Evaluation of Similarity and Reuse of Project Management Processes

Solvita Berzisa[1], Janis Grabis[2], [1-2]*Riga Technical University*

Abstract – **Project management is a knowledge-centric and experience-driven activity supported by the project management information system. In order to use the project management information system efficiently, it is necessary to configure it according to project requirements. The project manager is not always aware of the most appropriate configuration for the particular project. Adequate knowledge would help the project manager define the configuration requirements. Knowledge in the project management area is divided into two forms: data and processes. To generate suggestions concerning data, grouping, statistical analysis and ordering of data items as well as analysis of semantic ambiguities are used. But concerning knowledge about the processes, differences between various process representations should be studied as well as evaluation of process similarity is required. The objective of the paper is to elaborate an approach to evaluate the project management process similarity and reuse of the project-specific knowledge when defining the configuration requirements.**

Keywords – **process similarity, project management information system, PMIS configuration, project management processes**

I. INTRODUCTION

Project Management (PM) consists of a set of 42 processes (PMBOK [1]). These processes describe all PM activities and their execution order. The defined set of input and output data is associated with each of these processes. Project Management Information Systems (PMIS) are used to store project data and to organize PM processes. Different data structures and processes are used in each project. Therefore, it is necessary to configure PMIS according to the specific project circumstances (see Section II.A). At the beginning of the new project, a project manager is not always aware of the most appropriate configuration of PMIS. Appropriate knowledge would help to define the current PMIS configuration. This knowledge is collected from theoretical sources (methodologies, best practices and PMIS vendor supported configurations) and empirical sources (configurations used in previous projects) and includes information about data structures and processes. Information collected from different sources should be analyzed to get useful knowledge. Analysis of semantic name similarity, grouping and ordering are used to process information about PMIS data structure [2]. The same approach could be used for processing information about PM processes. It would use the analysis of process states and transactions, though the problem of comparing different process representations should be addressed (see Section II.B.). The comparison of processes is

another way of the PM process analysis, and in this case the process similarity analysis should be performed.

The objective of this paper is to elaborate an approach to the evaluation of PM process similarity and reuse of knowledge about PM processes in configuration of PMIS. This approach ensures formalization of the process data coming from different representations of the PM processes, evaluation of the process similarity, and the analysis of process information in similar process groups and reuse of the knowledge. Evaluation of the process similarity should be performed for different types of process representations (a detailed description in given in Section II.C). Therefore, all types of representation are initially transformed into state/transaction format. The semantic similarity metric is used to measure PM process similarity. The main contribution of the proposed approach is the grouping of similar processes and the analysis of the state and transactions used in these groups.

The paper is organized in five sections. Section II describes background of the research including description of the PMIS configuration approach, PM processes and the representation approaches and existing researches about the process similarity evaluation. The approach to PM process similarity evaluation and analysis is presented in Section III. Application of the approach in the reuse of knowledge during the PMIS configuration and the discussion of results are provided in Section IV. Section V draws conclusions.

II. PRELIMINARIES

This section introduces the research problem (Subsection B), application (Subsection A) and also related investigations in the field of the process similarity analysis (Subsection C).

A. Knowledge-Based Configuration of PMIS

According to the project requirement (R_{j+1}, where $j+1$ is an index of the new project), an approach to PM knowledge-aided configuration of the chosen PMIS has been developed in [3]. The configuration process is shown in Fig.1. The PM knowledge is stored in the PM knowledge repository and organized as the cases (H_j (theoretical case, $i = 1, .. p$) and C_k (empirical case, $k = 1 .. m$)) based on methodologies, projects and other situations. Each case is described according to the XCPM schema (XML schema for Configuration of PMIS [4]) and includes PM data ($Da = \{da_i \mid i=1..a\}$) and process descriptions ($Pr = \{pr_i \mid i=1..b\}$), $H_j=(Da_j, Pr_j)$ and $C_k=(Da_k, Pr_k)$. Appropriate knowledge cases are searched using principles of the case-based reasoning [5] according to characteristics of the project environment (A_{j+1}). Knowledge (K_{j+1}) is given to the user in a form of the suggestions that are

Scientific Journal of Riga Technical University
Computer Science. Information Technology and Management Science

2011
_____ *Volume 49*

obtained by analyzing information from the similar cases found. The prepared description of the PMIS configuration is specified and stored in the configuration file (C_{j+1}) that is structured according to the XCPM schema. The configuration file is used for automated configuration of PMIS by transforming the configuration file to the chosen PMIS application structures.

