PROCESS RECONSTRUCTION AND VISUALISATION FOR COLLABORATIVE ENGINEERING PROJECTS

Lei Shi
Department of Mechanical Engineering
University of Bath
Bath
BA2 7AY, UK
l.shi @bath.ac.uk

Linda Newnes
Department of Mechanical Engineering
University of Bath
Bath
BA2 7AY, UK
l.b.newnes@bath.ac.uk

Steve Culley
Department of Mechanical Engineering
University of Bath
Bath
BA2 7AY, UK
enssjc@bath.ac.uk

Chris Snider
Department of Mechanical Engineering
University of Bristol
Bristol
BS8 1TR, UK
chris.snider@bristol.ac.uk

ABSTRACT

Many modern engineering projects are required to coordinate multiple project teams, utilise distributed resources and integrate knowledge across multiple disciplines. Hence, the execution of these projects needs to involve high volume of remote communications, synchronous or asynchronous interactions, complicated decision-making and control processes. By considering these factors, human-centred management approaches could be problematic at times to handle the large amount of project data and manage the complex project processes. In order to improve the management efficiency and effectiveness, this paper presents an automatic approach on process reconstruction and process visualisation. As shown in a case study, this approach has the potential to support the information reuse, capture process dynamics, enhance process comprehensibility, as well as reduce human intervention in general process management.

Keywords: collaborative engineering projects, process management, process reconstruction, process visualisation.

1 INTRODUCTION

Collaborative working plays a vital role in the operation and management of modern engineering projects (Patel et al., 2012). Most engineering projects are required to coordinate multiple project teams, utilise distributed resources and integrate knowledge across multiple disciplines. Under these circumstances, remote communications, synchronous or asynchronous interactions, complicated decision-making and control processes are considered to be critical for executing the projects. Meanwhile, project members need to perform remote data sharing with project partners, thus large amount of distributed data including communication related, operation-related and management-related could be generated during the project execution.

In practice, human-centred management approaches have been applied in various organisations for the purposes of managing the essential production processes and improving the user-centred design efficiency (Earthy et al., 2001). However, due to the information overload problem caused by the large data amount, as well as by the frequently changed user requirements and market demands, the application of these approaches become to be inefficient or problematic at times, especially when complex processes are excessively involved by the project or the project members are under high working pressures (Zika-Viktorsson et al., 2006).

From process management point of view, the process of an engineering project can be represented by a collection of ordered activities (see Figure 1). Each activity is supposed to generate a single or multiple outcomes in either virtual or physical form. These activities are performed at different stages of the project life-cycle, and they are typically determined by the factors such as project objectives, resource availability, execution outcomes and relevant decisions (Aguilar-Savén, 2004). During project execution, a project process could have a dynamic structure that reflects the initial data input or current project status, e.g., the initial requirements of customer, decisions from project member or outcomes of previous activities. For each project process, the complexity of its structure can be measured based on certain factors such as the involved project members/teams and consumed time or resources. To determine subsequent activities of an ongoing process, the outcomes generated by the previous activities, together with the current project status should be taken into account collectively.

