

International Journal of Computer Science and Mobile Computing



A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

IJCSMC, Vol. 3, Issue. 8, August 2014, pg.135 – 145

RESEARCH ARTICLE

A Study on Matching of the Process Data and Operational Data for a Deep Business Analysis. Ontology Design for Possible Matching and Recommendations to Improve it

Humaira Majeed¹, Ramzan Talib², Madeeha Aslam³, Hira Majeed⁴

^{1, 2, 3, 4} College of Computer Science and Information Studies, Government College University, Faisalabad, Pakistan

majeed.humaira@hotmail.com

Abstract— Companies have long been taking the assistance of only the operational data of a business process, when analyst make their decisions just on the basis of operational data but the pitfall of this method was that context of the business process was missing therefore decisions made in the light of such analysis were not up-to-date and sometimes inaccurate because decision made just on the basis of operational data with the exclusion contextual information ultimately effects the decision making process. Therefore, amalgamation of these two data sources is the need of time, consolidation of the operational data (needed to perform business procedures, processes and collected during the operational implementation of an operational process) and business process data (consists of a set of activities that are executed in some enterprise or administration according to some rules in order to achieve certain goals) is inevitable. With this convergence evaluation and the decision making result will be more valuable and will result in the form of deep business analysis. How this convergence can be done seamlessly automatically considerable by the any system that can perform deep business analysis? Its answer ultimately needs for the design of ontology for the possible matching of the process data and operational data for a deep business analysis so that the recommendations impeccably automatically available for any analysis system, data mining or any business intelligence tool. Deliverable of this research will be an ontology, however to develop ontology we have selected different database management system tools and process management system tools to collect event log of operational data and process log respectively.

Keywords: Operational data; Process data; Deep business analysis; ontology; Amalgamation of Process data and Operational data

I. INTRODUCTION

Information must include context

Data is in rough form whereas information is in refined form, when delivered through different technologies helps executives, managers, process owners, and analysts make better business decisions. Organizations comprise of the data, procedures, processes and the different operations performed on the processes and data. Process is a sequence of actions and certain steps that are interdependent but linked within order to perform some operations or any instance that is being executed to achieve a particular end is called a process. A process may use various resources to convert the input into the output and further these outputs can be the input of the next action or operation. Business process consists of a set of activities that are executed in some business according to some rules in order to achieve certain goals.

Data Mining

Data mining techniques, algorithms and methods have long stick up alone, they are consuming data that is stemming from various applications, analyse data, and provide rules and models supported by that data. However, an important part was and is still missing the “context” of the business process in which data was created.[1] However context is everything in any business decision-making process, and when data is taken out of the context, the results are, at best, but limited, if not downright deceptive.[2]

Process Mining

Similarly Process mining has emerged as a new research field since last decade for the analysis of process using event data. Classical data mining techniques such as classification, clustering, regression, association rule learning, and sequence mining do not focus on business process models and are often only used for the analysis of specific steps in the overall process [4] whereas process mining focuses on end-to-end processes and it is possible just because of the high availability of event data. Previously process models were usually made by hands without using event data. However, tasks and activities executed by people, machines, and software dispense trails in so-called “event logs”. Process mining techniques use such logs to discover, analyse, and improve business processes activities that have been executed. Process mining is based on facts rather than opinions also based on observed behaviour and intelligent techniques that are used to extract knowledge recorded in event logs.

Convergence of Process data and operational data

Therefore, by combining data mining and business process management technology, organizations can leverage from domain relevant information in mining methods for producing more concise knowledge. This information can be of utmost important, not only for decision makers in order to make better decisions and to boost up business benefits but also for stimulating operational quality of automated systems like Workflow Management System (WfMS), a WfMS is a software system that automates processes and manages the flow of work between workflow participants (i.e. resources) according to the formal description of business processes, called workflow. A workflow (or workflow model) defines process steps, their order, under which conditions and when they will

be carried out, by whom (resources i.e. human agents) within an organization, with which tools (i.e. applications), and define the flow of data within these process steps [8].

