (Figure 2). However, this value is merely meant to provide us with a reference value, and should not be taken into account in terms of absolute values.

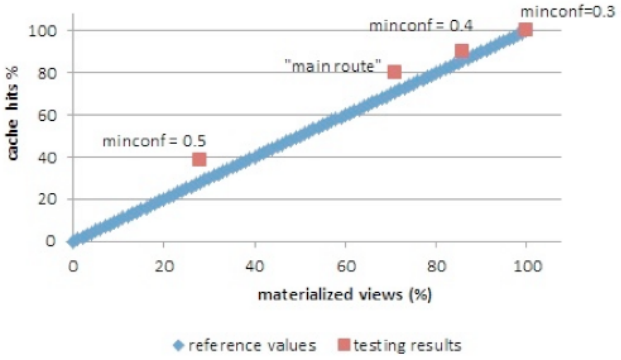


Figure 2. Test results graph for the first dataset

Observing Figure 2, it will lead us to note two key values of *minconf* – 0.3 and 0.5 –, which shows the most relevant (best and worst) test results. As for the value 0.5, it means that only 28% of all possible views were pre-materialized and, even in that case, the cache hits came around 38.3%, which represents a 10% increase in system performance when compared to our reference values. On the other hand, the usage of 0.3 for *minconf* resulted in no view being simplified and, consequently, the values of cache hits were measured at 100%.

Later, other tests were conducted with a real set of data retrieved after a simulation of an OLAP Server usage. This second dataset contains a total of 59 queries being issued to the server, and the values of *minconf* used to simplify the generated rules were 0.02, 0.03, 0.4 and 0.6. The results of this

 ISBN: 978-1-61208-223-3

second experience can be found in Table 2 and Figure 3. The results obtained in this second round of tests shows us that, even though the differences between the different values of *minconf*, they don't yield great differences in the percentage of cache hits – nor in the percentage of materialized views.

TABLE II. TEST RESULTS FOR THE SECOND DATASET

Minconf	**0.02**	**0.3**	**0.4**	**0.6**
Pre-materialized views (%)	54	52	50	46
Cache Hits (%)	89	88	87	86

The gains relative to the reference values were quite relevant, staying approximately between 35% and 40% (for values of *minconf* equal to 0.02 and 0.6, respectively).

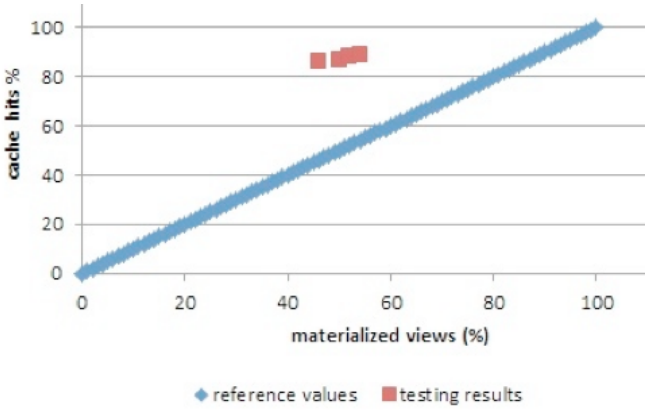


Figure 3. Test results graph for the second dataset

VI. CONCLUSIONS AND FUTURE WORK

The main goal of this study was to investigate in what conditions a predictive caching system could be used in a typical OLAP environment. In order to reach such goal, we studied several known cache techniques, e.g. [4] [7] [10] [11] [12] [13] [14] [16] [17], trying to establish the basis to propose a different form to know *a priori* the contents of an OLAP cache in a near future. All those techniques were crucial to the development of our work, for both the ideas of exploring the log files present in the OLAP Server and the simplification of the rules generated after the application of mining algorithms to that information.

All the tests performed showed satisfactory improvements in the ratio between materialized views and cache hits. In our perspective, they also showed that this approach has the necessary pre-requisites to be applied to a more real scenario with advantages for the overall system's global performance. Even though some important questions remain, both for the period of logs that should be analysed and for the values of *minconf* to be used. This last, is an issue that should be addressed on a case-by-case approach, and should be included in a typical tuning-phase after finishing system implementation.

Finally, we think that with larger datasets feeding the mining algorithm, results should be even better. For that reason we plan in a near future to extend the current study, comparing it with other similar approaches and including some work concerning the exploration of multidimensional queries. We will give particular attention to the less busy periods of an OLAP server, in order to pre-materialize some specific multidimensional views that can be used latter when a user logs in – the log in periods can, as well, be subject of prediction.

REFERENCES

[1] Albrecht, J., Bauer, A., Deyerling, O., Günzel, H., Hümmer, W., Lehner, W. and Schlesinger, L. (1999). Management of Multidimensional Aggregates for Efficient Online Analytical Processing. In Proceedings of the 1999 International Symposium on Database Engineering \& Applications (IDEAS '99). IEEE Computer Society, Washington, DC, USA, 156-,.

[2] Abelló, A. and Romero, O., On-Line Analytical Processing (OLAP). In Encyclopedia of Database Systems (editors-in-chief: Tamer Ozsu & Ling Liu). Springer 2009. Pages 1949-1954.

[3] Agrawal, R. and Srikant, R. (1994), Fast algorithms for mining association rules in large databases. In Jorge B. Bocca, Matthias Jarke, and Carlo Zaniolo, editors, Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, pages 487-499, Santiago, Chile, September 1994.

[4] Cao, P., Zhang, J. and Beach, K., Active Cache: Caching Dynamic Contents on the Web. Distributed Systems Engineering, 6(1):43-50, 1999.

[5] Cherkasova, L. (1998). Improving WWW Proxies Performance with Greedy-Dual-Size-Frequency Caching Policy, HP Laboratories Technical Report HPL, 1998.

[6] Chrobak, M. and Noga, J. (1999). LRU is Better than FIFO. Algorithmica, Springer New York 23: 180-185.

[7] Deshpande, P., Ramasamy, K., Shukla, A. and Naughton, J. (1998). Caching multidimensional queries using chunks. ACM.SIGMOD Rec. 27, 2 (June 1998), 259-270..

[8] Gribble, S., Halevy, A., Ives, Z., Rodrig, M. and Suciu, D. (2001). What can databases do for Peer-to-Peer. In Proc. of WebDB, 2001.

[9] Howard, R., Dynamic Programming and Markov Processes. MIT Press, June, 1960.

[10] Kalnis, P. and Papadias, D. (2001). Proxy-server architectures for OLAP. In Proceedings of the 2001 ACM SIGMOD international conference on Management of data (SIGMOD '01), Timos Sellis (Ed.). ACM, New York, NY, USA, 367-378.

[11] Kalnis, P., Ng, W., Ooi, B., Papadias, D. and Tan, K. (2002). An adaptive peer-to-peer network for distributed caching of OLAP results. In Proceedings of the 2002 ACM SIGMOD international conference on Management of data (SIGMOD '02). ACM, New York, NY, USA, 25-36.

[12] Lawrence, M., Dehne, F. and Rau-Chaplin, A. (2007). Implementing OLAP Query Fragment Aggregation and Recombination for the OLAP Enabled Grid, Parallel and Distributed Processing Symposium, 2007 (IPDPS 2007), IEEE International, 26-30 March 2007

[13] Lehner, W., Albrecht, J. and Hümer, W. (2000). Divide and Aggregate: caching multidimensional objects. Proceedings of the Second Intl. Workshop on Design and Management of Data Warehouses (DMDW 2000), Stockholm, Sweden.

[14] Loukopoulos, T., Kalnis, P., Ahmad, I. and Papadias, D. (2001). Active Caching of On-Line-Analytical-Processing Queries in WWW Proxies, In Proceedings of the International Conference on Parallel Processing (ICPP '01). IEEE Computer Society, Washington, DC, USA, 419-.

[15] Mookerjee, V. and Tan, Y. (2002). Analysis of a least recently used cache management policy for Web browsers. Operations Research, 50, 2 (March 2002), 345-357.

[16] Sapia, C., PROMISE: Predicting Query Behavior to Enable Predictive Caching Strategies for OLAP Systems, In Proceedings of the Second International Conference on Data warehousing and Knowledge

 ISBN: 978-1-61208-223-3

Discovery (DAWAK 2000), Greewich, UK, Septeber 2000, Springer LNCS, 2000.

[17] Yao, Q. and An, A. (2003). Using user access patterns for semantic query caching. Database and Expert Systems Applications, 14th International Conference. Prague, Czech Republic.

[18] Zhenyuan, W. and Haiyan, H. (2010). OLAP Technology and its Business Application, Intelligent Systems, WRI Global Congress on, pp. 92-95, 2010 Second WRI Global Congress on Intelligent Systems.

 ISBN: 978-1-61208-223-3

Cultural Intelligence Decision Support System for Business Activities

Zhao Xin Wu, Roger Nkambou
Computer Science Department, University of Quebec in Montreal, Case postale 8888, Succursale Centre-ville
Montreal, QC, H3C 3P8, Canada
wu.zhao_xin@courrier.uqam.ca
nkambou.roger@uqam.ca

Jacqueline Bourdeau
LICEF, TÉLÉ-UNIVERSITÉ, 100 Sherbrooke, O.
Montreal, QC, H3C 3P8, Canada
Jacqueline.bourdeau@licef.ca

Abstract— **With the rapid growth of globalization, cross-cultural business activities are on the rise. These activities are of increasing importance for individuals and organizations. As a result, culture has become a decisive factor in the success of cross-cultural business. In today's global marketplace, individuals and organizations must be culturally intelligent. The concept of cultural intelligence has been introduced in the literature on global management. Researches on cultural intelligence have provided a new perspective, and, hence, have presented a new way to alleviate cross-cultural business challenges. However, these researches mainly rely on questionnaires to find solutions to the problems of cultural intelligence in cross-cultural business. To date, no research on cultural intelligence has been empirically computerized. This study aims first to create a new cultural intelligence computational model in order to process the soft data in cross-cultural business activities. This will enable individuals and organizations to make better decisions in cross-cultural business activities. Second, this study will implement the model in a system called the CIDSS through the use of artificial intelligence techniques. The purpose of the CIDSS is to allow individuals and organizations to make intelligent decisions in culturally diverse business activities and to solve the intercultural adaptation problems faced in a variety of authentic cross-cultural situations.**

Keywords - Cross-Cultural; Cultural Intelligence; Business Intelligence; Decision Making; Soft-Computing.

I. INTRODUCTION

The globalization of productive forces is rapidly increasing. In order to face this new reality, individuals and organizations must form global strategic alliances to deal with worldwide competitors, suppliers and customers. When confronted with cultural diversity, some individuals and organizations are able to adapt successfully to the new cultural business environment [1], while others are not. What is the decisive factor for these opposing responses? How can good decisions be made in culturally diverse business environments? What skills can be improved to enable cultural adaptation? [2]

In recent years, researchers have shown great interest in globalization and intercultural management. Cultural intelligence has, therefore, been presented as a new phenomenon as a means to answer the above questions [3]. Organizational psychology and human resource management have paid a great deal of attention to cultural intelligence since its introduction. These fields of study have yielded valuable results that apply to the real business world.

However, since Earley and Ang put forward the concept of cultural intelligence in 2003, there has been no research model on cultural intelligence with business intelligence technology, and most current studies pertaining to cultural intelligence do not integrate any artificial intelligence computational aspect [4]. In addition, traditional business intelligence has encountered two challenges: the first involves determining the means of adapting to cultural diversity; the second pertains to the measures required to treat soft data for decision making. Our claim is that when cultural intelligence is applied to individuals and organizations in the fields of business, it should be computationally modeled.

This research attempts to offer effective solutions to the aforementioned problems. It is the first attempt to create a new computational model of cultural intelligence implemented in an intelligent system called the CIDSS (Cultural Intelligence Decision Support System), to resolve cross-cultural business problems. The main reason to create such a system is that, in the real business world, there are not enough qualified cultural experts to help users make better business decisions, and these experts may lose some of their effectiveness after long consecutive hours of work. Moreover, the sphere of application has been confined to cultural experts and researchers. From a user's point of view, the system appears as if an efficient team of top cultural experts is always working together with him/her. This system has the potential to achieve better performance results than human experts.

There are three goals behind such a system, when helping individuals and organizations to cooperate more effectively with people from different cultural backgrounds: 1) to assist individuals and organizations in their business decision-making processes involving cultural affairs; 2) to assist people in improving their use of a specific form of intelligence based on their capacity to understand, to reason correctly, and to adapt to culturally diversified situations [5]; and 3) to facilitate the work of researchers and to equip them with more effective tools in their studies on cultural intelligence.

 ISBN: 978-1-61208-223-3

This paper consists of six sections. In Section I, we state the research question and research objectives. In Section II, we briefly discuss the concept of cultural intelligence and its dimensions. In Section III, we introduce the relationship between business and cultural intelligence. In Section IV, we present an overview of the system's architecture, we identify the main modules in the architecture, and we explain how these modules work. In Section V, we discuss the fundamental cultural intelligence soft-computing model inference techniques that are applied in the architecture. Finally, in Section VI, we present the contributions of this research.

II. CULTURAL INTELLIGENCE AND ITS DIMENSIONS

In the research literature, researchers have different opinions regarding the concept of cultural intelligence. Earley and Ang presented cultural intelligence as a reflection of people's ability to collect and process information, to form judgments, and to implement effective measures in order to adapt to a new cultural context [6]. They also indicate that cultural intelligence should predict performance and adjustment outcomes in multicultural situations, when an individual is faced with diversity. Earley and Mosakowski redefined cultural intelligence as the ability of managers to deal effectively with different cultures [7]. They suggested that cultural intelligence is a complementary form of intelligence, which may explain the capacity to adapt to cultural diversity, as well as to operate in a new cultural setting. Peterson interpreted cultural intelligence in terms of its operation [8]. He believes that the concept of cultural intelligence is compatible with the cultural values of Hofstede and their five main dimensions, i.e., individualism versus collectivism, masculinity versus femininity, power distance, uncertainty avoidance, and short- and long-term orientation [9]. Brisling et al. defined cultural intelligence as the level of success people obtain when adapting to another culture [10]. Thomas explained cultural intelligence as the ability to interact efficiently with people who are culturally diverse [11] [12]. Ng and Earley presented cultural intelligence as the ability to be effective in all cultures [13]. Johnson et al. defined cultural intelligence as the ability of an individual to integrate a set of knowledge, skills and personal qualities so as to work successfully with people from different cultures and countries, both at home and abroad [14].