B. Project Management Processes

Processes describe different PM activities at different levels: all project lifecycle, knowledge area processes (e.g., change management, risk management, communication management), work item processes (e.g., change request, risk, issue, bug, requirement and task) and specific situation processes (e.g., occurrence of the risk). These processes in the knowledge repository have been stored in the XPDL [6] schema notation as a part of XCPM schema [4].

Object-centric and activity-centric process representation approaches can be used for describing PM processes. In case of the object-centric approach [7], the process is described with states and transactions; tasks in the process belong to the object lifecycle, and the object status is changed by tasks. The object-centric process is represented with UML Statechart and State Machine [8], Petri net [9] and others. This approach is used by PM software applications such as Team Foundation Server [10], JIRA [11] and others. In the case of the activity-centric approach [7], the process is described with tasks, activities, functions, events and other elements depending on the notation used. In this approach, the tasks change the object state, and the states and transactions are hidden in the attributes. The activity-centric process is modeled with UML Activity Diagram [8], Event-driven Process Chains (EPCs) [12], Business Process Modeling Notation (BPMN) [13], YAWL [14] and others. This approach is used, for example, in MS Project Server [15].

An example of the change control process with six states (open, impact analysis, in progress, implemented, rejected and closed) described using different representation approaches is shown in Fig.2. At first, these processes seem different; however, they are similar in relation to the change request. In the PM knowledge repository, these processes are stored in the XPDL format (Fig.3). In the object-centric process, 'Participant' describes the object and the process is defined with 'Activities' and 'Transitions'. Activity-centric process is described with 'Participants' (roles), 'Artifacts', 'Associations', and the process is defined with 'Activities' and 'Transitions'.

C. Process Similarity Evaluation

A number of investigations have been conducted for evaluation of business process similarity. These researches have been performed using different business process notations: EPC, Workflow nets, BPMN and UML Activity Diagram and State Chart. Only one of these notations is used in each investigation, but in most cases results could also be applied to other notations. Metrics for measuring similarity among process elements and for measuring similarity between processes have been defined.

Three types of metrics have been defined for measuring similarity among elements/nodes of the processes: syntactic (typographical), semantic (linguistic) and contextual. The syntactic similarity metrics evaluate only the syntax of the labels and return the degree of similarity as measured by the string-edit distance (number of atomic string operations necessary to get from one string to another) [12]. The semantic similarity metrics abstract from the syntax and analyze semantics of the words and return the degree of similarity based on equivalence between the words (identical and synonyms) [12]. The contextual similarity metrics do not evaluate only labels of elements but also the context in which these elements occur by the analysis of preceding and succeeding process elements [12]. These metrics have been defined for processes represented as EPC [12], Petri net [16] and statechart [17].

Different metrics/approaches have been defined to measure similarity between processes: label matching similarity, structural similarity, behavioral similarity; similarity evaluation uses OWL (Web Ontology Language) and statechart matching.

The label matching similarity [12] is based on pairwise comparisons of node labels. The label matching similarity score is the sum of the label similarity scores of matched pairs of nodes divided by the total number of nodes. The label similarity of matched pairs of nodes can be evaluated by syntactic or semantic similarity metrics or a weighted average of them.

The structural similarity [12] evaluates the whole structure of process models ([12] reviews EPC) as a labeled graph. The
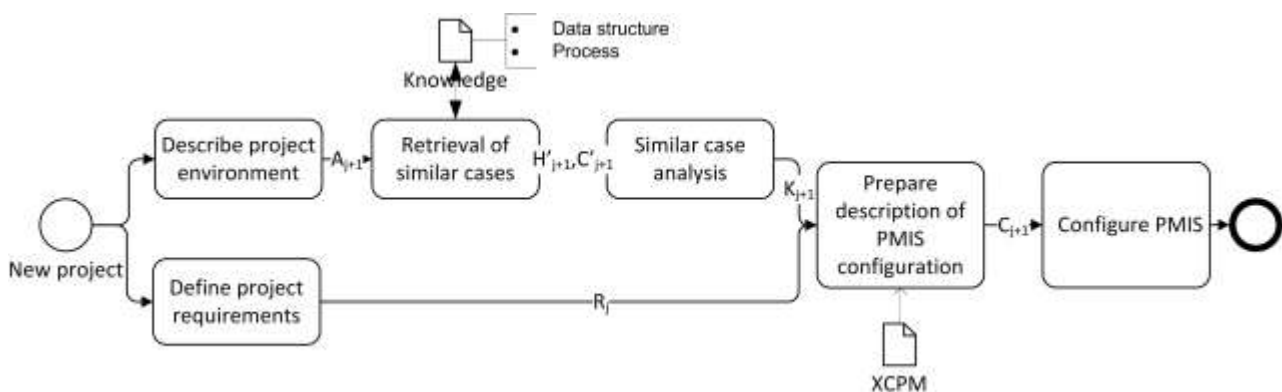

