Journal of Universal Computer Science, vol. 20, no. 3 (2014), 406-430
submitted: 24/3/13, accepted: 4/2/14, appeared: 1/3/14 © J.UCS

A Framework for Cost-Aware Process Management: Cost Reporting and Cost Prediction

Moe Thandar Wynn
(Queensland University of Technology, Brisbane, Australia
m.wynn@qut.edu.au)

Wei Zhe Low
(Queensland University of Technology, Brisbane, Australia
w4.low@qut.edu.au)

Arthur H. M. ter Hofstede
(Queensland University of Technology, Australia
Eindhoven University of Technology, The Netherlands
a.terhofstede@qut.edu.au)

Wiebe Nauta
(Eindhoven University of Technology, The Netherlands
w.e.nauta@alumnus.tue.nl)

Abstract: Organisations are constantly seeking efficiency gains for their business processes in terms of time and cost. Management accounting enables detailed cost reporting of business operations for decision making purposes, although significant effort is required to gather accurate operational data. Process mining, on the other hand, may provide valuable insight into processes through analysis of events recorded in logs by IT systems, but its primary focus is not on cost implications. In this paper, a framework is proposed which aims to exploit the strengths of both fields in order to better support management decisions on cost control. This is achieved by automatically merging cost data with historical data from event logs for the purposes of monitoring, predicting, and reporting process-related costs. The on-demand generation of accurate, relevant and timely cost reports, in a style akin to reports in the area of management accounting, will also be illustrated. This is achieved through extending the open-source process mining framework ProM.

Key Words: Business process management, management accounting, process mining, cost reporting, cost prediction.
Category: H.4

1 Introduction

Organisations are constantly seeking efficiency improvements for their business processes in terms of time and cost. Enterprise cost reduction and business process improvement both emerge as important topics in Gartner's 2011 survey amongst more than 2000 CIO's, ranking at places 3 and 5 respectively [Gartner, 2011]. Business process management (BPM) provides an innovative approach to solving long-standing business challenges. This is achieved by promoting a process-centric view of an organisation through end-to-end management of business processes. An executable process model (or a workflow) contains descriptions of what tasks need to be performed, when and by whom, what information they require and what information they produce [Weske, 2007]. A BPM system (BPMS) supports planning, execution, (re)design and deployment of workflows [Weske, 2007]. Process mining provides techniques to discover, monitor and improve processes by extracting knowledge from event logs [van der Aalst, 2011]. A large number of process mining techniques exist that can discover what really happened during process execution using timing and resource information etc. but there are no process mining techniques that can directly support the cost-perspective as of yet. *An increased uptake of BPMSs in recent years has increased the potential for real, evidence-based analysis of the cost associated with the execution of business processes in a timely manner.*

1.1 Problem Definition

The motivation for this work comes from challenges encountered in the Management Accounting discipline, some of which could be alleviated by making use of detailed process knowledge available from the BPM discipline.

The management accounting discipline explores ways to capture the real cost of operations accurately and transforms this fine-grained cost information into a form suitable for making decisions as highlighted in the reports by the International Federation of Accountants [Professional Accountants in Business Committee, 2009a, Professional Accountants in Business Committee, 2009b]. Based on a detailed literature review and interviews with a number of management accounting experts, the following three key characteristics of the discipline are identified.

– Multiple management accounting techniques: Unlike the financial accounting discipline, whereby the accounting techniques used for external reporting are prescribed by legislation, an organisation has flexibility to use any suitable accounting technique for internal decision making. One of the popular management accounting techniques is activity-based-costing (ABC) which emphasises the per (activity) unit cost of all possible activities [Cooper and Kaplan, 1988]. However, the ABC technique requires a substantial effort to implement and to keep up-to-date [Kaplan and Anderson, 2003]. A new technique, time-driven ABC (TDABC), was introduced to address the shortcomings of the ABC technique. TDABC focuses on the per (time) unit cost of all possible activities [Kaplan and Anderson, 2003]. Another contender for next generation cost management techniques is Resource Consumption Accounting (RCA) [Tse and Gong, 2009, White, 2009]. The RCA technique focuses on the resource perspective, highlights the per unit cost of the resources that are involved in process execution and is interested in determining the over/under utilisation of resources [White, 2009]. In most cases, management accountants tailor one or more variants of these techniques to produce their management accounting reports based on an organisation's needs.

– Multiple cost reporting templates. As each management accounting technique looks at different aspects of cost at different levels of an organisation, there is no standard format for management accounting cost reports which is in contrast to the standard format available for external financial reports such as balance sheets and profit and loss reports.

– Cost Estimation and Variance Analysis. Many organisations follow a typical procedure for costing and budgeting. First, a budget with estimated costs for the next financial cycle is produced. Decisions regarding business operations are then made and the activities are carried out. After the financial cycle has ended, variance analyses are conducted to analyse the accuracy, reliability, and suitability of the previous budget [Kaplan and Atkinson, 1998, Professional Accountants in Business Committee, 2009a]. One of the main drawbacks of the traditional costing/budgeting approach is that ongoing operations in the current financial cycle cannot be easily influenced by changes in the cost data.

As the main purpose of management accounting is operational decision making, cost reports are of limited use to managers if the data contained in them cannot be relied upon and if the reports are produced *"too late, too aggregated or too distorted"* [Johnson and Kaplan, 1987]. Hence, the following are two main challenges concerned with data collection.

– Information gathering is time consuming. Management accountants who design, use and collect cost information typically work with others from many different parts of an organisation as they need to delve below the information recorded in financial ledgers and external financial reports to produce management accounting reports [Professional Accountants in Business Committee, 2009a]. Information gathering for decision making remains a costly and time-consuming exercise for organisations.

– The need for accurate cost and process information. Management accountants typically work with budget figures (e.g., volume of raw materials required and their associated costs, volume of production etc.,) which are then compared against the actual production figures/expenditures as part of the variance analyses carried out after the fact. There is a huge reliance on the accuracy of the budget figures as operational decisions are being made based on these figures, rather than the actual figures. It is also recognised that in many cases variance analysis highlights a significant discrepancy between budget and actual figures. Hence, access to accurate data in a timely manner is one of the key challenges facing the discipline.

The premise of this paper is that smart automation of the principles behind management accounting techniques may serve to lessen the cost of data gathering and provide data for decision making within a short time frame. Provided that an organisation uses an IT system or a BPMS which records historical information about what is happening during the execution of the organisation's processes in sufficient detail, the detailed operational information that a management accountant needs for cost control is available in some form.

1.2 Research Goals

Although the fields of management accounting and business process management look at different aspects of an organisation, there is a huge potential for managing cost in a structured manner based on an explicit link with business processes, since anything done though the application of BPM has an impact on processes and resources within an organisation which in turn influences cost and therefore management accounting decisions. In the field of BPM, organisations focus on systematically documenting, managing, automating and optimising their business processes. Process logs typically contain detailed records of activities carried out by resources within an organisation which is then analysed to identify suitable process improvement opportunities.

The *main goal* of our work is to devise an approach to combine relevant cost data with the historical information recorded during process executions to provide *accurate, relevant and timely information* for decision making.