Researchers in the field of culture also use different dimensional structures to measure cultural intelligence. Each of these researches is associated with conceptual models. These structures seek first to explain the attributes that enable people to develop their abilities in various cultural contexts, and then to determine how people can improve these capabilities. Earley and Ang presented the first structure of cultural intelligence, which integrates the following three dimensions: cognition, motivation and behaviour [5]. While Thomas agrees with Earley and Ang that there are three dimensions to cultural intelligence [12], he does not share their point of view regarding what these three dimensions should be. He therefore advocated another tridimensional structure. His belief is founded on the theory of Ting-Toomey, which states that the structure of cultural intelligence should be based on the skills required for intercultural communication, that is to say, knowledge, vigilance and behaviour [15]. Vigilance acts as a bridge connecting knowledge and behaviour, which is the key to cultural intelligence. Tan believes that cultural intelligence has three main components: 1) strategic thinking about culture; 2) dynamics and persistence; and 3) specific behaviours [1]. Tan stressed the importance of behavior as being essential to cultural intelligence. If the first two parts are not converted into action, cultural intelligence is meaningless.

Ang and Van Dyne [5] suggested a four-dimensional cultural intelligence structure. This structure is based on the general intelligence structure of Robert and Douglas [16]. Ang and Van Dyne used the framework of Robert and Douglas, which divides cultural intelligence into metacognitive cultural intelligence, cognitive cultural intelligence, motivational cultural intelligence and behavioral cultural intelligence. This structure has been widely used in the following cultural researches and studies.

III. BUSINESS AND CULTURAL INTELLIGENCE

We believe that cultural differences have a greater impact on business efficiency than previously thought. Cultural backgrounds influence how people think, act and interpret information during business activities. Business is becoming increasingly globalized, and partnerships are a means to gain a competitive advantage. Thus, the potential for success or failure depends on the ability of organizations and leaders to make appropriate decisions within a framework of cultural diversity. Businesses and leaders must understand and become proficient in intercultural communication. In this regard, cultural intelligence offers strategies to improve cultural perception in order to understand the culturally motivated behavior of individuals and organizations. Many articles address the importance of cultural intelligence [17] [18] [19] and culture in the context of international business [20] [21] [22]. Huber indicated that the performance of an international business, in terms of efficiency and effectiveness, is determined by the quality of its organizational intelligence [23]. Ang and Andrew [24] specified that organizational intelligence is the cultural intelligence of businesses. Cultural intelligence in business is based on the research on psychology concerning the cultural intelligence of individuals, as well as on the views of the organizations. Cultural intelligence permits businesses to collect a set of resources and to develop their capabilities. Ang and Andrew suggest that, when organizations venture into foreign territories, cultural intelligence is a necessary predictor of organizational performance. The involvement in international trade offers significant advantages and challenges to the business development of a company. A business may be successful at home because of its cultural sensitivity. However, this does not guarantee that it will be able to attract international suppliers, partners and customers. If the business does not learn to adapt to cultural differences, it risks losing and missing business opportunities. A business approach that is culturally inappropriate may be detrimental

 ISBN: 978-1-61208-223-3

when doing business abroad. Knowledge and sensitivity toward other cultures result in increased business success. Consequently, cultural intelligence is of the utmost importance when engaging in business practices.

IV. CIDSS COGNITIVE ARCHTECTURE

Sternberg and Detterman [16] determined that general intelligence has four dimensions, i.e., metacognition, cognition, motivation and behaviour. Cultural intelligence should also include these four dimensions. We believe that the diverse structures of cultural intelligence should be considered collectively in order to integrate the elements required to respond to the cultural knowledge acquired. Therefore, we built the cultural intelligence architecture of the CIDSS based on the specific four-dimensional cultural intelligence structure of Ang and Van Dyne [5]. It represents a comprehensive overview of the multi-aspects of the researches on cultural intelligence.

The architecture of the CIDSS uses both the symbolic and the connectionist approaches of artificial intelligence. The CIDSS respects the cognitive concepts of Ang and Van Dyne [1] regarding the theories of global cultural intelligence, as well as other important aspects, such as the Hofstede's theory of the five dimensions of culture [9]. The CIDSS also relies on engineering concepts in its solutions when designing and implementing software. It offers learning mechanisms which emulate human intelligence. The CIDSS is a distributed and modular architecture. It relies on the functional "*consciousness*" mechanism for much of its operation [25]. Its modules communicate and offer information to each other.

The CIDSS can be considered as a cognitive agent with an eleven-step cognitive process. This agent recognizes business-related information in natural language from its environment. By using its cognitive cycle, the agent influences its environment by offering a recommendation to the users. The following describes these steps, which correspond to the numbers inside the rectangles in Fig. 1.

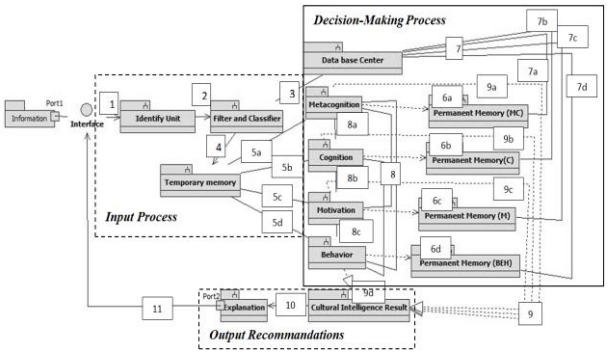


Figure 1. Architecture of CIDSS

Step 1: The business information is in natural language and expresses a problem, a question or a requirement of the user. It is input through the user interface. The information enters the *Identify* module, which identifies the information used to determine what the user requires.

Step 2: The business information goes to the *Filter and Classifier* module. In this module, the information is classified. Useful information is filtered from non-useful information. The useful information is culturally analyzed in the following steps.

Step 3: To perform this classification, the module is associated with the *Cultural Intelligence Database Center*. This center has the necessary data required by the system, such as countries, religions, languages and laws.

Step 4: The classified business cultural data are ready to be sent to the *Temporary Memory* module. This module keeps the data temporarily and, at the same time, interacts with the other modules.

Step5: Modules 5a-*Metacognitive*, 5b-*Congnitive* 5c-*Motiviational* and 5d-*Behavioral* collect the business cultural data belonging to them in the *Temporary Memory*.

Step 6: Each module depends on the consultation of its own *Permanent Memory*. These permanent memory modules are 6a for metacognition, 6b for cognition, 6c for motivation and 6d for behaviour. Each permanent memory represents a complete and specific cultural database that is used by its associated module to analyze the business cultural information stored in the *Temporary Memory*.

Step 7: 7a, 7b, 7c and 7d analyze the business cultural information. If data are missing, *Permanent Memory* modules go to the *Cultural Intelligence Database Center* to assist in the cultural analysis of the respective modules.

Step 8: After the analysis has been completed in each module, the four modules interact with each other to adjust their respective cultural decisions. This interaction enables each module to make a complete and effective decision before continuing to the next step.

Step 9: Following the interaction among the modules of the different dimensions of cultural intelligence, the four modules in steps 9a, 9b, 9c and 9d send their final cultural decisions to the *Cultural Intelligence Result* module. In this module, the decisions of these four modules are generalized and offer significant information to the user.

Step 10: The *Explanation* module justifies and explains in detail using natural language understandable to the user why these decisions were presented.

Step 11: The explanations are sent to the *User Interface*.

V. DESIGNING AND IMPLEMENTING THE CIDSS

A. The Cultural Intelligence Model

Business intelligence generally has two types of data: the first type consists of traditional crisp values, or numbers; the second type is uncertain, incomplete and imprecise. This information is presented in a manner that reflects human thinking and is called "soft data." When we introduce the cultural concept to cross-cultural business activities, we usually use soft information represented by words rather than traditional crisp numbers. The traditional computational technique, known as "hard" computing, is based on Boolean logic and cannot treat business cultural soft data. In order to enable computers to emulate a way of thinking that resembles that of humans, we used a neuro-fuzzy technique to design the CIDSS. This soft-computing technique is capable of operating with uncertain, imprecise and incomplete information. It attempts to model a human-

 ISBN: 978-1-61208-223-3

like understanding of words in decision-making processes.

This hybrid neuro-fuzzy technique makes use of the advantages and power of fuzzy logic and of the artificial neural network. Fuzzy logic and the artificial neural network are complementary paradigms. 1) The *Fuzzy Logic* technique is used for three reasons. First, the cultural intelligence concepts are described in natural language containing ambiguous and imprecise linguistic variables, such as "*this person has low motivation*" and "*that project is highly risky because of this religion.*" Second, fuzzy logic is well-suited to modeling human decision-making processes when dealing with "soft criteria." These processes are based on common sense and may contain vague and ambiguous terms [26]. Third, fuzzy logic provides a wide range of business cultural expressions that can be understood by computers. 2) The *Artificial Neural Network:* Although the fuzzy logic technique has the ability and the means to understand natural language, it offers no mechanism for automatic rule acquisition and adjustment. The artificial neural network presents a viable solution for processing incomplete and imprecise business cultural information. The artificial neural network can manage the new business cultural data input and the generalization of acquired knowledge. The hybrid neuro-fuzzy technique represents the essence of our soft- computing model.

In Fig. 2, we explain the neuro-fuzzy inference model, which is part of the *Main Cycle* in the CIDSS architecture, as shown in Fig. 1. Essentially, the hybrid neuro-fuzzy technique is a neural network with fuzzy inference model capabilities. The network can be trained to develop IF-THEN business cultural fuzzy rules and determine membership functions for input and output variables. The model is represented with a neural network composed of five layers in the CIDSS. It has four inputs: metacognition (MC), cognition (C), motivation (M) and behaviour (BEH), and one output: Cultural Intelligence.

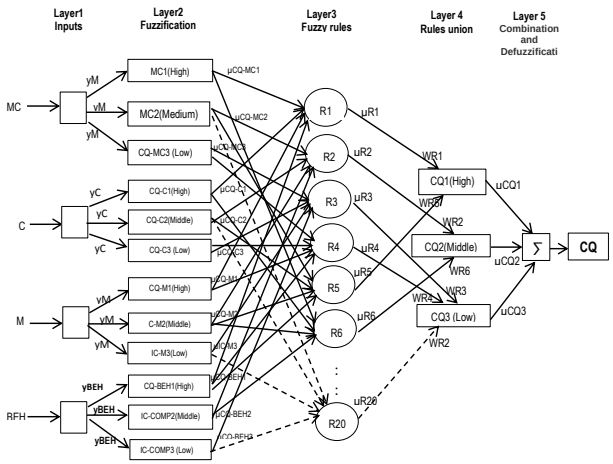


Figure 2. Soft-Computing Model of Cultural Intelligence

Layer 1 - Input: No calculation is made at this layer. Each neuron corresponds to an input cultural variable. These input values are transmitted directly to the next layer.

Layer 2 - Fuzzification: Each neuron corresponds to a business cultural linguistic label (e.g., high, medium and low) associated with one of the input cultural variables in Layer 1. In other words, the connection of the output, which represents the inclusion value, specifies the degree to which the four input cultural values belong to the neuron's fuzzy set. The connection is computed at this layer.

Layer 3 - Fuzzy Rule: The output of a neuron at Layer 3 is the cultural fuzzy rules. Each neuron corresponds to one cultural fuzzy rule. The cultural fuzzy rule neurons receive inputs from Layer 2 (fuzzification neurons), which represent cultural fuzzy sets. For example, neuron R1 represents cultural Rule 1 (Rule 1: *IF metacognition is high AND cognition is high AND motivation is high AND behavior is high THEN Cultural Intelligence is high).* Neuron R1 receives input from the neurons MC1 (*Metacognition High*) and C1 (*Cognition High*). The weights (WR1 to WR20) between layers 3 and 4 are the normalized degree of confidence of the corresponding cultural fuzzy rules. These weights are adjusted when the system is trained.

Layer 4 - Rule Unions (or consequence): This neuron has two main tasks: 1) to combine the new precedent of cultural rules, and 2) to determine the output level (High, Medium and Low). The output level belongs to the cultural linguistic variables. For example, $\mu R1$, $\mu R5$ are the inputs of *Cultural Intelligence High*, and $\mu 1^{(4)}$ is the output of the neuron *Cultural Intelligence High*.

Layer 5 - Combination and Defuzzification: This neuron combines all the consequential rules and computes the crisp output after defuzzification. The composition method "sum-product" [27] is used. It computes the outputs of the membership functions defined by the weighted average of their centroids. We apply, in this case, the triangle calculation in our neuro-fuzzy system, which is the simplest calculation of the fuzzy set as shown in Fig. 3. The calculation formula (see (1)) of the weighted average of the centroids of the clipped fuzzy sets *Cultural Intelligence 3(Low), 2(Medium)* and *1(High)* are calculated.

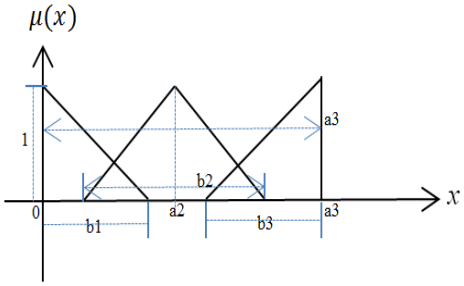


Figure 3. General Cultural Intelligence Fuzzy Sets

$$y\,(Cultural\ Intelligence) = \frac{\frac{1}{3}b_1^2\mu 1 + a_2 b_2 \mu 2 + \left(a_3 - \frac{1}{3}b_3\right) b_3 \mu 3}{b_1\mu 1 + b_2\mu 2 + b_3\mu 3} \quad (1)$$

where a_2 is the center and a_3 is the end of the triangle. b_1, b_2 and b_3 are the widths of fuzzy sets which correspond to *Cultural Intelligence 3*, *2* and *1*.

B. Implementing CIDSS as an Expert System for Cultural Recommendations

We would like the CIDSS to be capable of acquiring, extracting and analyzing the new knowledge of the cultural

 ISBN: 978-1-61208-223-3

experts. First, it should be able to: 1) express knowledge in a form that is easily understood by users, and 2) deal with simple requests in natural language rather than a programming language. Second, the CIDSS should act as would an efficient team of cultural experts who are able to make decisions and provide explanations in the decision-making process in culturally diverse settings. Hence, we integrated the model into an expert system. Fig. 4 illustrates the architecture of the CIDSS.

