17^{TH} EUROPEAN CONCURRENT ENGINEERING CONFERENCE 2011

ECEC'2011

7^{th} FUTURE BUSINESS TECHNOLOGY CONFERENCE

FUBUTEC'2011

EDITED BY

Ciprian Dobre

APRIL 18-20, 2011

LONDON, UNITED KINGDOM

A Publication of EUROSIS-ETI

Printed in Ghent, Belgium

17TH EUROPEAN CONCURRENT ENGINEERING CONFERENCE

7TH FUTURE BUSINESS TECHNOLOGY CONFERENCE

2011

LONDON, UNITED KINGDOM

APRIL 18-20, 2011

Hosted by

BITE
British Institute of Technology and Ecommerce

Organised by

ETI

Sponsored by

EUROSIS

BITE

University of Skovde

Higher Technological Institute
Tenth of Ramadan City

Ghent University

ECEC International Programme Committee

7.Formal Methods
Mauro Iacono, Seconda Universita di Napoli, Caserta, Italy
Yan Luo, NIST, Gaithersburg, USA
Stefano Marrone, Seconda Universita di Napoli, Caserta, Italy
Rachel Moreau, Hasselt University, Diepenbeek, Belgium

8.Collaborative CE environments for virtual teams
Khalil Drira, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Paul Stratil, Stuttgart, Germany
Roman V. Yampolskiy, University of Louisville, Louisville, USA

9.CE Enhanced Lean Manufacturing
Jean-Claude Hochon, AIRBUS, Blagnac, France .
Rik Van Landeghem, Ghent University, Zwijnaarde, Belgium .
Rob Walker, Anglia Ruskin University, Chelmsford, United Kingdom

10.Practical Applications and Experiences
James Gao, University of Greenwich, Chatham Maritime, Kent, United Kingdom
Irfan Manarvi, Center of Advanced Studies in Engineering, Islamabad, Pakistan

11. Factory Planning and Control for Small and Medium sized Enterprises (SME's)
Sílvio do Carmo Silva, University of Minho, Portugal
Luís António Nunes Lourenço, University of Beira Interior, Portugal
Paulo J.S. Gonçalves, Polytechnic Institute of Castelo Branco, Portugal
Nuno O. Fernandes, Polytechnic Institute of Castelo Branco, Portugal
António M. Fernandes, Polytechnic Institute of Castelo Branco, Portugal

12. Digital Direct Manufacturing
Philippe Geril, ETI, Ostend, Belgium
Peter K. K. Loh, Nanyang Technological University, Singapore

13. Patient Centered Health Care with Meaningful Use of IT
Gurdev Singh, State University of New York at Buffalo, Buffalo NY, USA

14. Student Session
Prof Algirdas Pakstas, London Metropolitan University, London, UK

FUBUTEC International Programme Committee

Statistical Analysis and Data Mining of Business Processes

Paulo Alexandre Ribeiro Cortez, University of Minho, Guimaraes, Portugal
Selwyn Piramithu, University of Florida, Gainesville, USA
Inna Pivkina, New Mexico State University, USA
Danny Van Welden, KBC Bank, Brussels, Belgium

Discrete Event Simulation and Queueing Systems in Business and Economics

J. Manuel Feliz Teixeiro, FEUP-University of Porto, Porto, Portugal
Leonid Novickis, IDC Information Technologies, Riga, Latvia

Agents in Business Automation and Economics

Martin Fredriksson, Blekinge Institute of Technology, Karlskrona, Sweden
Markus Koch, Orga Systems GmbH, Paderborn, Germany
Isabel Praca, Polytechnic of Porto, Porto, Portugal

Simulation in Business and Economics

Thomas Hanne, University of Applied Sciences Northwestern Switzerland, Olten, Switzerland
Sudhir Mudur, Concordia University, Montreal, Canada
Yan Luo, NIST, Gaitherburg, USA
Simon See, Oracle Systems, Singapore

Simulation in Business Games

Renate Sitte, Griffith University, Gold Coast, Australia
David Wortley, Coventry University, Coventry, United Kingdom

Simulation in OR and Knowledge Management

Chrissanthi Angeli, Technological Institute of Piraeus, Athens, Greece
Imed Kacem, Université Paul Verlaine Metz, Metz, France
Christophe Roche, Universite de Savoie, Le Bourget du Lac Cedex, France

Emergency Management and Risk Analysis Management

Kamil Drira, LAAS-CNRS, Université de Toulouse, Toulouse, France
Konstantinos Kirytopoulos, University of the Aegean, Greece
Gurdev Singh, UB Patient Safety Research Center, State University of New York at Buffalo, Buffalo NY USA
Lode Vermeersch, KBC Goldstate FMC Co, Shanghai, China P.R.

Simulation in E-Management, E-Government, E-Commerce and E-Trade

Vrassidas Leopoulos, Nat. technical University of Athens, Athens, Greece

Simulation Tools for Business Management and Business Intelligence Simulation

Carlo Meloni, Politecnico di Bari, Bari, Italy

Measuring the Costs and Effects of Environmental Policies for Business

Dietmar Boenke, Hochschule Reutlingen, Reutlingen, Germany
Elias A. Hadzilias, Universite Catholique de Lille, Lille, France
Roberto Razzoli, University of Genova, Genova, Italy

Ecological sustainable development and Innovative technologies for a bio-based economy

Philippe Geril, ETI Bvba, Ostend, Belgium

All articles published in these Proceedings have been peer reviewed

EUROSIS-ETI Publications are ISI-Thomson and INSPEC referenced
Legal Repository: Koninklijke Bibliotheek van België, Keizerslaan4, 1000 Brussels, Belgium

For permission to publish a complete paper write EUROSIS, c/o Philippe Geril, ETI Executive Director, Greenbridge NV, Wetenschapspark 1, Plassendale 1, B-8400 Ostend Belgium

EUROSIS is a Division of ETI Bvba, The European Technology Institute, Torhoutsesteenweg 162, Box 4, B-8400 Ostend, Belgium

Printed in Belgium by Reproduct NV, Ghent, Belgium
Cover Design by Grafisch Bedrijf Lammaing, Ostend, Belgium

EUROSIS-ETI Publication

ISBN: 978-90-77381-61-8
EAN : 978-90-77381-61-8

Preface

It is our great pleasure to welcome you in London, UK, for the FUture BUsiness TEChnology Conference (FUBUTEC'2011) and the European Concurrent Engineering Conference (ECEC'2011). As simulation and integrated knowledge management are regarded as the driving forces behind some of the world's largest and most successful business organizations, this year the conference is proud to host together many specialist and expert in business technology and computer science. The conference is held on April 18th -20th, 2011, at the British Institute of Technology & E-Commerce in London, UK.

The conference is meant to integrate business technology research of present day business practices such as "Operations Research" or "Business Process Simulation" into an even higher level enterprise wide framework with its new work roles, responsibilities, reward systems methods and tools. In other words, attaining true knowledge management is about radical and fundamentally new ways to create retain share and leverage knowledge of people and organizations in ways that were simply not possible before.

Next to the integral simulation part, the conference provides a strategic business overview of knowledge management in all its varied applications. In this aspect, the conference focus is on the latest knowledge strategies that business leaders need in order to become a Knowledge Organization and to withstand the forces of the financial and management markets in the present day precarious society, which constitutes the global environment.

The conference also embraces aspects such as Risk Analysis and Intelligence data Analysis in the certainty that the cross-fertilization of military thought and business environments can only lead to simulation product improvement. This year also features for the third year a special track on Sustainable Development.

The conference provides an open forum for researchers from academia and other research community to present, discusses, and exchange related ideas, results, and experiences in these areas. It aims at stimulating synergies between these new approaches, business technologies and knowledge management, risk analysis and intelligent data analysis, and the traditional models in this domain.

We would like to thank all the authors for submitting their research works to the workshop as well as to the authors of accepted papers for their participation and presentation of the papers in the conference.

We look forward to meet you all again in the 2012 edition of the conference!

General Conference Chair
Ciprian Dobre
University POLITEHNICA of Bucharest, Romania

CONTENTS

CONTENTS

DECISION SUPPORT TOOLS

AI APPLIED TO CONCURRENT ENGINEERING

FACTORY PLANNING AND CONTROL FOR SMALL AND MEDIUM SIZED ENTERPRISES

SCIENTIFIC PROGRAMME

BUSINESS INTELLIGENCE

MODELING OF STANDARDS AND THE OPEN WORLD ASSUMPTION

Mihnea Muraru, Alexandru Agache, Matei Popovici, Ciprian Dobre
Computer Science Department
POLITEHNICA University of Bucharest, Romania
email: {mmihnea,alexandruag,pdmatei,cipsmm}@gmail.com

KEYWORDS

Standards, BACnet, Modeling, Automated verification, Ontology, OWL, Open World Assumption

ABSTRACT

To be able to connect with the emerging technologies, real world concepts must be translated into machine-processable formats. Standards represent such collections of widely-accepted concepts and conventions, applicable in certain areas of interest. We believe that adopting formal, machine-understandable representations of standards generates a series of advantages, that we describe in this paper. We use the BACnet standard as a concrete example, and we present a possible approach to modeling it in an ontology. We highlight various design decisions and pitfalls encountered within the design process.

INTRODUCTION

Standards represent means to describe widely-accepted conventions in certain areas of interest. In many situations, it is of utmost importance that specifications are free of any inconsistencies. Usually, this is achieved through a continuous revision process, which doesn't rule out the presence of errors completely, but rather makes it unlikely. Furthermore, checking the compliance of different products and services is also an important issue. This has relevant applications in the field of Quality Management.

An alternative approach to checking inconsistencies is the use of automatic reasoning mechanisms. However, such techniques can only be applied if the specifications are translated into rigorous and unambiguous representations. This can be done by adopting different formalisms for describing factual data and their corresponding relations. A possible result is an ontology, that is a collection of such formal specifications. Many ontology modeling methods, such as the Ontology Web Language (OWL) [3], define precise semantics, thus allowing automated reasoning.

Besides checking for inconsistencies, a resoner may also be used to infer new knowledge, starting from what is explicitly represented in the ontology. Regardless of its final purpose, the reasoning process will often face difficulties generated by incomplete information. In such scenarios, a certain perspective must be adopted. These aspects, regarding Open and Closed World Assumptions, are discussed in more detail in a following section.

Formal representations offer a more concise perspective over the matter at hand. Written standards expose information in a linear manner, thus making connections between concepts more difficult to identify. A formal specification of such a standard induces some degree of structuring, by individualizing specific concepts and the relations in which they are involved. Visual tools would be able to exploit this kind of representations, in order to facilitate navigation throughout the problem universe.

Using ontologies to aggregate information allows publication, exploration and interconnection with other resources on the web, according to the *Linked Data* [1] principle. Such an established vocabulary could constitute the common ground for the automation of business processes that operate with standard-related information.

The paper is structured as follows. *Related work* presents similar initiatives related to modeling standards in ontologies. We then turn our attention towards the BACnet standard, described briefly in the section bearing the same name. In the *BACnet Standard Ontology* section we propose an ontological representation of the standard. The last section comprises our conclusions and future work.

RELATED WORK

Traditionally, ontologies provide a conceptual model of a domain of discourse, ocasionally simplifying or reformulating elements from the domain. Rigorous descriptions such as the one proposed in the article, are, to our knowledge, inexistent. Ontological representations of devices such as [5] and [6] provide a high-level view, focused on functionality from the user standpoint as well as device behavior. There is no explicit adherence to any particular standard. [7] gives a uniform representation for devices complaint with BACnet, KNX, LonWorks and ZigBee communication standards. While this approach is rather low-level, and concerned with technical, device-related details, it focuses on creating a common model, compatible with any of the enumerated standards. The emphasis is less on expressing standard-

related constraints and more on flexibility and integration. While this approach is succesful in certain device-related applications, it is insufficient in describing the complex structure of standards such as BACnet.
An ontology-based approach for detecting specification inconsistencies is described in [8], and applied on software configuration management. In a similar manner to ours, reasoning is applied on OWL ontologies in order to identify invalid test configurations. A notable difference relies in the representation domain. Software configurations discussed in [8] have a rather straightforward behavior and structure, and limited number of restrictions. Such structures can be adjusted in certain regards, to fit to representational needs specific to OWL. In contrast, human-developed standards raise complex representation problems, and are more rigid. The issues raised by OWA are such an example. Also, the development of representations for standards is useful for detecting standard inconsistencies. This issue is not addressed in approaches such as [8].

THE BACNET STANDARD

BACnet (Building Automation and Control Network) [2] is an open standard for communications between building automation devices, that was developed by ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineering). Its aim is to allow interoperation of various devices developed by different manufacturers. In order to attain such a target, BACnet proposes a complex model for describing building systems, irrespective to implementation details specific to manufacturers, or to the physical communication medium. At the heart of this model lie three concepts: *object*, *property* and *service*.
Depending on its purpose, a BACnet device may be broken down into a series of BACnet *objects*, that represent the fundamental control and communication units. It is important to note that the objects contained in a device merely describe capabilities of interfacing with that device. They do not offer means of describing its internal structure.
Each BACnet object possesses a standard set of *properties*, that reflect its current state. They represent the only means through which other devices can interact with a specific object. Every property is assigned a *datatype* and a *conformance code*. The *datatype* determines the internal representation and interpretation of the value of a property. The *conformance code* dictates how the value of the property may be accessed and whether its presence is mandatory (read-only, writable, optional). The standard imposes certain constraints on groups of properties, such as the necessity that a *Schedule* object should possess at least one of the *Weekly_Schedule* or *Exception_Schedule* properties, which are optional.
BACnet *services* are mechanisms that allow operations such as: reading and writing property values, generating commands on other devices and signaling events.
The BACnet standard also addresses many other aspects regarding device communication, which are not presented here, due to the lack of relevance to the modeling issues discussed in this paper.

THE BACNET STANDARD ONTOLOGY

This section covers the main design considerations involved in the development of a BACnet ontology, written in OWL. We focus on a specific part of the BACnet standard (2004 version), namely the fundamental concepts and their relations, leaving out very technical details such as the description of low-level communication protocols. By fundamental we refer to the building blocks of the BACnet modeling paradigm: objects, properties and services.
We start by introducing the overall structure of the ontology and then we delve into more specific issues.

Overall structure

Classes
The fundamental concept in the BACnet ontology, designed to represent any BACnet-related notion, is *BACnetEntity*. Its direct subclasses are:

- *BACnetDevice*: represents the 6 BACnet device types, such as *BACnet Operator Workstation* or *BACnet Building Controller*;
- *BACnetObject*: represents the 25 object types in the BACnet standard, e.g., *BinaryOutputObject* or *ScheduleObject*;
- *BACnetProperty*: represents the various properties the objects may possess, such as *ObjectIdentifierProperty* and *OutOfServiceProperty*;
- *BACnetService*: represents the 5 categories of services described in the BACnet standard: *AlarmEventService*, *FileAccessService*, *ObjectAccessService*, *RemoteDeviceManagementService*, *VirtualTerminalService*;
- *BACnetDatatype*: represents the possible datatypes for object properties. *Simple datatypes*, like *BACnetObjectIdentifier*, are individuals of the subclass *BACnetSimpleDatatype*. *Complex datatypes*, such as arrays and lists, are represented as individuals of another subclass, *BACnetComplexDatatype*, that is further divided into *BACnetArrayDatatype* and *BACnetListDatatype*;
- *BACnetConformanceCode*: represents the possible conformance codes for the object properties. It contains only two individuals: *R(eadable)* and *W(ritable)*.

Relations
We use relations to describe the connections between various concepts, according to the standard. From the OWL perspective, relations are divided into *object properties*[1] and *data properties*.
The object properties are:

- *hasProperty*[2]: associates an object with one of its properties. Its signature is: $BACnetObject \rightarrow BACnetProperty$;
- *hasDatatype*: associates a property with its datatype. Its signature is: $BACnetProperty \rightarrow BACnetDatatype$;
- *hasConformanceCode*: associates a property with its conformance code. Its signature is: $BACnetProperty \rightarrow BACnetConformanceCode$;
- *hasElementDatatype*: associates a complex type with the datatype of its elements. Its signature is: $BACnetComplexDatatype \rightarrow BACnetSimpleDatatype$[3].

The data properties are:

- *hasTheValue*: associates a property to its value, represented as a *string*. Its signature is: $BACnetProperty \rightarrow string$. It wasn't named *hasValue* so it wouldn't conflict with the OWL *hasValue* property restriction.
- *hasElementCount*: associates an array with its size, represented as an *integer*. Its signature is: $BACnetArrayDatatype \rightarrow integer$.

Naming conventions

Classes
The classes in the ontology have *suffixes* that correspond to their BACnet significance, e.g., *BinaryOutputObject*, *ObjectIdentifierProperty* and *BACnetArrayDatatype*.

Individuals
Individuals representing complex datatypes are named using the following conventions:

- *array* datatypes, that are instances of the *BACnetArrayDatatype* concept, appear as *NxElementType* (read *N times ElementType*). If the array size is fixed, N is replaced by the actual number of elements. If the array size is variable, the letter N is used directly. For example, the individual *3xBACnetTimeStampDatatype* stands for the type *BACnetARRAY*[3] *of BACnetTimeStamp*;
- *list* datatypes, that are instances of the *BACnetListDatatype* concept, appear as *ListofElementType*. For example, the individual *ListOfBACnetCalendarEntryDatatype* stands for the type *List of BACnetCalendarEntry*.

[1] Although the expression *object properties* is perfectly valid in the BACnet context, we use it here to refer to OWL properties, in the ontological sense.
[2] Disambiguation is needed: *hasProperty* is a property in OWL sense, that links a BACnet object to a BACnet property.
[3] We have assumed that values having complex types can only be made out of values having simple types.

Property attributes

Any property of a given object is characterized by a datatype and a conformance code. They are not intrinsically associated to a property and must be defined for *every* object.
For example, the BACnet specification of the *Calendar* object includes the *ObjectIdentifierProperty*, having the *BACnetObjectIdentifier* datatype and the *R* conformance code. This is represented as a *superclass* of the *CalendarObject* concept, as shown in Fig. 1. Note the use of the Manchester syntax [4].

Open vs. Closed World Assumptions

According to the Open World Assumption[4] (OWA), assertions are generally used to infer new knowledge rather than constrain existing knowledge. For example, the restriction (*hasProperty exactly* 1 *ActiveTextProperty*) produces the following behavior:

- if a particular object isn't explicitly asserted to possess an *Active_Text* property, the reasoner will conclude that in fact it possesses one that we don't know about, and no inconsistency is signalled;
- if a particular object is asserted to possess 2 *Active_Text* properties, the reasoner will infer that the two properties must represent the same entity. However, all the property instances can be declared different, in which case the reasoner will signal an inconsistency.

As one can notice, both scenarios are counterintuitive. Although in the second case, the constraint can be effectively enforced, in the first case it is *impossible* to impose the explicit presence of at least one property.
According to our current knowledge, reasoner extensions exist, that use the Closed World Assumption[5] (CWA).

[4] If something isn't explicitly stated to be true, it doesn't mean that it is false.
[5] If something isn't explicitly stated to be true, then it is false.

```
hasProperty exactly 1
  (ObjectIdentifierProperty
    and (hasDatatype value
          BACnetObjectIdentifierDatatype)
    and (hasConformanceCode value R))
```

Figure 1: Superclass of *CalendarObject*, denoting the attributes of its *Object_Identifier* BACnet property

We might find good use of such a reasoner, in order to enforce constraints in a more intuitive way. This is similar to the role of database schemas. For instance, defining a type for a table column is a means for restricting the possible values of that field, and not a mechanism for inferring types.

Disjointness

Because of the OWA, used by OWL, we must explicitly state as many facts as possible. Thus, all the concepts on the same hierarchical level are made *disjoint*. For example, the top-level concepts describing devices, objects, properties, are disjoint. Furthermore, within each concept, its direct subclasses are also disjoint, and so on.

Similarly, all the individuals that are direct or indirect instances of top-level concepts must be declared *different*.

Closure axioms

As stated in the *Property attributes* section, the definition of a BACnet object concept comprises qualified[6] cardinality restrictions for each associated property. The *ObjectIdentifierProperty*, mentioned in Fig. 1, was such an example. However, due to the OWA, stating that an object must possess certain properties does not enforce the absence of other properties, that are not explicitly mentioned. As an example, the *Active_Text* property could be added to a *Calendar* instance, although it is not part of the BACnet standard specification. In order to implement this constraint, a *closure axiom*, such as the one in Fig. 2, must be added as a superclass of the object in consideration.

This means that whatever properties our object possesses, they must belong to one of the concepts separated by the *or* operator.

[6] A *qualified* cardinality restriction constrains the range class of the property involved.

```
hasProperty only (
  (DateListProperty
    and (hasDatatype value
      ListOfBACnetCalendarEntryDatatype)
    and (hasConformanceCode value R))
  or
  (DescriptionProperty
    and (hasDatatype value
          CharacterStringDatatype)
    and (hasConformanceCode exactly 1
          BACnetConformanceCode))
  or ...)
```

Figure 2: Closure axiom for the *Calendar* object

Covering axioms

Similarly to the mechanisms described in the previous section, we also define *covering axioms* for the concepts in the ontology. Their purpose is to state that a class possesses only the explicitly defined subclasses and that no other unknown subclasses exist. The subclasses are said to *cover* the superclass.

We apply this principle to every concept in the ontology, by stating that each class is *equivalent* to the union of its subclasses. An example regarding the *BACnetEntity* concept is given in Fig. 3.

A particular situation regards those classes that contain only individuals. In this case, instead of being covered by its subclasses, each class is defined to be *equivalent* to an enumerated class containing the individuals. For example, the *BACnetConformanceCode* concept is equivalent to the class $\{R, W\}$, where R and W are individuals presented in the *Conformance codes* section.

Functional properties

Certain relations are inherently functional. This means that an individual is associated along these relations with *at most one* other individual. For example, a property associated with an object cannot have more than one datatype or conformance code.

If this characteristic was absent then, due to the OWA, one would be able to assign multiple datatypes and conformance codes to the same property within the same object.

The *Any* datatype

Several objects do not assign a specific datatype to some of their properties. This is encoded within the BACnet standard using the *Any* datatype. An example is the *Present_Value* property of the *Schedule* object, as shown in Fig. 4. The freedom in type choice is specified using the constraint (*hasDatatype exactly* 1 *BACnetDatatype*). This is different from properties having a fixed datatype, that are defined using the OWL *hasValue* property restriction, as in (*hasDatatype value CharacterStringDatatype*).

```
BACnetConformanceCode
or BACnetDatatype
or BACnetDevice
or BACnetObject
or BACnetProperty
or BACnetService
```

Figure 3: Covering axiom for the *BACnetEntity* concept

```
hasProperty exactly 1
  (PresentValueProperty
    and (hasDatatype exactly 1
          BACnetDatatype)
    and (hasConformanceCode value R))
```

Figure 4: Specifying the *Any* datatype of the *Present_Value* BACnet property

Conformance codes

In the BACnet standard, there are three conformance codes associated with properties: *R*(*eadable*), *W*(*ritable*) and *O*(*ptional*). The first two imply that the property is mandatory, while the last one allows its absence.
We have decided to explicitly model only the *R* and *W* conformance codes. Instead of explicitly stating that an object possesses a property having the *O* conformance code, we simply choose to allow at most one property of that kind. This will be reflected in the superclasses of an object. For example, the restriction in Fig. 1 will change from *exactly* to *max*, as depicted in Fig. 5. Also, notice that the conformance code isn't explicitly specified anymore.

Property constraints

The standard imposes certain constraints on groups of properties. Some examples are given below:

1. At least one of these properties is required.
2. If one of the optional properties *Inactive_Text* or *Active_Text* is present, then both of these properties shall be present.
3. If *Present_Value* is commandable, then it is required to be writable. This property is required to be writable when *Out_Of_Service* is TRUE.

Constraint #1 is applicable for the *Schedule* object and its optional *Weekly_Schedule* and *Exception_Schedule* properties. It is encoded as a superclass of the *ScheduleObject* concept, as shown in Fig. 6.
Constraint #2 is modeled as a superclass of the *BACnetObject* concept, as shown in Fig. 7. Unlike constraint

```
hasProperty max 1
  (DescriptionProperty
    and (hasDatatype value
          CharacterStringDatatype)
    and (hasConformanceCode exactly 1
          BACnetConformanceCode))
```

Figure 5: Specifying the *optional* conformance code of the *Description* BACnet property

```
(hasProperty exactly 1
    ExceptionScheduleProperty) or
(hasProperty exactly 1
    WeeklyScheduleProperty)
```

Figure 6: Specifying the mandatory presence of at least one of the *Weekly_Schedule* and *Exception_Schedule* properties for the *Schedule* object

```
((hasProperty exactly 0
    ActiveTextProperty) and
 (hasProperty exactly 0
    InactiveTextProperty))
or
((hasProperty exactly 1
    ActiveTextProperty) and
 (hasProperty exactly 1
    InactiveTextProperty))
```

Figure 7: Specifying the requirement of having the *Active_Text* and *Inactive_Text* properties both present or absent

#1, that was specified as a superclass of a particular object concept (*ScheduleObject*), constraint #2 was modeled as a superclass of the general *BACnetObject*, as it applies to several BACnet objects. The objects that do not possess these two properties are also members of the same superclass, because their closure axioms forbid the presence of both properties.
Constraint #3 proves different, as it involves dynamic conditions, that may change during the functioning of a device, e.g., the value of the *Out_Of_Service* property. This cannot be explicitly modeled in the ontology, using an approach similar to the one in the first two examples. Thus, we introduce a special conformance code, that will implicitly indicate this constraint, and assign it to the *Present_Value* property.

CONCLUSIONS

Although the OWA has its own advantages in knowledge representation, it proves inappropriate when modeling a standard, as they are closed in nature. As stated in the dedicated section, reasoners based on CWA may be used, but the available options are quite limited.
Using the proposed ontology, device specifications may be written and automatically validated against the standard. However, regarding the correctness of the ontology itself, specific tests have been used. We plan on developing an unit testing methodology that will allow automated verification of the ontology. The testing should verify that correct specifications are accepted (*positive tests*) and that invalid ones are rejected (*negative tests*).
The ontology was developed as part of research work for FCINT project (*Framework for service composition*

based on ontologies for the aggregation of knowledge and information for intelligent buildings).

ACKNOWLEDGMENT

The research presented in this paper is supported by national project “TRANSYS Models and Techniques for Traffic Optimizing in Urban Environments”, Project “CNCSIS-PD” ID: 238.

REFERENCES

[1] Linked Data. http://linkeddata.org/.

[2] BACnet Standard, 2004. http://www.bacnet.org/.

[3] OWL 2 Web Ontology Language, October 2009. http://www.w3.org/TR/owl2-overview/.

[4] OWL 2 Web Ontology Language Manchester Syntax, October 2009. http://www.w3.org/TR/owl2-manchester-syntax/.

[5] Dario Bonino and Fulvio Corno. DogOnt — Ontology Modeling for Intelligent Domotic Environments. pages 790–803. 2008.

[6] Harry Chen, Filip Perich, Tim Finin, and Anupam Joshi. SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications. *International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 258–267, August 2004.

[7] Christian Reinisch, Wolfgang Granzer, Fritz Praus, and Wolfgang Kastner. Integration of heterogeneous building automation systems using ontologies.

[8] Hamid Haidarian Shahri, James A. Hendler, and Adam A. Porter. Software configuration management using ontologies.

A SIMULATION-BASED EXPERIMENT ABOUT IRRATIONAL CHOICES IN BUSINESS

Emmanuel Dion
Audencia Nantes School of Management
8 route de la Jonelière
F - 44300 Nantes
France
edion@audencia.com

KEYWORDS

Decision, management, strategy, inefficiency, irrationality

ABSTRACT

Most managers make decisions rationally, and can justify their efficiency from the viewpoint of the firm's performance. The purpose of this paper is to examplify that in some cases though, managers' decisions do not comply with such an ideal scheme. In some cases, managers may be led to make irrational and underperforming choices, and to repeat them over time regardless of their inefficiency.

An experimental method has been set up to test this assumption. Based on a simulation game used with 578 participants, it shows that a certain useless budget, named the "X factor", is repeatedly spent away although it has no effect on the firms' performance. However, since these expenses are most of the time paid by the most profitable companies, which can afford it, and since these companies tend to remain profitable over time for other reasons, the useless expenditures are often maintained although they bring back nothing but additional costs.

1 - INTRODUCTION

Managerial decisions are usually based on a solid empirical understanding of their core business by the managers involved in the decision making process, in addition with a rational analysis of the situation in which they occur. Moreover, they are doubly guided by the historical and economic context on the one side, and on the other side by the explaining discourse that aims to justify them to the stakeholders. All those factors lead to the fact that most managerial decisions are strongly connected to reason and process conformity.

The purpose of this paper is to show that in some cases though, managerial decisions tend to escape from this double determination. When the context allows it (lack of relevant information, new situation, available funding), managers can be induced to make irrational and under-performing choices, and to reproduce them over time regardless of their inefficiency.

2 - THEORETICAL FRAMEWORK

The study of decision making has always been an important topic in management research. From Bowman (1963) to Kunc et al. (2010) and Papadakis et al. (2010), it was at the heart of a large number of publications in the most recognized journals in the field, maintaining an ongoing discussion between the different currents of thought involved, without ever the debate being interrupted.

Originally derived from both behavioral psychology (Bromiley, 2005) and decision theory in economics, this research theme has gradually acquired a certain autonomy (Schwenk, 1995). However, as noted by Eisenhardt and Zbaracki, the paradigm of this research is still to be refined, particularly regarding of its founding assumptions (Eisenhardt and Zbaracki, 1992).

Specifically, the role played by the context in strategic decision making (Bateman and Zeithaml, 1989; Bryson and Bromiley, 1993; Rajagopalan et al., 1993, 1997, Schneider and De Meyer, 1991) remains to be studied. This context can be examined under an organizational, environmental, or managerial perspective (Rajagopalan et al., 1997).

However, one of the biggest weaknesses of this field of research remains the scarcity of empirical studies connecting management theory and actual decision making, which is none the less remarkable than on the other hand, theoretical models tend to multiply (Papadakis et al. 1998). More specifically, management research would probably benefit from a closer interaction with applied economics on this topic. The present research aims to contribute to this effort.

3 – RESEARCH HYPOTHESES

The research hypothesis that we have made is that some inefficient managerial decisions tend to occur and to breed in organizations, invalidating any assumption of strong organizational efficiency.

This assumption is firstly based on the fact that a company's performance is typically measured in discrete time. Be it for accounting or organizational reasons, reporting is not done in continuous time, but periodically, at regular intervals (weekly, monthly, every year). This hypothesis is also consistent with the classical representation of the process of decision making in organizations, which is sequential in nature (Fredrickson, 1984, Mintzberg, Raisinghani, and Theoret, 1976). In this respect, we could seek to represent both decisions and outcomes in a given business as a series of numbers interpreted using the theory of Markov chains, but we do not need such sophistication in the context of this research. We only need to retain a representation scheme which postulates that both decisions and corporate performance tend to repeat over time by a stochastic process, and that decisions have an impact on results the corresponding period, and indirectly over the following periods (Fig. 1).

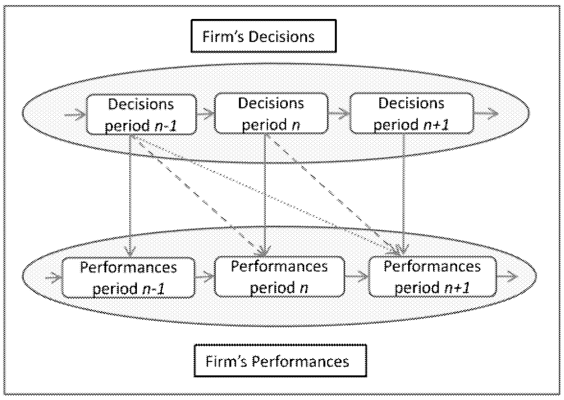


Figure 1 - Decisions and Business Results as an Iterative Process

We can also propose to integrate in our model the *context* in which decisions are taken, context that may have an effect on both decisions and results. For reasons of readability, we will at this stage start focusing on only two successive terms. In addition, and to make the model more explicit, we will mention on the scheme a number of supporting comments regarding the mentioned relations. For example, a link between the results of a given period and those of the next period may occur because some explaining factors are naturally recursive: financially, good results increase the cash flow, thus reducing debt and borrowing costs, thus contributing positively to earnings in the next period; commercially, increasing the market share enhances the competitive advantage and reputation of the company, which is an asset for the future.

It should also be noted that if the results of a company are partly determined by past results and effects of context, the strategic discourse will often comment on other aspects. It will more likely focus on that on which the managers may have a differential effect: decisions, their rationale and their consequences (Fig. 2).

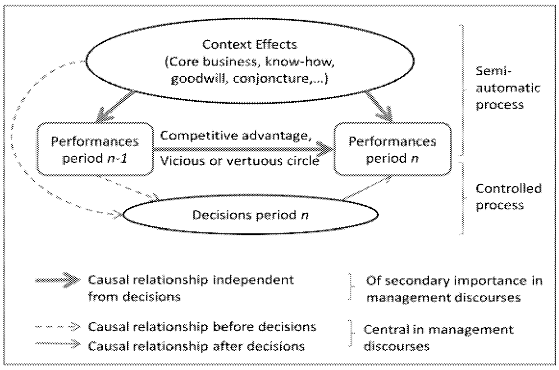


Figure 2 – Causal Relationships between Decisions and Performances

Although it is likely that for most strategic decisions, the process of decision making is largely rational, we can also assume that some decisions of lesser significance may emerge in a more random way. Decision makers not always having the means to ensure that they contribute significantly to the company's performance, some suboptimal choices may arise following an error of appreciation, a subjective preference, etc. However, we can also assume that such inefficient decisions are more likely to occur within specific organizations. In particular, one can postulate that the most profitable companies are likely to harbor the most decisions inefficiencies, partly because they are the only ones that can afford it. And we can also assume that these inefficient decisions can last a while, simply because the little impact (even though negative) on the firm's performance is more than compensated by the fact that the organizations hosting them being the most profitable, they somehow occupy a position of "stowaway" in the company: even if they are the subject of some publicity as decisions, they remain unnoticed from the viewpoint of their specific contribution to the performance. This may well be the case, in real economy, with certain expenses in public relations, sponsorship, patronage and even organizational or strategic consulting, or more recent measures concerning the social responsibility of the firm.

If we focus on trendy decisions that have not been sharply tested over time, we can then propose a simplified model similar to that shown in Figure 3. In this model, the link that appears between the decisions and performances of period *n* is the inevitable consequence of the relationship between these two variables taken separately and the results of the previous period. It's sort of a mathematically inevitable artifact, but it can be misinterpreted. Thus, it would be fair to say that companies making trendy decisions are more profitable than others, but it would be wrong to think or believe that it is *because* of these decisions that the outperformance appears (King, 2007; Gavetta and Levinthal, 2000).

To make our research as clear as possible, we have chosen to define our research hypotheses accordingly:

H1: Results of a given period are linked to those of the previous period

H2: The most profitable companies have a stronger propensity to engage into trendy decisions

H3: There exists a relationship between trendy decisions and performances that is only an artifact of H1 and H2

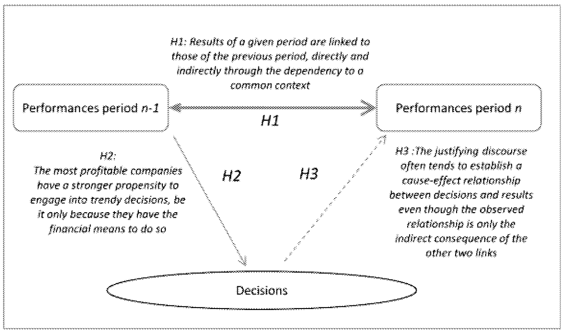


Figure 3 – Simplified Model

4 - METHOD

Following an old methodological tradition in the field (Remus, 1978, Sterman, 1987, Gist et al. 1998), we have chosen to address this research with an experimental method. Indeed the variety of situations encountered in real business life and the difficulties of categorizing and comparing make it tricky to resort to empirical studies. Therefore it is not surprising to observe that on the topic of decision making in management, research papers abound that propose purely theoretical approaches or models, but are significantly more scarce regarding the actual data and statistical analysis.

Our model of reference is that of behavioral psychology. This model has inspired many experiments conducted mainly from the 1950s to 1980s, under the impetus of Watson and Skinner. It has however so far been much less used in strategic management, because of the difficulty to sketch an experimental framework properly reflecting how a management decision is made, taking into account all that such a decision requires in terms of finesse and ability to integrate information of various kinds.

We furthermore chose to work, for reasons of convenience, on a public of business school students. Such a population is clearly not representative of all managers and decision makers in business, but we can assume that their reactions are inspired by the same logic, knowing that a few months after graduation, these students will usually find themselves in decision-making positions. In addition and for the sake of validation, the experiment was repeated on a number of continuing education students, who were taking their classes in while still assuming managerial responsibilities.

The experiment was grafted onto a strategic marketing course relying on a business simulation game. The simulation used (Jessie) has been known for a long time (more than ten years) by the teachers who participated in the experiment. According to a classic business game scheme, each class was divided into teams of 4 to 6 students which were asked to make a number of strategic decisions (investment, choosing a type of activity, product positioning, pricing, budgets, business, etc.) with the aim of maximizing their cumulative profit. These decisions are then confronted in a computer program carefully developed to produce the most possibly accurate simulation of the micro-economic reality supporting the game, and lead teams to achieve results they then try to interpret so that they can improve their subsequent performance. The simulation is repeated a dozen times, the whole sequence being characterized by a learning curve in participants' understanding of the model and a corresponding improvement in their level of performance.

The sample consists of 17 classes of 34 students on average, for a total of 578 participants. These 578 participants were grouped into 118 independent teams, which had the opportunity to make 310 decisions in the context of the experiment.

The experimental design consisted, after using the simulation for several weeks in order to accustom the subjects to the game (learning phase), to introduce a new variable, named "X Factor Budget" (test phase).

During the learning phase, students learn about the leverages available to them. They manipulate them at will and can get an idea of the effect of each variable on their performance by an elaborate feedback system. For example, regarding their selling price or quality level, they can access documents allowing them to assess their performance on this criterion, as well as their margin of possible improvements (local elasticities). The analytical capabilities of the teams are well tested, and performance tends to improve with time, indicating some learning curve in the order of rational decision.

During the test phase, the X Factor Budget is introduced by the professor who presents it explicitly as a budget whose singular effect may be difficult to measure. To limit some known biases in experimental psychology (Hawthorne effect, Pygmalion effect), the justification of this introduction is presented as neutrally as possible with explanation cards distributed to students (Fig. 4).

> Ranging between 0 and 10000, the « X Factor » budget corresponds to strategic consultancy operations aiming at improving the firm's global performance. In your professional environment, a consistent buzz has developed around this factor: must new source of performance according to some, useless management fashion according to others. Until the end of the game, you will never get a specific feedback on this particular point. The only way for you to try to assess its impact will be through the overall commercial performance of your firm, all products being equally concerned, whatever their type or age.

Figure 4 - Instructions Delivered to the Subjects of the Experiment

For the subjects of the experiment, it is not possible to know much about the effect of the X Factor Budget before they actually test it. Thus some teams choose to allocate a certain amount of money to this factor, and then draw conclusions about its effectiveness, given the overall results of their company. They then choose to cease this effort, gradually or abruptly, or on the contrary to renew it for a certain time span. Other subjects pay less attention to that budget and lose interest in it shortly after its introduction.

In the computer model, the effect of X Factor Budget can be set arbitrarily by the experimenter, with all degrees of possible scale between zero and critical importance.

In the experiment presented in this paper, the effect of the X Factor Budget was always set to zero. This means that this budget had no impact (either positive or negative) on the commercial performance of simulated firms, and consequently its only effect on the financial performance can be represented by a loss exactly matching the budget being squandered.

5 - RESULTS

Various computational methods were considered, the most significant in the end being also the most intuitive and easiest to understand at first. It simply consists in calculating the linear correlation coefficient between the model's variables, and to assess their significance level conventionally by using a Student *t* test.

The performance of simulated companies was measured through the account "cumulative retained earnings" of their balance sheet, which aggregates the results accumulated since the beginning of the simulation. This value being the one on which students were assessed at the end of the course, their aim to maximize it made little doubt.

Hypothesis H1

In the experimental setting, figures reported on 1413 results of 118 teams studied have shown a correlation of 0.91, which leads to a significance level way under 1%. We can therefore establish with near certainty a link between the results of a given period and those of the previous period, which perfectly validates hypothesis H1.

Hypothesis H2

Hypothesis H2 was more difficult to test. First, the logical link that can apply between outcomes and the propensity to invest in a decision with unknown consequences is not an unequivocal link. Some companies, well ahead of their competitors when the X Factor has been introduced into the game, did not see the point to take such an unnecessary risk. Conversely, some companies that were in trouble may have considered this option as a last resort. In addition, the introduction of the X Factor only mid-way through the game (imperatively dictated by the need for players to acquire sufficient expertise in simulation before they undergo the experience) has led to limit the number of observed decisions, and thus weaken the statistical power of the test.

The results were however consistent with the hypothesis tested. On 310 occasions to focus on the X Factor, simulated companies have chosen a non-zero expenditure 158 times (in 152 cases, the budget was set at 0). The linear correlation coefficient between the performance of the previous period and the amount spent for the X Factor established at 0.14. Although this is a modest figure, the *t* test shows that it is still significant with a 5% error margin. So we have a solid presumption that there exists a correlation which is not zero, but shows a relatively weak link between the past performance and the propensity to invest in an unknown expense. Without being able to present evidence in this research, we can also assume that this link is probably largely causal, and almost certainly predictive. Indeed, the perfect knowledge of the computer model proves without any doubt that there can be no reverse causal relationship (spending in the X factor can have no retroactive effect on past performance).

The hypothesis H2 is validated, but with a degree of certainty lower than H1.

Hypothesis H3

The latter hypothesis was the most difficult to defend. It should be understood that the X Factor, not only had no positive effect on results (which would have been sufficient to consider it logically inappropriate), but had quite a significant negative effect on them. It could indeed lead to wasteful spending in amounts corresponding often a significant fraction of the expected benefit, sometimes all of it, and on average one third (the maximum allowed budget was 10000, the average spent budget was 1800 when it was not zero, and the average yearly profit throughout the simulation was 5757).

Despite this, we observed a slight correlation of 0.09. However, this correlation is only significant with an error margin of 15%. So we have a real probability, though clearly unproven, that there exists a positive relationship between expenditures on the X Factor and the results of the corresponding period, despite the significant negative weight brought by these expenditures.

Since we know that there exists no "hard" link in the computer program between the X factor and the companies' performances, we can only attribute the observed link to some third party cause, most probably connected with H1 and H2. This can unfortunately not be proven because it is impossible to exhaustively test all possible third party causes.

H3 is confirmed, but with a low degree of certainty.

	H1	H2	H3
r	0,91	0,14	0,09
t	28,28	1,76	1,11
p	0,00	0,04	0,13
Error margin	0,01	0,05	0,15

Table 1 - Summary of Results

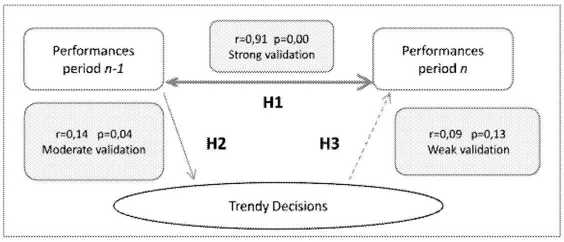


Figure 5 - Summary of Results

6 - DISCUSSION AND LIMITATIONS

The first criticism that can be addressed to the present research regards the type of method that has been used. One can indeed consider that the experimental method, scarcely used in management, is rare for some good reasons. The context of decision making is artificial, the stakes are limited to their academic impact and have no connection with real economy. These flaws are real and well known, but we can consider that they are the price that we need to pay for being able to have, as in experimental psychology, a rich and perfectly under control field experiment.

The second criticism, regarding the quality of experimental subjects (students rather than working executives) is probably less important. On the one hand, students who have been subjects of the experiment will be in positions of responsibility a few months or years after the test, probably with a mindset that will have changed little in the interval, and on the other hand, many participants were already executives with many years of experience, and showed no tendency to act differently from students with less experience.

Finally, the published results depend largely on the model used. It is obviously easy to change the settings of a computer simulation model to reach the results of an experiment in the desired direction. For example, to increase the level of significance of H3, it would have been easy to reduce the maximum budget allowed for the X Factor so as to minimize its negative impact on results.

Despite these limitations, it is satisfying to observe that the pattern of assumptions initially made, simple and robust, seemed properly validated by the results of the experiment, and we believe that the method can be repeated in the future to try examine in more detail some secondary propositions, such as the propensity of a wrong decision to repeat itself over time, or the propensity of inefficient decisions to emerge more easily as their consequences are benign.

REFERENCES

Bateman, T. S. and C. P. Zeithaml (1989). 'The psychological context of strategic decisions: A model and convergent experimental findings', *Strategic Management Journal*, 10, (1), pp. 59-74.

Bowman, E.H. (1963). 'Consistency and optimality in managerial decision making', *Management Science*, vol. 9.

Bromiley, P. (2005), *The behavioral foundations of strategic management*, Oxford, UK : Blackwell.

Eisenhardt, K. M. and M. J. Zbaracki (1992). 'Strategic decision-making', *Strategic Management Journal*, Winter Special Issue, 13, pp. 17-37.

Fredrickson, J. W. (1984). 'The comprehensiveness of strategic decision processes: Extension, observations, future directions', *Academy of Management Journal*, 27 (3), pp. 445-466.

Gavetti, G.; Levinthal, D. (2000), 'Looking forward and looking backward: Cognitive and experiential search', *Administrative Science Quarterly*, Vol. 45.

Gist, M. E.; H. Hopper; D. Daniels (1998), 'Behavioral simulation: Application and potential in management research', *Organisational Research Methods*, Vol. 1.

King, A. W. (2007), 'Disentangling interfirm and intrafirm causal ambiguity: A conceptual model of causal ambiguity and sustainable competitive advantage', *Academy of Management Review*, Vol. 32, p. 156-178.

Kunc, M. H.; Morecroft, J. D. W. (2010), 'Managerial decision making and firm performance under a resource-based paradigm', *Strategic Management Journal*, Nov 2010, Vol. 31 Issue 11, p. 1164-1182.

Langley, A. (1990). 'Patterns in the use of formal analysis in strategic decisions', *Organization Studies*, 11(1), pp. 17-45.

Lyles, M. A. (1987). 'Defining strategic problems: Subjective criteria of executives', *Organizational Studies*, 8 (3), pp. 263-280.

Mintzberg, H., D. Raisinghani and A. Theoret (1976). 'The structure of the 'unstructured' decision processes', *Administrative Science Quarterly*, 21, pp. 246-275.

Papadakis, V., Thanos, I., and Barwise, P. (2010). 'Research on Strategic Decisions: Taking Stock and Looking Ahead', *in* Nutt P. C. and Wilson D. C. (Éd.), *Handbook of Decision Making*: 31-69. Chichester: John Wiley.

Papadakis, V. , Lioukas, S. and Chambers, D. (1998), Strategic decision making processes : the role of management and context, *Strategic Management Journal*, Vol. 19, 115-147

Rajagopalan, N., A. M. A. Rasheed, D. K. Datta and G. M. Spreitzer (1997). 'A multi-theoretic model of strategic decision making'. *In* V. Papadakis and P. Barwise (eds.), *Strategic Decisions*. Kluwer, New York.

Remus, W. E. (1978), 'Testing Bowman's managerial coefficient theory using a competitive gaming environment', *Management Science*, Vol. 24.

Schwenk, C. R. (1995). 'Strategic decision making', *Journal of Management*, 21 (3), pp. 471-493.

Sterman, J. D. (1987), 'Testing Behavioral Simulation Models by Direct Experiment', *Management Science,* Vol. 33.

DATA MINING APPROACH FOR ANALYZING STUDENT PROFILES TO IMPROVE THE UNIVERSITY MARKETING POLICY

Dorina Kabakchieva
Sofia University “St. Kl. Ohridski”
125 Tsarigradsko shosse Blvd.
Sofia, Bulgaria
dorina@fmi.uni-sofia.bg

Kamelia Stefanova
University of National and World Economy
Students Town “Hr. Botev”
Sofia, Bulgaria
kamelia@fmi.uni-sofia.bg

KEYWORDS
Data Mining, Educational Data Mining, University Data Mining

ABSTRACT

Modern universities are collecting and keeping large volumes of data, referring to their students, the organization and management of the educational process, and other managerial issues. In most cases, these are unique types of data and their proper analysis could significantly support the university management in the decision making process and in the university policy development.
The data mining project, initiated at one of the biggest and most prestigious Bulgarian universities, is presented in this paper. The project objective is to develop the profile of university students according to the university performance results based on their pre-university characteristics. The obtained knowledge will support the university management in better defining the university marketing strategy, targeting and attracting the most appropriate and promising students.
Data mining is a new discipline in the ICT field. The implementation of sophisticated data mining methods and techniques for data analysis is a very innovative approach for Bulgarian organizations. The initiated project is the first attempt to apply data mining for processing university data in order to support decision making in the Bulgarian educational sector.

INTRODUCTION

Information technologies today are enabling organizations of all kinds to collect and store large quantities of data resulting from every day operational transactions. In most cases, however, this data is not efficiently used for supporting organizational managers in decision making. The available data is usually used for producing simple queries and traditional reports, but this information is rarely reaching the right people at the right time for making informed decision.
There is a continuously growing need to transform data into information and knowledge, considered very important organizational assets in the today’s knowledge-based economy and society. In response to that need, a great number of new information technologies, methods and tools, have been introduced and widely used during the last 15 years, known under different terms including Decision Support Systems, Executive Information Systems, etc., most of them being referred now as Business Intelligence (BI). Business Intelligence is a conceptual framework for decision support, combining architecture, databases (or data warehouse), analytical tools and applications, aimed at facilitating closing the gap between the current and desired performance of an organization as expressed in its mission, objectives and goals, and the strategy for achieving them.
The most sophisticated analytical methods and tools for achieving organizational “intelligence” are referred to as Data mining. Data mining is defined as the process of discovering meaningful patterns in large quantities of data which can be used for predicting future behavior thus leading to gaining some advantage, usually an economic advantage. Data mining is nowadays widely used in various areas – e.g. in business for marketing, sales and Customer Relationship Management (CRM), in medicine for diagnostics and image processing, in science for complex research studies involving processing of large data volumes (physics, astronomy, chemistry, etc.), to mention just a few.
The implementation of data mining methods and tools for analyzing data available at educational institutions, recently defined as “educational data mining” (EDM) (Romero et al. 2007) is a new stream in the data mining research field and the software industry. The educational data mining research community is constantly growing, starting by organizing workshops since 2004, then conducting an annual International Conference on EDM beginning since 2008, and now already having a Journal on EDM (the first issue being published in October 2009).
There are already a large number of research papers discussing various problems within the higher education sector and providing examples for their successful solutions reached by using data mining. Extensive literature reviews of the EDM research field are provided by Romero and Ventura in 2007 (Romero et al. 2007), covering the research efforts in the area between 1995 and 2005, and by Baker and Yacef in 2009 (Baker 2009), for the period after 2005. The problems that are most often attracting the attention of researchers and becoming the reasons for initiating data mining projects at higher education institutions are focused mainly on retention of students, improving institutional effectiveness, enrollment management, targeted marketing, and alumni management.
The rationale behind the research work described in this paper is based on the great potential that is seen in using data mining methods and techniques for effective usage of university data, and the identified needs for supporting

university management in decision making. Knowing the students better is expected to contribute to more efficient university enrollment campaigns, to attracting the most desirable students and to improving the organization and implementation of the educational and management processes at the university.

The paper describes the data mining project initiated at one of the most famous and prestigious Bulgarian universities. According to the authors' knowledge, this is the first attempt to apply data mining for processing university data in order to support decision making. The project is still in its early implementation phase and therefore the intended research approach and the first research activities are presented.

UNIVERSITY DATA MINING PROJECT RATIONALE

Universities today are challenged to meet the society and business requirements for adequately educating the future specialists that could fulfill the needs for advanced professional knowledge, technological expertise and open minded human behavior. Universities are becoming strong pillars of the knowledge-based society development and the most comprehensive institutions not only for delivering sophisticated knowledge and skills, but also for introducing innovative management approaches, supported by progressive methods and techniques.

Modern universities are organizations that are collecting and keeping large volumes of data, referring to their students, the organization and management of the educational process, and other managerial issues. The analysis of these unique types of data and the discovering of promising trends by applying data mining methods and tools, could substantially contribute to achieving flexible universities management and optimal performance results.

The EDM survey (Romero et al., 2007) reveals that the data in the educational sector is coming from two types of educational systems – traditional classroom and distance education. The implementation of data mining techniques for these two types of educational systems differs because of the different data sources and information systems used, and the specific goals and objectives followed. Most Bulgarian universities are still primarily organized based on the traditional classroom educational system model. University data is collected at the admission of new students, during the organization and implementation of the educational process, and for management issues. The collected data is always stored in paper format, and some or most of it is also available in electronic format. However, there are still no integrated university electronic systems developed. Data is usually accumulated in different text documents and electronic tables, in the best cases it is organized in relational databases, but there is no single data source available to provide university-wide, clean, and consistent data.

The application of data mining in educational systems can be addressed for contribution to different stakeholders – students, educators, administrators, managers, governmental institutions (Romero 2007; Baker 2009). Students could be supported by being recommended different learning resources, activities, tasks or even different learning paths. Educators can get more objective feedback and insight about the educational process that could help them to improve the content of the courses, to select adequate methods for the content provision, to differentiate students, based on their needs in guidance and monitoring, taking into consideration their learning abilities and peculiarities. Administrators could benefit from getting the right information at the moment when it is needed, thus supporting the decision-making process with adequate analysis. University managers could be assisted in their strategic tasks by receiving deep multi dimensional analysis that could reveal trends and opportunities for improving the effectiveness and efficiency of the university management. Governmental institutions could benefit by better analyzing the educational environment and making informed decision when defining the educational policies.