The *innovative* nature of this approach is that by utilising detailed process knowledge from IT/BPM systems, the approach makes it possible to generate insightful customised process-based cost analyses in a timely manner. By providing the ability to generate detailed cost reports automatically at any time (e.g., every hour/day should an organisation wish to do so), the approach provides a rare opportunity to monitor cost and resource utilisation within process executions in a cost efficient manner. In addition, having timely access to detailed process-related cost information in a timely manner allows for accurate predictions of the cost behaviour of current and future processes with similar characteristics.

The *main contributions* of the paper are three-fold:

- an architecture for the cost analysis framework using event logs,
- the conceptual data models for incorporating the cost perspective in BPM systems, and

– a first implementation of cost reporting and cost prediction functionalities in the process mining framework ProM.

The remainder of the paper is organised as follows. First, we describe the related work to our approach. Next, we present the proposed cost mining framework. This is followed by a discussion on the implementation of the cost mining framework as a number of cost-related plug-ins within the process mining framework ProM. The evaluation strategy used for validation is then discussed. A brief summary and a discussion of possible future work is then presented.

This paper is an extended version of the conference paper presented at the Ninth Asia-Pacific Conference on Conceptual Modelling [Wynn et al., 2013a]. It brings the results of the cost analysis [Wynn et al., 2013a] one step further by supporting cost prediction capabilities within the framework. A detailed discussion related to the requirements of cost estimation/prediction, and a first implementation of the cost prediction plug-in has been presented (see Section 3.6 and Section 4).

2 Related Work

Our work combines aspects of management accounting, business process management, and process mining. Some of the most relevant contributions from these broad areas are reviewed below.

2.1 Management Accounting

Management accounting techniques provide managers with an ability to plan ahead of time and to make informed strategic and operational decisions based on cost [Johnson and Kaplan, 1987]. The ABC costing technique provides an accurate way of assigning the costs of indirect and support resources to activities, business processes, products, services and customers [Kaplan and Atkinson, 1998]. Fixed, variable, and overhead costs are taken into consideration and different cost types such as resource costs and activity costs identify appropriate cost drivers. The time-driven ABC(TDABC) technique identifies fixed, variable, and overhead costs as well, but instead of assigning cost to individual products or services like ABC, it was assigned to a time frame and cost per time unit is calculated [Kaplan and Anderson, 2003]. Hence, TDABC only requires estimates of capacity of committed resources, their cost rates, and activity duration. The resource consumption accounting (RCA) technique [Tse and Gong, 2009, White, 2009] is concerned with over/under utilisation of resources [White, 2009]. Based on the cost rate of a resource, idle costs are attributed to those resources that are underutilised. Furthermore, time or capacity-related information (for example, activity duration, utilised resources, and maximum allocated resource capacity), is essential for utilisation-based cost reports such as the idle resource cost report. The RCA technique can be viewed as evolution of the ABC technique in Enterprise Resource Planning (ERP) systems [Keys and Merwe, 2001]. However, information gathering for the purpose of cost reporting and control remains time and resource intensive.

2.2 Business Process Management

Business process management defines a methodology and tools to design, configure, execute and diagnose processes within organisations [van der Aalst et al., 2003, Weske, 2007]. Through the iterative application of BPM techniques, processes are improved in terms of quality, flexibility, time and/or cost [Reijers and Mansar, 2005]. Process mining facilitates continuous process improvement by extracting knowledge from event logs [IEEE Task Force on Process Mining, 2011]. An event log is a data store where a historical record of process execution is kept. Information such as task events (offer, start, complete), data attributes, utilised resources and task durations can be extracted from an event log [van der Aalst, 2011]. Depending on the level of details recorded in event logs, various analyses of different perspectives (e.g., control, organisational, case, and time) are possible. As a result, process mining techniques can provide valuable insight into control flow dependencies, data usage, resource utilisation and various performance related statistics. Process mining is enabled through the application of software. One of the leading tools for process mining is the ProM framework [van der Aalst et al., 2007]. ProM provides a generic open-source framework for process mining/analysis tools and additional plug-ins can be developed to extend its functionalities [Verbeek et al., 2010].

2.3 Cost Considerations for Process Improvement

There is also a body of work that brings the notions of cost and processes together. Cost has always been one of the key factors under consideration in terms of process improvements [Reijers and Mansar, 2005] and process simulation [Laguna and Marklund, 2005]. Irani et al. discuss eight different cost taxonomies for IT/IS-related cost [Irani et al., 2006]. A qualitative cost analysis methodology to better understand the cost of Process-Aware Information Systems development projects is described in [Mutschler et al., 2007]. There is also a body of work that focuses on measurement and aggregation of Quality of Services in workflow and web services including cost [Jaeger et al., 2004, Cardoso et al., 2004, Wynn et al., 2004, Mohabbati et al., 2011]. Since the introduction of ERP systems, studies have been conducted on the effects of ERP systems on traditional management accounting practices [Booth et al., 2000, Granlund, 2001, Granlund and Malmi, 2002, Grabski et al., 2009, Hyvönen, 2010]. An information model to link the ARIS accounting structure with ARIS process semantics using Event Driven Process Chains (EPC) is proposed in [vom Brocke et al., 2011]. In [Rosa-Velardo, 2012], the authors presented an approach to determine the correctness of workflow processes while cost information is taken into account. We also reported our proposal for a framework to support cost-informed process execution in [Wynn et al., 2013b]. This illustrates a growing interest in this interdisciplinary area of accounting and business process management and motivates us to undertake our research in the area of cost-aware process management.

3 Cost Mining Framework

In this section, we first present the research approach undertaken to realise the notion of cost mining (see Section 3.1 and then provide a motivating example for our work in Section 3.2. The relationship between these cost elements and different process elements such as activities, data objects, resources is described using conceptual data models in Section 3.3. The requirements for annotating event logs with cost elements are explained in Section 3.4, the cost reporting requirements are described in Section 3.5, and the requirements for cost prediction are described in Section 3.6.

3.1 Research Approach

To achieve our objective of making process mining 'cost-aware', we followed the research approach involving literature review, architecture design, implementation and validation. Based on the findings from a detailed literature review, we identified a number of key requirements from the management accounting perspective and the business process management perspective. These requirements are then used to develop an integrated approach for process-based cost mining as well as conceptual data models to capture the data requirements. The approach is then validated through the development of a prototype as well as interviews with experts from the management accounting and the BPM disciplines.

Figure 1 illustrates the proposed integrated approach for process-based cost mining. The approach is kept simple and high-level to enable both disciplines to see the potential of an integrated approach. To perform cost analysis, three separate steps are proposed to be carried out: creation of a cost model, annotation of event logs with cost information, and analysis of process-based costs.

