

Chapter 21

Identifying Batch Processing Features in Workflows

Jianxun Liu

Hunan University of Science and Technology, China

Yiping Wen

Hunan University of Science and Technology, China

ABSTRACT

The employment of batch processing in workflow is to model and enact the batch processing logic for multiple cases of a workflow in order to optimize business processes execution dynamically. Our previous work has preliminarily investigated the model and its implementation. However, it does not figure out precisely which activity and how a/multiple workflow activity(s) can gain execution efficiency from batch processing. Inspired by workflow mining and functional dependency inference, this chapter proposes a method for mining batch processing patterns in workflows from process dataflow logs. We first introduce a new concept, batch dependency, which is a specific type of functional dependency in database. The theoretical foundation of batch dependency as well as its mining algorithms is analyzed and investigated. Based on batch dependency and its discovery technique, the activities meriting batch processing and their batch processing features are identified. With the batch processing features discovered, the batch processing areas in workflow are recognized then. Finally, an experiment is demonstrated to show the effectiveness of our method.

INTRODUCTION

The aim of batch processing in workflow is to improve the execution efficiency of business processes by modeling and enactment of batch logic for multiple workflow cases, i.e. vertically combining multiple *workflow activity* (hereafter activity) cases together and submitting for execution according to batch logics. In our previous literatures(Liu 2005, Liu 2007), we proposed a dynamic batch processing scheduling model and discussed the design and implementation of dynamic batch processing in WfMSs.

DOI: 10.4018/978-1-60566-669-3.ch021

However, there are still many problems need to be investigated. The following three problems are not solved yet: 1) determination of which *activity merits batch processing* (hereafter batch-efficient activity); 2) if it is a batch-efficient activity, how the multiple cases of the activity are batch-processed; and 3) the setting of the batch processing areas (batch processing patterns). In fact, just like the definition of workflow models, the design and modeling of batch-efficient activities as well as batch processing areas are also a time-consuming and error-prone task. They can be easily influenced by the perception of *business process designer* (hereafter designer). Moreover, designers may not know exactly which activity deserves batch processing (batch efficient activities) at workflow building time since there is no real data at that time to let us make a right decision. Even more designers may ignore these kinds of batch processing features due to their unconsciousness about the existence of batch efficiency in tasks. To optimize business processes, therefore, it is important to explore a way to identify and model both batch-efficient activities and batch processing areas automatically.

Inspired by data mining and workflow mining, this chapter proposes a method for identification of batch-efficient activities and their batch processing features as well as recognition of batch processing areas of a workflow from workflow logs. The basic idea of this method is to employ the control and relevant data in workflow logs to automatically discover the batch-efficient activities and their batch processing features, which is one or a set of input parameters of an activity and on which activities are batch-processed. To solve this problem, we introduce a new concept, batch dependency, which is a specific type of functional dependency in database. With the batch-efficient activities and their batch processing features, batch processing areas in workflow can be recognized and can be set automatically then. It is shown from the simulation experiments that our method works effectively.

The remainder of this chapter is organized as follows. First, in Section II we give an introduction of the background knowledge, notations, concepts and the problem itself for simplicity and clarification in identification and recognition algorithms. Section III is the theoretical foundation of our method, in which the batch dependency and its mining algorithm is introduced and investigated. In Section IV, we design two algorithms, one for identification of batch-efficient activities as well as recognition of their batch processing features and another for recognition of batch processing areas. Simulation experiments are done in Section V. Section VI is the review of related work. Our contribution is also pointed out here. Finally, Section VII concludes this chapter and points out some future directions.

PROBLEM DEFINITION

In this section we will give a brief introduction to what batch processing in workflow is, what data we will employ and how mining steps start and proceed.

Concept of Batch Processing in Workflow

Figure 1 is an example of processes with batch deserving activities (Liu 2007). It consists of six activities: “Application”, “Examination and Approval”, “Denial Informing”, “Renting Car”, “Informing”, and “Charging”. Table 1 shows the dataflow logs of activity “Renting Car” of the process in Figure 1. Let every car for renting with the same car type, e.g. Honda Civic, and can accommodate up to 4 passengers each time. In Table 1, tuple t_1 , t_3 , t_5 and t_8 are with the same destination, *Changsha*, and almost

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/identifying-batch-processing-features-workflows/36583

Related Content

Tech Industries Pattern Case Study in Modern Society Based on Political Economy

Neeta Baporikar (2022). *International Journal of Applied Management Sciences and Engineering* (pp. 1-13).

www.irma-international.org/article/tech-industries-pattern-case-study-in-modern-society-based-on-political-economy/302901

A Link-Node Nonlinear Complementarity Model for a Multiclass Simultaneous Transportation Dynamic User Equilibria

Mohamad K. Hasanand Xuegang Ban (2013). *International Journal of Operations Research and Information Systems* (pp. 27-48).

www.irma-international.org/article/link-node-nonlinear-complementarity-model/76671

E-marketing for SMEs

Neeta Baporikarand Rosalia Folela (2018). *International Journal of Applied Management Sciences and Engineering* (pp. 11-28).

www.irma-international.org/article/e-marketing--for-smes/196584

The Business Model of a Public Social Partnership: Contextual Determinants

Aldona Frczkiewicz-Wronka, Monika Tkaczand Saioa Arando (2017). *Public Sector Entrepreneurship and the Integration of Innovative Business Models* (pp. 108-138).

www.irma-international.org/chapter/the-business-model-of-a-public-social-partnership/174783

Opportunities and Challenges for Sustainable Business Strategic Planning in Small and Medium Enterprises (SMEs)

Mauricio Quintero-Angel, Claudia C. Peña-Montoya, Carlos Hernán Fajardo-Toroand Andres Aguilera-Castillo (2018). *Green Production Strategies for Sustainability* (pp. 153-167).

www.irma-international.org/chapter/opportunities-and-challenges-for-sustainable-business-strategic-planning-in-small-and-medium-enterprises-smes/192834