REVEALING THE REAL BUSINESS FLOWS FROM ENTERPRISE SYSTEMS TRANSACTIONS

Jon Espen Ingvaldsen, Jon Atle Gulla
Department of Computer and Information Science,
Norwegian University of Science and Technology, Trondheim, Norway

Ole Andreas Hegle, Atle Prange
Businesscape AS, Trondheim, Norway

Keywords: Information Systems Analysis, Change Projects, Business Modeling, Enterprise Resource Planning.

Abstract: Understanding the dynamic behavior of business flows is crucial when improving or reengineering organizations. In this paper we present an approach and a tool to business flow analysis that helps us reveal the real business flows and get an exact understanding of the current situation. By analyzing logs of large enterprise systems, the tool reconstructs models of how people work and detects important performance indicators. The tool is used as part of change projects and replaces much of the traditional manual work that is involved.

1 INTRODUCTION

For an enterprise to stay competitive or gain a competitive advantage it must continuously evaluate its performance and quickly adapt to changes at all levels of the organization. To meet this challenge enterprises are developing a growing interest in streamlining their business processes and implementing comprehensive *Enterprise Information Systems* (EIS). This process orientation allows them to concentrate on value-creating activities and be more responsive to changes in the environment. *Enterprise Resource Planning* (ERP) systems are now in wide-spread use among both large and midsize companies (van Everdingen et al., 2000).

Enterprise systems projects today typically involve a change element that ranges from narrow improvements of current processes to full reengineering of the whole business. Fundamental to all change projects, though, is the need to understand how the current business is run. The key issue is to identify weaknesses of the current business flows and work out better ways of performing the necessary tasks. Assessing the current business flows of the organization is also important for more restricted planning or coordination tasks. A well-known problem is the mismatch between the intended business processes and their actual execution. The process of identifying such gaps is called delta analysis (van der Aalst et al., 2003) and is continuously important in a process aware organization. In general, an accurate account of current business flows is needed for proper resource planning and IT investments.

In this paper we present an approach to business flow analysis that helps us understand the current situation and detect weaknesses and bottlenecks of the existing processes. A newly developed tool called *Enterprise Visualisation Suite* (EVS) automatically extracts business process models from the transactions of the enterprise system and provides a number of analyses of these transactional data. The tool is used as part of change projects, in which accurate and detailed descriptions of the AS-IS situation are needed for the subsequent design of new and optimized processes. The EVS tool replaces much of the traditional manual work of workshops and interviews with various stakeholders in the organization. It speeds up the early phases of change projects with more accurate accounts of existing workflows and can later be used to monitor the execution of the improved business processes.

2 THE CHALLENGES OF ASSESSING CURRENT BUSINESS FLOWS

Many recent works have pointed out the importance of analyzing the current business flows as part of enterprise systems projects or change projects (e.g. (Bancroft et al., 1997), (Offen, 2002)). If the existing

Espen Ingvaldsen J., Atle Gulla J., Andreas Hegle O. and Prange A. (2005).
REVEALING THE REAL BUSINESS FLOWS FROM ENTERPRISE SYSTEMS TRANSACTIONS.
In *Proceedings of the Seventh International Conference on Enterprise Information Systems*, pages 254-259
DOI: 10.5220/0002530902540259

processes and structures are not fully understood, decisions may be based on the wrong assumptions and lead to solutions that do not address the underlying weaknesses of the existing system. The issue is addressed by some recent key success factors for enterprise systems projects (Esteves-Sousa and Pastor-Collado, 2000) and also reflected in typical problem indicators for such projects (Gulla, 2004).

However, analyzing the current business processes is a tedious and challenging task. Workshops, interviews and on-the-job observations are all used to assess the way people work and the interdependencies between their activities. This involves a number of people with different backgrounds and possibly different perceptions of the organization's processes. The project needs to allocate time and resources to the assessment, and the quality of the outcome is hard to verify. In many cases there is also documentation available that can be consulted and presented to the stakeholders. Old project documentation or various kinds of user procedures describe how the various steps of the processes are to be carried out.

Except for the time and resources needed to run these labor-intensive analyses of existing processes, there are more fundamental problems related to the reliability and accuracy of the results. It is not obvious that people working in different parts of the organization have the same perception of their processes. People's opinions differ, and hidden agendas or power struggles may hamper the resolution of these differences. Their attitude towards the change project may also affect the way they present the current business processes. Even if they should agree on the overall processes, they may still find it difficult to drill down to the level of detail needed in the project. The project needs to know about potential bottlenecks, abilities to adapt to changing external events, process imbalances, and use of resources in general. Most of these issues are not documented properly, even though the people are often aware of these deficiencies.