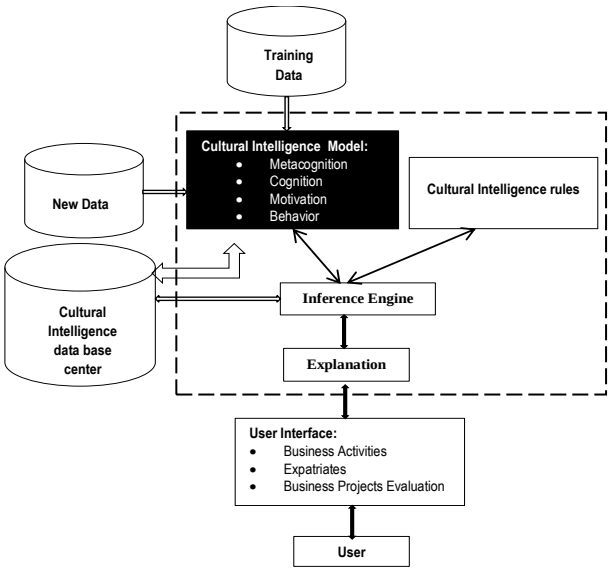


Figure 4. Deployment Structure of CIDSS

The CIDSS has three application domains (Business Activities, Expatriates Assignment and Business Projects Evaluation). The *Cultural Intelligence Model* in this structure is represented by a trained neural network. This structure includes four main modules:

1) The *Cultural Intelligence Model* contains cultural intelligence knowledge that is useful for solving business cultural problems. The soft-computing technique used in this model enables the system to reason and learn in an uncertain, incomplete and imprecise business cultural setting. It supports all the cultural decision-making steps in the system. This module connects with three different units: *New Data*, *Training Data* and the *Cultural Intelligence Database Center*. *New data* include users' requests for solving a given problem that involves cultural business affairs. *Training Data* are a set of training examples. They are used for training the neuro-fuzzy network during the learning phase. *The Cultural Intelligence Database Center* predominantly contributes to the knowledge gathered from the data about different cultural aspects which have been collected from different countries.

2) *The Cultural Intelligence Rules* examines the cultural intelligence neural knowledge base and produces neuronal rules which are implicitly "buried" in the CIDSS network.

3) *The Inference Engine* is the core of the CIDSS. It controls the flow of business cultural information in the system and initiates inference reasoning from the knowledge base in the *Cultural Intelligence model*. It also concludes when the system has reached a decision.

4) *The Explanation* clarifies to the user why and how the CIDSS has achieved the specific business cultural results. These explanations include analyses, advice, conclusions and other facts required for deep reasoning.

The CIDSS possesses generic cultural intelligence and is not specific to a particular culture (such as the United States or China). The system shows great capabilities of cultural adaptation by modeling the human decision-making process in situations characterized by cultural diversity. Furthermore, due to its elaborate cultural schemas and analytical abilities, the system can help users identify and understand key issues in cultural judgment and decision making. It also gives them the corresponding explanations.

Fig. 5, Fig. 6 and Fig. 7 present an example of three outputs of the *Expatriate Assignment* application domain of how the CIDSS can help a user make decisions by taking into consideration his/her inputted request. The CIDSS prototype system follows the decision-making cycle process shown in Fig. 1. The input data are specific business questions in natural language from the users. The system provides two outputs as answers to the question. Output 1 (Fig. 5) gives a general decision to answer the question put by the user.

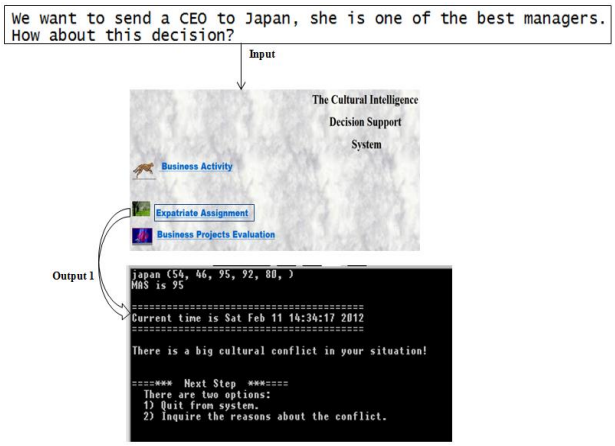


Figure 5. Example of CIDSS Prototype System (Output 1)

Output 2 (Fig. 6) gives more detail explanations which clarify to the user why the system reached this decision.

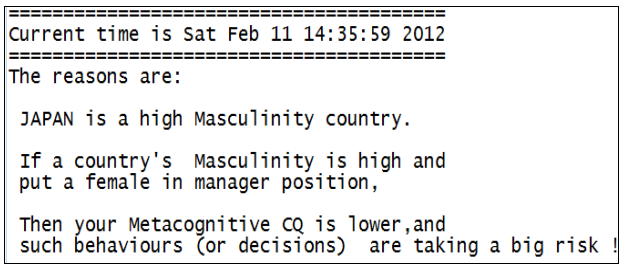


Figure 6. Example of CIDSS (Output 2)

Output 3 (Fig. 7) gives useful suggestions for self-improvement to the user whenever cultural intelligence is required.

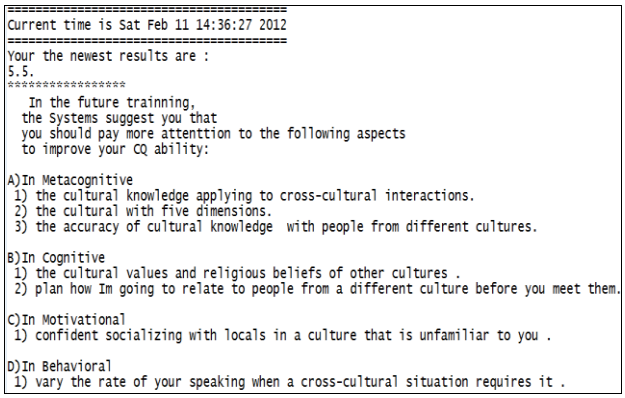


Figure 7. Example of CIDSS (Output 3)

Three cultural experts have validated our computational cultural intelligence model, as well as the CIDSS prototype system. This validation ultimately reflects the consistency between the real world and the artificial CIDSS system. The CIDSS prototype system was also tested with two hundred people. Based on the results of the validation, the cultural experts compared the CIDSS results with their own. These experts concluded that the cross-cultural business decisions recommended by CIDSS are similar to the ones suggested by a human expert.

VI. CONCLUSION

Cultural intelligence is the human ability to capture and reason properly in culturally diverse settings. Cultural intelligence can be measured with four dimensions. Thus, we build a cultural intelligence model based on a soft-computing technique so as to integrate these dimensions and to embody an expert system called the CIDSS. The CIDSS acts as an intelligent expert assistant which helps users make better decisions in cross-culture business activities; it also enables users to solve cultural problems that would otherwise have to be solved by cultural experts. This paper shows how the CIDSS can be used as a "culturally aware" system which assists both individuals and organizations in the decision-making process. It enables users to be more efficient and "intelligent" as they develop their cultural skills. The contribution of our research is, first, to fill that gap between cultural intelligence and artificial intelligence. Second, it improves the application of cultural intelligence theories in the field of cognition. The research focuses on modeling four cultural intelligence dimensions that are an interdependent and integrated body. As a result, the theories are more complete, more efficient, and more precise in their applications. Third, we added to the field of artificial intelligence by computerizing cultural intelligence. As a result, new research topics and directions relevant to this research have arisen, and the range of computational intelligence possibilities, such as soft-computing in business decisions, training and expatriates assignments, has been expanded. Fourth, our research is groundbreaking as it simplifies the work of the researchers by freeing them from heavy, complex, repetitive tasks, which were normally carried out manually in cultural intelligence studies. The algorithms and techniques used in this research may offer some enlightenment as to how they can be applied to other branches of business intelligence in order to improve model design and system performance.

REFERENCES

[1] J.S. Tan, “Cultural intelligence and the global economy,” Leadership in Action, 24, pp.19–21, 2004

[2] G. Zhong Hua and L. Chao Ping, “Research on Cultural Intelligence: Review and Prospect,” Advances in Psychological Science, Vol. 17, No. 1, pp.180–188, 2009

[3] M. Janssens and J. M. Brett, “Cultural intelligence in global teams: A fusion model of collaboration,” Group and Organization Management, 31, pp.124–153, 2006

[4] W. Zhao Xin, N. Roger, and B. Jacqueline, “The Application of AI to Cultural Intelligence, ” The 2012 World Congress in Computer Science Computer Engineering and Applied Computing-ICAI 2012, The International Conference on Artificial Intelligence, Las Vegas, U.S.A. 2012

[5] S. Ang and L. Van Dyne, “Handbook of Cultural Intelligence,” 1st ed. M.E. Sharpe. Armonk . 2008, 2010

[6] P.C. Earler and S. Ang, “Cultural intelligence: Individual interactions across cultures,” Stanford, CA: Stanford University Press, 2003

[7] P.C. Earley and E. Mosakowski, “Cultural intelligence”. Harvard Business Review, 82, pp.139–146, 2004

[8] B. Peterson, “Cultural intelligence: A guide to working with people from other cultures,” Yarmouth. ME: Intercultural Press, 2004

[9] G. Hofstede, “ Cultures and Organizations: Software of the Mind,” McGraw-Hill, New York, 1991

[10] R. Brisling, R. Worthley, and B. MacNab, “Cultural intelligence: understanding behaviors that serve people’s goals,” Group and Organization Management, 2006

[11] D. C. Thomas and K. Inkson, “Cultural Intelligence People Skills for a Global Workforce,” Consulting to Management, 16 (1). pp. 5-9, March, 2005

[12] D. C. Thomas, “Domain and development of cultural intelligence: The importance of mindfulness,” Group and Organization Management, 31, pp.78–99, 2006

[13] K.Y. Ng and P. C. Earley, “Culture + intelligence: Old constructs, new frontiers,” Group and Organization Management, 31, pp. 4–19, 2006

[14] J. P. Johnson, T. Lenartowicz, and S. Apud, “Cross-cultural competence in international business: Toward a definition and a model,” Journal of International Business Studies, 37, pp. 525–543, 2006

[15] S.Ting-Toomey, “Communicating across cultures,” NewYork. Guilford, 1999

[16] J.S. Robert and K.D. Douglas, “What is Intelligence?,” Contemporary Viewpoints on Its nature and definition, Ablex publishing corporation, 355 Chestnut Street, Norwood, NewJersey 07648, 1986

[17] S. Ang, L. Van Dyne, C.K.S. Koh, K. Y. Ng, K. J. Templer, C. Tay, and N. A. Chandrasekar, “Cultural Intelligence: Its Measurement and Effects on Cultural Judgment and Decision Making, Cultural Adaptation, and Task Performance,” Management and Organization Review 3 335–371, 2007

[18] L. Ottavi, “Cultural Intelligence in the Globalized Work Environment, ” December 1, 2009

[19] G. Hofstede, “The cultural relativity of organizational practices and theories,” Journal of International Business Studies, 1983

 ISBN: 978-1-61208-223-3

[20] F. Trompenaars and C. Hampden-Turner, "Riding the waves of culture: Understanding cultural diversity in business," London: Nicholas Brealey Publishing, 1997

[21] T. Kern and L.P. Willcocks, "Exploring information technology outsourcing relationships: Theory and practice," Journal of Strategic Information Systems, 2000

[22] S. Ang, L. Van Dyne, C. Koh, and K.Y. Ng, "The Measurement of Cultural Intelligence," Paper presented at the 2004 Academy of Management Meetings Symposium on Cultural Intelligence in the 21st Century, New Orleans, L.A. 2004

[23] G. Huber, "A theory of the effects of advanced information technologies on organizational design, intelligence, and decision-making, " Academy of Management Review, 1990

[24] S. Ang and C. I. Andrew, "Cultural Intelligence and Offshore Outsourcing Success: A Framework of Firm-Level Intercultural Capability," Journal compilation, Decision Sciences Institute, Decision Sciences Volume 39, Number 3, 2008

[25] D. Dubois, "Réalisation d'un agent doté d'une conscience artificielle: application à un système tutoriel intellgent," Montreal, University of Quebec in Montreal, 2007

[26] M. Negnevitsky, "Artificial Intelligence: A Guide to Intelligent Systems," ISBN 0-321-20466-2, British Library Cataloguing-in-Publication Data, 2nd Edition, 2005

[27] J.S.R. Jang, C.T. Sun, E. Mizutani, "Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence," Prentice Hall, Englewood Cliffs, NJ 1997

 ISBN: 978-1-61208-223-3

Using Network-based Business Process Analysis as a Tool for Delta Analysis of Process Models

Olga Levina
Department of Systems Analysis and IT
Berlin Institute of Technology
Olga.levina@sysedv.tu-berlin.de

Abstract- Business process models are often crucial for business software development. Using reference models or merged processes, software producers intend to optimize the development effort for enterprises of the same domain. Thus, it is important to know how to assess the similarity between the process and the reference or merged process model to assess the effort needed for their adaption. This paper uses network-based business process analysis approach to quantitatively assess process similarity based on information structure of the business process model. Application of the approach as well as metrics used for similarity assessment is presented using a case study from the health management domain.

Keywords- Business process analysis; process merging; network analysis.

I. INTRODUCTION

Merging organizations, a new business unit created from existing business departments, a common information system for different business units or domains, all these aspects require merging of processes. In each of these cases similarities and differences between business sub-processes have to be identified and the differences must be resolved. Thus, before process merging it is necessary to assess the effort that will be needed changing them. This evaluation need is similar in cases when organization's processes are evaluated towards their compliance with standard processes.

Process merging is defined by [1] as the proceeding when "two or more process models have to be brought together in order to create an improved business process". Solutions relying on information technology (IT) are often derived based on process models as they provide an initial insight into the business procedures. Process merging is a technique that is similar to process configuration but it requires the preservation of certain parts of the processes in focus as well as the preservation of the underlying structure of the IT systems.

In this paper, the process merging technique is used to derive a common consignation process for two hospitals that need to be supported by a common information system. Thus, process merging is applied here in the context of business driven software development including business process analysis and enhancement. The research question here is, how to determine similarity of the as-is processes with regard to the to-be or merged process. Answers to this question will provide the enterprises, or in this case hospitals, as well as business software developers with an estimate on the effort as well as potential adaptation of the developed tool.

The question on process model similarity has already been addressed in research from different angles, e.g., by [2]. Often, compliance with reference models or semantic compliance is in the focus of these activities. In this paper, quantitative analysis and comparison of process information structure in the process model is performed using metrics derived from network analysis. This approach allows repeatable and more objective results. Furthermore, the question of the defining dimensions of the fit of a particular process to the merged, reference or improved process arises and is also addressed using network-based business process metrics. Externalization of these aspects supports implementation of a new IT support tool and process by providing comparable and objectified results that otherwise had to be derived in expert interviews or a complementary process analysis.