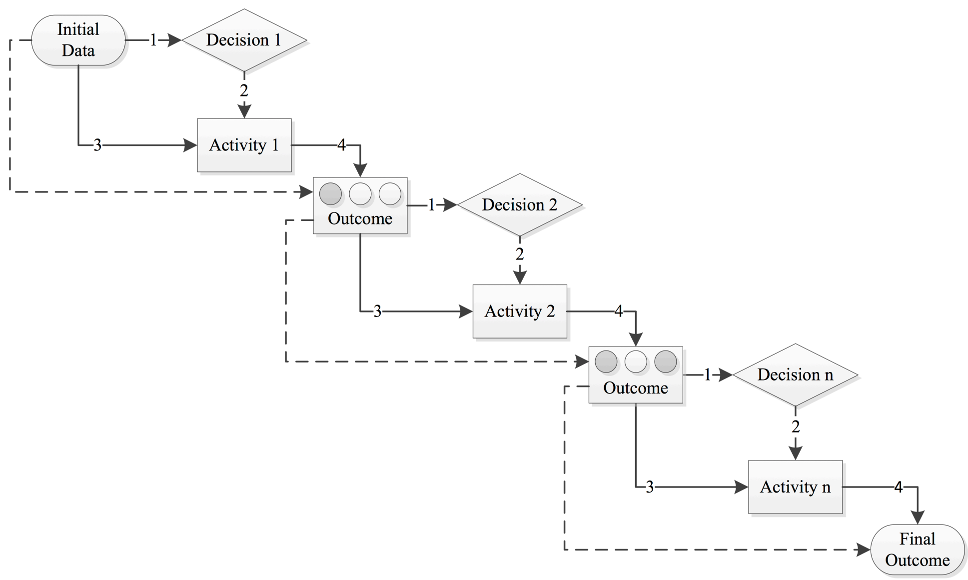

Figure 1 A process sample of engineering projects

The process development for an engineering project requires multiple iteration steps. The outcome of current iteration step is applied to determine the activity of next iteration step. Meanwhile, such outcome can also be applied to assess whether the objective(s) related to the current activity has been achieved. In other words, giving a project with a dataset, the actual process of the project can be reconstructed and represented by analysing the content of the data. Based on this idea, a data-driven approach on process reconstruction and process visualisation is proposed in this paper. It aims to improve the automation of process management, and also to reduce the human intervention in related decision-making processes.

The remainder of this paper is organised as follows. Section 2 reviews the related work in the fields of process management. Section 3 introduces the proposed approach. Section 4 describes a case study using industrial data. Section 5 concludes the paper.

2 RELATED WORK

Process management is considered to be an effective way to organise the project activities and to reduce the product development cost/risk (Ford and Sterman, 1998, Browning and Eppinger, 2002). The aspects of process management mainly cover process modelling, process reuse, complexity identification, process standardisation and process optimisation (Aguilar-Savén, 2004).

A major consideration of applying process management to engineering projects is about increasing the capability of understanding the inner relations among the activities, and reducing the time cost and human intervention in project execution. Pino et al. (2008) indicated that the application of process management could improve the execution efficiency of software engineering projects. The improvements cover the aspects of general management, product documentation, user requirements management, process establishment, configuration management and requirements elicitation. Zantek et al. (2002) stated that the management of manufacturing processes in an enterprise could facilitate the identification of product quality variations. Lerner et al. (2010) revealed that process management is particularly useful for identifying the exception patterns contained by project processes, and it could also be applied to capture the relations between the exception patterns and the normative processes.

In production environments, the application of process management could be restricted by certain constraints. For example, a process management model might be applied in early project stages, and then ignored intentionally at the following stages, if the time or resource constraints become tight (Michael Gnatz, 2004). In order to improve the usability of process management models, a critical requirement is to improve the automation level. As a solution, computer aided technologies are necessary to be integrated with the models. The main reason is that the digitalised data generated by electronic collection methods or ICT-based tools can be directly applied to implement project process monitoring (Elazouni and Salem, 2011). In automotive industry, ICT-based process management has been used to reduce the manual work and support process development related tasks (Müller et al., 2006). Similar technologies are also considered as useful for supporting process planning and product life-cycle management (Ming et al., 2008). To further improve the automation level of process management, the application of data mining and machine learning are necessary. Meier et al. (2006) proposed an automatic approach for sequencing product design processes based machine learning algorithms. By using data mining, Shi et al. (2014b) proposed an automatic approach on process structure comparison and process normality identification. That research demonstrated that certain patterns discovered from project data could facilitate the understanding of inner relations among the process-related activities.

Based on this review, increasing the automation level of process management is a major concern in practice. To achieve this, a data-driven approach on process reconstruction and visualisation is proposed in the following section.

3 DATA-DRIVEN PROCESS RECONSTRUCTION AND VISUALISATION

Understanding large amount of project data on a detailed level is difficult. From the information management perspective, the application of data mining could be a solution to solve this issue. In engineering domain, data mining can be applied to perform various types of analysis including predictive maintenance, fault detection, quality control and customer relationship management (Köksal et al., 2011). Furthermore, it can also be applied to support large-scale information management and knowledge discovery (Choudhary et al., 2009). Under this context, the use of data mining in process management could enable project members to comprehensively understand the large amount of project data without consuming lots of time or resources.