Universities in Europe are facing new problems in their policy development today. On one hand, they build the curricula profile based on the needs of businesses and society, on the other, they should attract the right students that will be able to perform best and fulfill the university objectives. Taking into consideration the whole changing environment of educational processes within the enlarged Europe, universities are competing strongly to identify their own uniqueness and to select the most appropriate students. Bulgarian universities are focusing on the future priorities and understand the importance of introducing innovative approaches in collecting and processing the required data for improving the enrollment and educational processes and the university performance as a whole.

Based on the review of the data available at the university and the identified university management needs (Kabakchieva et al. 2010), our research team decided to start the data mining project by focusing on problems concerning the university managers, and more specifically, to support them in better knowing their students and performing more effective university marketing policy. The literature review reveals that these problems have been of interest for various researchers during the last few years. Luan (2002) discusses the potential applications of data mining in higher education and explains how data mining saves resources while maximizing efficiency in academics. Understanding student types and targeted marketing based on data mining models are the research topics of several identified scientific papers (Ma et al 2000, Luan 2002, Luan 2004, Antons 2006). The implementation of predictive modeling for maximizing student recruitment and retention is presented in the study of Noel-Levitz (2008). These problems are also discussed by DeLong (2007). The development of enrollment prediction models based on student admissions data by applying different data mining methods is the research focus of Nandeshwar (2009).

THE RESEARCH APPROACH

The main goal of the initiated research project is to reveal the high potential of data mining applications for university management, referring to the optimal usage of data mining methods and techniques to deeply analyze the collected historical data.

The project specific objective is to define the profile of university students according to the university performance results, based on their pre-university characteristics available in the university admission campaigns data. The obtained knowledge is expected to support the university management in better defining the university marketing strategy, targeting and attracting the most appropriate and promising students.

The initiated data mining project at the university will be realized following the CRISP-DM (Cross-Industry Standard Process for Data Mining) model (Chapman et al. 2000). The CRISP-DM is chosen as a research approach because it is a non-propriety, freely available, and application-neutral standard for data mining projects, and it is widely used by researchers in the field during the last ten years. It is a cyclic approach, including six main phases – Business understanding, Data understanding, Data preparation, Modeling, Evaluation and Deployment. There are a number of internal feedback loops between the phases, resulting from the very complex non-linear nature of the data mining process and ensuring the achievement of consistent and reliable results.

The software tools that are considered to be used for the project implementation are the data mining open source software packages WEKA and RapidMiner.

The main project phases and activities, to be accomplished during the project implemented and based on CRISP-DM, are presented on Fig.1.

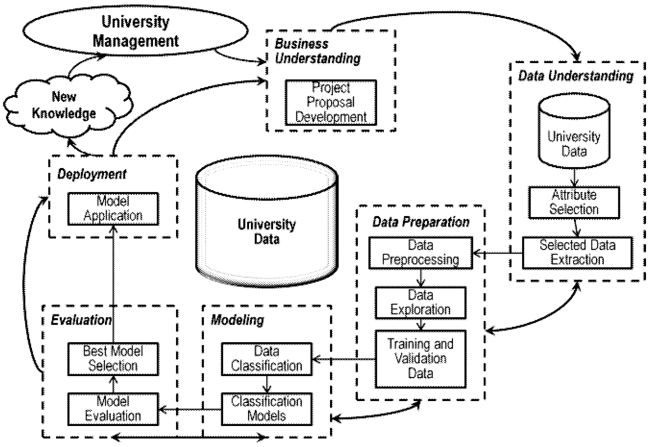


Figure 1: University Data Mining Research Project Implementation

The defined project goal and objectives are based on the results of the preliminary research performed, including an extensive literature review in the educational data mining field, formal interviews with representatives of the University management at university, faculty and departmental levels, and gathering insights from informal talks with lecturers, students and representatives of the administrative staff (IT experts and managers).

THE IMPLEMENTED RESEARCH ACTIVITIES

The university data mining project is still in its early stage of implementation. The performed research activities related to the data collection, selection and preparation, are described in this section.

The main project objective, to define the profile of university students according to the university performance results based on their pre-university characteristics, in data mining terms is considered a task for classification to be performed on the available student data.

The application process for student enrollment at the University was studied first, including the formal procedures and application documents, in order to identify the types of data collected from the university applicants and stored in the university databases in electronic format. The work continued with discussions with representatives of the administrative staff who are responsible for the university data collection, storage and maintenance. It was established that university data is basically stored in two databases. All the data related to the university admission campaigns is stored in the first database, including personal data of university applicants (names, addresses, secondary education scores, selected admission exams, etc.), data about the organization and performance of the admission exams, scores achieved by the applicants at the admission exams, data related to the final classification of applicants and student admission, etc. All the data concerning student performance at the university is stored in the second database, including student personal and administrative data, the grades achieved at the exams of the different subjects, etc.

Standard reports are usually generated from the two databases, supporting the University administrative management. The admission data and the student performance data are separately kept and queried. There have been no attempts for data integration and extensive analysis yet. However, the university management has realized the growing need for receiving accurate and timely information for making informed decisions. A discussion has already started for the necessity of building a university data warehouse to integrate all the available historic data, which is an important prerequisite for effective data analysis in order to get an overall picture of the university performance and a better insight about the educational and management processes.

The data mining task to be realized is to build a model that would classify the students into several classes, depending on their university performance and based on the student pre-university data. This is a task for supervised learning because the classification models are constructed from data where the target (or response) variable is known. The classification modeling task within the project will be implemented as presented on the logical diagram (Fig.2):

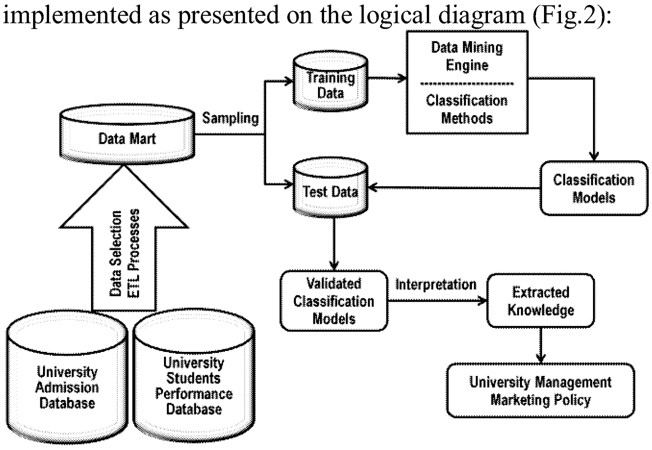


Figure 2: Logical Diagram for the implementation of the Data Mining Classification Task

The input data that is selected for the data mining task implementation is extracted from the University Admission and the University Students Performance Databases. It is then being pre-processed (Extract Transform Load – ETL processes), integrated and organized in a data mart. The data in the constructed data mart is distributed in two data sets by applying a selected sampling method. The first data set will be used as training data for the classification model generation, by applying different data mining methods for classification. The second data set will be used as test data for evaluation and validation of the generated the classification models. Based on the best performing classification models, the extracted knowledge will be described and presented to the university management to support them in the university marketing policy development and execution.

Input and Output Variables Selection

In any data mining exercise, one of the first tasks is to identify the input variables and the output (predicted) variable(s). Essentially, the challenge in the presented data mining project is to predict the student university performance based on the collection of attributes providing information about the student pre-university characteristics. The selected output variable in this case, or the concept to be learned by data modeling, is the "student class". It is a categorical target variable that has five distinct values (categories) - "excellent", "very good", "good", "average" and "bad". The five categories (classes) of the target (class) variable are determined from the total university score achieved by the students. A six-level scale is used in the Bulgarian educational system for evaluation of student performance at schools and universities. "Excellent" students are considered those who have a total university score in the range between 5.50 and 6.00, "very good" – in the range between 4.50 and 5.49, "good" – in the range between 3.50 and 4.49, "average" – in the range between 3.00 and 3.49, and "bad" – in the range below 3.00.
The input data attributes that are selected for the data mining task implementation include two types of data about the university students, personal data and pre-university data.

- Personal data: *gender*, *age* (calculated as the difference between the year of applying to the university and the year of birth contained in the personal identification number) and *special feature* (revealing any available advantages or peculiarities of the applicant, e.g. a person with disability who takes only oral exams, a contest winner who automatically receives the highest possible score for the admission exam, a regular student who is applying for a second specialization or is willing to change its specialization, twins, etc.);
- Pre-university data: *secondary school location* (city/village, region, country), *secondary school profile* (economic, technical, language, arts, sports, etc.), *secondary education scores*, *successfully passed admission exam* (selected between Literature, Mathematics, History, Geography, Economics), *successfully passed admission exam score*, *total score for admission/rejection*.

The selected input and output features include both, categorical and numerical variables. The generated classification models will provide an opportunity to extract knowledge about the profiles of the students according to their university performance, defining their gender, age, secondary school location and profile, scores from the secondary education, successfully passed admission exams and scores achieved at admission. That knowledge could support the university management in the decision making processes concerning the university marketing policy and marketing campaigns organization, for targeting and attracting applicants that have the highest chances of performing well at the university.
The study is limited to the student data for three university admission campaigns (for the time period between 2007 and 2009).

Data Mining Classification Methods

The defined data mining task for determining the profile of students according to the university performance results based on their pre-university characteristics could be realized by using a great variety of data mining methods, including decision trees, regression, neural networks, Bayesian classification, K-nearest neighbour method, etc. Several different classification algorithms will be implemented during the performed research work, selected because they have potential to yield good results and are accessible to a wider audience. They will then be evaluated and compared, and the best performing models will be selected for the knowledge extraction.
One of the software tools that will be used for the university data mining project implementation is the open source software WEKA, offering a wide range of classification methods for data mining (Witten, 2005). The study will be performed by using *OneR* classifier, two common decision tree algorithms CART (*SimpleCart*) and C4.5 (*J48*), and a Bayesian classifier (*BayesNet*).
The *OneR* classifier generates a one-level decision tree expressed in the form of a set of rules that all test one particular attribute. It is a simple, cheap method that often produces good rules with high accuracy for characterizing the structure in data. This classifier will be used as a baseline for the comparison between the other classification models, and as an indicator of the predictive power of particular attributes.
Decision trees are powerful and popular tools for classification. A decision tree is a tree-like structure, which starts from root attributes, and ends with leaf nodes. Generally a decision tree has several branches consisting of different attributes, the leaf node on each branch representing a class or a kind of class distribution. Decision tree algorithms describe the relationship among attributes, and the relative importance of attributes. The advantages of decision trees are that they represent rules which could easily be understood and interpreted by users, do not require complex data preparation, and perform well for numerical and categorical variables.
Bayesian classifiers are statistical classifiers that predict class membership by probabilities, such as the probability that a given sample belongs to a particular class. Several Bayes' algorithms have been developed, among which Bayesian networks and naive Bayes are the two fundamental methods. Naive Bayes algorithms assume that the effect that an attribute plays on a given class is independent of the values of other attributes. However, in

practice, dependencies often exist among attributes; hence Bayesian networks are graphical models, which can describe joint conditional probability distributions. Bayesian classifiers are popular classification algorithms due to their simplicity, computational efficiency and very good performance for real-world problems. Another important advantage is also that the Bayesian models are fast to train and to evaluate, and have a high accuracy in many domains.

The same classification methods will also be implemented in the RapidMiner software for data mining on the same data. The results achieved by applying similar methods when using different software tools will finally be compared.

Classification Model Evaluation

As already mentioned above, the supervised learning data mining exercise requires the use of at least two data sets – training data for the classification model generation and test data for evaluation of the classification error rates of the models. Stratified 10-fold cross-validation technique will be used when applying each of the selected classification algorithms, which is the standard way of predicting the error rate of a learning technique given a single, fixed sample of data. Data stratification will ensure that the data is divided randomly into 10 parts in which the class is represented in approximately the same proportions as in the full dataset. The 10-fold cross-validation technique will ensure re-sampling with multiple modeling experiments and different samples of the data. The learning procedure will be executed 10 times for each application of the selected classification algorithm, using in turn each of the ten parts as test data and the other nine-tenths as training data, providing an overall error estimate as an average of the 10 error estimates. Other techniques for selecting the training and testing data sets when applying the different data mining methods will also be consideted.

The accuracy of the generated classification models with the different data mining methods will be compared. The extracted knowledge will be presented to the university management to support them in the university marketing policy definition and execution.

CONCLUSIONS

A research data mining project is initiated at one of the most prestigious Bulgarian Universities in order to reveal the opportunities that data mining applications could provide to support the university management in better defining the university marketing policy and strategy. The project implementation is based on the CRISP-DM approach considered standard for data mining project development and fulfillment. The rationale for the project initiation and the initial research activities performed, including the data mining task definition, the data preparation and the selection of data mining methods for carrying out the research, are described. The project is still in its early implementation phase and it is the first attempt to apply data mining techniques and methods within the Bulgarian educational sector.

The future steps will be related to the implementation of the next project phases, including the application of the selected data mining methods on the prepared data and the generation of classification models, the model evaluation and knowledge extraction. The achieved results will be presented to the university management to support them in the decision making processes concerning the university marketing policy and actions.

REFERENCES

Antons, C., Maltz, E. (2006). Expanding the role of institutional research at small private universities: A case study in enrollment management using data mining. New Directions for Institutional Research, 2006: 69–81. doi: 10.1002/ir.188

Baker, R., Yacef, K. (2009). The State of Educational Data mining in 2009: A Review and Future Visions. Journal of Educational Data Mining, Vol.1, Issue 1, Oct. 2009, pp.3-17.

Chapman, P., et al. (2000). CRISP-DM 1.0: Step-by-step data mining guide. 2000 SPSS Inc. CRISPWP-0800. Available at: http://www.crisp-dm.org/CRISPWP-0800.pdf

DeLong, C., Radclie, P., Gorny, L. (2007). Recruiting for retention: Using data mining and machine learning to leverage the admissions process for improved freshman retention. Proc. of the Nat. Symposium on Student Retention 2007.

Educational Data Mining Research Community, http://www.educationaldatamining.org/

Dekker, G., Pechenizkiy, M., Vleeshouwers, J. (2009). Predicting Students Drop Out: A Case Study. Conference Proceedings of the 2nd International Conference on Educational Data Mining (EDM'09), 1-3 July 2009, Cordoba, Spain, pp.41-50.

Kabakchieva, D., Stefanova, K., Kisimov, V., Nikolov, R. (2010). Research Phases of University Data Mining Project Development. Proceedings of the Second International Conference on Software, Services and Semantic Technologies (S3T), 11-12 Sept. 2010, Varna, Bulgaria, pp.231-239.

Kumar, N., Uma, G. (2009). Improving Academic Performance of Students by Applying Data Mining techniques. European Journal of Scientific Research, ISSN 1450-216X, Vol.34, No.4 (2009), pp.526-534.

Luan, J. (2002). Data Mining and Its Applications in Higher Education. New Directions for Institutional Research, Special Issue titled Knowledge Management: Building a Competitive Advantage in Higher Education, Vol. 2002, Iss.113, pp.17–36.

Luan, J. (2004). Data Mining Applications in Higher Education. SPSS Executive Report, SPSS Inc,

Ma, Y., Liu, B., Wong, C. K., Yu, P. S., Lee, S. M. (2000). Targeting the right students using data mining. Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, Boston, pp 457-464.

Nandeshwar, A., Chaudhari, S. (2009). Enrollment Prediction Models Using Data Mining. Available at: http://nandeshwar.info/wp-content/uploads/2008/11/DMWVU_Project.pdf

Noel-Levitz White Paper (2008). Qualifying Enrollment Success: Maximizing Student Recruitment and Retention Through Predictive Modeling. Noel-Levitz, Inc., 2008.

Romero, C., Ventura, S. (2007). Educational Data Mining: A Survey from 1995 to 2005. Expert Systems with Applications 33, 2007, pp.135-146.

Shyamala, K., Rajagopalan, S. (2006). Data Mining Model for a Better Higher Educational System. Information Technology Journal 5 (3), 2006, pp.560-564.

Vialardi, C., Bravo, J., Shafti, L., Ortigosa, A. (2009). Recommendation in Higher Education Using Data Mining Techniques. Conference Proceedings of the 2nd International Conference on Educational Data Mining (EDM'09), 1-3 July 2009, Cordoba, Spain, pp. 190-199.

Witten, I., Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers, Elsevier Inc. 2005.

Formulation of hybrid methodology to develop a Management System based on Business Intelligence to support decision-making in Restaurants (RMS)

Ginna Largo Ordoñez
Technological Faculty
Distrital University Francisco Jose of Caldas
Highway 70 B # 73 a 35 South
Bogota, Colombia
E-mail: ginna_largo_ud@yahoo.com

KEYWORDS
Business intelligence, customer relationship management, on-line analytic process, management system for restaurants.

ABSTRACT

Despite the delicacies, comforts, and amenities that many restaurants offer, many of them are overwhelmed by taking orders and in most cases, retyping information numerous times throughout different departments, which makes work cumbersome and redundant.

This paper focuses on an analysis and implemented methodology, used to formulate a new management system that supports decision-making in restaurants. To accomplish this, we took into account the Customer Relationship Management (CRM) through stored data and business intelligence tools, with the customer's information and On-Line Analytical Processing (OLAP). Next, a restaurant management system (RMS) was developed. Finally, a hardware infrastructure was designed to support all implementations in real situations from "Top Down".

INTRODUCTION

Nowadays, many Colombian restaurants provide customers with their best dishes offered by the waiters, who are responsible for taking orders on sheets of paper and delivering the food when the meals are ready. Although these roles are carried out adequately, most restaurants do not have a process of bill optimization or an automatic method of sending purchase orders to the kitchen. As a result, a lot of downtime is incurred in taking orders, correcting them and then sending them to the kitchen personally. Consequently, the extra time incurred in completing this process might interrupt the customer's busy schedule and change their general perception of the service offered.

Therefore as we can see from the previous example, processing information without the use of technology makes things very difficult and inaccurate. It becomes challenging for managers to reach expected results, which in turn makes it harder for the business to grow and compete in the business environment. Current solutions to managing restaurants are concentrated on operational processes and most lack a model that allows them to gain aggregate knowledge about their customers which in turn could be transformed into income for the business.

All of this weakens sales, as there is no support infrastructure to keep a database or data warehouse with real time access to support business management in order to get statistics or reports, which are useful for making decisions.

Finally, without a management system that allows for remote monitoring of data, it is very difficult to control a restaurant if manager is not physically on the compound. Current advances in management systems, allow managers and staff to access information from anywhere, by using a computer or a device. These systems should possess the necessary scalability, where the system can be adapted as the business grows, to help with the decision-making process and offer solutions regarding the collection of customers' information.

APPROACH THE SOLUTION

Taking into account previous knowledge, a management system (RMS) was developed to make decisions based on the knowledge we have of customers. All data is stored in a database and is accessed through a mobile device or a computer that is connected to the Internet. Additionally, the database allows the restaurant manager to store customer information and knowledge captured from operations they perform (Sommerville 2002).

A RMS can be used in two ways according to the infrastructure. The first proposes a waiter, acting as an intermediary between the RMS and the client through a mobile device (PDA, IPAQ, handheld or smart phone). The second way proposes the client as the only actor who interacts with the RMS through a mobile device (tablet or slate pc). Thus, they only have to access physical distribution and select their tables, order meals, personalize dishes to their preference, see featured dishes, as well as the status of their order. They can also read daily newspapers or play games while awaiting their order.

Finally, information stored in the database may be accessed by the restaurant manager, who through systematic implementation of business intelligence and CRM services may offer improved customer service such as sending greeting cards to customers on special dates and newsletters. RMS have online services like text messaging and promotions and a system of complaints and suggestions which helps to improve service and transform the loyalty of customers, thereby motivating them to visit the restaurant again. It makes a big difference in restaurant businesses as it generates a cost-benefit ratio higher than normal operational applications.

CUSTOM RELATIONSHIIP MANAGEMENT MODEL

This chapter dealt with interactions management between the business and its customers. The focus was on optimizing the customer life cycle as a whole and a model based on CRM was formulated. This model offered a solution to all of the development perspectives, strategy implementation and implementation of solutions, in order to support decision making based on customer information, thereby generating added value through business intelligence techniques. It consists of four main phases: analysis, design, implementation of the strategy and implementation of the solution, Figure 1.

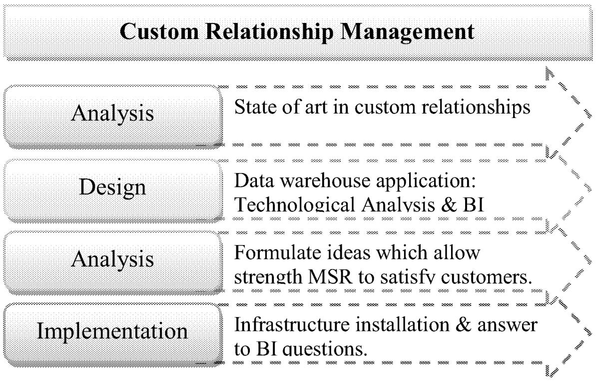


Figure 1: CRM Phases

Analysis

The first phase included a study of the current relationship with customers, competition, the status of the industry and the relationship with suppliers. This research was based on primary sources, which is a practice consistent with the formulation and evaluation of projects, in order to ensure a stable research and ascertain viability. Approximately 100 surveys were applied in different locations around Bogota, Table 1.

Table 1: Showing results from locations

Survey results			
Antonio Nariño	19	Ciudad Bolívar	6
Bosa	14	Teusaquillo	9
Chapinero	18	Usaquén	10
La Candelaria	14	Kennedy	10

Design

Considering the previous phase, we collected information and identified the factors that would be later used to design a strategy to respond to the needs expressed by the restaurants. The method was designed to obtain information about the customers, to obtain information about the way in which information is exchanged amongst different departments in the company as well as to decipher the appropriate technological platform that would be needed. To achieve this, we considered the five motivational factors which strengthen customer relationships. These were then implemented through two CRM tools: Data warehouse and Business Intelligence.

1. Data Warehouse

The purpose of implementing this tool was to centralize all company data, so that it could be easily, efficiently and effortlessly access. A prerequisite to this implementation was the search for technology that allowed us to build a collection of data. This was necessary to facilitate easy access to information, for the decision making process (Golfarelli et al. 1998). To make a decision and ensure that the solution met all needs, the premises of four of the five motivational factors were taken into account:

- <u>Information based on the client:</u> "partial knowledge of the client can be dangerous".
- <u>Personalization:</u> new technology allows for the customization of the customer relationship - CEP (customer experience personalization)[1].
- <u>Communications:</u> constant interaction with customers require 24/7 access to information, thereby taking advantage of new technologies with easily accessible and flexible transactions.
- <u>Transactions:</u> based on the need to generate reliable transactions to maintain trusted relationships with customers, taking advantage of moments of credibility and avoiding the excesses of information.

As a result, a technical analysis was done prior to the development of the management system for restaurants. Listed below are justified needs and their definitions by modules:

Table 2: RMS main modules

Billing	Where customers' orders are recorded.
Home orders	Management of door to door orders requested by Internet or by telephone.
Clients	Management of clients' personal and strategic information.
Stock	Management of hardware devices which are used in restaurant.
Stock analysis	Management of goods and products used to prepare dishes
Purchases from Supplies	Information relating to purchases is recorded in order to maintain stock levels in the restaurant.
Suppliers	Management of suppliers' information.
Sales reports	Management of reports related to IVA, ICA, other deductions on orders and other activities carried out in the restaurant.
Human resources	Management of employees' information.
Prices	Where real time prices are generated online.

[1]One should also personalize the customers' experiences and partially appeal to the emotions.

Additionally, the application has personalized customer services, taking into account the information from them and using online analytical processing:

Table 3: Modules oriented Customers

Dishes Customization System	Allows the customer to customize their order and the size of the portion.
Entertainment System	Allows the client to surf the internet, read newspapers or simply have access to games while awaiting their order.
Complaints tracking system	Allows management to deal with complaints or suggestions from customers by creating a feedback processes to improve services.
Frequent Customer Club	Offers personalized online services such as promotions, sales, E-mail birthday greetings and many others.

The implementation of these modules was carried out through the Rational Unified Process (RUP) software development methodology (Fraude 2003). Below describes the sequence of activities and devices that must be followed and used in order to achieve the desired result, in this case the restaurant management system, for each workflow (Jacobson et al. 2000).

- <u>Business modeling</u>: the business processes were defined, the domain model was determined as well as the glossary of terms.
- <u>Requirements</u>: the functional and non-functional system requirements were determined. The case usage model, system actors and the preliminary list of usage cases were obtained (OMG 2008).
- <u>Analysis</u>: the internal view of the system was defined through sequence diagrams, activity, collaboration, states and model classes.
- <u>Design</u>: the physical model of the system was defined through the initial listing of classes, their responsibilities, interface models, logical and physical structure of the data warehouse.
- <u>Implementation</u>: an initial prototype was obtained through system programming languages. It was conceptualized through the deployment diagram and components (QVT-Partner 2008). User manuals were generated.
- <u>Tests</u>: individual tests were performed and the integration of different modules and applications that make up the system were also tested.

The tools that would be necessary to develop the management system were defined through the previous methodology. Open source code tools to support the economic feasibility of the project were also chosen (Pressman 2002), to avoid licensing costs and allow cross-platform operation. For further details see Table 4.

Table 4: RMS Open Source Platform

Technological Platform	
Java eclipse	Use of java programming language on Eclipse software development environment.
Java Enterprise Edition	Usage of JEE 6 framework with JPA to improve communication with database.
RichFaces	JSF Technology with implementation of component library.
GlassFish	Application's server using Top Link (Java Persistence API).
PostgreSQL	Load and storage database through ORM (Object Relational Mapping).
Microsoft Windows debian redhat	SGR runs on Windows and Linux Operative Systems.
	Mozilla 3.6, IE 8 and Chrome 5 are some of web browsers in which SGR works.

2. Business Intelligence

Business intelligence (BI) is a broad category of applications and technologies which are used to capture, store, analyze and provide access to data that can help users to make better business decisions (Kimball et al. 2002). In this case the BI application tries to include activities such as a decision support system and consultation, reporting and online analytical processing. The final motivational factor for Data warehouse was classified within this category.

- <u>Segmentation and analysis</u>: uses an analysis tool for customer segments. These include Online Analytical Processing (OLAP) and Statistics and Data Mining (Data Mining or Data Extraction). Each of these techniques has strengths relating to certain types of customer information. In order carry out the segmentation and analysis, the following guided steps to OLAP were carried out on the management system (Herrera 2010).

First of all business questions which they wanted to resolve though the information analysis were defined. The RMS was defined around 10 questions which included "Who are my best clients?" a. By profitability b. By loyalty and c. For acquisition of new products.

Then the *Extract Transform Load (ETL)* process was implemented on our database tables where the information was listed. It began with the *extraction* stage, where the data was converted to a particular format. Next, *transformation,* where some business rules were applied to the extracted data, in order to convert them to data that would be uploaded and finally the *loading* process where information was uploaded and stored for safekeeping (Vassliadis 2002).

This new depository is made up of *data marts* or subsets of data that would help support decision making in specific areas of business, in this case for customer oriented RMS (Sapia 2008).

These *data marts* generated a logical model which made the design of data warehousing and the physical model which was used to define the logical storage structures and configuration of servers that would support the data warehouse using ROLAP, (Tryfona et al. 1999).

Finally the *Pentaho BI Server* was implemented based on these models, giving the results for each of the defined dimensions and metrics generated.

Strategy

For this third phase, when we speak about strategy implementation we are referring to the implementing of CRM which does not deal with software installation nor the wait for results but rather, with the complete change of the company's philosophy, where everything should be fully centered on customer satisfaction. As a result, all personnel should be trained and sensitized to the importance of having integrated departments where everyone has the same focus.
Within this context some aspects through which quality customer service could be strength were defined. One of them was to increase awareness amongst restaurant staff in order to strengthen the customer service philosophy. This was achieved through training in ISO 9001:2001 and the Colombian industry standard NTS-USNA 008: Categorization of restaurants: which allows holders to know the service parameters, infrastructure and good manufacturing practices based on international standards of operations procedures and administration, which was created in collaboration with the Colombian Association of Food Industry (Acodres). Other aspects are the analysis of customer preferences and market-leaders. These can be solved directly by restaurant management by periodically performing a market analysis in order to preserve and increase the number of customers, improve the delivery of service and innovation with respect to other local restaurants and markets.

Implementation

This is the last phase in the CRM which dealt with implementing the management system. However, a very important part for optimizing operation was still missing: the hardware infrastructure. As a result, we designed the network to support the operation of the RMS. It included the following characteristics: functionality, scalability, performance and capacity. At the same time, issues such as security, redundancy and management, were also taken into account to ensure consistency with the corporate goals of the restaurant (Oppenheimer 2010). The application of this methodology carefully followed the following steps:

Characterizing network: information which was administrative and technical in nature was collected and analyzed.

New customer requirements: the business constraints, security requirements, administration and application performance were all identified.

Topology design: a design based on a topology map according to customer needs was developed. This included a high-level view of internetworking devices and average interconnection. The hierarchical model for security and redundancy were also generated.

Hardware and media for LAN: to identify the appropriate hardware, aspects such as: the evolution of the services of layer 2 and 3, the solution to the internetworking problems, switching vs. routing, network segmentation using switches, broadcast radiation and scalability constraints in flat networks were all taken into account.

Hardware and media for WAN: reliable services, optimizing efficiency and minimizing costs were considered.

Model name and address: a hierarchical structure was designed for routing (autonomous system, areas, networks, subnets, and end stations), route summarization (aggregation), the distribution of administrative authority for addressing and naming.

Selection of the protocols: the routing and bridging protocols were selected.

Strategy management: the strategy for meeting design goals such as connectivity, security, cost optimization and manageable growth. In addition, we defined the network management process and the management and development of proactive management plan.

Documentation: all of the information generated was then gathered in order to create the network design.

Validation of the design: the design was validated according to the resources that the restaurant had and the size of the network. As a result a pilot plan was generated according to the methodology for small and medium enterprises, Figure 2.

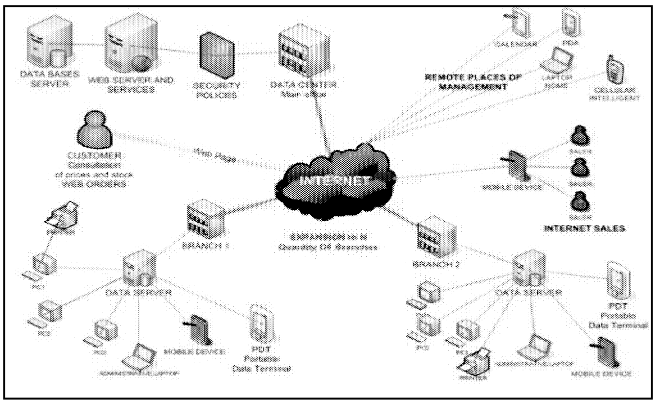


Figure 2: Generated pilot plan

The acquired infrastructure was comprised of lower-middle ranged servers, switches with embedded solution with VPN for remote administration, 2 Internet Service Providers in the market, handhelds, tablets, touch screens, printers and workstations.

Once all of the stages of the CRM are completed, the restaurant can start operating with a new philosophy. However it should always be controlled and monitored closely to react quickly to market needs and solve any errors that may arise in the future.

Figure 3: Management System for Restaurants (RMS) 'A new way to manage your business'

CONCLUSIONS

Creating independent software solutions to cater to business needs can become a headache for those who manage them and over time these may generate additional costs. The successful implementation of this methodology has already been proven. This can be seen in restaurant management systems aimed at improving customer relationships as well as exploiting and bring added value through information. Using the proposed methodology in this document other businesses in the sector can truly mitigate their problems. By having the answers to the different processes they can implement the tools to create robust systems and hardware infrastructure to support their solutions. The expected future work is to implement this project on a sample of restaurants and perhaps translate into other languages to be distributed in another country.

REFERENCES

Fraude, E. 2003. *Ingeniería de Software una perspectiva orientada a objetos.* Alafaomega.

Golfarelli, M., Maio, D. and Rizzi, S. 1998. 'The Dimensional fact Model: A Conceptual Model for Data Warehouses', *International Journal of Cooperative Information systems* (IJCIS), 7 (2-3): 215-247.

Herrera, E. *"Diseñando un modelo de inteligencia de negocios con UML: Estado de Arte".* 1-6. Added By: Erik José Enríquez Carmona, Universidad Veracruzana, Mexico, on October 14, 2010.

Jacobson, I., Booch, G., Rumbaugh, J. 2000. *"El Proceso Unificado de Desarrollo Software"*, Addison Wesley.

Kimball, R and Ross, M. 2002. "*The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*". Wiley. [Online]. Available: http://www.sigmod.org/sigmod/record/ issues/0309/R17.AnisimovBookReview.pdf

Object Management Group (OMG). 2008. *Common Warehouse Metamodel (CWM), version 1.1,* Internet: http://www.omg.org/technology/documents/modeling_spec_catalog.htm

Oppenheimer, P. 2010. 'Top Down Network Design', *A systems analysis approach to enterprise network design*. 250-342.

Pressman, R. 2002. *Ingeniería del Software*. 5^{TH} Edition. Mc Graw Hill.

QVT-Partner. 2008. "*Revised MOF 2.0 Query/ Views/ Transformations RFP".* Internet: http://www.omg.org/cgibin/doc?ad/05-03-02.

Sapia, C. Blaschka, M., Hofling, G. and Dinter, B. 'Extending the E/R Model for the Multidimensional Paradigm', *In Proceeding of the 1^{ST} International Workshop on Data Warehouse and Data Mining,* (DWDM'98), Vol. 1552 of Lecture Notes in computer Science, 105-116, Singapore. Springer- Velang.

Sommerville, I. 2002. *Ingeniería de Software*. Pearson Education. 6th Edition.

Tryfona, N., Busborg, F. and Christiansen, J.G.. 1999. 'starER: A Conceptual Model for Data Warehouse Desing', *In proceedings of the ACM 2^{nd} international Workshop on Data Warehousing and OLAP* (DOLAP`99), Kansas City, USA. ACM, 3-8.

Vassliadis, P., Simitsis, A. and Skiadopoulos, S. 2002. 'Conceptual Modeling for ETL Processes', *In proceeding of the ACM 5th International Workshop on Data Warehousing and OLAP* (DOLAP 2002). McLean, USA. ACM, 14-21.

BIOGRAPHY

GINNA LARGO ORDOÑEZ was born in Bogota, Colombia. There, she studied Technology in Systematization Data and Telematic Engineering, at the Distrital University Francisco Jose of Caldas, where she obtained her degrees in 2007 and 2011. Her collaborative research projects afforded her the opportunity to travel to various countries such as Miami, Buenos Aires, Lisbon, Villach and Barcelona. She worked for a few years at the Distrital University Francisco Jose of Caldas and in 2008 she was offered a position in the Minister of National Education. Currently, she is working in E-Government for Colombian Websites on Education.

A BUSINESS INTELLIGENCE SOLUTION FOR PUBLIC TRANSPORTATION SECTOR

Paulo Sampaio[1]
Guilherme Pereira[1]
Maria Sameiro Carvalho[1]
José Telhada[1]
António Paisana[1]
Paulo Paixão[2]
Albertino Fonseca[1]
[1]University of Minho, Systems and Production Department, Campus Gualtar, 4710-057 Braga, Portugal
E-mail: {paulosampaio, gui, sameiro, telhada, apaisana}@dps.uminho.pt; tino.fonseca@gmail.com
[2]EFACEC Sistemas de Electrónica, Rua Eng. Frederico Ulrich – Guardeiras, Apartado 3081, 4471-907 Moreira da Maia
E-mail: paulo.paixao@efacec.com

KEYWORDS
Business intelligence, performance measurement, intelligent transportation system, decision support system.

ABSTRACT

Smart companies in the 21st century use business intelligence solutions to gain a clearer picture of their internal operations, customers, supply chain and financial performance. In short, they are using business intelligence solutions to become intelligent about the way they do business. Our paper reflects the result of a project that we conducted in Portugal related to the development of a business intelligence solution for the intelligent transportation system (ITS). This project was developed on a close partnership between the University of Minho and Efacec, that is one of the most important and large Portuguese companies. More specifically, our aim was to: (1) perform a strict identification and selection of the information requirements; (2) develop technical specifications; (3) develop a business intelligence solution that will be used to support the companies top management decisions. Our approach consists on integrating in the computer-aided support system all the relevant information by mixing up different visualization forms (tables, graphs and maps), customized according to each performance indicator, thus enhancing reading efficiency and correct interpretation of every depicted situation being analyzed.

INTRODUCTION

This paper is focused on providing a systematization of the performance measures that are relevant to the transport system of EFACEC (public transport for passengers) – section 2, on proposing a data base model to support the information management system – section 3, and on proposing an adequate BI solution, creating some forms of visualization to enhance the effective and efficient use of the data and information available – section 4.

The company providing BI solutions for the transportation sector

With a history of already 100 years, the Efacec Group had its origin in "A Moderna", a company established in 1905.

Formed in 1948, Efacec, the largest Portuguese Group in the field of electricity, employs around 4500 people and has a volume of orders that already exceeded 1000 MEuros, being present in more than 65 countries.

The portfolio of Efacec business activities is composed by:

- Energy Solutions:
  - Transformers.
  - High and Medium Voltage Switchgear.
  - Service.
- Engineering solutions and Services:
  - Engineering.
  - Automation.
  - Maintenance.
  - Environment.
  - Renewable.
- Transport and Logistics:
  - Transport.
  - Logistics.

The Efacec portfolio is based on a Systems Integrating Contractor philosophy that covers the present needs of the market and turns the different Group capabilities profitable.

Efacec's strategy, focusing on International market, as well as important investments made in Innovation and New Technologies Development, backed by the resources of Efacec traditional technology led the Group to a position at the forefront of the Portuguese Industry and overseas markets.

The business intelligence concept

Business Intelligence (BI) solutions create learning organizations by enabling companies to follow a virtuous cycle of collecting and analyzing information, devising and acting on plans, and reviewing and refining the results. To support this cycle and gain the insights that BI delivers, organizations need to implement a BI system comprised of data warehousing and analytical environments.

According to Eckerson (2003), smart companies recognize that the systems that support BI solutions are very different from other systems in the company. Well-designed BI systems are adaptive by nature and they continually change to answer new and different business questions. The best way to adapt effectively is to start small and grow organically. Each new increment refines and extends the solution, adjusting to user feedback and new requirements. Additionally, the BI solutions combine data with analytical tools in order to provide information that will support top management strategic decisions (Cody and Kreulen 2002).

The main goal of a BI solution is to collect data, store the data collected and analyze the data gathered in order to produce knowledge (Santos and Ramos 2006).

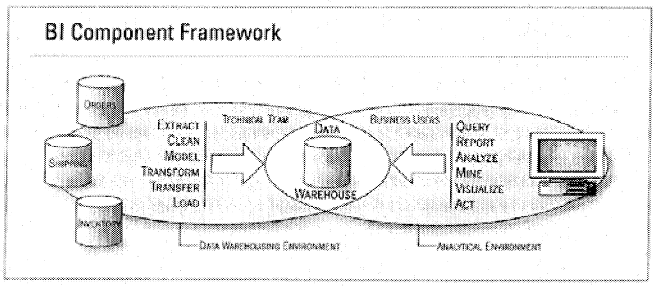


Figure 1: BI environment (Eckerson, 2003)

It is important that organizations understand the key indicators of success so they can surmount the challenges associated with every BI project. The successful BI solutions should have the following characteristics (Eckerson 2003):

1. Business sponsors highly committed and actively involved in the project.
2. Business users and the BI technical team working together closely.
3. The BI system viewed as an enterprise resource and given adequate funding and guidance to ensure long-term growth and viability.
4. Firms should provide users both static and interactive online views of data.
5. The BI team should have prior experience with BI.
6. All the company should be committed and involved with the BI solution.

Organizations that have deployed BI solutions cite many tangible and intangible benefits. Although it's difficult to associate a concrete Return on Investment (ROI) resulting from these benefits, most enlightened executives place huge value on having a "single version of the truth", better information for strategic and tactical decision making and more efficient processes (Eckerson 2003). The most common benefits cited in the literature are the followings:

1. Time savings.
2. "Single version of truth".
3. Better strategic and plans.
4. Better tactics and decisions.
5. More efficient processes.
6. Cost savings.
7. Greater customer and supplier satisfaction.
8. Greater employee satisfaction.
9. Return on Investment.
10. New revenues.
11. Total cost of ownership.
12. Shareholder value.

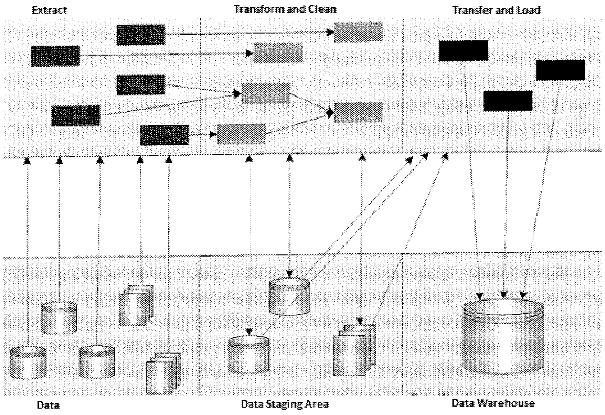


Figure 2: Data Extraction, Transformation and Cleaning, and Transfer and Load (Santos and Ramos, 2006)

PERFORMANCE MEASUREMENT FOR TRANSPORTATION SECTOR

Performance measurement is a way of monitoring progress towards a result or goal (Sousa *et al.*, 2005, Juran and Godfrey, 1998). It is also a process of gathering information to make well-informed decisions. Performance measures are valuable and provide several useful benefits (NCHRP 2006):

1. Greater accountability to policy-makers, organization customers and other stakeholders.
2. Improved communication of information about the transportation system to customers, political leaders, the public and other stakeholders.
3. Increased organizational efficiency in keeping agency staff focused on priorities and enabling managers to make decisions and adjustments in programs with greater confidence that their actions will have the desired effect.
4. Greater effectiveness in achieving meaningful objectives that have been identified through long-range planning and policy formulation.
5. A better understanding of the impacts of alternative courses of action that performance measures can provide.
6. Ongoing improvement of business processes and associated information through feedback.

Performance measures traditionally have been largely technical in nature. However, today transportation executives and managers must address an increasingly complicated and wide-ranging set of issues regarding the best solutions on balance to transportation problems, the cost-effectiveness of proposed projects and anticipated impacts of those projects.

Based on the literature related to performance measurement systems for transportation, there is a large number of measures within the following categories:

1. Preservation of assets.
2. Mobility and accessibility.
3. Operations and maintenance.
4. Safety.

The Transportation Research Board – TRB (NCHRP 2006) had developed a framework that transportation organizations should use to:

1. Identify performance measures that are most useful for asset management: assessing existing performance measures that are in place, identifying gaps, and considering new measures to fill gaps.
2. Integrate these performance measures effectively within the organization: engaging stakeholders to ensure buy-in, designing families of measures that can be used at different organizational levels and for different types of decisions, ensuring consistency across measures, identifying needed improvements to data collection and analytic tools, designing communication devices, and documenting measurement and reporting procedures.
3. Set performance targets: establishing both long-term goals and short-to medium-term targets for performance measures.

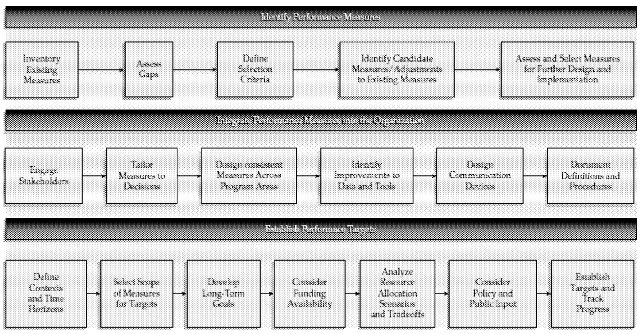


Figure 3: Guidance for Performance Measures and Targets (NCHRP 2006)

Rather than a single, prescriptive set of measures that might provide a good textbook example but surely would not be suitable for all companies, the framework is in the form of guidance on the preferred types and forms of measures and the process by which companies should identify and implement them. The framework was developed with the recognition that each company will have a different set of circumstances and needs to consider in selecting and implementing performance measures (NCHRP 2006).

One could not define the best set of performance measures for transportation companies, because each company has their own characteristics. However, the Transportation Research Board suggests a set of "good measures" that if implemented properly, will produce good results. The performance measures categories that have been defined by TRB are the following:

- Preservation – measures the condition of the transportation system and actions to keep the system in a state of good repair.
- Mobility and Accessibility – measures the ease of movement of people and goods.
- Operations and Maintenance – measures the effectiveness of the transportation system in terms of throughput and travel costs and revenues from a system perspective and maintenance level of service measures focused on the customer experience of the system.
- Safety – measures the quality of transportation service in terms of crashes or incidents that are harmful to people and damaging for freight, vehicles and transportation infrastructure.
- Economic development – measures direct and indirect impacts of transportation on the economy.
- Environmental impacts – measures effects on environment.
- Social impacts – measures effects on broader society or on population groups.
- Security – measures protection of travelers, freight, vehicles and system infrastructure from criminal and terrorist actions.
- Service Level – measures the delivery of transportation projects and services to customer.

The absence of measurement limits organizations ability to evaluate the changes and therefore precludes systematic improvement. Thus, good performance measures are a necessity for any progressive organization to recognize successful strategies and discard the unsuccessful. The performance measures categories defined for the model developed were supported in the TRB framework.

DATA BASE CONCEPTUAL MODEL

As it was already stated, this project aim was to develop a business intelligence solution for passenger transportation operators companies from the public sector.

Based on the literature review carried out and on the indicators identified, our proposal was to create a model based in the set of performance indicators that have been developed by the Transportation Research Board, integrated with the performance indicators defined and used by Efacec.

The model developed is supported in Data Base Model philosophy with the aim of organizing and managing the company information.

The most common model frameworks developed for business intelligence solutions could be based on a pure star schema design (one fact table connected with multiple dimension tables) that could evolve to a star schema design with multiple fact tables (one for each group of performance indicators) connected with multiple dimension tables, that, additionally, could evolve to a star scheme design with multiple fact tables sharing different dimension tables, usually known as constellation scheme design (Moody and Kortink, 2003).

The adopted scheme will have a significant impact in the rationalization of data storage space, concerning the information access and in terms of hardware necessities.

The scheme that best fit to the project goals was the constellation scheme design.

One important limitation of the constellation model is related to its dimension. It is important to point out that, regardless being the best solution, this option will require a higher computation performance that will be reflected in terms of higher software and hardware costs.

The constellation model proposal is based on one Data Warehouse that is composed by three Data Marts – Courses, Events and Incidents, with multiple dimension tables, some of them shared between more than one fact tables (Date and Time). These Data Marts intend to register historical information on facts occurring during delivering of transportation services (Figure 4).

Courses Data Mart is composed by 10 dimensions tables:

- Date
- Time
- Route
- Alternative Route
- Path
- Link
- Stop
- Schedule
- Vehicle
- Driver

Concerning Events Data Mart, there are 5 dimensions tables defined:

- Date
- Time
- System
- Equipment
- Failure type

The Incidents Data Mart is composed by 7 dimensions tables:

- Date
- Time
- Place
- Operator
- Tender
- Road
- Incident type

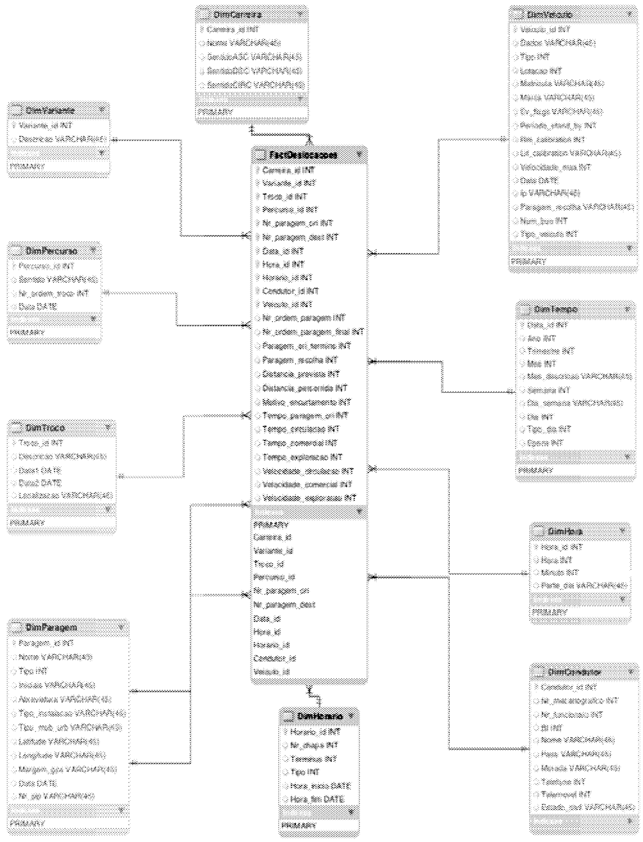


Figure 4: Conceptual Model developed for Courses Data Mart

BUSSINESS INTELLIGENCE SOLUTION

Intelligent transportation Systems (ITS) integrate a broad range of IS/IT elements for gathering on-site data and transfer it to the central database(s) of the enterprises or organizations. ITS also integrate the required hardware and software to analyze and manage the information extracted from data.

In general, the amount of data available (e.g. detailed traffic operations events and conditions) is huge, and only a relatively small fraction of that is adequately translated into information that is used by managers. In most cases, this is restricted to the information that is easily obtained by applying traditional query database operations and graphical displays, including standard statistical elements (e.g. historical means and deviations of traffic volumes) in general-purposed tables, graphs and maps. But, in fact, much more information can be potentially gathered from data: (1) further performance measures can still be obtained by traditional forms; (2) deeper information somehow hidden in the databases that can be highlighted by applying more advanced analytical methods, such as data mining, optimization and simulation, as well as advanced forecasting (e.g. auto-correlative, multi-regressive) and statistical techniques; and, very important, (3) most of all this information can be better highlighted by using alternative, more adequate forms of visualization such as well-designed graphs and maps; also, some information (e.g. spatial-temporal trends) can be exclusively evidenced by specific forms of visualization. It is widely recognized that it is (often) easier to detect a pattern from a "picture" than from a textual or numeric explanation.

From the above, it is obvious that one of the most important problems of ITS is the analysis and management of information. This problem is becoming further relevant due to the permanent increase on the availability of data and their related inexpensive storage and processing power as a result of the widely implementation of modern IS/IT systems.

The BI solution developed to Efacec is based on commercial available software, and most of the performance measures implemented were easily translated into the form of traditional tables, graphs and reports. This includes making use of simple dashboards. Most of these elements can be easily customized by end users, according to their specific needs and options.

Data visualization can be seen as the process of translating information into a visual form (usually graphs and maps), enabling users (scientists, engineers and managers) to perceive the features of the data available, and, from that point, to help them to make more effective decisions at different levels: strategic, tactical or operational. For example, deciding on the design of service operations strategies or, simply deciding on the realization of a functional regulatory (e.g. corrective) measure to ensure the normal functioning of the system or to minimize the impact of an anomalous event (incident) that has occurred.

In the following paragraphs three different examples show the importance of adequate visualization patterns according to specific nature of data recorded and information needed.

In order to detect peak periods and understand correlation to service delays may be worth using simple line charts of average traffic volumes (widely available from public data sources) and actual delays observed in the system, as a function of day-time (Figure 5). This should be done by taking into account the average figures for relevant zones or intersections that affects each PT service. Such information (patterns) can be also incorporated into adequate predictive models of time arrivals at stops (e.g. for public panel information).

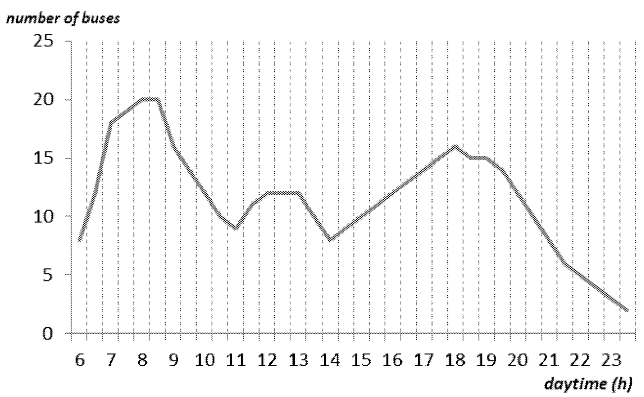


Figure: 5 Line chart of the average volume at a given location (e.g. PT stop or servive zone).

In order to perceive and monitor service levels in different directions from the city center (e.g. incoming PT services bringing people to work at the morning peak), may be worth using a rose-like diagram (Figure 6). This can be done per service or group of services in each direction (e.g., north, west, south and east bands).

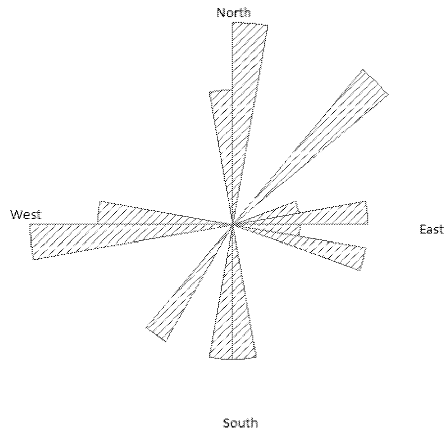


Figure 6: Rose-like graph of frequencies for 12 outbound PT services at the evening peak: each route is represented by a triangle illustrating its direction and relative frequency.

In order to infer the patterns of transit volumes in each direction per time of day (in average), may be worth using clustering analysis, e.g. putting in evidence similar directional service patterns among different time intervals. In these cases, the use of colored "graph-image" may be useful (Figure 7).

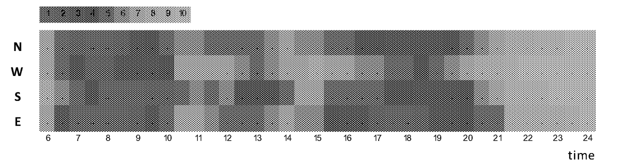


Figure 7: Average inbound and outbound public transit volume by direction and time of day (30 min clusters).

CONCLUSIONS

The Efacec BI solution developed produce performance reports for the transportation operators companies. The reports have two different levels. There are strategic reports that are only used by the companies' top management and there are operational reports to be used by the medium level management.

In this paper we do present the literature review carried out to this project and the model developed to support the Efacec business intelligence solution. The model adopted is supported in the framework developed by the Transportation Research Board for transportation companies, which define nine performance indicators categories.