Another interesting issue is the potential discrepancy between the way processes are designed and the way they are actually run. People tend to find ways around the job descriptions if they for some reason are not happy with the way they are supposed to work. Hence, user procedures cannot be assumed to provide the complete picture of the organization's processes.

Due to these challenges, many projects do not take AS-IS analyses of business processes serious enough. They risk developing the wrong system, or fail to understand the new processes' impact on the organization and the needs to train employees.

Interestingly, there should be ways of coming up with more accurate analyses. People leave traces when they do their jobs, and many of these traces are stored in enterprise information systems logs. As more and more of these processes are supported by the information systems, these systems should be able to tell us how the real business flow is. A fundamental problem, though, is how to interpret the logs in terms understandable to people. There is too much detailed information in the logs to present to people, and there is no obvious way of aggregating this information to a level and a format that can be understood by people involved in change projects.

3 THE ENTERPRISE VISUALIZATION SUITE

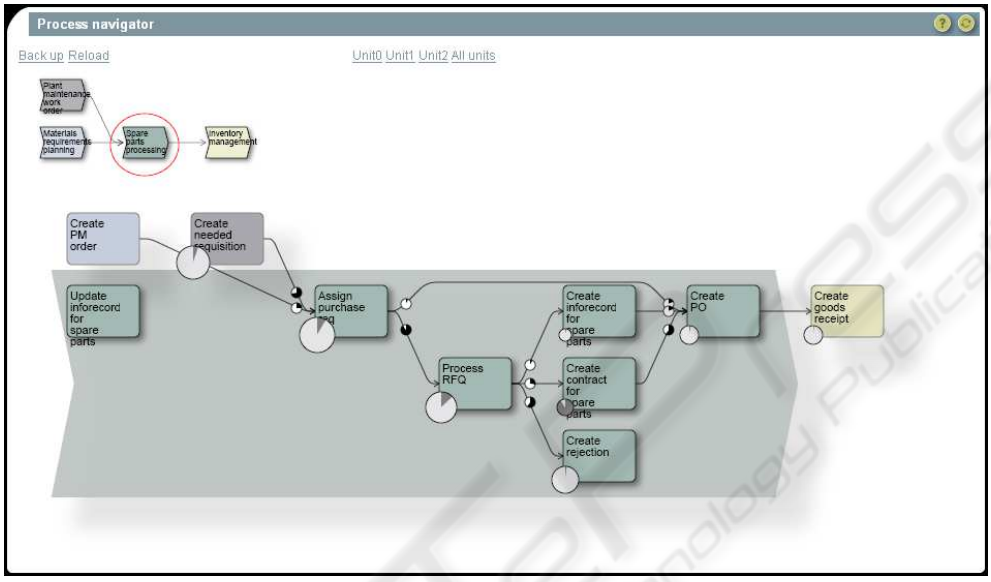
Enterprise Visualization Suite (EVS), from *Businesscape AS*, is a web based process mining package that enables enterprises to extract explicit process models and key performance indicators from event logs in their EIS. EVS aims at giving an exact picture of the AS-IS situation and a foundation for defining measurable goals for the TO-BE situation. EVS is implemented in Java, has an underlying *MySQL* database, and uses *Scalable Vector Graphics* (SVG) to show graphical models in the web browser.

Businesscape has currently developed an adapter for SAP R/3 systems, but EVS is built up of separated layers and the framework is independent of both specific EIS vendors and solutions. The *EIS adapters* extract and identify elements like applications, process hierarchies, document categories, users, user roles, and departments. Through analysis of event logs and collections of both document headers and positions, EVS is able to extract information about when these model elements where created or modified, and how they interrelate. An example of an event log is the CDHDR (*Change Document Header*) table in SAP R/3. As it keeps the history of almost all the document changes in the system, it is one of the very largest database tables in SAP R/3.