Basically the difference between the term “process” and a “workflow” in [9] is, A process is an abstract description and therefore cannot be executed whereas a workflow it is “somehow” derived from a process and in contrast it is executable. However in this paper, since we mostly focus on the execution and evaluation aspects for performance management and deep business analysis, we want to take them synonymous. However, when a workflow is executed in a workflow application system, the main task of WfMS is to coordinate all aspects of a process. A WfMS has to take care that all process steps are executed in the right order, consuming and producing the right data, applying the right applications, and the most important to select the right agents.

In order to execute a certain process, different WfMSs offer various ways or manners for the selection of appropriate means (resources, persons) and tasks from the organizational database, based on the certain criteria defined by assignment policies. For instance, WfMS can use workflow execution history and data mining techniques to first learn which actions and/or what combination of resources are most efficient and then use this knowledge in allocating them for upcoming processes which is unavoidable not only for performance improvement but also for effective process management [1].

In most conditions business process determines the context for information and yet the various business intelligence technologies used to manipulate and report on data do not comprise a native mechanism to define or act on processes. Instead, processes are custom-coded through the proprietary scripting languages. Similarly, many business process management suite (BPMS) products do not contain a native mechanism to analyze business data, such as loan amounts, alongside process-related data, such as process cycle time [2]. In order to make the better and timely decisions most of the business companies focus on the business processes. Business processes are a chain of connected business activities, which usually made a contribution to the value of a company decision making process [3]. A profound analysis of all relevant information in the company is the need of time and for the deep optimization of the process all the factors involved in the business have to be considered. These factors can be the process data, operational data, different resources and people who execute different processes. In order to improve the businesses, unnecessary activities should be removed and replaced by more efficient ones. Classical data mining techniques do not support in managing the performance of processes specifically for the efficient allocation of resources to the business processes. However, Workflow technologies can be helpful to play vital roles for an automatic allocation of the resources to the different business process. Normally Work flow Management Systems are used to sustain the modelling, execution and overall management of business processes. One of the crucial roles of a WfMS is assigning tasks to different users (resources, employees, or human agents as they are the driving forces) for the execution of the processes whereas issue is how to allocate resources to their processes of workflow execution on the basis of history. WfMS describes policies for assigning tasks to eligible persons to their processes that is mostly expressed in terms of “roles” and these roles are defined during design time of the process for the successful enactment.

Aim of business process analysis is to discover the analysis information about processes or redesigning and improving the business processes. The starting point of any business process analysis technique is the event log [5]. Log files contain the abundant information about different perspectives (different perspective of any business may be the security perspective, control perspective information perspective and history perspective) and a combined approach toward these perspectives will provide the efficient evaluation and analysis. Usually,

these event logs provide the information about the start to complete steps of a process along with related context data (e.g. actors and resources) that execute the corresponding tasks. It is possible to get the process log to collect the data in the order in which they take place, only if those tasks are recorded. However, the denotation and processing of data perspective in business process become more and more important, so the event log can add plentiful information about data perspective to support the analysis of a process.

Contemporary information systems like, WFM, CRM, ERP, SCM, and B2B systems, record business events in so-called event logs. Business process mining takes these logs to find out process, control, organizational, data and social structures [6]. These days, many organizational information systems store relevant events in some well-defined form. For instance, workflow management systems usually record the start and completion of activities [7]. ERP systems like SAP log all transactions, e.g. changing documents, users filling out forms etc. Business-to-business (B2B) systems log the exchange of messages with other parties. Call center packages but also general-purpose CRM systems log interactions with customers. These examples show that many systems have some kind of event log often referred to as "history", "transaction log", "audit trail, etc. The event log typically contains information about events referring to an activity and a case. The case (also named process instance) is the "thing" which is being handled, e.g. a customer order, a job application, insurance claim etc.

