

Footprint based Workflow Model for Analytical Task Automation

Minjae Park*, Jin Ma**, Hyunah Kim***, Hyun Ahn***, Kwnaghoon Pio Kim***

*Engineering Innovation Center, BISTel, Inc., Seoul, South Korea

**National Institute of Supercomputing and Networking, KISTI, Daejeon, South Korea

*** Dept. of Computer Science, Kyonggi University, Suwon, Kyonggi-do, South Korea

mjpark@bistel.com, majin@kisti.re.kr, {hyuna486, hyahn, kwang}@kgu.ac.kr

Abstract— Generally, analytical tasks target some specific problem that requires engineering expertise to be solved efficiently. However, it is wasteful to use an engineer's expertise to address repetitive tasks that are performed iteratively. Therefore, in this paper, we propose a method to automate analytical tasks using the footprint of an approach previously followed by an engineer. Using this method, we can discover and create executable workflow models and use it to automate analytical tasks in manufacturing data analysis.

Keywords— Workflows, Processes, Workflow/Process Modeling, Footprint, Workflow Discovery

I. INTRODUCTION

Workflow modeling methods have been studied extensively [1, 2, 3, 4, 5, 6, 7, 8]. These methods typically model workflows using modelers or designers, but there are also methods based on workflow/process mining [8, 9, 10, 11, 12, 13] that do not require human intervention.

Analytical tasks usually consist of the following steps—visualization, articulation, conceptualization, and decision making for simple/complex problem solving using available information. They are used in a variety of fields, such as manufacturing, finance, and chemistry. Anyway, it depends on the skill of the analyst engineer of specific area.

In this study, we focus on analytical tasks associated with manufacturing. We propose a novel workflow modeling method for automating such tasks.

II. WORKFLOWS

The workflow concept is a well-known technique used to automate a variety of tasks. Workflows are typically divided into two phases: modeling during the build time, and execution at run time. A workflow management system defines, creates, and manages the execution of workflows using software that uses one or more workflow engines to interpret the process definition, interact with workflow participants, and invoke the use of IT tools and applications, if required.

These workflows and workflow management systems can be used for manufacturing analysis. In particular, we are interested in the automation of analytical tasks that can be executed

without human involvement. Once the workflow model is created, the engine in the workflow management system will automate it.

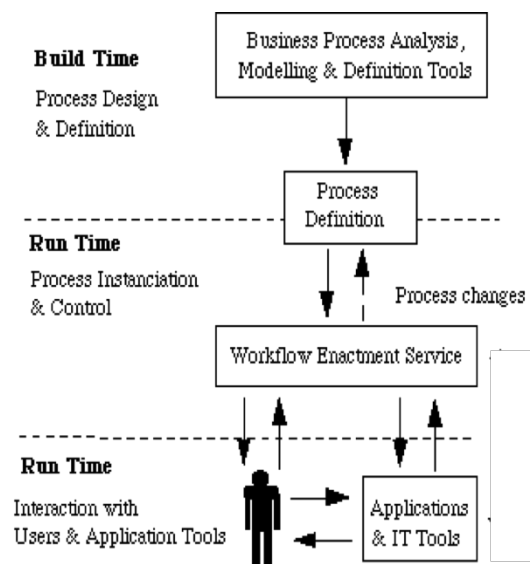


Figure 1. A Workflow Management System

Analysts carry out analytical tasks that are associated with a specific application and set of parameter data. We collect their footprint for analytical tasks and use it to create a workflow model. There are other workflow discovery methods such as the log-driven [11, 12] and model-driven methods [9, 10]. However, in contrast to the proposed method, these techniques mine the workflow process to generate the model from the scratch, and understand its meaning. In our method, we automate the repetitive analysis using footprint information associated with the specific task.

III. FOOTPRINT BASED WORKFLOW MODELING METHOD FOR ANALYTICAL TASK AUTOMATION

We use analytical data to create the workflow model. Three important aspects are considered—the task type, user action, and task history. The task type describes a work step of an analysis task. The user action describes a set of tasks related to each other. The task history refers to the records associated with

the user action. Depending on the task history, executable workflows are created. In this case, we use the BPMN 2.0 standard [14, 15], a well-known workflow model standard that can be executed on compatible workflow management systems.