Fig. 1. PMIS configuration approach

Scientific Journal of Riga Technical University
Computer Science. Information Technology and Management Science

2011
_____ *Volume 49*

similarity score of two processes have been evaluated by their graph-edit distance (minimal numbers of graph edit operations that are necessary to get from one graph to the other).

The behavioral similarity [12] of two process models ([12] reviews EPC) is computed by their distance in the document vector space that can be constructed from their causal footprints. Causal footprint is causal relations between activities in the process model (activity look-ahead and look-backs links).

In [16], an approach is presented that for similarity evaluation uses Petri nets described with OWL. The similarity score of two processes is calculated using combined element similarity values. This combined element value is aggregation of weighted syntactical, linguistic and structural similarity measures.

Statechart matching [17] uses combined measurement from typographic, linguistic, depth and behavioral similarity metrics for identification of similar state.

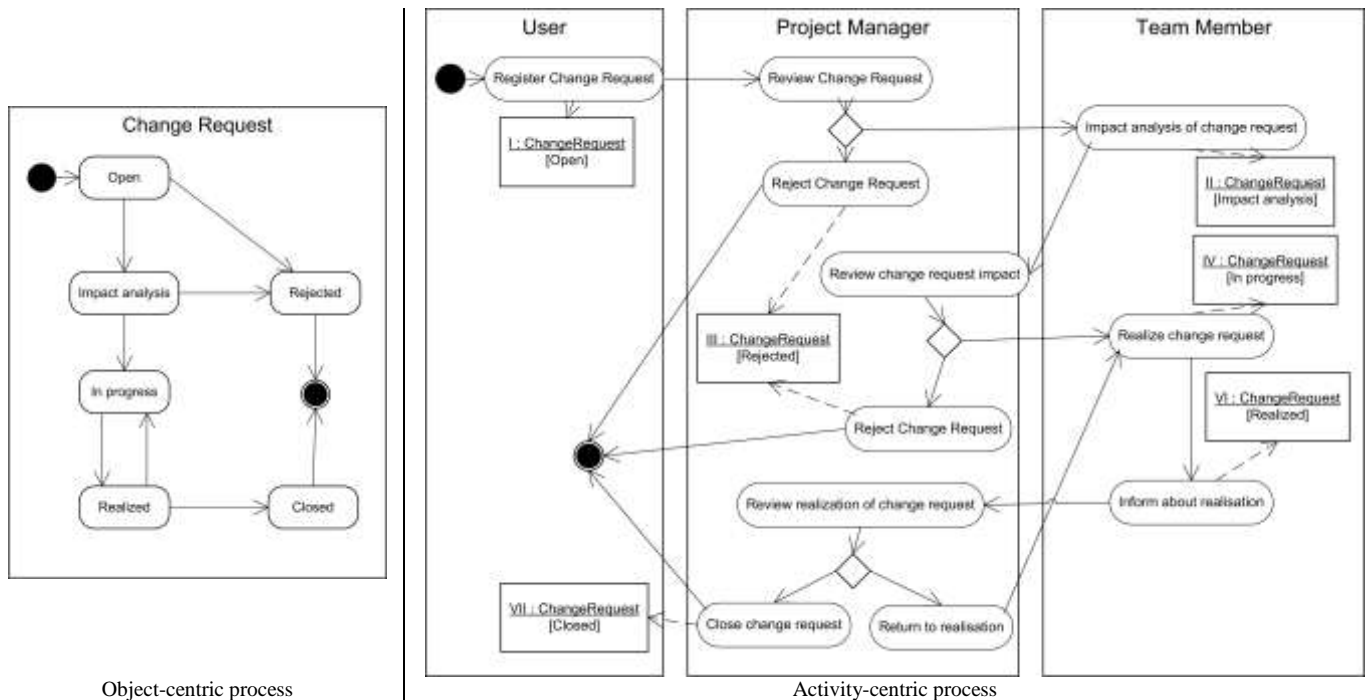Process element or full process similarity score/value is within the interval [0..1].