Being based in the information system research this paper presents the use of an approach for quantitative assessment of business process characteristics and inter-process comparison using network-based business process analysis [3]. This approach is based on the view of a business process as a network, with process activities as nodes and information and message flows as links. It provides insights for theory and practice. From a practical point of view developing application systems or software in a specific industry requires elevation and implementation of similar processes. This is a rather cumbersome and resource consuming process. Assessment of similarity and differences between processes of the same domain but in different enterprises helps to estimate the effort needed for their implementation or change. Other techniques to measure process differences, e.g., process simulation, require more data than comprised in the process model, e.g., exact data on activity time and costs, which also needs to be elevated.

In this paper, structural and communicational effects induced by the new merged process are assessed using the analysis of the to-be process model. A consignment process from two different hospitals is analyzed with the goal to develop supporting software for its execution. These processes are merged to provide an improved and more general process that can be integrated into the two hospitals.

 ISBN: 978-1-61208-223-3

The merged (to-be) process has been presented and discussed with the process managers and owners from the two hospitals who approved the feasibility and design of the resulting process. Network-based business process analysis is used here to quantitatively assess the differences and similarities of the two processes as well as the improvement effects and fit of the merged process providing comparable and repeatable results. The fit of an as-is process to the merged process is measured here using the difference between metrics of the both process networks.

The paper unfolds as follows. In the following Section related work on process similarity assessment is reviewed. The case study and the processes are presented in Section 3. Section 4 briefly reviews the network-based business process analysis approach, while the analysis results are presented and discussed in Section 5. Conclusion and outlook finish the paper.

II. RELATED WORK

Research on merged processes intersects several areas of business process management such as reference modeling and process comparison.

Business process reference models are standard business process models for a specific business domain or industry that can be tailored to the needs of a specific company. Reference models can be tailored by describing their configuration options [2, 4]. These are patterns that pre-define what differences can exist between possible users. These aspects are also used in the research on process differences and their classification by [5].

Process merging was introduced by [1] as a variation of adaption of business processes to reference models by establishing correspondence between the two process models. The authors discuss process merging in the context of business driven software development [6]. Business driven development is a methodology for developing IT solutions that directly satisfy business requirements. Therefore, it strongly relies on the information captured in the business process model.

Dijkman develops patterns for business process model differences as well as techniques to identify the differences by using a detection algorithm that was developed by formalizing frequently occurring differences discovered in practice in [5, 7]. Van der Aalst [2] uses process mining to detect the conformance of an executed business process with its descriptive model. Thus, he (quantitatively) compares the real behavior of a business process tracked in logs of the executing information system with the intended or expected behavior.

Rozinat and van der Aalst [8] operationalize the fitness of the prescriptive and actually executed business models also using the process mining technique. The authors define process fitness as "[the] extent to which the log traces can be associated with valid execution paths specified by the process model". Their approach allows detecting conformance problems in real-life scenarios. Nevertheless, the authors do not provide a possibility for quantitative assessment of the detected problems.

Gerke et al. [9] developed an algorithm, which allows measuring the compliance, i.e., the degree to which a process model behaves in accordance to a reference model, of process models with reference models using the process structure. They developed an algorithm (using data obtained by process mining), called sequence- based compliance, which is based on the analyzing task sequences and comparing them to accordant sequences in the reference model.

Described research activities approach process similarity on the level of semantics or tasks. The approach described in the next sections analyzes the process structure and derives accordant metrics.

III. CASE STUDY

The case study presented here in situated in the domain of health management. A consignment process in two hospitals (KHA and KHB) is in the focus of analysis. The case study is based on a process analysis project. The goal of the project was to analyze and subsequently improve the processes as well as to design a general consignment process that can be supported by an accordant software tool in both hospitals. Actors involved in the process (see Figures 1 and 2) are the hospital, represented by ward, procurement and finance departments, a medical worker (in ward sub-process of KHB), an IT-system as well as a supplier. The goal of the process is to purchase the needed medical equipment and supplies and deliver them to the requesting ward. Due to the confidential agreement as well as the focus on structural analysis focus the process is only sketched here.

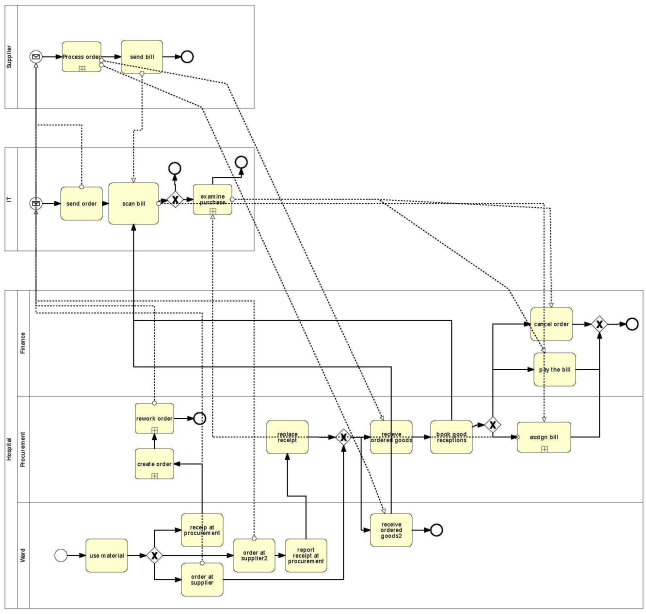


Figure 1. As-is process-KHA

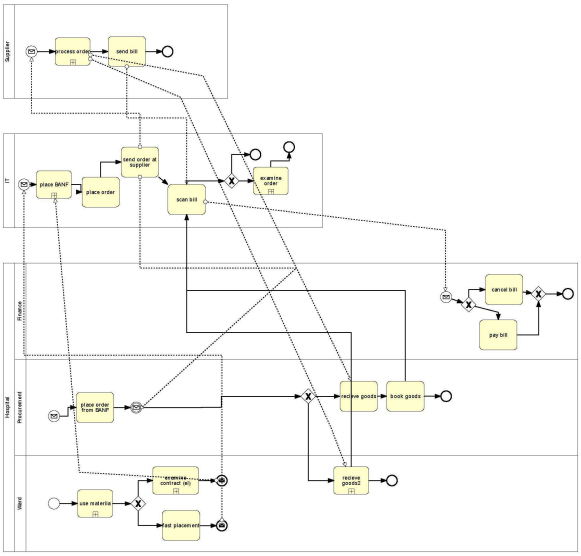


Figure 2. As-is process- KHB

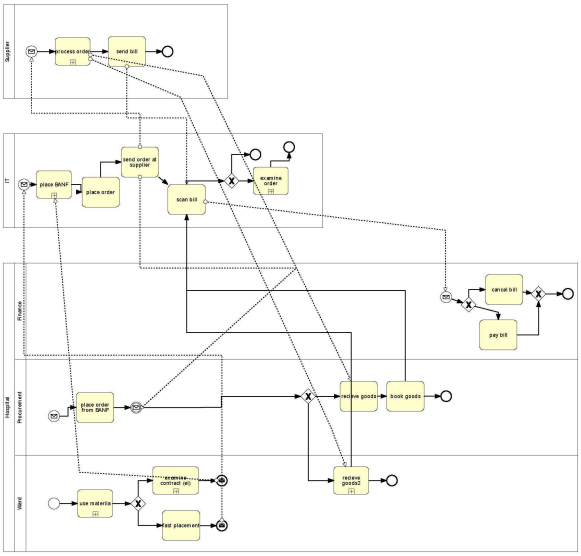


Figure 3. Merged process: KHPlanned

Processes from the two hospitals slightly differentiate in their use of the IT-system as well as involvement of the procurement department. These differences are partly due to the local circumstances, e.g., staff capacity, workload, etc., but the processes also show potential for workflow automation and involvement of the medical personnel. Hence, the processes were analyzed and merged to a more general and optimized process.

The merged process was created on the basis of the two as-is processes but also as a results of the process analysis for process improvement (see Figure 3). Thus, its content and structure reflects the as-is processes but also integrates the suggested process improvements. Improvement efforts were directed towards a higher automation degree of the process, i.e., a higher involvement of IT system for document and collaboration management, as well as towards a more efficient management of medical material to ensure its fast and timely delivery. Nevertheless, performance measurement of the merged process will not be in focus of this paper. Here, the structure of each of the two as-is processes is compared to the structure of the merged process using network metrics described in the following section. Based on these metrics, the fit between as-is and the merged process is detected enabling assessment of changes introduced by the new process. In the context of software development this measurement enables an estimation of the effort that will be needed for the case dependent software implementation.

IV. NETWORK-BASED BUSINESS PROCESS ANALYSIS

To address the research question on how to quantitatively measure process similarity using a business process model, network-based business process analysis as presented in [3] was chosen. This approach allows analysis of process structure and its characteristics using network metrics. Thus, business processes under analysis were transformed into accordant business process networks with business process activities as nodes and control as well as information flows as links. To analyze dynamic process characteristics, the notion of linkevent has been introduced. A linkevent is a relation between one or more nodes. Each linkevent can have one sender as well as no, one or more than one recipient(s) [10].

Network metrics used for the analysis have been calculated using Commetrix® network analysis software and inserted into the discriminant functions and metrics (see Table 1 and 2) to define the process type and activity roles (see [3] for details on process definitions and approach).

Process Type	Discriminant function	Value interval
Core	D_c=8.894-6.447LS+0.144conn-0.034reach -0.08PL-0.016CluCo	<0
Automatable	D_a=0.514-5.923LS +0.146conn +0.04reach-0.08PL-0.006CluCo	>0
Distributed	D_v=-21.016 +18.839LS + 0.022conn +0.028reach - 0.338PL - 0.05CluCo	<0
Information intensive	D_i=-10.421 +4.473LS - 0.106conn +0.06reach+ 0.204PL+0.03CluCo	>0
Decision intensive	D_e=-20.554 +16.805LS - 0.081conn + 0.045reach +0.104PL +0.062CluCo	>0
Communication intensive	LE/n	>1
Event intensive	L/n	>1.2

TABLE I. DISCRIMINANT FUNCTIONS: OVERVIEW

Following process types are distinguished here and can be identified using discriminant functions and metrics from Table 1: core, automatable, distributed, information, decision, communication and event intensive processes. Core processes are defined here as value-adding processes

 ISBN: 978-1-61208-223-3

therefore involving direct customer interaction. A core process model includes an internal or external process customer. Automatable processes are defined here as repetitive, predictable [11], and indicating low level of variance, i.e., decisions or exceptions, in the control flow. Level of abstraction of the modeled process needs to be considered here, as processes presented on a high level of detail or containing multiple sub-processes can be considered as non-automatable in their full length. Information intensive processes are defined here as processes, which activities use information as a main resource, implying frequent information exchange within the process flow. A decision in this context can be identified as a choice between several process variants rather than a choice between two alternatives [12]. Thus, decision intensive processes were identified as processes containing an increased number of operative decisions [13]. Distributed processes were defined here as processes, which sub-processes are executed by different actors situated in different geographical locations. Communication intensive processes are defined as processes that require increased cooperation and communication activity from the process actors. To assess this process property, number of interrelation activities, i.e., linkevents LE, is normalized by the number of process activities n. Event intensive processes are processes that are defined by the state changes of business objects, i.e., the state of the business object defines the process flows. This characteristic can be considered as a feature of process complexity. Thus, the event intensiveness is assessed here similar to the coefficient of network complexity as defined by [14] by the quotient of links L, i.e., control and message flows within the process, and nodes n, number of process activities.

Discriminant functions in Table 1 include following network metrics as variables defining behavioral characteristics of a process: average path length (PL), average clustering coefficient (CluCo), average connectivity and reach. Average path length indicates the average number of steps along the shortest paths (geodesic) for all possible pairs of network nodes, while reach denotes the degree of any member of the network, to which it can reach other network members. PL can be interpreted as a measure for the efficiency of information transport within the network. Connectivity (conn) of a network indicates how many nodes need to be removed to separate the network in several groups [15]. Clustering coefficient measures the interaction of nodes within an ego-network including transitive connections and indicates the transitivity of the node, i.e., its ability to distribute information directly with its neighbor nodes. Link strength (LS) is a measure based on the concept of linkevents and denotes how many linkevents have been exchanged on one link weighting the relationships between two nodes. Density is defined as the ratio of links present in the network and the maximum number of possible links [15]. Thus, it can also be used to refer to the stability of the network with respect to structural changes. Sparse density can be observed in communication networks as not all nodes are connected with each other [16]. Closeness centrality metric calculates the sum of the distance between a specific node and every other node in the network [17]. Thus, its value is inverse proportional to the distance between related nodes. The average closeness centrality of a network can provide insight on the collaboration and information distribution productivity within the network [18].

Business process activity roles were also analyzed here using the network-based business process analysis approach. These roles are: information sink and source, process sink and control activity. An information sink is a process activity that requires an increased amount of information for its execution. This role is defined using the number of received linkevents per node. Activities with an amount of received linkevents that is found in the third quartile of received linkevents is considered as an information sink. An activity that delivers information to other activities is considered as information source. These are activities from the third quartile according to number of sent linkevents. A process sink is an activity that is considered as the process goal, i.e., the activity that captures the control and information flows of the process. This is the activity with the highest number of linkevents received in the process. Control activities are identified using the metric of betweenness centrality [15]. Betweenness centrality indicates a node that lies in between of other pair of nodes [19], i.e., the node with the highest number of shortest paths passing from all vertices to all others through that node.

V. APPLYING NETWORK-BASED PROCESS ANALYSIS

Network-based business process analysis approach is used here to assess the changes in the process structure of the merged process as opposed to the as-is process structure. The goal is to provide process managers of KHA and KHB processes as well as software developers with a possibility to evaluate the future effort for process adjustement. Additionally, the use of (network) metrics allows a comparison between the as-is and merged process to show the impact of the improvement measures providing performance related insights to the process managers. Thus, process owners can assess whether the merged process enhances the process characteristics and in what ways. First, the network analysis results are presented and then discussed in this section.

A. Results Overview

Using the network-based business process analysis approach process network metrics have been derived (see Table 2) and allow detection of the differences between the process in the two hospitals on process as well as activity levels. Thus, network metrics derived from the process models that were transformed into networks were inserted into the discriminant functions in Table 1 to define and compare the process types and informational characteristics of the processes. Further metrics as described in Section 4 were used to define the roles of process activities in the information transportation within the process. Additionally, communication load of the process participants has been investigated using the count metric of the linkevents (see Figure 4).