Process mining or Business process mining, focuses at the automatic construction of models explaining the behavior observed in the event log. For instance, based on event log, one can construct a process model expressed on Petri Net (process management tool) and the key objective of process mining is to mine information about processes from audit logs. Suppose that it is possible to enlist events such that (i) each event refers to an activity (i.e. structured step in the process), (ii) each event refers to a case (i.e., a process instance), (iii) each event can have a performer (the person instigating the activity), and (iv) events have a timestamp and are totally ordered.

II. LITERATURE REVIEW

Identified Issues

The missing ingredient: Information must include context

Data is in rough form whereas information is in refined form, when delivered through different technologies helps executives, managers, process owners, and analysts make better business decisions. Organizations comprise of the data, procedures, processes and the different operations performed on the processes and data. Process is a sequence of actions and certain steps that are interdependent but linked within order to perform some operations or any instance that is being executed to achieve a particular end is called a process. A process may use various resources to convert the input into the output and further these outputs can be the input of the next action or operation. Business process consists of a set of activities that are executed in some business according to some rules in order to achieve certain goals.

Data Mining

Data mining techniques, algorithms and methods have long stick up alone, they are consuming data that is stemming from various applications, analyze data, and provide rules and models supported by that data. However, an important part was and is still missing the “context” of the business process in which data was created.[1] However context is everything in any business decision-making process, and when data is taken out of the context, the results are, at best, but limited, if not downright deceptive.[2]

Process Mining

Process mining has emerged as a new research field since last decade for the analysis of process using event data. Classical data mining techniques such as classification, clustering, regression, association rule learning, and sequence/episode mining do not focus on business process models and are often only used for the analysis of specific steps in the overall process [4], whereas process mining focuses on end-to-end processes and it is possible just because of the high availability of event data. Previously process models were usually made by hands without using event data. However, tasks and activities executed by people, machines, and software dispense trails in so-called “event logs”. Process mining techniques use such logs to discover, analyze, and improve business processes activities that have been executed. Process mining is based on facts rather than opinions based on observed behavior, intelligent techniques are used to extract knowledge recorded in event logs.

Matching of Process data and Operational data.

Therefore, by combining data mining and business process management technology, organizations can leverage from domain relevant information in mining methods for producing more concise knowledge. This information can be of utmost important, not only for decision makers in order to make better decisions and to boost up business benefits but also for stimulating operational quality of automated systems like Workflow Management System (WfMS), a WfMS is a software system that automates processes and manages the flow of work between workflow participants (i.e. resources) according to the formal description of business processes, called workflow. A workflow (or workflow model) defines process steps, their order, under which conditions and when they will be carried out, by whom (resources i.e. human agents) within an organization, with which tools (i.e. applications), and define the flow of data within these process steps [8].

Basically the difference between the term “process” and a “workflow” in [9] is, A process is an abstract description and therefore cannot be executed whereas a workflow it is “somehow” derived from a process and in contrast it is executable. However in this thesis, since we mostly focus on the execution and evaluation aspects for performance management, we want to take them synonymous. However, when a workflow is executed in a workflow application system, the main task of WfMS is to coordinate all aspects of a process. A WfMS has to take care that all process steps are executed in the right order, consuming and producing the right data, applying the right applications, and the most important to select the right agents.

In order to execute a certain process, different WfMSs offer various ways or manners for the selection of appropriate agents from the organizational database, based on the criteria defined by assignment policies. For instance, WfMS can use workflow execution history and data mining techniques to first learn which actions and/or what combination of resources are most efficient and then use this knowledge in allocating them for upcoming processes which is unavoidable not only for performance improvement but also for effective process management [1].

Deep Business Analysis

In most conditions business process determines the context for information and yet the various business intelligence technologies used to manipulate and report on data do not comprise a native mechanism to define or act on processes. Instead, processes are custom-coded through the proprietary scripting languages. Similarly, many business process management suite (BPMS) products do not contain a native mechanism to analyze business data, such as loan amounts, alongside process-related data, such as process cycle time [2]. In order to make the better and timely decisions most of the business companies focus on the business processes. Business processes are a chain of connected business activities, which usually made a contribution to the value of a company decision making process [3]. A profound analysis of all relevant information in the company is the need of time and for the deep optimization of the process all the factors involved in the business have to be considered. These factors can be the process data, operational data, different resources and people who execute different processes. In order to improve the businesses, unnecessary activities should be removed and replaced by more efficient ones.