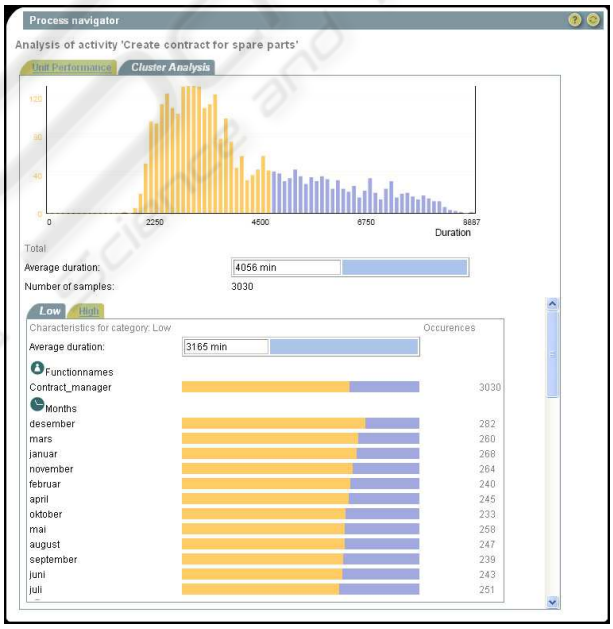
Table 1 shows an example of records and columns in the CDHDR table. As we can see, documents, users, executed applications, and their interrelations leave traces in this log. For a certain document, i.e., the HANDL_UNIT (Handling Unit) document with id 0001455075, we can see that it is altered twice by two different users and applications. By analyzing the change number we can see that these two alterations are subsequent. As a result, we know that the application LM46 produces a HANDL_UNIT that VL02N consumes. We are also able to calculate duration measures as the timestamp for each event is present. In SAP R/3, LM46 is the transaction code for *"Pick and Pack by Delivery"* and VL02N is the transaction code for *"Change Outbound Delivery"*. An EIS typically includes dozens of such event logs, including logs for application executions and document alterations. However, the structure of how similar information is

Table 1: Examples of data the the CDHDR (*Change Document Header*) table in SAP R/3

Doc.Type	Doc.Id	Change Nr.	Date	Time	User	Application
HANDL_UNIT	0001477919	0008992520	01.04.2004	09:00:49	BARG	VLMOVE
HANDL_UNIT	0001455075	0008992571	01.04.2004	09:00:52	BARG	LM46
HANDL_UNIT	0001455075	0008992572	01.04.2004	09:01:26	CI_TOR	VL02N
LIEFERUNG	0180118548	0008992579	01.04.2004	09:01:25	VIBGF	VL02N
MATERIAL	0000000128444	0008992564	01.04.2004	09:04:42	ANTER	MR21



(a) Constructed model of the process *"Spare parts processing"*



(b) Cause-analysis of execution time variations
Figure 1: Examples of screenshots from EVS

logged may vary within a single EIS. To cope with this inconsistency and still keep a general framework, EVS accesses its data foundation through a meta model description. The meta model says which event logs that should be analyzed and where information attributes within the event logs are located. These information attributes typically include current and previous activity, executing application, involved users, timestamps and textual descriptions. The meta models may be individually specified for each installation. However, for users with similar EIS, i.e. users with SAP R/3, the structure of how the log information is located is mostly identical. For such users, meta models can be reused and only minor modifications are necessary. The EIS Adapter is also responsible for datacleansing. There are several issues and challenge related to the data quality of event logs, including redundant information and missing values.

The *modelstore* database is the foundation for both process model construction and statistical performance analysis. Model visualization and analysis functionality is made available in a web based user interface. As a result, EVS can be used and accessed from multiple sites within the organization and by external consultants. Based on the output from EVS and other information sources, a change team can evaluate and modify business processes and EIS configurations. Executions of the redesigned business processes are then again input for EVS and future change projects.

4 VISUALIZING BUSINESS FLOWS

Process models tend to be both large and complex when many processes are involved. An example of such complexity appearances is the reference models of SAP R/3. These models are represented as Event Process Chains (EPC) (Curran and Keller, 1997), which lack proper abstraction mechanisms and duplicate event information. When the EVS model formalism was created, two aspects were in focus. First, the models had to be presented with such simplicity that any user can grasp the meaning instantly. Second, empirical performance indicators should be presented in the models, providing users with information about load distribution and execution times. EVS defines a hierarchy of processes and the resource flow between them. In the model represenation, a process consists of sets of sub processes or activities. Processes are shown as arrowed boxes while activities are shown as boxes with rounded corners. Activities are connected to the executable applications in the EIS, and their executions are traceable in the event logs. The dependencies between activities are found by identifying which resources the activities consume and produce. Dependencies in the higher process layers are defined as the aggregation of dependencies between their activities.

Figure 1(a) shows screenshots of how constructed business process models and empirical performance indicators are presented in EVS. The user can navigate through the set of layers in the process hierarchy. The present layer is shown in the main map of the model view. The super processes and their relationships are shown in a navigation map (upper left corner of the model view). Both the main model and the navigation map are interactively responsive. By selecting a process in the navigation map, its sub processes and their resource flow can be shown in the main map. If the user expands a super process in the main map, the content of the present main map is represented in the navigation map and main map becomes replaced by sub processes. These can then again be expanded, and so forth. Users do also have the option of navigating upwards in the process hierarchy. In figure 1(a), the process ”Spare parts processing” is selected, and its sub activities and their relationships are presented in the main map.