One of the most important problems of ITS is the analysis and management of information. The BI solution presented, based on commercial available software, implemented relevant performance measures, translating into the form of traditional tables, graphs and reports, and allowing easy customization by end users. Thus, this work promoted the use of general proposed BI tools that, along with the integration of advanced analytical techniques, support the inference of both strategic and operational performance measures helping decision makers in the transportation sector – specific emphasis was devoted to visualization techniques that allow the detection of spatial-temporal patterns of behavior.

It is obvious that PT services must be monitored by means of dynamic (digital and vectorial) maps inter-operating with real-time data gathering devices (e.g. in-vehicle equipment including GPS). This sub-system of the ITS is commonly recognized as being a geographic information system (GIS). Most spatial-temporal patterns will be only detected by displaying the related information on the map, so this sub-system constitutes a must-be in a BI for transportation systems. Further work consists on establishing the integration of the BI application with a GIS model and software in order to enhance the usefulness of the overall decision support tool.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the European Commission and the Portuguese Government through the research grant QREN, in compliance with SmartCenter.

REFERENCES

Cody, W F, Kreulen, J T, 2002. The integration of business intelligence and knowledge management. *IBM Systems Journal* 41:697-713.

Eckerson, W, 2003. *Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligence Solutions*. USA. The Data Warehousing Institute.

Juran, J and Godfrey, A B, 1998. *Juran's Quality Handbook*. USA. McGraw-Hill.

Moody, D L and Kortink M A R, 2003. From ER Models to Dimensional Models: Bridging the Gap between OLTP and OLAP Design, Part I. *Business Intelligence Journal* 8: 7-24.

NCHRP, 2006. *Performance Measures and Targets for Transportation Asset Management*. USA. Transportation Research Board.

Santos, M Y, Ramos, I, 2006. *Business Intelligence: Tecnologias da Informação na Gestão do Conhecimento*. Portugal. FCA-Editora de Informática.

Sousa, S, Aspinwall, E, Sampaio, P, Rodrigues, A G, 2005. Performance Measures and Quality Tools in Portuguese Small Medium Enterprises: Survey Results. *Total Quality Management and Business Excellence* 16: 277–307.

AUTHORS BIOGRAPHY

PAULO SAMPAIO is graduated in Industrial Engineering and Management (University of Minho – Portugal, 2002) and holds a PhD degree in Systems and Production Engineering (University of Minho – Portugal, 2008). His main research interests are Quality Management, Operational Research and Applied Statistics.

GUILHERME PEREIRA was born in 1961 in Porto, Portugal. He graduated in Industrial Engineering and Management in the University of Minho, Portugal. He holds an MSc degree in Operational Research and a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. His main research interests are Operational Research and Simulation.

MARIA SAMEIRO CARVALHO is graduated in Computer and Systems Engineering (University of Minho – Portugal, 1986) and holds a MSC degree in Transportation Planning and Engineering (University of Leeds - UK, 1990) and a PhD degree in Transport Planning and Engineering (University of Leeds - UK, 1996). Her main research interests are in Transportation, Supply Chain Management and Operational Research.

JOSÉ TELHADA holds a PhD degree in Transport Planning and Engineering (ITS-University of Leeds – UK, 2001). His main research interests are Operational Research and Logistics applied to transportation systems.

ANTÓNIO PAISANA is graduated in Economics (University of Birmingham, 1978) and holds a PhD degree in Production Engineering from the Loughborough University, UK (1992). His main research interest is Strategic Management.

PAULO PAIXÃO is graduated in Electrical Engineer and Computers at Oporto University. He has been responsible for several National and European R&D projects. At present, Paulo is R&D Director of EFACEC Transports - Telecommunications and Signalling Division and Project Manager of SmartCenter activities.

ALBERTINO FONSECA is graduated in Informatics Engineering (University do Minho, 2010). At present he works at EFACEC.

Petri Net Based Simulation for SPI

Jan Kožusznik, Svatopluk Štolfa, David Ježek, Štěpán Kuchař
Department of Computer Science
VSB - Technical University of Ostrava
17.listopadu 15, Ostrava-Poruba
Czech Republic
E-mail: {jan.kozusznik, svatopluk.stolfa, david.jezek, stepan.kuchar}@vsb.cz

KEYWORDS
Business Modeling, Business Process, SPI, Simulation, BPM Method, Petri net

ABSTRACT

Specification of a software process, as well as its improvement, is a fundamental landmark and tenet that any successful software company must follow. One way how to support a software improvement process is using modelling and simulation. In this paper we focus on applying a discrete event based modeling and simulation to the real case study of the software process improvement activity. After a brief description of a modeling and simulation possibilities and the description of our tool and method we present the results of his particular case study.

INTRODUCTION - SPI AND ASSESSMENT

Describing and maturing the software process is a key element of a company's strategy, because a more mature software process means higher quality and more inexpensive software. Maturity of the software process is recognized through the assessment and its evaluation. According to the reference standards, a company's software process maturity is rated at a level from 0 to 5 (ISO/IEC 15504). If the company wants to have a more mature process, the process must follow appropriate good practices for a higher level. One of the possible ways how achieve the improvement is to model the process (Makinen and Varkoi 2008) or/and simulate the model to check the improvement's goals (Raffo and Wakeland 2008).
In this paper we are going to describe an application of a discrete event simulation based approach to the area of software process improvement.
This paper is organized as follows: section 2 describes the ways of software process modeling and simulation; section 3 summarizes the opportunities for modeling and simulation in the software process improvement process; section 4 presents our tool and method for discrete event based modeling and simulation. In section 5 we present the case study of software process improvement using our tool. Section 5 concludes and discusses future work.

SOFTWARE PROCESS MODELING AND SIMULATION

Generally, modeling means simplified description of reality and software process modeling encompasses description of processes connected to software development. Business processes represent the core of company behavior. They define activities which companies (their employees) perform to satisfy their customers. Software process is also business process (Štolfa et al. 2010). There exist many modeling technics for process modeling as is mentioned in (Vergidis, Tiwari, and Majeed 2008). On the other hand, the software process is quite specific (Raffo 1996) and it has been characterized as "the most complex endeavor humankind has ever attempted" (Brooks 1987). However, software process could be modeled formally (Curtis, Kellner, and Over 1992)
There were defined main objectives and goals for software process modeling (Curtis, Kellner, and Over 1992):

- facilitate human understanding and communication;
- support process improvement;
- support process management;
- automated guidance in performing process;
- automated execution support.

Created models will be processed with computer and that's the reason for their precise or formal specification. Simulation models of software process were used in last two decades (Abdel-Hamid and Madnick 1991) and they provide precise approach.
Main reasons why to simulate is defined in (Raffo and Wakeland 2008):

- strategic management (Raffo and Wernick 2001; Raffo and Kellner 2000),
- planning (Raffo and Kellner 2000),
- control and operational management (Raffo and Kellner 2000),
- process improvement and technology adoption (Raffo and Kellner 2000),
- understanding, (Raffo and Kellner 2000) and
- training and learning (Raffo and Kellner 2000).

Generally, simulation helps to achieve the optimal balance among quality, budget and duration – simulation help achieve balance(Rus, Collofello, and Lakey 1999). Simulation helps forecast and quantify process trajectories in respect to their actual performance (Scacchi 1999).
There are these leading paradigms used for software process simulation (Raffo and Wakeland 2008):

- discrete event simulation (Raffo, Vandeville, and Martin 1999) – controlled by discrete events occurrence, useful for modeling of activity orchestration;
- continuous simulation System Dynamics (Raffo, Vandeville, and Martin 1999; Ruiz, Ramos, and Toro 2004) – controlled by continuous time and change of

parameters in process is modeled as system of differential equations;
- state-based simulation – not widely used;
- hybrid model (Donzelli and Iazeolla 2001; Rus, Collofello, and Lakey 1999) – combines approach of discrete event approach and continuous simulation system dynamics;
- knowledge-based simulation(Scacchi 1999) - textual and primarily used for process understanding and educational purposes;
- agent-based simulation (David, Sichman, and Coelho 2003) - just starting to be applied to system and software development;
- qualitative (Zhang et al. 2006);

OPPORTUNITIES FOR MODELING IN ASSESSMENT

As was defined in introduction, our primary objective and goal is *supported process improvement*. There are defined main opportunities of simulation models using in (Raffo and Wakeland 2008):
- architect, compose, and document processes;
- estimate project costs from the bottom up;
- improve process planning and tradeoff decisions2;
- analyze and evaluate process improvement opportunities quantitatively;
- assess costs and benefits of applying new tools and technologies on a project;
- manage and control projects quantitatively;
- optimize development processes;
- train project staff;
- evaluate strategic issues.

PROCESS SIMULATION BASED ON PETRI NETS

Petri Nets are one of the formal mechanisms used for the description, verification and simulation of processes. The classical Petri Nets were founded by Carl Adam Petri ((Petri 1962)) as a basic tool to describe chemical processes, but since then they have evolved into very strong process modeling technique that supports temporal aspects of the process, stochastic behavior, hierarchisation of the models and even description of process resources and data (High-Level Petri Nets (Aalst 1994)). Properties of Petri Nets have been studied for over 40 years and this research makes Petri Nets a very well defined mechanism for verification and simulation of processes.

Petri Nets describe the process as a set of transitions (activities, tasks) that are connected with places. A place in the Petri Net can contain any number of tokens that represent available resources or occurrences of appropriate event. Whenever all the places preceding any transition are active (they contain at least one token), the transition can be performed. By firing the transition one token from each preceding place is taken away and at the same time one token is generated to every following place. This changes the state of the process that is described by the number of tokens in each place. Simulation in Petri Nets is a sequence of these atomic state changes and thus corresponds to the discrete event type of simulation.

Modeling and simulation of software processes and business processes in general presents some specific problems which lead to the creation of specialized modeling methods. BPM Method ((Vondrak, Szturc, and Kruzel 1999)) is one of these methods that is based on the Petri Nets approach and is used in the case study presented in this paper. BPM Method looks at the three elemental views of the process – architecture of the process, objects and resources utilized in the process and the behavior of the process. Each of these aspects is described by one of the models included in the BPM Method.

Functional model identifies the process architecture and its customers and products. The primary focus of the functional model is to answer which processes are cooperating with the main process and which subprocesses are used to perform specific tasks in the main process. This model isn't needed for the simulation because it just describes the overall architecture of the process, not the exact sequence of steps performed in the process.

Object model identifies static structure of all objects and resources that are essential for the enactment of the process. This model captures all workers employed in the process and their communication channels. It also contains information about all artifacts (documents, products, material, etc.) that are manipulated or created in the process. Each worker and artifact can be characterized by various optional attributes. This model is useful for the simulation purposes because it contains important information about the abilities and properties of the workers and artifacts, that can be utilized in the simulation.

Coordination model specifies the behavior of the process as a sequence of activities, what resources the activities demand and which artifacts are consumed and produced. Alternative flow in the coordination model is enabled by multiple activity scenarios and concurrency of the activities can also be modeled. Complex activity chains can be substituted by sub processes allowing hierarchisation that helps to clarify the process. This model is based on the Petri Nets formalism and is the most important part of the simulation.

Each of these three models can be visualized by appropriate type of diagram and because the coordination model is the most important one for performing simulations of the modeled process, it would be appropriate to show how its diagram looks. A small part of the Coordination diagram used in the case study presented in this paper is depicted in Figure 1.

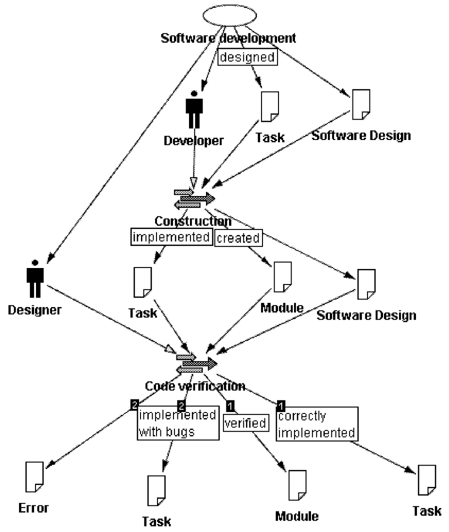


Figure 1: Coordination diagram

This diagram describes a part of the software construction subprocess. The first modeling element at the top of diagram shows the interface with the main process called Software development. This modeling element allows for the hierarchisation of the model. The Designer and Developer places are the human resources that come from the main process, and Task and Software design places define the appropriate artifacts. When the Task and Software Design documents are ready and the Developer resource is available the Construction activity can be performed. By completing this activity the state of the Task has changed to implemented and new module was created. The next activity is Code verification that is performed by the Designer and this activity can end up in two ways (two scenarios). The first scenario signifies that the constructed code is correct and it's outputs are marked by number 1. The second scenario shows that there were errors in the implementation and the process will continue by reporting and repairing the error. Outputs of the second scenario are marked by number 2.

We had to extend this basic BPM Method by a number of new mechanics that were needed for performing more relevant and more accurate simulations needed for software process assessment and improvement. These extensions are all supported by various types of Petri Nets and their properties and so they don't disturb the formal nature of the method.

The first extension was the addition of stochastic parameters to activities in the coordination model. Stochastic properties were added to the time duration of activities, their waiting time duration and their scenario probability. These parameters and their conversion to the Petri Nets properties were described in our last paper (Kuchař and Kožusznik 2010). In this paper we also introduced a very important mechanic for running automatic simulations for modeled processes to test their performance for multiple concurrently running projects. This requirement was fulfilled by adding the stochastic parameter to objects in the coordination model that defined time intervals between individual entries of new customers to the process (Kuchař and Kožusznik 2010).

But to complete the possibility to run the performance testing simulations another extension was also needed – sharing of limited human and artificial resources in the process. We managed to implement this feature by introducing the pools of limited resources to the objects in the coordination model. Objects with the same shared pool are then linked together and when one of the resources in the pool is used to perform an activity, it is taken from all linked objects. When the activity finishes, the used resource is returned to the pool and thus returns to all linked objects. This behavior can be converted to the Petri Nets formalism using the fusion places (Huber, Jensen, and Shapiro 1991). Each shared pool corresponds to one fusion place that is shared between all concurrent simulation instances of the process. This fusion place is then set as the input and output place for all activities that require shared resources from the appropriate shared pool.

Notation of the coordination diagram can also describe *chained execution* (Aalst 1998) if one object acts as an output place of any activity and at the same time as an input place of another activity. When this happens the same shared resource should be used for both the activities. This can be easily done by adopting some colored properties to the resulting Petri Net (Colored Petri Nets (Kurt 1997)), coloring the shared resources that should be chained and then matching the colors while using the resources for all activities in the chain.

The last extension that we added to the coordination model is the notion of weighted arcs. Arcs in the Petri Net can be weighted by any natural number and this defines how many tokens are consumed from the input place or produced to the output place when the transition is fired. BPM Method didn't have this property but it is very useful in describing the software processes.

We implemented all these extensions into the BPStudio software tool (Vondrák 2000) used for modeling, simulation and enactment of the processes based on the BPM Method approach and used it for the simulation of the software process and its improvements described in the following case study.

CASE STUDY

We described software development process of a software development company from the Czech Republic. The company employees about 100 employees and has been established in 90'ths of last century.

The company software process has been modeled and simulated with BPStudio software tool. The model of the currently performed process was taken as an initial point and was enhanced in several ways. Simulation results of the currently performed process and the enhanced process was compared and the result was discussed with the management of the company.

The process consists of several sub processes that can be recognized as standard parts of a common software development process.

The process starts with a *Requirement Elicitation and Analysis* sub process where a project manager is mainly utilized. That sub process produce document *Analysis* of requirement with specified and described futures requested by the customer.

Second sub process is *Architecture Design*. *Analyst* is the main role for this sub process. The main purpose of it is the

design of new system architecture. System is divided into smaller parts called system blocks.
Designer creates design for each system block in sub process called *Software Design* and define task for developers.
After previous, *Software Construction* follows and developers create parts of source code based on task from designer. Developed source code is reviewed against design by *designer*. If error is found developer has to rework source code.

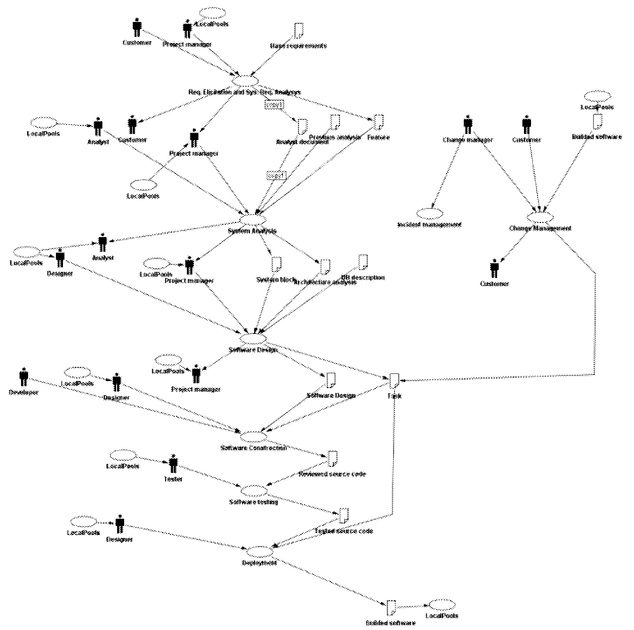


Figure 2: Overall schema of the software process model

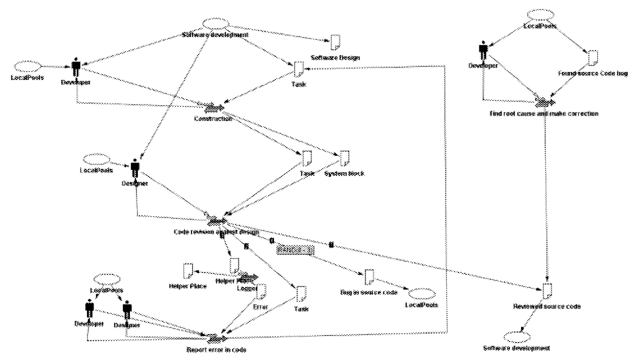


Figure 3: Schema of the sub process Software construction

First experiment examines influence of revision to the process. Common parameters for both simulations were numbers of employees in project. Concrete numbers for each role are in Table 1.

Table 1: Numbers of employees

Role	Number of employees
Project manager	1
Analyst	1
Designer	2
Developer	5
Tester	1
Change manager	1

There are produced artifacts in the every part of the process And they are used in the different one part. During this production, there are also generated faults[1]. Rate of the artifact and faults production are in Table 2. Generated faults could cause a failure in the delivered system. Failures are reported as *incidents* by customer. Faults from different sub process of the development process have different probability of detection by the customer (Table 2). Sub process *Software testing* is quite different because it does not generate any fault but try to find them. The success rate of the *Software testing* sub process is in Table 2, there is also defined probability of found specific fault base on its type.
Software process extended by revision has the same error rate as mentioned before, but extends software process by revision in sub process *Requirement Analysis*, *System Analysis*, and *Software Design*. Each revision has defined rate of errors removed from reviewed artifact (Table 3).

Table 2: Reate of artifacts and errors production

Phase	Artifacts	Errors	Probability of incident
Requirement Elicitation	5 – 12	0 – 3	70:30
Architecture Design	2 – 4	0 – 2	40:60
Software Design	2 – 5	0 – 2	50:50
Software Construction	1	0 – 3	70:30
	Revision success rate		
Software Testing	50:50 (20:10:10:60)		

Table 3: Success rate of revisions

Phase	Revision success rate
Requirement Analysis	0-3
System Analysis	0-2
Software Design	0-2
Software Construction	60:40

The software process with parameters, as mentioned above, was simulated in the tool BPStudio. Each simulation was run 6 times and numbers of artifacts, faults and other ranged values were randomly generated for each run. Random number generator with normal distribution was used for all values.
Simulation results are in Table 4 and Table 5. There is total cost of project, duration of software development, and utilization for each role. The total cost of project is calculated from the cost of all activities with specified resources (employees). Utilization is related to the duration of all activities in projects.

Table 4: Simulation results for process without revisions

Utilization [%]	Minimum	Average	Maximum

[1] A terminology for describing bugs is defined on the page 6 in (Pfleeger and Atlee 2009)

Project manager	4	4,34	5
Analyst	4	4,00	4
Designer	15	17,67	19
Developer	29	32,50	36
Tester	12	12,50	13
Change manager	3	4,34	6
Project duration	32,8	38,0	42,4
Cost of activities	284413	311206	388198

Table 5: Simulation results for process with revisions

Utilization [%]	**Minimum**	**Average**	**Maximum**
Project manager	7	7,83	8
Analyst	7	7,17	8
Designer	18	19,50	21
Developer	34	36,50	42
Tester	13	14,67	16
Change manager	4	4,83	6
Cost of activities	231 241	254 470	278 581
Project duration	31,15	35,20	38,22

These results clearly shows that overall cost and duration was not increased by revisions but simulation with revisions finish with a little bit lower overall average cost and little bit lower average duration. The main difference is in number of errors in released version of software.

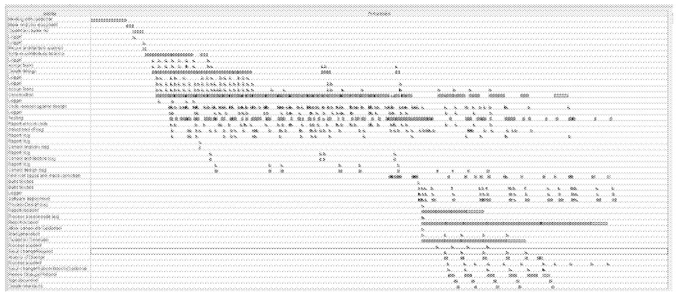


Figure 4: Gantt chart example for simulation process without any enhancements

Second experiment examines the influence of the number of stuffs in roles Designer and Developer (most utilized roles in the process) in the software process without any enhancements. The number of developers and designers was changed in range 1-8 (developers) and 1-4 (designers). Six simulations were performed for each combination and average cost and duration are shown in graph (Figure 5 and Figure 6).
That graphs clearly shows that number of designer in presented version of software process does not have relevant influence to duration of the simulated process, o the other hand number of developer has significant influence to the duration especially in range 1-4. If number of developers is over 4 the influence to the duration of process is not as significant as in range of 1 – 4.

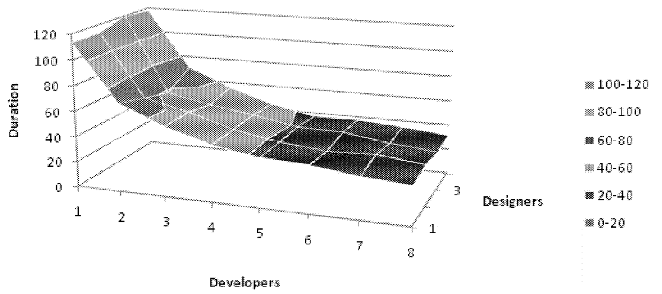


Figure 5: Duration of project in relation to the number of designers and developers

Both parameters, number of designer and developers do not have influence to cost of process, because cost of process is calculated from activities duration in BPStudio. Therefore numbers of designers and developers do not influence the cost because do not influence the amount of work.

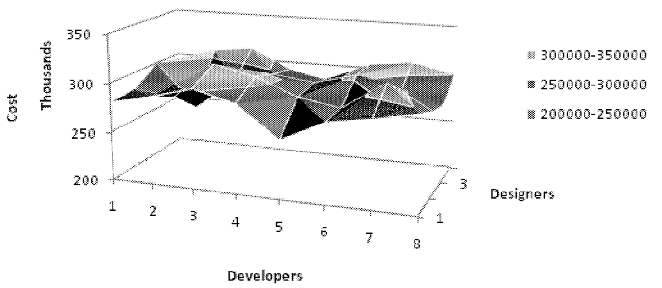


Figure 6: Cost of project in relation to number of designers and developers

Third experiment examines if process with revisions has the same result in case of changed number of developers as the software process without revision. Both processes were simulated with two designers and number of developers was changed in range from 1 to 8.
Results depicted in the graph on the Figure 7 clearly show the same behavior of both processes.

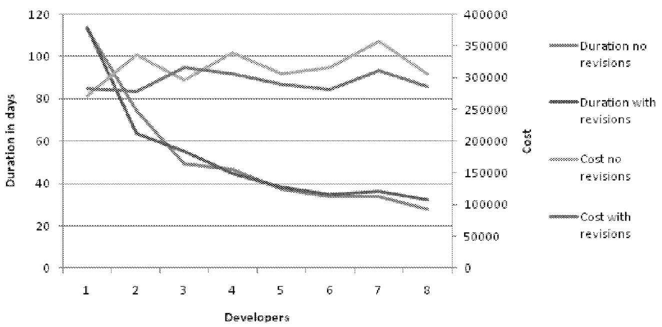


Figure 7: Cost and duration of both processes in relation to the number of developers

USEFULNESS OF THE BPM METHOD FOR SPI

The case study presented in this paper showed how the BPM Method can be used to support basic notions of SPI and

simulate the improvements made in the process. These simulations can also be used to help the management estimate the resources needed for successfully completing the project and meet the deadlines and budget.
We used the BPM Method for simulations because we wanted to prove the usefulness of the Petri Nets approach to software process improvement. We also had access to the method's foundations and source codes of the BP Studio software tool because we continue the research started by Prof. Vondrak. While preparing and conducting the experiments we identified a few advantages and disadvantages of this approach that can be applied to the Petri Nets based simulations in general.

Advantages

- Diagrams that are used for simulation directly describe the sequence of activities performed in the process. Therefore no special effort has to be made to create the models if the process is correctly defined.
- Basic simulation attributes (activity durations, scenario probabilities, etc.) can be directly deduced from the measurement and monitoring of the process.
- Even though the Petri Nets approach is mathematically formal, the resulting model is easy to understand and simulations performed in this model make the understanding even easier. Tokens in the Petri Net are copying the real progress of the process.
- Petri Nets formalism provides the means to verify and enact the modeled process.

Disadvantages

- Models used in simulations have to be accurate and has to be described in detail. These models can be very complex and confused. The diagrams grow in size and number and are hard to manage.
- Many extensions to the basic Petri Nets mechanism have to be applied to model all the properties used in the simulations of business processes.
- Dynamic parameters (such as knowledge, skill and continuous learning and improvement of human resources, error rate, etc.) and related parameters (such as relation of the workers' skills and activity duration, relation of the workers' skill and error rate, etc.) are hard to describe.

CONCLUSION AND FUTURE WORK

This paper described a discrete event modeling and simulation in the domain of software process improvement. We used our tool and method that use Petri-nets as its formal background to model and simulate software process for the purpose of process improvement. This case study shows that this type of simulation can be successfully used for the software process improvement, but have some limitations.
Our future work is concerned with the elimination of some disadvantages of this presented approach. Since we have our own tool for discrete event simulation and modeling, we are able to adjust it for the specific conditions and environments. We would like to use this advantage and implement hybrid model (Donzelli and Iazeolla 2001; Rus, Collofello, and Lakey 1999) that seem to be more promising for the domain of software process improvement than the simple discrete event modeling approach.
Acknowledgements. This research has been supported by the internal grant agency of VSB-TU of Ostrava - SP2011/56 Knowledge approach to the modeling, simulation and visualization of software processes.

REFERENCES

Aalst, W. M. P. van der. 1994. Putting high-level Petri nets to work in industry. *Computers in Industry* 25 (1):45-54.
Aalst, W. M. P. van der. 1998. The Application of Petri nets to Workflow Management. *The Journal of Circuits, Systems and Computers* 8 (1):21-66.
Abdel-Hamid, Tarek, and Stuart Madnick. 1991. *Software Project Dynamics: An Integrated Approach*: Prentice Hall.
Brooks, F. P. 1987. No Silver Bullet - Essence and Accidents of Software Engineering (reprinted form information processing 86, 1986). *Computer* 20 (4):10-19.
Curtis, Bill, Marc I. Kellner, and Jim Over. 1992. Process modeling. *Commun. ACM* 35 (9):75-90.
David, N., J. S. Sichman, and H. Coelho. 2003. Towards an emergence-driven software process for agent-based simulation. In *Multi-Agent-Based Simulation Ii*, edited by J. S. Sichman, F. Bousquet and P. Davidsson. Berlin: Springer-Verlag Berlin.
Donzelli, P., and G. Iazeolla. 2001. Hybrid simulation modelling of the software process. *Journal of Systems and Software* 59 (3):227-235.
Huber, Peter, Kurt Jensen, and Robert Shapiro. 1991. Hierarchies in coloured petri nets. In *Advances in Petri Nets 1990*, edited by G. Rozenberg: Springer Berlin / Heidelberg.
ISO/IEC 15504:2003-2010 – Information technology – Process assessment.
Kuchař, Š., and J. Kožusznik. 2010. BPM Method Eextension for Automatic Process Simulation. In *FUBUTEC*, edited by G. Lencse and L. Muka. Budapest, Hungary: EUROSIS.
Kurt, Jensen. 1997. *Coloured Petri nets: basic concepts, analysis methods and practical use*. 2nd corrected printing ed, *Monographs in Theoretical Computer Science*: Springer-Verlag.
Makinen, T., and T. Varkoi. 2008. Assessment driven process modeling for software process improvement. Paper read at Management of Engineering & Technology, 2008. PICMET 2008. Portland International Conference on, 27-31 July.
Petri, C. A. 1962. Kommunikation mit Automaten. PhD, Institut für instrumentelle Mathematik, Bonn.
Pfleeger, Shari Lawrence, and Joanne M. Atlee. 2009. *Software Engineering: Theory and Practice*: Prentice Hall.
Raffo, D. M., and M. I. Kellner. 2000. Empirical analysis in software process simulation modeling. *Journal of Systems and Software* 53 (1):31-41.
Raffo, D. M., J. V. Vandeville, and R. H. Martin. 1999. Software process simulation to achieve higher CMM

levels. *Journal of Systems and Software* 46 (2-3):163-172.
Raffo, D., and P. Wernick. 2001. Software process simulation modelling. *Journal of Systems and Software* 59 (3):223-225.
Raffo, David Mitchell. 1996. Modeling software processes quantitatively and assessing the impact of potential process changes on process performance. Ph.D., Carnegie Mellon University.
Raffo, David Mitchell, and Wayne Wakeland. 2008. Moving Up the CMMI Capability and Maturity Levels Using Simulation.
Ruiz, M., I. Ramos, and M. Toro. 2004. Using dynamic modeling and simulation to improve the COTS software process. In *Product Focused Software Process Improvement*, edited by F. Bomarius and H. Iida.
Rus, I., J. Collofello, and P. Lakey. 1999. Software process simulation for reliability management. *Journal of Systems and Software* 46 (2-3):173-182.
Scacchi, W. 1999. Experience with software process simulation and modeling. *Journal of Systems and Software* 46 (2-3):183-192.
Štolfa, Svatopluk , Jan Kožusznik, Michal Košinár, Marie Duží, Martina Číhalová, and Ivo Vondrák. 2010. Building Process Definition with Ontology Background. In *CISIM 2010*. Krakow, Poland.
Vergidis, K., A. Tiwari, and B. Majeed. 2008. Business Process Analysis and Optimization: Beyond Reengineering. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 38 (1):69-82.
Vondrák, I. 2000. Business Process Studio, version 3.0. . VŠB – Technical University of Ostrava.
Vondrak, I., R. Szturc, and M. Kruzel. 1999. Company Driven by Process Models. In *European Concurrent Engineering Conference ECEC '99*. Erlangen-Nuremberg, Germany.
Zhang, H., M. Huo, B. Kitchenham, and R. Jeffery. 2006. *Qualitative simulation model for software engineering process*, *2006 Australian Software Engineering Conference, Proceedings*.

AUTHOR BIOGRAPHY

JAN KOŽUSZNIK was born in Havířov, Czech Republic and went to VSB – Technical University of Ostrava, at their Ostrava site, where he studied computer science and obtained his degree in 2002. Afterward he worked for Disnetsoftware a.s. and in 2004 moved back to VŠB – Technical University of Ostrava, where he has been working in the Department of Computer Science.

DAVID JEZEK was born in Ostrava, Czech Republic and went to VSB – Technical University of Ostrava, at their Ostrava site, where he studied computer science and obtained his degree in 2000. Afterward he worked for PVT a.s. Ostrava and in 2001 moved back to VSB – Technical University of Ostrava, where he has been working in the Department of Computer Science.

INFORMATION MANAGEMENT AND SECURITY

INDUSTRIAL INFORMATION SITE MAPPING (I²SM)

Janek Götze[1], Egon Müller[1], Andreas Rutsch[1,2], Christian-Andreas Schumann[2]

[1]Chemnitz University of Technology
Institute of Industrial Management and Factory Systems
Chemnitz, Germany

[2]University of Applied Sciences Zwickau
Institute of Management and Information
D-08056 Zwickau, Germany

E-mail: Christian.Schumann@fh-zwickau.de

KEYWORDS
Augmented Reality, Logistics, Mashup, Smartphone

INITIAL SITUATION

Logistics systems should be high-performing, robust, flexible, and yet versatile measured by the current market requirements and future expectations. The importance of efficient logistics systems has long been recognized as a critical factor of success. Within logistics, the information factor emerged increasingly for being responsible of competitive advantage. Benefiting from the latest manufacturing trends such as postponement, mass customization, and shorter product life cycles, the ubiquitous information extraction, processing and representation is an important capability of state-of-the-art logistics systems.

The significance of this ability is especially evident at more complex logistics structures and processes, where autonomous –self-controlling– software agents are making the complexity of the system manageable, but heavily rely on a high level of detail and timeliness from the system's point of view. This is due to the fact that agents can only respond to something which is recognizable. Inventory, resources and information about material movements can be mentioned here as examples. Another very significant opportunity is the ability to give fast responds if ad-hoc information is needed. Today it is expected that at each time and stage companies are able to give a statement about current stocks, ongoing transportation, handling, storing processes or a resulting delivery status of goods.

I²SM's AIM

The presented project does not focus on predictable processes. Especially ad-hoc processes, which may or may not occur, are being addressed. These processes are not adding any value, but account for high operating cost as they often come along with interruptions of value-added processes. Ad-hoc processes include, for example, search routines, control and monitoring processes or ad-hoc information which needs to be provided immediately to clients or customers. The range of such tasks depends primarily on the timely accessibility of the required information. However, operational practice is characterized by media disruptions and a lack of accessibility.

Figure 1: Exemplarily Use Case

Therefore, the objectives of the joint project are better accessibility, connectivity, and integration of action-relevant information into logistic routines and their mobile and intuitive deployment. The central element of the to-be developed solution is the latest generation of mobile devices. Those devices serve as an information unit, which on-site provides the network location information, sensor data and backend data information in an intuitive way. Initially, the following features will be implemented:

(1) SEARCH feature

- Where is the person in charge, the product, the tool, the batch of material, the order or a specific component?

Executes a search routine with different object or subject categories and dissolves keywords through further detailing to an object ID. Upon this object ID the POSITION INFORMATION SYSTEM (PIS) passes the specific spatial coordinates to the AUGMENTED REALITY (AR) ENGINE. The AR ENGINE projects a label at the right point in the display of the mobile device. By option, more object-related information will be available via a link to BUSINESS INFORMATION SYSTEMS by expanding the displayed label.

(2) VISUALIZATION feature

- What is there? (e.g. clarification of certain bins on a tour of inspection)

At VISUALIZATION all visible objects collected from the PIS are displayed in the image, in contrast to the SEARCH feature, and could be further detailed accordingly. A pre-selection of object categories for the VISUALIZATION is planned to, for instance, display only tools or just the equipment. The jump label in turn contains actionable context information.

(3) Special feature SPACE LOGICS

The two (2) and three (3) dimensional (2d/3d) layout design tool VISTABLE is used to plan and design the layout of manufacturing, assembly and logistics areas. The software design features can determine transport links and security clearances of machines. For the parking of objects, these and other floor areas are restricted for various reasons. Location coordinates of the PIS within restricted areas of the 2d/3d layouts will provoke that messages will be send to the appropriate area manager claiming the potential of a pending collision or safety issue in the restricted place.

(4) Special feature PROCESS LOGICS

The central production control supervises all relevant even distributed processes for the production and assembly. Information includes, for example, the sequence of orders to be processed, the processing sequence at the stations, and the necessary raw material. A target motion profile for raw material or tools could be derived from the production schedule. Again the violation of those movement profiles is being sent to a person in charge.

IMPLEMENTATION STEPS

The research and development project is intended to be implemented by a national consortium of three (3) research institutes and five (5) companies. The project work is done in five (5) stages and nine (9) work packages. The first stage focuses on the definition of system requirements from a technical and operational point of view. The framework of concepts comes in addition to the application development. Stages two (2) to four (4) –analysis, design, and implementation– contain the system development, which is planned iteratively by time. The 5th Phase is the testing and validation of the resulting system. The division into Work Packages (WP) is pragmatically according to the development of the system components as shown below.

WP1: REQUIREMENTS ANALYSIS

WP2: APPLICATION DEVELOPMENT (foremost location-based services)

WP3: BUSINESS INFORMATION SYSTEMS

The BIS form the informational foundation of the I²SM application, aiming to link two (2) corporate information systems: firstly, the ERP system Microsoft Dynamics for master, transaction, and process information and secondly, the container management system IBOX for information such as content, abode, and availability.

WP4: 2D/3D DESIGN SYSTEM

Targeting the integration and expansion of the software VISTABLE, a 2d/3d design system for layout design of factory systems. The resulting 3d-model of the software forms the base for the AR-projection in WP7.

WP5: POSITION INFORMATION SYSTEM

The aim of WP5 is the development of a PIS based on a tracking approach using a transit time measurement in the 2.4 GHz range. Furthermore, alternative indoor localization methods are validated against the background of the system requirements.

WP6: RULES & MASHUP

The Component Rules & Mashup includes two (2) main tasks. First, it sets rules for areas and monitors them. This needs to include the transfer of process logics from the BIS rules into space. Second, it solves information requirements of jump labels from the mobile application to query routines in relation to the BIS, determines the current entries and passes them according to a feature-specific mashup to the AR ENGINE.

WP7: AR ENGINE

The aim of WP7 is the superposition of the camera image with different content (kinds of information). These include the projection of anchor links to contextual information (each label stands for an object and is based on location coordinates represented in the correct position), layer views of the VISTABLE model (e.g. transportation infrastructure), and spatial objects for less precise location-based information.

WP8: INDUSTRIAL SITE MAP

WP8's objective is to develop an application for mobile devices with extensive sensors and state-of-the-art communication interfaces. The application will consist mostly of its frontend, as the user interface. The information processing is done server-side.

WP9: VALIDATION & FIELD TESTING

EXPLOITABILITY

The research and development work on the topic of mobile information retrieval using AUGMENTED REALITY based on spatial coordinates is still at a very early stage of industrial attention. However, in the context of location-based services increasingly business models emerge, which are proved to generate specialized site-specific value-added for users. Although, currently the focus of those developments is mainly focused on the consumer goods market, the development of applications for industrial site applications has already begun.

The joint project takes an innovative approach to improve the mobile information provision. INDUSTRIAL INFORMATION SITE MAPPING (I^2SM) creates a dynamic factory plan containing infrastructure (layout) and scheduled flows (materials, containers, tools, etc.) as well as real-time transaction data. The application of this map via a Smartphone in the world of AUGMENTED REALITy is at low-cost and could be intuitively used safe. It accounts for the printing of lists, phone calls or the navigation through complex menu structures on small handheld devices. Nonetheless, to achieve this objective carries a significant research risk, as only the successful interaction of all components described delivers the potential for this development. For that reason, the developmental work in the context of industrial research, notably the validation of the technology base is necessary.

The range of benefits of a well functioning INDUSTRIAL INFORMATION SITE MAPPING is remarkable. Process transparency and ubiquitous information acquisition will accelerate logistics ad-hoc processes. Today almost disregarded, such processes amount for a large part of daily operations. In addition, the automated, intelligent processing of information in the context of sensor data fusion is in its infancy. Stage one (1) of the project ensures that the development of the project is based on industrial requirements. The alignment of the research project based on industry application potentials, but also the awareness of our industry partners for this technological development, is part of our commitment and project understanding.

The development objective of this joint project is a validated prototype. There are split realisation plans for several components of the system solution. Some of them will pass over into continuous operation at three (3) application partners already during the project. The continuation of those sub-pilots is then performed through traditional realisation plans such as buying, renting or leasing licenses and/or other manufacturing-related services. By working closely with our partners, such as the Federal Logistics Association (BVL) and the Association of German Engineers (VDI), also captures multiplier effects for spreading the results in a scientific and economic sense.

BIOGRAPHIES

JANEK GÖTZE, born 1974 in Karl-Marx-Stadt has a Joint Masters Degree in Industrial Engineering with Business Studies from the Chemnitz University of Technology. His majors have been Accounting, Logistics and Quality Management. Currently he does his doctorate at the Chemnitz University of Technology.

EGON MÜLLER was born in 1952, studied Mechanical Engineering and earned a doctorate at the Engineering School of Zwickau. Between 1978 and 1992, he held several positions in the automotive industry and the Engineering School of Zwickau. In 1992 he became full professor for Factory Planning at the University of Applied Sciences Zwickau. Since 2002, he is holding a professorship for Factory Planning and Factory Management at the Chemnitz University of Technology.

ANDREAS RUTSCH was born in 1980 (Plauen, Germany). He earns a Master of Science in Systems Engineering from the Chemnitz University of Technology where he still does his doctorate. Currently he is a member of research staff at the Universities of Leipzig and Zwickau.

CHRISTIAN-ANDREAS SCHUMANN, born 1957 in Chemnitz (Germany), studied Industrial Engineering at the Chemnitz University of Technology, doing his first doctor's degree in 1984 and second doctor's degree in 1987. He became associate professor for Plant Planning and Information Processes at the Chemnitz University of Technology in 1988. In 1994, Christian-Andreas Schumann became full professor for Business and Engineering Information Systems at the University of Applied Sciences Zwickau.

FRACTAL GEOMETRY AND INFORMATION MANAGEMENT

Hamid Reza Dezfoulian
Department of Industrial Engineering
Bu-Ali Sina University
Hamadan, Iran
E-mail: hrdezfoolian@yahoo.com

Mehrdad Garousi
No. 153, Second Floor, Block #14
Maskan Apartments, Kashani Ave
Hamadan, 6517639891, Iran
E-mail: mehrdad_fractal@yahoo.com

Mehdi Garousi
Industrial Engineering Department
Tarbiat Modares University
Tehran, Iran
E-mail: mehdigarousi@gmail.com

KEYWORDS

Fractal Geometry, Complexity, Information Management, Market Analysis.

ABSTRACT

Fractal geometry, as one of the most important and significant modern mathematical discoveries, through the last half century has caused a vast variety of evolutions in diverse science species. This paper aims to survey the role of fractal patterns, shrinking complex mathematical calculations, geometrically in arrangement of groups of relevant information which are needed to be easy to manage, understand or conclude from.

INTRODUCTION

Fractal geometry provides a powerful approach for the quantitative description of complex, highly irregular and random, i.e. disordered systems. Moreover, it can be used to describe the processes leading to the formation of such systems and their physical behavior. Fractal geometry relates to structures that cannot be described by Euclidean whole number dimensions of 1, 2, or 3, but instead they have fractional dimensions (Senesi and Wilkinson 2008).

Fractals and chaos are closely intertwined and often occur together. In theory, virtually anything that happens over time could be chaotic (Midgley 2001). Virtually every branch of the sciences, engineering, economics, and related fields now discusses or refers to chaos (Williams 1997).

Unlike previous economic revolutions, in which the input was energy and matter, the present revolution in economic productivity is based on information and knowledge (Freeman 1988; Caraça 1993; Fonseca 2002).

In a modern society, surrounded with information and telecommunication, management of bunches of information in order to have control over the current and future of a system is extremely vital. In each unit like a society, market, factory, firm or …, there are diverse kinds of information belonging to different areas and levels with different degrees of importance; conclusion through which is a complex phenomenon. Our information of the unit most of the time is not symmetric nor does act linearly and Euclidean mathematics can not arrange it perfectly to a certain goal or need. On the contrary, Fractal geometry is the geometry of the Demiurge. Unlike Euclidean geometry, it thrives on roughness and asymmetry (Peters 1994).

MANAGEMENT AND FRACTAL IMAGES

Basically, the most important thing in management of a group of information sources, in every system, is their arrangement. Information architecture describes the structure of a system, i.e., the way information is grouped, navigation methods and terminology used within a system. Flexible and effective visible information architecture enables people to step logically through a system confident that they are getting closer to the information they require (Link 2008; Kirikova 2009).

An arrangement always can evoke a graphic or visual depiction. Through the history picturing has been the easiest and most handy way of having control over complexity. Because, comparison of bunches of diverse information is actually difficult numerically. The mind of human is more convenient with images than numerical information, in spite of the computers that it does not differ to them.

Graphics is wonderful for matching models with reality. When a chance mechanism agrees with the data from some analytic view point but simulations of the model do not look at all real, the analytic agreement should be suspect. A formula can relate to only a small aspect of the relationship between model and reality, while the eye has enormous powers of integration and discrimination. In addition, graphics helps find new uses for existing models (Mandelbrot 1983).

Fractal pictures are the accurately depicted outcomes of fractal behaviors which contain all properties of fractals together as a whole visually. Mandelbrot in his renowned book *The Fractal Geometry of Nature* (1983) says: in particular, in the theory of fractals “to see is to believe”.

Self-similarity is an underlying property of fractals. Based on the self-similarity between part and whole, fractal theory establishes the global property from the part, along the direction from microscopic to macroscopic. As the whole is made up by the part, we can understand the whole through part (Cui and Yang 2009).

This strategy can be also used in terms of arrangement of information. Massive information of constructively complex systems and manufactures can be classified as boxes of information regarding different levels of branchings in correspondence with other branches or subsets of the whole system. It is while traditional charting and modeling of information work linearly and can only present properties and information without depicting the important interactive relationship between them. However, fractals settle all kinds of information interactively in defined neighborhoods in a certain magnifiable figure with certain levellings and branchings. Fractal images due to their self-similarity, as can be evidently seen in most of standard fractals, start from a

bounded basic shape which contains several other similar parts and components inside or outside of it in a certain boundary. With this property, all information of a system is placed inside a certain perimeter inside or outside of the body of the initiator (the first shape by which our fractal starts to be shaped) in different ways. What is important is the generator (the pattern whereby our iteration starts to be applied) that defines how we should place the information inside or outside the basic form. Generator, in each iteration, is in fact our arrangement of information in each level of our system. For example, if our system compounds of *n* main smaller components, we will define the generator the way that we have *n* smaller copies of the initiator inside or outside of it, as the first iteration. This current continues according to the branching of the system to the latest and lowest levels. Standard fractals are only some instructions which act like tutors for us and each project requires its own fractal pattern with specific properties.
The interesting part of this kind of information settlement is the way of analyzing through which. Because every small component has placed inside a higher branch, it has a weight inside not only its one stage higher branch, but also all higher branches which it is a part of them and at last, the whole system. The other very useful property is the ability of depicting the balances of weights. The weights of subdivisions can be shown as numbers or colors or on the other hand by the size of the generator in that iteration. In other words, if a subset has a heavier weight will have a larger number, warmer color or more comprehensive area. If we use of larger area, our figure will present a more interactive face. If the heavier branch owns a larger area, smaller areas will be divided between other light branches and the visual balance of the shape collapses. Thus, where the massive area is centered, one essential property or the average of corps of properties has a higher degree of importance qualitatively or quantitatively, negatively or positively.

INSTANCES

Although basic standard fractals might not be adapted to most of everyday activities, being the groundwork of fractal geometry, play a significant role in inspiring new ideas well adapted to different situations. As it can be seen, in figure 1 we have presented an altered example according to the clustering of the standard T-Square fractal. The central black square is our main system which has four offshoots. Like a shoes manufacturer which presents its products through four main shops A, B, C and D.
What is of priority to this manufacture is the mass of the sale. So it has considered ten colors between mere red and mere yellow corresponding ten distinct percentages of selling the products sent to their storages from the manufacture in the month.
Mere red is the color representation of under 10 percent and yellow is an representation of between 90 and 100 percents. Here, each square is a shop center with some offshoots. The color of each square is the sum of what it takes out of its storage, including the sale to both customers and offshoots which shows its average activity.

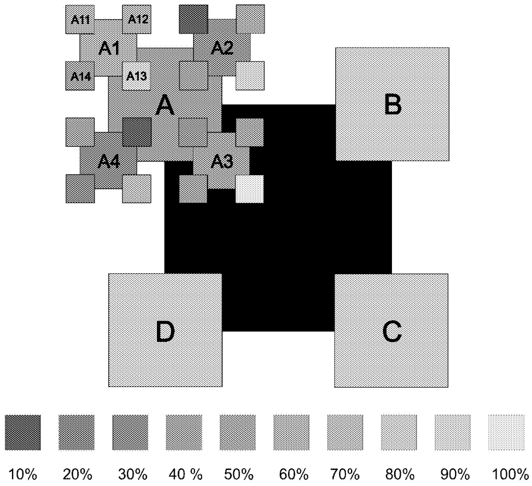


Figure 1: Depicting the sale percentage of branch A of a manufacture according to geometrical shape of T-Square fractal strictly.

Thus, a hasty look at the pictorial fractal analysis of the A branch of this factory clarify that it has an average productivity between 30% and 40%. In such a model due to its fractal branching property the quality or quantity of each unit is an average of its own work and its offshoots. So, looking at the color of for example square A, we are looking at the analysis result of surveying work and productivity of all its branches and subbranches to the smallest ones. We can change the number of branches or subsets according to our needs and magnify as much as possible to see smallest subbranches at the lowest levels of our market. Depicting market analysis of all other main branches B, C and D with different numbers of subbranches reveals the easiness and comfy of management according fractal geometry. A closer instance of previous model to reality can be seen in figure 2.

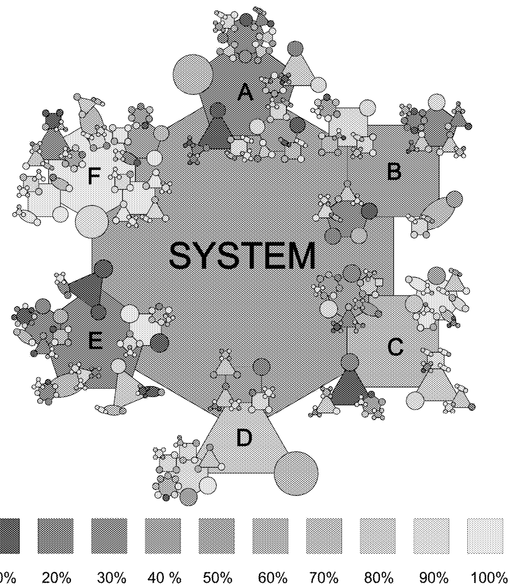


Figure 2: A more complex T-Square pattern of activities of different branches of a central shop in distinct districts. The more red the color is, the less sale that certain branch has had.

This complex fractal pattern displays the situation of different components or activities of subagents of a central system which here is assumed as a shop and its several branches in six different states (A to F). It is evident that this image is not constructed based on a certain generator. Initially, corresponding to the number of A to F main branches of the shop, the initiator is a hexagon. However, in next iteration as second level of offshoots according to the number of central subbranches in each state, different polygons have been used and it continues similarly to the lowest levels of shops which have no further subsets (circles). Therefore, in accordance to our needs we can exploit different initiators and generators in constructing a fractal pattern in order to have a fractal analysis of our activity. In the above company if the sale information of each subbranch is reflected to the upper branch in certain periods, the matrix will be automatically in change and up to date. Therefore to collect such detailed groups of information does not need any further employee, time or energy. By means of providing a simple interactive software on the web to which every branches of the market has reach, everything could be updated instantly, without even any need to an operator.

The main two essential aspects of each analysis are, collecting information and arranging them properly to analyze according to their comparison. All of these are automatically done in above feature and what the analyzer need do is to have a look at the image and compare the colors with each other. For example is clearly evident that the sale of the company in state F is the best and in states A and E the worst. Also, it can be analyzed that in a sate like E with a very bad sale, one subbranch and most of its subsets have had convincing sales. A complete matrix is one with different factors, features and interfaces relating to the needs of the user. For example each pattern may contain different faces of information. By preparing different kinds of information, each time, distinct analysis of them can be displayed. For example one feature with the percentage of sale, one with the percentage of damaged stuff, one with the number of permanent customers, one with the differences of sales and costs and so on. Different factors must be analyzed in order to have a clear prospect of the productivity of a set, all of which would be easily accessible with fractal system of information collection and featuring.

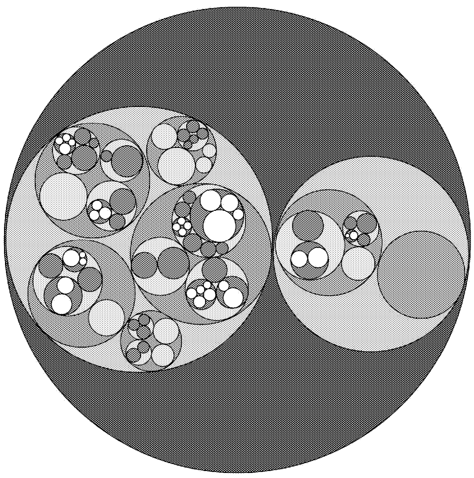


Figure 3: A circular fractal analyzing matrix basing Apollonian fractal.

There are diverse kinds of fractals with different geometrical figure which can be exploited basing the needs of the set and the kind of demanded analysis. Other very useful fractals in this order are different Apollonian sets. Apollonian feature because settle everything in form of circles inside of them tangibly, provide a better display of the power or comprehensive of the activities in comparison to each other. Because, enlarging a circle pressures others to get smaller. These are more appropriate for analysis of indicators which work complementarily. Below a derivation of Apollonian fractal can be seen in which every component has a complementary relationship at least with two of its branch mates (in a dual group both circles are inevitably tangent).

CONCLUSION

The most essential goal of management is to be able to control the future of the system according to its past and current. Therefore, between numerous ways of arrangement of information of past and current, one is the best that by displaying a connection between the past and now, and different consequents of the system, provides a clearer satisfaction of its future. Fractal visual analysis due to their interconnectivity, inner complexity, and easiness to fast analysis can be a suitable answer to this need of management. However, it needs more study and investigations in parallel in mathematics, management and computer sciences.