Classical data mining techniques do not support in managing the performance of processes specifically for the efficient allocation of resources to the business processes. However, Workflow technologies can be helpful to play vital roles for an automatic allocation of the

resources to the different business process. Normally Work flow Management Systems are used to sustain the modeling, execution and overall management of business processes. One of the crucial roles of a WfMS is assigning tasks to different users (resources, employees, or human agents as they are the driving forces) for the execution of the processes whereas issue is how to allocate resources to their processes of workflow execution on the basis of history. WfMS describes policies for assigning tasks to eligible persons to their processes that is mostly expressed in terms of "roles" and these roles are defined during design time of the process for the successful enactment.

Aim of business process analysis is to discover the analysis information about processes or redesigning and improving the business processes. The starting point of any business process analysis technique is the event log [5]. Log files contain the abundant information about different perspectives (different perspective of any business may be the security perspective, control perspective information perspective and history perspective) and a combined approach toward these perspectives will provide the efficient evaluation and analysis. Usually, these event logs provide the information about the start to complete steps of a process along with related context data (e.g. actors and resources) that execute the corresponding tasks. It is possible to get the process log to collect the data in the order in which they take place, only if those tasks are recorded. However, the denotation and processing of data perspective in business process become more and more important, so the event log can add plentiful information about data perspective to support the analysis of data perspective in process.

Event Logs

Contemporary information systems like, WFM, CRM, ERP, SCM, and B2B systems, record business events in so-called event logs. Business process mining takes these logs to find out process, control, organizational, data and social structures [6]. These days, many organizational information systems store relevant events in some well- defined form. For instance, workflow management systems usually record the start and completion of activities [7]. ERP systems like SAP log all transactions, e.g. changing documents, users filling out forms etc. Business-to-business (B2B) systems log the exchange of messages with other parties. Call center packages but also general-purpose CRM systems log interactions with customers. These examples show that many systems have some kind of event log often referred to as "history", "transaction log", "audit trail, etc. The event log typically contains information about events referring to an activity and a case. The case (also named process

instance) is the “thing” which is being handled, e.g. a customer order, a job application, insurance claim etc.

Process mining or Business process mining, focuses at the automatic construction of models explaining the behavior observed in the event log. For instance, based on event log, one can construct a process model expressed in terms of a Petri net (process management tool) and the key objective of process mining is to mine information about processes from audit logs. Suppose that it is possible to enlist events such that (i) each event refers to an activity (i.e. structured step in the process), (ii) each event refers to a case (i.e., a process instance), (iii) each event can have a performer (the person instigating the activity), and (iv) events have a timestamp and are totally ordered.

Indeed, after a thorough analysis of business process management technology we observed that process models are suitable source for domain knowledge especially when business intelligence is required within the application of a workflow. We realized that even though both technologies (process management and data mining) seem dissimilar, but still both support directly related promises:

- Data mining technology delivers knowledge that leads to “actions” that are incorporated and utilized within business environments for example by means of processes either manually by decision makers or automatically through an application.
- Process management technology defines the “context awareness” of business processes within the process models. Since, it defines syntax and semantics for different perspectives of processes like data being utilized, eligible agents who execute them, when and in which order processes are executed, and applications required for executing them etc. Therefore, this context awareness explicitly defined for each process can be utilized to support domain knowledge in data mining methods for delivering promising knowledge or actions again which are useful for process.
- Thus process management technology delivers context awareness that supports domain knowledge for data mining techniques. Also, data mining delivers actions that are again useful and proactively deployed in processes.