In order to reconstruct a project process, two types of information need to be identified from the project data, which are semantic feature and process-related feature (Shi et al., 2014b, Shi et al., 2014a). Most of the descriptive information contained in project data has textual format, thus data mining in conjunction with natural language processing are applied to perform feature identification. Other applied technologies include named entity recognition, frequency analysis and sequence analysis. During process management, the knowledge related to the project scope is captured, and then modelled as knowledge bases. The knowledge concepts contained in a knowledge base could reflect the expertise of the domain experts. It is can be treated as a high-level guidance for the analytical tasks such as feature classification, feature weighting and dependency identification. In other words, the knowledge base is applied to filter, weight and organize the identified features. It ensures that the features with less importance can be eliminated, so that only the essential ones will be applied for process reconstruction.

To demonstrate the functionality of proposed approach, a case study based on industrial data is included in the following section.

4 CASE STUDY

In this case study, a dataset captured from a manufacturing organisation is applied. It contains 1000 maintenance projects that are implemented between 2012 and 2014. The project data is in textual format, which contains the information about project objectives, technical details, proposed solutions, performed evaluations, communications and workflow. The case study is demonstrated as two parts,

i) feature modelling, ii) process reconstruction/visualisation, and the detailed information is introduced as below.

4.1 Feature Modelling

Given a project, the feature modelling process employs data mining and natural language processing to analyse the project data. During this process, a feature set is generated for the project, whilst the timestamp of each feature is also identified. The domain-specific knowledge is captured from the experienced staff in the organisation, and then is applied to guide the analytical tasks to filter and weight the identified features. In this case study, a list of features (with a total number 21) are considered critical. These features cover the following aspects, i) indicating the key activities, ii) representing the essential communications, and iii) representing the technical details. Figure 2 shows the quantity of the features regarding the projects, and Figure 3 shows the distribution of such features[1].

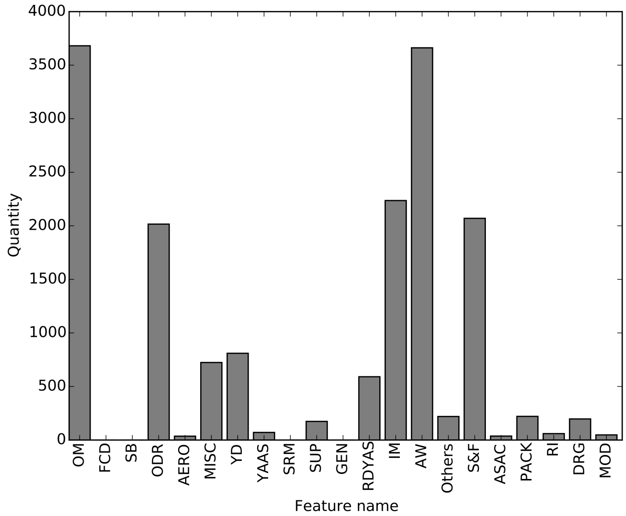

Figure 2 Feature quantity

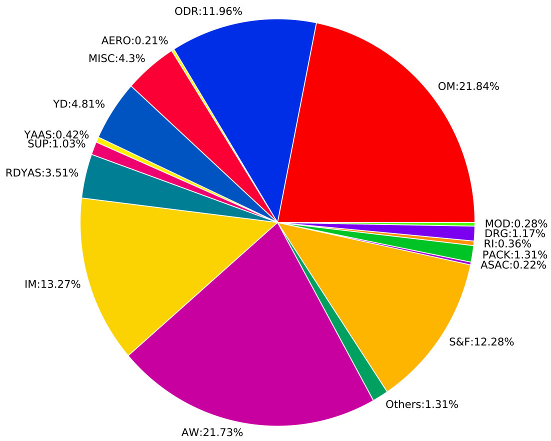

Figure 3 Feature distribution

The generation of feature distribution is a pre-process for automating process management. It aims to draw a sketch of the dataset structure from a bottom-up perspective. The information contained in a feature distribution is considered as an additional guidance for optimising the feature weighting and feature filtering related tasks.