Fig. 2. Example of PM process representation based on different approaches

Fig. 3. Fragments of example processes in the XPDL format

Scientific Journal of Riga Technical University
Computer Science. Information Technology and Management Science

2011
Volume 49

III. APPROACH TO PM PROCESS SIMILARITY EVALUATION AND ANALYSIS

The PM process similarity evaluation and reuse is a part of the similar case analysis (Fig.1) in the PMIS configuration approach. The process of the PM process similarity evaluation and reuse is shown in Fig.4. It consists of three phases: representation formalization, similarity evaluation and process analysis. The sets of similar theoretical and empirical cases (H'_{j+1} and C'_{j+1}) are returned after the similar case retrieval. The set of related processes $Pr_{j+1}= \{Pr_{k,i}/Pr_{k,i}\in Pr_k$ and $Pr_k\in H_{j+1}\} \cup \{Pr_{k,i}/Pr_{k,i}\in Pr_k$ and $Pr_k\in C_{j+1}\}$ is retrieved from these cases. The set of processes Pr_{j+1} is input data for the approach. The processes (pr_i) are defined with XPDL.

A. Process Representation Formalization

In the first phase, all processes are formalized ($Pr'_{j+1}=F(Pr_{j+1})$), i.e., all activity-centric processes are transformed to object-centric processes. The objects in the activity-centric process are defined with 'DataObject' element inside 'Artifact'. One 'DataObject' could be used in more than one artifact, but with different 'stage' values. The object-centric process is transformed by analysis change of 'stage' in 'DataObject' within the activity-centric process. In the new process, 'Participant' is 'DataObject', 'Activities' are 'DataObject', stages and stage changes are described with 'Transactions'. All processes are object-centric after the process formalization, and they describe definite object stages used in the process and the way object stages are changed within the process. The transformed activity-centric process in Fig.3 will be equal to the object-centric process in Fig.3.

B. Process Similarity Evaluation

Evaluation of the process semantic similarity is sufficient for evaluation of the formalized PM. The object or work item statuses (states or stages) of the PM processes are compared during the similarity evaluation, and the synonym values of these statuses are stored in the synonym dictionary. The synonym dictionary is used for the data structure analysis of a similar project, and it includes information about similar meanings of data item attributes and statuses [2]. The semantic similarity metric is also defined for different notations and is used in the different measurements combinations (see Section II.C.) that demonstrate good results in practical evaluation [12]. The principle defined in label matching similarity is used for the process similarity score calculation.

Semantic score ($sem(a_{i,s}, a_{k,t})$) of two process activities (pr_i, $pr_k \in Pr'_{j+1}$) is calculated with Formula 1, where the similarity score is 1 if names of the activities are identical or synonymous. In other cases, the similarity is calculated with a formula derived from the formula defined in [12].

$$sem(a_{i,s}, a_{k,t}) = \begin{cases} 1, if\ a_{i,s} = a_{k,t} \\ 1, if\ syn(a_{i,s}, a_{k,t}) = 1 \\ \dfrac{(a_{i,s} \cap a_{k,t}) + \sum\limits_{\substack{c\in a_{i,s}/a_{j,t} \\ b\in a_{k,t}/a_{i,s}}} syn(c,b)}{max(|a_{i,s}|, |a_{k,t}|)} \end{cases}, (1)$$

where:

$a_{i,s} \in pr_i$ (*i* process *s* activity);
$a_{k,t} \in pr_j$ (*k* process *t* activity);
$(a_{i,s}\cap a_{k,t})$ is count of identical names in the activity name;

$$syn(c,b) = \begin{cases} 1, if\ c\ and\ b\ is\ synonums \\ 0, if\ c\ un\ b\ is\ not\ synonums \end{cases};$$

$|a_{i,s}|$ is count of words in the name of $a_{i,s}$.