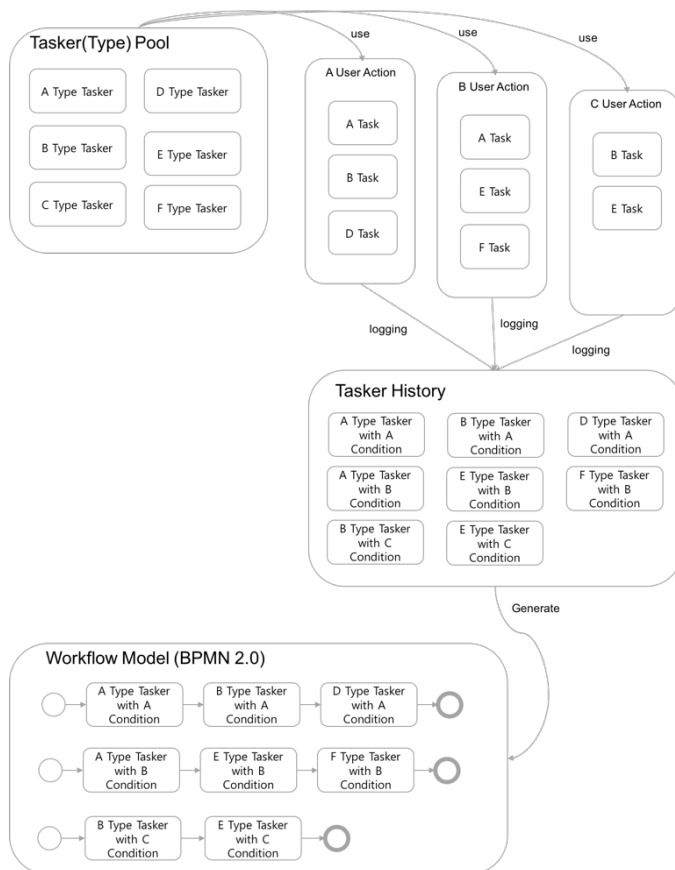


Figure 2. Workflow model creation using user actions

IV. AN EXAMPLE OF FOOTPRINT BASED WORKFLOW MODEL (DISCOVERED WORKFLOW MODEL)

We used the Activiti [16] workflow management system to create and run the BPMN model. The BPMN model and workflow related information is managed using a database system.

NAME	DEPLOYMENT_ID	BYTES
1609_2756_2016061615031016.bpmn20.xml	267	(BLOB)
1609_2756_2016061615033654.bpmn20.xml	268	(BLOB)
1609_2756_2016061617163555.bpmn20.xml	272	(BLOB)
1609_2756_2016062018061119.bpmn20.xml	298	(BLOB)
1609_2756_2016062018065727.bpmn20.xml	299	(BLOB)
1609_2756_2016062115591094.bpmn20.xml	306	(BLOB)
1609_2756_2016062115592213.bpmn20.xml	307	(BLOB)
1609_2756_20160621171844107.bpmn20.xml	315	(BLOB)
1609_2756_2016062117235617.bpmn20.xml	317	(BLOB)
1609_2756_2016062117244726.bpmn20.xml	318	(BLOB)

Figure 3. Workflow list managed by Activiti DB

Figure 3 shows a list of workflows created using the task history and the structured data tables. Figure 4 shows the executable model information. Using this information, analysis tasks that are performed manually and repeatedly by engineers can be automated and implemented as workflow models. In

addition to this model information, details about the implemented application and related parameters are also needed. The generated workflow model is executed on the engine. Figure 5 shows an example of the run history generated by the workflow engine.



Figure 4. An Analytic workflow model (BPMN model on Activiti DB)