4.2 Process Reconstruction and Visualisation

During process reconstruction, the identified features need to be further processed. A main task here is to weight the features, and then filter out certain ones based on defined thresholds. By considering the feature distribution and knowledge base, some features will be assigned higher weightings than the others, if the former ones have higher correlations with the key activities than the latter ones. Only the features with acceptable level of weightings are applied to reconstruct the project process. Figure 4 shows the visualisation of some processes generated from the project data. In this visualisation, each row indicates a single process, and each ***Tx*** in the process indicates a single feature.

According to the approach proposed by the previous study (Shi et al., 2014b), the similarity between any two reconstructed processes is measureable. For example, projects P1 and P2 have a high similarity value 0.7185. As shown in the figure, common patterns appeared frequently in both of them: in the early project stage, the pattern is [***T1***, ***T9***]; in the middle project stage, the pattern is [***T1***, ***T9***, ***T13***, ***T15***, ***T1***]; and in the late project stage, the patterns are [***T9***, ***T11***] and [***T15***, ***T13***, ***T15***, ***T9***]. By inspecting the data content, the reason of high similarity value between P1 and P2 is that both of them are about maintaining a mature product, so that they have similar objectives, design processes

[1] due to confidentiality reasons, some feature names have been modified.

and technical requirements. In contrast, the similarity value between project P6 and the others is close to zero. As shown in the visualisation, very limited common patterns are contained in its process when comparing to the others. The reason is that P6 is about maintaining a new product, so that the implementation requires additional preparation and different procedures, leading to its process being different from the others.

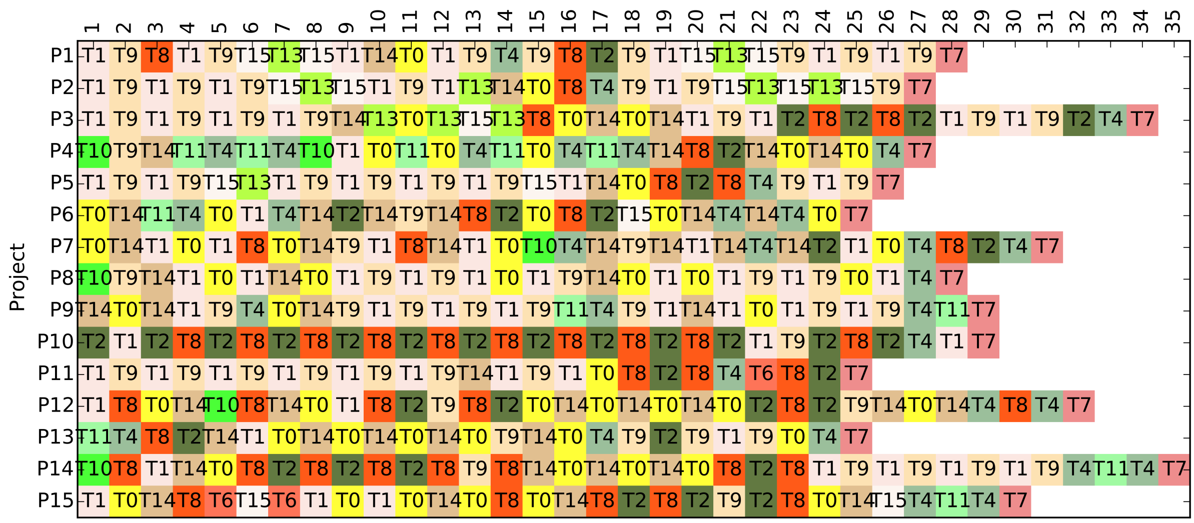

Figure 4 A visualisation of modelled project processes

The reconstructed process can be treated as a compressed representation of project execution process, because it mainly represent the information generated by the key operation and management activities. Due to the fact that the information with low relevance or minor importance has been eliminated, this visualisation can be used to facilitate the understanding of project processes with complex structures. Furthermore, it can also help project members to assess project characteristics, process evolutions, and enable them to compare multiple project processes simultaneously.