Convergence of data mining and process management technology is ideal – but still limited. By integrating data mining techniques inside business process management technology organizations, one can leverage from domain (knowledge) relevant information in mining methods for producing more concise knowledge useful for processes. Obviously, the APE

framework[10] is an example of such a convergence that addresses the problem of static assignment strategies of workflow technology. It uses the “goal definition” from process models not only as a success criteria but also as a source of domain knowledge, and then integrates data mining methods for performing process centric analytical analysis of employee’s performance thus enabling a WfMS to support proper resource management [10]. So the amalgamation of data mining techniques within the process layer to perform process centric analytical analysis of employees’ performance to support “continuous resource development”

So the amalgamation of data mining techniques within the process layer to perform process centric analytical analysis of employees’ performance to support “continuous resource development” will be helpful.

Indeed, after a thorough analysis of business process management technology we observed that process models are suitable source for domain knowledge especially when business intelligence is required within the application of a workflow. We realized that even though both technologies (process management and data mining) seem dissimilar, but still both support directly related promises:

- Data mining technology delivers knowledge that leads to “actions” that are incorporated and utilized within business environments for example by means of processes either manually by decision makers or automatically through an application.
- Process management technology defines the “context awareness” of business processes within the process models. Since, it defines syntax and semantics for different perspectives of processes like data being utilized, eligible agents who execute them, when and in which order processes are executed, and applications required for executing them etc. Therefore, this context awareness explicitly defined for each process can be utilized to support domain knowledge in data mining methods for delivering promising knowledge or actions again which are useful for process.
- Thus process management technology delivers context awareness that supports domain knowledge for data mining techniques. Also, data mining delivers actions that are again useful and proactively deployed in processes.

Convergence of data mining and process management technology is ideal – but still limited. By integrating data mining techniques inside business process management technology organizations, one can leverage from domain (knowledge) relevant information in mining methods for producing more concise knowledge useful for processes. Obviously, the APE framework is an example of such a convergence that addresses the problem of static assignment strategies of workflow technology. It uses the “goal definition” from process models not only as success criteria but also as a source of domain knowledge, and then integrates data mining

methods for performing process centric analytical analysis of employees’ performance thus enabling a WfMS to support proper resource management [Tal10a].

Fundamentals of these concepts and their interrelationships are based on the followings notions:

- Processes are “assigned” to organizational resources in order to execute them (link-1).
- Resources are responsible to “achieve” the goal (link-2).
- Achievement of a goal “demonstrates business success” thus validates the credibility of the link between resource and process (link-3).

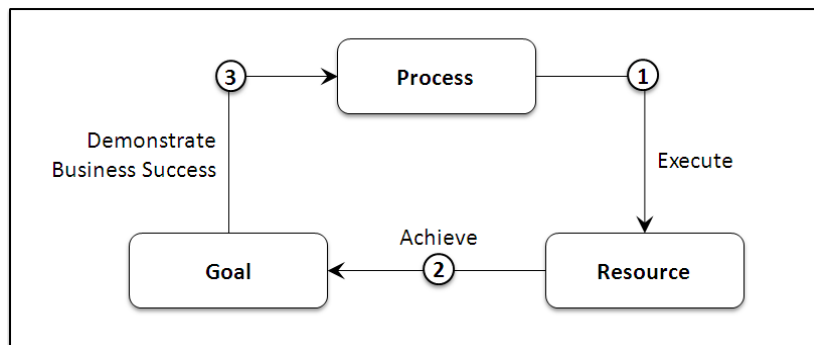


Figure: Process resource success conceptual relationship model

III. CONCLUSIONS

Indeed, after a thorough analysis of business process management technology we observed that process models are suitable source for domain knowledge especially when business intelligence is required within the application of a workflow. We realized that even though both technologies (process management and data mining) seem dissimilar, but still both support directly related promises [11]:

- Data mining technology delivers knowledge that leads to “actions” that are incorporated and utilized within business environments for example by means of processes either manually by decision makers or automatically through an application.
- Process management technology defines the “context awareness” of business processes within the process models. Since, it defines syntax and semantics for different perspectives of processes like data being utilized, eligible agents who execute them, when and in which order processes are executed, and applications required for executing them etc. Therefore, this context awareness explicitly defined for each process can be utilized to support domain knowledge in data mining methods for delivering promising knowledge or actions again which are useful for process.
- Thus process management technology delivers context awareness that supports domain knowledge for data mining techniques. Also, data mining delivers actions that are again useful and proactively deployed in processes.