REFERENCES

Caraça, J.M. 1993. *Do Saber ao Fazer: porquê organizar a ciência*. Gradiva, Lisbon.

Fonseca, J. 2002. *Complexity and Innovation in Organizations*. Taylor & Francis, NY, 6-72.

Freeman, C. 1988. "Introduction". In *Technical Change and Economic Theory*, G. Dosi, C. Freeman, R. Nelson, G. Silverberg, and L. Soete (Eds.). Pinter Publishers, London, 1–8.

Kirikova, M. 2009. "Towards Flexible Information Architecture for Fractal Information Systems". In *eKNOW 2009: International Conference on Information, Process, and Knowledge Management*, A. Kusiak and S.G. Lee (Eds.). Cancun, Mexico, 135-140.

Link, J.A. 2008. *How to Develop and Implement Your IT Strategy*. Matador, UK, 150.

Mandelbrot, B.B. 1983. *The Fractal Geometry of Nature*. W. H. Freeman, NY, 21, 22.

Midgley, M. 2001. *SCIENCE AND POETRY*. Routledge, NY, 41.

Senesi, N. and K.J. Wilkinson. 2008. "Introduction to the Study of Environmental Fractals". In *Biophysical Chemistry of Fractal Structures and Processes in Environmental Systems,* N. Senesi and K.J. Wilkinson (Eds.). A John Wiley and Sons, Ltd, UK, 1-10.

Peters, E.E. 1994. *Fractal market analysis: applying chaos theory to investment and economics*. John Wiley & Sons Inc., USA, 4.

Williams, G.P. 1997. *Chaos Theory Tamed.* Taylor & Francis Ltd, UK, xi.

Cui H. and L. Yang. 2009. "Short-Term Electricity Price Forecast Based on Improved Fractal Theory". In *ICCET 2009: International Conference on Computer Engineering and Technology (Singapore, January 22 - 24)*, J. Zhou and X. Zhou (Eds.). Volume I, 347-351.

ENHANCING WORKFLOW SECURITY FOR LARGE SCALE DISTRIBUTED SYSTEMS

Florin Pop, Alexandru Corneliu Drenea, and Valentin Cristea

University POLITEHNICA of Bucharest, Faculty of Automatic Control and Computers, Department of Computer Science
Splaiul Independentei, 313, Bucharest 060042, Romania
E-mails: florin.pop@cs.pub.ro, alexandru.drenea@cti.pub.ro, valentin.cristea@cs.pub.ro

KEYWORDS
Workflows, Distributed Systems, Security, Digital Signature, QoS.

ABSTRACT

Information security is a very delicate matter because more and more information is sent electronically and some of it requires special handling and quality standards. The strength and security offered by system using PKI is proven by the multiple examples of existing solutions that are built the same way and are around us for some time. Paper documents that need to be filled out and signed are always around us and often can cause problems and delays when poorly managed. With its characteristics, the proposed solution described in this paper hopes to eliminate all the inconvenient of the paper documents helped by the document library and workflows, while keeping the security part, now represented by hand signatures with the implementation of the digital signatures. The novelty of proposed solution is represented by creation of an offline application capable of using digital certificates to digitally sign or encrypt digital document. We studied the application requirement and came up with a modular and scalable architecture allowing for a quick and easy future development. We present the deploying of online portal for managing the document libraries and workflows in large scale distributed environment.

1. INTRODUCTION

The strength and security offered by system using PKI is proven by the multiple examples of existing solutions that are built the same way and are around us for some time. When reading an email you can be very sure that you are reading exactly what the other person wrote before hitting the send button, or you can be sure that no other person but you has read the information, features enabled by the PKI products, the user digital certificates. When accessing an online-banking account to operate a payment or money transfer or simply when buying anything on the internet, you are sure that your most sensitive credit card information is secure, thanks to the PKI enabled SSL certificates (Biddle, 2009). When enabling a VPN connection to a remote place, all the network traffic flow is encrypted using digital certificates (Rajaravivarma, 2009). All the above examples stand as proof that the PKI is a viable option for our design.

For an enterprise, implementing the solution enables better management for internal documents. Again, dealing with digital versions of a document is much desirable than handling paper documents that are very prone to being lost, damaged and even forged. Digital signatures evolved to a point where it is easier to forge a hand signature than forging a digital signature produced even by a medium security PKI environment in the same time eliminating the human factor in deciding if the signature is indeed authentic.

The project development consists in the first step in the creation of an offline application capable of using digital certificates to digitally sign or encrypt digital document. We studied the application requirement and came up with a modular and scalable architecture allowing for a quick and easy future development. The last part consists in deploying an online portal for managing the document libraries and workflows. We decided to use a solid existing framework for the library – a Microsoft SharePoint portal – with a good workflow support both integrated and extensible through programming. We also built example workflows for a couple of test scenarios (Microsoft Corporation, 2001).

The main benefit it brings to the client is that it automates the signing and approval process to any kind of document it uses inside or outside the company. The signature system allows signing on multiple levels (counter-signatures) and multiple signatures per level (co-signatures) for perfectly mimicking a plain document. Also common workflows for every client can be defined in the system when deploying and the users should be able to use them out of the box.

The papers is structured as follow: Section 2 presents some related work in the field, then Section 3 covers he requirements and architecture, Section 4 is about the last main part of the proposed solution and talks about the digital library implementation and document workflows. Section 5 is dedicated to revealing conclusions regarding the use of digital certificates to enable online document handling trough workflows.

2. RELATED WORK

Digital signatures are a “product” of digital certificates (Chang, 2005). They enable message integrity and authentication services. It is important to mention that a digital signature does not offer message encryption. They can be used to sign any type of document and several other entities. In our days they are most commonly used for: email, documents, application setup kits (jar and msi files). At the base of digital signatures stay two common algorithm types: hashing and encrypting.

The most common PKI implementations use SHA1 (Lee, 2006) for the hash and RSA for the encrypting but others like MD5 or Diffie-Hellman can be found in some signatures (Sasaki, 2008).

Standard format for digital signatures is PKCS#7 was developed by RSA (Davis, 2008). It is an open format enabling anybody to develop applications for creating and validating digital signatures. PKCS#7 formats can be used to transport any kind of information no matter its format, destination or contents. It uses ASN.1 (Brad, 1999) encoding, as digital certificates, in order to enable applications to easily process it. The signed file encoded in PKCS#7 will contain extra information along with the required information data and signature bytes, such as: digital certificates used to sign the data and hashing algorithm used to sign.

A very important feature of the format is that it allows for multiple signatures attached linked in the output file, divided into two options: co-signatures that sit one next to another, signing the same information, and counter-signatures. A small note on counter-signatures, they do not actually sign the original information rather they sign the signature byes of the signature they counter-sign. The reasoning for that is that a counter-signature is supposed to approve a signature and not the data. The data is still protected because corrupting the original data will corrupt the signatures directly applied to it thus corrupting all counter-signatures.

Specialized applications dealing with documents and messages started including their own internal support for digital signatures and do not rely on PKCS#7 only. Such applications, like Adobe Reader for PDF files, Microsoft Office for DOC files, allow you to sign a document without the need of an external application that creates the signature file. The main advantage of this approach is that the files are still readable trough their original application without the need of the external application to verify and unpack them from their format (Phillip, 2010).

CryptoBOT e-Workflow is a commercial application part of a bigger solution, also including a signing application (e-Crypt) and a Certification Authority server. It also works with other externally issued certificates. The e-Workflow solution is composed from two main components (Hongbin, 2008): The server component is simply an application running on a designated computer that listens for workflow requests and forward alerts to users that need to take action in a workflow. Also interacts with the database for logging. The client part is a more complex application where the user logs in and can see a list of documents and running workflows. The application seems complicated and difficult to use and manage and does not have an attractive look and feel (Pop, 2011).

3. REQUIREMENTS AND ARCHITECTURE

Following is the main, requirements plan for the application:

1. *Sign a document*: having a local document, the user must be able to quickly sign it using he's own digital certificate. Should be able to create multiple types of signatures: simple signature, co-signature or counter-signature. Also the user must be able to sign the same document multiple times.
2. *Verify a signed document*: having a local signed document, the user must be able to verify the existing signatures. There should be two levels of signature checking one for each verification step: the signature integrity check and the certificate validity check. Should be able to check the entire signature tree in both a quick visual representation of the signatures and detailed information about them.
3. *Encrypt a document*: having a local plain or signed document, the user must be able to encrypt the document. The application must allow adding multiple recipients for the encrypted message.
4. *Decrypt a document*: having local encrypted document, the user must be able to decrypt the document, using his private key as long as he is listed as a recipient in the encrypted document.
5. *Create output documents* in a public format so they can be used in places that do not share the same solution and application for managing the documents as well as being able to read digitally signed or encrypted documents originating from outside the solution.
6. *Secure* the passwords that are entered in the application

The application architecture is highly modular, allowing for easy extensibility or future improvements. It was designed in accordance with recommended design patters and following strict design principles as described below.

Classes are always strict representation of concepts. They are built as an abstraction over that concept and their methods do not break the abstraction thus providing with high cohesion classes. All implementation specific classes are designed with an interface contract for low coupling between classes. Coding to an interface and not to an implementation is an old design principle enabling the use of dependency injection. This greatly improves testability giving the tester the ability to separate the currently tested class from all dependencies by loading dummy objects.

For some classes, generally for those at the service Layer we implemented the Singleton pattern along with the Factory pattern allowing me to create and use similar instances of the class.

Business Logic Layer. This is the core layer of the application defining and implementing all the actual cryptographic operations. As stated above all classes have corresponding interfaces to sustain and improve extensibility. The layer is physically implemented in a separate assembly, with references only to the services layer and using only interfaces to access any needed dependencies. This proves to be an efficient and modular design allowing me to use the library in any other application needing signature or encryption services.

An intensely used object in the BL (Business Layer) is the X509Certificate2 object, provided by the .NET framework. This offers a very good access to any kind of certificates, providing easy access to any needed properties of any certificate.

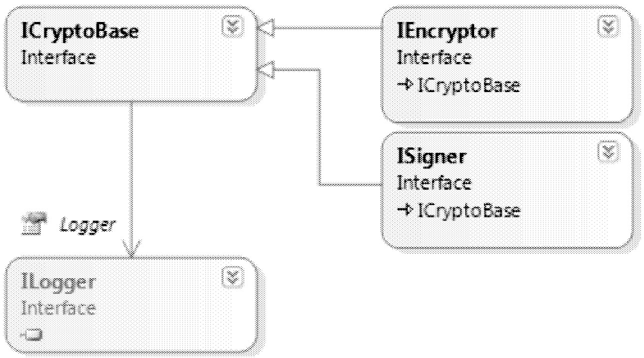


Figure 1. Business Layer interface design

Services. Responsible with providing different functionalities to entire application. The most common elements usually found in a service layer are the data access layer object, the logger and the object builder. Since our application does not use any databases and we don't have a very big diversity of objects, the only common service we implemented was the Logger service. Along with it I also included a static helper class for certificate manipulation.

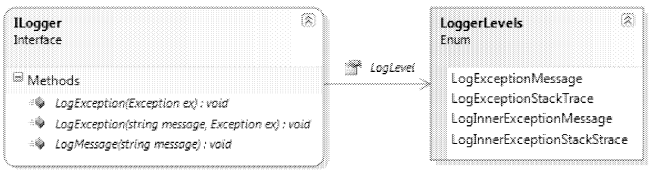


Figure 2. ILogger interface diagram

- *ILogger*: Provides all the logging functionalities throughout the application. It is referenced in the lowest units of each layer of the application for high availability: in the BL it's referenced from the ICryptoBase interface and in the Presentation Layer is referenced in the IFormParent interface. The only property available is used to set the level of details that will be included in a log message when an error occurs. The functions exposed in the interface are self-explanatory and don't really need further description.
- *X509CertificateHelper*: The class offers only three functions for the moment but can be extended if needed. Most important function is the one that searches for certificates installed on the local computer, under current user stores. It can return all certificates found or only the ones including a private key, as a simple array of certificates. The other important function creates a ListItem based on a certificate. The need for this function appears in the presentation layer where we often present the certificates in a List format and it is useful to be able to get a ListItem quickly based on a certificate. Last but not less important is a function that creates a friendly string out of the certificate subject name to be displayed in common places such as signature caption or ListItem text.

Presentation. Responsible with displaying all the information to the user in the best possible manner. The presentation layer (or User Interface Layer – UI) is located at the highest functional level, closest to the user. Because the nature application needs the user to pass through multiple steps in order to perform the desired action, I chose to implement the Graphical User Interface (GUI) with a masked MDI format. As with the BL, all is done through interfaces, helping us to extend if needed. There are two main interfaces, IFormParent and IFormContent which are coupled only through a property in the last one, as shown in the diagram below:

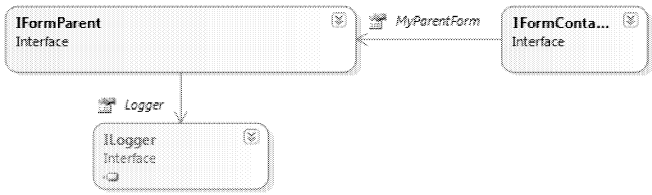


Figure 3. Presentation Layer diagram

4. ENHANCING WORKFLOW SECURITY

In this section we will present the solution proposed and give details about its implementation. SharePoint technology is a vast domain and only how to customize and use it could make a whole book and this is not the purpose for the current paper. As a result to that I will just cover minimal details about SharePoint services representing key points in enabling a company's digital library.

The signature is the end-result for all the trouble in designing and implementing a PKI in the first place. Creating a signature is actually a very simple process. The main ingredients when creating a signature are the data that needs to be signed and a digital certificate with an available private key. The digital certificate should contain all information about what algorithms to be used so the application performing the actual signing will know and use them.

To create a signature, take the input data bytes and create a message digest (hash); encrypt the hash with the private key and obtain a byte array representing the signature byte. All that remains to be done is pack them in a signature file format, like PKCS#7, and release the signed file.

Workflows are the last ingredient to complete the online document library we designed. A workflow can be compared with an application that is performing actions at specific moments. Workflows can be associated with almost any action that can happen in the SharePoint environment. For starters, in our solution we will use the most common action:

assigning a new task when an event is triggered (such as a document being uploaded). Other actions include sending an email.

A workflow is always associated to a list in the web-site and a workflow instance is associated to an item from that list. However, it can access items from any list for example a workflow can read the user that created the item that started the workflow, based on this variable it can search another list for a set of users that are associated with him with some criteria and assign tasks for each user in the resulting set. The possibilities are countless, making workflows a very important feature of SharePoint.

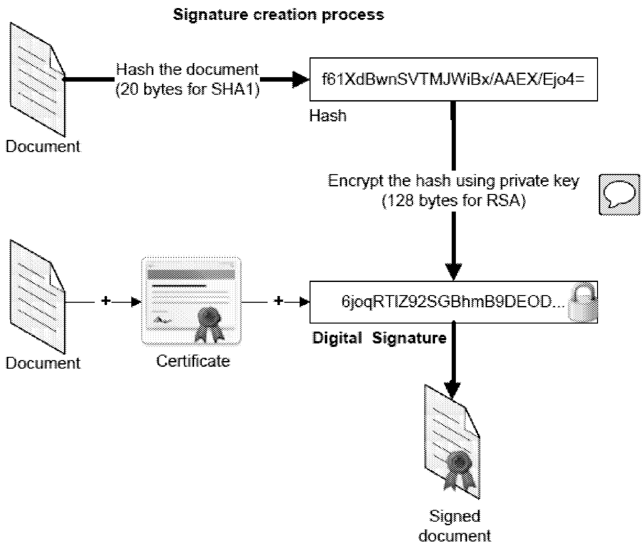


Figure 4. Digital Signing process

We will present the steps required to build the most basic workflow that can be further extended when needed. We start by opening SharePoint Designer and then open the web site containing the library we want to create the workflow for. We add a new workflow instance and select the required information from the wizard that appears to guide the process. After selecting a name for the workflow and the library it's associated with, we get to a wizard step similar to that presented in the figure below. There, I've added an action that will assign an instace of "new task" to user when the workflow is started. Additional actions can be added to the workflow, the complete list of default actions being also presented in the image below. As we can see, creating a near standard workflow is an easy task, the harder task being desingnig it, as we will present in the following section.

A workflow is at its base a file defined with the Extensible Object Markup Language (XOML) that promises to provide a common syntax for the various types of workflows used in Windows applications programming. The file is compiled and processed on the fly when required by the Workflow Foundation Runtime (Matheus, 2005).

Next, we will present a real-life scenario, where we could successfully apply our solution. There are two main categories of clients as identified earlier on in the paper: the medium-large company or a public institution. We will present a scenario for each of these, in more detail and a few others just as example.

There are multiple standardized scenarios involving multiple users and at least one type of document that are running daily throughout the enterprise. To note just a few of them: Expenditure request from a user to his direct superior requesting permission to use company money to buy something.

Vacation request. Let's consider an employee that needs a day off and check the steps that he needs to take to perform the request.

- Prints a document, signs it and takes it to the manager.
- The manager approves it (or not) and signs it.
- The employee or the manager then takes the signed document to the HR department where it will be processed and archived.
- The manager or employee has to manage responsibility delegation (redistribute tasks to other employees during the absence) and then inform the ones involved.

The whole process can take a few minutes in a small sized company but for larger companies it can take days. The situation is more complicated when several managers need to approve a request (the case of a programmer whose request should be approved by both team leader and direct manager). Here's the same example using out proposed solution. The employee logs in to the SharePoint portal and downloads a template of the absence form. After filling it with his information, he signs it with the application and uploads the doc back to the portal and the workflow is started. The manager responsible with approving the request receives a mail informing him of a new task that he needs to perform. He now downloads the file, and approves it by applying a counter signature to the employee signature. When uploading back the file, another workflow starts sending notification emails to both the HR department and the employee delegated with the absentee responsibilities. The comparison of different scenarios is presented in the figure bellow. We can observe that the electronic solution offers a good results then real life. The distributed systems offers the support of enhancing security and time execution for workflows.

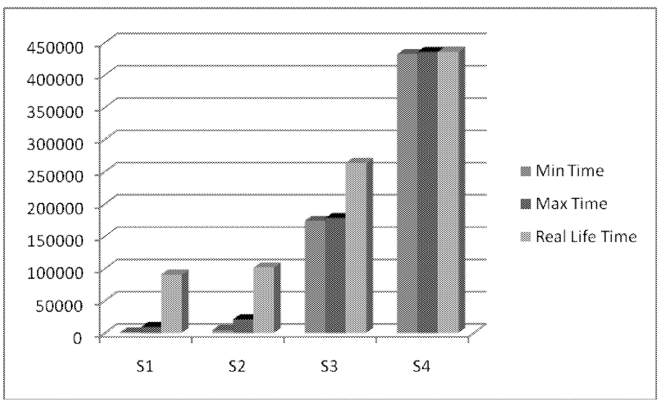


Figure 5. Time comparison

Public institutions have way more examples when it comes to bureaucracy. Some of them are very complex and they require very much work in designing the specific workflows, especially because they might include cross-institution handling such as **opening a new company,** where the applicant has to pass through multiple stages and multiple institutions. Other is simpler and may be easily implemented with our solution. From my experience I found out that yearly fiscal reports such as **yearly income report** can be very annoying because it almost always involves trips and queues at the financial administration.

5. CONCLUSIONS

We propose in this paper a solution for digital signing for documents involved in workflows. The workflows was considered in a large scale distributed environments for possible optimization of key distribution and interoperability.

The novelty of proposed solution is represented by creation of an offline application capable of using digital certificates to digitally sign or encrypt digital document. We studied the application requirement and came up with a modular and scalable architecture allowing for a quick and easy future development for future business technologies. We present the deploying of online portal for managing the document libraries and workflows in large scale distributed environment.

ACKNOWLEDGMENTS

The research presented in this paper is supported by national project: "SORMSYS - Resource Management Optimization in Self-Organizing Large Scale Distributes Systems", Contract No. 5/28.07.2010, Project CNCSIS-PN-II-RU-PD ID: 201. The work has been co-funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the Financial Agreement POSDRU/89/1.5/S/62557.

REFERENCES

Robert Biddle, P. C. van Oorschot, Andrew S. Patrick, Jennifer Sobey, and Tara Whalen. 2009. Browser interfaces and extended validation SSL certificates: an empirical study. In *Proceedings of the 2009 ACM workshop on Cloud computing security* (CCSW '09). ACM, New York, NY, USA, 19-30.

Veeramuthu Rajaravivarma. 2009. Open source virtual private network experience in classroom.*J. Comput. Small Coll.* 24, 3 (January 2009), 85-92.

Microsoft Corporation. 2001. *Microsoft® Sharepoint^{Tm} Portal Server 2001 Resource Kit*. Microsoft Press, Redmond, WA, USA.

Ya-Fen Chang and Chin-Chen Chang. 2005. Schemes for Digital Gift Certificates with Low Comp. Complexity. *Informatica* 16, 4 (December 2005), 503-518.

Mun-Kyu Lee, Jung Ki Min, Seok Hun Kang, Sang-Hwa Chung, Howon Kim, and Dong Kyue Kim. 2006. Efficient implementation of pseudorandom functions for electronic seal protection protocols. In *Proceedings of the 7th international conference on Information security applications: PartI* (WISA'06), Jae-Kwang Lee, Okyeon Yi, and Moti Yung (Eds.). Springer-Verlag, Berlin, Heidelberg, 173-186.

Yu Sasaki, Lei Wang, Kazuo Ohta, and Noboru Kunihiro. 2008. Security of MD5 challenge and response: extension of APOP password recovery attack. In *Proceedings of the 2008 The Cryptopgraphers' Track at the RSA conference on Topics in cryptology* (CT-RSA'08), Tal Malkin (Ed.). Springer-Verlag, Berlin, Heidelberg, 1-18.

Don Davis. 2001. Defective Sign & Encrypt in S/MIME, PKCS\#7, MOSS, PEM, PGP, and XML. In*Proceedings of the General Track: 2002 USENIX Annual Technical Conference*, Yoonho Park (Ed.). USENIX Association, Berkeley, CA, USA, 65-78.

Brad Williamson and Craig Farrell. 1999. Independent active program representation using ASN. 1. *SIGCOMM Comput. Commun. Rev.* 29, 2 (April 1999), 69-88.

Florin Pop, Ciprian Dobre, Decebal Popescu, Vlad Ciobanu, and Valentin Cristea. 2010. Digital certificate management for document workflows in E-government services. In *Proceedings of the 9th IFIP WG 8.5 international conference on Electronic government* (EGOV'10), Maria A. Wimmer, Jean-Loup Chappelet, Marijn Janssen, and Hans J. Scholl (Eds.). Springer-Verlag, Berlin, Heidelberg, 363-374.

Hongbin Sun and Yongsheng Ding. 2008. An Adaptable Method of E-Workflow Composition Based on Distributed Awareness. In *Proceedings of the 11th Asia-Pacific Symposium on Network Operations and Management: Challenges for Next Generation Network Operations and Service Management* (APNOMS '08), Yan Ma, Deokjai Choi, and Shingo Ata (Eds.). Springer-Verlag, Berlin, Heidelberg, 503-506.

Phillip J. Brooke, Richard F. Paige, and Christopher Power. 2010. Document-centric XML workflows with fragment digital signatures. *Softw. Pract. Exper.* 40, 8 (July 2010), 655-672.

Andreas Matheus. 2005. How to Declare Access Control Policies for XML Structured Information Objects using OASIS' eXtensible Access Control Markup Language (XACML). In *Proceedings of the Proceedings of the 38th Annual Hawaii International Conference on System Sciences - Volume 07* (HICSS '05), Vol. 7. IEEE Computer Society, Washington, DC, USA, 168.1-.

BIOGRAPHY

FLORIN POP PhD, is lecturer with the Computer Science and Engineering Department of the University Politehnica of Bucharest. His research interests are oriented to: Contextualized Services in Distributes Systems, Predictive Systems and Intelligent Optimization Thecniques. He received his PhD in Computer Science in 2008 with "Magna cum laudae" distinction. He has received an IBM PhD Assistantship in 2006 (top ranked 1st in CEMA out from 17 awarded students) and a PhD Excellency grant from Oracle in 2006-2008.

ELECTRONIC SERVICE FOR DISTRIBUTED INTERACTION IN E-GOVERNMENT ENVIRONMENTS

Nirvana Popescu, Decebal Popescu, Vlad Ciobanu, and Florin Pop

University POLITEHNICA of Bucharest, Faculty of Automatic Control and Computers, Department of Computer Science
Spl. Independentei, 313, Bucharest 060042, Romania
E-mails:{nirvana.popescu, decebal.popescu, florin.pop}@cs.pub.ro, vlad.ciobanu@cti.pub.ro

KEYWORDS
Software communication processor, e-Government, Security, Digital Signature.

ABSTRACT

In the context of European Service Directive, the development of web services for government needs is a great challenge. Companies need to periodically report their fiscal, social and statistics information to various institutions. Almost all of these reports are required by legislation. We proposed and device a safe and flexible solution for communication between companies and public agencies. The proposed solution is a data flow processor that supposes three steps. First, the Public Institution sends to the Central Authority declarations model and the completing methodology. In the second step, the Central Authority implements the model in the programmatic language. Finally, the Central Authority starts the configuration procedure for all companies' processors. The retrieval of primary and transactional data can be realized in automatic manner. The processor was designed and implemented based on Microsoft technologies involving filling out of forms designed in InfoPath and, than parts of the forms can be filled out automatically. The process implies that the form data are digital signed. The proposed processor offers Active Directory Domain Services by means of the user personal information and the user public key. Other important security services are Active Directory Certificate Services that refers to creation and management of public key certificates and to bind person identity to corresponding private key.

INTRODUCTION

Nowadays, the development of web services for government needs has become a great challenge. For public administrations the support offered by this type of services is an opportunity for process optimization. Companies need to periodically report their fiscal, social and statistics information to various institutions. Almost all of these reports are required by legislation.

The need for distributed is vital for in E-Government, because of SOA bases application. SOA can potentially address those needs and provide software architecture for the interaction of existing and new distributed systems in E-Government. Through analyzing the characteristics of SOA and Web Services, Web Services is considered as the most appropriate technology for implementing SOA. In general, the e-Government applications address the complex of the integration and interaction and heterogeneous systems in E-Government system, especially interoperability, and may extend into other areas such as E-Commerce, E-Healthcare areas (Nahid, 2010).

On the other hand, E-government diffusion is an international phenomenon (Lemuria, 2008). So, the development of e-government in most countries is still primarily aimed at developing electronic services that customers can access via the Internet (Mateja, 2004). The more developed countries are therefore increasingly tailoring their e-government strategies in the direction of customer-orientation and instead of persisting with rigid organisational structures are working on integrating services and processes across individual administrative bodies and institutions and even include private businesses. So, the distributed and flexibale integration of front end services with backend solution represents the key to this problem.

In this context, we proposed a safe and flexible solution for communication between companies and public agencies. The proposed solution is a data flow processor that supposes three steps. First, the Public Institution sends to the Central Authority declarations model and the completing methodology. In the second step, the Central Authority implements the model in the programmatic language. Finally, the Central Authority starts the configuration procedure for all companies' processors. The retrieval of primary and transactional data can be realized in automatic manner. The automatic method implies that the company processor is interconnected with other specialized software systems. For data modeling the processor uses metadata database that keep track of all installed report templates, together with report proprieties as the date on which it was added to the current processor instance, the periodicity at which it should be run, the start and the end date of the report.

The processor, called in this paper PRO, was designed and implemented based on Microsoft technologies involving filling out of forms designed in InfoPath and than, parts of the forms can be filled out automatically. The process implies that the form data are digital signed. An important aspect regards security and will be detailed in this paper. The PRO is also a solution for future bussines technology as a support for e-government environment, having aggregated funcionality and security issues.

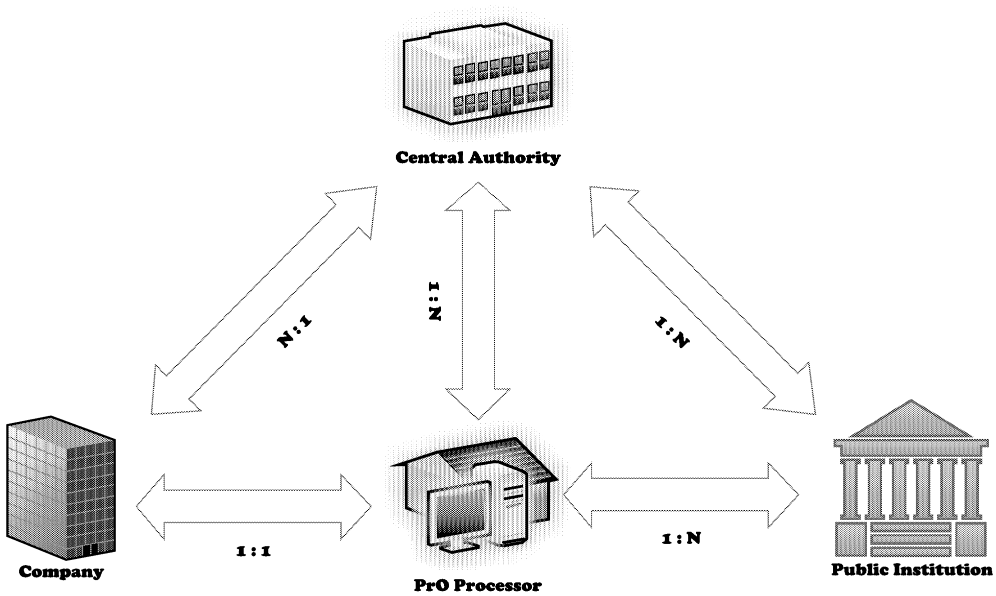
The paper is structured as follow. Section 2 present the PRO features and capability and the architecture of this processor is presented in Section 3. The use case scenario is presented in Section 4. Next, we detalies the structure of PRO: security infrastructure in Section 4, databases in section 5 and implementation details in Section 6. We presents in the last section the conclusions of this paper and the possible future work.

PRO PROCESSOR FEATURES

PRO Processor is a software communication processor that has as main purpose to create a transparent and automated link between the companies and the public institution regarding the legal duties of reports and declarations. The processor appeared due to the need of companies to optimize the current workflows in order to eliminate the redundant and low quality data and also to reduce costs for the local maintenance of current existing systems. The presented solution is flexible, safe and simple.

Threre are several features of the proposed procesessor that offers the possibility to manage online the process of periodically reports to authorities. One of them is that fact that it involves filling out of forms designed in InfoPath and parts of the forms can be filled out automatically. Also, the form data is digitally signed. All these features are better emphasized when describing the involved modules and the whole workflow under which the PRO processor works.

A report is a legal declaration that is required by a public institution. From technical point of view, a report consists of an InfoPath form generated over an XML Schema that has the same visual design as the paper declarations but is improved by using validation and navigation inside the electronic form.



Figures 2: The PROinfrastructure

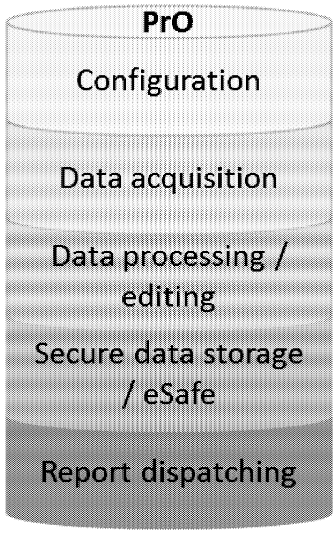
PRO ARCHITECTURE

The main modules involved in this processor can be seen in figure 1. The first one is the configuration module and is responsible for configuring the master data inside the PrO. The second module is the data acquisition one, which allows the PROto interconnect to other local systems inside the companies. Next, the data processing module generates the reports accordingly to the business rules implemented. The eSafe module stores all the required data securely. The last module is the report dispatching module that accomplishes the transportation of reports to the corresponding public institutions (Pop et al. 2010).

As a whole, the PRO system has a centralized architecture consisting of two main entities: the PRO Processor and the Central Authority as it can be seen in figure 2.

The PRO Processor is in a one-to-one relation to the customer company, being deployed in the company premises and being interconnected to the internal systems from where it gets the required data. The workflow under which the PRO processor works is the following: First of all, after being configured, the PRO will import all the required data from the external company systems. After that accordingly to what reports are installed and their properties, PRO generates the declarations accordingly to the business rules. Once this step is accomplished the reports are sent to a predefined person inside the company responsible for digitally signing the documents. After that PRO sends the report via a preconfigured interface to the Public Institution.

In order to accomplish its role, the PRO Processor uses a number of software: Microsoft BizTalk Server – is the main core of the system that has different orchestrations defined in order to execute the workflow of generating a report, Microsoft Form Server – that contains a Windows Sharepoint Services combined with an InfoPath Services and it is used as a report repository and a number of custom applications that integrate all the servers.



Figures 1: The modules of PRO processor

The Central Authority is a unique institution responsible for translating the manual report filling into automated BizTalk orchestration (Chappell 2009). This institution keeps track of all the reports that are required by the current legislation and maintains a unified data model. Another important role played by the Central Authority is to distribute these automated procedures to all the PrOs that are in production. From the technical point of view, the Central Authority consists of multiple portals created in Microsoft Form Server, each having its own role, as follows: Public Institution portal – is the link between the Central Authority and the Public Institution, where all the paperwork regarding a report is uploaded. The Customer portal allows new or registered users to keep track of their accounts, to update their master data or subscribe to reports.

On the Central Authority there is also a development environment, where a certified employee translates the manual steps of completing a report into electronic format: create the InfoPath template, create the BizTalk business rules and orchestration and configure the outgoing ports. Once the development process is finished for a report, the corresponding BizTalk application is prepared for synchronization in the development environment. From this environment all the BizTalk applications must be replicated on all PrOs in production.

Figure 3 presents the components of PRO design and their roles (Ciobanu et al. 2010). The **Portal** is responsible for data entry for both master and transactional data and consists of a number of sections, each section with a corresponding web form and a corresponding XML mapping considering the PRO internal XML representation of master and transactional data. Each section has one or more persons responsible for filling the data. There are two types of sections: master sections for master data and transactional sections for transactional data. There may be a third section for "user and role management". Where some admin must be able to authorize the "persons" to do their work.

The **Data Warehouse** is responsible for storing both master and transactional data. Here, a module stores and selects the different versions of the business rules.

The **Central Management Unit** is responsible for job scheduling, deciding when a job should be run and for job starting, testing if all the required data is available and if so start the job. If the data is not available yet, PRO sends alerts via email to the persons responsible for providing the missing data. This unit is also responsible for logging and functional reports, linking with the Central Authority according to the communication protocol defined for updates and deployment. It alerts, sending email alerts to responsible persons when data needs to be entered and gets confirmations before sending the reports for the auto generated data. Also it sends confirmed reports to the Public Institutions.

The **Business Rules Processing Unit** is responsible for executing the jobs.

The **Security Module** is responsible with managing the security issues and certificate management

The **Report archiving module** is responsible for archiving the XML reports following predefined rules.

Two types of adapters are shown in figure 3. The **enterprise adapter module** implements the enterprise specific interfaces to PrO. It is responsible for mapping the PRO internal XML representation of reports to enterprise specific forms and vice versa. It also maps the enterprise specific representations of master and transactional data to PRO representations.The **public institution adapter** is a module that implements the public institution specific interfaces to PrO. As in the case of the previous adapter, it is responsible for mapping the PRO internal XML representation of reports to the public institution specific forms and vice versa.

USE CASE SCENARIO FOR PRO

The dataflow configuration consists of three steps presented in figure 4.

Step 1: The Public Institution sends to the Central Authority model declarations together with the completing methodology.

Step 2: Central Authority will implement the model in the programmatic language.

Step 3: Central Authority starts the configuration procedure for all the PRO processors.

The retrieval of primary and transactional data can be realized in two methods:

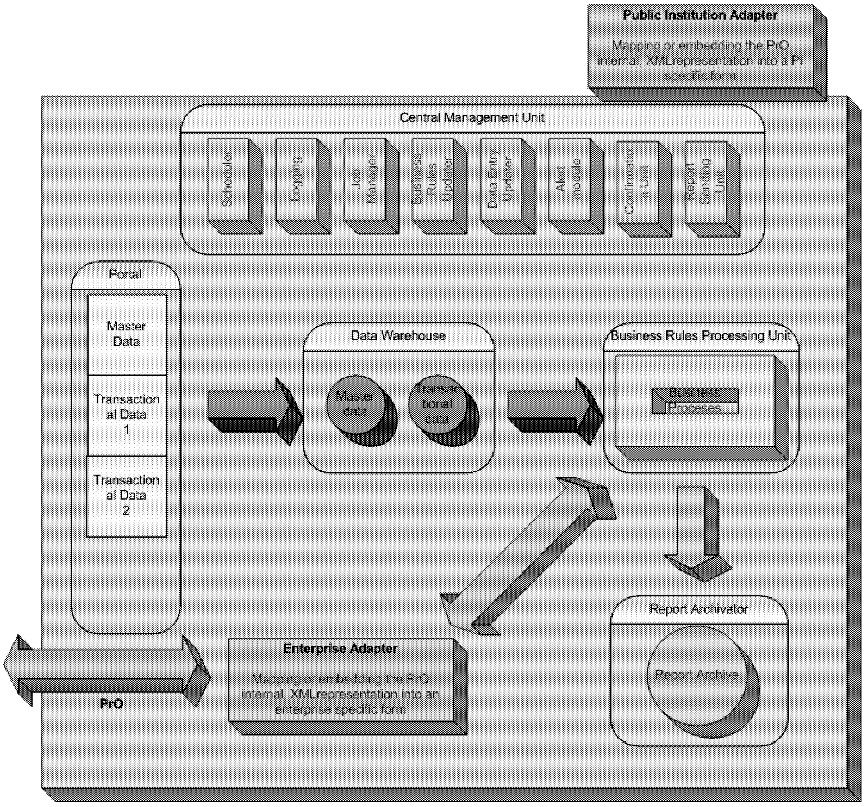
- Automatic – PRO will be interconnected with other specialized software systems.
- Manual – data will be manually introduced in a portal.

The generating declarations workflow implies three stes (see figure 5):

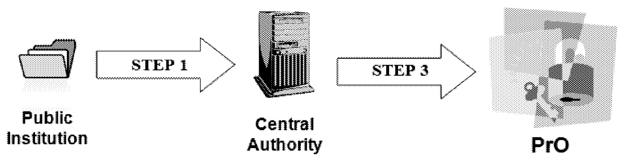
- Step 1: PRO generates a report using existing data and sends it to the responsible person from the company for validation and digital signature.
- Step 2: Responsible person modifies the report, signs it and sends it back to PrO.
- Step 3: PRO archives the report and sends it to the Public Institution.

The PRO Customer Portal is a Sharepoint site with many options that can help the customer to work with PRO processor. Thus PRO can be upgraded, updated and removed from this portal.

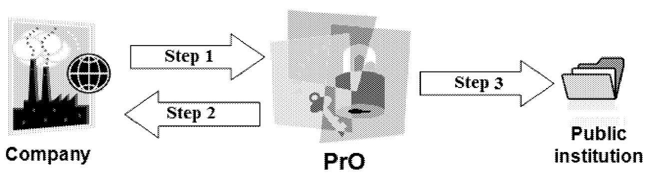
When a public institution or other report requesting authority sends a form template to Central Authority together with other form documentation, the responsible Central Authority studies all documentation and defines and implements the data model, the form template, the business process, the business rules and the Output Data Adapters. Then the Central Authority creates the deployment package. If a company makes new request for a PRO processor, it enters master data inside the PRO Customer Portal. The Central Authority prepares PRO processor and sends it to the company which fills a request form on PRO Customer Portal for different modules. Central Authority develops input data adapters accordingly to the needs of the company.



Figures 3: PRO design



Figures 4: The dataflow configuration



Figures 5: Generating declarations workflow

The synchronization process between the Central Authority and all PrOs is executed when some persons are defined responsible for new roles and the adapters are confugured together with the responsible Central Authority.

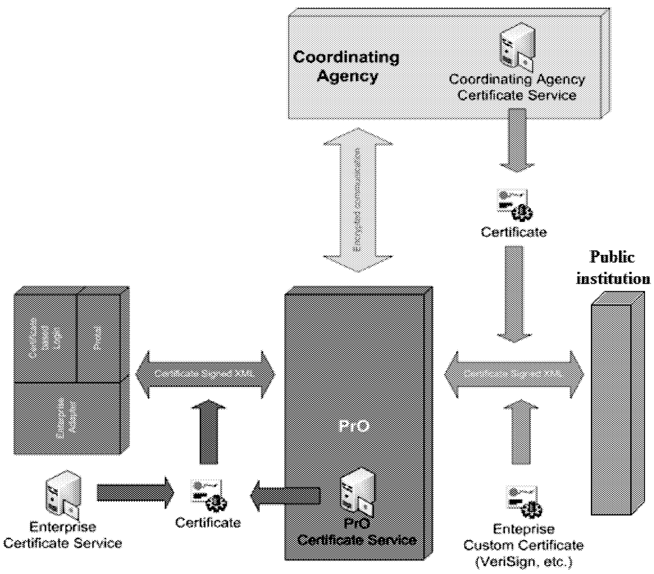
PRO imports raw data directly from Company environment by using data input adapters. Then PRO generates report based on raw data following all the business rules defined by the Central Authority. The auto-generated report is sent for eventual modification and confirmation (digital signature) to the responsible person, defined by the company. This responsible company modifies the report and/or signs the final document and returns it to PrO. PRO stores report inside the Report Repository and sends it to its destination by using the predefined output data adapter.

THE SECURITY INFRASTRUCTURE

The most important feature of PRO solution regards its security infrastructure. The security domains are:

- the Central Authority-the public institution;
- the Central Authority – the PRO processor;
- the Company- the PRO processor;
- the public institution – the PRO processor.

In these security domains, the communication is encypted (https) and the digital signature is used (Dai-Rui, 2010). The procedure implying the digital signature is presented in figure 6.



Figures 6: Digital signature

PRO offers Active Directory Domain Services (John, 2009) by means of the user personal information and the user public key. Other important security services achieved by PRO are Active Directory Certificate Services that refers to creation and management of public key certificates and to bind person identity to corresponding private key (Mitch, 2008).

PRO DATABASES

The PRO solution consists of several databases used for different tasks. One of them is the PRO Metadata database and the other is the report database (Carlos, 2007).

PRO Metadata database is used for keeping track of all installed report templates, together with report proprieties as the date on which it was added to the current PRO instance, the periodicity at which it should be run, the start and the end date of the report. On the other hand, each installed report has its own database that is automatically generated from the template xsd.

In order to modify data inside the database the Microsoft SQL Management Studio tool must be used.

Data synchronization is another important aspect of PrO. Database transfer employee master data is used to transfer employee master data from the customer's database to the PRO database. The customer can submit mapping rules, freely choosing which data to transfer by mapping columns and tables from both databases. The relationship between the PRO databases with the portal and the Business Rules Processing Unit can be seen in the PRO design presented in figure 3.

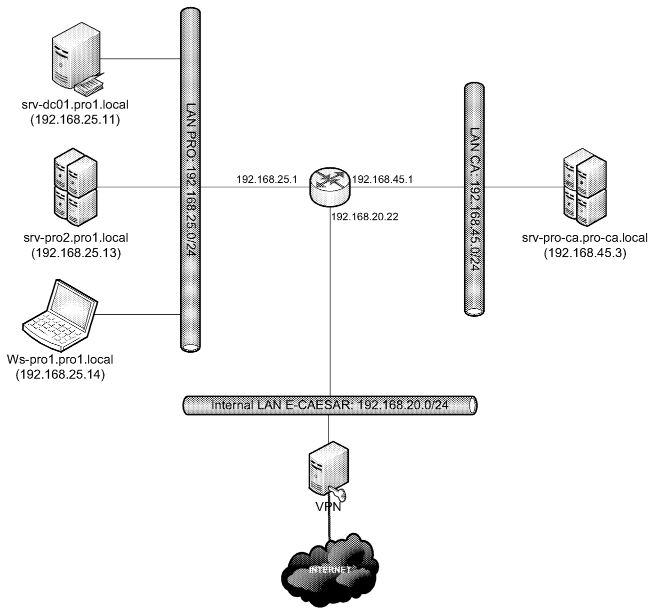
IMPLEMENTATION DETAILS

The processor was designed and implemented based on Microsoft technologies involving filling out of forms designed in InfoPath and than, parts of the forms can be filled out automatically. The solution was developed by using Microsoft .NET Framework 3.5 having Microsoft C# as programming language (Marshall 2008). The Extensible Markup Language (XML) was used thus assuring general interconnectivity between applications written in different programming languages (Esposito 2002). Another advantage provided by XML is the easiness in debugging due to the readability property of the language.

Description of the implementation must be realized taking into account each involved element.

For PRO client site, on both servers, the operating system is Windows Server 2008 SP1 Standard Edition. On the first server, the following roles are installed: Active Directory Domain Services, Active Directory Certificate Service and Internet Information Services (IIS) 7.0.

In the case of the second server, Internet Information Services (IIS) 7.0 is installed and the services that run on it are: Microsoft SQL Server 2005, Microsoft BizTalk Server 2009 and Microsoft Office Sharepoint Services 2007.



Figures 7: Network infrastructure

For the Central Administration, the operating system of the server is Windows Server 2008 SP1 Standard Edition. Here, the following roles are installed: Active Directory Domain

Services, Active Directory Certificate Service and Internet Information Services (IIS) 7.0. The services that run on this server are: Microsoft SQL Server 2005, Microsoft BizTalk Server 2009 and Microsoft Office Sharepoint Services 2007.

For the Public Institution, the installed operating system is Windows Vista SP2 and the software neededherae are Microsoft Outlook 2007 and Microsoft Infopath 2007. For the client computer, the server must be installed in the same manner as in the case of the public institution.

The network infrastructure is presented in figure 7, where three subnets cen be seen: one for the PRO site, one for the Central Administration and one for the Public Administration.

CONCLUSIONS

The paper presented a flexible solution for communication between companies and public agencies. The proposed PRO Processor is a software communication processor that has as main purpose to create a transparent and automated link between the companies and the public institution regarding the legal duties of reports and declarations. The processor appeared due to the need of companies to optimize the current workflows in order to eliminate the redundant and low quality data and also to reduce costs for the local maintenance of current existing systems.

The processor was designed and implemented based on Microsoft technologies involving filling out of forms designed in InfoPath and, than parts of the forms can be filled out automatically. The process implies that the form data are digital signed. As it was described in the paper, the processor offers Active Directory Domain Services by means of the user personal information and the user public key. Other important security services are Active Directory Certificate Services that refers to creation and management of public key certificates and to bind person identity to corresponding private key.

Using this solution, the companies overcome all the troubles when they must periodically report their fiscal, social and statistics information to various institutions.

ACKNOWLEDGMENTS

The research presented in this paper is supported by national project: "SORMSYS - Resource Management Optimization in Self-Organizing Large Scale Distributes Systems", Contract No. 5/28.07.2010, Project CNCSIS-PN-II-RU-PD ID: 201. The work has been co-funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the Financial Agreement POSDRU/89/1.5/S/62557. The research presented in this paper was implemented in the international project called “PRO”, developed by Centre for Advanced Studies on Electronic Services (e-CAESAR).

REFERENCES

Nahid Rashid and Shams Rahman. 2010. An investigation into critical determinants of e-government implementation in the context of a developing nation. In *Proceedings of the First international conference on Electronic government and the information systems perspective*(EGOVIS'10), Kim Normann Andersen, Enrico Francesconi, and Tom M. van Engers (Eds.). Springer-Verlag, Berlin, Heidelberg, 9-21.

Lemuria Carter and Vishanth Weerakkody. 2008. E-government adoption: A cultural comparison.*Information Systems Frontiers* 10, 4 (September 2008), 473-482.

Pop, F., C. Dobre, D. Popescu, V. Ciobanu and V. Cristea. 2010. „Digital Certificate Management for Document Workflows in E-Government Services Infrastructure”, in *Proc. of The 9th IFIP WG 8.5 International Conference on Electronic Government* (EGOV 2010) Lausanne, Switzerland.

Mateja Kunstelj and Mirko Vintar. 2004. Evaluating the progress of e-government development: A critical analysis. *Info. Pol.* 9, 3,4 (December 2004), 131-148.

Ciobanu, V., F. Pop, D. Popescu and V. Cristea. 2010. „A Distributed Approach to Business Intelligence Systems Synchronization”, in *Proc. of The 12th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, Timisoara, Romania.

Chappell, D., 2009, *“Introduction to BizTalk Server 2009*”, Microsoft.

E-CAESAR, “*PRO Project”,* accessed on June 4th, 2010, from http://www.e-caesar.ro/index.php/en/projects/pro

Esposito, D., 2002*, “ Applied XML Programming for Microsoft .NET”*, Microsoft Press.

Smith, J..2007, *“Inside Windows® Communication Foundation*”, Microsoft Press.

Microsoft,2010. *“Microsoft BizTalk Server 2009 Help*”, accessed on June 10th, 2010, from http://msdn.microsoft.com /en-US/library/aa548004(v=BTS.10).aspx

Marshall, D. , 2008, *Programming Microsoft Visual C#*, Microsoft Press.

Dai-Rui Lin, Chih-I Wang, Zhi-Kai Zhang, and D. J. Guan. 2010. A digital signature with multiple subliminal channels and its applications. *Comput. Math. Appl.* 60, 2 (July 2010), 276-284.

John Policelli. 2009. *Active Directory Domain Services 2008 How-To* (1st ed.). SAMS, Carmel, IN, USA.

Mitch Tulloch. 2008. *Windows Server 2008 Server Core Administrator's Pocket Consultant*. Microsoft Press.

Carlos Ordonez, Zhibo Chen, and Javier Garcia. 2007. Metadata management for federated databases. In *Proceedings of the ACM first workshop on CyberInfrastructure: information management in eScience* (CIMS '07). ACM, New York, NY, USA, 31-38.

DECISION SUPPORT TOOLS

IT-BASED SYSTEMS APPROACH TO ENGAGE PATIENTS IN MEDICAL ERROR REPORTING: IN CONCURRENCE WITH STAFF

Ranjit Singh
Ashok Singh
Gurdev Singh

462 Grider Street, Clinical Center, Buffalo NY 14215. USA
e-mail: {gsingh4 / rs10}@buffalo.edu

KEYWORDS
IT-Based, Meaningful Use of HIT, Patient Empowerment, Patient Friendly, Quality Management, Visualization.

ABSTRACT
Various consumer surveys show that a large proportion of US adults have had experience of injury caused by medical treatment and yet a very small proportion actually report medical errors. Although reporting efforts to date have almost exclusively focused on staff as the source of reports, it is now being recognized that patients can potentially play a significant and vital role in improving their own safety. At the same time there is an international move towards exploiting the powers of electronic health records (EHR) to improve the quality, effectiveness and efficiency of health care. The authors have therefore proposed a computer touch-screen patient-reporting system that has great potential for patients to (a) help prevent adverse events in the first place, (b) detect errors, (c) ameliorate their effects, and (d) *empower themselves* to propose interventions for safety improvement. This system should contribute to enriching *meaningful use of HIT* in all healthcare settings through concurrent input from patients.

INTRODUCTION AND BACKGROUND

Medical errors contribute to significant harm and costs to individuals and nations. Formation of a culture of safety is widely accepted as the most effective strategy in the war against medical errors. Figure 1 summarizes the eight major contributors to the formation of a safety culture [1]. As can be seen, the creation of a 'Learning environment' through error reporting is one of these major contributors, which are mutually re-enforcing. Advantage needs to be taken of the cybernetic loops between these contributors to the formation of a culture of safety. The Patient Safety and Quality Improvement Act (2005) was introduced in the US to stimulate increased error reporting through the creation of so-called Patient Safety Organizations (PSOs).

Outpatient settings are where most citizens receive most health care [2,3]. Various consumer surveys show that a large proportion of US adults have had experience of injury caused by medical treatment and yet a very small proportion actually report medical errors. Although reporting efforts to date have almost exclusively focused on staff as the source of reports, it is now being recognized that patients can potentially play a significant and vital role in improving their own safety. An important barrier to their input is the fact that patients often have different perceptions than clinicians about what constitutes an unsafe episode. Patients are likely to find the vocabulary of patient safety even more confusing - than do the providers [4,5]. The association between patient perceptions of safe care, medical injury, and service quality merit further study.

Phillips et al., Kuzel et al. and Weingart et al express the need for developing methods of inviting/engaging/enabling patients to increase the quality and frequency of their reports to help improve their own safety [4,6,7]. Patients are increasingly being recognized as partners in healthcare, particularly with the advent of the 'patient-centered medical home' in the US. They are a major source of knowledge about the state of healthcare safety. Previous attempts to elicit error reports from patients have had limited success, due to a variety of factors, particularly health literacy. Little is known about how best to take advantage of the valuable knowledge and experience of patients and their health givers. There is an international move towards exploiting the powers of electronic health records (EHR) to improve the quality, effectiveness and efficiency of health care. England has invested £12.8 billion and recently the US has committed $38 billion towards this end. It is expected that the EHRs will be user friendly, and will engage patients concurrently in their own care via 'patient portals'.

INCORPORATING PATIENTS' NEEDS AND PERSPECTIVES IN THE PROCESSES OF ERROR REPORTING

Patients and their families have special and particular needs. Specific aids/'lenses' need to be developed for facilitating:

- clearer perceptions of what a medical error is [4,5]
- what the cascades and consequences of errors can be
- how to narrate the 'story' of an event
- how to make suggestions for avoiding future errors
- removing/minimizing the fear of reporting due to emotive and cognitive biases
- how and where to report
- how to receive feedback for their valuable input
- comparison between various health care settings

Figure 2 shows the conceptual Model for the Error Reporting Cycle and the Barriers Faced by Patients.

CONSIDERATIONS FOR THE DESIGN OF THE PATIENT REPORTING INSTRUMENT WITH SPECIAL ATTENTION TO HEALTH LITERACY

Patients must therefore be *invited, encouraged and enabled* to take responsibility as vital members of the health care team i.e. participate at the front line [8]. Whilst assigning responsibility of reporting to patients particular attention needs to be paid to the huge national health literacy gap. According to the US Institute of Medicine nearly half of the American adults have difficulty understanding and using health information [9]. The rate of hospitalization and use of emergency services is higher among patients with limited literacy [10]. Even well educated people with strong reading and writing skills may have trouble comprehending a medical form or doctor's instructions. The Health Literacy Institute calls for 'vibrant plain language' (visualization is a great example) to help patients acquire ability to partner with their clinicians and to increase safety and quality of care. It is imperative that any reporting system, particularly one for patients, should be functionally efficient, effective, equitable, (bearing in mind the special needs of older persons and socio-economically disadvantaged and minority populations) and be cost effective.