The approach highlights the importance of information sharing between management accountants and BPM professionals. The top half of the diagram (see Figure 1) depicts the role played by management accountants in providing accounting knowledge to the framework. This may involve the supply of static cost

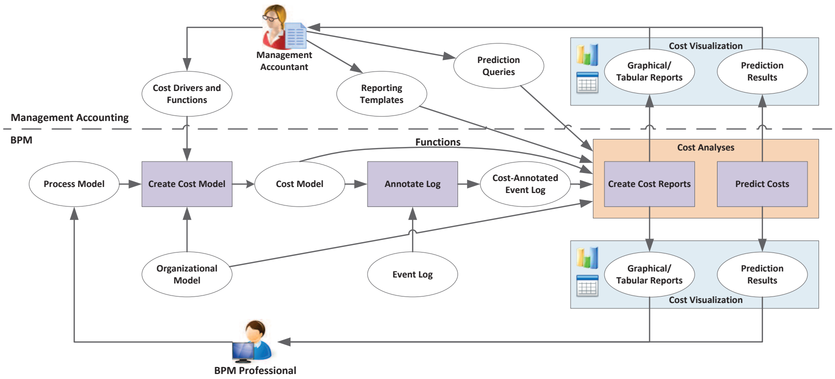

Figure 1: Overview of the integrated approach to log-based cost analysis

data collated from HRM, CRM or other systems, management accounting techniques used in the organisation, cost report templates and cost prediction queries. The bottom half of the diagram shows how existing knowledge from the BPM systems can be leveraged for cost analysis. This is achieved by making use of process knowledge from BPM systems in the form of executable process models, organisational models and event logs. Detailed process-related cost and resource utilisation cost can be automatically calculated and stored within event logs in the form of cost-annotated event logs. Relevant cost reports can be generated in a timely manner using report templates supplied by management accountants. In addition, graphical reports for BPM professionals and other stakeholders can also be generated from the data recorded in cost-annotated event logs. The framework also supports the ability to predict process-related costs for currently running cases based on cost-annotated event logs.

3.2 Motivating Example

A simple telephone repair process is being used throughout the paper to illustrate our approach (see Figure 2). The repair process starts by registering a faulty telephone sent in by a customer. After registration, the telephone is sent to the Test Department. There, the problem is analysed and its defect is categorised by an employee with the tester role. Once the problem is identified, the telephone is sent to the Solve Department and a letter is sent to the customer to inform him/her about the problem. The Solve Department has two teams. One of the teams fix simple defects (role: Solver Simple) and the other team repairs complex defects (role: Solver Complex). Once the repair has been completed, the telephone is sent back to the Test Department in order to test whether the repair was successful. If the defect has not been repaired properly, the telephone is sent back to the Solve Department. If the repaired telephone passes the test, the telephone is sent to the customer and the case is archived. In addition, it is the policy of the company that it only tries to fix a defect in a particular telephone for a maximum of five times. Furthermore, after the problem with the telephone has been analysed and in parallel with repairing and testing, the customer is informed about the result of the defect analysis.

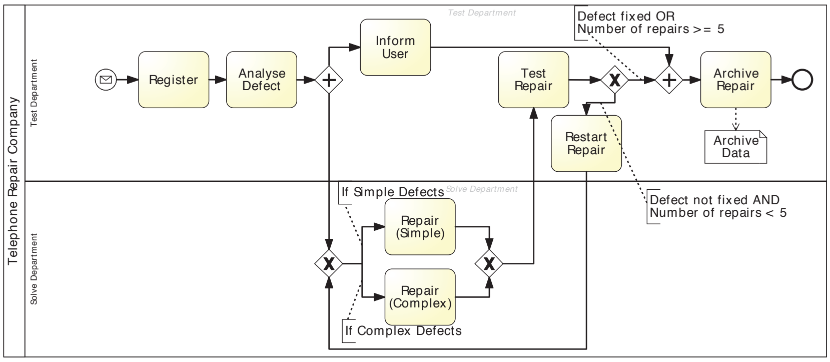

Figure 2: A BPMN model of the telephone repair process

There are altogether 12 employees in the telephone repair company of whom six work as testers in the Test Department and the other six work as problem solvers in the Solve Department. The resource assignments in the process are role-based and shown together with cost data in Figure 3. To be able to calculate cost incurred during activity executions of this process, we need to know relevant cost rates for this process. For instance, from the HR system we can find out the per hour rate for labour costs of an individual employee or a role (i.e., tester, solver). In Figure 3, we assign one or more cost rates for each activity in the telephone repair process. We make use of the notions behind Time-driven ABC costing, whereby, the cost of an activity is calculated based on the time it takes to carry out an activity by one or more resources. It is possible for an activity to have different cost rates for different cost types (e.g., the *Analyse defect* task has three different cost rates for three different cost types - fixed cost, labour and overheads). We also included simple/static cost rates (e.g., $65 per hour for an employee with the role SolverS) as well as more complex/dynamic cost rates where the cost rate is dependent on the characteristics of a case (e.g, the type of phone and/or the type of defect). Please note that we intentionally use many different types of cost rates for illustrative purpose, although in reality, an organisation may only use a selection of these cost types. Next, we discuss how

Activities/ Tasks	Role Assignments	Cost Type	Cost Rate (AUD)
Register	System	Fixed Cost ➢ Administration	2.00 per invocation
Analyse Defect	Tester	Fixed Cost ➢ Administration + PhoneType[1] [1]Only PhoneType "T3" costs AUD 20.00 extra.	10.00 per invocation + 20.00[1] per invocation
		Labour ➢ Tester	25.00 per hour
		Variable Overhead ➢ 1.00 * DefectType ➢ 0.01 * DefectType * minutes(s)	Varies per invocation Varies per minute
Inform User	System	Nil	0.00
Repair (Simple)	SolverS	Labour ➢ SolverS	65.00 per hour
		Materials Cost ➢ 5 * DefectType[2] [2]E.g. DefectType "1" costs AUD 5, hence AUD 5 * 5 = AUD 25.	Varies[2] per invocation
		Variable Overhead ➢ Repair (Simple)	10.00 per hour
Repair (Complex)	SolverC	Labour ➢ SolverC	75.00 per hour
		Materials Cost ➢ 5 * DefectType[2]	Varies[2] per invocation
		Variable Overhead ➢ Repair (Complex)	10.00 per hour
Test Repair	Tester	Labour ➢ Tester[3] [3]Only Tester4 costs AUD 35.00.	25.00/35.00[3] per hour
		Fixed Cost ➢ DetectFixed = False	10.00 per invocation
		Variable Overhead ➢ Test Repair	15.00 per hour
Restart Repair	Tester	Labour ➢ Tester	25.00 per hour
		Variable Overhead ➢ Restart Repair	10.00 per invocation
Archive Repair	System	Fixed cost ➢ Administration	2.00 per invocation
		Material Cost ➢ (RepairAttempt >= 5) AND (DetectFixed = False)	100.00 per invocation

Figure 3: Cost data associated with the telephone repair process

this cost data is mapped into a cost model of a process.

3.3 Cost Model

The creation of a cost model requires cost information from management accounting and process and resource information from a BPMS. A *cost model* is considered as the complete specification of cost-related data for a business process. The data requirements for cost models are collated from the literature and from conversations with domain experts. Three main costing techniques (ABC, TDABC and RCA) are analysed in terms of informational input and output of each technique. We then generalise these requirements to design the conceptual data model of a cost model artefact(Figure 4). This enables us to design a conceptual data model that can cater for *multiple management accounting techniques* - the first characteristic mentioned within the problem definition section.

We propose a cost model with three core elements: cost driver, cost function and mapping.