Similarity between activities "in progress" and "implementation" is evaluated as an example. These two activities are equal in meaning, and these values are stored in the synonym dictionary. The similarity score is 1 for these two activities. However, if only strings "progress" and "implementation" are synonyms, then these two activities have the similarity score equal to (0+1)/2=0.5. The similarity score value depends on the values that have been saved in the synonym dictionary.

The similarity score of two PM processes ($sim(pr_i,pr_k)$) is calculated with Formula 2 that finds the maximum activity similarity score for each activity of pr_i combination to all activities of pr_k. Full process similarity score is calculated by summing the maximum activity similarity score of activities of pr_i and dividing it by the maximum count of activities in pr_i and pr_k.

$$sim(pr_i, pr_k) = \frac{\sum_{s=1}^{x} max(V_{t=1}^{y} a_{k,t}|sem(a_{i,s}, a_{k,t}))}{N}, \quad (2)$$

where:

x is the count of activities in pr_i;
y is the count of activities in pr_k;
$N=max(x,y)$.

The similarity evaluation is performed for each pair of processes in the Pr'_{j+1}. The result is a two-dimensional matrix M[z,z] with the process similarity scores (*z* is count of process in Pr'_{j+1}). The number of the process pairs that are needed to calculate the similarity score can be reduced by taking into consideration that $sim(pr_i,pr_i)=1$ and $sim(pr_i,pr_k)= sim(pr_k,pr_i)$.

Two PM processes are considered to be similar if the

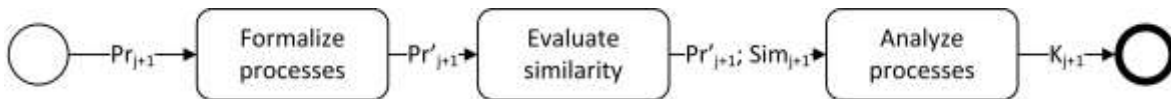Fig. 4. The approach to PM process similarity evaluation and analysis

Scientific Journal of Riga Technical University
Computer Science. Information Technology and Management Science

2011
_____ *Volume 49*

similarity score is greater than or equal to the similarity threshold sim_{lim}. The similar process group consists of the processes whose mutual similarity is greater than or equal to sim_{lim}. One process can be included in more than one group of similar processes.

The result of the process similarity evaluation phase is the set Sim_{j+1} that consists of subsets of the similar processes.

C. Process Analysis

The process analysis phase gets the set of formalized processes (Pr'_{j+1}) and the set with similar process groups (Sim_{j+1}) as the input data. The process analysis is performed in the similar process groups. During the analysis, semantic similarity of activities is analyzed, grouping and ordering are performed, and transactions among activities are analyzed. This analysis is performed for all similar process groups in Sim_{j+1}.

The processes without the similar processes ($pr_i \in Pr'_{j+1}$ and $pr_i \notin Sim_{j+1}$) are represented as a list of independent processes. If the set of processes Pr'_{j+1} does not have similar processes ($Sim_{j+1} = \emptyset$), the process analysis is performed for all sets of the processes.

The result of process analysis is knowledge K_{j+1} that contains groups of similar processes with process identifiers and statistics about activities (stages or state) and transactions. This knowledge is presented to the user.

IV. DISCUSSION

To demonstrate process similarity evaluation and analysis, six processes describing the change management process are considered. This set includes two processes from Fig.2, as well as four other processes.

These processes are formalized, and the list of activities used in each process is shown in Fig.5.

In the first step of the second phase, similarity is evaluated among the processes. In the synonym dictionary it is defined that synonyms are values: "in progress" and "implementation", "verification" and "testing". The similarity evaluation result is the matrix M:

$$M = \begin{bmatrix} 1 & 1 & 0{,}67 & 1 & 0{,}42 & 0{,}50 \\ 1 & 1 & 0{,}67 & 1 & 0{,}42 & 0{,}50 \\ 0{,}67 & 0{,}67 & 1 & 0{,}67 & 0.58 & 0.83 \\ 1 & 1 & 0{,}67 & 1 & 0{,}42 & 0{,}50 \\ 0{,}42 & 0{,}42 & 0.58 & 0{,}42 & 1 & 0{,}70 \\ 0.50 & 0.50 & 0.83 & 0.50 & 0{,}70 & 1 \end{bmatrix}$$

The second step in this phase is identification of the similar process groups with the similarity threshold $sim_{lim}=0{,}70$. The resulting set Sim_{j+1} consists of three similar process groups: {pr_1, pr_2, pr_4}, {pr_3, pr_6} and {pr_5, pr_6}.