What is needed is a common vision that is clearly understood by both patients and providers, thus creating team spirit between them. A natural question therefore is: can visualization provide a clear common language and a common vision of safety in the healthcare system. As the saying goes: *"A picture is worth a thousand words."* The authors attempt to address that above issues by proposing a "Patient-centered Visual Error Reporting System" (P-VERS) using a visual approach that aids reporting, meaningful and rapid feedback. The P-VERS concept is based on clear visual modeling of healthcare at macro-system and micro-system levels [5]. This micro-model *portrays* (1) all the entities (e.g. doctor, patient and her/his home), and (2) interactions between them (e.g. communication between patient and doctor). Lists of possible errors in each of these are made available to patients to choose the appropriate one. Figure 3 (A&B) illustrate what the patients see and how they are helped with understanding the scope of reporting as well as choosing the appropriate error/s, through out the system. The error lists are an excellent aid to reducing cognitive and emotive biases. Further patients need to understand the patterns of errors in the system. Figure 4 shows a proposed format for summarizing multiple errors using visual models – in this case, the micro-model of primary care is annotated to show the frequencies of errors reported at each point in the system. The use of a touch-screen system is proposed, that is specially designed from the patient perspective. These concepts will be simulated at the conference. Audience response will inform future developments. Our preliminary experiences of engaging lay persons using a visual approach have been encouraging, supporting the potential effectiveness of P-VERS for soliciting error report from patients.

FUTURE WORK

The P-VERS is currently under development and plans for testing are in place, including formal comparative effectiveness studies of P-VERS vs. traditional text-based patient-reporting.

CONCLUSIONS

Based on preliminary experiences, the P-VERS concept appears to have potential for identification and reporting of errors as well as meaningful compilation of data on these errors and adverse events. The authors believe that a computer touch-screen P-VERS has great potential for patients to (a) help prevent adverse events in the first place, (b) detect errors, (c) ameliorate their effects, and (d) *empower patients* to propose interventions for safety improvement. P-VERS can be particularly helpful in reducing the age, income, ethnicity, race and related disadvantages through its use of universal 'language,' which in turn should lead to patient satisfaction. The P-VERS should contribute to enriching *meaningful use of HIT* in other healthcare settings as well.

References

1. Singh G, Singh R, Thomas EJ, et al. Measuring safety climate in primary care offices. In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in patient safety: new directions and alternative approaches.* Vol 2. Rockville, MD: Agency for Healthcare Research and Quality; 2008:59-72.
2. Green LA, Dovey SM. Practice based primary care research networks. They work and are ready for full development and support. *Bmj.* Mar 10 2001;322(7286):567-568.
3. Phillips RL, Dovey SM, Hickner JS, Graham D, Johnson M. The AAFP patient safety reporting system: development and legal issues pertinent to medical error tracking and analysis. *Advances in Patient Safety: from Research to Implementation.* Vol 3. Rockville, MD: AHRQ Publication No. 05-0021-CD; 2005:121-134.
4. Weingart SN, Price J, Duncombe D, et al. Patient-reported safety and quality of care in outpatient oncology. *Jt Comm J Qual Patient Saf.* Feb 2007;33(2):83-94.
5. Singh R, Pace W, Singh S, et al. A concept for a visual computer interface to make error taxonomies useful at the point of primary care. Informat Prim Care 2007; 15: 221-229.
6. Phillips RL, Dovey SM, Graham D. Learning from different lenses: reports of medical errors in primary care by clinicians, staff, and patients: a project of the American academy of family physicians national research network. *J Patient Safety.* 2006;2(3):140-146.
7. Kuzel AJ, Woolf SH, Gilchrist VJ, et al. Patient reports of preventable problems and harms in primary health care. *Ann Fam Med.* Jul-Aug 2004;2(4):333-340.
8. Conway PH, Clancy C. Transformation of health care at the front line. *Jama.* Feb 18 2009;301(7):763-765.

9. Nielsen-Bohlman L, Institute of Medicine (U.S.). Committee on Health Literacy. *Health literacy : a prescription to end confusion*. Washington, D.C.: National Academies Press; 2004.
10. AHRQ. *Literacy and health outcomes*: Agency for Healthcare Research and Quality, Rockville, MD.; 2004.

BIOGRAPHIES

Ranjit Singh MB BChir, MA (Cantab), MBA, MABFM, MD, a graduate of the University of Cambridge (England) where read Medicine and Management. He is Associate Director and Professor, Patient Safety Research Center, State University of New York at Buffalo, USA. He has authored numerous papers on patient safety and simulation and served on the US National Expert panel on reporting in primary care.

Ashok Singh, MB BChir, MA (Cantab), MABFM, MD, is a graduate of the University of Cambridge (England) where he read Medicine and Management. He is a Partner of The Niagara Family Medicine Associates, Niagara Falls. He has authored papers on patient safety and simulation.

Gurdev Singh MScEng PhD (Birm), is the founding Director of Patient Safety Research Center at State University of New York at Buffalo, USA. In 1996 he was recognized by the US Department of Justice as an 'Alien of Extraordinary Ability This Classification is awarded to a person who has demonstrated that s/he has reached the top of her/his profession at the international level for a sustained period of time, or to a Nobel Laureate or equivalent. He is Fellow of the Royal Society of Medicine (London). His experience and themes of his publications include: development and use of hybrid decision support systems, multidisciplinary appraisal of risk and reliability in projects and structures based on Monte Carlo simulation, multi-objective and multi-resource optimization and 'satisfisation', development of Singh-Markowitz efficient frontier method, concurrent engineering, stochastic and visual simulation of construction projects, and computer-aided learning of risk and reliability analysis.

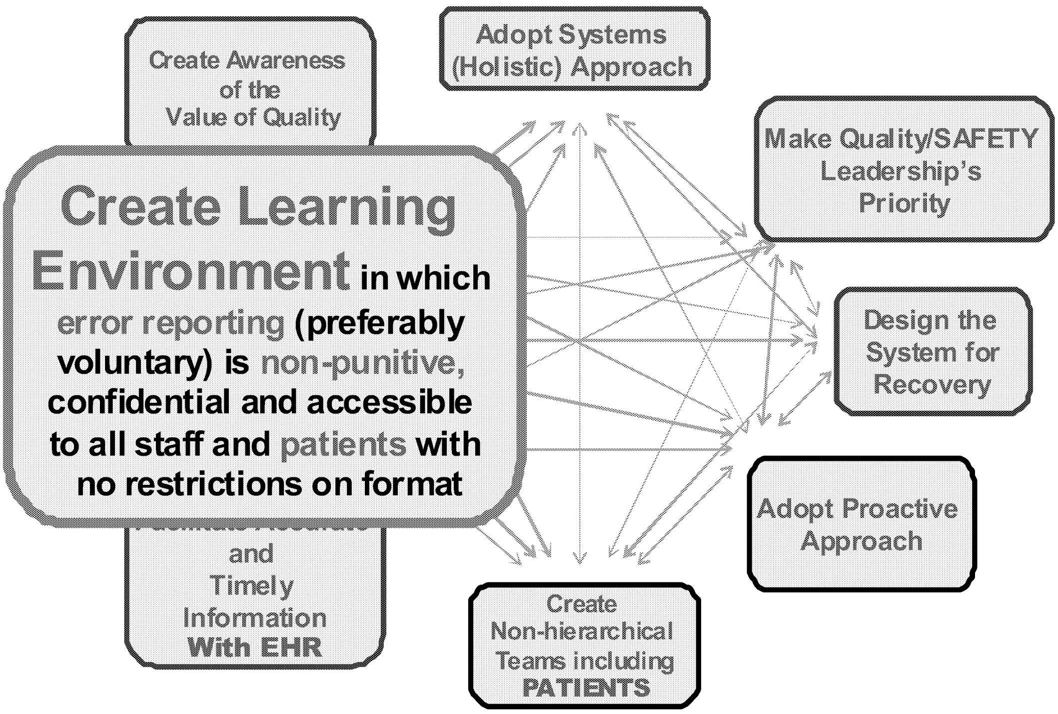


Figure 1: Interactive Contributors to the Culture of Patient Safety Singh et al. [1]

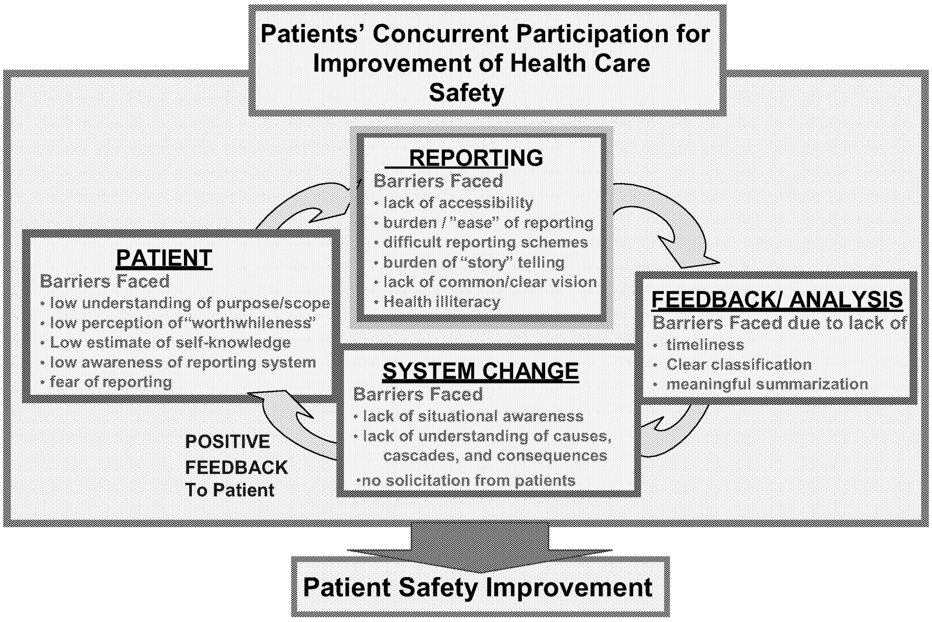


Figure 2: Conceptual Model for the Error Reporting Cycle and the Barriers Faced by Patient.

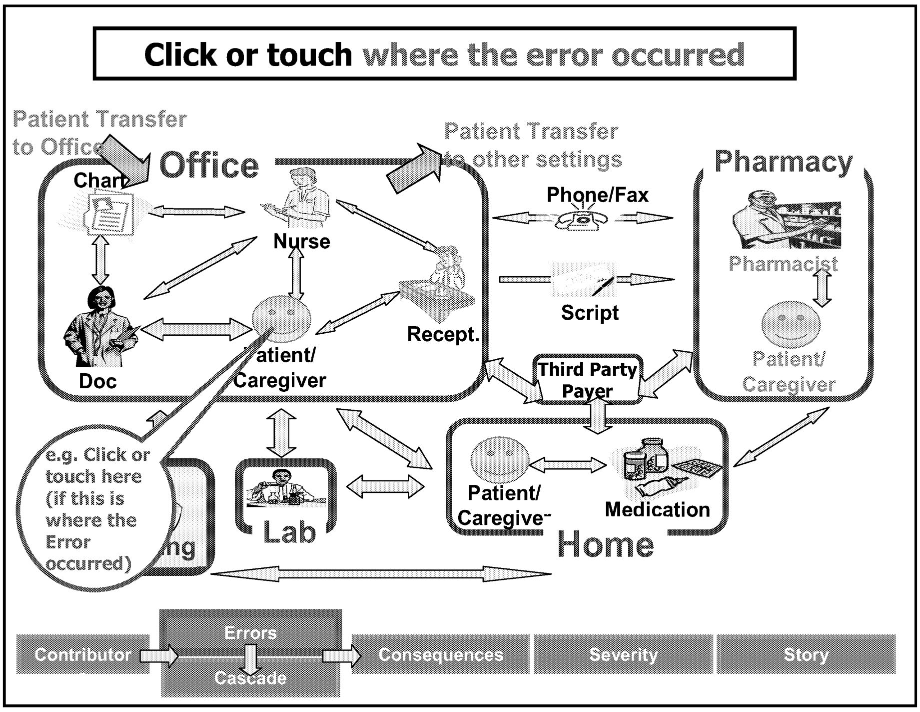


Figure 3A: Portrayal of the Office Designed to Invite Patient Input.

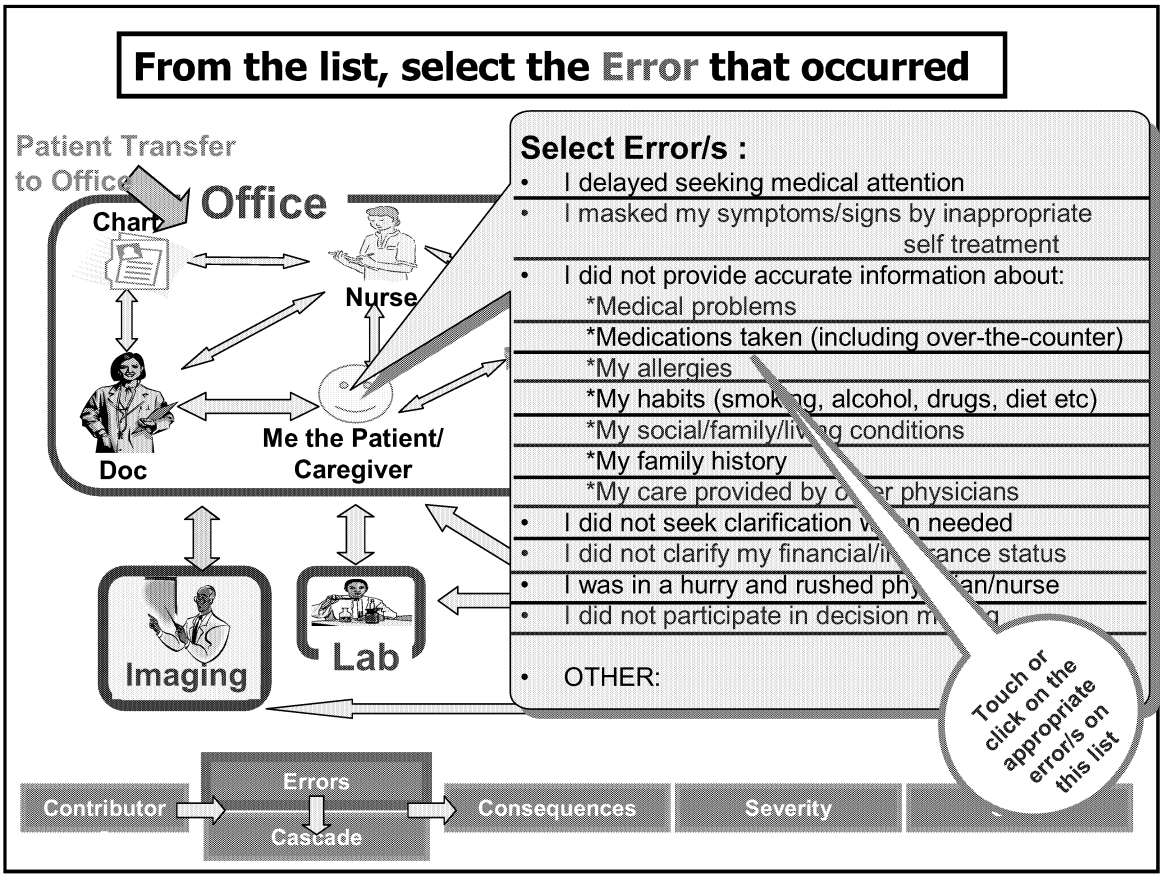


Figure 3B: What the Patients See on the Screen with Error Lists to Minimize Cognitive and Emotive Biases.

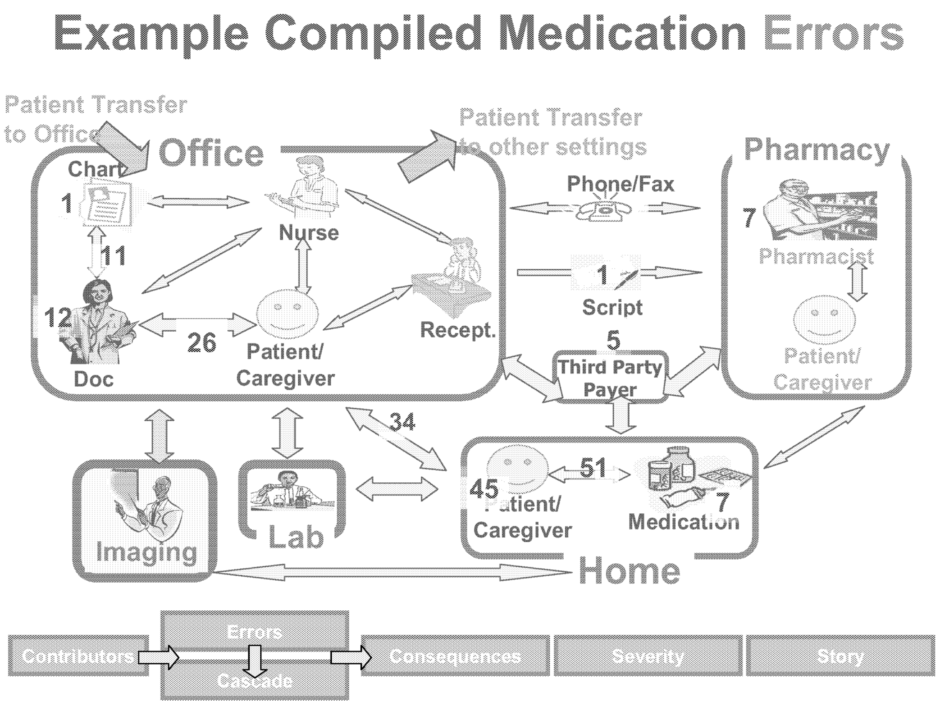


Figure 4: Easy to Understand Summary of the Error Reports Collected Over a Study Period.

SIMULATION AS A DECISION SUPPORT TOOL IN MAINTENANCE FLOAT SYSTEMS

Francisco Peito
Polytechnic Institute of Bragança
Higher School of Technology and Management
Industrial Management Dpt
Campus de Santa Apolónia
5301-857 Bragança
Portugal
pires@ipb.pt

Guilherme Pereira
University of Minho
School of Engineering
Production and Systems Dpt
Campus de Gualtar
4710 057 Braga
Portugal
gui@dps.uminho.pt

Armando Leitão
Polytechnic Institute of Bragança
Higher School of Technology and Management
Industrial Management Dpt
Campus de Santa Apolónia
5301-857 Bragança
Portugal
afleitao@ipb.pt

Luis Dias
University of Minho
School of Engineering
Production and Systems Dpt
Campus de Gualtar
4710 057 Braga
Portugal
lsd@dps.uminho.pt

KEYWORDS

Simulation, Discrete Event Simulation, Maintenance, Preventive Maintenance, Queueing Theory, Float Systems.

ABSTRACT

This paper is concerned with the use of simulation as a decision support tool in maintenance systems, specifically in MFS (*Maintenance Float Systems*). For this purpose and due to its high complexity, in this paper the authors explore and present a possible way to construct a MFS model using Arena® simulation language, where some of the most common performance measures are identified, calculated and analysed.

INTRODUCTION

Since the industrial revolution, the world has seen an accelerated growth of the complexity of enterprises and organizations. This considerable increase of the complexity and specialisation, either in production, or management, has raised several problems, namely those associated with the complexity in the distribution and the optimum use of resources (generally restricted), in order to improve the effectiveness within the organization. The need to resolve more efficiently these and other problems led to the appearance of operational research (O.R.) as a science during the second World War. There are two types of models in O.R. : *optimization* and *simulation*. The *optimization models* are more suitable when all the variables within the system are deterministic and structured towards the choice of a single alternative, which will be considered optimal according to a predetermined criterion. The other type, the *simulation models,* would be more adequate in systems which have stochastics variables. These models, in turn, allow the analysis of various scenarios in the process of decision. Each scenario can be seen as a specific configuration of the analysed system. Thus, simulation does not produce an optimum and unique solution, but instead, a response from the system to a change of its configuration (Mamede 1984; Ingalls 2001; Shannon 1998). If the model is not a valid representation of the system in study, the results of the simulation bring few useful information about the real system (Rodrigues and Carvalho 1984; Rubinstein and Melamed 1998).

According to (Pegden et al. 1990), the simulation can be understood as the process of construction of a representative model of a real system, as well as experiments with this model with the goal of better understanding their behavior and assess the impact of alternative strategies of operation. Thus, simulation may also be considered as a decision support tool that allows to project and to analyze the performance of complex systems and processes as they are in many real systems. These real systems, such as factories, hospitals, transport fleets, etc., by their complexity and requirements imposed by the increase of the competitiveness or maintenance of a high level of operational availability, involve a large number of variables whose management is complex and has a high impact on its performance. However, with the use of simulation we acquired a capacity to forecast and achieve quickly the importance of taking some decisions about the system under analysis.

Mainly due to the non-existance of a specific simulator for the maintenance field, we had a great difficulty in choosing an appropriate simulation tool. However, (Dias et al. 2005) had a definite contribution as far as the simulation tool decision is concerned.

In fact, the choice of Arena® as a simulation language was based on the fact that its hierarchical structure offers different levels of flexibility, thus allowing the construction of extremely complex models, allied to a strong visual component.

As far as float systems maintenance models is concerned, (Lopes 2007) refers some studies where simulation has been used to produce results based on specified parameters. Due to the fact that these simulation models were only concerned with the input/output process, without dealing with what is happening during the simulation data process, some metamodels have emerged (Madu and Kuei 1992a; Madu and Kuei 1992b; Madu and Lyeu. 1994; Kuei and Madu 1994; Madu 1999; Alam et al. 2003). The metamodels express the input/output relationship through a regression equation. These metamodels can also be based on taguchi methods (Madu and Kuei 1992a; Kuei and Madu 1994) or on neuro networks (Chen and Tseng 2003). These maintenance system models were also recently treated on an analytical basis by (Gupta and Rao 1996; Gupta 1997; Zeng and Zhang 1997; Shankar and Sahani 2003; Lopes 2007). However, the model proposed by (Lopes 2007) is the only one that deals, simultaneously, with three variables: number of maintenance teams, number of spare equipments, and time between overhauls, aiming the optimization of the system performance. Although this proposed model already involves a certain amount of complexity it may become even more complex by adding new variables and factors such as: a) time spent on spare equipment transportation, b) time spent on spare equipment installation; c) the introduction of more or different ways of estimating efficient measures; d) allowing the system to work discontinuously; e) speed or efficiency of the repair and revision actions; f) taking into account restrictions on workers timetable to perform the repair and revision actions; g) taking into account the workers scheduling to perform the repair and revision actions; h) taking into account the possibility of spare equipment failure; etc. Anyway these mentioned approaches would aim at ending up with MFS models very close to real system configurations. In fact, the literature review showed that most of the works published, involving either analytical or simulation models, concentrate on a single maintenance crew, or on a single machine on the workstation or even considering an unlimited maintenance capacity – thus overcoming the real system complexity and therefore not quite responding to the real problem as it exists. This way, the authors believe this paper will make an important contribution to this issue.

We must refer that the proposed tool is intended exclusively to give a response to a type-standard configuration of *Maintenance Float Systems*. This way, the resulting MFS model aims to fill a gap in terms of computer solutions currently existing for this specific type of maintenance systems.

MAINTENANCE FLOAT SYSTEM DESCRIPTION

A typical *Maintenance Float System* is composed of a workstation, a maintenance center with a set of maintenance crews to perform overhauls and repair actions and a set of spare machines (Fig.1). The workstation consists of a set of identical machines and the repair center of a limited number of maintenance crews and a limited number of spare machines. However, the model we have adopted, being a typical MFS, presents certain specificities both as far the philosophy of the maintenance waiting queues are concerned, and related to the management of the maintenance crews.

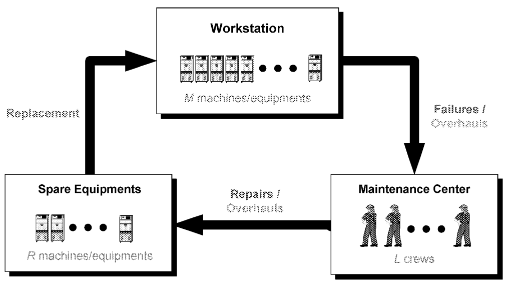


Fig. 1 –Typical Maintenance Float System

This model follows the one proposed and developed by (Lopes 2005; Lopes et al. 2006; Lopes et al. 2007), considering *M* active machines, *R* independent and identical spare machines and *L* maintenance crews The active machines considered operate continuously. Machines that fail are taken from the workstation and sent to the maintenance park waiting queue, where they will be assisted according to arrival time (Fig.2). Machines that reach their optimal overhaul time are kept in service until the end of a period *T* without failures. However they will be also kept on a virtual queue to overhaul. If the number of failed machines plus the number of machines requiring overhaul is lower than the number of maintenance crews available, machines are replaced and repaired according to FIFO (*First In First Out*) rule. Otherwise if it exceeds the number of maintenance crews, the machines will either be replaced (while there are spare machines available) or will be sent to the maintenance queue. The machines that complete a duration period *T* or time between overhauls in operation without failures are maintained active in the workstation, where they wait to be assisted, and they are replaced when they are retired of the workstation to be submitted to a preventive intervention. Its replacement is assured by the machine that leaves the maintenance center in the immediately previous instant. If an active machine happens to fail it awaits for the accomplishment of an overhaul, then it will be immediately replaced, if a spare machine is available or as soon it is available. (Fig.2)

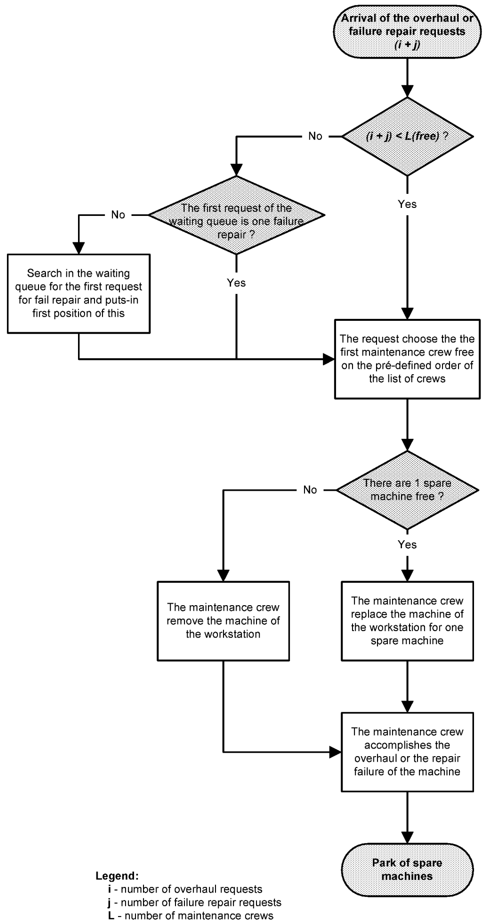


Fig. 2 - Procedure for the management of requests for revision or fail repair by the maintenance crews

In the MFS analized, it is assumed that the *M* active machines of the workstation have a constant failure rate for the following reasons:

- When the failure rate is increasing, the models for calculating the optimal number of spare machines have little interest. In these cases, one must be concerned with the causes rather than on the consequences - the optimal number of standby units, the optimal number of maintenance crews and the optimal time interval between overhauls should be discussed.
- When the failure rate is decreasing, and until the system reaches the steady state, it would be premature to plan maintenance - this mainly occurs in the early life of the machines.

Time between failures are assumed as independent and identically distributed following an *Exponential* Distribution for all machines (failures occur under a *Homogeneous Poisson Process*). However, during a simulation run, this value could be adjusted based on time between overhauls. Obviously a smaller time between overhauls implies greater time between failures.

As far as time to overhaul and time to repair are concerned, we have assumed the *Erlang-2* distribution, eventhough considering overhaul time significantly lower than the repair time.

For this MFS, the following parameters and variables are identified:

Parameters

1. Number of active machines (M);
2. Number of maintenance crews (L);
3. Number of spare machines (R);

Variables

4. Machine- Overhauls rate (λ_{rev})*;
5. Machine-Initial Failures rate (λ_f)*;
6. Crews-Repair rate (μ_{rep})*;
7. Crews-Overhaul rate (μ_{rev})*;
8. Failure cost (C_f);
9. Repair cost (C_{rep});
10. Overhaul cost (C_{rev});
11. Replacement cost (C_s);
12. Cost due to loss production (C_{lp});
13. Holding cost per time unit (h);
14. Labour cost per time unit (k);
15. Time to convey and install spare machine ($T_{ConvInst}$).

(*) This variable can be adjusted during the simulation run.

SIMULATION MODEL

The Arena® simulation language environment was used in the development of the simulation model for this MFS (Kelton 2004; Pidd 1989; Dias 2006 and Pidd 1993).

The steps for the development of the simulation model are presented in the following figure (Fig.3).

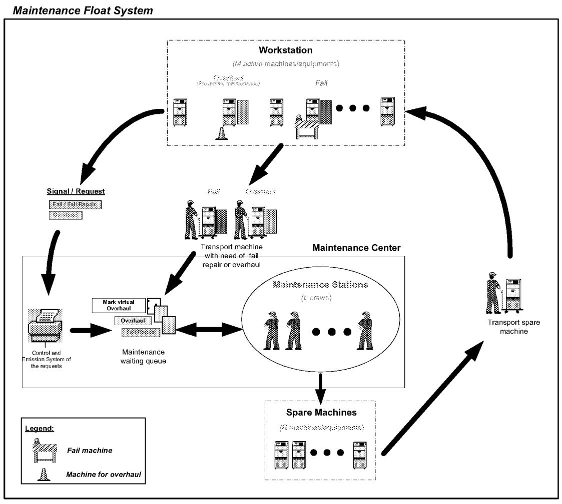


Fig. 3 - Steps for simulation model development

The logical model configuration choice (Fig.4) for the MFS intends to provide a clear global visualization of the undergoing operations and a great flexibility as far as the definition of its basic structure (M, L and R) is concerned.

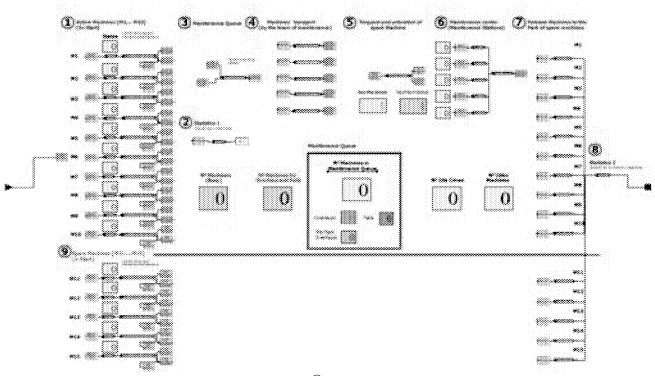


Fig. 4 - Arena® Logic Model

Figure 4 explicits the global logical simulation model, underlining its different components developed:

1. Active machines *(workstation)*;
2. Statistics 1 *(Recording Machines T_{up})*;
3. Maintenance queue;
4. Machines' transportation *(by the maintenance crews)*;
5. Spare machine request;
6. Maintenance center *(set of maintenance Stations)*;
7. Release machines to the set of spare machines;
8. Statistics 2 *(Recording Machines T_{up} and T_{down})*;
9. Spare machines *(in the start of the system)*.

The components 1 and 9 include a generation and control system for repair and overhaul requests for each machine (Fig.5).

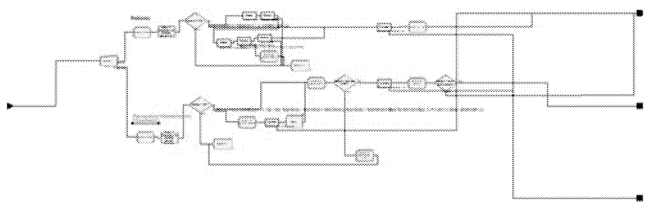


Fig. 5 - Generation and control system for repair and overhaul requests

This repair and overhaul generation and control system is implemented for each machine in order to assure total independence of each occurrence of each request.

As far as the overhaul requests are concerned, they were defined according to the following situations:

1) Based on the total number of requests in the waiting queue (repairs and overhauls), there exists, at least, a maintenance crew and a spare machine available;
2) Or, based on the total number of requests (in the waiting queue (repairs and overhauls), there is no maintenance crew or spare machine available.

In the first situation, the action for overhaul is made immediately and the machine is removed and replaced by the maintenance crew. In the second situation, it is only made a virtual request that is performed only after confirmation. In this case, the machine remains in operation and confirmation will only occur when the system identifies a maintenance crew and a spare machine available.

For all the machines in the system, the repair request has priority over the overhaul request. In this respect it should be emphasized the specific situation where a machine fails, while active but already waiting for an overhaul. In this case the overhaul request is removed and replaced by a repair request. Furthermore, the time between failures will be adjusted according to the overhaul rate.

The waiting queue rule in the maintenance queue is defined in component 3, together with the "control mechanism" which only allows a request to proceed if there is a free maintenance crew (Fig.6).

FIFO (First In First Out) is the rule for the maintenance queue management, except for the case when the total number of maintenance crew requests (overhauls plus repair actions) exceeds the number of maintenance crews – in this case, machines requiring repair action heve priority over machines requiring overhauls.

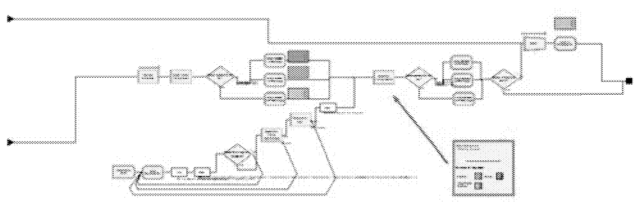


Fig. 6 - Maintenance waiting queue

Component 4 is responsible for requesting transportation action for a spare machine while seizing the respective maintenance crew

Component 5 (Fig.7) represents the request of a spare machine, performed by a maintenance crew. If a spare machine is available, the maintenance crew takes and installs the machine, replacing the one requiring maintenance crew action.

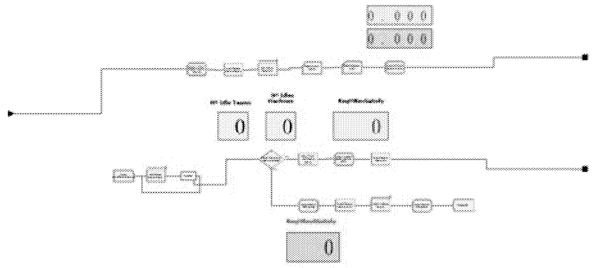


Fig. 7 - Request and activation of spare machines

Component 6 distinguishes overhaul actions and repair actions (Fig.8). This will permit to clearly understand whether a maintenance crew is performing one action or the other. This component is also important as to produce true performance measures related to both maintenance operations.

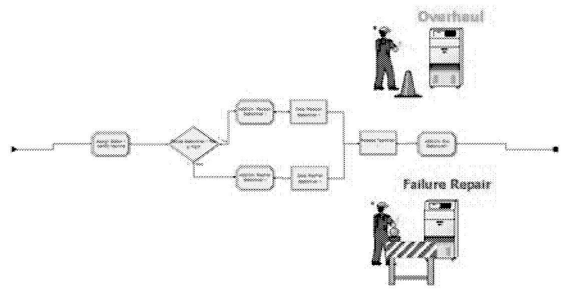


Fig. 8 – Identification and statistics of the states of the maintenance crew

Component 7 is then responsible for releasing machines under maintenance crew actions, whenever they finish their work, either repairing or performing overhauls.

Finally, components 2 and 8 are recording fundamental statistical data to calculate adequate efficiency measures. In fact component 2 is storing operation times for each machine (Time-up), while component 8 is storing down times for each machine, thus estimating and updating performance measures.

Fig. 9 – Record statistics

Figures 10 and 11 highlight both input parameters window and output updates – numerically and graphically. Figure 12 shows an application screenshot including simulation animation.

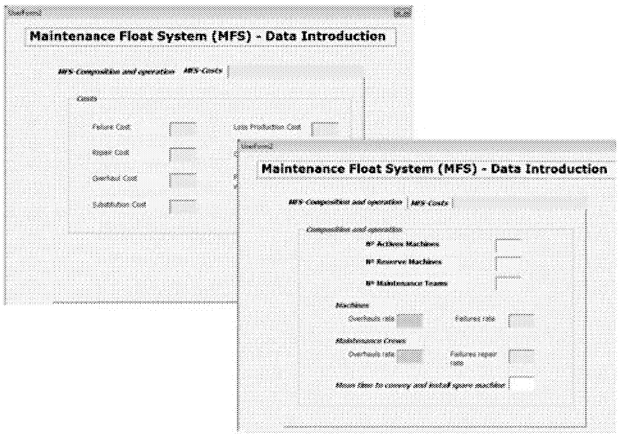


Fig. 10 – Data intoduction area sample screenshot

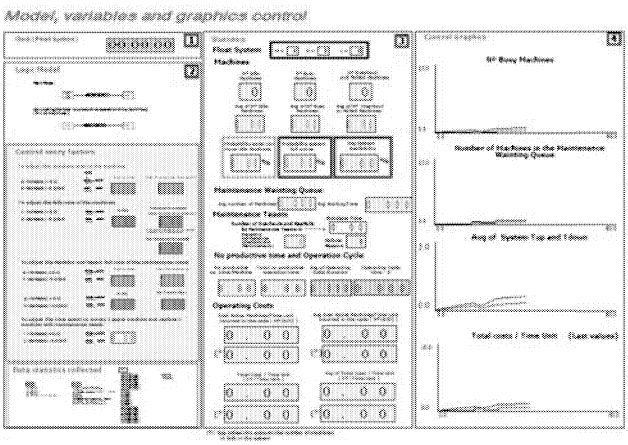


Fig. 11 - Variables and graphics control

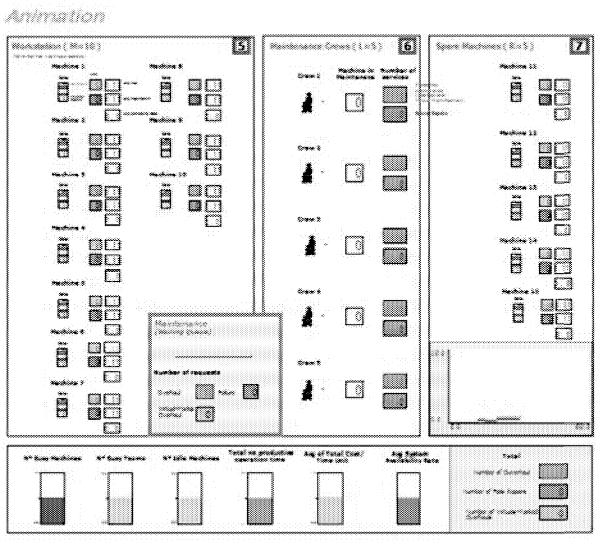


Fig. 12 – Animation area sample screenshot

DISCUSSION

The developed simulation application for a Maintenance Float System allows the estimation of the following global efficiency measures:

a) Average system availability *(AvgSAv)*;
b) Total maintenance cost per time unit *(AvgTCu)*;

However, some other performance measures are also analysed, such as:

c) Average number of missing machines at the workstation $(AvgM_{eq})$,
d) Average number of machines in the maintenance waiting queue *(AvgLq)*;
e) Average waiting time in the maintenance waiting queue (*Avg Wt)*;
f) Average operating cycle time (*AvgD)*;
g) Probability of existing 1 or more idle Machines $(Prob_{im})$;
h) Probability of the system being fully active $(Prob_{s})$;

And, finally, the simulation model also computes some individual efficiency measures per machine or maintenance crew, i.e.,

i) Utilization rate per machine;
j) Utilization rate per maintenance crew;
k) Number of overhauls and repair actions performed per maintenance crew;
l) Average availability per machine.

The MFS decision support system developed model has three major characteristics:

- Setting Parameters

  This means that the user can interact with the simulation system through the initial introduction of various parameters such as repair rate, overhaul rate, repair cost and others. The user can, this way, evaluate the system possibilities under different operating conditions.

- Flexibility

  This means that the user can interact with the running simulation system through the modification of various parameters values. The user can, this way, evaluate the system behaviour under different maintenance strategies.

- Interactivity

  This means that the computer screen continuously display the system status as where as its time evolution allowing a better communication between the model and the user. Indeed, the strong visual aspect offered by the developed model clarifies the actual process inside the system. This allows a better understanding of the different interactions in the model and of the simulation results.

FUTURE DEVELOPMENTS

The simulation model here presented, incorporating analysis of usual performance measures, also drives its concern towards new efficiency measures, enabling new trends for the analysis and discussion of the best decisions as far as a specific Maintenance Float System is concerned. Nevertheless the authors are now aiming to the development of an advanced simulation model, incorporating flexibility. This target would be reached by developing and incorporating new modules in our simulation tool, following past experiences found on literature (Luís S Dias, 2005, 2006 and Vilk, P., 2009, 2010) where the automatic generation of simulation programs enables desired model flexibility, i.e., making the model generating specific simulation programs for specific Maintenance Float Systems. These mentioned future developments also intend to potentiate the known capability of simulation to efficiently communicate with managers and decision makers, even if they are not simulation experts.

REFERENCES

Alam, Fasihul M.; McNaught, Ken R. and Ringrose, Trevor J. (2003), “Developing Simulation Metamodels for a Maintenance Float System”, AMORG, Engineering Systems Department, Cranfield University, RMCS Shrivenham, Swindon SN6 8LA, UK

Chen, M. C. and Tseng, H. Y. (2003), "An approach to design of maintenance float systems," Integrated Manufacturing Systems, vol. 14, pp. 458-467.

Dias, Luis M. S; Rodrigues, A. J. M. Guimarães; Pereira, Guilherme A. B. (2005), An Activity Oriented Visual Modelling Language with Automatic Translation to Different Paradigms, Proceedings of the 19th European Conference On Modelling And Simulation (ECMS 2005), Riga, Letónia. Ed. Yury Mercuryev et al. Junho de 2005. pp. 452-461

Dias, Luis S; Pereira, Guilherme A. B. and Rodrigues, A. J. M. Guimarães (2006), “A Shortlist of the Most ‘Popular’ Discrete Simulation Tools”, ASIM 2006 - 19th Symposium on Simulation Technique. SCS Publishing House. Ed. M. Becker and H. Szczerbicka. Hanover, Alemanha. pp. 159-163.

Dias, Luis S.; Pereira, Guilherme A. B. and Rodrigues, A. J. M. Guimarães (2006), “Activity based modeling with automatic prototype generation of process based arena models”, EMSS 2006 - 2nd European Modeling and Simulation Symposium. Barcelona, Espanha. pp. 287-296

Gupta, V. and Rao, T. (1996), "On the M/G/1 machine interference model with spares," European Journal of Operational Research, vol. 89, pp. 164-171.

Gupta, S. M. (1997), "Machine interference problem with warm spares, server vacations and exhaustive service," Performance Evaluation, vol. 29, pp. 195-211.

Ingalls, Ricki G. (2001), "Introduction to Simulation", in Proceedings of 2001 Winter Simulation Conference, B.A. Peters, J.S. Smith, D.J. Medeiros, and M. W. Rohrer, eds.

Kelton, W. David; Sadowski, Randall P. and Strurrok, David T. (2004), "Simulation With Arena", (3rd edition), McGraw-Hill, (1998-2004).

Kuei, C. H. and Madu, C. N. (1994), "Polynomial metamodelling and Taguchi designs in simulation with application to the maintenance float system," European Journal of Operational, Reseach, vol. 72, pp. 364-375.

Lopes, Isabel S. (2007), "Técnicas Quantitativas no Apoio à Decisão em Sistemas de Manutenção", Tese de Doutoramento, Universidade do Minho.

Lopes, Isabel S.; Leitão, Armando L. F. and Pereira, Guilherme A. B. (2005), "Modelo de Custos de Manutenção para um Sistema com M Unidades Idênticas", in C. G.Soares, A. P. Teixeira e P. Antão (eds), Análise e Gestão de Riscos, Segurança e Fiabilidade 2: p 603-620, Lisboa: Edições Salamandra (ISBN 972-689-230-9).

Lopes, Isabel S.; Leitão, Armando L. F. and Pereira, Guilherme A. B. (2006), "A Maintenance Float System with Periodic Overhauls", in Guedes Soares & Zio (eds), Safety and Reliability for Managing Risk 1: p 613-618, London: Taylor & Francis Group (ISBN 0-415-41620-5).

Madu, C. N. and Kuei, C. H. (1992a), "Group screening and Taguchi design in the optimization of multi-echelon maintenance float simulation metamodels," Computers and Operations Research vol. 19, pp. 95-105.

Madu, C. N. and Kuei, C. H. (1992b), "Simulation metamodels of system availability and optimum spare and repair units," IIE Transactions, vol. 24, pp. 99-104.

Madu, C. N. and Lyeu, P. (1994), "On the use of simulation metamodeling in solving system availability problems," Microelectronics Reliability, vol. 34, pp. 1147-1160.

Madu, I. E. (1999), "Robust regression metamodel for a maintenance float policy," International Journal of Quality & Reliability Management, vol. 16, pp. 433-456.

Mamede, Nuno (1984), "Simulação Digital de Processos", Tese de Mestrado em Engenharia Electrotécnica e de Computadores (Telecomunicações e Computadores), Instituto Superior Técnico.

Pegden, C.D.; Shannon, R. E.and Sadowski, R. P. (1990) "Introduction to Simulation Using SIMAN, McGraw-Hill", New York, USA. v. 2. 1990.

Pidd, Michael (1989), "Computer Modelling for Discrete Simulation", (Editor), Wiley.

Pidd, Michael (1993), "Computer Simulation in Management Science", Third Edition, Wiley.

Rodrigues, Guimarães and Carvalho, Valério (1984), "CAPS - ECSL, Experiência de modelagem e simulação aplicada a um sistema de elevadores", Relatório Técnico, Universidade do Minho.

Rubinstein, Reuven Y. and Melamed, Benjamin (1998), "Modern Simulation and Modeling", Wiley Series in Probability and Statistics, Applied Probability and Statistics Section, A Wiley-Interscience Publication, John Wiley & Sons, INC., ISBN 0- 471-17077-1.

Shannon, Robert E. (1998), "Introduction to the Art and Science of Simulation", in Proceedings of 1998 Winter Simulation Conference, D.J.Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, eds.

Shankar, G. and Sahani, V. (2003), "Reliability analysis of a maintenance network with repair and preventive maintenance," International Journal of Quality & Reliability Management, vol. 20, pp. 268-280.

Vik, P.; Pereira, G. and Dias, L. (2009), “Software Tools Integration for the Design of Manufacturing Systems”, Proceedings of the Industrial Simulation Conference, Loughborough, eds. Diganta Bhusan Das, Vahid Nassehi and Lipika Deka, pp 127-131

Vik, P.; Dias, L. and Pereira, G. (2010), “Automatic Generation of Computer Models Through the Integration of Production Systems Design Software Tools”. ASMDO2010 - Third International Conference on Multidisciplinary Design Optimization and Applications. Paris (France)

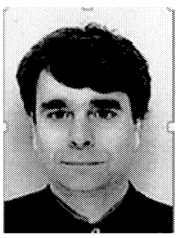
Vik, P.; Dias, L.; Pereira, G. and Oliveira, J. (2010), “Improving Production and Internal Logistics Systems – An Integrated Approach Using CAD and Simulation”. ILS2010 - 3rd International Conference on Information Systems, Logistics and Supply Chain - Creating value through green supply chains. Casablanca (Morocco), April 14-16.

Zeng, A. Z. and Zhang, T. (1997), "A queuing model for designing an optimal three-dimensional maintenance float system", Computers & Operations Research, vol. 24, pp. 85-95.

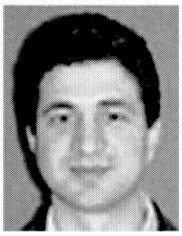
AUTHOR BIOGRAPHY

FRANCISCO PEITO was born in 1966 in Macedo de Cavaleiros, Portugal. He graduated in Mechanical Engineering-Management of Production in Polytechnic Institute of Porto. He holds an MSc degree Industrial Maintenance in University of Porto. His main research interests are Simulation and Maintenance.

GUILHERME A B PEREIRA was born in 1961 in Porto, Portugal. He graduated in Industrial Engineering and Management at the University of Minho, Portugal. He holds an MSc degree in Operational Research and a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. His main research interests are Operational Research and Simulation.



ARMANDO LEITÃO was born in 1958 in Porto, Portugal. He graduated in Mechanical Engineering at University of Porto, Portugal. He holds an MSc degree in Production Engineering and a PhD degree in Production Engineering from the University of Birmingham, UK. His main research interests are Reliability and Quality Maintenance.



LUÍS M S DIAS was born in 1970 in Vila Nova de Foz Côa, Portugal. He graduated in Computer Science and Systems Engineering at the University of Minho, Portugal. He holds an MSc degree in Informatics Engineering and a PhD degree in Production and Systems Engineering from the University of Minho, Portugal. His main research interests are Simulation, Operational Research and Systems Visual Modeling.

A DECISION SUPPORT SYSTEM FOR A FLEXIBLE TRANSPORT SYSTEM

Ana Dias
José Telhada
Maria Sameiro Carvalho
University of Minho, Systems and Production Department, Campus Gualtar, 4710-057 Braga, Portugal
E-mail: anacecidias@hotmail.com
{sameiro, telhada}@dps.uminho.pt

KEYWORDS
Flexible transportation systems, demand responsive transportation, transportation, decision support system.

ABSTRACT

The provision of traditional public transport services in rural areas have shown to be very inefficient and ineffective. In fact, rural areas are typically characterized by low levels of population density leading to complex demand patterns (low levels and high spatial and temporal dispersion), which in turn leads to low levels of service of conventional transport services (low frequencies, usage of old vehicles, etc). Demand Responsive Transport (DRT) systems have been seen as an effective alternative solution already adopted in several countries. In most cases, however, the services provided by the implementation of the system have revealed to be somewhat inadequate or even unsustainable, and therefore they needed to be either substantially redefined or abandoned. The principal reason for this is the lack of a proper design of the system before undergoing its implementation. The problem is that the viability and sustainability of a DRT system is, in general, highly dependent on the correct tuning of some organizational and functional parameters, namely all those related to the level of flexibility of the services provided. This design work must be carried out before the implementation phase of the system, by using adequate demand and supply models and related methodological approaches to accurately estimate the effects of alternative DRT solutions. Despite the existence of a vast literature concerning DRT systems, there is currently a lack in terms of comprehensive approaches to tackle this problem. In this way, this research proposes an integrated multi-disciplinary framework to support decision-makers in the design and planning of flexible transportation systems.

INTRODUCTION

Traditionally, public transport systems in rural areas have been based on static services: fixed routes, fixed stops and fixed schedules. However, the low levels of density observed in some of these areas leads to high levels of inefficiency and efficacy. In fact, low rates of occupation of vehicles and high rates of population dissatisfaction are frequently observed. Moreover, vehicles used are the oldest of operator fleet, does not attracting many potential users.

Therefore, many potential users have difficulty to reach their desired destination, because either they are too far away from fixed route or they want to travel out of predefined schedule. As a result rural area inhabitants are seeing their mobility limited, increasing social exclusion of some population sectors, such as children, elderly and/or mobility impaired people.

One way to overcome some of these problems is a flexible transportation system based on flexible services and called Demand Responsive Transport (DRT). DRT has been adopted over the last decades all over the world and some cases have been reported in the literature. Authors are unanimous in stressing that the success of these transportation systems depends on the use of intelligence solutions to process trip requests, optimizes routes and schedules in order to respond in real time to users mobility needs. Such intelligent solutions require the use of modern information and communication technologies and operational research optimizations approaches either at planning and operational level. According to literature, one of the gaps associated to DRT systems (already implemented) is a framework that could provide an integrated decision support system to help decision makers on devising intelligent strategic solutions at the design phase.

In this research project, a decision support system to design and plan a DRT system in rural areas has been developed. The objective is to reproduce and test different decision-making alternatives in order to assess, in advance, the quality of alternative design scenarios or management strategies.

In fact, such a tool can provide what-if analyses required to achieve better planning decisions and will allow evaluating operating strategies prior to the implementation of such as a complex system as it the case of a DRT system. Furthermore, the tool will ultimately assure the adoption of a sustainable DRT system, by properly adequate supply (e.g. fleet of vehicles and typology of services provided) to estimated demand levels and patterns, taking into account financial, economic and social decision criteria.

The decision support system is part of a general framework that combines data with analytical tools in order to provide information that will support top management strategic decisions.

This paper is structured as follows. Section 2 presents a brief literature review on flexible transportation systems in order to highlight the main design issues concerning the design of DRT systems. In Section 3, a new framework to support strategic decisions concerning the design and planning of DRT systems is proposed. Finally, section 4 presents the main conclusions and final considerations regarding future developments.

LITERATURE REVIEW

Flexible transportation systems have been adopted over the last decades, as reported in some studies (Brake *et al.*, 2004; Mulley and Nelson, 2009). In fact, there are many successful cases of DRT systems all over the world. Most of the cases reported are serving rural areas, small towns or special populations (elderly and people with disabilities) in large cities.

These systems provide transport on demand from users, using flexible schedules and routes to pick up and drop off users as required. A DRT system receives trip requests either for an immediate service or as an advanced reservation and organizes routes and schedules to accommodate trip requests aiming to respond in real time to user's mobility needs. Its implementation typically involves the use of information and communication technologies as shown in Figure 1.

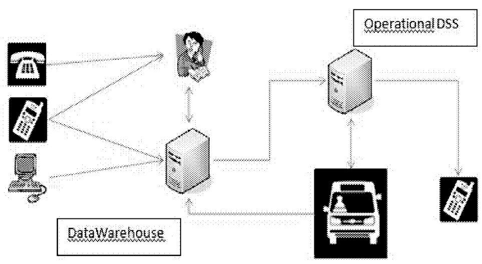


Figure 1: Elements of a demand responsive transport system.