ID	PROC_INST_ID	BUSINESS_KEY	PROC_DEF_ID	START_TIME	END_TIME
1 200001 200001	(null)	3699_4478_201607011642424:1:351	16/07/04 14:24:32.439000000	16/07/04 14:24:33.331000000	
2 200001 320001	(null)	3699_4478_201607011642424:1:351	16/07/04 15:04:06.891000000	16/07/04 15:04:06.895000000	
3 322501 322501	(null)	3699_4478_201607011642424:1:351	16/07/04 15:04:06.891000000	16/07/04 15:04:06.895000000	
4 327501 327501	(null)	55_73_2016060118045977:1:197	16/07/13 12:28:39.285000000	16/07/13 12:28:40.466000000	
5 282501 282501	(null)	55_73_2016060118045977:1:197	16/06/21 16:47:18.542000000	16/06/21 16:47:19.386000000	
6 187501 187501	(null)	standardanalyticprocess_2016051216563092:1:24	16/05/12 16:56:33.271000000	16/05/12 16:56:33.285000000	
7 110001 110001	(null)	standardanalyticprocess_2016051314100856:1:25	16/05/13 14:10:11.851000000	16/05/13 14:10:11.861000000	
8 247501 247501	(null)	55_73_2016060118045977:1:197	16/06/15 15:11:43.897000000	16/06/15 15:11:43.919000000	
9 252501 252501	(null)	55_73_2016060118045977:1:197	16/06/15 15:23:01.789000000	16/06/15 15:23:01.814000000	
10 285001 285001	(null)	55_73_2016060118045977:1:197	16/06/21 16:51:36.869000000	16/06/21 16:51:37.725000000	
11 302501 302501	(null)	3699_4478_201607011642424:1:351	16/07/04 14:26:23.325000000	16/07/04 14:26:24.081000000	
12 317501 317501	(null)	3699_4478_201607011642424:1:351	16/07/04 15:02:52.185000000	16/07/04 15:02:52.899000000	
13 112501 112501	(null)	standardanalyticprocess_2016051711050824:1:14	16/05/17 11:05:11.420000000	16/05/17 11:05:11.516000000	
14 252501 252501	(null)	55_73_2016060118045977:1:197	16/06/15 15:23:01.788000000	16/06/15 15:23:01.812000000	
15 252501 252501	(null)	55_73_2016060118045977:1:197	16/06/15 15:24:00.449000000	16/06/15 15:24:00.451000000	
16 27501 27501	(null)	3699_4478_201607011642424:1:351	16/07/13 12:22:04.891000000	16/07/13 12:22:05.677000000	
17 330001 330001	(null)	55_73_2016060118045977:1:197	16/06/21 16:45:04.540000000	16/06/21 16:45:05.978000000	
18 36501 36501	(null)	3699_4478_201607011642424:1:351	16/07/04 14:58:29.634000000	16/07/04 14:58:30.512000000	
19 36501 36501	(null)	3699_4478_201607011642424:1:351	16/07/20 15:11:26.734000000	16/07/20 15:11:27.520000000	
20 312501 312501	(null)	3699_4478_201607011642424:1:351	16/07/04 15:02:52.185000000	16/07/04 15:02:52.895000000	
21 335001 335001	(null)	3699_4478_201607011642424:1:351	16/07/04 15:02:52.185000000	16/07/04 15:02:52.895000000	
22 245001 245001	(null)	55_73_2016060118045977:1:197	16/06/15 15:18:24.892000000	16/06/15 15:18:24.913000000	
23 245001 245001	(null)	55_73_2016060118045977:1:197	16/06/15 15:18:24.892000000	16/06/15 15:18:24.913000000	
24 250001 250001	(null)	55_73_2016060118045977:1:197	16/06/15 15:18:24.892000000	16/06/15 15:18:24.913000000	
25 115001 115001	(null)	standardanalyticprocess_2016051711140293:1:142	16/05/17 11:14:04.396000000	16/05/17 11:14:04.410000000	
26 245001 245001	(null)	55_73_2016060118045977:1:197	16/06/15 15:09:53.651000000	16/06/15 15:09:53.726000000	
27 307501 307501	(null)	3699_4478_201607011642424:1:351	16/07/04 14:59:47.148000000	16/07/04 14:59:47.950000000	
28 335001 335001	(null)	3699_4478_201607011642424:1:351	16/07/13 17:43:37.096000000	16/07/13 17:43:37.950000000	

Figure 5. Workflow execution history using the footprint based workflow model

V. CONCLUSIONS

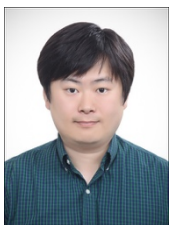
In this paper, we proposed a novel automated workflow modeling method based on the footprint of analytical task history. In order to automate the analytical tasks performed by an engineer, we created a workflow model. The modeled workflows can be automated by using workflow management systems. We expect this technique to be used for automotive analytical tasks without having to repeat manual tasks. It can also be applied directly to the industrial field to facilitate efficient analysis.

ACKNOWLEDGMENT

This research was supported by the contents convergence software research center at Kyonggi University funded by the GRRC program of Gyeonggi Province, South Korea.