The process analysis is performed in the third phase, and the result is shown in Fig.6.

To demonstrate the impact of process grouping, Fig.7 shows the result of process analysis without grouping. The results in Fig.6 and Fig.7 are quite different. The list of

1st group	3 processes ({pr₁, pr₂, pr₄}) Use the following activities: • open (100%) • impact analysis (100%) • in progress/implementation (100%) • implemented (100%) • rejected (100%) • closed (100%)
2nd group	2 processes ({pr₃, pr₆}) Use the following activities: • open (100%) • in progress (100%) • resolved (100%) • verification/testing (100%) • closed (100%) • rejected (50%)
3rd group	2 processes ({pr₅, pr₆}) Use the following activities: • open (100%) • testing (100%) • closed (100%) • in progress (50%) • in implementation (50%) • on hold (50%) • resolved (50%)

Fig. 6. The result of process analysis with similar process groups

6 processes Use the following activities: • open (100%)   • closed (100%)   • in progress/implementation (83%)   • rejected (67%)   • impact analysis (50%)   • verification/testing (50%)   • implemented (50%)   • resolved (33%)   • in implementation (17%)   • on hold (17%)

1st process (pr₁)	2nd process (pr₂)	3rd process (pr₃)
open, impact analysis, in progress, implemented, rejected, closed	open, impact analysis, in progress, implemented, rejected, closed	open, in progress, resolved, verification, rejected, closed
4th process (pr₄)	5th process (pr₅)	6th process (pr₆)
open, impact analysis, implementation, implemented, rejected, closed	open, in implementation, on hold, testing, closed	open, in progress, resolved, testing, closed

Fig. 5. List of activities used in the processes

Scientific Journal of Riga Technical University
Computer Science. Information Technology and Management Science

2011
_____ *Volume 49*

activities and their frequency is obtained (Fig.7) by analyzing processes without evaluation of their similarity. But if the same process analysis is performed using the grouping, then the user also gets information about variations of processes and trends of process status and transaction. The result of example (Fig.6) has shown that in case of a similar project three variations of processes and some activities (status), e.g. "impact analysis" and "implemented", are used in only one type of processes. Using this result, a user can choose one variation of the process as a basis for a new process and extend it. This information cannot be obtained from results in Fig.7.

V. CONCLUSIONS

This paper describes the approach that helps to analyze and interpret knowledge stored in the PM processes. This approach ensures formalization or transformation of activity-centric processes to object-centric processes; identification of the similar process groups by the evaluation of process similarity and analysis of the process information using the similar process groups. After the process formalization, all processes are reviewed from the prospective what object stages are used in the process and how object stages are changed by the process. This information about the PM processes is more useful from the PMIS configuration point of view. The PM process similarity is evaluated by measuring the semantic similarity between activities or object stages of the formalized processes. The processes are considered to be similar if their similarity score exceed the similarity threshold. Mutually similar processes create a group of similar processes. The process analysis is performed separately for each process group. Grouping yields more information about process knowledge than the simple process analysis, because a user gets more information about the variations of processes and the trends of statuses and transactions.

Two factors affect the results of the approach. The first factor is the similarity threshold. Changes of the threshold value affect grouping results. The use of high threshold values results in very small groups of similar process, while low threshold values result in a few very large groups. The effect of the inappropriate threshold is the unidentified significant variation or trends of the PM processes. The other factor is the synonym dictionary that impacts the process similarity score. The values that are or are not stored in the synonym dictionary have strong effect on the similarity of process activities, as well as on the similarity of process. Use of the other similarity metric instead of semantic is also possible.