- A **cost driver** defines how cost is associated with one or more process elements (resource, activity, case data) and the cost rate. The rationale behind our conceptual design is to keep the model as general as possible so that we are able to capture both simple and complex cost rates (drivers). The following are some example cost drivers supported in our approach.
  - the cost rate of resource 'R1' is $50 AUD per hour (variable labour cost)
  - each invocation of activity 'A' costs $20 AUD (fixed processing cost)
  - each process instance has an overhead of $100 AUD (fixed overhead cost)
  - the cost rate of a resource with a 'manager' role performing activity 'A' is $30 AUD per hour (labour cost dependent on a task)
  - the fixed cost of activity 'A' with the data attribute 'D1' being 'standard' is $25 AUD while activity 'A' with the data attribute 'D1' being 'premium' is $50 AUD. (fixed cost dependent on a data attribute)
  - the variable cost rate of two resources 'R1' and 'R2' working together on an activity 'A' with data attribute 'D1' being 'premium' is $200 AUD per hour. (variable/labour cost of multiple resources working together dependent on task and data attribute)
- A **cost function** defines a formula for aggregating various cost elements. It is possible to specify cost functions that incorporate both fixed and variable cost components. Variable cost is defined in the cost driver using a cost function that takes into account the duration of a task or the duration of a resource working on that task and/or relevant case data. For instance, we can specify a cost function as follows: each invocation of activity 'A' incurs $20 AUD of cost type fixed cost plus $50 per hour for each resource working on the task (labour cost) plus $10 per kilogram of the raw materials used (materials cost).
- A **mapping** provides a way to relate terms used in management accounting to terms used in a BPMS. The terms used to identify resources, activities, and data may differ within a BPMS and an accounting system (e.g., an

employee might have two different identifiers (an employeeID in the HR system and a different username in IT systems). Mappings can be created for every workflow element (task, resource or case data) related to a cost driver.

Figure 4 shows the conceptual cost model, modelled using the Object-Role Modelling version 2 notation. For those readers who are not familiar with the ORM notation, we now briefly discuss the ORM 2 notation used in the paper. An ORM model captures relationships between entities [Halpin, 2005]. An entity type in the model is depicted as a round-corner rectangle (e.g., Cost Model, Cost Driver, Cost Function). An entity type has an identifier label type (e.g., cost driver ID). An entity type can be associated with one or more label types which are depicted as ovals with dash lines (e.g., CD Name, CD description). An entity type can have relationships with one or more entity types (e.g., "linked to" is a binary fact type between a cost driver and a cost technique and it is modelled using a fact type with two roles connected to the two entities involved). A bar above a fact type represents a uniqueness constraint that applies to that fact type (e.g., one to one, one to many, many to many for binary fact types). A black dot attached to the connector between an entity and a role indicates whether this role is mandatory for that entity (e.g., every cost driver must have a cost rate associated with it). A solid line with an arrow head describes a generalisation/specialisation relationship between two entities (e.g., a resource cost driver is a specialisation of a cost driver entity type).

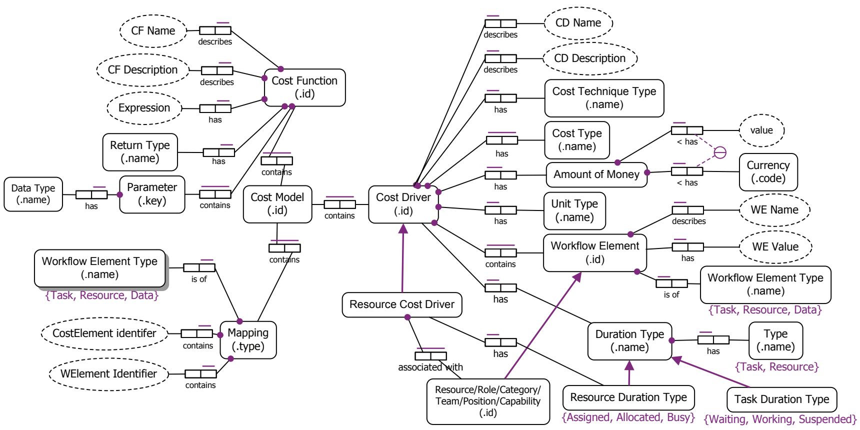

Figure 4: Data model for different cost elements

An XML schema has been defined based on this conceptual model and is used as input in the cost mining framework. At this stage, an XML document that represents a cost model for a process is manually created. We envision that there will be tool support for cost associations within a process modelling environment in the future. This will enable a BPM professional to specify a cost model in consultation with a management accountant.

3.4 Cost Annotation

A cost-annotated event log is one where events recorded in a log (e.g., the start event of a case, start and complete events of a task instance) are enriched with detailed cost information related to that event. The cost annotation itself is an association between a case/task instance, cost type, cost driver and a monetary value together with its currency (e.g., carrying out the approval task for caseID 001 incurs $35 AUD of variable cost type using costdriverID c12). A cost element can be associated with a case and/or with a task. Variable costs are calculated by multiplying a cost rate with relevant durations derived from timestamps of events in a log. We support six different duration types: waiting, working, and suspended for tasks and assigned, allocated and busy for resources to cater for advanced resource and task lifecycles supported by BPM systems. The ORM 2 representation for cost annotation is shown in Figure 5.