Convergence of data mining and process management technology is ideal, but still limited. By integrating data mining techniques inside business process management technology organizations, one can leverage from domain (knowledge) relevant information in mining methods for producing more concise knowledge useful for processes. Obviously, the APE framework [10] is an example of such a convergence that addresses the problem of static assignment strategies of workflow technology. It uses the “goal definition” from process models not only as a success criteria but also as a source of domain knowledge, and then integrates data mining methods for performing process centric analytical analysis of employee’s performance thus enabling a WfMS to support proper resource management [10]. So the amalgamation of data mining techniques within the process layer to perform process centric analytical analysis will provide the support for “continuous resource development” and deep business analysis.

REFERENCES

- [1] Ramzan Talib, Bernhard Volz, Stefan Jablonski, “*Agent Assignment for Process anagement: Competency-driven Dynamic Resource Management Methodology*” Applied Informatics IVUniversity of Bayreuth Bayreuth, Germany {Ramzan.Talib, Bernhard.Volz, Stefan.Jablonski} @ uni-bayreuth.de
- [2] Keith Gile and Colin Teubner with Connie Moore and Lucy Fossner “*Business Intelligence Meets BPM In The Information Workplace*”. This is the sixth document in the “Information Workplace” series.by Keith Gile and Colin Teubner with Connie Moore and Lucy Fossner
- [3] Sylvia Radeschtitz, Bernhard Mitschangl and Frank Leymanrl “*Matching of Process Data and Operational Data for a Deep Business Analysis*”. UniversiHit Stuttgart, UniversiUitsstr. 38, 70569 Stuttgart, Gelmany firstname.lastname@{ipvs1-iaas2}.uni-stuttgart.de
- [4] WIL VAN DER AALST, “*Process Mining: Overview and Opportunities*”, Eindhoven University of Technology.
- [5] A.K. Alves de Medeiros, C.W. Gunther. “*Process Mining: Using CPN Tools to Create Test Logs for Mining Algorithms*,” In: K. Jensen, editor,Proceedings of the Sixth Workshop on the PracticalUse of Coloured Petri Nets and CPN Tools (CPN 2005), Aarhus, Denmark: University of Aarhus,2005.
- [6] W.M.P. van der Aalst et al, “*Business Process Mining: An Industrial Application*”.
- [7]W.M.P. van der Aalst and K.M. van Hee. “*Workflow Management: Models, Methods, and Systems*.” MIT press, Cambridge, MA, 2002.
- [8]Jablonski, S., Bussler, C. “*Workflow management.Modeling concepts,architecture and implementation*“. Thomson, London, UK, 1996.
- [9]Jablonski, S. “*Process Based Data Logistics: Data Integration for Healthcare Applications*.” 1st European Conference on eHealth (ECEH06), Fribourg, Switzerland, 10.2006.
- [10]Ramzan Talib, Stefan Jablonski, Stefan Jablonski, “*Agent Assignment for Process Management: Agent Performance Evaluation Framework*.” 4th Int’l. Workshop on Domain Driven Data Mining (DDD2010) in Conjunction with the 10th Int’l IEEE International Conference on Data Mining (ICDM2010) 14-17 December 2010. Sydney, Australia.
- [11] M. Ramzan Talib “ Agent Assignment for Process Management: Resource Management Support for Skill Intensive Applications of Workflow Technology “Thechnical and Methodological Issues.Von der Universität Bayreuth zur Erlangung des Grades eines Doctors der Naturwissenschaften (Dr. rer. nat.) Genehmigte Abhandlung.