5 DISCUSSION AND CONCLUSION

This paper proposed an data-driven approach to reconstruct and visualise engineering project processes, by applying data mining and natural language processing technologies with modelled knowledge. This approach aims to increase the automation of process management and the understanding capability of process structure. It also has the potential to support project information reuse, capture process dynamics, enhance process comprehensibility, and reduce human intervention in general process management. In the case study, the usability of the approach has been demonstrated by using industrial data.

ACKNOWLEDGMENTS

The research reported in this paper is funded by Engineering and Physical Sciences Research Council (EP/K014196/1). The authors would like to thank the industrial collaborators and their engineers for their input and support on this project.

REFERENCES

AGUILAR-SAVÉN, R. S. 2004. Business process modelling: Review and framework. *International Journal of Production Economics,* 90**,** 129-149.

BROWNING, T. R. & EPPINGER, S. D. 2002. Modeling impacts of process architecture on cost and schedule risk in product development. *Engineering Management, IEEE Transactions on,* 49**,** 428-442.

CHOUDHARY, A. K., HARDING, J. A. & TIWARI, M. K. 2009. Data mining in manufacturing: a review based on the kind of knowledge. *Journal of Intelligent Manufacturing,* 20**,** 501-521.

EARTHY, J., JONES, B. S. & BEVAN, N. 2001. The improvement of human-centred processes—facing the challenge and reaping the benefit of ISO 13407. *International Journal of Human-Computer Studies,* 55**,** 553-585.

ELAZOUNI, A. & SALEM, O. A. 2011. Progress monitoring of construction projects using pattern recognition techniques. *Construction Management and Economics,* 29**,** 355-370.

FORD, D. N. & STERMAN, J. D. 1998. Dynamic modeling of product development processes. *System Dynamics Review,* 14**,** 31-68.

KÖKSAL, G., BATMAZ, İ. & TESTIK, M. C. 2011. A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications,* 38**,** 13448-13467.

LERNER, B. S., CHRISTOV, S., OSTERWEIL, L. J., BENDRAOU, R., KANNENGIESSER, U. & WISE, A. 2010. Exception handling patterns for process modeling. *Software Engineering, IEEE Transactions on,* 36**,** 162-183.

MEIER, C., YASSINE, A. A. & BROWNING, T. R. 2006. Design Process Sequencing With Competent Genetic Algorithms. *Journal of Mechanical Design,* 129**,** 566-585.

MICHAEL GNATZ, M. D., MICHAEL MEISINGER 2004. Towards an integration of process modeling and project planning. *IET Conference Proceedings***,** 22-31.

MING, X. G., YAN, J. Q., WANG, X., LI, S., LU, W. F., PENG, Q. & MA, Y. 2008. Collaborative process planning and manufacturing in product lifecycle management. *Computers in Industry,* 59**,** 154-166.

MÜLLER, D., HERBST, J., HAMMORI, M. & REICHERT, M. 2006. *IT support for release management processes in the automotive industry*, Springer.

PATEL, H., PETTITT, M. & WILSON, J. R. 2012. Factors of collaborative working: A framework for a collaboration model. *Applied Ergonomics,* 43**,** 1-26.

PINO, F. J., GARCÍA, F. & PIATTINI, M. 2008. Software process improvement in small and medium software enterprises: a systematic review. *Software Quality Journal,* 16**,** 237-261.

SHI, L., GOPSILL, J., SNIDER, C., JONES, S., NEWNES, L. & CULLEY, S. Towards Identifying Pattern in Engineering Documents to Aid Project Planning. DS 77: Proceedings of the DESIGN 2014 13th International Design Conference, 2014a.

SHI, L., GOPSILL, J. A., NEWNES, L. & CULLEY, S. 2014b. A Sequence-Based Approach to Analysing and Representing Engineering Project Normality. *2014 IEEE 26th International Conference on Tools with Artificial Intelligence (ICTAI).* Limassol, Cyprus.

ZANTEK, P. F., WRIGHT, G. P. & PLANTE, R. D. 2002. Process and product improvement in manufacturing systems with correlated stages. *Management Science,* 48**,** 591-606.

ZIKA-VIKTORSSON, A., SUNDSTRÖM, P. & ENGWALL, M. 2006. Project overload: An exploratory study of work and management in multi-project settings. *International Journal of Project Management,* 24**,** 385-394.