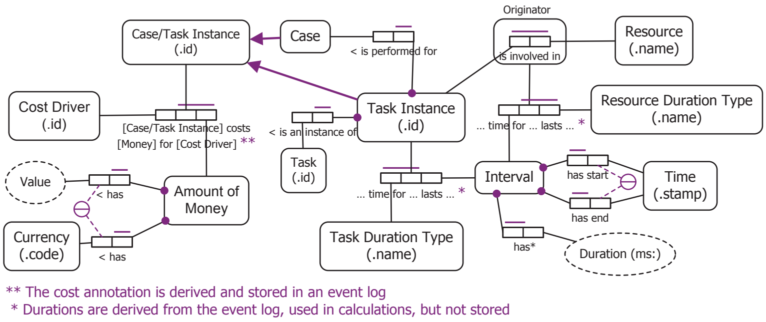

Figure 5: Data model for annotating cost in event logs

The concept of a cost annotated log is realised as an XML schema. XES is an XML-based format for capturing event information in a structured way [Günther, 2009]. The XES format uses the perspective of a case (or trace in XES) for grouping events, whereas the time perspective, control flow perspective and organisational perspective can be implemented as optional extensions of events in a case. Figure 6 illustrates the cost-annotated event log in the XES format where three cost types (fixed cost, variable overhead and labour) are associated with a completed event (lifecycle:transition - complete) of the Analyze Defect Task (concept:name - Analyze Defect).

```
<log xes.version="1.0" xes.features="nested-attributes"
  openxes.version="1.0RC7" xmlns="http://www.xes-standard.org/">
  <trace>
    <string key="concept:name" value="1"/>
    ..........
    <event>
      <string key="concept:name" value="Analyze Defect"/>
      <string key="org:resource" value="Tester4"/>
      <string key="lifecycle:transition" value="complete"/>
      <date key="time:timestamp" value="1970-01-02T17:10:00.000+10:00"/>
      <string key="phoneType" value="T3"/>
      <string key="defectType" value="6"/>
      <string key="cost:currency" value="AUD"/>
      <float key="cost:total" value="31.80">
        <string key="d2f4ee27-686c-48ef-a7ac-dcb053408245" value="">
          <float key="cost:amount" value="21.00" />
          <string key="cost:driver" value="d2f4ee27-686c-48ef-a7ac-dcb053408245" />
          <string key="cost:type" value="Fixed Cost" />
        </string>
        <string key="a7edbc32-e038-492a-a2e7-305ee1ed5d3e" value="">
          <float key="cost:amount" value="4.20" />
          <string key="cost:driver" value="a7edbc32-e038-492a-a2e7-305ee1ed5d3e" />
          <string key="cost:type" value="Labour" />
        </string>
        <string key="b0fac208-2c0f-4fe9-baec-d0180aa21df2" value="">
          <float key="cost:amount" value="6.60" />
          <string key="cost:driver" value="b0fac208-2c0f-4fe9-baec-d0180aa21df2" />
          <string key="cost:type" value="Variable Overhead" />
        </string>
      </float>
    </event>
    ..........
  </trace>
  ..........
</log>
```

Figure 6: A cost-annotated event log where different cost types were recorded against one of the completed events of the Analyze Defect task

3.5 Cost Reporting

Fine-grained cost information from cost annotated event logs can be used for analysis. Cost reports can provide different views (i.e., process, resource, organisational/departmental) of cost data in different formats. A number of key functional requirements have been identified in order to support on-demand cost reporting. They include:

– **Support for well-known management accounting techniques.** As different management accounting techniques focus on different cost aspects, the generated reports must adhere to the management accounting's reporting requirements. The framework was developed to support the principles of common and well-known management accounting techniques such as ABC, TDABC and RCA and their reporting requirements.

– **Support for customisable report templates.** One crucial difference between management accounting reports and external financial reports is that there is no standard format for a management accounting report. Instead, the report format must suit its purpose within an organisation. To cater for the customisation requirement, we propose the use of report templates that can be tailored for different management accounting reports. The framework uses the XSLT stylesheet technology[1] for this purpose and existing cost report templates can easily be converted to XSLT stylesheets for automatic report generation.

– **Ability to filter report output.** Another requirement is the need to support various filtering and display options. Some parties may desire to filter out unnecessary information, or to assign information into different groupings. Thus, we support filtering and grouping of the cost report information based on organisational hierarchy (e.g., according to departments or roles), on time (monthly, quarterly, or financial year) and different display options (e.g., tabular and graphical display). This is achieved using different XSLT stylesheets where XML data can be formatted, filtered and summarised according to the specification provided in an XSLT stylesheet template.

– **Ability to accept multiple file inputs with their designated file formats.** In order to generate comprehensive cost reports, extra information such as resource and cost rates are required in addition to cost annotated event logs. Therefore, the ability to input multiple files with different formats is supported in the framework.

– **Support for resource utilisation reports.** Resource utilisation reports used in the RCA technique highlight under-utilised resources and attribute idle cost to such resources based on operational data [White, 2009]. A resource's idle capacity is calculated by deducting the used capacity from the maximum/allocated capacity. We discovered during the requirements analysis of existing workflow organisational models that such capacity information

[1] http://www.w3.org/TR/xslt

is not being recorded. Hence, we propose a data model to describe the capacity of resources (see Figure 7). Each resource, human or non-human, has a maximum working capacity. The capacity can be consumed within a specific duration or at any desired time (i.e., indefinite duration). For example, James can work a maximum of 50 hours (max. capacity) per week (time unit). In some cases, a resource has an unlimited capacity (i.e., it is not a consumable resource). For example, a workdesk or an operation theatre can have an unlimited capacity.

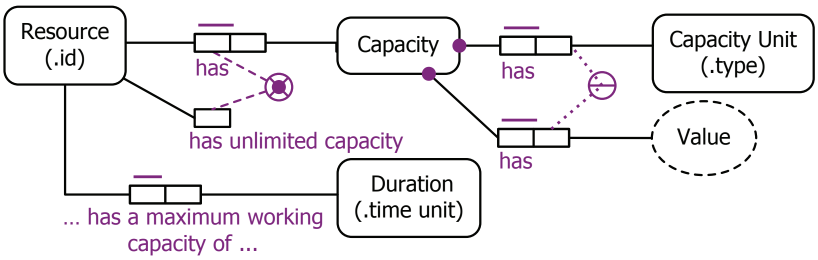

Figure 7: Data model for resource capacity

3.6 Cost Prediction

Cost reporting only analyses the data after the completion of an activity or a case. Based on this analysis, little can be done to influence costing decisions of business operations that are currently being carried out. Cost prediction, on the other hand, looks for cost patterns and characteristics from cost-annotated event logs so that one can predict the possible cost consumption of current, ongoing business processes based on historical data. This complements the traditional approach to cost estimation using budgets and variance analysis.

A number of key functional requirements have been identified in order to support cost prediction. They include:

– **Support for different cost query types.** There are two dimensions that can be considered when describing the query types. The first dimension is to do with the nature of the query itself. It can range from a 'simple' query (e.g., what is the average cost of a case?) to a 'compare' query (e.g. does case 1 cost more than case 2?) to a 'predict/recommend' query (e.g., what is the likelihood of this running case costing more than $150?, or who is the cheapest resource that can execute activity A?) [Nakatumba et al., 2012]. The second dimension depends on the object of interest: case, activity or resource (i.e., are we interested in estimating the cost of a case, an activity within a process or a specific resource?).

– **Support for generation of cost statistics.** The very first step in cost prediction is to carry out detailed cost analysis of completed cases to identify cost patterns and characteristics of the completed cases and to store relevant cost statistics (such as average, minimum, maximum, standard deviations). The framework supports this correlation of cost data with event log data using the notion of cost-annotated event logs where process-based costs are captured in its detailed from. To realise the cost prediction capabilities, an existing approach to predict the completion time of a current case has been extended [van der Aalst et al., 2011]. The existing approach involves the generation of a transition system from an event log and the storage of time-based statistics such as average, minimum, maximum, standard deviation for each state within the transition system. In a similar manner, we generate a transition system from a *cost-annotated event log* and store cost-based statistics.

– **Ability to incorporate different perspectives.** Cost consumption within an organisation is complicated. Cost may vary depending on the attributes of processes and activities. Attributes such as the different types of resources, the different states of the resources (working or idle), the details of the activities, and many more factors, can influence the cost of activities and cases. The incorporation of extra attributes allows a more sophisticated analysis by taking into account the role of resources and data in a process. The framework provides sophisticated cost analyses and predictions for the user while taking into account multiple aspects including control flow, resource and case attributes.

– **Ability to predict the cost of a ongoing case.** By incorporating the characteristics of an ongoing case (i.e., resources, case attributes) in the cost analysis, one can predict the cost outcomes for that case more accurately. For instance, we can answer questions such as what is the expected total cost of case #1 with case attributes - ‘VIP’ customer and PhoneType ‘T4’. The framework supports the generation of a cost-annotated transition system tailored to suit the case characteristics of an ongoing case and as a result, a more accurate cost prediction result based on the case characteristics of an ongoing case can be provided.