REFERENCES

[1] Project Management Institute, *A Guide to the Project Management Body of Knowledge (PMBOK® Guide) Fourth Edition*. Newtown Square: Project Management Institute, 2008.
[2] S. Bērziša and J. Grabis, "Knowledge Reuse in Configuration of Project Management Information Systems: A Change Management Case Study" in: *Proceedings of the 15th IEEE International Conference on Intelligent Engineering Systems 2011*. Budapest: IEEE, 2011, pp.: 51-56.
[3] S. Bērziša and J. Grabis, "A Framework for Knowledge-Based Configuration of Project Management Information Systems" in: *The 17th International Conference on Information and Software Technologies*, Butleris, R., Butkiene R., Ed. Kaunas: Kaunas University of Technology, 2011, pp. 31-38.
[4] S. Bērziša, "XML-based Specification of the Project Management Domain and Its Application" in: *Databases and Information Systems VI. Volume 224 Frontiers in Artificial Intelligence and Applications*, J. Barzdins and M. Kirikova, Ed. Amsterdam: IOS Press, 2011, pp. 213-226.
[5] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches", AI Communications, vol. 7, pp. 39–59, 1994.
[6] Workflow Managemen Coaltion, "Process Definition Interface -- XML Process Definition Language", Workflow Management Coalition Workflow Standard, WFMC-TC-1025, 2008. [Online]. Available: http://www.wfmc.org/xpdl.html. [Accessed: September 28, 2011].
[7] G.M. Redding, "Object-centric Process Models and the Design of Flexible Processes," PhD thesis, Queensland University of Technology, Brisbane, Australia, 2009.
[8] OMG, "OMG Unified Modeling Language (OMG UML), Superstructure", OMG Specification, formal/2010-05-05, 2011. [Online]. Available: http://www.omg.org/spec/UML/2.3/Superstructure. [Accessed: September 28, 2011].
[9] W. Reisig and G. Rozenberg, Eds. *Lectures on Petri Nets I: Basic Models*, LNCS vol.1491, Berlin/Heidelberg: Springer, 1998.
[10] E. Blankenship, M. Woodward, G. Holliday and B. Keller, *Professional Team Foundation Server 2010*. Indianapolis: Wrox, 2011.
[11] Atlassian, "JIRA". [Online]. Available: http://www.atlassian.com/software/jira/ [Accessed: September 28, 2011].
[12] R. Dijkman, M. Dumas, B.V. Dongen, K. Reina and J.Mendling, "Similarity of Business Process Models: Metrics and Evaluation", Information Systems, vol. 36(2), pp. 498-516, Apr. 2011.
[13] OMG, "Business Process Model and Notation (BPMN)", OMG Specification, formal/2011-01-03, 2011. [Online]. Available: http://www.omg.org/spec/BPMN/2.0. [Accessed: September 28, 2011].
[14] W.M.P. van der Aalst and A.H.M. ter Hofstede, "YAWL : Yet Another Workflow Language", Information Systems, vol. 30(4), 245-275, Jun. 2005.
[15] D. Gochberg and R. Stewart, *Microsoft Office Project Server 2007 : The Complete Reference*, New York: McGraw-Hill Professional Publishing, 2008.
[16] M. Ehrig, A. Koschmider and A. Oberweis "Measuring Similarity between Semantic Business Process Models" in: *Proceedings of the Fourth Asia-Pacific Conference on Conceptual Modelling*, Darlinghurst: Australian Computer Society, 2007, pp.71-80.
[17] S. Nejati, M.Sabetzadeh, M. Chechik, S. Easterbrook and P. Zave, "Matching and Merging of Statecharts Specifications", in: *Proceedings of the 29th International Conference on Software Engineering*. Washington: IEEE Computer Society, 2007, pp. 54-64.

Solvita Berzisa is a Doctoral student of Riga Technical University, at the program of Information Technology. She obtained a BSc (2005) and MSc (2007) in Computer Science and Information Technology from Riga Technical University. Her research areas are IT project management, configuration of project management information systems, as well as project management knowledge.
She currently works at Riga Technical University as an Assistant and at Exigen Services Latvia as a System Analyst – Programmer.

Janis Grabis is a Professor at the Faculty of Computer Science and Information Technology, Riga Technical University, Latvia. He obtained his Doctoral degree from Riga Technical University in 2001 and worked as a Research Associate at the College of Engineering and Computer Science, the University of Michigan-Dearborn. He has published articles in major academic journals including OMEGA, European Journal of Operational Management, International Journal of Production Research, Human Systems Management and others. He has been a guest-editor for two top academic journals and member of the program committee of several academic conferences. His research interests are enterprise applications, project management and optimization.