The size of the circles on each activity indicates the number of executions, while the size of darker pie within the circles indicivates the average execution time. In figure 1(a), we can see that the activity ”Assign purchase req.” is executed most frequently, while the activity ”Create contract for spare parts” has the longest average execution time. The pie charts on the connectors indicate relative distribution of multiple input or output trafic.

To enhance the expressiveness, EVS also exploits interactive opportunities of the web. Information that can be shown by clicking on a process or activity includes produced and consumed resources, and involved user roles and departments.

5 ANALYZING BUSINESS PROCESS PERFORMANCE

In order to get a detailed understanding of processes and their execution history, statistical reports that supplement the model map are necessary. In EVS, users can generate such reports for any selection of interesting processes or activities. Graphs that show the distribution of measures over time enable users to identify peaking periods, and tendencies over time. For process analysis, such measures include throughput or execution time. By aligning a set of processes in the same diagram the user can identify if correlations, propagations of peak values, etc. are present.

Two important issues for process change projects is to identify causes of variations in process execu-

tions. In order to perform such analyses, data from the relational database, *modelstore*, is exported to a flattened data set structure. In these data sets, process paths of subsequent activities are individually measured with respect to start time, end time, user, department, etc.The flattened data sets can further be extended by including new features that are found in other information sources or extracted from clustering analysis or combinations of the existing features. Having a defined data sets available, we are able to identify typical feature co-occurrences and find relationships between certain process execution behaviors.

Figure 1(b) shows an example of such cause analyses for the activity "Create contract for spare parts". Here we are identifying how features are co-occuring with *clusters* of execution time values. Two apparent clusters are identified, namely *high* and *low*. The *K-means* algorithm is applied to find appaerent clusters of the continuous variable. The list below the frequency diagram show how values of other features are co-occuring with the respective cluster assignments. As we can see, only one user role has executed the activity, and about 2/3 of the executions are in the *low* category. We can also see that *low* execution time is more apparent in the winter months than in the summer months.

6 DISCUSSION OF CASE

EVS has been tried out on data from a SAP R/3 implementation at a midsize Norwegian company. SAP R/3 is delivered with a reference model that shows best practice business processes. However, this reference model is not sufficient for documenting their real business flows, as their implementation included several add-ons, specific configurations, and interfaces to external systems. The reference model of SAP is neither able to show the load distribution of actual and past execution instances. When EVS was applied to give a picture of their real business flow and AS-IS situation, we faced several challenges.

One of the main challenges is how to deal with the many to many relationships between applications in the event log and the processes they belong to. In SAP R/3, transactions belongs to multiple business processes. The challenge arises as the event logs do not express which of the business processes that are involved in a specific execution of the transaction.

As long as EVS just collected data through the developed SAP adapter, it could not access event logs from business processes that were executed in external enterprise system. SAP and other EIS vendors have recently developed process oriented middleware solutions that ensures process oriented integration of separated enterprise systems. The solution from SAP is SAP Netweaver. Future work and effort will include examining if event logs from such middleware can enable EVS to show the real business flow across separated enterprise systems, and even across organizations.

In spite of these problems, EVS has shown itself to provide valuable insight into the business flow of the organization. It can deal with the complexities of a real case, and some of the analysis results were both interesting and surprising.

7 RELATED WORK

Our approach is related to work both in the academia and commercial area. We will group these activities into two categories, namely *process discovery* and *Business Activity Monitoring* (BAM). Both categories aim at extracting information from event logs, but the first category focuses on extraction of process models while the latter category focuses on extraction of key performance indicators.

There are several academic research activities that are working within the areas of process discovery. Cook and Wolf have shown the applicability of process discovery within different domains (Cook and Wolf, 1998)(Cook et al., 2004). In (Cook and Wolf, 1998) they investigate how neural nets, Markovian approaches and purely algorithmic approaches can be used to discover models of software engineering processes.Van der Aalst, et al., describes how the process mining tool *EMiT* is used to create *Petri nets* from event logs that are represented in a standardized XML format (van der Aalst et al., 2003). To support the practical application of EMiT, various adapters have been developed that allow for the translation of system-specific event logs to their XML format. Having a Petri net model as a foundation, it is possible to map the content to less expressive model representations like EPC. Their approach with use of adapters and a standardized format for process model constructions is similar to EVS' use of EIS adapters and the modelstore database. *Process Miner* is another tool that is created to visualize process models from event logs (Schimm, 2004)(van der Aalst et al., 2003). Its approach is to identify nested building block of sequences, parallels, alternatives, and loops in the event logs. Based on these blocks, Process Miner creates a process model representation. Similar to EVS, Process Miner has also a hierarchical view of its processes.