4 Realisation in the Process Mining Framework

The process mining framework ProM is a well-known open source platform which can be extended with additional plug-ins for process analysis. ProM version 6 receives the input of event logs in the Extensible Event Stream (XES) format in order to perform various analyses [Günther, 2009]. ProM also allows process mining operations to be performed through chaining (or stacking) of plug-ins. Figure 8 illustrates the architecture to support cost analysis in ProM, whereby, the output from the log annotation plug-in is used as one of the inputs for the cost reporting and prediction plug-ins. A first implementation of this framework has been developed in ProM 6. These plug-ins can be found in the "CostBased-Analysis" package in the ProM code repository.

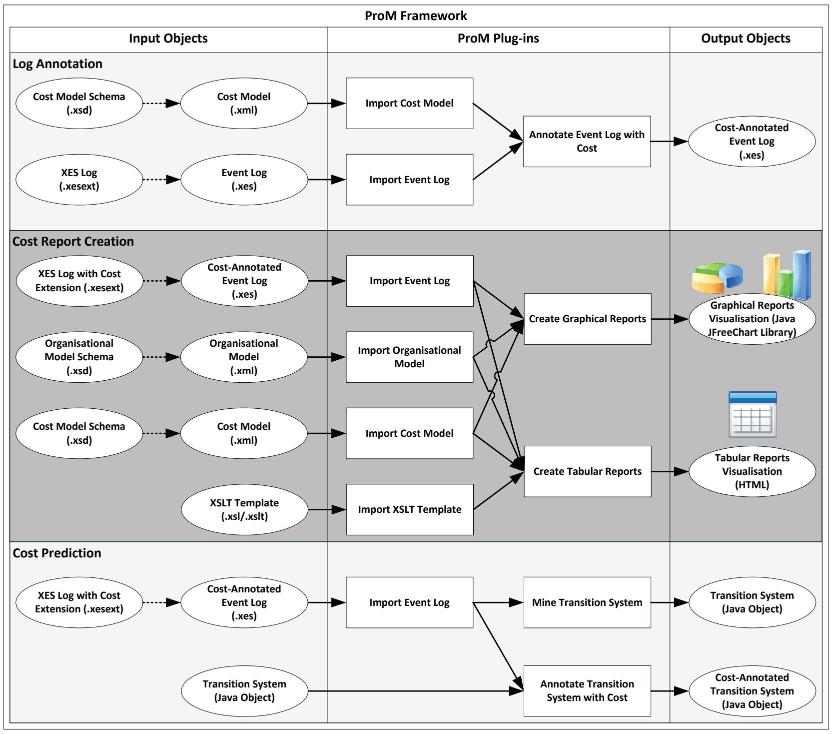

Figure 8: Cost annotation, reporting, and prediction in ProM

We used a dataset with 1000 completed cases of the telephone repair process generated using simulation. The log contained many records of timestamps for when tasks were started and completed, by whom and the values of data attributes where applicable. Four data variables that represent phone type, defect type, the frequency of repairs, and the repair outcome are recorded. A cost model, an organisational model and a number of XSLT templates were defined for this example.

4.1 Cost Annotation Plug-in

The cost annotation plug-in annotates an event log produced by a BPMS with cost information supplied from a cost model. Cost annotation of task instances is based on the task itself, resource(s) participating in the task instance execution and/or case data variables supplied during execution of the task instance. One or more applicable cost drivers are selected from a set of cost drivers given in a cost model based on the information contained in the log. The amount of the cost annotation is determined by the cost rate in the driver, and two variable components. First, the cost amount can be dependent on the case data values using a formula specified in the cost driver. Secondly, the duration of the task or duration of resource participation in the task is evaluated influencing the cost amount.

4.2 Cost Reporting Plug-ins

A number of cost reporting plug-ins that show the potential for using cost annotated logs for reporting purpose have been implemented. Visualisation of cost associated with process executions in the form of charts and graphs showing cost from different perspectives gives a quick overview of cost incurred during process execution. The Cost Visualization plug-in makes use of cost-annotated event logs to visualise resource usage in the form of 2D and 3D pie/bar charts, waterfall charts, and multiple bar charts. Customised cost reports in a tabular format can be generated using different cost reporting plug-ins with XSLT templates. Basic filtering and validation functions for XSLT templates are also available within these plug-ins[2].

4.3 Cost Analyser Plug-in

The transition system cost analyser plug-in, "Transition System Cost Analyzer", has been implemented to predict process-related costs of ongoing cases based on the cost information known about completed cases. The transition system approach proposed in [van der Aalst et al., 2011] to predict the completion time for cases (time prediction) has been extended to realise the concept of cost prediction. A cost-annotated event log is the starting point for any cost prediction. If an event log does not contain any cost information, then a cost model is required to annotate an event log with cost information. In such cases, a cost-annotated log can be generated using the cost annotation plug-in discussed in the previous section. The transition systems miner plug-in [van der Aalst et al., 2011] uses the cost-annotated log to generate a transition system. The newly developed

[2] Due to space limitations, a detailed discussion on the cost reporting plug-ins has been omitted, please see [Wynn et al., 2013a] for full details.

cost analyser plug-in then annotates the transition system with cost information, generating a cost-annotated transition system, which can then be analysed for cost prediction.

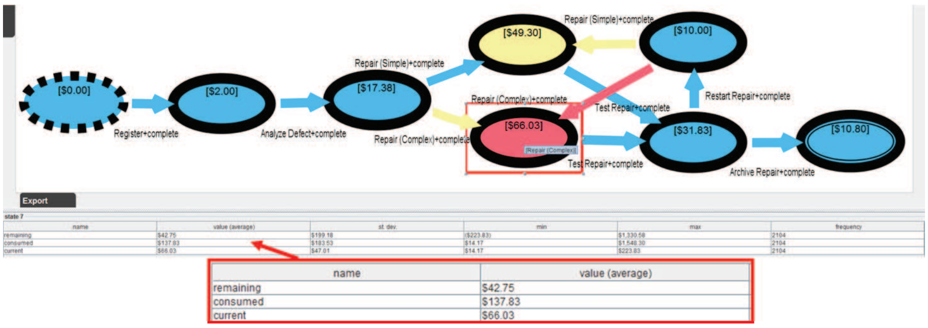

Figure 9: A screenshot of the cost-annotated transition system (activities only) generated from the telephone repair example

As discussed in Section 3.6, a cost-annotated transition system can be generated with different perspectives in mind. For instance, a log can be filtered to only consider activities while ignoring the resources who executed activities. Figure 9 shows the resulting cost-annotated transition system for the telephone repair example that only takes into account activities in a process. The ovals represent different states a case can be in and the arcs represent the transitions between two states. Each state has its own set of cost statistics, which includes minimum, maximum, standard deviation and frequency for three types of costs (the cost to date (consumed), the cost of the current activity (current), the cost to completion (remaining)). The colour-coding of the states and transitions provides a means of categorising cost behaviour. For instance, based on a certain user defined range, we can visualise cost that is below the defined range (blue), within the range (yellow), and above the range (red). Several cost-related queries can be answered by interpreting the generated cost-annotated transition system.