REFERENCES

- [1] Ellis, C. A. (1979, August). Information control nets: A mathematical model of office information flow. In *Proceedings of the Conference on Simulation, Measurement and Modeling of Computer Systems* (Vol. 3670).
- [2] Ellis, C. A., & Nutt, G. J. (1993). *The Modelling and Analysis of Coordination Systems*. University of Colorado, Boulder, Department of Computer Science.
- [3] Ellis, C. A. (1983, September). Formal and Informal Models of Office Activity. In *INFIP Congress* (pp. 11-22).
- [4] Sharp, A., & McDermott, P. (2009). *Workflow modeling: tools for process improvement and applications development*. Artech House.
- [5] Liu, D. R., & Shen, M. (2003). Workflow modeling for virtual processes: an order-preserving process-view approach. *Information Systems*, 28(6), 505-532.
- [6] Kappel, G., Lang, P., Rausch-Schott, S., & Retschitzegger, W. (1995). Workflow management based on objects, rules, and roles. In *IEEE Bulletin of the Technical Committee on Data Engineering*.
- [7] Van der Aalst, W. M. (1998). The application of Petri nets to workflow management. *Journal of circuits, systems, and computers*, 8(01), 21-66.
- [8] Dumas, M., Van der Aalst, W. M., & Ter Hofstede, A. H. (2005). *Process-aware information systems: bridging people and software through process technology*. John Wiley & Sons.
- [9] Kim, K. H., & Ellis, C. A. (2006). Workflow reduction for reachable-path rediscovery in workflow mining. In *Foundations and Novel Approaches in Data Mining* (pp. 289-310). Springer Berlin Heidelberg.
- [10] Park, M., & Kim, K. P. (2008). Control-path Oriented Workflow Intelligence Analyses. *J. Inf. Sci. Eng.*, 24(2), 343-359.
- [11] Van der Aalst, W. M. (2011). Getting the Data. In *Process Mining* (pp. 95-123). Springer Berlin Heidelberg.
- [12] Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... & Burattin, A. (2011, August). Process mining manifesto. In *International Conference on Business Process Management* (pp. 169-194). Springer Berlin Heidelberg.
- [13] Weijters, A. J. M. M., van Der Aalst, W. M., & De Medeiros, A. A. (2006). Process mining with the heuristics miner-algorithm. *Technische Universiteit Eindhoven, Tech. Rep. WP*, 166, 1-34.
- [14] Chinosi, M., & Trombetta, A. (2012). BPMN: An introduction to the standard. *Computer Standards & Interfaces*, 34(1), 124-134.
- [15] Silver, B., & Richard, B. (2009). *BPMN method and style* (Vol. 2). Aptos: Cody-Cassidy Press.
- [16] Rademakers, T. (2012). *Activiti in Action: Executable business processes in BPMN 2.0*. Manning Publications Co..



Minjae Park is a senior member of research staff at the solution R&D research center of BISTel, Inc., South Korea. He received B.S., M.S., and Ph.D. degrees in computer science from Kyonggi University in 2004, 2006, and 2009, respectively. His research interests include groupware, workflow systems, BPM, CSCW, collaboration theory, process warehousing and mining, workflow-supported social networks discovery and analysis, and process-aware factory automation systems.



Jin Ma is a Researcher at the Supercomputing R&D Center of KISTI (Korea Institute of Science and Technology Information), South Korea. He received B.S., M.S. degrees in computer software and computer science from Kwangwoon University in 2010, 2012, respectively. He had worked as member of research staff at the solution R&D research center of BISTel, Inc., South Korea. His research interests Big Data, Analysis System and Visualization, distributed system, BPM.



Hyunah Kim is a senior member of research staff at BIEE, Inc., and she is an adjunctive professor of the computer science department at Kyonggi University, South Korea. She received B.S. degree in computer science and engineering from Korea Nazarene University, and M.S. and Ph.D. degrees in computer science from Kyonggi University in 2005, and 2009, respectively. Her research interests include groupware, workflow systems, SCORM based e-Learning process modeling and management systems, e-Learning process mining techniques, and workflow-supported social networking knowledge discovery and analysis.



Hyun Ahn is a full-time Ph.D. student of computer science department and a graduate member of the collaboration technology research laboratory at Kyonggi University, South Korea. He received B.S. and M.S. degrees in computer science from Kyonggi University in 2011 and 2013, respectively. His research interests include workflow systems, BPM, scientific workflow systems, workflow-supported social and affiliation networks discovery, analysis, and visualization.



Kwanghoon Pio Kim is a full professor of computer science department and the founder and supervisor of the collaboration technology research laboratory at Kyonggi University, South Korea. He received B.S. degree in computer science from Kyonggi University in 1984. And he received M.S. degree in computer science from Chungang University in 1986. He also received his M.S. and Ph.D. degrees from the computer science department at University of Colorado Boulder, in 1994 and 1998, respectively. He had worked as researcher and developer at Aztek Engineering, American Educational Products Inc., and IBM in USA, as well as at Electronics and Telecommunications Research Institute (ETRI) in South Korea. In present, he is a vice-chair of the BPM Korea Forum. He has been in charge of a country-chair (Korea) and ERC vice-chair of the Workflow Management Coalition. He has also been on the editorial board of the journal of KSII, and the committee member of the several conferences and workshops. His research interests include groupware, workflow systems, BPM, CSCW, collaboration theory, Grid/P2P distributed systems, process warehousing and mining, workflow-supported social networks discovery and analysis, process-aware information systems, data intensive workflows, and process-driven Internet of Things.