 ISBN: 978-1-61208-223-3

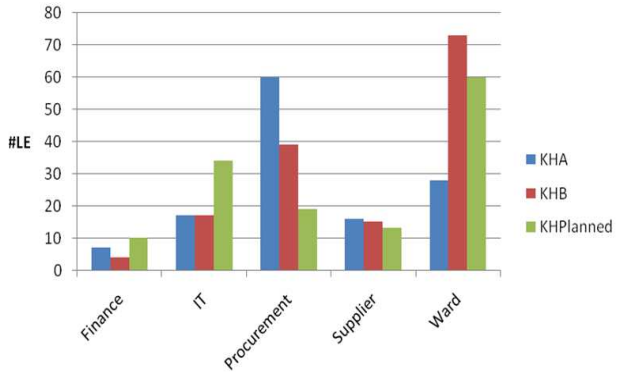


Figure 4. Activity of process actors

Analysis of the process types showed that KHB process is more automated, more distributed, more information and decision intensive than the KHA process according to the results of the discrimination functions in Table 2. It is less customer-oriented (measured by using the core process type definition) with the hospital as customer.

	KHA	KHB	KHPlanned	dKHA	dKHB
# Nodes	47	55	58	-11	-3
# Linkevents	64	74	68	-4	6
# Links	63	73	68	-5	5
Diameter	10	15	25	-15	-10
Link strength	1.02	1.01	1	0.02	0.01
Density	5.83	4.92	4.11	1.72	0.81
avg. Degree Centr.	5.83	4.92	4.11	1.72	0.81
avg. Betweeness Centr.	7.07	8.07	14.12	-7.05	-6.05
avg. Closeness Centr.	30.23	26.01	20.29	9.94	5.72
avg. Connectivity	0.92	3.26	11.88	-10.96	-8.62
avg. Reach	100	100	100	0	0
avg. Path Length	4.32	5.44	9.19	-4.87	-3.75
Clustering Coefficient	4.08	9.8	4.76	-0.68	5.04
L/n	1.34	1.33	1.17	0.17	0.15
LE/n	1.36	1.35	1.17		
Automatable	-1.71	-1.37	-0.38		
Core	-1.36	-1.14	-0.05		
Information intensive	1.05	1.15	0.81		
Decision intensive	-1.36	-1.14	-0.05		

TABLE II. PROCESS ANALYSIS RESULTS: OVERVIEW

Process sink, i.e., process goal of the KHA and KHB is the "search for the material" activity executed by the procurement department. Control and core activities in KHA process is "scan the bill" and in KHB "send the order to supplier", both executed by the IT tool. Figure 4 shows the number of linkevents executed by each process participant documenting their communication and information exchange activity. In the hospital KHA procurement is the actor with the major load of communication activity in the process while in KHB it is the ward. The IT tool is only sparsely involved in communication, supporting the identified process type definition as non-automatable.

The to-be process for both hospitals (KHPlanned) is also analyzed using network-based process analysis. It is less complex than both of the as-is processes according to the increased diameter of the process network. Its increased average path length indicates less effective information transportation as well as the fact that the process is now more distributed according to the results of the D_v function.

The value of D_a for the to-be process suggests that the process can be broadly supported by IT. It is also less information and decision intensive than both of the as-is processes. Figure 3 also indicates that the communication activity is now shifted towards IT and the ward (for the KHA process). The main process activities remain the same comparing to the KHA and KHB processes.

Delta analysis of the KHA and KHPlanned as well as KHB and KHPlanned processes is performed using the derived network metrics and is summarized in the \triangle metric in Table 2. This metric can also be used to assess the fit of the process structures to the planned process. \triangle indicates higher values for KHA and KHPlanned comparison than for KHB and KHPlanned comparison. KHA metrics are significantly different comparing to the KHPlanned metrics on the level of significance of α=10%, while KHB metrics are not significantly different.

B. Discussion of the Results

Comparing the network metrics shows that KHA process is less information and decision intensive as well as smaller in terms of number of nodes, links and linkevents comparing to the KHB process, suggesting a lower need of IT support involvement. It also shows a highly sensitive structure in terms of process changes, as indicated by the value of the average connectivity metric. Productive information communication is indicated in KHA by a high value of the average closeness centrality and density indicating a communication network, i.e.,, message-based communication between process tasks and actors. The smaller clustering coefficient value than the one of the KHB process suggests a smaller tendency of the actors to share information directly. Analysis of the KHB process shows that the ward personnel is strongly involved within the consignation process (see Figure 4). This activity restrains the capacity of the personnel to involve in their medical tasks.

KHPlanned resulted as being a process that has a higher automation potential than the two as-is processes being more distributed as well as less communication intensive, as indicated by the decreased density as well as event intensity and complexity metrics. More distributed process with longer paths requires a stronger focus on information management by e.g., implementation of a document management system with access for all process participants. Communication activity analysis shows that on one hand the communication activity has been partly shifted to IT (indicating potentially faster response times) but also the fact that the communication load of the ward personnel increased for KHA and slightly decreased for KHB. This aspect needs to be analyzed in more detail to ensure that the (increased) involvement of the personnel is necessary and complies with their primer duty.

The comparison of the network metrics showed that the to-be process merges and extends the characteristics of the

 ISBN: 978-1-61208-223-3

two as-is processes. This fact suggests that the to-be process has been extended and improved to accommodate the draw-backs of both processes. Communication analysis indicates that finance and procurement departments will need to adjust more to the new process. Density metric of the planned process suggests a communication network but also a loss in the potential for direct information distribution according to the only slightly higher clustering coefficient value comparing with the KHA process.

Significant difference in the \triangle values indicates that the KHA process will need more effort for its adjustment to the merged process than the KHB process. Increased connectivity metric of the KHPlanned process indicates a less sensitive structure in case of process changes. Also the planned process might face a decrease in information accuracy as indicated by the decline of the average closeness centrality metric [18].

As the communication activities were now shifted towards the IT system, statements suggested by these metrics need to be examined further using the actual performance related data to assess the actual gain or loss in performance related process indicators. Nevertheless, the trend shown by the metrics needs to be taken into account by process owner and manager. These results derived using network-based process analysis focus on the (informational) structure changes in the process without taking time or costs implications of the new process into account. These aspects need to be evaluated using further analysis techniques, e.g., business process simulation.

VI. CONCLUSION AND OUTLOOK

In this paper metrics derived using network-based business process analysis approach have been applied to assess structural process characteristics and therefore the delta between as-is and merged processes. This analysis allows valuable insights for process managers, owners and IT tool developers as they enable them to assess the induced process changes on one hand and the potential adjustment effort on the other hand.

Presented approach for business process comparison can be applied on processes from different business domains to compare merged, as-is, to-be or reference processes on the level of information structure using their process models. Its application has been shown on a case study from the health management domain. To enrich the results with further insights, a performance comparison based on business process simulation including costs and time analysis can be added.

The goal of the paper was to present the application of the quantitative comparison approach. Nevertheless, it will need further evaluation in future, more exhaustive case studies.

VII. REFERENCES

[1] Kuester, J., Koehler, J. and Ryndina, K. Improving Business Process Models with Reference Models in Business-Driven Development in BPM 2006 Workshops. 2006.

[2] van der Aalst, W., 2005, 10, 198 -211, *Business Alignment: Using Process Mining as a Tool for Delta Analysis and Conformance Testing* Requirements Engineering, 2005. **10**: pp. 198-211.

[3] Levina, O., Hillmann, R. *Network-based Business Process Analysis* in *45th Hawaii International Conference on System Sciences*. 2012.

[4] Rosemann, M. and van der Aalst, W., *A Configurable Reference Modelling Language.* Information Systems, 2007. **32**: pp. 1 -23.

[5] Dijkman, R. *A Classification of Differences between Similar Business Processes*. in *11th IEEE International Enterprise Distributed Object Computing Conference*. 2007.

[6] Mitra, T., *Business-driven development.* IBM developerWorks, 2005.

[7] Dijkman, R. *Diagnosing Differences between Business Process Models* in *Business Process Management.* 2008.

[8] Rozinat, A. and van der Aalst, W., *Conformance checking of processes based on monitoring real behavior*. Information Systems, 2008. **33**: pp. 64 -95.

[9] Gerke, K., Cardoso, J. and Claus, A. *On the Move to Meaningful Internet Systems*. in *OTM 2009*. 2009.

[10] Bender-Demoll, S. and McFarland, D., *The art and science of dynamic network visualization.* Journal of Social Structure 2006. **7**(2): pp. 1-38.

[11] Georgakopoulos, D., Hornick, M. and Sheth, A., *An overview of workflow management: From process modeling to workflow automation infrastructure.* Distributed and Parallel Databases, 1995. **3**: pp. 119-153.

[12] Bromberg, D., *BPM for Knowledge Workers Inside Decision Intensive Processes (DIPs): Knowledge, Practice, Context, and Characteristics*, in *BPTrends*. 2007.

[13] Levina, O., *Design of Decision Service Using Cause-and-Effect Business Process Analysis.* Computer Science- Scientific Journal of Riga Technical University, 2011. **46**: pp. 19-27.

[14] Cardoso, J., Mendling, J., Neuman, J. and Reijers, H.A. *A discourse on complexity of process models*. in *BPM 2006 workshops*. 2006: Springer.

[15] Wasserman, S. and Faust, K., *Social Network Analysis - Methods and Applications*. 18. ed. 2009: Cambridge University Press.

[16] Trier, M., *IT-supported Visualization and Evaluation of Virtual Knowledge Communities*. 2005, TU Berlin: Berlin.

[17] Xu, J. and Chen, H., *Criminal network analysis and visualization.* Communications of the ACM, 2005. **48**: pp. 100- 107.

[18] Okamoto, K., Chen, W. and Li, X.-Y., *Frontiers in Algorithmics* in *Ranking of Closeness Centrality for Large-Scale Social Networks*. 2008, Springer. pp. 186-195.

[19] Jamali, M. and Abolhassani, H. *Different Aspects of Social Network Analysis.* in *2006 IEEWICIACM International Conference on Web Intelligence*. 2006: IEEE.

 ISBN: 978-1-61208-223-3

Using Business Process Simulation to Assess the Effect of Business Rules Automation

Olga Levina
Department of Systems Analysis and IT
Berlin Institute of Technology
Olga.levina@sysedv.tu-berlin.de

Abstract-- Application of the business rules approach is often argued in research and practice to result in more efficient and enhanced processes. As the extent of this statement has yet to be shown, in this paper the process simulation method is applied to assess the effect of business rules externalization and automation on business process performance. A real-world decision and information intensive process is used here as a case study. Business process simulation is applied here to assess process performance difference between as-is and enhanced to-be process. Comparison of the obtained simulation results of the re-engineered to-be process that includes automated business rules and the as-is process indicate an increase of process performance measured in time and costs while attributing this effect to the automation of identified business rules.

Keywords- Business rules; business process management; business process simulation.

I. INTRODUCTION

This paper addresses the question put in [1] on how automation of business rules changes business process performance and extends its methodology by using a more extensive case study, multiple performance indicators and business process simulation.

Business process simulation has gained a lot of attention in the context of business process re-engineering in 1990ies as a tool to support managerial decision and illustrate potential effects of a newly designed business process without significant financial effort. Business process simulation aims at assisting the process of modeling and analyzing organizational structures. Use of simulation in the context of business process re-engineering is based on the approach to computer-aided analysis expressed by Simon [2]. He argues that one of the most important uses of computers is "to model complex situations and to infer the consequences of alternative decisions to overcome bounded human rationality". Being an operational research technique simulation has a major advantage as it allows experimentation with any element of a business system [3] being used in order to measure, understand and predict the metrics of process improvement and quality [4]. Thus, one of the major application areas for simulation is the exploration of effects introduced support by the information technology (IT) will have on the process performance (see e.g., [4]).

In this paper, the effect of business rules automation on defined process performance indicators is measured using a case study: a real-life business process of a health fund. Business rules automation implies rule identification, e.g., from software code or expert knowledge, their externalization, i.e., documentation and description in machine-readable format, and automated execution, i.e., enactment within a business rules management system or rule engine of a workflow management system. This research is based on the discussion in the area of business process management stating that IT-supported business rules management enables a better process management in terms of maintainability [5] as well as performance and efficiency [5]. Thus, the research question here is: can the impact of business rules automation on business process performance be captured and if yes, is this impact positive or negative on pre-defined business process indicators. The research method applied here is a case study-based simulation of a real-life business process. The effect of business rules automation is measured by comparing metrics from the as-is and to-be process that integrates business rules automation. This research method is used for artifact evaluation purposes in design science research [6] and business process re-engineering as well as hypothesis or theory testing in social sciences. Here it is applied in the area of assessment of business process automation and thus, re-engineering.

The remainder of the paper is structured as follows: first a short introduction into the business rules and business process management context is provided. In Section 3, a short overview of related work in business process management is given. Simulation experiment design as well as the case study process are described in Section 4. Section 5 provides an overview and discussion of the simulation results. Conclusion and outlook finish the paper.

II. BUSINESS RULES AND BUSINESS RULES MANAGEMENT

Business rules are defined by the Business Rules Group as "constrains [....] that guide or limit the business" [7]. Business rules are specific instantiations of rules and capture process and enterprise specific knowledge. Business rules are often categorized in different types such as, e.g., constraints, derivations or process rules [7-10]. Constraints are statements concerning the enterprise or its philosophy that need to be always true [7, 11]. Derivation rules are statements that are derived from other information or given facts, e.g., calculations. Process rules describe the process logic, i.e., sequence of process steps execution, and are often

 ISBN: 978-1-61208-223-3

known as ECA (event, condition, action) rules. One of the main concerns in business rules management is their extraction and identification from the business process context. Here, the process by [12] is described and applied for business rules extraction from a business process model.

First, the area of analysis, i.e., the scope of business rules extraction, needs to be defined. Here, the area is well limited being the self-contained business process model. Further, identifying the goal of analysis helps the evaluation of the importance of involved potential business objects that can be business rules' sources or sinks. For the given experiment, the goal is to capture all business rules involved for efficient and accurate process execution using the existing business process model. The goal of this process can be identified as providing an optimal care for the patient from medical, financial and temporal points of view.

Further, a mandatory business vocabulary of the terms used in the analyzed context needs to be established. Process- guiding rules can often be identified by the use of a gateway in a modeled process flow, which can be seen as the first step for business rules identification [13]. Questioning the implicit (or explicit) pre- and post conditions of a process steps realization can also provide a indication on business rules that are needed to achieve successful automated process execution. For identification of further potential business rules sources pre- and post-conditions for the process goal achievement can be helpful. Next step of the process suggests the identification of the process stakeholders. Possible restrictions coming from the stakeholders need to be specified and their impact on the process or analysis goal evaluated. An important structuring criterion for found business rules candidates is their affiliation to the goal of the analysis. Stakeholders' restrictions can be summarized as expectation of: maximum transparency (including clear and comprehensible decision guidelines), minimal possible processing time (maximum possible process automation), maximum benefit (including clear accordance of requirements and possibilities), maximal possible accuracy (meaning clear decision guidelines and clear requirements assessment).