- What is the cost of a telephone repair case? This is a simple query type and the answer is available by using the cost statistics generated from the final state. Figure 9 shows that completed cases cost \$122.99 on average, with a lowest total cost of \$2, and a highest cost of \$1591.05.
- What is the predicted cost of an ongoing case (for instance, after the repair (complex) activity has been completed)? This can be answered using the cost statistics from the state of interest (see Figure 9). In this case, we can see

that the remaining (predicted) cost is $42.75 on average, the consumed cost to date is $137.83 on average, and the cost of the repair (complex) activity is $66.03 on average.

- Which is the cheapest activity that could be executed after the activity “Analyze Defect” is completed (disregarding other constraints)? By comparing the transitions that extend out from the “Analyze Defect” activity (the third state from the left in Figure 9), the “Repair (Simple)” activity is cheaper (the arc is coloured blue) with an average cost of $49.3, while the “Repair (Complex)” activity costs $66.03 on average (the arc is coloured yellow).
- Who can perform the activity “Analyze Defect” the cheapest? To answer this question, we generate another cost-annotated transition system which takes both activities and resources into account (see Figure 10). By comparing different “Analyze Defect” activities performed by six different resources (possibly facilitated by colour-coding of the states), we can observe that Tester 6 (represented by the right most oval) is (slightly) cheaper than other resources, costing only $16.41 on average.

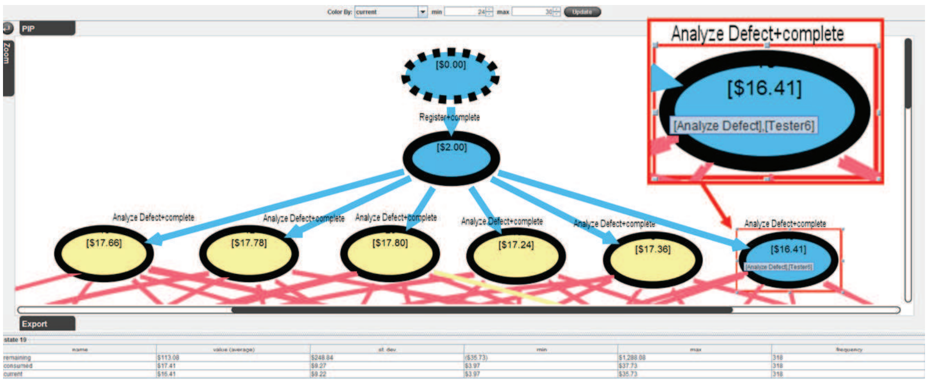

Figure 10: A screenshot of the cost-annotated transition system (activities and resources) generated from the telephone repair test case example

4.4 Evaluation Approach

In terms of evaluation, we are first and foremost interested in demonstrating that the general cost mining framework is valid and that the proposed data models are sufficient and necessary. To this end, we realised our approach as plug-ins of the well-known state-of-the-art Process Mining platform ProM and evaluated the implemented framework using an event log of the process that was publicly available together with a complex cost model that we generated for the process. This demonstrates that our approach works as expected. Furthermore, we used the time-driven ABC accounting approach for annotating cost information in event logs. A number of resource-based cost reporting plug-ins also illustrated how RCA-style cost reports can be generated. The cost prediction plug-in demonstrates how the processing cost of an activity or an ongoing case can be estimated using logs.

The proposed XES extension for cost annotations was distributed to members of the IEEE Taskforce on Process Mining to get their feedback and consensus on how cost information can be incorporated into event logs. Their comments were taken into account and changes were made to the XES extension. The cost extension was accepted as one of the standard extensions of XES. This further provides assurance that our approach of annotating event logs with cost information is general and that it can be used by others to associate cost information with process logs.

To evaluate the suitability our approach, we interviewed four people with knowledge of both the management accounting discipline and the BPM discipline to get their views on the proposed framework and the data models. All four interviewees provided positive feedback on the research and mentioned that they can see value in the integrated approach to process-based cost analysis. Regarding the framework, interviewees found the framework to be "comprehensive", "concise", "systematic", and "high-level". In terms of the data models, they mentioned that the concepts captured in the data models are "sufficient" and "more generic that what might be needed". For instance, interviewees found the cost drivers supported by the framework to be more extensive than typical utilisation rates used in the management accounting discipline (e.g., hourly rates for employees, materials/overhead costs depending on usage). The cost types used in the cost model of the example (for instance, fixed, variable, overhead) were found to cover the majority of cost types supported in cost reports. In addition, the generic nature of the cost model is also shown to support tailored accounting techniques and tailored cost reports required by the management accounting discipline. In terms of further improvements, one interviewee mentioned the need to demonstrate cost reconciliation between the cost information provided in the reports and the typical cost information available from the general ledger. Another interviewee mentioned that the cost reports could be more generalised, for example, comparison between cost and revenue, budget vs. actual costs, cost comparisons between different processes etc. Overall, this positive feedback provides us with an early indication of the suitability of our approach, although this is subjective in nature. To further validate the usefulness of our approach, we are looking into carrying out a case study within an organisation that has detailed event logs and makes use of sophisticated cost models for their management accounting decision making.

5 Conclusion and Future Work

There is a huge potential for managing cost in a structured manner based on an explicit link between cost and business processes. In this paper, we focus our attention on developing an overall framework for incorporating the cost perspective in the BPM Systems with a view to enable cost-aware process mining. To this end, we proposed a generic data format for a cost model and a data format for the cost extension of an event log (an XES extension for the cost perspective). We also illustrated how different management accounting cost reports can be generated using cost annotated event logs and how detailed cost estimations can be made using the process mining framework ProM. We also presented the findings from a first round of evaluations.

We acknowledge that potential limitations of the approach can come from the nature of the approach which is bottom-up whereby cost analysis is carried out based on the information recorded in an event log of a process. As a result, there may be a gap between high-level management accounting reports which not-only rely on process-based cost information but also on other types of information for decision making. Regardless, we believe many organisations can benefit from having real-time process-based cost information at their fingertips at the design, implementation, monitoring, and evaluation phases of a business process's lifecycle.

Other forms of cost-aware process mining that make use of our proposed cost-annotated logs can be conceived of. Simulation and visualisation of cost information enables managers to evaluate their process-related costs in a timely manner. Incorporating the cost dimension as part of a BPMS solution will enable cost-informed resource allocation and control flow decisions. We are currently exploring a number of these new research areas, using the work presented in this paper as their foundation.

Acknowledgments.
The authors would like to especially thank Wil van der Aalst, Boudewijn van Dongen, Eric Verbeek, Michael Adams, and Jochen de Weerdt for their input to this research. This work is partly supported by an ARC Discovery grant with number DP120101624.

References

[Booth et al., 2000] Booth, P., Matolcsy, Z., and Wieder, B. (2000). The impacts of enterprise resource planning systems on accounting practice–the Australian experience. *Australian Accounting Review*, 10(22):4–18.

[Cardoso et al., 2004] Cardoso, J., Sheth, A., Miller, J., Arnold, J., and Kochut, K. (2004). Quality of service for workflows and web service processes. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1(3):281–308.