Trip requests are, in general, made using telephone or internet and stored in a data warehouse which holds all the relevant data concerning the transportation network. A fleet of vehicles are available: buses, mini buses, taxis provided by a variety of providers (taxi owners, bus operators, community transport, etc). Services can be operated on their own or integrated with traditional transportation systems, acting as feeder services for buses or rail services. The heterogeneous fleet of vehicles is coordinated by a Travel Dispatch Centres (TDC) with advanced information and communication technologies such as on-board integrated GPS, continuous GPRS connection to TDC. An operational decision support system (DSS) incorporating mathematical models (optimization, simulation and statistical methods), will provide the most adequate transport solution according to the area characteristics and demand pattern.

If possible, a flexible transportation system will combine the advantages of a private car solution, characterized by high levels of commodity and responsiveness, with economies of scale of public transportation systems. In practice, different levels of flexibility can be adopted and customer demand determines the route and vehicle used at varying timescales prior to travel as illustrated in Figure 2. This figure shows the range of options for these aspects of demand responsiveness.

It is essential to realize the role of each flexible typology of transport service as part of the overall public transport system. DRT system is flexible in terms of route, vehicle allocation, vehicle operator, type of payment and passenger category. The flexibility of each element can vary along a continuum of demand responsiveness from services where all variables are established a considerable time before the operation (*e.g.* a conventional fixed public transport bus route) to services where variables are determined close to the time of operation (Brake *et al.*, 2007).

The development of Intelligent Transport Systems (ITS) tools, as well as the availability of mobile communications, has allowed new public transport service options to be developed whereby the service is more responsive to customer demand in terms of time and space (Mageean and Nelson, 2003; Ambrosino *et al.*, 2004; Brake *et al.*, 2004).

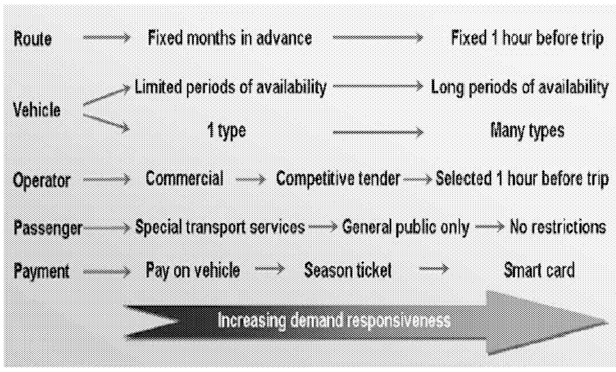


Figure 2: Different levels of flexibility (adapted from Brake *et al.,* 2007).

Different types of contributions can be found in the literature concerning flexible transportation systems. In general, they can be grouped in three categories: general articles, case studies, and analytical models. Although, there are some review papers reporting good practices and difficulties, the bulk of the literature is related to general issues and case studies (*e.g.* Enoch *et al.*, 2006).

General issues papers address advantages and disadvantages of DTR systems or conceptual questions related with flexible transportation systems, as well as critical factors for their success or failure in practice (*e.g.* Giannopoulos, 2004; Mulley and Nelson, 2009).

Case studies papers report experiences of DRT implementation and discuss main results achieved benefits, problems and limitations of adopted solutions (*e.g.* SAMPLUS, 2000; Gray *et al.*, 2001; McDonagh J., 2006).

Finally, modeling papers are, by far, the most numerous, and propose mathematical models to address some of the most complex problems faced in the operational management of these transportation systems such as network routing and scheduling, demand forecasting models, simulation models, etc (*e.g.* Parragh *et al.*, 2008; Quadrifoglio and Li, 2009).

Interesting review articles have been proposed by Savelsbergh (1995), Desaulniers *et al.* (2000), and Cordeau *et al.* (2007).

From some DRT systems reported in literature, its implementation obeys to several conditions, encompassing high dynamic levels of both planning and coordination processes. The use of modern information and communication technologies, transport telematics/ Intelligent Transport Systems or even, Flexible Transportation Services (FTS) allied to adequate strategy planning services has been pointed out as the solution to improve the costs-effective performance of DRT services (Mulley and Nelson, 2009).

Although DRT is a very interesting solution since it offers a very user-friendly solution to their potential passengers (small and comfortable vehicles on demand) and overcomes some of the shortcomings of the traditional transport services, its use has been limited due to the relatively high costs of operation as a result of high levels of resources (vehicles and drivers) and to the relatively high complexity associated to planning and operational issues. Additionally other issues have been pointed out by several authors (Brake *et al.*, 2004) which can have a significant impact on its implementation: the lack of legal and regulatory framework in most countries to accommodate flexible transport services; technological issues including the selection of algorithms to provide operationsl solutions; sustainability of DRT services and the lack of support to select the most appropriate service and system design.

In this research project a framework is proposed to address some of the main difficulties associated to the design and management of a DRT system. A methodology is suggested involving some of the key issues that must be taken into account when defining the most adequate transportation solution according to the area characteristics, population demographics and legal and regulatory framework.

CONCEPTUAL FRAMEWORK

As it was already stated, the aim of the underlying project of this paper is the development of a methodology to help the design and management of efficient and sustainable flexible transportation systems, whose services combine the flexibility of a private car with the economies of scale of a public transportation system. Such a DSS will be essential to minimize the increased operational costs connected to the requirements associated to a flexible transport system.

In the proposed framework, the conception of a DRT system involves a concurrent engineering approach since it allows to integrate and analyze all the main design, planning and operational issues (economical, social, environmental) at its conceptual stage. This includes establishing user requirements and expectations, by considering perspectives altogether from the beginning of the project, running computational models (software has a huge role in the design process) and creating prototypes.

By adopting a concurrent engineering approach, the DRT system solution (for a particular operating area) will integrate technical and non-technical cross-functional cooperation in order to produce better services for users and more efficient and sustainable operations for operators.

Based on demand and supply models of the area where a DRT system is to be implemented, for a given level of resources and a set of operating rules, the system simulates the real system and will produce a set of performance indicators associated to the tested scenario.

Additional analysis is then performed to assess solution viability and encompassing several dimensions such as: technical, financial and economical. A sustainable transportation system must be achieved to guarantee that population needs are met now and for generations to come. The result of the assessment process will provide guidelines and the required feedback to adjust system resources and operating parameters as illustrated in Figure 3.

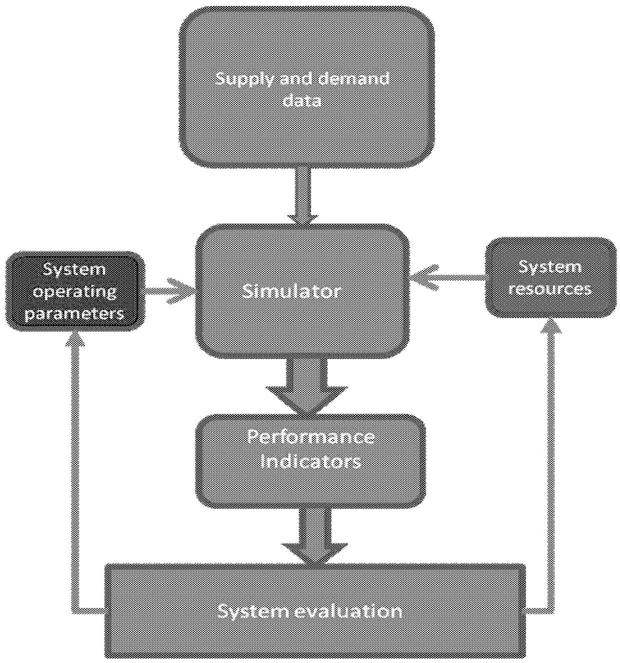


Figure 3: Conceptual framework.

The simulator constitutes one of the key elements of this framework since it will be responsible for reproducing the real system behavior and provide essential information for system evaluation process. The simulation model (Figure 4) incorporates a data base with data concerning travel demand – origin/destination information and geo-referenced data on the network.

Based on these data, a trip request generator reproduces users mobility needs for a given period.

Then, operational configuration must be defined by setting objectives, operating rules and level of resources. Objectives can include total cost minimization (a generalized distance/time cost function or multi-objective function can be used), minimizations of delays, minimization of vehicles, among others.

Different operation rules can be adopted associated with different levels of flexibility: fixed routes, fully flexible routes or a mixed solution, minimum advance booking time, level of articulation between DRT and fixed transportation systems, pick up and drop off points, dynamic or static planning, etc. In terms of resources, it can be assumed that available vehicles can have different characteristics, such as taxis, mini buses or buses.

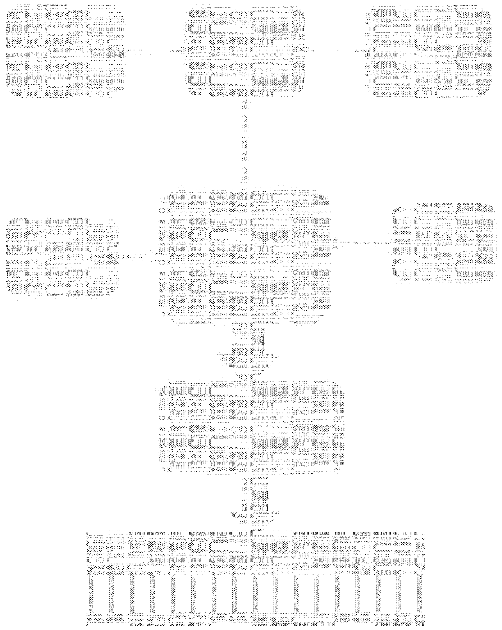


Figure 4: Simulation model of the DSS

The simulation model uses optimization routines to produce solutions (routes/schedules/drivers plans/users notifications) as well as, printed reports on systems performance. In fact, several performance indicators can be produced in order to provide insight on systems operations: total generalized costs of trip plan, medium delay of each vehicle; mean users delay time; mean waiting time; vehicles utilization; requests not satisfied; level of service

Two different types of information can be generated: strategic information crucial in a design stage of DRT system to set up the correct level of resources and there is operational information required to define the most adequate operational rules to be adopted in a daily operations context.

The last step, and undoubtedly one of the most challenging ones, is to carry out systematic system evaluation of the transportation system, setting up the most important evaluation criteria and their quantification. Specificities of theses transportation systems must be taken into account: spatial aspects of public transport such as the characteristics of the local population, transport network, the patterns of commuters, and the framework within which the system works, determine the demand and operational scale of system and, as such, can affect performance and efficiency. Eventually, the primary purpose of a public transportation system is to serve people who chose not to use their own vehicle due to their low income, to some resource constraints or to other personal reasons. Globally, these projects aim to improve customer satisfaction, with regard to their mobility.

Performance measurement is a procedure to monitor progress toward a result or goal. Based on the literature on transport performance measurement systems (NCHRP, 2006), there are a large number of measures that can be grouped in four categories: preservation of assets; mobility and accessibility; operations and maintenance and safety.

Traditional methods to economic evaluation of transport projects have been based on two approaches: public and private. The financial assessment, which evaluates the point of view of the expected transport operators' profitability, naturally refers to private operators. It is therefore assumed that the DRT must be efficient, and the use of telematics combined with optimization models is fundamental. This results on the consideration of Intelligent Transportation Systems (ITS). On the other hand, public or economic evaluation focuses on Regional, Local or National interest of a project. The environmental and social impacts must then be considered. Particularly, in DRT projects, social assessment stands out as fundamental, addressing the efficiency for all stakeholders, including aspects like customer acceptance, impact on traffic, environmental and socio-economic impacts, and other externalities. This evaluation is often based on cost-benefit analysis.

According to the literature, efforts have been made in order to identify the potential benefits with the least impact on costs, but the convergence of methods for costs and benefits assessment turns difficult to measure or evaluate impacts in monetary terms. This obviously implicates a set of performance indicators that can be ideally translated into a monetary scale. However this is a complex task and additional research is needed to develop models and better evaluate the socio-economic development of ITS. Furthermore, some authors refer the absence of sufficient information to make a quantitative analysis of transport services. This information can be obtained through surveys of potential users of the service.

CONCLUSIONS

Prior to the implementation of a DRT system there are a lot of issues that must be properly addressed at the design and planning phases of its project. Many DRT projects has been implement world-wide without taken adequate care of such issues, and therefore some related failures (that have led to system re-engineering or even project withdrawal) have been reported. In addition, there is currently a lack of comprehensive methodologies to address the problem of designing and planning such systems.

In this research, an attempt is made to develop a framework that will provide an integrated decision support system to enable decision makers to perform systematic analysis leading to intelligent strategic solutions. The proposed approach includes a realistic micro-simulator of alternative transportation scenarios and functional parameterization (*e.g.* flexibility and operating rules) allowing planners to estimate accurately the correspondent performance indicators. Additionally, there is an evaluation module to incorporate sustainability of alternative transportation systems, including financial and economic viability and sustainability. Following a concurrent engineering approach, the adoption of this methodological framework will certainly contribute to a better understanding of the key issues in the design of DRT systems.

Further developments of this project includes the enrichment of the simulation model by incorporating more efficient routing and scheduling algorithms, using different solution methods (exact algorithms, heuristics, meta-heuristics), and providing the system with the cleverness to adopt the most appropriate models according with the decision context. Furthermore, this simulation tool will be integrated to Geographic Information Systems (GIS) in order to enhance the graphical displaying of the solutions produced, and to allow further statistical analyses of spatio-temporal indicators. A case study in a small rural area of the north of Portugal will be used to validate the framework.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the European Commission and the Portuguese Government through the research grant FCT/TRA/72871-2006.

REFERENCES

Ambrosino, G., Nelson, J., Romanazzo, M. (Eds.), 2004, "Demand Responsive Transport Services: Towards the Flexible Mobility Agency" ENEA, Rome.

Brake, J., Nelson, J., and Wright, S., 2004, "Demand responsive transport: towards the emergence of a new market segment", *Journal of Transport Geography*, 12, 323-337.

Brake, J., and Nelson, J., 2007, "A case study of flexible solutions to transport demand in a deregulated environment", *Journal of Transport Geography*, 15, 262-273.

Brake, J., Mulley, C., Nelson, J., and Wright S., 2007," Key lessons learned from recent experience with Flexible Transport Services", Transport Policy 14, 458–466.

Cordeau, J., and Laporte, G., 2007, "The dial-a-ride problem: models and algorithms", *Ann Operational Research*, 153, 29–46.

Desaulniers, G. and Villeneuve, D., 2000, "The shortest path problem with time windows and linear waiting costs", *Transportation Science*, 34 (3), 312-319.

Enoch, M., Ison, S., Laws, R., and Zhang, L., 2006, "Evaluation Study of Demand Responsive Transport Services in Wiltshire", Transport Studies Group Department of Civil, Building Engineering, Loughborough University, Leicestershire LE11 3TU, United Kingdom.

Giannopoulos G., 2004, "The application of information and communication Technologies in transport", *European Journal of Operational Research*, 152, 302-320.

Gray D., Farrington J., Shaw J., Martin S., and Roberts D., 2001, "Car dependence in rural Scotland: transport policy, devolution and the impact of the fuel duty escalator", *Journal of Rural Studies*, 17, 113–125.

Mageean, J., and Nelson, J., 2003, "The evaluation of demand responsive transport services in Europe", *Journal of Transport Geography* 11 (4), 255–270.

McDonagh J., 2006, "Transport policy instruments and transport-related social exclusion in rural Republic of Ireland", *Journal of Transport Geography*, 14, 355-366.

Mulley C., and Nelson J., 2009, "Flexible transport services: A new market opportunity for public transport", *Research in Transportation Economics*, 25, 39-45.

NCHRP, 2006. *Performance Measures and Targets for Transportation Asset Management.* USA. Transportation Research Board.

Parragh N., Doerner K., and Hartl R., 2008, "A survey on pickup and delivery problems, Part II: Transportation between pickup and delivery locations", *JfB*, 58, 81-117.

Quadrifoglio, L., and Li, X., 2009, "A methodology to derive the critical demand density for designing and operating feeder transit services", *Transportation Research Part B: Methodological*, 43, 922-935.

SAMPLUS, 2000, "A Basic System Architecture and Technical Solutions for DRT", WP4, Deliverable 4.1, version 4, In: ftp://ftp.cordis.europa.eu/pub/telematics/docs/tap_transport/samplus_d4.1.pdf. Date: 28 July 2008.

Savelsbergh, M. W. P., and Sol, M. 1995. The General Pickup and Delivery Problem, *Transportation Science*, 29, 17-30.

AUTHORS BIOGRAPHY

Maria Sameiro Carvalho is graduated in Computer and Systems Engineering (University of Minho – Portugal, 1986) and holds a MSC degree in Transportation Planning and Engineering (University of Leeds - UK, 1990) and a PhD degree in Transport Planning and Engineering (University of Leeds - UK, 1996). Her main research interests are in Transportation, Supply Chain Management and Operational Research.

José Telhada holds a PhD degree in Transport Planning and Engineering (ITS-University of Leeds – UK, 2001). His main research interests are Operational Research and Logistics applied to transportation systems.

Ana Dias is graduated in Matematics (University of Minho (2004) and Master in Industrial Engineering (University do Minho, 2008). At the present she works as research student at University of Minho.

AI APPLIED TO CONCURRENT ENGINEERING

Representation of Uncertainty in Distributed Product Development

Andre Sprenger
Lucia Mosch
Kai Mecke
Reiner Anderl

Dept. of Computer Integrated Design
Faculty of Mechanical Engineering
Technische Universität Darmstadt
Darmstadt
Germany

email: {sprenger|l.mosch|mecke|anderl}@dik.tu-darmstadt.de

KEYWORDS
Uncertainty, Engineering Data Management, Ontology, Product Development, Processes.

ABSTRACT

The knowledge about future process properties of a product which is in development is essential for the success of the product. The scattering and bias of the real products process properties leads to uncertainty. This uncertainty has negative economic and ecologic consequences, such as failure and oversizing.
In this paper we present a new approach which allows the representation of uncertainty of processes in different product life phases for the distributed product development. Based on semantic technology an ontology for the uncertainty in structural mechanic systems is developed.
Furthermore possible approaches to present uncertainty information in 3D-CAD-systems will be outlined.

INTRODUCTION

The increasing competition forces manufacturers of complex products to develop their products faster on a high quality level. The product development process is a very skill and knowledge intensive task. A lot of precise information is needed. High complexity levels and strong interdisciplinary product development enforce the need of sophisticated data and information management. The available information has different characteristics due to its information content, e.g. from rough estimations to high sophisticated stochastic experimental data. This information is valuable but expensive to gather. On the other hand a maximum of information is needed to reduce failure or high safety coefficients. The information needed in the product development refers to real load scenarios and real manufacturing processes and their properties. There are numerous properties in different domains which interact. Semantic technologies and ontologies are a new approach for the management of information and knowledge, and are a subject of current research activities. These technologies allow to represent relationships between entities in a graph and allow to add a meaning to these relationship which can be evaluated by an inference engine. This contribution gives a brief overview of the contemporary product design process and the use of ontologies in this field. The state of the art in the field of mechanical engineering ontologies is presented as well as a system to represent and exchange uncertainty based on these ontologies. Further a system supporting future product development with detailed information and knowledge about uncertainties and their effects is proposed.
Process representation is based on an advanced process model, which offers the ability of considering uncertainty. The elements of this process model will be introduced. Following this, the benefit for an engineer using this ontology-system for a visualization will be illustrated in an example.

STATE OF THE ART

Lifecycle

In **product development** uncertainty about future process properties of the product is a common problem. For a comprehensive view over the products life different lifecycle models are used. The information product life cycle focuses the information generated during the products life phases like development, production, usage and recycling. This holistic approach makes it possible to collect data from the production of raw materials, production, usage and finally the recycling of the product (Abele et al. (2005)).
The product development is a complex process where

information is processed to make decisions about the future product. In modern distributed product development there are a lot of participants which are distributed spatial and by their disciplines. These people have to collect, share and process information and make decisions. The quality of the base data is often unknown to the subsequent engineer. So the consequences of a bias or scattering of a parameter cannot be regarded in the decision process although tools and methods are available (Yassine et al. (2008)).
The solution for this problem is to get ensured information about future process properties in order to have a more realistic product development process. Contemporary product development uses numerous methods and tools to analyze the impact of failures and the reduction of complexity (Lindemann et al. (2009), Pahl et al. (2007)).
In the field of **manufacturing** many research activities target the enhancement of the production processes. The bias and scatter of the produced products is regarded with tolerancing and process capability methods (Roy (1990)). Takeguchi introduced a loss function which couples scatter and deviation with loss. From this approach several methods called "'six sigma"' methods were developed e.g. the "Design for Six Sigma" (Reh et al. (2005)).
During the **utilization** of products new monitoring techniques allow to capture real processes and the products behavior with different operational conditions. New technologies and monitoring techniques increase the need for the enhancement of data management in this field (Bertsche (2008)).

Uncertainty in Engineering

Uncertainty in the field of engineering is defined as the gap between the present state of knowledge and absolute certainty. There are several approaches to classify uncertainty in the engineering domain. Common to these is that with the growth of knowledge in the regarded case the uncertainty decreases. Two mayor types of uncertainty are regarded in the field of engineering (Lindley (2006)). The aleatoric uncertainty is the part of uncertainty which arises from stochastic effects like random noise in measurement. The common opinion about these says that this uncertainty can only be reduced by changing the basic processes. It cannot be reduced by additional information. Epistemic uncertainty arises from a lack of knowledge and is reducible by additional knowledge (Knetsch (2003)). This kind of uncertainty can arise through weak simulation models. Knowledge about deficits and further improvement can reduce this uncertainty (Nikolaidis (2005)).
In our approach a three tier model (presented in Engelhardt et al. (2010)) is used which was developed in the Collaborative Research Centre (CRC) 805 "Control of Uncertainties in Load-Carrying Structures in Mechanical Engineering" founded by the German Research Foundation (refer to Hanselka and Platz (2010)). This model defines three types of uncertainty which are derived on engineering practice. With "unknown uncertainty" the uncertain property and its relevance for the regarded process is ignored, or the possible relevance is only guessed. With the type "estimated uncertainty" a mechanism of cause and effect is known but could not be quantified exactly by stochastic data. In engineering this kind of uncertainty is often represented as tolerances or intervals. Following the distinction between epistemic and aleatoric uncertainty the lowest level of uncertainty can be described by a distribution. This is called "stochastic uncertainty" in the model (refer to Engelhardt et al. (2010) for details).

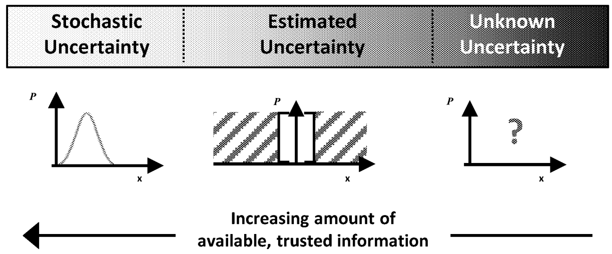


Figure 1: Uncertainty Model, with respect to Engelhardt et al. (2010)

Semantic Technology

According to (Gruber (1993)) Ontologies are formal models of selected aspects of the real world . So ontologies are a tool to represent things (real or abstract) and their relations in a computer processable way. These models consist of a collection of concepts, properties and relationships between the knowledge concepts (Kashyap et al. (2008)). Ontologies allow to assign higher semantic relations to describe knowledge. The utilization of defined knowledge renders the possibility to apply rule-based deduction of information (Kashyap et al. (2008)). The Web Ontology Language (OWL)[1] is an ontology language standardized by the World Wide Web Consortium (W3C). OWL defines a set of relations which can be interpreted by an interference engine. Depending on the amount of supported/used relations there are three types of OWL expressiveness levels. OWL lite and OWL DL are decidable. Most interference engines support the OWL DL syntax. OWL Full as the third type is not decidable. A OWL ontology consists of Classes, Individuals, Data Properties and Object Properties. Data Properties assign values to an Instance and follow the XML Scheme. Object properties assert Individuals to Individuals. In order to ease the work with ontologies and to enhance the expressivity rules can be included.

[1] http://www.w3.org/2007/OWL/

The rule interchange format was recommended by the W3C to exchange different rules for ontologies.
For the representation of processes different ontologies have been modeled. For manufacturing processes the American "National Institute of Standards and Technology" specified the "Process Specification Language" which allows the formal representation of processes. The W3C released a recommendation for an ontology which defines the functions of web services called OWL-S which contains a service model which can also be used as process model. OWL2 is the successor of OWL and has the state as an W3C recommendation. A common editor for OWL and OWL2 is Protegé in the version 4.1beta. This editor implements the OWL2 recommendation (Stanford School of Medicine (2010)). Other tools and Application Programming Interfaces only support the OWL DL specification.

Related Work

In the domain of mechanical engineering several ontology based approaches were published. In this knowledge intensive field an effective knowledge management is necessary. In (Pavkovic and Storga (2006)) four modeling levels for product development ontologies are presented. The epistemological level builds the upper level with a general ontology, the second level builds the domain ontology, which contains the product design and the development process ontology. The application level contains domain specific information like configuration or assembly. The last level is the project level which contains project specific knowledge.
Another concept is presented in (Lee and Suh (2007)). It proposes a three level product ontology architecture with an Meta-Product Ontology, Generic Product Ontology with a particular ontology which describes a specific product.
A generalized upper ontology is the SUMO approach from (Niles and Pease (2001)). Another upper ontology approach is the DOLCE ontology from (Masolo et al. (2003)). A light version of of the DOLCE ontology is the "DOLCE-UltraLite". An overview over upper ontologies is given in (Mascardi et al. (2007)). In (Patil et al. (2005)) a product semantic based on OWL called (PSRL) is presented. The PSRL serves as an interlingua for the integration of product semantics into PLM especially for the 3D-CAD domain. The field of assembly design is the scope of (Kyoung et al. (2006)). They present an approach for the exchange of product assembly semantics in the product development.
For the support of the methodology in product development process an approach for the FMEA, called OntoFMEA is presented in (Dittmann (2007)). It describes the support of the usage of the Failure Mode and Effects Analysis with an ontology based system. There are more general ontologies which do not describe a special aspect of product development but are useful in general. The "Semantic Web for Earth an Environmental Terminology" (SWEET)[2] from the NASA Jet Propulsion Laboratory provides ontologies which describes a wide variety of aspects e.g. properties processes and states. A simmilar approach is the QUDT Ontology QUDT-Ontology[3]. It represents values and quantities in an ontology. A comparison of ontologies in engineering domain is given in (Kim and Bracewell (2008)).

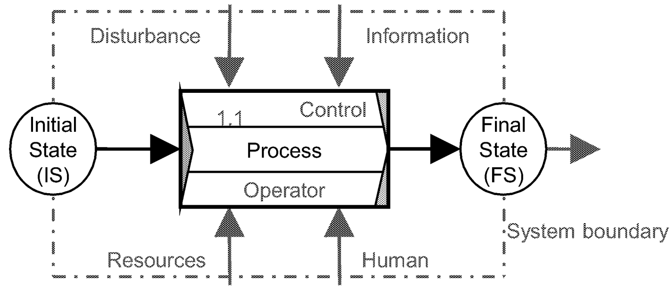


Figure 2: Process Example

UNCERTAINTY PROCESS ANALYSIS

To show where and how uncertainty during conducting an activity or a sequence of activities is occurring, a representation of processes is necessary. The goal of this representation is to visualize and analyze technical processes and their occurrence of uncertainty, to consider their interdependencies and afterwards to control the uncertainty. In the Collaborative Research Centre 805 a strategy for describing processes and systems as well as their properties including uncertainty was developed. The developed strategy emphasizes the working hypothesis of the CRC 805: Uncertainty occurs in processes and can be interpreted as a deviation of process results which is described as final states (Hanselka and Platz (2010), Engelhardt et al. (2010)). The strategy describes a process model considering uncertainty.
After analyzing current modeling methodologies for describing processes like Structured Analysis and Design Technique (SADT) (Marca and McGowan (1987)), especially IDEF0 (IDEF0 (1993)), EPK (Keller et al. (1992)) and eEPK (Scheer (1996)), the modeling language for description of distributes systems like Petri nets (David and Alla (2005)) the process model of Heidemann (Heidemann (2001)), the Uncertainty Process Analysis (UPA) was developed (see Fig. 2). It contains elements describing **states**, **processes** and **influencing variables**. UPA uses a top-down approach; a hierarchy presentation is possible.
Elements describing the **states** represent inputs and outputs of processes and are described by a state vector containing all properties.

[2] http://sweet.jpl.nasa.gov/
[3] http://www.qudt.org/

Elements describing **processes** include:

- Process activity (describing the act of engineering),
- Control (representing commands),
- Operator (characterizes objects to operate a process) and
- System indicator (symbols the level of the process).

Influencing variables describe the source of uncertainty and are divided into:

- Disturbance,
- Information,
- Resources and
- Human (action).

In this context two different views for the presentation of the process model are possible: process-oriented and state-oriented view. Depending on the focusing it is to choose between the process-oriented view, focused on the process level with its sub-elements and actuating variables, and the state-oriented view, focusing on the specific state.
During virtual product creation the information flow is based on single values. During real processes like production and usage these values can scatter (deviation), that means it exists a range of possible values or a distribution function is given. The goal is to reach a virtual product creation close to reality by providing these information of real processes.
In this fact, the process model has the goal, to represent and collect information about processes, states, uncertainty and the effect of uncertainty. The process model gives a look on processes effected by process properties, like disturbance, information, resources and human. If these process properties are uncertain an effect on product properties (states) exists.

APPROACH

As described in the previous section the representation of uncertainty makes it necessary to regard the whole product life cycle. Uncertainty arises in the processes of the product life, according to manufacturing and usage. Existing approaches do not provide the ability to cope with this problem. Thus it is necessary to provide trusted information about these processes as early as possible in the product development process to cope with the uncertainty. To get this information it is necessary to know in which processes the product is involved in the future, and to have information how the processes behaved in the past, in order to infer to the future processes (figure 3). This allows a more realistic product and process development. The knowledge about the

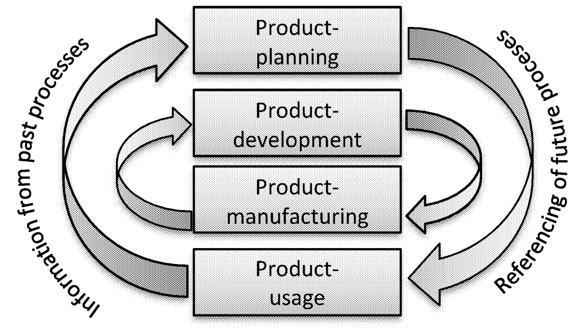


Figure 3: Process Information

processes allows to analyze which process properties or product properties influence a certain property. This can be used to analyze chains of uncertainty, here called uncertainty aspects.
An ontology based model was developed to represent and collect information about processes, states, uncertainty and the effect of uncertainty. The ontology has three main parts describing the product, the processes and uncertainty arising in those processes.
As described before the ontology represents generalized information about processes and uncertainty like project specific information, as well as different parts. This is realized by different sub-ontologies which are imported into the current project ontology. This distributed approach allows easy maintenance and further development of the information base. Fig. 4 depicts an overview of the ontology architecture. The presentation layer builds a bridge from the different CAx-Tools to the Ontology Layer. The Ontology Layer contains the ontologies and builds the base for the inference engine. The Data Layer stores the non-semantic data, referenced by the Ontology Layer.

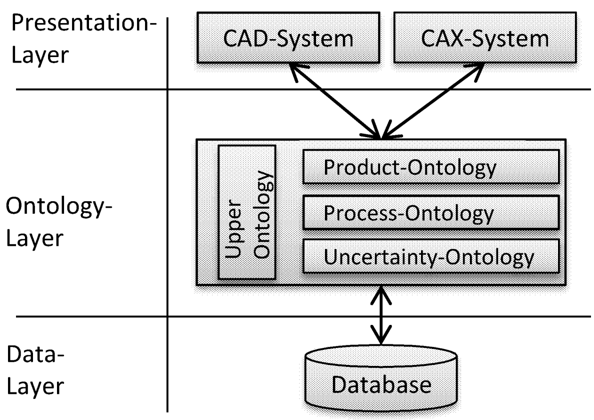


Figure 4: Ontology Architecture

The product ontology represents the emerging product by referencing the basic product architecture. This architecture is a skeleton of the particular product and is completed during the design process towards the final product structure. The usage of an architecture skeleton builds an initial semantic structure, and is possible within variant and adaptive designs. The product structure, namely the parts and assemblies of the product, is build and concretized by instantiating the skeleton and referencing the instances to assemblies respectively the product. Each part of the derived product structure is referenced to its geometry model created in a 3D-CAD System. To link processes and uncertainties to this structure it is necessary to have the possibility of referencing certain features and selected properties of the geometry. Thus uncertainty does not only arise in a single feature or geometry element, these can be aggregated to collections called design zones.
The process ontology describes the manufacturing and the usage process by a generalized information base. It contains the information about input, output, simulation model, critical process properties. This information base is populated by the monitoring of real processes so that a reuse or prediction for future processes is made possible. These elements build the foundation for an specific process model for a certain product.
Necessary processes in manufacturing are linked to their geometric pendants e.g. hole and drilling. Usage processes refer to the whole product, but only a set of properties is regarded. Each process has a preceding and a successive state which contain product an process properties. The states change from process step to process step. This serves as basic information for the reuse of information about uncertainty.
The uncertainty ontology contains generalized "uncertainty aspects" for a specific domain, in this case for load carrying systems. A product specific "single uncertainty" references the uncertain property with its value, the process which defines the property and the product geometry. This allows to track the direct sources of uncertainty by the use of an inference engine further indirect influences to the single uncertainty can be identified. The digital representation of the uncertain value of a parameter in the field of engineering is realized by an schema which contains one of five different representations shown in figure 5.

The value of an uncertain property can be represented as distribution, histogram, tolerance, interval, and a single value. The distribution and the histogram represent stochastic uncertainty in different qualities. Tolerance, interval and single value represent incertitude in different levels of uncertainty.
CAx-Tools and workflows do not consider the effects of uncertain processes in an comprehensively way. In product design for example CAD-systems often use and present exact product properties, like diameter or length of a component or feature like bore hole or axle. Information about real processes and the cause and effect of uncertainty is not given or supported. An approach to integrate information about uncertainties by annotations in a CAD-system was already presented in Mosch et al. (2010). The ontology based model was connected with a CAD-system. Within this work a known and given ontology was assumed. The presentation of uncertainty was realized on basis of a three-dimensional model in a CAD-system and extended by an uncertainty-browser. Part of the concept was the methodology of parametric and knowledge-based design. The uncertainty-browser acts as a graphical user interface to categorize the information about uncertainty. A view-concept enables a limitation of the number of information. Beyond that, information about uncertainty will be visualized as annotations referring to chosen property.

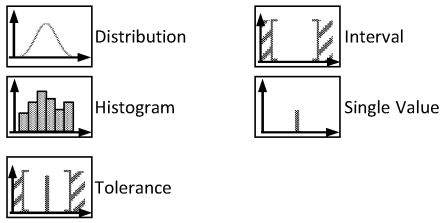


Figure 5: Informational Representation of Uncertain Property Values

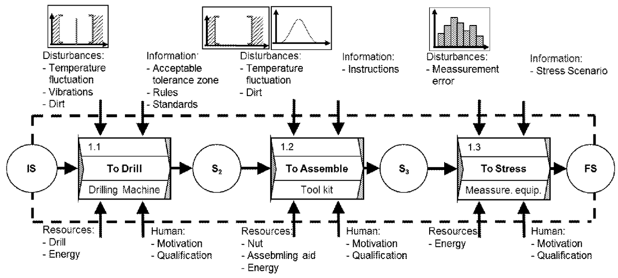


Figure 6: Process-oriented view

Illustrative Example

The validation of the concept is based on the developed process model. As example for a process chain the processes "to drill", "to assemble" and "to stress" are chosen. As described in the previous section the example contains all elements describing states, processes and influencing variables. Figure 6 shows the process-oriented view, focused on the process level. The influencing variables, mainly the disturbances like temperature fluctuation or dirt represents the uncertain process properties and the symbols illustrate the types of uncertainty. These properties are the source of uncertainty. For example, dirt in the process drilling can cause displacement of the drill. As shown in figure 6 ignorance has no concrete value; with incertitude a range of possible values is given; in case of stochastic uncertainty a distribution function or a histogram is given. Another example of this process model is shown in Platz et al. (2010). In the state-oriented view, the states, represented in the product properties (figure 7), create the context to the three tier model developed in the CRC 805, too. The icons also symbolize the types of uncertainty. Remain to the example in the process drilling with its disturbance dirt, in figure 8 the deviation of the product properties can be seen. The state describing the diameter of the bore hole is scattering, for example.

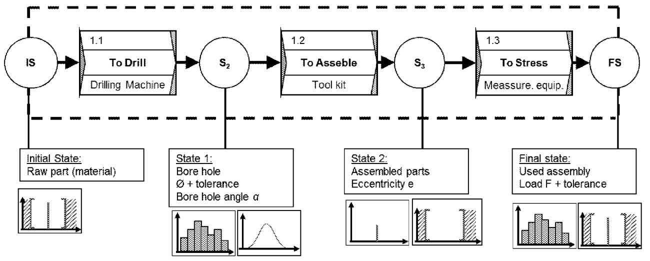


Figure 7: State-oriented view

The visualization is relized in a CAD-system an an annotation to the geometric model. Further information about processes and uncertainty are also given.

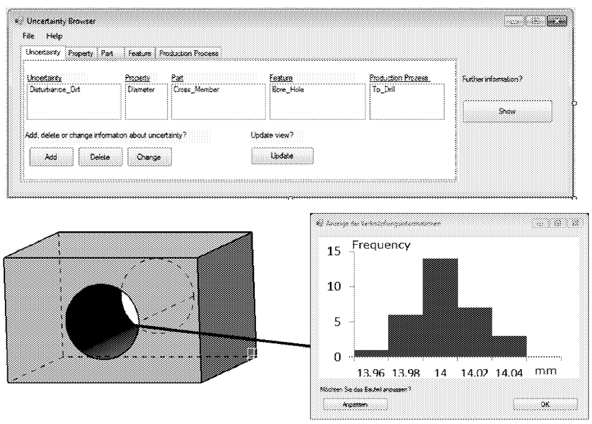


Figure 8: Graphical User Interface of the implemented System

CONCLUSION

In this paper an approach to represent uncertainty of process properties in complex product development processes is proposed. Uncertainty is a problem which affects all phases of the product life cycle and has to be regarded as early as possible for the development of save and economic products. In product creation the amount of information an its interdependencies is difficult to handle. With the regard of uncertainty this problem will increase. The approach shown in this paper addresses this problem with the evolving web-technology of ontologies. To get this information to the subsequent engineer a visualization of uncertainties in context is also shown. .

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG) for funding this project within the Collaborative Research Centre (CRC) 805. The development of an ontology based information model for uncertainty is part of the subproject A5 "Data Model for the Representation and Visualization of Uncertainties" of the Collaborative Research Centre 805 "Control of Uncertainty in Load-Carrying Structures in Mechanical Engineering" funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). For further information visit www.sfb805.tu-darmstadt.de.

REFERENCES

Abele E.; Anderl R.; and Birkhofer H., 2005. *Environmentally-friendly product development: methods and tools.* Springer Verlag.

Bertsche B., 2008. *Reliability in Automotive and Mechanical Engineering: Determination of Component and System Reliability.* VDI-Buch. Springer-Verlag, Berlin.

David R. and Alla H., 2005. *Discrete, continuous, and hybrid Petri nets.* Springer Verlag.

Dittmann L., 2007. *OntoFMEA: Ontologiebasierte Fehlermöglichkeits- und Einflussanalyse.* Deutscher Universitäts-Verlag — GWV Fachverlage GmbH Wiesbaden, Wiesbaden.

Engelhardt R.A.; Koenen J.F.; Enss G.C.; Sichau A.; Platz R.; Kloberdanz H.; Birkhofer H.; and Hanselka H., 2010. *A Model to Categorise Uncertainty in Load-Carrying Systems.* In *Proceedings of the 1st Internatinal Conference on Modelling and Management of Engineering Processes.* Springer Verlag.

Gruber T., 1993. *A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition Academic Press Inc*, 5(2).

Hanselka H. and Platz R., 2010. *Ansätze und Maßnahmen zur Beherrschung von Unsicherheit in lasttragenden Systemen des Maschinenbaus. Konstruktion*, 2010.

Heidemann B., 2001. *Trennende Verknuepfung: Ein Prozessmodell als Quelle fuer Produktideen.* VDI-Verlag.

IDEF0, 1993. *Integration definition for function modelling (IDEF0). Federal information processing standards publication*, 183.

Kashyap V.; Bussler C.; and Moran M., 2008. *The Semantic Web - Semantics for Data and Services on the Web*. Springer-Verlag.

Keller G.; Nuettgens M.; and Scheer A., 1992. *Semantische Prozessmodellierung auf der Grundlage" Ereignisgesteuerter Prozessketten (EPK)"*. Inst. füer Wirtschaftsinformatik.

Kim S. and Bracewell R. Wallace K., 2008. *Some Reflections On Ontologies In Engineering Domain*. In I. Horvth and Z. Rusk (Eds.), *Proceedings of TMCE 2008 Symposium*.

Knetsch T., 2003. *Unsicherheiten in Ingenieurberechnungen*. Ph.D. thesis, Fakultät für Verfahrens- und Systemtechnik der Otto-von-Guericke-Universität Magdeburg.

Kyoung Y.; Manley D.; and Yang H., 2006. *Ontology-based assembly design an information sharing for collaborative product development*. In *Computer-aided Design 38*.

Lee J. and Suh H., 2007. *OWL-Based Product Ontology Architecture and Representation for Sharing Product Knowledge on a Web*. In *ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2007)*.

Lindemann U.; Maurer M.; and Braun T., 2009. *Structural Complexity Management*. Springer, Berlin Heidelberg.

Lindley D.V., 2006. *Understanding uncertainty*. Wiley, New York.

Marca D. and McGowan C., 1987. *SADT: structured analysis and design technique*. McGraw-Hill, Inc. New York, NY, USA.

Mascardi V.; Cordi V.; and Rosso P., 2007. *A comparison of upper ontologies*. In *Atti del Workshop Dagli Oggentti agli Agenti, WOA, M. Baldoni and et al.*

Masolo C.; Borgo S.; Gangemi A.; Guarino N.; and Oltramari A., 2003. *WonderWeb Deliverable D18 (Ontology library)*.

Mosch L.; Sprenger A.; and Anderl R., 2010. *Approach for Visualiuation of Uncertainty in CAD-Systems based on Ontologies*. In *Proceedings of the ASME 2010 International Mechanical Engineering Congress & Exposition (IMECE)*.

Nikolaidis E., 2005. *Engineering design reliability handbook*. CRC Press, Boca Raton, Fla.

Niles I. and Pease A., 2001. *Towards a standard upper ontology*. In *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*. ACM.

Pahl B.; Beitz W.; Feldhusen J.; and Grote K., 2007. *Engineering Design*. Springer-Verlag, New York.

Patil L.; Dutta D.; and Sriram R.D.a., 2005. *Ontology Formalization of Product Semantics for Product Lifecycle Management. US Department of Commerce NIST*, NISTIR 7274.

Pavkovic N. and Storga M., 2006. *Product Development Ontology in Tracability Implementation Framework*. In *Proceedings of International Dseign Cconference - DESIGN 2006 Dubrovnik - Croatia*.

Platz R.; Ondoua S.; Habermehl K.; Bedarff T.; Hauer T.; Schmitt S.; and Hanselka H., 2010. *Approach to validate the influences of uncertainties in manufacturing on using load-carrying structures*.

Reh S.; Beley J.D.; Mukherjee S.; and Khor E.H., 2005. *Design for Six Sigma und Robust Design*. Springer, Berlin.

Roy R.K., 1990. *A Primer on the Taguchi Method*. Society of Manufacturing Engineers.

Scheer A., 1996. *ARIS-House of Business Engineering: Von der Geschäftsprozeßmodellierung zur Workflow-gesteuerten Anwendung; vom Business Process Reengineering zum Continuous Process Improvement. Veröffentlichungen des Instituts für Wirtschaftsinformatik*, 133.

Stanford School of Medicine, 2010. *The Protégé Ontology Editor and Knowledge Acquisition System*. http://protege.stanford.edu/.

Yassine A.A.; Sreenivas R.S.; and Zhu J., 2008. *Managing the exchange of information in product development. European Journal of Operational Research*, 184.

AN EVOLUTIONARY ALGORITHM WITH ACTIVE PLANS FOR THE RCPSP

José A. Oliveira, Luís S. Dias, Guilherme Pereira
Dept. Produção e Sistemas, Universidade do Minho
Campus de Gualtar
4710-057 Braga – P, Portugal
E-mail: {zan|lsd|gui@dps.uminho.pt}

KEYWORDS
Resource Constrained Project Scheduling Planning, Scheduling, ERP, Heuristics, Optimization, Evolutionary Algorithm

ABSTRACT

This paper presents a genetic algorithm for the Resource Constrained Project Scheduling Planning (RCPSP). The Evolutionary Algorithm is based on a random keys chromosome that is very easy to implement and allows using conventional transformation operators for combinatorial optimization problems. A project is formed by a set of activities that has to be performed using a set of resources. Each activity uses a specific set of resources, and it is also necessary to guarantee that there is no overlap in the time it takes to process activities in the same resource. The objective of the RCPSP is to minimize the makespan.
The use of exact algorithms for the RCPSP is still limited to instances of small size. The alternative in solving the Real-Large Resource Constrained Project Scheduling Planning is the use of heuristic procedures. This evolutionary algorithm includes specific knowledge of the problem to improve its efficiency. A constructive algorithm based on Giffler-Thompson's algorithm is used to generate active plans. The constructive algorithm reads the chromosome and decides which activity is scheduled next.
The Evolutionary Algorithm will be tested by using some benchmark problems.

INTRODUCTION

Some Combinatorial Optimization Problems are very hard to solve and therefore require heuristic procedures. Two of them are the Resource Constrained Project Scheduling Problem (RSPSP) and its particular variant, the Job Shop Scheduling Problem (JSSP). The use of exact methods to solve both problems is limited to the instances of small size. According to (Zhang et al. 2007), the Branch and Bound methods do not solve instances larger than 250 operations in a reasonable time. As stated in (Liu et al. 2008), in practical manufacturing environments, the scale of job shop scheduling problems could be much larger - in some big textile factories, the number of jobs may be as much as 1,000.
Heuristic methods have become very popular and have garnered much success in solving scheduling problems. In the last twenty years, a huge quantity of papers was published presenting several metaheuristic methods. From Simulated Annealing to Particle Swarm Optimization (Lian et al. 2006), there are several variants of the same method class. A very popular method between the researchers is the Evolutionary Algorithms (EA). This work presents a new idea for improving the effectiveness and efficiency of an Evolutionary Algorithm to solve the RCPSP. Since a difficulty with the RCPSP is the change from instance to instance, a procedure is implemented to take knowledge from the instance and transfer it to the initial population of the EA.
The paper is organized as follows: an introductory section defines the RCPSP problem and its representation with the use of an AoN graph; a central section describes the methodology and a numerical example supports the explanation of constructive algorithm; finally some conclusions and a dis cussion on future work are presented.

RCPSP

A project can be represented as a network of n activities in a graph, where there are links between pairs of activities, representing processing precedence between them. The order to process the set of activities must respect the precedence set. There are two modes to represent a project: the activity-on-arc (AoA) and the activity-on-node (AoN).
In the AoA mode, the set of nodes represents the "events" (start/end processing) and the set of arcs represents the "activities." In the AoN scheme, the set of nodes represents the "activities" and the set of arcs represents the precedence between the "activities." In general, each activity requires the simultaneous use of several resources (Demeulemeester et al. 2003). In this paper, the AoN scheme has been adopted since it allows a direct correspondence with the disjunctive graph that is usually used to represent the JSSP.
For a given r renewable resources $R1,..,Rr$, a constant amount of bk units of resource Rk is available at any time. Activity j must be processed for pj time units, where preemption is not allowed. During this time period, a constant amount of rjk units of resource Rk is occupied. The objective is to determine the starting times Sj for the activities j =1,..,n in such a way that:
- at each time t, the total resource demand is less than or equal to the resource availability for each resource type;
- the given precedence constraints are fulfilled;
- the makespan $C_{\max} = \max_{j=1}^{n} C_j$, where $C_j = S_j + p_j$, is minimized.

The RCPSP is a highly complex optimization problem due to its combinatorial nature. This problem belongs to the NP-hard

class (Brucker et al. 1998), and thus an efficient algorithm for obtaining exact solutions is unknown.
The research on the RCPSP has been extensive along the years and it involves both exact algorithms and heuristic methodology. Generally, the exact methods are based on branch and bound algorithm, where the branch strategies are related to the choice of activities in the disjunctive graph that represents the project. Also, in this field of research, the continuous development of new and better lower bounds for the optimal solution is pointed out, which is fundamental auxiliary information for the heuristic methods.
The state-of-the-art RCPSP performed by (Hartmann and Kolisch 2000) and updated later (Kolisch and Hartmann 2006) shows a greater number of publications on heuristics than exact methods. In this study, the authors explain that in the specialized bibliography, there are several heuristic methodologies to deal with RCPSP. These methods could be classified as "single pass" (over activities set) or "multiple pass," based in priority dispatch rules to construct the solution, local iterative improvement and meta-heuristics.
In the metaheuristics set, several approaches have arisen such as: evolutive strategies, taboo search, ant colony system, variable neighbourhood search and simulated annealing. Some of those metaheuristics include hybridization with local search procedures with single pass or multiple pass.
The process to construct a solution is a fundamental component in the heuristic methodology. With RCPSP in particular, this component is called Schedule Generation Scheme (SGS). For the RCPSP, two SGS models exist: the serial model and the parallel model. It is possible to find in the literature heuristic procedures that use either one scheme or the other, but it is also possible to find some methodologies that use both, or even a combination of the two models.
Analyzing the computational results presented in the literature performed with test instances with 30, 60 and 120 activities, it is possible to point out that the better approximation methodology is the population-based heuristic with representation of list of activities and serial SGS. The parallel SGS seems to be better when dealing with large instances.
It seems that further research for approach methods will concentrate on the development of algorithms which integrate the Forward–Backward Improvement strategy (FBI), since four of the best six heuristics already use this local search strategy. The FBI strategy has two phases: the first phase gives a feasible schedule using a priority dispatch rule to sequence the activities in a bidirectional way; the second phase tries to improve the resources assignment performing successive bidirectional passes without increasing the project duration obtained in the first phase.
The performance of a heuristic can be improved by integrating specific knowledge of the problem. With the RCPSP, the integration of a constructive algorithm in an Evolutionary Algorithm (EA) constructs only active schedules. This strategy is implemented similarly to that performed for the particular case of the Job Shop Scheduling Problem.

EVOLUTIONARY ALGORITHM

This work adopts a method based on Evolutionary Algorithms. EA can be viewed as a metaheuristic search techniques that imitates the principle of evolution in nature i.e., EAs use a set of competing potential solutions of the problem which evolve according to rules of selection and transformation.
It is possible to briefly describe the functioning of an EA in the following way. An EA proceeds in iterations called generations. In each generation, a set of solutions called population is generated. Usually, the first population is generated randomly. A new population is generated from a current population using selection and transformation operators. The members of a population are called individuals. The fitness-function assigns a value called fitness to each individual. Given an individual and its fitness, a selection operator decides if an individual from the current population is used as an input of a transformation operator. A transformation operator creates a new element of a new population from an arbitrary number of elements of the current population. It is expected that during the search process, increasingly better solutions are found. To achieve this goal, it is necessary that the transformation operators focus attention on high fitness. Very common transformation operators are reproduction (copying an individual to the new population unaltered), mutation (an individual is changed in a random fashion) and crossover (where parts of two individuals are combined).
The simplicity of an EA to model more complex problems and its easy integration with other optimization methods were factors that were considered for its choice. The algorithm proposed was conceived to solve the classical RCPSP, but it is possible to use the same method to solve other variants of the RCPSP, like RCMPSP, and Complementary of Resources in a Project Planning (Silva et al. 2010).
One of the features that differentiates conventional Evolutionary Algorithms is the fact that the algorithm does not deal directly with the problem's solutions, but with a solution representation, the chromosome. The algorithm manipulations are done over the representation and not directly over the solution (Goldberg 1989).
Traditionally, genetic algorithms used bit string chromosomes. These chromosomes consisted of only '0s' and '1s.' Modern genetic algorithms more often use problem-specific chromosomes, as the balance between flexibility and raw efficiency tends away from the latter, and with evidence that use of real-valued chromosomes often outperformed bit string chromosomes anyway. Another alternative is the Gray code that is a binary numeral system where two successive values differ in only one digit (Goldberg 1989).
The permutation code was adequate for permutation problems. In this kind of representation, the chromosome is a literal of the operations sequence on the machines. In the RCPSP case, the chromosome is composed by n genes, one for each activity. The i gene of the chromosome corresponds to the activity i in the disjunctive graph. The allele identifies the activity's priority to be considered in the Constructive Algorithm.

Nevertheless, in this work, the random key code presented by (Bean 1994) is used. As (Gonçalves et al. 2005) state, the important feature of random keys is that all offspring formed by crossover are feasible solutions, when it is used as a constructive procedure based on the available operations to schedule and the priority is given by the random key allele. Through the dynamics of the genetic algorithm, the system learns the relationship between random key vectors and solutions with good objective function values. Another advantage of the random key representation is the possibility of using the conventional genetic operators. This characteristic allows the use of the genetic algorithm with other optimization problems, adapting only a few routines related to the problem.

A chromosome represents a solution to the problem and is encoded as a vector of random keys (random numbers). In this work, according to (Cheng et al. 1996), the problem representation is indeed a mix from priority rule-based representation and random keys representation.

The solutions are decoded by an algorithm, which is based on Giffler and Thompson's algorithm (Giffler and Thompson 1960). While the Giffler and Thompson's algorithm can generate all the active plans, the constructor algorithm only generates the plan in agreement with the chromosome.

A plan is called active if the starting times of the activities are such that no activity can be started earlier without delaying some other activity or violating the constraints.

As advantages of this strategy, we have pointed the minor dimension of solution space, which includes the optimum solution and the fact that it does not produce impossible or disinteresting solutions from the optimization point of view. On the other hand, since the dimensions between the representation space and the solution space are very different, this option can represent a problem because two chromosomes can represent the same solution.