Business rules are especially useful in business domains with high decision and policy intensiveness because they enable expressing, managing and updating pieces of business knowledge [14]. As the definition of the process type requires significant analysis effort, Levina and Hillmann [15] show a quantitative approach for business process type identification. In the context of business rules management, information and collaboration intensive processes are in the focus of interest as they would benefit from business rules-oriented management approach.

The business rules management approach [10, 16] is often seen as a part of business process management [17] and is also intended as an enabler of process improvement. Among the main reasons for integration of business rules management following business domains are cited by practitioners and researchers: decision automation to reduce process complexity and transaction costs of and decision errors within a process task; process flexibility due to the ability to update or extend the business logic using a business rules management system; spreading and harmonizing business policies and agreements throughout the enterprise [18-20]. Nevertheless, the claimed benefits of the business rules approach are rarely empirically investigated in research. The major part of business rules management research is focused on structural and operational construction of business rules for accurate response to changes in business environment [21] or is often related to complex event processing.

In this paper, process cycle costs and time are analyzed aiming at the investigation of the effect of business rules automation on process efficiency by comparing performance indicator values of the as-is and to-be process. Therefore, a real-life business process in the as-is state as well as with automated business rules is simulated using a business process modeling and simulation software. The assumption here is that business rules and workflow management systems create an integrated environment for process execution.

III. RELATED WORK

Vendors of business process and business rules management solutions often provide case studies on the positive effect of business rules automation based on their customer's experience (see e.g., [5]). Nevertheless, this data cannot be rigorously evaluated, as these case studies are described on an abstract level. Furthermore, data collection methods and analysis are neither discussed, nor presented in the corresponding publications. Thus, these case studies cannot be considered in full in this state-of-the-art review of current research on business rules management assessment.

In the research area of business process management, business process simulation has been widely used for business process re-engineering [4, 22, 23] as well as organizational design [3] and quality assessment [24] among other domains. Simulation is furthermore a component of a business rules engine, thus, business rules can be simulated to provide a better robustness and validate the process and rule design.

Research on business rules often focuses on business rule description [17], data management aspects and technical implementation [21, 25, 26]. Only little research has been done in the area of business rules capturing and extraction methods [12, 19, 27, 28] as well as evaluation of the business process management approach on business process performance. Levina et al. [1] explore the impact of business rules automation in the e-business domain using complex event processing, i.e., event-action/process business rules. The indicator used to show the advantage of business rules in [12] is process cycle time. The case study also focuses on a short business process that involves only few decisions.

In this paper, a case study that requires extended process knowledge as well as frequent decision taking is analyzed. The effects of business rules automation are measured using process costs and time indicators. Next section describes the experiment design.

 ISBN: 978-1-61208-223-3

IV. EXPERIMENT DESIGN

To assess the business rules automation effect, a simulation of a business process from a German health fund has been performed. The process was elcited using interviews and modeled using Business Process Modeling Notation (BPMN) as well as enriched with detailed information on its organizational, temporal and financial structure. This proceeding provided a way to achieve an understanding about the current and future potential design of organizational processes without the risk of disruption to the real system itself.

Derived information as well as the business processes have been modeled using Adonis® 2.01 community edition software. This open source tool allows BPMN modeling as well as process simulation. To derive the effect of business rules automation on the business process, the process has been modeled and simulated using two different scenarios. In the first scenario, the as-is process described in interviews has been represented. In the second scenario, identified business rules have been implemented in the model. Hereby, process rules as well as constrain rules have been modeled as automated tasks, while rule-centered information gathering and management have been implemented in the first scenario as manual tasks. Duration of the simulation was a year in working days with five working days per week and 8 hours per week per day for both scenarios. Duration of automated tasks has been set corresponding to the values elevated in interviews. Activity costs were assigned in accordance with process actors' levels of expertise. Activities requiring deeper medical knowledge were assigned higher costs comparing to less knowledge intensive activities.

A. Methodology

A case-based research methodology was chosen with the aim to provide an example of practice and test the proposition that the supporting tools of process mapping and business process simulation can illustrate the effect of a rule-based approach to change in the context of business process management.

Although, a single-site study has obvious limitations with respect to the generalizability of the findings, the case is not aimed at being representative, but rather exemplary. Thus, the simulation results that were retrieved using the simulation are not exhaustively representative of all similar situations (same reasoning has been successfully applied in [29]). The study is used here to explore the potential effects of business rules externalization and automation on business process performance. Therefore, it is an explanatory example and an exploration study which simulation results indicate the effect of business rules automation. The process used here shows attributes of business rules-intensive processes as well as the potential of automation of its sub-processes (see [15] for business process type definition and identification).

Thus, the approach for business process simulation suggested by Hlupic and Robinson [22] is adopted in the presented research. To derive research results following steps are performed:

- Define modeling objectives;
- Decide on model boundaries;
- Collect and analyze data;
- Develop business process simulation model;
- Test model;
- Model experimentation;
- Analyze output;
- Provide business process change recommendations.

Hlupic and Robinson [22] suggest this proceeding for evaluation and decision support in the context of re-engineering, thus the last phase is omitted here, because the goal of the research is exploratory rather than recommendatory. The case study roots in a business process analysis project where the process under consideration had to be defined, modeled and explored concerning its potentials for automation. The resulting as-is model and process information was used here as input for simulation design.

Thus, the modeling objective is to document and analyze the as-is as well as to-be processes. The model boundaries have been set according to the project scope, i.e., the patient management process by health and medical managers. The as-is process model, including execution time and costs of the tasks as well as the designed to-be process model have been discussed and evaluated with process workers and process manager. As mentioned above, the process flows have been modeled and task characteristics such as their execution costs and times have been included to enable the simulation of the quantitative characteristics and thus performance measurement. After the short process and simulation set-up descriptions in the following sub-sections, the accordant simulation results are presented and discussed.

B. Process Description

The process under analysis aims at the provision of a suitable control of patient treatment. Process goal is to provide efficient and suitable treatment corresponding to the patient's diagnostics for the patient and the health fund. The challenge for the health manager at the health fund is to understand the diagnosis of the patient and to evaluate which treatment can be suitable for an efficient recovery. Thus, a certain amount of medical knowledge is needed to guide the treatment process.

 ISBN: 978-1-61208-223-3

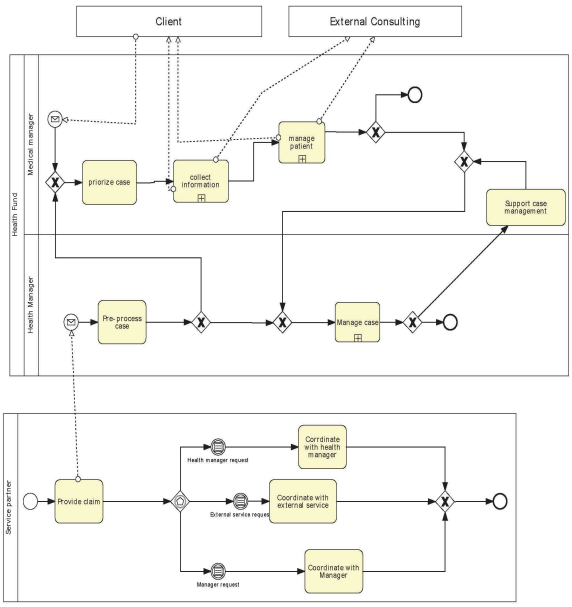


Figure 1. Health fund process

The process presented in Figure 1 starts with a message or request about the medical case filed by the patient or treating facility in cases when the patient had to be delivered to the hospital or is already under medical treatment. This case is analyzed by either by the health or medical manager with the regard to sufficiency of the included data. If all needed information is given, either within the case or in the patient history available at the fund, the case is evaluated concerning its complexity. If health manager classifies the case as clear, he or she supervises further treatment. If he or she doubts whether her or his knowledge is sufficient for further handling the case, he or she can pass the case to a worker who has an indicated higher medical knowledge, i.e., medical manager. In this case, medical manager is responsible to govern the treatment. Thus, medical manager often acts in a consulting position for the health manager or reworks cases that could not be managed by the health manager.

Governance of the treatment is composed of several steps. Diagnostics of the patient needs to be comprehensive and his or her nursing assessment needs to be identified. According to this data a general treatment proceeding can be planed. This proceeding is discussed with the patient. According to the diagnostics a specialist or a treatment facility, such as a hospital or a rehabilitation institution, needs to be chosen. The choice depends on the current list of partner institutions, their availability or agreements between the health fund and medical specialists. Furthermore, according to the treatment plan, costs of the treatment need to be reviewed and monitored along the treatment process. External consulting provides support upon request in cases where the diagnosis or treatment activities are not clear. Thus, we can identify following actors involved into this process: patient, treatment facility (that can also be a medicine or a hospital), external consulting, the operative health and medical manager from the health fund. The goal of this process can be identified as: provide an effective patient treatment. The term efficiency implies best possible recovery of the patient and efficient resource involvement (such as costs and time) on the side of the health fund.

While in the described as-is process the information collection and decision making have been performed manually, in the course of to-be process design, business rules as well as process automation potentials were derived and implemented into the process model. The goal of the to-be process was to reduce the process time and costs for the health fund while providing the right treatment, i.e., efficient process for the patient and health fund. This goal is consistent with the as-is process goal. Therefore, the work load of the medical manager has been reduced using a more effective decision making and information collection. Identified business rules were supposed to be externalized in a business rules management system to reduce process time. The simulation of the as-is and to-be processes is now applied to assess the performance difference between the two scenarios using process time and cost metrics.

C. Simulation Set- up

Conducted interviews with process workers and managers were led to capture the process as well as to provide information on process times and indications for derivation and constrain business rules. Business rules that are relevant for process execution have been identified as mostly being of type constraints and process rules indicted by control flow gateways. Additional business rules as well as process rules have been derived using the business process model (see BPMN rule events in Figure 1) and business rules extraction process described above. Additionally, decision points within the process were identified and the accordant information has been gathered. Decision points indicate process stages or tasks that require additional data for decision making [30] and are thus potential sources of business rules enactment.

Furthermore, the as-is process has been modeled including the as-is process implementation, i.e., manual information gathering has been represented as manual tasks. This as-is model has been simulated using Adonis® 2.01 community edition software and the defined performance indicators such as time and estimated costs in money units (MU) of the process have been observed. The process has been simulated for a time-span of a year to gather average costs and frequencies of the single activities, as well as to take into account that a patient recovery needs to be accompanied for a longer period of time, that is, until he or she likely does not need any medical attention for the original diagnosis anymore. Thus, monitored frequencies and costs of activities represent the real process without being exhaustive and based on simulation rather than on the actual execution patterns.

In a next step, the as-is process model has been modified to a to-be process to include identified and automated business rules. Thus, constrain rules have been treated as information collection points [30], i.e., information gathering, as well as constrain estimation has been

 ISBN: 978-1-61208-223-3

automated. In the actual realization of this automation e.g., a document management or business rules management system can be used. Process rules have been modeled as automated, by e.g., a workflow management system. Subsequentially, this to-be process has also been simulated with the focus on the described performance indicators. While the as-is process data on time and cost management have been gathered in interviews, data in the to-be process have been estimated and evaluated using experience and expert interviews. As the result, two of the manual process activities of the as-is process have been fully automated in the to-be process.

V. SIMULATION RESULTS

Simulation of the as-is process resulted in an average time for the case support (in working days) of 8 days and approx. 7 hours with estimated costs of 69.62 MU (see Table 1). As the most frequent activities in the process "manage case" activity performed by the health manager as well as "prioritize case", "gather information" and "manage patient" by the medical manager have been identified. "manage patient" was identified as the activity with the highest costs (25.2 MU) as well as with a rather high average frequency of execution in a process cycle (0.84).

Simulation of the to-be process resulted in a reduced process time by 37.5% as well as reduced estimated costs by 17.3%. Although, in the to-be process the number of activities for health manager increased, his or her workload is now more focused on the efficient case processing and his or her area of expertise. "manage case" activity is now performed with a lower frequency while health manager also performs "manage patient" activity. This fact is also beneficial for process efficiency as the time of the medical manager is more expensive comparing to the time of the health manager due to their different levels in medical knowledge. In the to-be process medical manager is less involved in the support of health (frequency of the "support case management" activity is decreased), while being slightly more involved in patient management.

TABLE I. SIMULATION RESULTS

Indicator	As-is values	To-be values
Process time	8d 6h 47min	5d 7h 42min
Process costs	69.62 MU	59.37 MU
Performance Quotient (MU/min)	0.00584	0.00775
Manage patient (frequency; costs)	0.89; 25.20	1.19; 35.55
Gather information (frequency; costs)	0.84; 25.20	0.92; 18.39
Support case management	0.43; 6.5	0.21; 3.15
Priories case (frequency; costs)	0.84; 8.4	0.8; 1.6
Pre-manage case (frequency; costs)	0.4; 2	0.4; 0.8

In the to-be process activities "gather information" and "manage patient" are the most frequent and expensive activities performed by medical manager. They are "routine" activities for the medical manager with frequencies of 0.92 and 1.19 respectively, as well as respectively requiring the costs of 18.39MU and 35.55 MU. An additional, though a rather obvious observation, is that the rise of frequency of these activities is accompanied by a nearly proportional rise in costs (frequency rise of 42% was accompanied by rise of the costs by 41.1%). Another effect of business rules automation can be observed on the two process tasks that have been automated: "priorise case" and "pre-manage case". Their occurrence frequency remained nearly the same but their costs were reduced to a fraction of the original costs. Hence, the automation of the to-be process resulted, among other, in a higher number of activities for health manager as well as the supporting information system.

Finally, to be able to compare the performance of the two processes, their performance quotient has been calculated. The performance quotient of the process has been defined here as the relation of the process cycle costs (in MU) and process cycle time (in minutes). As the result of this analysis, the to-be process performance increased by 24.65% comparing to the as-is process performance. This rise can be attributed to the business rules automation regarding the experiment design and simulation results.

VI. CONCLUSION AND OUTLOOK

In this paper, simulation-based experiment and its design for the assessment of the effects of business rules automation on business process performance has been illustrated using a process modeling and simulation software. The focus of this business rules simulation was put on constrain as well as process rules as they provide the highest process enhancement potential when realized using IT-supported execution. Experiment results show that the process performance of a to-be process involving automation of the business rules that were identified and externalized from the as-is process using a systematic approach was increased. This performance rise is attributed here to the business rules externalization and automation.