[Cooper and Kaplan, 1988] Cooper, R. and Kaplan, R. (1988). Measure costs right: Make the right decisions. *Harvard Business Review*, September - October:96–103.

[Gartner, 2011] Gartner, C. . e. l. r. g. (2011). Reimagining IT: The 2011 CIO Agenda.

[Grabski et al., 2009] Grabski, S., Leech, S., and Sangster, A. (2009). *Management accounting in enterprise resource planning systems*. CIMA Publishing.

[Granlund, 2001] Granlund, M. (2001). Towards explaining stability in and around management accounting systems. *Management Accounting Research*, 12(2):141–166.

[Granlund and Malmi, 2002] Granlund, M. and Malmi, T. (2002). Moderate impact of ERPS on management accounting: a lag or permanent outcome? *Management accounting research*, 13(3):299–321.

[Günther, 2009] Günther, C. (2009). Extensible event stream xes standard definition.

[Halpin, 2005] Halpin, T. (2005). ORM 2. In Meersman, R., Tari, Z., and Herrero, P., editors, *On the Move to Meaningful Internet Systems 2005: OTM 2005 Workshops*, volume 3762 of *Lecture Notes in Computer Science*, pages 676–687. Springer.

[Hyvönen, 2010] Hyvönen, T. (2010). *Exploring Management Accounting Change in ERP Context*. PhD thesis, University of Tampere.

[IEEE Task Force on Process Mining, 2011] IEEE Task Force on Process Mining (2011). Process Mining Manifesto. In Daniel, F., Dustdar, S., and Barkaoui, K., editors, *BPM 2011 Workshops*, volume 99 of *Lecture Notes in Business Information Processing*, pages 169–194. Springer-Verlag, Berlin.

[Irani et al., 2006] Irani, Z., Ghoneim, A., and Love, P. (2006). Evaluating cost taxonomies for information systems management. *European Journal of Operational Research*, 173(3):1103–1122.

[Jaeger et al., 2004] Jaeger, M., Rojec-Goldmann, G., and Muhl, G. (2004). Qos aggregation for web service composition using workflow patterns. In *Enterprise Distributed Object Computing Conference, 2004. EDOC 2004. Proceedings. Eighth IEEE International*, pages 149–159. IEEE.

[Johnson and Kaplan, 1987] Johnson, H. and Kaplan, R. (1987). *Relevance lost: The rise and fall of management accounting*. Harvard Business School.

[Kaplan and Anderson, 2003] Kaplan, R. and Anderson, S. (2003). Time-driven activity-based costing.

[Kaplan and Atkinson, 1998] Kaplan, R. and Atkinson, A. (1998). *Advanced Management Accounting*, volume 3. Prentice Hall New Jersey.

[Keys and Merwe, 2001] Keys, D. and Merwe, A. (2001). The case for RCA: excess and idle capacity. *Journal of cost management*, 15(4):21–32.

[Laguna and Marklund, 2005] Laguna, M. and Marklund, J. (2005). *Business process modeling, simulation and design*. Dorling Kindersley (India) Pvt. Ltd.

[Mohabbati et al., 2011] Mohabbati, B., Gašević, D., Hatala, M., Asadi, M., Bagheri, E., and Bošković, M. (2011). A quality aggregation model for service-oriented software product lines based on variability and composition patterns. *Service-Oriented Computing*, pages 436–451.

[Mutschler et al., 2007] Mutschler, B., Reichert, M., and Rinderle, S. (2007). Analyzing the dynamic cost factors of process-aware information systems: a model-based approach. In *Advanced Information Systems Engineering*, pages 589–603. Springer.

[Nakatumba et al., 2012] Nakatumba, J., Westergaard, M., and van der Aalst, W. M. (2012). A meta-model for operational support. *BPM Center Report BPM-12-05, BPMcenter. org*.

[Professional Accountants in Business Committee, 2009a] Professional Accountants in Business Committee (2009a). Evaluating and improving costing in organizations.

[Professional Accountants in Business Committee, 2009b] Professional Accountants in Business Committee (2009b). Evaluating the costing journey: A costing levels continuum maturity model.

[Reijers and Mansar, 2005] Reijers, H. and Mansar, S. (2005). Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. *Omega*, 33(4):283–306.

[Rosa-Velardo, 2012] Rosa-Velardo, M. M.-S. F. (2012). Cost soundness for priced resource-constrained workflow nets. In Haddad, S. and Pomello, L., editors, *Application and Theory of Petri Nets*, volume 7347 of *Lecture Notes in Computer Science*, pages 108–127. Springer Berlin Heidelberg.

[Tse and Gong, 2009] Tse, M. and Gong, M. (2009). Recognition of idle resources in time-driven activity-based costing and resource consumption accounting models. *Journal of Applied Management Accounting Research*, 7(2):41–54.

[van der Aalst, 2011] van der Aalst, W. (2011). *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer.

[van der Aalst et al., 2007] van der Aalst, W., Reijers, H., Weijters, A., van Dongen, B., de Medeiros, A. A., Song, M., and Verbeek, H. (2007). Business process mining: An industrial application. *Information Systems*, 32(5):713–732.

[van der Aalst et al., 2011] van der Aalst, W., Schonenberg, M., and Song, M. (2011). Time prediction based on process mining. *Information Systems*, 36(2):450–475.

[van der Aalst et al., 2003] van der Aalst, W., ter Hofstede, A., and Weske, M. (2003). Business process management: A survey. *Business Process Management*, 2678:1019–1019.

[Verbeek et al., 2010] Verbeek, H., Buijs, J., van Dongen, B., and van der Aalst, W. (2010). ProM 6: The process mining toolkit. *Proc. of BPM Demonstration Track*, 615:34–39.

[vom Brocke et al., 2011] vom Brocke, J., Sonnenberg, C., and Baumoel, U. (2011). Linking Accounting and Process-Aware Information Systems - Towards a Generalized Information Model for Process-Oriented Accounting. *European Conference on Information Systems*, pages 1–13.

[Weske, 2007] Weske, M. (2007). *Business Process Management: Concepts, Languages, Architectures*. Springer-Verlag, New York, Inc., Secaucus, NJ, USA.

[White, 2009] White, L. (2009). Resource consumption accounting: Manager-focused management accounting. *Journal of Corporate Accounting & Finance*, 20(4):63–77.

[Wynn et al., 2004] Wynn, M., Edmond, D., and Milliner, S. (2004). The estimation of invocation costs for composite services. *In the Transactions of the SDPS: Journal of Integrated Design and Process Science*, 8(1):31–47.

[Wynn et al., 2013a] Wynn, M., Low, W. Z., and Nauta, W. (2013a). A framework for cost-aware process management: Generation of accurate and timely management accounting cost reports. In *Proceedings of the Ninth Asia-Pacific Conference on Conceptual Modelling (APCCM 2013).*

[Wynn et al., 2013b] Wynn, M., Reijers, H., Adams, M., Ouyang, C., ter Hofstede, A., van der Aalst, W., Rosemann, M., and Hoque, Z. (2013b). Cost-informed operational process support. In *Proceedings of the 32nd International Conference on Conceptual Modelling (ER 2013) (short paper).*