Using the Giffler and Thompson's algorithm to construct a solution based on the priority values of each activity, a Serial Schedule Generation Scheme is adopted. It has been shown by (Kolisch 1996) that a serial scheme yields better results when a large number of schedules is computed for one project instance. Attending (Hartmann 1998), the parallel scheme might exclude all optimal solutions from the search space, while the search space of the serial scheme always contains an optimal schedule.

The constructive algorithm has *n* stages and in each stage an activity is scheduled. To assist the algorithm's presentation, consider the following notation existing in stage *t*:

P_t - the partial schedule of the $(t-1)$ scheduled activities;

S_t - the set of activities schedulable at stage t, i.e. all the activities that must precede those in S_t are in P_t;

Q_t - the set of activities queued at stage t, i.e. all the activities that are not in S_t nor in P_t;

σ_k - the earliest time that activity a_k in S_t could be started;

ϕ_k - the earliest time that activity a_k in S_t could be finished, that is $\phi_k = \sigma_k + p_k$;

M^* - the set of resources used by a_k in S_t which has $\phi^* = \min_{a_k \in S_k} \{\phi_k\}$;

S_t^* - the conflict set formed by a_k in S_t which use at least one resource of M^* and $\sigma_j < \phi^*$;

a_j^* - the selected activity to be scheduled at stage t.

Table 1 presents the constructive algorithm that is used to build a solution.

Table 1: A Constructive Algorithm

Step 1	Let $t=1$ with P_1 being null. S_1 will be the set of all activities with no predecessors.
Step 2	Find $\phi^* = \min_{a_k \in S_t} \{\phi_k\}$ and identify M^*. Form S_t^*.
Step 3	Select activity a_j^* in S_t^*, with the greatest allele value.
Step 4	Move to next stage by (1) adding a_j^* to P_t, so creating P_{t+1}; (2) removing a_j^* from the list of predecessors of activities in Q_t deleting a_j^* from S_t and creating S_{t+1} by adding to S_t the activities that a_j^* as the only predecessor; (3) deleting a_j^* from S_t and creating S_{t+1} by adding to S_t the activities with no predecessors; (4) incrementing t by 1.
Step 5	If there are any activities left unscheduled $(t < N)$, go to *Step 2*. Otherwise, stop.

In *Step 3*, instead of using a priority dispatching rule, the information given by the chromosome is used. If the maximum allele value is equal for two or more operations, one is chosen randomly.

NUMERICAL EXAMPLE

Consider the following example, formed with six activities (1,...,6). Consider also two resources *R*1, *R*2, available with four and two units, respectively. Table 2 presents data for this instance. For each activity *j*, it is presented the respective processing time *pj* and the number of each resource needed *rj*1 and *rj*2. For instance, activity 3 takes six units of processing time, it uses two units of the first resource and one unit of the second resource.

Table 2: Example with 6 activities and 2 resources

j	*1*	*2*	*3*	*4*	*5*	*6*
pj	3	4	6	2	1	4
R1	2	2	2	2	3	3
R2	1	1	1	1	0	1

Also, consider two dummy activities: activity 0 and 7 that represent the start and the end of the project. Figure 1 shows this example and also illustrates the precedence constraints that exist between activities.

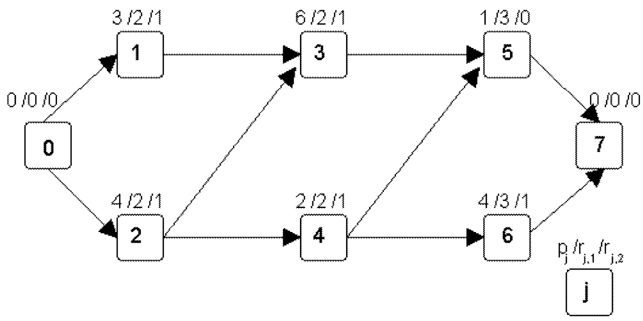


Figure 1: A project network with 6 activities

A chromosome to represent a solution for this instance has six genes, since there exists one gene for each activity. Table 3 presents a chromosome of random keys values for this instance. The values are generated randomly between 0 and 10.

Table 3: A chromosome with 6 genes

j	*1*	*2*	*3*	*4*	*5*	*6*
genes	7	5	8	5	3	2

Table 4 shows the sets P_1, S_1, Q_1, S_1^* at the first iteration. In the conflict set there are activity 1 and activity 2. To select one activity, the information given by the chromosome is taken into account (Table 3) and is selected activity 1, since gene 1 (7) is greater than gene 2 (5).

Table 4: Iteration 1

P	*S*	*Q*	*j*	1	2	3	4	5	6
--	1, 2	3, 4, 5, 6	•	0	0	--	--	--	--
	*S**		•	3	4				
	1, 2		••	3					
			Cj	3					

In the following, Tables are presented with the successive iterations until the solution is formed. Table 5 gives details of the second iteration. Since activity 2 is the only activity that could be scheduled, it is not necessary take in account the gene value.

Table 5: Iteration 2

P	*S*	*Q*	*j*	1	2	3	4	5	6
1	2	3, 4, 5, 6	•	0	0	--	--	--	--
	*S**		•		4				
	2		••		4				
			Cj	3	4				

In iteration 3, activity 3 is chosen instead of activity 4 according to the genes values in the chromosome. Table 6 presents the third iteration.

Table 6: Iteration 3

P	*S*	*Q*	*j*	1	2	3	4	5	6
1, 2	3, 4	5, 6	•	0	0	4	4	--	--
	*S**		•			10	6		
	3, 4		••				6		
			Cj	3	4	10			

Since activity 4 is the only activity that could be scheduled at iteration 4, the information on the chromosome is not relevant. Table 7 shows this situation.

Table 7: Iteration 4

P	*S*	*Q*	*j*	1	2	3	4	5	6
1, 2, 3	4	5, 6	•	0	0	4	4	--	--
	*S**		•				6		
	4		••				6		
			Cj	3	4	10	6		

In iteration 5, activity 5 is chosen instead of activity 6 according to the genes values in the chromosome. Table 8 resumes the fifth iteration.

Table 8: Iteration 5

P	*S*	*Q*	*j*	1	2	3	4	5	6
1, 2, 3, 4	5, 6	--	•	0	0	4	4	10	10
	*S**		•					11	14
	5, 6		••					11	
			Cj	3	4	10	6	11	

Finally, Table 9 presents the last iteration, where activity 6 is scheduled. Again, there is only one activity in the conflict set, the choice is mandatory and the chromosome does not decide.

Table 9: Iteration 6

P	*S*	*Q*	*j*	1	2	3	4	5	6
1, 2, 3, 4, 5	6	--	•	0	0	4	4	10	11
	*S**		•						15
	6		••						15
			Cj	3	4	10	6	11	15

The $\sigma_6 = 10$ defined at iteration 5 is defined not by the predecessor activity (activity 4) but by the set of resources that will be used. The constructive algorithm must take in account the values of availability of all resources and the conflict set S_i^* depends on these values. We believe that the conflict set information could be very important to implement local search procedures, since it is possible to change from one active plan to another active plan, swapping two activities that appear in the same conflict set.
The resources occupation is illustrated in Figure 2.

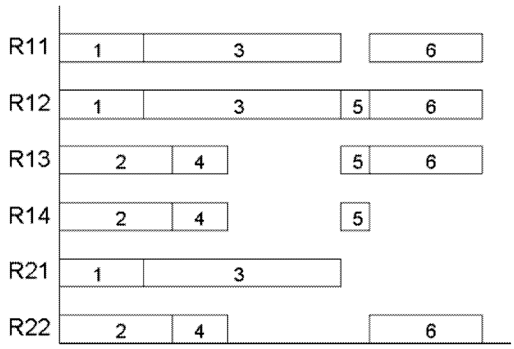


Figure 2: Resources occupation

COMPUTATIONAL EXPERIMENTS

We intend to implement this Evolutionary Algorithm for the Resource Constrained Project Scheduling Problem. In the present, the results are not yet available. However, a similar idea of Evolutionary Algorithm was implemented to the particular situation that is the job shop scheduling problem, and the computational results are promissory.

CONCLUSIONS AND FURTHER WORK

The goal of this paper was to provide a formal model to one important problem in project management, specifically the one related to constrained resources utilization and its implementation. Since the strategy to implement a constructive algorithm inside an Evolutionary Algorithm to generate active schedules was successfully tested in the Job Shop Scheduling Problem (which can be seen as a particular case of RCPSP), it is our intention to implement it on a more general situation.

Considering the feasibility of the model proposed, we believe it can provide to the user a new option to plan and to determine the best assignment of resources and the lowest project cost, pushing the planning phase and increasing the estimation ability of the companies.

This paper presents a structure to implement an Evolutionary Algorithm to solve the RCPSP. The schedules are built using information given by the Evolutionary Algorithm to sequence the activities. The approach will be tested on a set of instances taken from the literature on RCPSP.

To improve the efficiency and effectiveness of this Evolutionary Algorithm new transformation operators will be developed namely for crossover and mutation operations. This development aims to get better neighbourhoods to perform the search. In real-world situations the organizations frequently have several projects running simultaneously, sharing several resources. This problem is a generalization of RCPSP and is called Resource Constrained Multi-Project Scheduling Problem (RCMPSP). We intend to extend the proposed EA to the RCMPSP.

REFERENCES

Bean, J.C. 1994. "Genetics and random keys for sequencing and optimization." *ORSA Journal on Computing*, 6, 154–160.

Brucker, P.; S. Knust; A. Schoo; and O. Thiele. 1998. "A branch and bound algorithm for the resource-constrained project scheduling problem." *European Journal of Operational Research*, 107, 272-288.

Cheng, R.; M. Gen; and Y. Tsujimura. 1996. "A tutorial survey of job-shop scheduling problems using genetic algorithms part I: Representation." *Computers & Industrial Engineering*, 34 (4), 983–997.

Demeulemeester, E.; M. Vanhocke; and W. Herroelen. 2003. "RanGen: A Random Network Generator for Activity-on-the-node Networks." *Journal of Scheduling*, 6, 17-38.

Gao, J.; L. Sun; and M. Gen. 2008. "A hybrid genetic and variable neighborhood descent algorithm for flexible job shop scheduling problems." *Computers & Operations Research*, 35, 2892– 2907.

Giffler B. and G.L. Thompson. 1960. "Algorithms for solving production scheduling problems." *Operations Research*, 8, 487-503.

Goldberg, D.E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA.

Gonçalves, J.F.; J.J Mendes and M.G.C. Resende. 2005. "A hybrid genetic algorithm for the job shop scheduling problem." *European Journal of Operational Research*, 167, 77–95.

Hartmann, S. and R. Kolisch. 2000. "Experimental evaluation of state-of-the-art heuristics for the resource-constrained project scheduling problem." *European Journal of Operational Research*, 127, 2, 394-407.

Kolisch, R. 1996. "Serial and parallel resource-constrained project scheduling methods revisited: Theory and computation." *European Journal of Operational Research*, 90, 320–333.

Kolisch, R. and S. Hartmann. 2006. "Experimental investigation of heuristics for resource-constrained project scheduling: An update." *European Journal of Operational Research*, 174, 1, 23-37.

Laarhoven, P.J.; E. Aarts; and J. Lenstra. 1992. "Job shop scheduling by simulated annealing." *Operations Research*, 40, 1, 113-125.

Lian, Z.; X. Gu; and B. Jiao. 2006. "A similar particle swarm optimization algorithm for job-shop scheduling to minimize makespan." *Applied Mathematics and Computation*, 183, 1008–1017.

Liu, M.; J. Hao; and C. Wu. 2008. "A prediction based iterative decomposition algorithm for scheduling large-scale job shops," *Mathematical and Computer Modelling*, 47, 411–421.

Oliveira, J.A. "A genetic algorithm with a quasi-local search for the job shop problem with recirculation." *Applied Soft Computing Technologies: The Challenge of Complexity*. Springer, Berlin / Heidelberg, 221-234.

Park, B.J.; H.R. Choi; and H.S. Kim. 2003. "A hybrid genetic algorithm for the job shop scheduling problems." *Computers and Industrial Engineering*, 45, 597–613.

Silva, H.; J.A. Oliveira; and A. Tereso. 2010. "Um Algoritmo Genético para Programação de Projectos em Redes de Actividades com Complementaridade de Recursos." *Revista Ibérica de Sistemas y Tecnologías de la Información*,. 6, 59-72.

Zhang, C.; P. Li; Z. Guan and Y. Rao. 2007. "A tabu search algorithm with a new neighborhood structure for the job shop scheduling problem." *Computers & Operations Research* 53, 313-320.

BIOGRAPHY

JOSÉ A. OLIVEIRA was born 1966 in Matosinhos, Portugal. He studied Mechanical Engineering at the University of Porto, Portugal. He graduated with a Ph.D. in Production and Systems Engineering at University of Minho, Portugal. His main research interests are Optimization with Heuristic Methods in Systems Engineering.

Luís M S Dias was born in 1970 in Vila Nova de Foz Côa, Portugal. He graduated in Computer Science and Systems Engineering at the University of Minho, Portugal. He holds an MSc degree in Informatics Engineering and a PhD degree in Production and Systems Engineering from the University of Minho, Portugal. His main research interests are Simulation, Operational Research and Systems Visual Modeling.

Guilherme A B Pereira was born in 1961 in Porto, Portugal. He graduated in Industrial Engineering and Management at the University of Minho, Portugal. He holds an MSc degree in Operational Research and a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. His main research interests are Operational Research and Simulation.

A GENETIC ALGORITHM FOR PROJECT SCHEDULING IN ACTIVITY NETWORKS UNDER RESOURCE COMPLEMENTARITY

Helder Silva

IFAM - Instituto Federal de Educação
Tecnológica do Amazonas
Manaus, Brazil
E-mail: helder@ifam.edu.br

José Oliveira
Anabela Tereso
Dept. Produção e Sistemas,
Universidade do Minho
Campus de Gualtar
4710-057 Braga – P, Portugal
E-mail: {zan|anabelat@dps.uminho.pt}

KEYWORDS
Project Management, Scheduling, Complementarity of resources.

ABSTRACT

We address the issue of optimal resource allocation, and more specifically, the analysis of complementarity of resources (primary resource or *P*-resource and supportive resource or *S*-resource) to activities in a project. The concept of complementarity can be incorporated into the engineering domain as an enhancement of the efficacy of a "primary" resource (*P*-resource) by adding to it other "supportive" resources (*S*-resources). We developed a Genetic Algorithm capable of determining the ideal mixture of resources allocated to the activities of a project, such that the project is completed with minimal cost. This problem has a circularity issue that greatly increases its complexity.
In this paper we present a constructive algorithm to build solutions from a chromosome that will be integrated in a Genetic Algorithm, which we illustrate by application to a small instance of the problem. The Genetic Algorithm is based on a random keys chromosome that is very easy to implement and allows using conventional genetic operators for combinatorial optimization problems. A project is formed by a set of activities. Each activity uses a specific set of resources, and it is also necessary to guarantee that there is no overlap in the time it takes to process activities in the same resource.

INTRODUCTION

This paper is concerned with the optimal resource allocation in activity networks under conditions of resource complementarity. The concept of complementarity which has been discussed from an economic point of view (Kremer, 1993) can be incorporated into the engineering domain as an enhancement of the efficacy of a “primary” resource (*P*-resource) by adding to it other “supportive” resources (*S*-resources). Aspects related to performance improvement, short duration, quality improvement have been presented by Silva et al. (2010) as well as the effect of the “supportive” resource for project cost.
We developed a mathematical model capable of determining the ideal mixture of resources allocated to the activities of a project, such that the project is completed with minimal cost (Silva et al. 2010; Silva et al. 2010b). This problem has a circularity issue that greatly increases its complexity. We have developed a procedure which we illustrate by application to small instances of the problem, using complete enumeration over the decision space (Silva et al. 2010b).
The optimal resource allocation in activity networks under conditions of resource complementarity is a generalization of the well known RCPSP which belongs to the NP-hard class (Brucker et al. 1998). This problem is a highly complex optimization problem due to its combinatorial nature, and thus an efficient algorithm for obtaining exact solutions is unknown. The development of a more computationally efficient procedure is now presented and is implemented in a Genetic Algorithm.

PROBLEM DESCRIPTION

Consider a project network in the activity-on-node (AoN) representation: $G=(N,A)$ with the set of nodes $|N|=n$ (representing the “activities”) and the set of arcs $|A|=m$ (representing the precedence relations between the activities). In general each activity requires the simultaneous use of several resources (Tereso et al. 2008; Tereso et al. 2009; Tereso et al. 2009b).
There is a set of “primary” resources, denoted by *P*, with $|P|=\rho$. Typically, a primary resource has a capacity of several units (say workers, m/c’s, processors; etc.) (Mulcahy 2005). Additionally, there is a pool of “supportive” resources, denoted by *S*, with $|S|=\sigma$ (such as less-skilled labor, or computers and electronic devices; etc.) that may be utilized in conjunction with the primary resources to enhance their performance. The number of supportive resources varies with the activity and the primary resources required for its execution. The impact on the *P*-resource is evaluated using a variable $0<v(r_p,s_q)\leq 1$ that indicates the fraction by which the *S*-resource s_q improves the performance of *P*-resource r_p. Typically, $v(r_p,s_q)\in[0.1,0.5]$. Consider $x_a(r_p)$ as the level of allocation of (primary) resource r_p to activity a, and $x_a\left(r_p,\{s_q\}_{q=1}^{\sigma}\right)$ as the total allocation of resource r_p (including complementary resources) to activity a.

We assume that the impact of the *S*-resources is additive: if a subset $\left\{s_q\right\}_{q=1}^{v}$ of the *S*-resources is used in support of *P*-resource r_p in activity *a*, and only one unit of each *S*-resource is used, then the performance of the former is enhanced to,

$$x_a\left(r_p,\left\{s_q\right\}_{q=1}^{\sigma}\right)=x_a\left(r_p\right)+\sum_{q=1}^{\sigma}v\left(r_p,s_q\right) \qquad (1)$$

With $w_a\left(r_p\right)$ representing the work content of activity *a* for *P*-resource *r*, the primary resource $r_p \in P$ would accomplish activity *a* in time $y_a\left(r_p\right)$ (see (2)). If it is enhanced by the addition of support resources, then its processing time decreases to $y_a\left(r_p,s_q\right)$ (see (3)).

$$y_a\left(r_p\right)=\frac{w_a\left(r_p\right)}{x_a\left(r_p\right)} \qquad (2)$$

$$y_a\left(r_p,s_q\right)=\frac{w_a\left(r_p\right)}{x_a\left(r_p,s_q\right)} \qquad (3)$$

An activity normally requires the simultaneous utilization of more than one *P*-resource for its execution. The problem then becomes: "At what level should each resource be utilized and which supportive resource(s) should be added to it (if any) in order to optimize a given objective?"

Recall that the processing time of an activity is given by the maximum of the durations that would result from a specific allocation to each resource (see a previous discussion on the evaluation of the duration considering multiple resources in (Tereso et al. 2008; Tereso et al. 2009; Tereso et al. 2009b)).

$$y\left(a\right)=\max_{all\ r_p}\left\{y_a\left(r_p\right)\right\} \qquad (4)$$

To better understand this representation, consider a simple project in AoN mode of representation (see Figure 1).

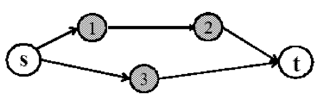


Figure 1: Project with 3 activities AoN

This project is formed by three activities, 1, 2 and 3, for which we will assume that it is required the utilization of four *P*-resources (in man-days); not all resources are required by all the activities (see Table 1).

Table 1: Work content of the activities of project

P-resource →	1	2	3	4
↓ Activity/Availability→	2	1	3	2
A1	16	0	12	12
A2	0	7	10	8
A3	20	0	22	0

The relevance and impact of the support resources on *P*-resources are represented in Table 2.

Table 2: The P-S matrix

	P-Resource →	1	2	3	4
↓*S*-Resource	↓ Availability				
1	1	0.25	ϕ	0.25	ϕ
2	2	0.15	0.35	ϕ	ϕ

We consider the cost of the resource utilization, a bonus for early completion, and a penalty for late completion of the project, after specifying a due date. A model was developed to minimize the total cost, considering that the activities should start as soon as they are sequence feasible (if there are enough primary resources to start them) (Silva et al. 2010). Some results were also reported using a procedure based on the analysis of the network and the concept of state space (Silva et al. 2010b).

GENETIC ALGORITHM

Since the Job Shop Scheduling Problem (JSSP) can be seen as a particular case of the Resource Constrained Project Scheduling Problem (RCPSP), we will extend a Genetic Algorithm (GA) developed for the JSSP (Oliveira et al. 2010) to the RCPSP, and particularly to the Project Scheduling in Activity Networks under Resource Complementarity. The GA is based on a random keys chromosome representation, that allows an easy reconfiguration to be applied in other problems.

The GA's simplicity to model more complex problems and its easy integration with other optimization methods were factors that were considered before it was chosen. Initially, the algorithm proposed was conceived to solve the classical JSSP (Oliveira 2007), but it is possible to use the same method to solve other variants of the JSSP (Oliveira 2006), or in this case to solve a generalization of JSSP, that is the RCPSP with complementary resources.

One of the features that differentiates conventional genetic algorithms is the fact that the algorithm does not deal directly with the problem's solutions, but with a solution representation - the chromosome. The algorithm manipulations are performed over the representation and not directly over the solution (Goldberg 1989).

We represent the project scheduling problem in a graph in AoN (Activity-on-Nodes) because it is similar to the disjunctive graph that is used to represent the JSSP (Roy and Sussmann, 1964). An activity can only be started if its predecessors are completed and if all the primary resources required are available. A project has a technological definition that determines a specific order to process some activities, and it is necessary to guarantee that there is no overlap in the time it takes to process such activities on the common resources. Considering this characteristic, we use the concept of the schedulable activity (an activity that could be started), and at each decision moment, it is only necessary to choose an activity from the set of schedulable activities. The choice of the activity is driven by the genetic algorithm attending to the alleles existent in the chromosome that give the priority of each activity.

Traditionally, genetic algorithms used bit string chromosomes. These chromosomes consisted of only '0s' and '1s.' Modern genetic algorithms more often use problem-specific chromosomes, as the balance between flexibility and raw efficiency tends away from the latter, and with evidence that use of real-valued chromosomes often outperformed bit string chromosomes anyway. Another alternative is the Gray code that is a binary numeral system where two successive values differ in only one digit (Goldberg 1989).

The permutation code was adequate for permutation problems. In this kind of representation, the chromosome is a literal translation of the operations sequence on the machines. In the classical JSSP case, the chromosome is composed by m sub-chromosomes, one for each machine, each one composed by n genes, one for each operation (Oliveira 2007). The i gene of the sub-chromosome corresponds to the operation processed in i place in the corresponding machine. The allele identifies the operation's index in the disjunctive graph (Roy and Sussmann, 1964).

For the Activity Networks under Resource Complementarity, we define a chromosome with $n(\rho+\sigma+1)$ genes. For each activity, the chromosome gives the quantity of each *P*-resource and the quantity of the complementary *S*-recourse. The chromosome also indicates the priority of each activity.

Nevertheless, in this work, the random key code presented by Bean (1994) is used. The important feature of random keys is that all offspring formed by crossover are feasible solutions, when it is used jointly with a constructive procedure based on the available operations to schedule and the priority is given by the random key allele. Through the dynamics of the genetic algorithm, the system learns the relationship between random key vectors and solutions with good objective function values. Another advantage of the random key representation is the possibility of using the conventional genetic operators. This characteristic allows the use of the genetic algorithm with other optimization problems, adapting only a few routines related with the problem.

A chromosome represents a solution to the problem and is encoded as a vector of random keys (random numbers). In this work, according to Cheng et al. (1996), the problem representation is indeed a mix from priority rule-based representation and random keys representation.

The genetic algorithm has a very simple structure and can be represented in the Algorithm 1. It begins with population generation and her evaluation. Attending to the fitness of the chromosomes the individuals are selected to be parents.

Algorithm 1: Genetic algorithm

begin
$P \leftarrow$ GenerateInitialPopulation()
Evaluate(P)
while termination conditions not meet **do**
    $P' \leftarrow$ Recombine(P)
    Evaluate(P')
    $P \leftarrow$ Select($P \cup P'$)
end while

The crossover is applied and it generates a new temporary population that also is evaluated. Comparing the fitness of the new elements and of their progenitors the former population is updated.

The Uniform Crossover (UX) is used this work. This genetic operator uses a new sequence of random numbers and swaps both progenitors' alleles if the random key is greater than a prefixed value. Figure 2 illustrates the UX's application on two parents (prnt1, prnt2), and swaps alleles if the random key is greater or equal than 0.75. The genes 3, 4 and 12 are changed and it originates two descendants (dscndt1, dscndt2). Descendant 1 is similar to parent 1, because it has about 75% of genes of this parent.

gene	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***
prnt1	0,89	0,49	0,24	0,03	0,41	0,11	0,24	0,12	0,33	0,30	0,27	0,18
prnt2	0,83	0,41	0,40	0,04	0,29	0,35	0,38	0,01	0,42	0,32	0,28	0,13
randkey	0,64	0,72	0,75	0,83	0,26	0,56	0,28	0,31	0,09	0,11	0,37	0,76
dscndt1	0,89	0,49	0,40	0,04	0,41	0,11	0,24	0,12	0,33	0,30	0,27	0,13
dscndt2	0,83	0,41	0,24	0,03	0,29	0,35	0,38	0,01	0,42	0,32	0,28	0,18

Figure 2: The UX crossover

CONSTRUCTIVE ALGORITHM

The solutions are decoded by an algorithm, which is based on Giffler and Thompson's algorithm (GT) (Giffler and Thompson 1960). While the GT algorithm can generate all the active plans for the JSSP, the constructive algorithm only generates the plan in agreement with the chromosome. As advantages of this strategy, we have pointed out the minor dimension of solution space, which includes the optimum solution and the fact that it does not produce impossible or disinteresting solutions from the optimization point of view. On the other hand, since the dimensions between the representation space and the solution space are very different, this option can represent a problem because many chromosomes can represent the same solution.

The constructive algorithm has n stages and in each stage an activity is scheduled. To assist the algorithm's presentation, consider the following notation existing in stage t:

P_t - the partial schedule of the $(t-1)$ scheduled activities;

S_t - the set of activities schedulable at stage t, i.e. all the activities that must precede those in S_t are in P_t;

σ_k - the earliest time that activity a_k in S_t could be started. This time respects the conclusion of all predecessors of a_k and the availability of all resources that a_k will use (primary and supportive resources);

ϕ_k - the earliest time that activity a_k in S_t could be finished;

M^* - the set of resources used by a_k in S_t which has

$$\phi^* = \min_{a_k \in S_k} \{\phi_k\};$$

S_t^* - the conflict set formed by a_k in S_t which use at least one resource of M^* and $\sigma_j < \phi^*$;

A^* - the activity where $\varphi^* = \min_{a_k \in S_k} \{\varphi_k\}$;

S_t^* - the conflict set formed by $a_k \in S_t$ that have $\delta_k < \varphi^*$;

a_k^* - the selected activity to be scheduled at stage t.

The constructive algorithm of solutions is presented in a format, similar to the one used by Cheng et al. (1996) to present the GT algorithm. In Step 3, instead of using a priority dispatching rule, the information given by the chromosome is used. If the maximum allele value is equal for two or more activities, one is chosen randomly.

Algorithm 2: Constructive algorithm

Step 1 Let $t=1$ with P_1 being null. S_1 will be the set of all activities with no predecessors; in other words those that are connected to start vetex.

Step 2 Find $\phi^* = \min_{a_k \in S_t} \{\phi_k\}$ and identify A^*. Form S_t^*.

Step 3 Select activity a_k^* in S_t^*, with the greatest priority.

Step 4 Move to next stage by

(1) adding a_k^* to P_t, so creating P_{t+1};

(2) deleting a_k^* from S_t and creating S_{t+1} by adding to S_t the activities that directly follows a_k^* and have all predecessors in P_{t+1};

(3) incrementing t by 1.

Step 5 If there are any activities left unscheduled $(t < N)$, go to *Step 2*. Otherwise, stop.

Consider the example presented in Figure 1, Table 1, and Table 2 with three activities (A_1, A_2, A_3). In this instance, there are $\rho = 4$ primary resources and $\sigma = 2$ supportive resources. A chromosome to represent a solution for this instance has 21 genes, since there are six genes for each activity to establish the number of elements of each resource that will be used, plus a gene A_k that defines the activity's priority. Table 3 presents a chromosome of random keys values for this instance. The values are generated randomly between 0 and 99.

Table 3: A chromosome

A1	P1	P2	P3	P4	S1	S2	A2	P1	P2	P3	P4	S1	S2	A3	P1	P2	P3	P4	S1	S2
94	51	88	76	52	23	68	36	73	60	61	53	75	35	47	7	15	42	86	16	16

Activity A_1 has a priority 94, which is the greatest, while A_2 has the lowest priority (36). To define the number of elements of each resource, we use Table 1. The availability of *P*-resource 3 is 3 units. We define three equal intervals between 0 and 99. For this resource, if the allele is a value between 0 and 32 one unit is assigned. For values between 33 and 66, we assign two units, and for values between 67 and 99, three units will be assigned. To establish the number of units for the supportive resources, the procedure is similar, but it also includes the possibility to assign 0 units, because it is not required to use at least one unit. Considering these rules, the chromosome presented in Table 3 defines the following units of resources to be used, which are presented in Table 4.

Table 4: Amount of units of resources to be used

A1	P1	P2	P3	P4	S1	S2	A2	P1	P2	P3	P4	S1	S2	A3	P1	P2	P3	P4	S1	S2
Units	2	0	3	2	0	2	Units	0	1	2	2	1	1	Units	1	0	2	0	0	0

The procedure assigns zero units of a *P*-resource to an activity, if the activity does not use that primary resource, which is the case of P_2 in the activity A_1, according to Table 1. The assignment of supportive resources to the primary resources is performed considering the "amount" of Work content existent after the assignment of primary resources. The first unit is assigned to the *P*-resource with the largest Work content. After the assignment, the amount of work is recalculated, and then the next assignment is made, and so on. Activity A1 has the following Work content (see Table 1):

P-resource →	1	2	3	4
A1	16	0	12	12

Assigning the units of primary resources defined by the chromosome, the duration is then:

P-resource →	1	2	3	4
A1	8	0	4	6

The first unit of supportive resource S_2 is assigned to P_1. Recalculating the durations by (3), we have:

P-resource →	1	2	3	4
A1	7.442	0	4	6

The second unit of supportive resource S_2 is also assigned to P_1. Recalculating the durations by (3) we have:

P-resource →	1	2	3	4
A1	6.957	0	4	6

Applying the same procedure to the remaining activities, we will have the durations presented in Table 5.

Table 5: Activities duration

P-resource →	1	2	3	4
A1	6.957	0	4	6
A2	0	5.185	4.444	4
A3	20	0	11	0

NUMERICAL EXAMPLE

Figure 3 illustrates the results of applying the constructive algorithm. It presents the evolution of the schedulable set $\overline{S}$ and the corresponding staring and conclusion times for each activity during the execution of the constructive algorithm. The final Gant Chart of the project is also presented, and it shows the occupation of all the resource units.

S_1	1	3
δ_k	0	0
φ_k	6.957	20
a_k^*	1	

S_2	2	3
δ_k	6.957	6.957
φ_k	12,14	26,96
a_k^*		3

S_3	2
δ_k	17,97
φ_k	23,14
a_k^*	2

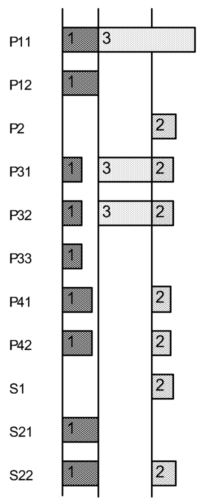


Figure 2: Constructive Algorithm execution

Consider the due date for the project equal to 24 units of time and the following parameters:

• each unit of *P*-resource costs 4 per unit of Work content;
• each unit of *S*-resource costs 1 per unit of Work content;
• the delay costs 60 per unit of time;
• the earliness costs 40 per unit of time.

The project is complete at time 26.96 with 2.96 units of lateness. The resources cost is equal to 845, and the delay cost is 177.39. The total cost of the solution is 1022.39.

CONCLUSIONS

The goal of this paper was to provide a formal model to some unresolved issues in the management of projects, especially as related to the utilization of supportive resources, and to its implementation. The relevance of the problem is the opportunity to shape a system that allows not only for improvement in the allocation of often scarce resource(s), but also results in reduced uncertainties within the projects, combined with increased performance and lower project costs. The model was first presented in (Silva et al. 2010), but it still needed to be implemented and applied to some project networks to demonstrate its validity. We presented the procedure developed to solve the mathematical model, and we applied it to two simple networks, obtaining the desired results through an initial implementation in C (Silva et al. 2010b).

Considering the feasibility of the model proposed, we believe it can provide the user a new option to plan and determine the best combination of resources and the lowest project cost, pushing the planning phase and increasing the estimation ability of the companies.

This paper presents a structure to implement a genetic algorithm to solve the Project Scheduling in Activity Networks under Resource Complementarity. The schedules are built using information given by the genetic algorithm to order the activities. We presented an example of the application of the model, and we achieved preliminary results for a small example. In future work, we intend to test this approach on a set of instances taken from the literature of RCPSP, which will be modified to accommodate the supportive resources.

REFERENCES

Bean, J.C. 1994. "Genetics and random keys for sequencing and optimization." *ORSA Journal on Computing*, 6, 154–160.

Cheng, R.; M. Gen; and Y. Tsujimura. 1996. "A tutorial survey of job-shop scheduling problems using genetic algorithms - I. Representation." *Computers & Industrial Engineering*, 30, 983-997.

Giffler, B.; and G.L. Thompson. 1960. "Algorithms for solving production scheduling problems." *Operations Research*, 8, 487-503.

Goldberg, D.E. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA.

Kremer, M. 1993. "The O-Ring Theory of Economic Development." *The Quartely Journal of Economics*, 108, 551-575.

Mulcahy, R. 2005. *PMP Exam Prep, Fifth Edition: Rita's Course in a Book for Passing the PMP Exam*. RMC Publications, Inc.

Oliveira, J.A.; L. Dias; and G. Pereira. 2010. "Solving the Job Shop Problem with a random keys genetic algorithm with instance parameters, In *Proceedings of the 2nd International Conference on Engineering Optimization*. Lisbon, Portugal.

Oliveira, J.A. 2007. "Scheduling the truckload operations in automatic warehouses." *European Journal of Operational Research*, 179, 3, 723-735.

Oliveira, J.A. 2006. "A genetic algorithm with a quasi-local search for the job shop problem with recirculation". *Applied Soft Computing Technologies: The Challenge of Complexity*. Springer, Berlin / Heidelberg, 221-234.

Roy, B.; and B. Sussmann. 1964. "Les problemes d'ordonnancement avec constraintes disjonctives." Note DS, No. 9 Bis, SEMA, Paris.

Silva, H.C.; A.P. Tereso; and J.A. Oliveira. 2010. "On Resource Complementarity." In *Proceedings of the 3rd International Conference on Information Systems*, Logistics and Supply Chain. Casablanca.

Silva, H.C.; A.P. Tereso; and J.A. Oliveira. 2010b. "On Resource Complementarity in Activity Networks – Further Results". In *Proceedings of the 2nd International Conference on Engineering Optimization*, Lisbon, Portugal.

Tereso, A.; M. Araújo; R. Moutinho; and S. Elmaghraby. 2008. "Project management: multiple resources allocation." In *Proceedings of the International Conference on Engineering Optimization*. Rio de Janeiro, Brazil.

Tereso, A.; M. Araújo; R. Moutinho; and S. Elmaghraby. 2009. "Duration Oriented Resource Allocation Strategy on Multiple Resources Projects under Stochastic Conditions," In *Proceedings of the International Conference on Industrial Engineering and Systems Management*. Montreal, Canada.

Tereso, A.; M. Araújo; R. Moutinho; and S. Elmaghraby. 2009b. "Quantity Oriented Resource Allocation Strategy on Multiple Resources Projects under Stochastic Conditions." In *Proceedings of the International Conference on Industrial Engineering and Systems Management*. Montreal, Canada.

BIOGRAPHY

HELDER SILVA was born in Minas Gerais, Brazil and went to the University of Uberlândia, where he studied Electrical Engineering and obtained his degrees in 1998. He holds an MSc in the same area (2001) and nowadays he is doing a PhD at University of Minho in Portugal. Professor of some universities in Brazil (UNIP, IDAAM and IFAM), he has several international publications related to Project Management, Resource Allocation in Projects, Complementarity and Project Cost Optimization. He is currently the Coordinator of Project Management Office in a Chinese Worldwide Company in Brazil.

JOSÉ A. OLIVEIRA was born 1966 in Matosinhos, Portugal. He studied Mechanical Engineering at the University of Porto, Portugal. He graduated with a Ph.D. in Production and Systems Engineering at University of Minho, Portugal. His main research interests are Optimization with Heuristic Methods in Systems Engineering.

ANABELA TERESO was born in Cantanhede, Portugal. She has a degree in Systems and Informatics Engineering (1990), an MSc in Informatics (1997) and a PhD in Production and Systems Engineering (2002) all from University of Minho – Portugal. She is Professor at University of Minho since 1995, and does research in the area of Project Management since her PhD.

FACTORY PLANNING AND CONTROL FOR SMALL AND MEDIUM ENTERPRISES

QUALITY, INNOVATION AND PERFORMANCE: AN EXPLORATORY STUDY

António Fernandes
Politecnic Institute of Castelo Branco
School of Tecnology

Avenida do Empresário
6000-767 Castelo Branco, Portugal
amfernandes@ipcb.pt

Luís Lourenço
University of Beira Interior
Department of Management and Economics
NECE – Núcleo de Estudos em Ciências Empresariais
Estrada do Sineiro
6200–209 Covilhã, Portugal
lourenco@ubi.pt

KEYWORDS
Quality; Innovation; Performance; ISO 9000.

ABSTRACT

The relationship between total quality management and innovation appears contradictory and complex. Literature suggests that conflicting arguments exist and while for some researchers, total quality management practices do not directly improve innovation, for others, there is a positive relationship between management practices of quality and innovation. Similarly, it remains to explain the influence of total quality management and innovation in the organizational performance.

In this research, our goal is to understand the synergistic relationships between the three variables: total quality management, organizational innovation and organizational performance. This study is based on two exploratory interviews in two ISO 9000 certified SMEs (Small and Medium sized Enterprises): a Logistics Company and a Fuel Company (household fuel and equipment installation and maintenance).

Based on the analysis of these interviews we present a set of propositions about the relationship between those variables. These propositions will support the formulation of hypotheses for future scientific studies in this area. Results of this research can be of high utility in understanding how the variables interact and their impact in different types of organizations.

INTRODUCTION

During the last two decades, the deregulation and globalization of markets, as well as internationalization of service firms, originated severe competition between firms. The concepts of quality and innovation have become guiding elements for what, in the business world, is known as management excellence. They constitute the centre of strategic management for formulating and implementing objectives, policies and performance. TQM and innovation management are elements that can increase the competitive advantage of firms (Garrido et al. 2007). The organizations have adopted TQM (Total Quality Management) as a strategy to respond to customer requirements (Han et al. 2007). TQM can be understood as a management philosophy, whose main concern is to meet the needs and expectations of customers through the integration of all functions and processes of the organization to achieve continuous improvement of quality of goods and services (Fuentes et al. 2006; Han et al. 2007; Lenka and Sweat 2008).

Despite the incorporation of innovation into management excellence models based on TQM and, although, consensus states that innovation offers a major source of sustained competitive advantage, research into the relationship between TQM and innovation performance remains scarce (Prajogo and Sohal 2003, 2004; Singh and Smith 2003). In this sense, it is of major relevance to analyse the relationship between TQM and innovation of firms in general and of small firms in particular.

Small and medium enterprises are very important for economic growth in all countries. They contribute with job opportunities and act as supplier of goods and services to large organizations. SME`s are defined by a number of factors and criteria, such as location, size, age, structure, organization, number of employees, sales volume and worth of assets (Rahman 2001).

The purpose of the current research, an exploratory study, is to understand the relationship between total quality management, innovation and organizational performance in small business. Specifically, the paper presents a set of propositions about the relations between them. It is expected that the results of this research may be helpful in understanding how the three variables interact and their impact on different types of service organizations.

After introducing the subject topic and having highlighted the main objective of this article, here we present a revision of the literature, the methodology used throughout the study, the achieved results, ending with the findings and limitations of the study.

LITERATURE REVIEW

Total Quality Management

The Total Quality Management (TQM) is a systematic approach that considers all the interactions between the various elements of the organization. TQM comprises a group of ideas and techniques for enhancing competitive performance by improving the quality of products and processes (Grant et al. 1994). TQM focuses on improving organizational effectiveness and responsiveness to customer

needs. The goals of TQM are the organizational distinction and the customer satisfaction (Han et al. 2007). From what was exposed it can be said that the strategy of total quality management is based on the global involvement of all collaborators who are encouraged to be more flexible, interactive and participatory in organizational activities.
Proponents of ISO 9000 argue that certification is the first step towards total quality, being an important contribution towards TQM (Anderson et al. 1999; Gotzamani at al. 2006). The adoption of ISO 9000 is a way of achieving competitive advantage through quality management (Anderson et al. 1999).

Innovation

Innovation can not be seen as something periodical that happens by accident or something that results from the action of an individual agent. Innovation is related with the adoption and application of new knowledge and practices, including the ability of an organization to adopt or create new ideas and implement these ideas in developing new products, services, and working processes and procedures and improving those already established (Bates and Khasawneh 2005). Innovation is considered an intangible resource. Silva et al. (2007) suggested that innovation is seen as the result of an interactive and nonlinear process between the firm and the environment. According to the Oslo Manual (OCDE, 2005), innovation in services is organized in a less formal way, it is less technological and is more incremental in nature.

Performance

In the context of global and highly competitive market, the evaluation of the performance has become an important component of the development strategies of organizations. Performance measurement can be defined as the process of quantifying the efficiency and effectiveness of production systems (Singh and Garg 2008). Thus, performance may relate to the individual, group, organization, organizational segment, function, activity, market segment, etc. In literature one can find various methodologies to measure performance, including qualitative and quantitative measures. These latter are of financial or non-financial nature.
According to Franco et al. (2006) strategic performance is generic, synthetic, and generally covers long periods (usually five years) and the whole organization; tactical performance is less generic and focuses on one area or on a specific organizational segment and does not encompasses long periods (usually 1 to 3 years); operational performance is more analytical, it has a short term time reference horizon (up to one year), it focuses, for instance, on the activities, processes and operations.
Newbert (2008) cites several authors and typifies the performance measures into 3 types that are regularly used in the strategic literature: objective financial performance, subjective financial performance and subjective nonfinancial performance.
Matsuno, Mentzer and Özsoner (2002) proposed as variables: market share, percentage of new products sales on total sales and return of the investment. Rodrigues (2004) measured the performance with three dimensions: economic results, successful new products and relative market share. Han et al. (2007) used two indicators, profit and market share, measured in the course of various items. Metts (2007) measured the performance through three distinct dimensions: financial performance (market share, redraft, growth rate) operational performance (product quality, delivery, ease of product change) and organizational performance (customer satisfaction, collaborators satisfaction).
For the above we can say that there are different ways of measuring the performance. Each researcher uses the most appropriate measures of performance for the study that he undergoes. In this study, the measurement, performance is measured through the adoption of underlying concepts of the excellence criteria of EFQM (European Foundation Quality Management).

RELATIONSHIP BETWEEN TQM, INNOVATION AND PERFORMANCE

TQM and Performance

TQM and the performance of the organization have been under study by several investigators. One of TQM's greatest benefits is its emphasis on continuous improvement of business processes. The aim of TQM is to improve competitiveness, effectiveness and flexibility (Evans and Lindsay 2004).Although some authors recognize that no robust and consolidated evidence concerning the positive relationship between TQM and performance exists, there is consensus regarding the empirical validity of a positive effect of TQM on operational-type performance, such as productivity, flexibility, on-time delivery of goods and services, quality and customer satisfaction in general (Kaynak 2003). Agus and Abdullah (2000) and Hanet al. (2007) consider that practices of TQM have not directly improved the performance of the organization. However, according to many other authors (Ahire et al. 1996; Prajogo e Sohal 2003, 2004; Costa and Lorente 2004; Pinho 2007; Beck and Walgenbach 2009) the implementation of TQM practices can benefit many aspects of the organization and result in an improved performance. The empirical studies assessing the relationship between TQM and organizational performance have indicated strong and positive results.

TQM and Innovation

There are no consistent findings concerning the relationship between the adoption of principles of TQM and innovation. In some investigations the management for the total quality sustains innovation and there is a positive relationship between management practices of total quality and innovation (Zaire 1997; Prajogo e Sohal 2003, 2004). However, in other investigations, TQM can prevent innovation. Pinho (2007), in a study of Portuguese small and medium sized enterprises (SME`s) did not confirm a positive relationship between TQM and innovation. Abrunhosa and Sá (2008), in a study regarding the relationship between principles of TQM and innovation, found that not all principles of total quality play a major role on innovation. Vijande and González (2008), in an investigation that involved 93 SMEs from the manufacturing and service sectors of the autonomous region of Asturias, concluded that TQM itself is able to promote innovation at the management level. However, regarding the TQM effects on products and processes innovation, it is mediated by the company's culture favorable to innovation.

Innovation and Performance

There is a general agreement that entrepreneurial orientation, of which innovation is an integral part, influences the performance of the organizations (Miller 1983; Covin and Slevin 1989, Barrett and Weinstein 1998; Ferreira 2003; Rodrigues 2004), and the entrepreneurial companies will have better performances and better levels of innovative products (Miller and Friesen 1982). Barrett and Weinstein (1998) concluded that the relationship between entrepreneurial orientation and performance is direct and positive. According Damanpour (1996), innovation includes activities such as improving products, processes or procedures that aim to increase the value and performance of products, processes or procedures. Pinho (2007) argued that innovation in organizational administrative processes, leading edge technologies and ability to produce differentiated products contribute to improving the firm's performance.

METHODOLOGY

In such studies, in spite its presence, theory is not necessarily a starting point. The theoretical assumptions are being presented as the research evolves. The selection of cases is made on the basis of theoretical criteria and occurs throughout the study (Brannen 2005).

The research work was developed through the study of two ISO 9000 certified organizations. Two separate interviews were conducted with the senior manager and quality coordinator of each firm. Each interview typically lasted for two hours.

ISO certification guarantees, to a certain extent, that organizations have knowledge and interest on total quality management practices. It also guarantees familiarity with the quality concepts used in the research and particularly useful in interviews.

In the qualitative approach researchers are required to show flexibility and engagement. This is particularly evident in interviews in which the role of the researcher, in personal terms, is unparalleled in all social research (Easterby-Smith, Thorpe and Jackson, 2008). For this reason, it was especially important to develop an interview guide composed of three parts: TQM; Innovation; Performance. In each part it was intended to assess organization's evolution since the beginning of the ISO 9000 certification process, at the level of three considered dimensions.

TQM evolution was assessed through a set of 24 questions based on the "enablers" criteria of the EFQM (European Foundation Quality Management) excellence model. The EFQM excellence model is widely recognized as a powerful tool for self assessment of organizations and has been internationally used as a model by total quality management (Pires, 2006). The enablers criteria (leadership; people; strategy; partnerships and resources; processes, products and services) represent what the organization does, therefore it was considered appropriate to assess the organization's evolution concerning the total quality management. The evolution of the level of the organization's performance was similarly evaluated through the criteria results of the EFQM excellence model. The criteria results (customer results; people results; society results; key results) represent what the organization achieves. The evaluation was performed by 20 questions. Evolution of Innovation was assessed using five questions related to change and evolution demonstrated at the level of innovation, based on the scale developed by Covin and Slevin (1989).

RESULTS

The interviews conducted at the Fuel Company and the Logistics Company allowed for the registering of a set of statements on which the five propositions presented here, were based.

Relationship between TQM and Performance

The senior manager of the Logistics Company pointed out as the main difficulties of implementation of the quality management system the involvement of individuals and departments within the organization. The senior manager has stated the following: "the main difficulty that we have had is related to the involvement of all departments in the implementation process system. We have had some employees who were hard to convince. However, they seem now very interested and satisfied". For its part, the quality coordinator at the Logistics Company reported that the main difficulty was to involve the top management: "The difficulty mainly lies in the poor involvement of the top management, CEOs do not care much about these issues".

The senior manager of the Fuel Company stated that the implementation process of quality management system was a success. He valued the speed of the implementation process, the strong involvement of all employees and recognized major changes in the level of procedures. The senior manager stated: "Total quality management helped in creating procedures and processes. It has forced the organization to be more effective and efficient and was relatively fast". He even added: "Concerning customers, there is a relevant relationship between total quality management and their satisfaction. Now, customers seem more satisfied and believe more in ourselves".

The quality coordinator at the Fuel Company has stated that: "the implementation of total quality in our organization has enabled think organization in terms of management according to objectives and continuous improvement. The implementation of the quality management system can continuously improve the organization. Now we have a lot of concern in monitoring the company's quality objectives".

The senior management of Fuel Company indicated that despite some initial resistance, people have realized the importance of quality management system implementation and participated actively. He said: "After some initial resistance, there was a voluntary participation in the activities required to implement the quality management system such as elaboration of documents and internal auditors. People seem now more satisfied ". The Fuel Company quality coordinator stated: "The monitoring of processes enables us to attack certain situations we want improve and where we have some problems. For us the system is an important management tool".

All four respondents referred important benefits as result of the implementation of quality management system. The senior manager of Logistic Company stated: "the organized image of the organization that can be passed to the client and the clear and objective definition of vision, mission and values of the organization are, for me, the main advantages of the quality system". The quality coordinator of Logistic Company ruled as follows: "we have now an important management tool that allows evaluating the performance of the organization and enables the implementation of best practices in monitoring the operations. The senior manager of Fuel Company stated: "With the implementation and certification of quality management system, we can say that the organization meets international standards". The quality coordinator of Fuel Company ruled as follows: "Now we can demonstrate for our customers that we have quality in our organization".

The interviews suggest that the total quality management has influence on organizational performance. According to what was said by respondents the implementation of quality management system has enabled an improvement in results at the organization's processes. Thus, there might be a possible relationship between total quality management and results at the level of the organization's processes.

PROPOSITION 1: The total quality management can positively influence the level of the operating results of the enterprise

Despite the fact that collaborators did not seem very convinced when implementation of quality management system starts, after some time, they are completely surrendered to TQM. So, a second proposition can be stated:

PROPOSITION 2: Total quality management can positively influence the level of the "people results" of the enterprise

The senior manager of Logistic Company stated that: "I must emphasize improvement at the level of economic results such as profitability and sales volume and, above all, at the productive efficiency and organizational level". The opinion of the senior manager of Fuel Company is more or less the same: "I think we have good results from quality management. Quality certification was announced in local media and that was important to have more customers". From what was exposed, the third proposition can be proposed:

PROPOSITION 3: The total quality management can positively influence the level of the financial results of the enterprise.

Relationship between TQM and Innovation

The interview with the fuel company's manager suggests that the total quality management provides continuous improvement of procedures and monitoring of quality objectives. He said the following: "I can say that the quality management system and its network of processes have promoted a spirit of change and innovation in the organization. The network allowed a different view and understanding of some problems that existed. We developed a set of monitoring indicators of activities and requirements and strategies are defined based on outcome indicators".

In the logistics company, the senior manager reported the following: "given the business and market nature where the organization operates, the risk is explored through small but safe steps, reducing the probability of wrong decisions. The system started providing indicators to take the risk with greater security". About this, but in a more operational context, the quality coordinator said the following: "Despite the value of the product, the brand, the fact that the organization system is already well known, there has been a predominant focus on research and development issues. Total quality management is a powerful task for this".

According to what was mentioned above, the implementation of quality system has implications at the level of innovation. Thus, the fourth proposition emerges:

PROPOSITION 4: The total quality management can positively influence the level of organizational innovation of the enterprise

Relationship between Innovation and Performance

In interview the senior manager of Fuel Company stated that the changes in the process of delivery of fuel caused a led to improved outcomes in terms of customers. He reported that: "The service delivery of fuel has significantly improved. With the changes made at the level of procedures, it is easier to provide for the delivery. The delays were reduced and customers are more satisfied. I think it is a result of operational changes made". About openness towards innovation, the coordinator of quality said: "the analysis of indicators and our approach to innovation allowed examining trends, non-conformities and changes. The results have been continuously improved. We think that it is the reflection of the improvements that we introduce in form to work".

According to the quality coordinator of Logistic Company, "the firm has noted a strong evolution at the level of the costumer's satisfaction either by the drop of the number of complaints, or by the reduction of the reaction time (deadliness). These improvements may be a reflection of our careful posture to continuous improvement. We have been introducing changes and improvements in the procedures". Due to what was stated, the fifth proposition arises where a relationship between organizational innovation and its operational performance is expected.

PROPOSITION 5: The organizational innovation can positively influence the operational performance of the enterprise.

FINAL CONSIDERATIONS

This study contributes to the existing literature on relationship between the following variables: total quality management, innovation and performance. First, it identifies significant research issues and addresses previously unanswered questions. Second, it investigates the role of TQM in supporting innovation and organizational performance and the role of innovation in organizational performance in the context of Portuguese small and medium sized enterprises.

Based on literature review, it is noted that research results have not arrived at common conclusions. There are, however, a large number of researchers that identify the existence of a positive relationship between TQM and organization's performance. According to the interviews that took place in two organizations, this study concluded that TQM has an influence on the operational results and collaborators' performance and TQM also enabled the improvement at the level of the financial results. It is reasonable to conclude that the TQM positively influences the organizations' performance.

About the relationship between TQM and organizational innovation, the previous investigations did not show agreement in the results. In this exploratory research, it was noticed that in both organizations, the implementation of TQM practices allowed for the improvement at the level of the innovation attitude. So, it can be said that TQM positively influences innovation.

A large number of researchers found out that the innovation contribute to improving the firm's performance. In both firms, the exploratory interviews suggest that improvement in performance results occurs as a consequence of innovation posture in processes and procedures. Continuous improvement of the quality management system had positive consequences in the operational performance results. So, a positive relationship between innovation and the operational performance of the firms is confirmed.