Applied research approach can be used as a basis for estimation of the business rules automation benefits. One major limitation of the presented results though, is the fact that the experiment design takes neither the investment, nor the implementation costs for business process and business rules automation infrastructure into account. Additionally a general limitation of business process simulation is that each modeled scenario represents only one possible outcome and over long periods of time these outcomes could be greatly different to actual performance. It is therefore essential to undertake a series of experiments and judge the sensitivity of model results to changes in input factors. This will provide some indication of how useful the simulation results are likely to be. Nevertheless, this paper assessed some of the claimed benefits of the business rules management approach. Here, its effects on process improvement or efficiency have been addressed.

This paper illustrated that the performance of a (rule and information intensive) business process can be enhanced considering the indicators time and cost using business rules automation. Considering business rules in this process also resulted in a process workload and responsibilities for the process actors that are now more fitted to the job description.

 ISBN: 978-1-61208-223-3

This fact needs to be examined in future research, as it may provide an indication towards a measurable effect of psycho-social dimensions such as higher job satisfaction on process performance. Further research directions are to explore the effects of business rules management approach on process and data maintenance costs, compliance, update related efforts as well as decision accuracy.

VII. REFERENCES

[1] Levina, O., Kubicki, A., Holschke, O. and Rake-Revelant, J., Exploring the Impact of Business Rules Consideration on Processes in E-Business, in 2nd IEEE International Conference on Information Management and Engineering. 2010, IEEE: Chendgu, China.

[2] Simon, H.A., Applying Information Technology to Organization Design. Public Administration Review, 1973. **33**(3): pp. 268-278.

[3] Giaglis, G., Paul, R. and Hlupic, V., Integrating simulation in organisational design studies. International Journal of Information Management, 1999. **19**(3): pp. 219-236.

[4] MacArthur, P.J., Crosslin, R.L. and Warren, J.R., A strategy for evaluating alternative information system designs for business process reengineering. International Journal of Information Management, 1994. **14**(4): pp. 237-251.

[5] Robbe, N. The Business Rules Advantage for Process Improvement. in BrainStorm New York. 2010.

[6] Hevner, A.R., March, S.T., Park, J. and Ram, S. , Design Science in Information Systems Research. MIS Quarterly, 2004. **28**(1): pp. 75-105.

[7] BusinessRuleGroup. The Business Rules Group. Defining Business Rules – What Are They Really? [cited June, 2012]; Available from: http://www.businessrulesgroup.org/first_paper/br01c0.htm.

[8] Ross, R., Principles of the Business Rule Approach. Management Information Systems. 2003: Addison- Wesley.

[9] Taveter, K. and Wagner, G. Agent-Oriented Enterprise Modeling Based on Business Rules? in 20th Int. Conf. on Conceptual Modeling. 2001. Yokohama, Japan: Springer.

[10] von Halle, B., Business Rules Applied. 2002: Wiley & Sons.

[11] Schacher, M., Grässle, Patrick, Agile Enterprises through the Business Rules Approach (In German). Xpert. press Series. 2006: Springer.

[12] Levina, O., Holschke, O. and Rake-Revelant, J. Extracting Business Logic from Business Process Models. in 2nd IEEE International Conference on Information Management and Engineering. 2010. Chengdu, China: IEEE.

[13] von Halle, B., Goldberg, L. and Zachman, J., Business Rule Revolution. Running Business the Right Way. 2006: Happy About.

[14] Charfi, A., Mezini, M. Hybrid web service composition: business processes meet business rules. in 2nd international conference on Service oriented computing. 2004. New York, USA: ACM.

[15] Levina, O. and Hillmann, R. Network-based Business Process Analysis in 45th Hawaii International Conference on System Sciences. 2012.

[16] Ross, J., W., Creating a Strategic IT Architecture Competency: Learning in Stages. MIT Sloan Working Paper 2003..

[17] Muehlen, M.Z., Indulska, M. and Kamp, G. . Business process and business rule modeling languages for compliance management: a representational analysis. in 26th international conference on Conceptual modeling. 2007. Auckland, New Zealand

[18] Eijndhoven, T.V., Iacob, M. and Ponisio, M. L. Achieving Business Process Flexibility with Business Rules. in IEEE International Enterprise Distributed Object Computing Conference. 2008.

[19] Gottesdiener, E., Harvesting Business Rules in Workshops. DataToKnowledge Newletter (now Business Rules Journal), 1999 27(2).

[20] Graml, T., Bracht, R. and Spies, M. Patterns of Business Rules to Enable Agile Business Processes. in 11th IEEE International Enterprise Distributed Object Computing Conference. 2007.

[21] Resch, O., Six Views on the Business Rule Management System. e-Journal of Practical Business Research, 2010. 11.

[22] Hlupic, V. and Robinson, S. Business process modelling and analysis using discrete-event simulation. in Proceedings of the 30th conference on Winter simulation. 1998. Los Alamitos, CA, USA: IEEE Computer Society Press.

[23] Swami, A. Building the business using process simulation. in Simulation Conference. 1995: IEEE.

[24] Wong, K., Parkin, R. and Coy, J. , Integration of the Cimosa and high-level coloured Petri net modelling techniques with application in the postal process using hierarchical dispatching rules. Journal of Engineering Manufacture,, 2007. 221(5): pp. 775-786.

[25] Rosenberg, F. and Dustdar, S. , Towards a Distributed Service-Oriented Business Rules System. Web Services, European Conference on, 2005: pp. 14-24.

[26] Rosenberg, F. and Dustdar, S. Design and Implementation of a Service-Oriented Business Rules Broker. in: Proceedings of the Seventh IEEE International Conference on E-Commerce Technology Workshops. 2005. Washington, DC, USA: IEEE Computer Society.

[27] Bajec, M. and Krisper, M., A methodology and tool support for managing business rules in organizations. Journal of Information Systems, 2005. **30**(6).

[28] Bubenko, J.A. and Wangler, B. Objectives Driven Capture of Business Rules and of Information Systems Requirements. in Systems, Man and Cybernetics, 1993. Systems Engineering in the Service of Humans, Conference Proceedings.

[29] Stuart, I., McCutcheon, D., Handfield, R., McLachlin, R., and Samson, D., Effective case research in operations management: a process perspective. Journal of Operations Management, 2002. **20**: pp. 419–433.

[30] Levina, O., Design of Decision Service Using Cause-and-Effect Business Process Analysis. Computer Science- Scientific Journal of Riga Technical University, 2011. **46**: pp. 19-27.

 ISBN: 978-1-61208-223-3

Twitter Search Methods using Retweet Information

Jaeyoung Chang
Department of Computer Engineering
Hansung University
Seoul, Korea
jychang@hansung.ac.kr

Han-joon Kim
School of Electrical and Computer Engineering
University of Seoul
Seoul, Korea
khj@uos.ac.kr

***Abstract*— Recently, as social network services such as *Twitter* and *FaceBook* are becoming more popular, a large number of researches have been carried out with various approaches. However, since social network services have been launched recently, its related search methods are still at an early stage of practical service. Thus, most of current web search sites provide a simple search service for social network service posting articles in the order of their upload time. In this paper, we present a novel way of searching informative posting data in *Twitter*. The proposed method uses both the frequency of *retweets* and the number of users' followers as major factors of ranking function in order to evaluate the quality of postings.**

Keywords-Twitter; social network service; ranking; search

I. INTRODUCTION

Currently, Web 2.0 provides a variety of advanced services such as information sharing, information generation and user-friendly service. One of the popular Web 2.0 services is the social network service such as *Twitter* and *FaceBook*. Especially, the *Twitter* service begun in 2006, and the number of its members increases exponentially and is now over 200 millions.

The posting articles of Twitter are different from the general Web contents in two aspects. First, the posting articles (shortly, *tweet*) are limited to be 140 bytes in length; thus, the Twitter users should compose their thoughts and information concisely and clearly. Secondly, the tweets can be disseminated extremely fast due to the *retweet* and *mention*; the *mention* allows Twitter users to deliver their opinions for each tweet in a question/answer form. The *retweet* is to automatically transmit the tweets to users' *followers*, and thus it allows users to share significant tweets with the followers. As a result, twitter users can read their followers' tweets in nearly real time.

However, in order to obtain new and valuable information in Twitter, users have only to depend upon their followers, even though more than 200 million tweets in every single day are spread out in Twitter social network [1]. Thus, twitter users come to expect some effective search methods in tweets as in Web document search. Of course, many of Twitter web pages provide a sort of search service, but they show only the tweets containing given search keywords in the order of posting time. By contrast, the current Web search engines can find out Web documents relevant to given keywords by using geometric (or semantic) similarity functions and analyzing hyperlink structures, and they show the search results in the order of similarity values.

The approach to searching tweets is very different from the one to searching Web documents. Since tweets are very short, it is not easy to evaluate the relevance of the articles only with search keywords. In addition, many of tweets contain users' sentiments or opinions, and so such subjective articles (or sentences) should be excluded for search service. With considering these characteristics, one better solution to searching tweets is to use their meta-information rather than their contents; the meta-information includes the number of a user's followers, the frequency of *retweets*, the frequency of *mentions*, link information, and so on.

This paper proposes a novel way of searching tweets utilizing the *retweet* meta-information. Recent research work in [2-4] tackled the search problem for tweets, and however it is still at an early stage of practical service. Particular tweets to be *retweeted* are informative ones that users have judged to be worthy of dissemination. Thus the *retweet* frequency is very useful for evaluating the importance of initial tweets. The proposed method considers not only the number of followers but also the retweet frequency.

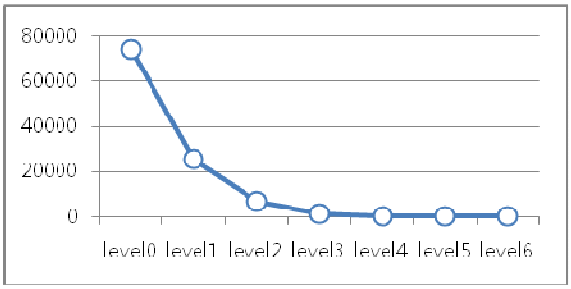
The rest of this paper is organized as follows. In Section II, we review several related research work on analyzing social network data. In Section III, we describe interesting and useful patterns for the number of followers and *retweet* frequency. In Section IV, we present three search ranking functions based on the observed patterns. In the last section, we summarize our work and introduce future work.

II. RELATED WORK

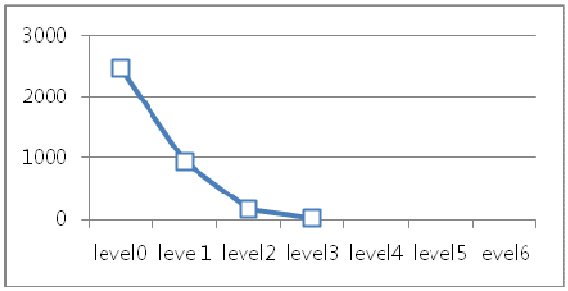
The conventional TF-IDF (Term Frequency–Inverse Document Frequency) based search methods are not suitable for search methods of tweets [5-6]. It is because the TF-IDF weighing scheme is effective for the case that the term frequency is a good indicator for the importance of documents. Since a tweet normally does not have more than two significant words, it is very hard for TF and IDF to be metrics suitable for tweets.

Until now, several Twitter search methods have been proposed, which are usually based on the meta-information. A. Sarma, At. Sarma, S. Gollapudi, and R. Panigrahy [3] proposed a ranking method that exploits users' feedback information, and its ranking function was implemented by analyzing the users' feedback score data. However, this method is dependent all upon users' feedback, and thus it is

 ISBN: 978-1-61208-223-3



(a) Posting articles about '*the death of Jungil Kim*'



(b) Posting articles about '*the service shut down of a Korean broadcaster*'

Figure 1 The frequency of *retweet* operations

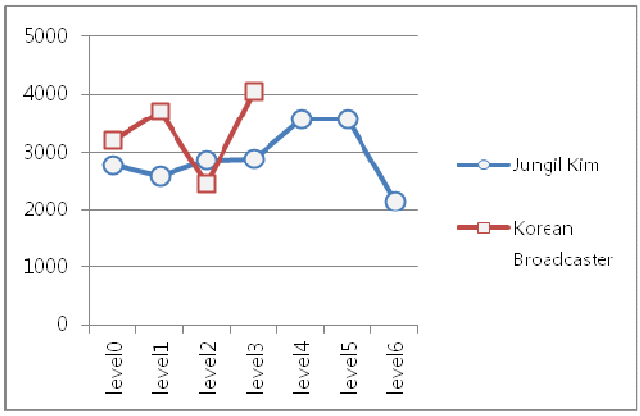
impractical to search tweets. H. W. Lauw, A. Ntoulas, and K. Kenthapadi [4] proposed another search method that actively uses meta-information such as the *retweet* frequency, the number of followers, and the posting time interval. This method is more or less similar to our proposed method. However because it estimates the *retweet* frequency only with the number of followers, the estimated retweet frequency can be incorrect in many cases. Our proposed method tries to use the precise *retweet* frequency. Also, R. Nagmoti and M. D. Cock [2] proposed a ranking method that uses the number of tweets, the number of followers, and the length of tweets.

Actually, most of studies on Twitter have focused upon searching influencing users instead of searching interested tweets [7-11]; for example, in [10], the Google's PageRank method was applied to the Twitter service in order to evaluate users' effect for each other. However, we believe that future systems will require their sophisticated search methods for social network data.

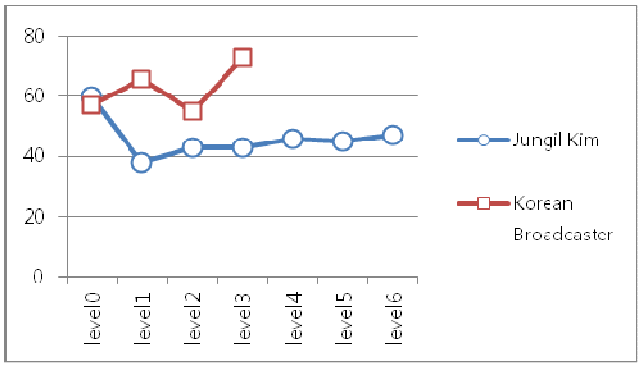
As mentioned before, some previous studies consider the *retweet* frequency and the number of followers to evaluate the quality of tweets [2][4], which are also used as major factors in our work. We show that the factors are highly important in evaluating the quality of tweets in this paper. Moreover, we consider the retweet step for original tweets, which is useful for ranking the tweets to be retrieved.