Activities within the BAM area are mostly dominated by commercial software vendors, but there is also some relevant research in the academic world. In (Grigori et al., 2004) they have developed a set

of tools that support analysis, prediction, monitoring, control and optimization of business processes. *ARIS Process Performance Manager* (PPM) is a software package by *IDS Sheer* that monitors business process performance indicators like flow times, bottlenecks, and utilization. ARIS PPM is a part of the ARIS tools, which also includes a design platform for business process modeling (Sheer, 1998). Other commercial activities within the area of BAM include Cognos (Notice Cast), FileNet (Process Analyzer), Hyperion (Business Process Management Suite), Tibco (Business Factor). Tools like SAP Reverse Business Engineer (RBE) and Intellicorp LiveModel can be used to monitor how frequent parts of the SAP system are being used.

8 CONCLUSIONS

The successful execution of change projects requires that the current processes are well understood and documented. With the manual approach, the task is time-consuming and the resulting models are often unreliable or inaccurate. EVS promises to speed up the assessment of the AS-IS situation and produce models and analyses that are at a more suitable level of detail. It is a tool that supports both incremental improvements of business processes as well as the more radical approach favored in business process reengineering projects.

The first version of EVS is now available, and we are analyzing the data from the SAP R/3 installations of two Norwegian companies. Both of these two cases show promising results, revealing some interesting process data that were not previously known. We are also considering taking part in reengineering projects that do not involve R/3 installations, though this requires that we implement adapters for other enterprise systems.

A problematic aspect of EVS is the construction of user-friendly hierarchical process models from these low-level enterprise systems logs. We need to make sure that the models do not get too detailed, but still include all the relevant information for assessing the quality of the processes. Building up semantically meaningful hierarchies is a non-trivial task.

Another issue here is how to deal with log transactions that can be part of several business processes. Future work concentrates on graphical presentations of important aspects of the business processes and more powerful data mining facilities. We would for example like to provide process cost estimates and predictions of *Service Level Agreement* (SLA) outcomes. The advantage of our analyses will be that we can relate the numbers directly to the organization's business flows.

REFERENCES

Bancroft, N. H., Seip, H., and Sprengel, A. (1997). *Implementing Sap R/3 : How to Introduce a Large System into a Large Organization, 2nd Edition*. Prentice Hall.

Cook, J. E., Du, Z., Liu, C., and Wolf, A. L. (2004). Discovering models of behavior for concurrent workflows. *Computers in Industry*, 53(3):297–319.

Cook, J. E. and Wolf, A. L. (1998). Discovering models of software processes from event-based data. *ACM Transactions on Software Engineering and Methodology*, 7(3):215–49.

Curran, T. and Keller, G. (1997). *SAP R/3 Business Blueprint: Understanding the Business Process Reference Model*. Prentice Hall PTR.

Esteves-Sousa, J. and Pastor-Collado, P. (2000). Towards-the unification of critical success factors for erp implementations. *10th Annual Business Information Technology (BIT) 2000 Conference, Manchester*.

Grigori, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M., and Shan, M.-C. (2004). Business process intelligence. *Computers in Industry*, 53(3):321–343.

Gulla, J. (2004). Understanding requirements in enterprise systems projects. *Proceedings of the 12th IEEE International Requirements Engineerings Conference*, pages 163–172.

Offen, R. (2002). Domain understanding is the key to successful system development. *Requirements Engineering*, 7(3):172–175.

Schimm, G. (2004). Mining exact models of concurrent workflows. *Computers in Industry*, 53(3):265–281.

Sheer, A. (1998). *ARIS - Business Process Modeling*. Springer-Verlag.

van der Aalst, W. M. P., van Dongen, B. F., Herbst, J., Maruster, L., Schimm, G., and Weijters, A. J. M. M. (2003). Workflow mining: A survey of issues and approaches. *Data Knowl. Eng.*, 47(2):237–267.

van Everdingen, Y., van Hillegersberg, J., and Waarts, E. (2000). Enterprise resource planning: ERP adoption by European midsize companies. *Communications of the ACM*, 43(4):27–31.