Solvita Bērziša, Jānis Grabis. Projektu vadības procesu līdzības novērtēšana un atkalizmantošana

Projektu vadības procesi apraksta projektu vadības aktivitātes un to izpildes secību. Katram procesam ir definēta kopa ar ieejas un izejas datiem. Projektu vadības informācijas sistēmas tiek izmantotas, lai uzglabātu datus un koordinētu procesu izpildi. Projektu vadības informācijas sistēma tiek konfigurēta katra konkrētā projekta vajadzībām. Lai atvieglotu un uzlabotu konfigurācijas procesu, ir izmantotas uzkrātās zināšanas. Šīs zināšanas satur gan datus, gan procesus. Datu analīzei ir izmantota nosaukumu semantiskā analīze, grupēšana un kārtošana, bet procesu analīzei ir vajadzīga citādāka pieeja, kas ir apskatīta šajā rakstā. Raksta mērķis ir izstrādāt pieeju projektu vadības procesu līdzības novērtēšanai un atkalizmantošanai.

Rakstā piedāvātā pieeja ir daļa no projektu vadības informācijas sistēmu konfigurēšanas procesa un sastāv no trīs fāzēm: procesu formalizācijas, procesu līdzības novērtēšanas un analīzes. Procesu formalizācijā tiek veikta visu analizējamo procesu transformēšana uz statusu/transakciju formātu. Procesu līdzības novērtēšanai tiek izmantoti semantiskie mērījumi un sinonīmu vārdnīca, kurā ir saglabātas statusu sinonīmu vērtības. Par līdzīgiem ir uzskatīti procesi, kuru līdzības vērtība pārsniedz definēto slieksni. Savstarpēji līdzīgie procesi veido līdzīgo procesu grupas. Analīzes fāzē katrai līdzīgo procesu grupai tiek veikta statusu un transakciju semantiskā analīze, grupēšana un kārtošana. Rezultātā lietotājs saņem informāciju par procesu variācijām un statusiem un transakcijām katrā procesu variantā. Pieejas demonstrācijai ir izveidots piemērs, kurā iegūtie dati ir salīdzināti ar datiem, kādus iegūtu, ja netiktu ņemta vērā procesu līdzība. Rakstā ir apskatīta arī projektu vadības informācijas sistēmu konfigurēšanas pieeja, dažādās pieejas procesu attēlošanā un eksistējošie pētījumi un metodes procesu līdzības novērtēšanai.

Galvenais izstrādātās pieejas ieguldījums projekta vadības procesu zināšanu atkalizmantošanā ir plašākas interpretācijas iespējas par procesu variācijām un statusu/transakciju izmantošanas tendencēm. Pieejas iegūto iznākumu ietekmē izvēlētais līdzības slieksnis un procesu līdzības metrika.

Солвита Берзиша, Янис Грабис. Оценка сходства и повторное использование процессов управления проектами

Процессы управления проектами описывают деятельность по управлению проектами и порядок их выполнения. Для каждого процесса определяется множество входных и выходных данных. Информационные системы управления проектами используются для того, чтобы хранить данные и координировать выполнение процессов. Информационная система управления проектами настраивается для нужд каждого конкретного проекта. Чтобы облегчить и улучшить процесс настройки системы, используются накопленные знания. Эти знания содержат как процессы, так и данные. Для анализа данных используется семантический анализ названий, сортировка и группировка; в свою очередь, для анализа процессов необходим другой подход, который и описан в данной статье.

Описанный в статье подход является частью настройки процессов информационных систем управления проектами и состоит из трех фаз: формализация процессов, оценка схожести процессов и анализ. На этапе формализации процессов осуществляется трансформация всех анализируемых процессов в формат статусов/транзакций. Для оценки схожести процессов используются семантические измерения и словарь синонимов, в котором сохраняются значения синонимов статусов. Схожими считаются процессы, мера схожести которых превышает установленный порог. Схожие между собой процессы образуют группы схожих процессов. В фазе анализа для каждой группы схожих процессов производится семантический анализ статусов и транзакций, группировка и сортировка. В результате пользователь получает информацию о вариациях процесса, его статусах и транзакциях в каждом варианте процесса. Для демонстрации данного подхода разработан пример, данные которого сравниваются с данными, которые были бы получены в случае, если не принимать во внимание схожесть процессов. В статье также описан подход к настройке информационной системы управления проектами, различные подходы в отображении процессов, а также существующие исследования и методы оценки схожести процессов.

Главный вклад разработанного подхода в повторное использование знаний в процессах управления проектами - более широкие возможности интерпретации вариаций процессов и тенденций использования статусов/транзакций. На полученные данные наибольшее влияние оказывают выбранный порог схожести и мера схожести процессов.