III. EXPERIMENTAL ANALYSIS OF RETWEET SERVICE

In our work, we have conducted an interesting experiment to analyze the characteristics of *retweet* service.



(a) The average number of users' followers



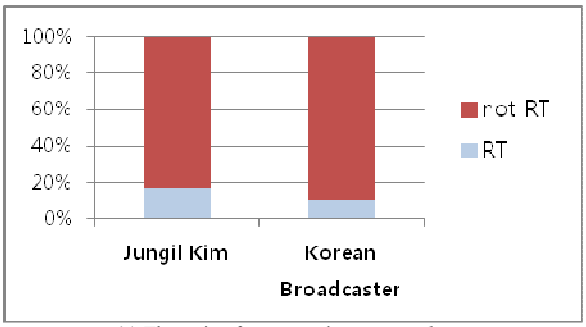
(b) The average number of posting articles

Figure 2 The average number of users' followers and posting articles

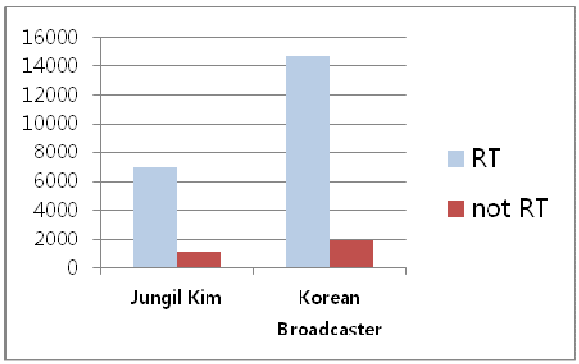
For the experiment, we have collected a large number of articles related to the current two hot issues: one is relevant to the death of Jungil Kim (who was a previous director of North Korea) with being collected for 36 hours from 2 PM, 17 Dec. 2011, and the other is relevant to an event of service shut down of a Korean major broadcaster with being collected for 36 hours from 9 AM, 17 Jan. 2012. People were much more interested in the former articles.

Figure 1 shows the frequency of *retweet* operations for tweets of the two hot issues. In this figure, *level-0* means the initial posting articles, and *level-i* means the *i*-th *retweeted* posting articles; that is, *level-i* means the frequency of *retweet* operations from *level-(i-1)*. As seen in the figure, the articles about 'the death of Jungil Kim' were *retweeted* at six steps, and the ones about 'the service shut down of a Korean broadcaster' at three steps. This implies that the articles related to hotter issues have more *retweet* operations.

Figure 2 shows the average number of a user's followers to perform *retweet* operations, and the average number of tweets posted by retweet users nearly at the *retweet* time. Figure 2(a) shows that both of the tweet sets are *not* obviously different in terms of the number of followers and the number of *retweet* steps. However, Figure 2(b) shows that the average number of tweets posted by users who have *retweeted* the articles related to "the service shut down of a



(a) The ratio of tweets to be *retweeted*



(b) The average number of users' followers for *retweeted* posting articles

Figure 3 The ratio of tweets and the average number of followers for *retweeted* posting articles

Korean broadcaster" is larger than the ones posted by users who have *retweeted* the articles related to "the death of Jungil Kim". This implies that the power users of Twitter are interested in various issues.

Figure 3(a) shows the ratio of the articles to be *retweeted* more than once to the articles not to be *retweeted*, in which RT and notRT mean the articles to be *retweeted* and not to be *retweeted*, respectively. In this figure, we have found that the 10~20% tweets among all the tweet articles are to be *retweeted*; of course, the ratio can change according to the degree of importance of articles. Figure 3(b) shows the average number of followers only for *retweeted* articles. As expected, we find that the average number of the followers of the user who has posted *retweeted* articles is larger than that of the user who has posted the articles not to be *retweeted*.

IV. THE SEARCH METHOD USING *RETWEET* INFORMATION

With considering the experimental patterns of Twitter data mentioned in Section III, we suggest the following assumptions, which are applied to our search function.

- The *retweet* operation is a way of dissemination for informative tweets worthy of sharing.
- As the number of a tweet user's followers increases, his/her articles are more likely to be *retweeted*. Consequently, the tweet user who has a lot of followers plays a great role in disseminating some valuable information.
- The tweets that have drawn great attention are more likely to be *retweeted*.

With the above assumptions, we propose a way of search method over tweets. Basically, searching tweets is to find out those having given query words. As in the Web search engines, it is necessary for Twitter search to rank the related articles in the order of some appropriate criteria such as their information values. As mentioned before, some tweet articles have very informative whereas others have very subjective or non-informative. Thus it is necessary to isolate highly informative articles by evaluating the frequency of *retweet*s.

Let $\overline{D_q}$ be the set of articles that include the search keyword q. $\overline{D_q}$ includes not only the initial articles but also their *retweeted* articles; that is, let $\overline{D_q^0}$ and $\overline{D_q^i}$ be the initial articles and the i-th *retweeted* articles, respectively, then $\overline{D_q} = \bigcup_{i=0}^{\infty} \overline{D_q^i}$.

Under this assumption, a feasible way of evaluating the value of tweet articles is to the *retweet* frequency irrespective of the steps of *retweets*.

As a result, among the articles collected for t time, a set of the initial articles d_q^0 including the search keyword q can be given the following impact value.

$$impact(d_q^0, t) = \sum_{i=1}^{N(t)} \left| \overline{d_q^i} \right| \quad (1)$$

where $\overline{d_q^i}$ denotes a set of articles *retweeted* at the i-th step from d_q^0, $N(t)$ the number of *retweet* for t time and $\left| \overline{d_q^i} \right|$ the number of articles in $\overline{d_q^i}$. This idea has been applied to several search systems over Twitter data. For example, Korean twitter service such as *Joinmsn* help users to show the interested articles that has been *retweeted* the most frequently for some time [12-13].

The alternative way is to integrate the factor of *retweet* steps into Equation 1. Successive re-sending (i.e., *retweet*-ing) *retweeted* articles means that such articles are valuable so as to share with each other. This idea can be realized as the following equation by giving weight values at each *retweet* step.

$$impact(d_q^0, t) = \sum_{i=1}^{N(t)} \alpha_i \left| \overline{d_q^i} \right| \quad (2)$$

where α_i is the weight value at i-th *retweet* step. Therefore, this equation means that as α_i is given a larger value, the articles that have been successively *retweeted* can be assigned greater priorities for searching.

 ISBN: 978-1-61208-223-3

Furthermore, we cannot help considering the number of followers in evaluating the tweets. In general, some people who have a strong influence on others (i.e., have a large number of their followers) tend to be circumspect in *retweet*-ing tweet articles. Each of them is a sort of *information hub* in disseminating posting articles. Considering such an issue, we devise the following final equation for searching.

$$impact(d_q^0, t) = \sum_{i=1}^{N(t)} \sum_{d \in d_q^i} \sum_{u \in pu(d)} \alpha_i \times \log_{10}(|\ followers(u)\ |) \quad (3)$$

where $pu(d)$ denotes the user who has posted the tweet d, and $followers(u)$ denotes the set of followers of the user u. In comparison with Equation 2, Equation 3 is considering the influencing power of a user who performs the *retweet* operation. Normally, there are wide variations in the number of followers. Thus, we use logarithm in Equation 3 to avoid undesirable fluctuation by too big (or small) values in computing the value of Equation 3

V. EMPRICAL RESULTS

We have performed a number of experiments in order to evaluate the proposed tweet search method. As for test data, we have prepared two kinds of tweets collections: one is a collection of more than 100,000 tweets about 'the death of Jungil Kim' and the other is a collection of more than 3,500 tweets about 'the service shut down of a Korean broadcaster'. With the test data, we have estimated the accuracy of ranking results by using Equations 1, 2, and 3. To evaluate the accuracy of ranking results, it is necessary to have correctly ranked lists for tweets. For this, we have made about 30 humans evaluate the quality of top 100 articles to be the most frequently retweeted.

As a measure of accuracy, we have used *nDCG* (Normalized Discounted Cumulative Gain) [14], which is commonly used to evaluate the effectiveness of Web search algorithms. Basically, this measure is to compare the current query results with perfect ranking results for different types of queries, which is defined for the query q as follows.

$$nDCG_q = M_q \sum_{i=1}^{K} (2^{rel(i)} - 1) / \log(1+i) \quad (4)$$

where $rel(i)$ is the graded relevance of the query result at the position i, and it is evaluated with four discrete values, i.e., 0, 1, 2, 3. As the relatively upper part of query results are evaluated as higher values for $rel(i)$, the *nDCG* values come to be higher. And, M_q is a normalized constant. Consequently, the *nDCG* value ranges from 0 and 1, and it has the value 1 when we obtain the best result.

Figure 4 shows the effect of the proposed methods in terms of *nDCG* measure, where *impact1*, *impact2*, and *impact3* denote the ranking methods by Equations 1, 2, and 3, respectively. In Equations 2 and 3, α_i which is the weight value at i-th *retweet* is set to 2^i as retweet step increases. As

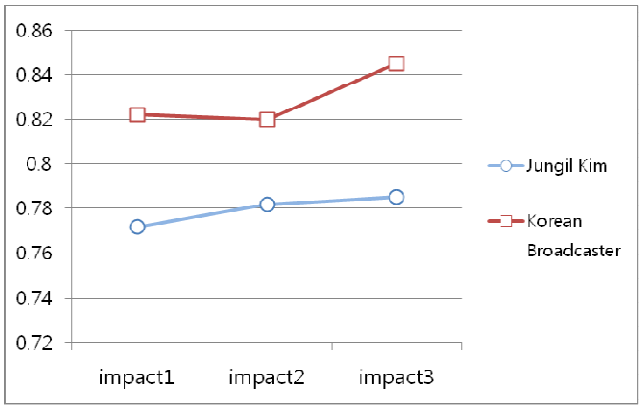


Figure 4 Accuracy of the ranking results in terms of nDCG

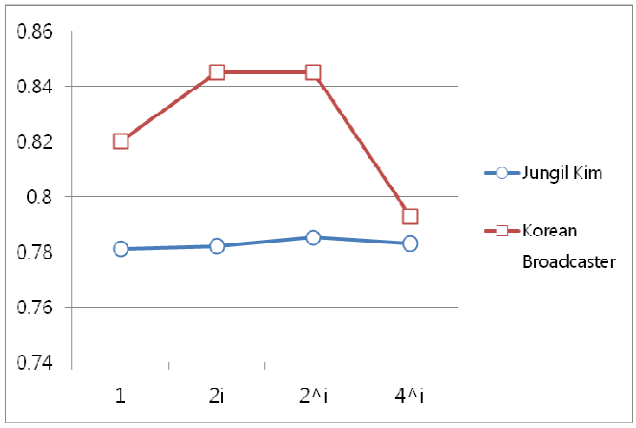


Figure 5 The changes of ranking accuracy from varying α_i

shown in the figure, the ranking method based on Equation 3 give the best results for both of the test collections. This means that the retweet frequency and the number of followers play an important role in evaluating tweets. Specifically, the *impact3* method outperforms the *impact1* method by 1.7% and 2.2% for the test collections about 'the death of Jungil Kim', and 'the service shut down of a Korean broadcaster', respectively. This is because the tweets related to the issue about 'the death of Jungil Kim' are mostly users' opinions rather than facts, and moreover some of them are the articles retweeted by followers.

Figure 5 shows the changes of ranking accuracy from varying α_i when using Equation 3. The value of α_i is allowed to increase by 1, $2i$, 2^i, 4^i according to the retweet step. As shown in this figure, the proposed method based on Equation 3 gives higher accuracy at the positions $2i$, and 2^i over the test collection about 'the service shut down of a Korean broadcaster'. However, the method does not give any significant change from varying α_i over the test collection about 'the death of Jungil Kim'. This is also because most of the tweets related to the issue about 'the death of Jungil Kim' are users' opinions or retweeted articles. From the empirical results, we have found that retweet steps could not play a

 ISBN: 978-1-61208-223-3

great role in evaluating the quality of tweets. We can expect that the proposed methods show relatively higher accuracy if the test collections contains more larger amount of informative news based on new 'facts'.

VI. SUMMARY AND FUTURE WORK

Recently, several studies on searching the Twitter data has been carried out as the social network services become very popular and influential. However, a lot of research issues need to be tackled since the Twitter data is extremely different from general Web documents.

In this paper, we proposed a novel method of searching Twitter data. The method considers two main factors: *retweet* and the number of users' followers. Now, we plan to apply machine learning algorithms to the currently proposed method; this is because for more accurate search, it is necessary to automatically adjust the weight value α_i .

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number: NRF-2011-0022445) and was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number: NRF-2011-0025212) .

REFERENCES

[1] TwitterEngineering, "200 million tweet per day", http://blog.twitter.com/2011/06/200-million-tweets-per-day.html, [retrieved: Jan., 2012].

[2] R. Nagmoti and M. D. Cock, "Ranking Approach for Microblog Search", Proc. of WI-IAT conference, pp. 153-157, 2010.

[3] A. Sarma, At. Sarma, S. Gollapudi, and R. Panigrahy, "Ranking Mechanisms in Twitter-like Forums", Proc. of WSDM conference, pp. 21-30, Feb. 2010.

[4] H. W. Lauw, A. Ntoulas, and K. Kenthapadi, "Estimating the Quality of Postings in the Real-time Web", Proc. of SSM conference, pp. 102-115, 2010.

[5] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition), ACM, 2011.

[6] http://lucene.apache.org/nutch/, [retrieved: Dec., 2011].

[7] J. Teevan, D. Ramage, and M. R. Morris, "#TwitterSearch: A Comparison of Microblog Search and Web Search", Proc. of WSDM conference. pp. 35-44, 2011.

[8] M. Char, H. Haddadi, F. Benevenuto, and K. Gummadi, "Measuring User Influence in Twitter: The Million Follower Fallacy", Proc. of International AAAI conference on Weblogs and Social Media, pp. 17-18, 2010.

[9] H. Kwak, C. Lee, H. Park, and S. Moon, "Finding Influentials Based on Temporal Order of Information adoption in Twitter", Proc. of WWW conference, pp. 1137-1138. 2010.

[10] TunkRank, http://tunkrank.com, [retrieved: May, 2011].

[11] J. Weng and Q. He, "TwitterRank: Finding Topic-sensitive Influential Twitterers", Proc. of WSDM conference, pp. 261-270, 2010.

[12] http://www.joinsmsn.com/, [retrieved: Jan., 2012].

[13] http://koreantweeters.com/, [retrieved: Jan., 2012].

[14] http://en.wikipedia.org/wiki/Discounted_cumulative_gain/, [retrieved: Mar., 2012].

 ISBN: 978-1-61208-223-3