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.

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.

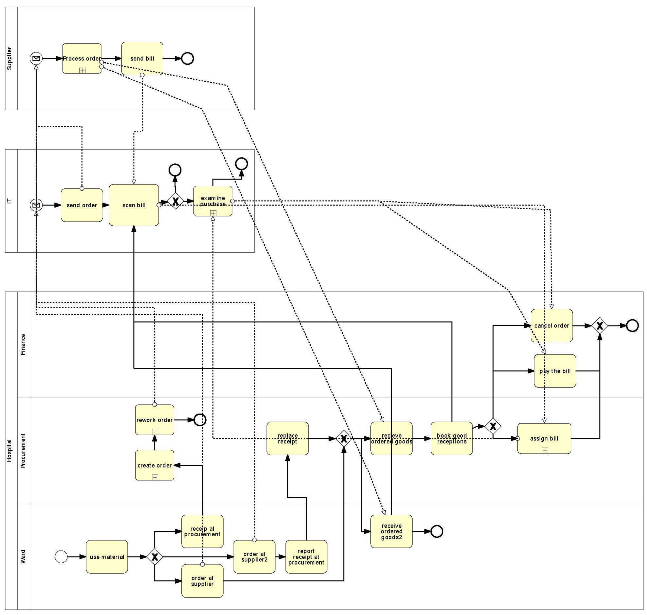

Figure 1. As-is process-KHA

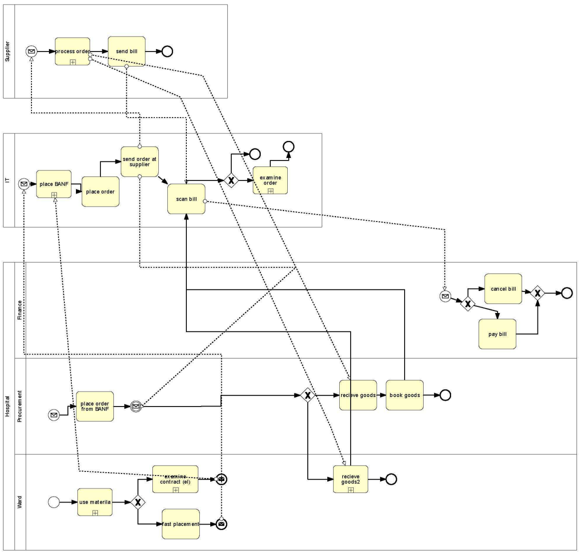

Figure 2. As-is process- KHB

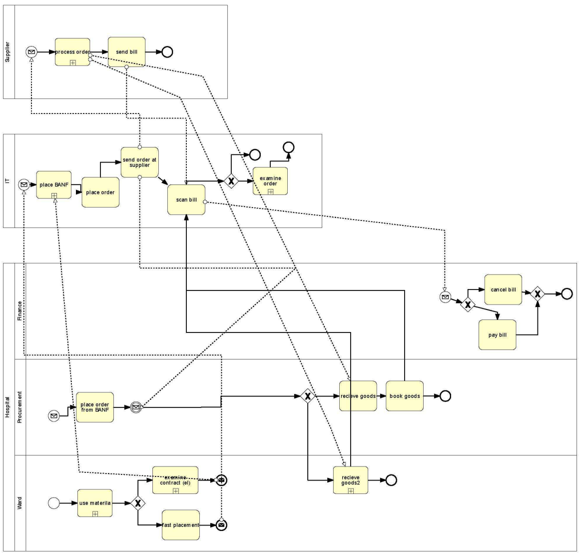

Figure 3. Merged process: KHPlanned

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

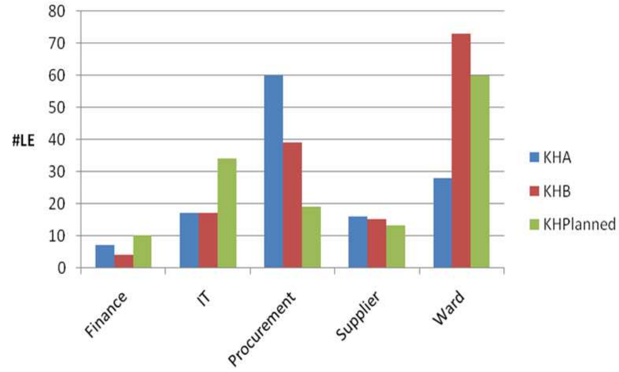

Figure 4. Activity of process actors

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 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.