Our study has several limitations. First, it focuses in only on two small and medium sized enterprises. It is unclear to what extend our results can be generalized. Second, it relies on perceptual data from senior manager and quality coordinator and might not give a complete and real picture of the organization. Third, it focuses on Portuguese firms. Given the internationalization of the standard ISO 9000 and the resulting quality management, it would be interesting to replicate the study in different regions and services.

For these reasons the establishment of a new exploratory study covering more and different organizations is suggested. In a second step, the propositions presented in this study and others must be empirically tested.

REFERENCES

Abrunhosa, A. and Sá, P. M. 2008. "Are TQM principles supporting innovation in the Portuguese footwear industry?", *Technovation*, Vol. 28, pp. 208-221.

Agus, A. and Abdullah, M. 2000. "The mediating effect of customer satisfaction on TQM practices and financial performance", *Singapore Management Review*, Vol. 22, No. 2, pp. 55-75.

Ahire, S., Golhar, D., and Waller, M. 1996. "Development and validation of TQM implementation constructs", *Decision Sciences*, Vol. 27, No. 1, pp. 23–56.

Anderson, S. W., Daly, J. D. and Johnson, M. F. 1999. "Why firms seek ISO 9000 certification: Regulatory compliance or competitive advantage?", *Production and Operations Management*, Vol. 8, No.1, pp. 28-43.

Barrett, H. and Weinstein, A. 1998. "The Effect of Market Orientation and Organizational Flexibility on Corporate Entrepreneurship", *Entrepreneurship Theory and Practice*, Vol. 23, No.1, pp. 57-70.

Bates, R. and Khasawneh, S. 2005. "Organizational learning culture, learning transfer climate and perceived innovation in Jordanian organizations, *International Journal of Training and Development*, Vol. 9, pp. 96–109.

Beck, Nikolaus and Walgenbach, Peter 2009. "The economic consequences of ISO 9000 certification in east and west German firms in the mechanical engineering industry", *Journal for East European Management Studies*, Vol. 14, No. 2, pp. 166- 186

Brannen, J. 2005. "Mixiing Methods: The Entry of Qualitative and Quantitative Approaches into the Research Process", *International Journal Social Research Methodology*, Vol. 8, No. 3, pp.173-184.

Costa, M. M. and Lorente, Á. R. M. 2004. "ISO 9000 as a tool for TQM: a Spanish case study", *The Quality Management Journal*, Vol. 11, No. 4; pp. 20-31.

Covin, J. G. and Slevin, D.P. 1989. "Strategic Management of Small Firms in Hostile and Benign Environments", *Strategic Management Journal*, Vol. 10, No. 1, pp. 75-87.

Damanpour, F. 1996. "Organizational complexity and innovation: developing and testing multiple contingency models", *Management Science*, Vol. 42 No. 5, pp. 693-716.

Easterby-Smith, M., Thorpe, R. and Jackson, P. R. 2008. *Management Research*, 3th Edition, London, UK, SAGE Publications.

Evans, James R. and Lindsay, William M. 2004. *The Management and Control of Quality*, South-Western College Pub, 6th Edition.

Ferreira, J. J. M. 2003. "Estudo do Crescimento e Desempenho da Pequena Empresa: A Influência da Orientação Estratégica Empreendedora", Tese de Doutoramento, Covilhã, PT: Universidade da Beira Interior.

Franco, V. S., Oliveira, Á. V. and Jesus, M. A. 2006. *Contabilidade de Gestão – Orçamento Anual e Instrumentos de Avaliação do Desempenho Organizacional*, Volume II, Lisboa, PT, Publisher Team.

Fuentes, M., Montes, F. e Fernández, L. 2006. "Total quality management, strategic orientation and organizational performance: the case of Spanish companies", *Total Quality Management and Business Excellence*, Vol. 17, No 13, pp. 303-323.

Garrido, E., Martín-Peña, M.L., and García-Muiña, F. 2007. "Structural and infrastructural practices as elements of content operations strategy. The effect on a firm's competitiveness", *International Journal of Production Research*, Vol. 45, No. 9, pp. 2119–2140.

Gotzamani, K. D., Theodoakioglou, Y. D. and Tsiotras, G. D. 2006. "A longitudinal study of the ISO 9000 (1994) serie`s contribution towards TQM in Greek industry", *The TQM Magazine*, Vol. 18, No.1, pp. 44-54.

Grant, R.M., Shani, R., and Krishnan, R. 1994. "TQM's challenge to management theory and practice", *Sloan Management Review*, Vol. 35, No.2, pp. 25–35.

Han, S. B., Chen, S. K. and Ebrahimpour, M. 2007. "The Impact of ISO 9000 on TQM and Business Performance", *The Journal of Business and Economic Studies*, Vol. 13, No. 2; pp. 1- 25.

Lenka, U. and Suar, D. 2008. "A holistic model of total quality management in services", *The Journal of Management Research*, Vol. 7, No. 3, pp. 56-72.

Matsuno, K., Mentzer, J. T. and Özsoner, A. 2002. "The Effects of Entrepreneurial Proclivity and Market Orientation on Business Performance", *Journal of Marketing*, Vol.66, No. 3, pp. 18-32.

Metts, G. A. 2007. "Measuring the effectiveness of managerial action in SME`s.: an empirical analysis of management`s response to industry competitive forces", *Management Research News*, Vol. 30, No. 12, pp. 892-914.

Miller, D. 1983. "The correlates of entrepreneurship in three types of firms", *Management Science*, Vol. 29, pp. 770-791.

Miller, D. and Friesen, P. 1982. "Innovation in Conservative and Entrepreneurial Firms: Two Models of Strategic Momentum", *Strategic Management Journal*, Vol. 3 (Jan-Mar), pp. 1-25.

Newbert, S. L. 2008. "Value, Rareness, Competitive Advantage, and Performance: A Conceptual-Level Empirical Investigation of the Resource-Based View of the Firm", *Strategic Management Journal*, Vol.29, pp. 745-768.

OCDE 2005. "Guidelines for Collecting and Interpreting Innovation Data", Oslo Manual, 3rd Edition, Paris, OCDE.

Pires, A. R. 2006. Qualidade - sistemas de gestão da qualidade, 3th Edition, Lisbon, PT: Edições Sílabo, Lda.

Pinho, J. C. 2007. "TQM and performance in small medium enterprises: the mediating effect of customer orientation and

Prajogo, D.I. and Sohal, A.S. 2003. "The relationship between TQM practices, quality performance, and innovation performance - an empirical examination", *International Journal of Quality and Reliability Management*, Vol. 20, No. 8, pp. 901–918.

Prajogo, D.I. and Sohal, A.S. 2004. "The multidimensionality of TQM practices in determining quality and innovation performance - an empirical examination", *Technovation*, Vol. 24, No. 6, pp. 443–515.

Rahman, S. 2001. "A comparative study of TQM practice and organizational performance of SME`s with and without ISO 9000 certification", *International Journal of Quality & Reliability Management*, Vol. 18, No. 1, pp. 35-49.

Rodrigues, R. 2004. "Orientação para o Mercado, Orientação Empreendedora e Desempenho nas PME Industriais Portuguesas: Exploração da Relação entre Marketing e Empreendedorismo", Tese de Doutoramento, Covilhã, PT: Universidade da Beira Interior.

Silva, M. J., Leitão, J. and Raposo, M. 2008. "Barriers to innovation faced by manufacturing firms in Portugal: how to overcome it for fostering business excellence?", *International Journal Business Excellence*, Vol. 1, No. 1/2 , pp. 92-105.

Singh, R. K. and Garg, S. K. 2008. "Strategy development by SMEs for competitiveness: a review", *Benchmarking: An International Journal*, Vol. 15, No. 5, pp. 525- 547.

Singh, P.J. and Smith, A.J.R. 2003. "Relationship between TQM and innovation: an empirical study", *Journal of Manufacturing Technology Management*, Vol. 15, pp. 394–401.

Vijande, M. and González, L. 2008. "Efectos de la Gestión de Calidad Total en la transformación en la innovación tecnológica y administrativa", *Cuadernos de Economia y Dirección de la Empresa*, Vol. 37, pp. 33-66.

BIOGRAPHY

ANTÓNIO FERNANDES is an assistant at School of Tecnology, Politecnic Institute of Castelo Branco. He is a PhD student of Management. His area of research is TQM and Innovation.

LUÍS LOURENÇO is an Associate Professor at Universidade da Beira Interior and research at NECE.

GT for CE: Operating Group Technology for Team Formation in Concurrent Engineering

Esra Aleisa
Dept. of Industrial Engineering
Kuwait University
e.aleisa@ku.edu.kw

Nallan C Suresh
Operations Management and Strategy
University at Buffalo
ncsuresh@buffalo.edu

Li Lin
Dept. of Industrial Engineering
University at Buffalo
lilin@eng.buffalo.edu

KEYWORDS
Concurrent Engineering, team formation, group technology (GT), cellular manufacturing, sequence-based clustering, fuzzy artificial neural networks

ABSTRACT

This research proposes a method for team establishment and organization in concurrent engineering (CE) context using cell formation techniques in Group technology. The team formulation relies on the resemblances in the overall structure of interdisciplinary designs teams in CE to multi-functional machine-cells in GT. This approach avails most cell-formation algorithms to be utilized for this purpose. To maintain task sequence in teams, we apply a sequence-based cell-formation algorithm to the task-team assignment in CE. The algorithm uses Fuzzy ART neural networks which is suitable for large scale CE applications. The approach was demonstrated on an illustrative manufacturing facility design problem.

INTRODUCTION

Concurrent Engineering (CE) has emerged as an essential design principle that facilitate rapid and efficient product development that are necessary to survive in today's fiercely competing environment. CE is by large is a managerial challenge (Ishii 1993; Yassine and Braha 2003). A successful CE application requires efficient and flexible formation and organization of design teams (Kusiak and Park 1990; Smith and Eppinger 1997; Yassine and Braha 2003; Chen and Lin 2004; Ke, Archetti et al. 2008) and utilization of the most advanced technology (Jin, Levitt et al. 1995; Singh 1996; Gardoni 2005). Trends in team formation for CE have been inclined towards supporting concurrent product development across geographically dispersed teams (Li and Qiu 2006) and encompassing human resource management, supply chain and marketing information into its formulation (Tseng and Abdalla 2006). Although there have been widely used tools to facilitate CE design for assembly (Boothroyd and Dewhurst 1983), design for serviceability (Makino, Barkan et al. 1989) and design for compatibility (Ishii 1993), there hasn't been tools specialized in interdisciplinary design teams' assignment and coordination suitable for CE (Jin, Levitt et al. 1995). In this regard, we intend to provide methods for efficient CE team formation, by utilizing the ample advances and tools from the field of Group Technology (GT). We achieve that by modeling design tasks as manufacturing parts and team-members as machines in cellular manufacturing (CM) settings (Aleisa and Lin 2009; Aleisa and Lin 2009). Such that, similar design tasks are grouped into clusters and modeled as part-families that need to be processed using a specific machine-cell. Where, a machine-cell corresponds to a particular design team in CE context. Sequence-based cell-formation in CM contains sophisticated methods that can tractably establish efficient team-tasks assignments without inducing feedback coupling, which will result in more realistic team formations. In this paper, we apply a Fuzzy Artificial Neural Network (FANN) algorithm that has been developed in Suresh *et al.*(Suresh, Slomp et al. 1999) and was originally intended for sequence-based part-families and machine-cell-formation. This algorithm was chosen mainly due to its ability to handle large industrial size problems without computational burden. In addition, we use this approach in this regard due to its added convenience to enabling analyst to adjust the similarity sequence threshold as appropriate and compare the results of team formations accordingly.

GROUP TECHNOLOGY AND CELLULAR MANUFACTURING

Originally proposed by Mitrofanov (1966) and Burbidge (Burbidge 1975), Group Technology (GT) is a manufacturing concept that is principled on grouping parts that are similar in their design, geometry or processing requirements (Irani 1999). Cellular manufacturing (CM) is the application of GT concepts on the facility shop-floor. CM is intended to facilitate and accommodate the norm of demands of greater variety and smaller order volumes (Salum 2000).

A basic step in CM is the recognition of machine components; in a process referred to as cell-formation. A manufacturing cell is an independent group of functionally dissimilar machines that manufacture a family of similar products (Waghodekar and Sahu 1984; Ham, Hitomi et al. 1985). Methods to form machine-cells and product families have been extensively addressed in the literature and usually employ two basic steps: defining a metric of similarity, and developing an algorithm for converting this measure to relationship among cluster (Mosier 1989). Along these lines a wide latitude of methods were developed for cell-formation which then has been classified to either product flow analysis (PFA), Array based methods (ABM), Similarity coefficient methods (SCM) and many others.

SEQUENCE-BASED CELL-FORMATION

Incorporating process sequencing requirements when identifying part-families and machine-cells has earned a considerable attention in the literature of cellular manufacturing (CM). This is due to its contribution in forming a more realistic machine-cells that have smoother flow patterns with less backtracking in the material handling system and reduction in work-in-process (WIP) levels (Suresh, Slomp et al.). Most research in sequence-based clustering proposes a mathematical formulation for incorporating more shop-floor operational characteristics such as set-uptimes, volumes, lot sizes, processing times, routings, etc. (Selvam and Balasubramanian 1985; Vakharia and Wemmerlov 1986; Dahel 1995; Hwang and Sun 1996; Jang and Rhee 1997; Nair and Narendran 1998; Yin and Yasuda 2002). Another, trend in this area is to propose a novel similarity measure based on sequence (Choobineh 1988; Tam 1990; Vakharia and Wemmerlov 1990; Kang and Wemmerlov 1993; Suresh, Slomp et al. 1999; Irani, Zhang et al. 2000). However, most sequence based clustering approaches are combinatorial and are tedious when applied to large-scale problems, which encouraged the resolution to AI-based techniques such as, neural networks and fuzzy theory (Kaparthi and Suresh 1991; Zhang and Huang 1995; Suresh, Slomp et al. 1999; Park and Suresh 2003; Ozdemir, Gencyilmaz et al. 2007; Won and Currie 2007; Mahdavi and Mahadevan 2008).

ADAPTING SEQUENCE–BASED CLUSTERING OF CM TO BE USED FOR TASK-TEAM ASSIGNMENTS IN CE

In this research we shall pursue Fuzzy neural network for sequence-based clustering developed by Suresh *et al.* (1999) to formulate task-team assignment problem in CE. This is due to its convenient matrix representation of processes' (task) sequences, which could be readily adapted for team formation. In addition, this approach enables the use of threshold values to manipulate the number of clusters. Also, it has the capacity to handle large industrial-type data sets with relatively low execution time, and superior solution quality. Experimental analyses that show the advantages of this approach are provided in (Park and Suresh 2003; Won and Currie 2007).

The Task-Team-Member Incidence Matrix

Here, we reformulate the classical cell-formation part–machine incidence matrix to represent task-team-member interaction in an attempt to utilize cell-formation techniques of CM to form teams in CE.

The classical part-machine incidence matrix is defined as:

$$a_{ij} = \begin{cases} 1 & \text{if part } i \text{ requires machine type } j \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

for all $i = 1,2,..n$ and $j = 1,2,..m$ we shall replace part i with task i and machine type with specialized team-member j. Accordingly, Eq.(1) will be modified to represent task-team-member incidence matrix $\mathbf{A}_{m\times n} = [a_{ij}]$:

$$a_{ij} = \begin{cases} 1 & \text{if task } i \text{ requires team member type } j \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

The above formulation allows general GT concepts and most common cell-formation methods such as rank ordering clustering algorithm (ROCA) (King 1980), Direct clustering algorithm (DCA) (Chan and Milner 1982), Bound energy algorithms (BEA), similarity coefficient based j approaches such as single linkage clustering algorithm (SLCA) (McAuley 1972), and others to be utilized for task-team assignment in CE (Pharn and Afify 2007). However, in CE, it is necessary to maintain task sequence requirements. Task sequence corresponds to part routings; therefore, cell-formation algorithms that emphasize part routing are necessary to create reasonable task-team assignments in CE.

Representation of Task Sequence Requirements

Routing sequence refers to the sequence of machine types required by a part (Vakharia and Wemmerlov 1990). To incorporate sequences, Suresh *et al.* (Suresh, Slomp et al.) have modified the traditional incidence matrix definition given in (1) as follows:

$$\bar{a}_{ij} = \begin{cases} k & \text{is the routing index by which part } i \text{ uses machine type } j \\ 0 & \text{if part } i \text{ does not require machine type } j \end{cases} \qquad (3)$$

Adapting the above equation according to the $\mathbf{A}$ matrix defined in Eq.(2), results the following sequenced-based task-team-member incidence matrix $\bar{\mathbf{A}}_{m\times n} = [\bar{a}_{ij}]$:

$$\bar{a}_{ij} = \begin{cases} k & \text{is the sequence index by which task } i \text{ need to be proccessed by team member } j \\ 0 & \text{if task } i \text{ does not require team member } j \end{cases} \qquad (4)$$

For instance, consider a task that requires the involvement of five team-members out of a total of seven individuals working on the same project ($m = 7$) according to the following sequence: {1,3,4,2,5}. Then, based to Eq.(4), the corresponding row of this task in the $\bar{\mathbf{A}}$ incidence matrix will be the following [1,4,2,3,5,0,0]. In addition, for convenience, we shall let $k = 1,2,..ni$, where ni denotes the final sequence index of a particular task i.

To prepare the above formulation for the task-team assignment algorithm we need to create a *precedence matrix* for each particular task i, which, is denoted by PM_i. Such that each PM_i is an $m \times m$ matrix, that depicts the directional relationships among all project individuals ($1 \le j \le m$) based on the premises of a particular task sequence using these team-members. Using the task-team-member framework, the entries of the $PM_i = [pm_{ipq}]$ is formulated as follows:

$$pm_{ipq} = \begin{cases} 1 & \text{if the contribution of team member } p \text{ have to precede the contribution of team member } q \text{ in operation sequence of task } i \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

Equation (5) forms a suitable platform to construct task-team-member assignments using CM sequence-based clustering such as Suresh *et al.* (1999) Fuzzy ART ANNs approach for sequenced based cell-formation or any other reasonable substitute thereof.

APPLYING FUZZY NEURAL NETWORK FOR SEQUENCE –DEPENDENT CLUSTERING TO TASK-TEAM ASSIGNMENTS OF CE

Fuzzy ART neural network is based on unsupervised learning (Carpenter, Grossberg et al. 1991). Therefore, the network exhibits no a priori knowledge of the number of task families (class) that exist nor that attributes of each cluster of similar tasks.

The layers of the Fuzzy ART network shown in Figure 1. As illustrated, it consists of two layers of neurons. The lower layer, the input or comparison layer, consists of neurons that are fully connected to the upper (output) layer through weight vectors. The lower layer interact with the routing input corresponds to the team-members ($1 \leq j \leq m$) that are required to work on some design tasks.

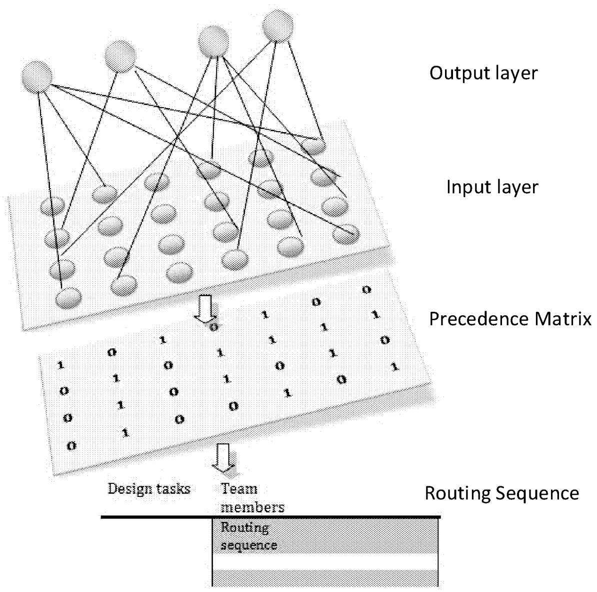


Figure 1: The layers for the Fuzzy ART network for sequence-based clustering

The Sequence-Based Clustering Fuzzy Art Algorithm

In this section, we explain the Fuzzy ART sequence based clustering algorithm that uses the PM_i to produce clusters of task families that require similar processing from team-members and accordingly generates suitable teams to assign to these tasks in CE settings. The algorithm starts by initializing: the weight vectors that connect upper and lower level neurons to the value of one; and the specification of the Fuzzy ART parameters. These are the choice parameter ($a > 0$), the learning parameter ($b = [0,1]$), and the vigilance parameter ($r = [0,1]$). In this application the vigilance parameter represents the threshold upon which exemplars are considered similar. That is the higher the value of r the larger he number of task families to be identified.

In the algorithm, $j = 1,2,.U$ correspond to the index of task families, where U is the total number of clusters of similar tasks identified. Within the algorithm, the array of task sequencing is fed to the algorithm and is converted to a precedence matrix, PM_i . Each PM_i is fed to the lower-level neurons (input neurons). Accordingly, the output value T_j is computed using:

$$T_j = \left[\left|\mathbf{PM} \wedge w_j\right|\right] / \left[a + \left|w_j\right|\right] \tag{6}$$

where $\wedge$ designate the Fuzzy AND operator: $(x \wedge y) = \min(x_i, y_i)$, and w_j denotes the weight of node j .

For every upper-level neuron, $j = 1,2,.U$, the neuron that exhibits the largest T_j and satisfies the pre-specified similarity parameter, r , is appended to the current family of tasks. The best-matching exemplar is updated accordingly by modifying the associated weight vector. This process is repeated until all task processing sequences are fed to the algorithm.

Experimental verification for this algorithm and the effect of the parameters a , b and r are provided in Suresh *et al.*(1999).

ILLUSTRATIVE EXAMPLE

In this section, we will apply the developed framework to of using sequence-based cell-formation Fuzzy ANT algorithm to the task-team assignment problem and team formation in the CE context. Consider the problem of a concurrent manufacturing facility design that consists of 43 tasks given in Table 1. These design tasks ($i = 1,2,...43$) need to be accomplished by thirty individuals ($j = 1,2,...30$) from eight departments ($D_1, D_2, ..D_8$), which respectively correspond to finance and administrative services, marketing, product design, process planning and scheduling, storage and warehouses management division, human resources and safety division, construction and project administration, and the architectural and infrastructure department. The distribution of individuals in each of the aforementioned departments is provided below:

$D_1 = \{1,2,3,4\}$ $D_2 = \{5,6\}$ $D_3 = \{7,8,9,10\}$

$D_4 = \{11,12,13,14\}$ $D_5 = \{15,16\}$ $D_6 = \{17,18,19\}$

$D_7 = \{20,21,22,23,24,26\}$ $D_8 = \{27,28,29,30\}$

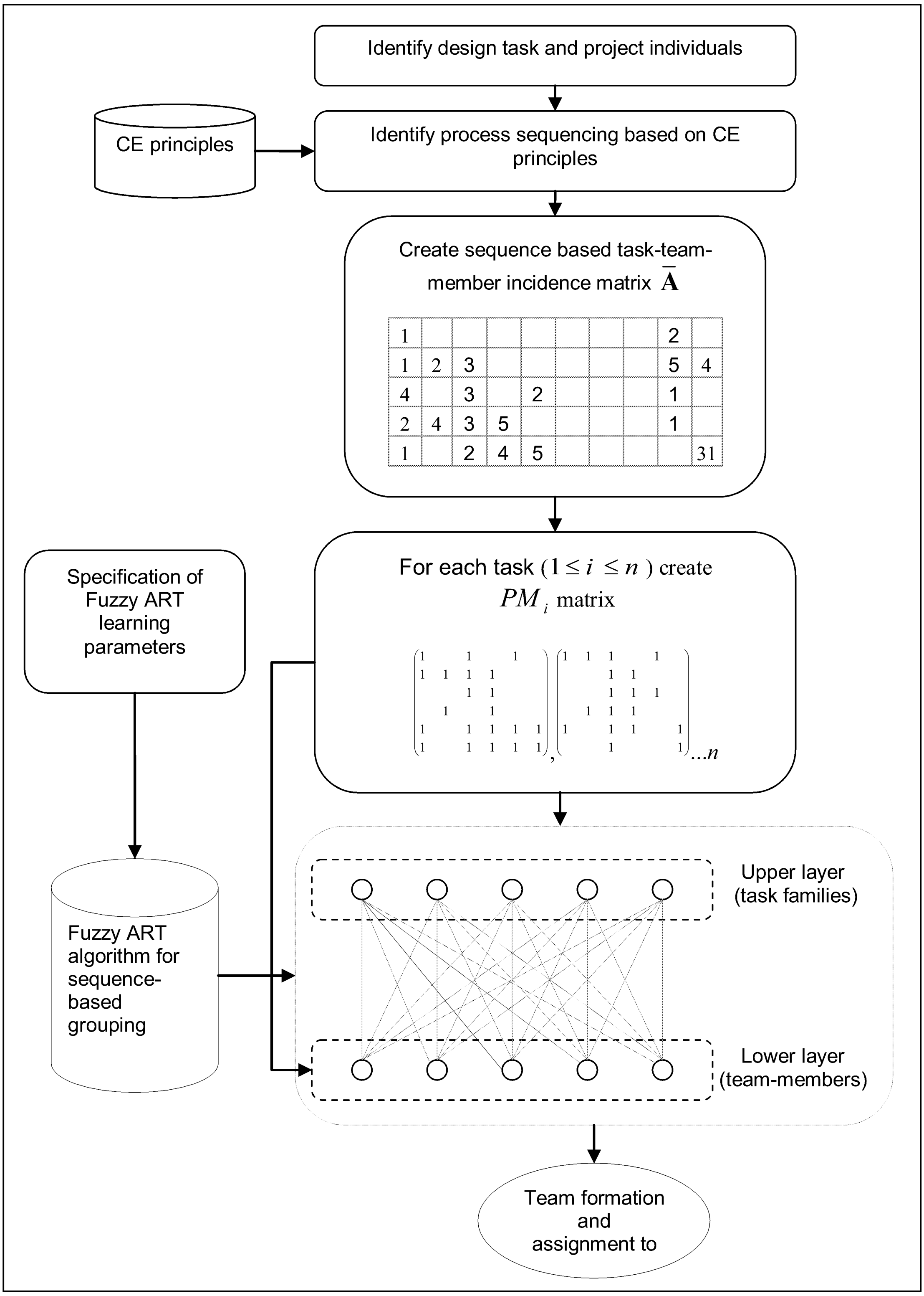


Figure 2: Team and task formation using sequence-based cell-formation fuzzy artificial network method

Application of the Fuzzy Art Neural Network Algorithm

In this section, we feed the routing sequence provided earlier to the Fuzzy Art neural network. The code for the neural network is written in Java programming language.

Prior to running the neural network, the choice parameter ($a > 0$), the learning parameter ($b = [0,1]$), and the vigilance parameter ($r = [0,1]$) need to be identified. As indicated earlier, the vigilance parameter represents the threshold upon which exemplars are considered similar. That is the higher the value of r the larger he number of task families to be identified.

Table 1: Tasks for designing a manufacturing facility

I	Task
1.	Definition of facility objectives, corporate strategy, customer, investment policy
2.	Specification of general product lines
3.	Planning of production and operation strategy
4.	Identification of product volumes and varieties and capacity growth planning
5.	Human Resource planning and labor force composition
6.	Component list of products
7.	CAD drawings of manufactured parts (dimensions, tolerances)
8.	material composition for manufactured parts
9.	Bill of Materials
10.	Design of manufacturing operations
11.	Type of workstations
12.	Number of workstations and preliminary assembly line balancing
13.	Determination of on-line/ off-line inspection methods
14.	Specification of shop floor control devices (Hand written report, manual data entry, bar codes, cameras, voice data entry systems, etc.)
15.	Production schedule design and policies (min WIP, idle time, setup time, average flow time or meet due dates, etc.)
16.	Layout type (product, processes, GT, CM, etc.)
17.	Conduct layout analysis (interrelationship , flow patterns, flow, layout method)
18.	Estimate WIP/Scrap (workstation)
19.	Layout simulation
20.	Determine non-production areas and their relationships
21.	Determination of required configuration non-production departments
22.	Transported material condition / orientation/ speed/ unit load size (workstation 1, workstation 2)
23.	Material handling equipment (unitizing, conveyors, sortation equipment, storage retrieval and stacking , transportation equipment and trucks)
24.	Degree of automation
25.	Product arrangement principles for warehouse and storage (popularity, similarity, size characteristics, etc.)
26.	Future space requirements for production area, warehouse and storage
27.	Amount of safety stock required for warehouse and storage
28.	Finished product material condition for warehouse
29.	Warehouse and storage operations
30.	Select storage equipment for warehouse and storage
31.	Aisle arrangement/ allowance, cubic space required for production area, warehouse and storage
32.	Shipping/receiving operations for warehouse and storage
33.	Design of docking areas for warehouse and storage
34.	Establishment of management policy for warehouse and storage
35.	Identify security precautions for plant
36.	Plan general accessibility means
37.	Design atmospheric system (regulating temperature, humidity, noise, etc.)
38.	Design structural system (steel skeleton, column spacing, height clearance, etc.)
39.	Design enclosure system (walls, roofs, doors, windows, etc.)
40.	Design electrical/lighting system
41.	Design sanitation system
42.	Prepare plot plans for facility total site
43.	Prepare blue print for facility total site

The result of using parameters $a = 0.1$, $b = 0.1$ and $r = 0.1$ results the creation of six work groups. The tasks within each work group need to be accomplished concurrently by a work team. In practical terminology, each work group corresponds to a work package that can be addressed separately i.e. by a separate team, department, company or a contractor.

Table 2: Task sequencing for design of manufacturing facility example

Task	Sequence	Task	Sequence
1.	<1,2,3,4,5,6>	2.	<1,2,3,4,5,6,7,8>
3.	<2,3,4,5,6,7,8,11,	4.	<1,2,3,4,5,6,7,8,11,12>
5.	<1,2,3,4,17,18,18	6.	<5,6,7,8,9,10>
7.	<7,8,9,10>	8.	<4,5,6,7,8,9,10>
9.	<5,6,7,8,9,10>	10.	<6,9,10,11,12,13,14>
11.	<,7,8,11,12,13,14	12.	<4,5,11,12,13,14>
13.	<9,10,11,12,13,14	14.	<3,7,8,11,12,17,18,19>
15.	<3,4,11,12,13,14>	16.	<3,4,6,7,8,9,11,12,13,14>
17.	<11,12,13,14>	18.	<11,12,13,14>
19.	<11,12,13,14>	20.	<17,18,18,20,21,27,28>
21.	<7,8,11,12,13,14,	22.	<9,13,14>
23.	<9,13,14,15,16>	24.	<4,7,8,11,12,16>
25.	<4,5,15,16>	26.	<15,16,1,2,5>
27.	<15,16,15,2,5,1>	28.	<7,15>
29.	<15,17,16,18>	30.	<15,17,16,18,2>
31.	<7,15,17,15,18,3,	32.	<11,15,12,16>
33.	<11,15,12,16>	34.	<2,3,15,17,16,18>
35.	<15,17,20,27>	36.	<15,17,27,20>
37.	<18,28,21>	38.	<19,29,22>
39.	<30,23>	40.	<19,29,22>
41.	<18,28,21>	42.	<28,26>
43.	<30,23>		

On the other hand, in another run using $a = 0.1$, $b = 0.1$ and $r = 0.4$ results in seven work groups/task families that are shown in Table 4 while using $a = 0.1$, $b = 0.1$ and $r = 0.7$ rusts in nine work groups that are shown in Table 5.

.

Table 3: Work groups for $a = 0.1$, $b = 0.1$ and $r = 0.1$

Work group or Task family	Tasks
1	1 2 3 4 5 8
2	6 7 9 10 11 13
3	12 14 15 16 17 18 19
4	20
5	21 22 23 24
6	25 26 27 29 30

Table 4: Work groups for $a = 0.1$, $b = 0.1$ and $r = 0.4$

Work group or Task family	Tasks
1	1 2 3 4 5 6 7 8 9 10 11 16 24
2	12 13 14 15 17 18 19 21 22 23 32 33
3	20 37 41
4	25 26 27 29 30 31 34 35 36
5	28
6	38 40
7	39 43

Table 5: Work groups for $a = 0.1$, $b = 0.1$ and $r = 0.7$

Work group or Task family	Tasks
1	1 2 3 4 5 8
2	6 7 9 10 11 13
3	12 14 15 16 17 18 19
4	20 35 36 37 41
5	21 22 23 24 32 33
6	25 26 27 29 30
7	28 31 34
8	38 40
9	39 43

From the preceding analyses, we can observe that by fixing a and b values and varying r between 0.1 and 0.7, the number of task families increased indicating that the degree of sequence similarity in more deliberate as r gets closer to 1. The values of these parameters need to be adjusted until a reasonable plan for work groups or task families is created based on reasonable work load balance that is distributed among work teams. For this particular example, with both a and b set to a value of 1.0 the rate of increase in the number of workgroups is depicted in Figure 3.

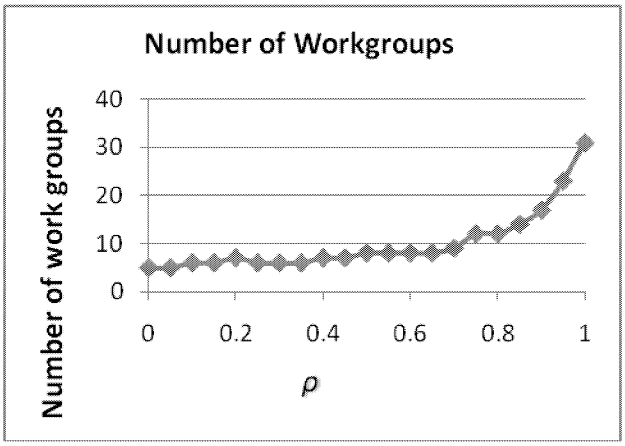


Figure 3: Number of workgroups with respect to parameters a and b set to a value of 1.0

RESULTS AND ANALYSIS

In this section, we conduct further analyses to investigate the effect of the learning, choice and vigilance parameters on the number of task families and the maximum number of tasks per family as these criteria mostly affects the workload balancing effort. The analyses are carried out using a multilevel factorial design that consists of three factors namely: a, b and r. Each of these factors consists of five levels: 0.1, 0.25, 0.5, 0.75 and 0.9. The design is replicated once. Two responses are associated, response 1: The number of task families and response 2: the maximum number of tasks per family. Thus the design is 5^3 full factorial design with 125 runs. Using a confidence level of 5%, the results shows that only b and r are significant with respect to both responses, while the a is insignificant. The regression model for the first response is:

$$\text{number of task families} = -1.69 + 0.09a + 14.4b + 21.6r \tag{7}$$

Where, S = 2.10612 R-Sq = 93.3% R-Sq(adj) = 93.2%

The regression model for the first response is:

$$\text{Max number of tasks in a family} = 18.1 - 0.297a - 6.10b - 11.7r \tag{8}$$

Where, S = 2.21707 R-Sq = 76.5% R-Sq(adj) = 75.9%.

Since only b and r proved to be significant with respect to both responses, here we show the surface plots for these factors with respect to both responses. These are shown in Figure 4 and Figure 5 respectively.

Figure 4 shows that higher values of b and r creates a larger number of parts families, while from Figure 5, smaller values of b and r results in bigger families. An expert investigation to the number and size of part families to promote a balanced workload among work teams for the illustrative example in hand indicates that 6-11 task families or teams is adequate. To obtain this level, we need to pick b and r values that will produce this range.

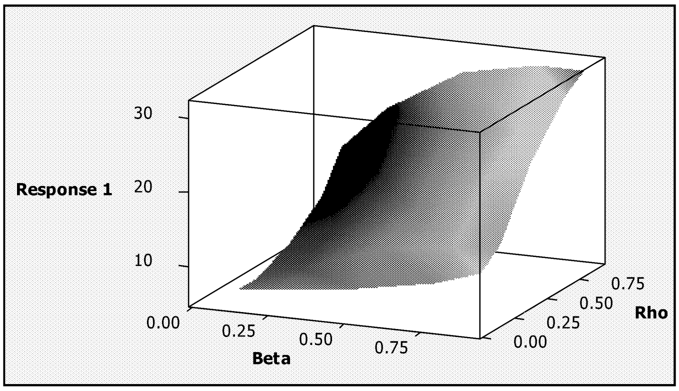


Figure 4: Surface plot for response 1 versus b and r

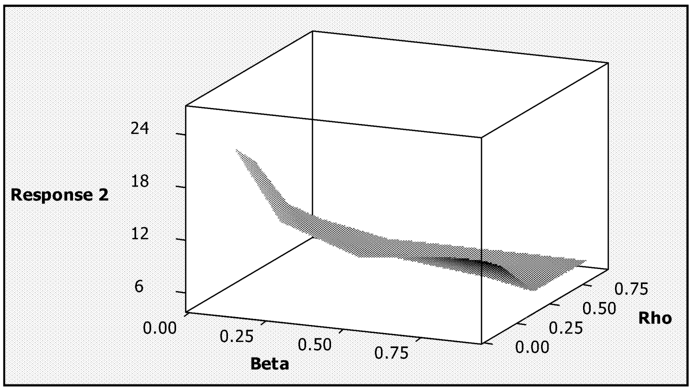


Figure 5: Surface plot for response 2 versus b and r

CONCLUSIONS

In this paper we have formulated the concurrent engineering task-team assignments as a cell-formation in cellular manufacturing, by revealing the relevance between the two problems' structures. Particularly, we have modeled interdisciplinary designs teams in concurrent engineering as multi-functional machine-cells in cellular manufacturing. This framework avails most cell-formation algorithms to be utilized for multidisciplinary team formation in CE. In addition, to create more realistic design teams, the paper applies a sequence-based cell-formation algorithm to the task-team assignment in CE. The algorithm uses Fuzzy ART neural networks which is suitable for large scale CE applications. The approach was demonstrated on an illustrative manufacturing facility design problem and the analyses were analyzed experimentally using a 5^3 full factorial design.

REFERENCES

Aleisa, E. and L. Lin 2009. "Accelerating Design Convergence Using the Degree of Permissible Concurrency". In *Proceeding of the International Conference on Information Science and Technology Applications (ISTA 2009)* Kuwait, 29-36.

Aleisa, E. and L. Lin 2009. "Sequencing of design tasks based on the degree of permissible concurrency". *International Journal of Applied Management Science (IJAMS)* No To appear.

Boothroyd, G. and P. Dewhurst 1983. *Design for assembly: A designer's handbook*. Boothroyd Dewhurst Inc., Wakerfield, RI.

Burbidge, J. L. 1975. *The introduction of group technology*. Wiley, New York,.

Carpenter, G. A., S. Grossberg and D. B. Rosen 1991. "Fuzzy ART. Fast stable learning and categorization of analog patterns by an adaptive resonance system". *Neural Networks* 7, No 6, 759-771.

Chan, H. M. and D. A. Milner 1982. "Direct Clustering Algorithm for Group Formation in Cellular Manufacture". No.

Chen, S.-J. G. and L. Lin 2004. "Modeling team member characteristics for the formation of a multifunctional team in concurrent engineering". *IEEE Transactions on Engineering Management* 51, No 2 May, 111-124.

Choobineh, F. 1988. "A Framework for the Design of Cellular Manufacturing Systems". *International Journal of Production Research* 26, No 7, 1161-1172.

Dahel, N. E. 1995. "Design of Cellular Manufacturing Systems in Tandem Configuration". *International Journal of Production Research* 33, No 8, 2079-2095.

Gardoni, M. 2005. "Concurrent engineering in research projects to support information content management in a collective way". *Concurrent Engineering-Research and Applications* 13, No 2, 135-144.

Ham, I., K. Hitomi and T. Yoshida 1985. *Group Technology: Application to production management*. Kluwer-Nijhoff), Hingham, MA.

Hwang, H. and J. U. Sun 1996. "A genetic-algorithm-based heuristic for the GT cell formation problem". *Computers & Industrial Engineering* 30, No 4, 941-955.

Irani, S. A. 1999. *Handbook of cellular manufacturing systems*. Wiley, New York, NY.

Irani, S. A., H. Zhang, J. Zhou, H. Huang, T. K. Udai and S. Subramanian 2000. "Production flow analysis and simplification toolkit (PFAST)". *International Journal of Production Research* 38, No 8, 1855-1874.

Ishii, K. 1993. "Modeling of concurrent engineering design".In: *Concurrent Engineering: Automation, Tools, and Techniques*. A. K. Kusiak, A. John Wiley & Sons, Inc., New York.

Jang, I. H. and J. T. Rhee 1997. "Generalized machine cell formation considering material flow and plant layout using modified self-organizing feature maps". *Computers & Industrial Engineering* 33, No 3-4, 457-460.

Jin, Y., R. E. Levitt, J. C. Kunz and T. R. Christiansen 1995. "Virtual Design Team: A computer simulation framework for studying organizational aspects of concurrent design". *Simulation* 64, No 3 Mar, 160-174.

Kang, S. L. and U. Wemmerlov 1993. "A Work Load-Oriented Heuristic Methodology for Manufacturing Cell-Formation Allowing Reallocation of Operations". *European Journal of Operational Research* 69, No 3, 292-311.

Kaparthi, S. and N. C. Suresh 1991. "Neural network system for shape-based classification and coding of rotational parts". *International Journal of Production Research* 29, No 9, 1771-1784.

Ke, L., C. Archetti and Z. Feng 2008. "Ants can solve the team orienteering problem". *Computers & Industrial Engineering* 54, No 3, 648-665.

King, J. R. 1980. "Machine-Component Group Formation in Group Technology". *Omega-International Journal of Management Science* 8, No 2, 193-199.

Kusiak, A. and K. Park 1990. "Concurrent Engineering - Decomposition and Scheduling of Design Activities". *International Journal of Production Research* 28, No 10, 1883-1900.

Li, W. D. and Z. M. Qiu 2006. "State-of-the-art technologies and methodologies for collaborative product development systems". *International Journal of Production Research* 44, No 13, 2525-2559.

Mahdavi, I. and B. Mahadevan 2008. "CLASS: An algorithm for cellular manufacturing system and layout design using sequence data". *Robotics and Computer-Integrated Manufacturing* 24, No 3, 488-497.

Makino, A., P. Barkan, L. Reynolds and E. Pflaff 1989. "Design for serviceability Expert System". In *SME Winter Annual Meeting: Concurrent Product and process design* San Fransisco, CA,

McAuley, J. 1972. "Machine grouping for efficient production". *Production Engineer* 51, No 53-57.

Mitrofanov, S. 1966. *The Scientific Principles of Group Technology*. National Lending Library Translation,

Mosier, C. T. 1989. "Experiment investigating the application of clustering procedures and similarity coefficients to the GT machine cell formation problem". *International Journal of Production Research* 27, No 10, 1811-1835.

Nair, G. J. and T. T. Narendran 1998. "CASE: A clustering algorithm for cell formation with sequence data". *International Journal of Production Research* 36, No 1, 157-179.

Ozdemir, R. G., G. Gencyilmaz and T. Aktin 2007. "The modified fuzzy art and a two-stage clustering approach to cell design". *Information Sciences* 177, No 23, 5219-5236.

Park, S. and N. C. Suresh 2003. "Performance of Fuzzy ART neural network and hierarchical clustering for part-machine grouping based on operation sequences". *International Journal of Production Research* 41, No 14, 3185-3216.

Pharn, D. T. and A. A. Afify 2007. "Clustering techniques and their applications in engineering". *Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science* 221, No 11, 1445-1459.

Salum, L. 2000. "The cellular manufacturing layout problem". *International Journal of Production Research* 38, No 5, 1053-1069.

Selvam, R. P. and K. N. Balasubramanian 1985. "Algorithmic Grouping of Operation Sequences". *Engineering Costs and Production Economics* 9, No 1-3, 125-134.

Singh, N. 1996. *Systems approach to computer-integrated design and manufacturing*. Wiley, New York.

Smith, R. P. and S. D. Eppinger 1997. "Identifying controlling features of engineering design iteration". *Management Science* 43, No 3, 276-293.

Suresh, N. C., J. Slomp and S. Kaparthi 1999. "Sequence-dependent clustering of parts and machines: a Fuzzy ART neural network approach". *International Journal of Production Research* 37, No 12, 2793-2816.

Tam, K. Y. 1990. "An Operation Sequence Based Similarity Coefficient for Part Families Formations". *Journal of Manufacturing Systems* 9, No 1, 55-68.

Tseng, K. C. and H. Abdalla 2006. "A novel approach to collaborative product design and development environment". *Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture* 220, No 12, 1997-2020.

Vakharia, A. J. and U. Wemmerlov 1986. "Designing a cellular manufacturing system: a cell layout approach using operation sequences". In *Proceesings of the Decision Science Institute* Hawaii, 1147

Vakharia, A. J. and U. Wemmerlov 1990. "Designing a Cellular Manufacturing System - a Materials Flow Approach Based on Operation Sequences". *Iie Transactions* 22, No 1, 84-97.

Waghodekar, P. H. and S. Sahu 1984. "Machine-Component Cell-Formation in Group Technology - Mace". *International Journal of Production Research* 22, No 6, 937-948.

Won, Y. and K. R. Currie 2007. "Fuzzy ART/RRR-RSS: a two-phase neural network algorithm for part-machine grouping in cellular manufacturing". *International Journal of Production Research* 45, No 9, 2073-2104.

Yassine, A. and D. Braha 2003. "Four Complex Problems in Concurrent Engineering and the Design Structure Matrix Method". *Concurrent Engineering Research & Applications* 11, No 3, 165-176.

Yin, Y. and K. Yasuda 2002. "Manufacturing cell's design in consideration of various production factors". *International Journal of Production Research* 40, No 4 Mar 10, 885-906.

Zhang, H. and S. Huang 1995. "Appliations of neural networks in manufacturing: a state-of-the-art survey". *International Journal of Production Research* 33, No 705-728.

ACKNOWLEDGEMENTS

The authors would thank the office of vice president for research and the research administration at Kuwait University for funding this project under Grant Number EI02/08.

AUTHOR BIOGRAPHY

ESRA ALEISA is an assistant professor of Industrial and Management Systems Engineering (IMSE) at Kuwait University. Her research interests include, Planning and design of large scale facilities, simulation and improvement of manufacturing and service systems, multilevel planning and design of complex engineering design, group technology (GT), and design structured matrices (DSM). She is a member of Omega Rho, the international operations research honor society, IEEE, INFORMS, IIE, ASEE.

NALLAN C. SURESH is UB Distinguished Professor and Chair of the Department of Operations Management & Strategy at UB School of Management. He specializes in the areas of manufacturing, supply chain & logistics management. His research has been in four broad streams, including: 1) Economic evaluation of flexible manufacturing systems (FMS) and robotics; 2) Design and performance evaluation of cellular manufacturing systems; 3) Design and Performance of virtual manufacturing cells, and 4) Design and performance of global manufacturing and sourcing networks, and avoidance of disruptions in global supply chains.

Li Lin is Professor of Industrial and Systems Engineering at University at Buffalo. His research areas include manufacturing and healthcare systems simulation and concurrent engineering and design, including environmentally conscious design and manufacturing. Dr. Lin research has been supported by the National Science Foundation (NSF), the Environmental Protection Agency (EPA) and the Agency for Healthcare Research and Quality (AHRQ).

THE INFLUENCE OF THE ORDER RELEASE FREQUENCY ON JOB SHOP LOGISTIC PERFORMANCE

Nuno O. Fernandes[1,2] & S. Carmo-Silva[2]
[1]Polytechnic Institute of Castelo Branco, School of Technology
Av. Empresário, 6000-767, Castelo Branco, Portugal
[2]CITEP, University of Minho
Campus de Gualtar, 4710-057, Braga, Portugal
E-mail: nogf@ipcb.pt; scarmo@dps.uminho.pt

KEYWORDS
Workload control, release frequency, simulation.

ABSTRACT

Workload Control is a Production, Planning and Control concept designed for Small and Median sized Enterprises in the make-to-order sector. Traditionally, release decisions within the WLC concept are taken periodically, e.g. weekly or daily. However, the periodicity of releases of new orders into the shop floor can have an important influence on system performance. This paper investigates the impact of the order release frequency on the logistic performance of a make-to-order job shop. Two dispatching rules, i.e. a flow conserving rule and due date oriented one, and two release strategies, unrestricted and workload controlled, are investigated using simulation. The experimental results show that the best job shop operation strategy, which combines order release and dispatching procedures, depends on the release frequency.

INTRODUTION

For manufacturing enterprises to stay competitive in the global marketplace of today, with an increasing demand for more customized and unique products, manufacturing strategies have to be focused on improving the logistics' performance objectives, namely reducing delivery times and improving due date adherence. Many approaches to improve performance are not practical for Small and Median sized Enterprises (SMEs) and/or make-to-order (MTO) production, which accordingly to Thurer et al. (2011) represents an important sector of the economy.

Workload Control is a Production, Planning and Control (PPC) concept primarily designed for SMEs in the MTO sector that can contribute to achieve the above referred objectives. WLC emphasis is on firmly controlling order flow times through the production system by means of input and output control decisions at four hierarchical levels: namely customer enquiry, order entry, order release and priority dispatching.

The focus of this paper is on the order release and dispatching decisions.

Order release has been described as an essential decision function and a core part of WLC (Missbauer 2009). Order release determines the amount, the type and the timing of release of new orders into the shop floor. A pre-shop pool, where the orders wait for a release decision, and a release decision mechanism are used for regulating the congestion on the shop floor. By setting an upper limit to the workload on the shop floor, WLC prevents the so called lead time syndrome (Plossl, 1988). While in the pool, unexpected changes to orders quantity and product design specification (minor changes) can be accommodated (Hendry et al., 2010).

Traditionally, release decisions within the WLC concept are taken periodically, e.g. weekly or daily. The length of the release period, i.e. the release frequency, influences the amount of work that is released into the shop each time order release is activated. Land (2006) and Fernandes and Silva (2011) explain that the choice of an appropriate period between releases is a delicate decision. Although a long release period leads to increased opportunities to find orders in the pool which fit workload norms and seek good load balancing, it also may unnecessarily delay orders in the pre-shop pool and increase the time orders spend in the entire system, i.e. the system flow time. A short release period, on the other hand, may hinder the release of orders that could contribute to stable workloads and of orders with high processing content and long routings, negatively affecting the timing of release. However, it has the advantage of leading to a more frequent updating of the shop floor situation which is likely to improve smoothing and stabilization of workload at capacity groups (e.g. machines or work centers), and also ordes due date adherence.

Once an order is released, processing routings determine the movement of orders through the shop floor. Here priority dispatching rules are used to control the progress of orders. Since WLC keeps the length of queues on the shop floor short and stable, and the most urgent orders are first considered for release, it has been suggested in the literature (Bechte, 1994) that WLC allows for the use of a simple dispatching rule such as First-Come-First-Served (FCFS). FCFS is a flow conserving rule that enables a smooth and constant progress of orders, supporting the predictability of flow times, which are used to establish more accurate planned release times for the orders. Due date oriented rules may also be used. These are intended to reduce the lateness variation across orders on the shop floor and thus to reduce the percentage of tardy orders. Note that reducing the percentage of tardy orders can result from both: (1) reducing the average flow times; and (2) reducing the variance of lateness across orders.

This paper investigates the impact of the release frequency of new orders on the logistics' performance of a make-to-order job shop. In particular it investigates its influence on the release strategy (unrestricted or workload controlled) and on the dispatching rules (flow conserving or due date oriented) to use in order to reduce the percentage of tardy orders. The main aim of the paper is to contribute to the theoretical development of the WLC concept providing valuable insights for its design in practical situations.

The remainder of the paper is structured as follows. The next section presents the simulation model and the experimental design and the following section analyses the results from the simulation experiments. In last section of the paper, concluding remarks and directions for future research work are put forward.

SIMULATION STUDY

The simulation study was carried out using the Arena® software. During simulation runs, data was collected under system steady-state. The length of each run was for 50,000 time units including a warm-up period of 10,000 time units. The average values of 100 independent replications are presented as results.

The following sections details the simulation model, the experimental design and the performance measures used in the study.

Simulation model

The study is based on a small job shop model used in several simulation studies of the WLC concept (e.g. Land and Gaalman 1998, Oosterman et al. 2000, Thurer et al. 2011). The shop consists of six work centres or capacity groups, each with a single machine. Machines capacities are equal and remain constant over time. A machine can perform only one operation at a time on any order (job) and an operation of an order can be performed by only one machine at a time. Job pre-emption is not allowed in the simulation. The routing lengths are uniformly distributed between one and six operations, without return visits. Processing times are stochastic, following a 2-Erlang distribution with a mean of one time unit per order. Processing times are assumed to be identical for all the operations of an order. The mean inter-arrival time of orders is set to result in a machine utilisation rate of 90%.

We assume that orders coming from a planning system or directly from costumers flow into the pre-shop pool. An exponential distribution was used to model the inter-arrival times of the orders to the pool. The due dates of orders are set by assigning to the order arrival time a discrete uniform distributed allowance between 35 and 60 time units.

Orders in the pool are reviewed periodically, with intervals of length T, and considered for release accordingly to its latest planned release date. This is determined for each order by backward scheduling from the due date using orders lead times for each work centre processing the order. An order is released only if, as consequence, the accounted workload of any of the work centers in its routing doesn't exceed an upper limit - usually referred as the work centre *workload norm.* The corrected aggregate load method suggested by Oosterman et al. (2000) was used for workload accounting. Table 1 resume model assumptions.

Table 1: Model assumptions.

Shop type	Job shop
Routing sequence	Random, no return visits
No. of machines	Six
Operations per order	Discrete uniform (1, 6)
Operations proc. times	2-Erlang, μ=1 time unit
Machine utilisation rate	90%
Interarrival times	Exponential distributed
Due-date allowance	Uniform (35, 60) time units

Experimental design

Table 2 summarises the experimental setting. Three experimental factors were evaluated, namely: the workload norm levels; the release period length and the dispatching rules.

Table 2: Experimental setting.

Factors	Levels	
Workload norm levels	Eleven levels of restriction	
Release period length	1 time unit	5 time units
Dispatching rules	FIFO	OSD

Workload norms were tested at eleven levels of restriction, including unrestricted release. These are deterministic parameters, setting the maximum workload that can be released to each capacity group.

The release period length (T) determines the time interval between order release activations, influencing the amount of work that is released into the shop floor each time order release is activated. It was tested at two levels, namely: 1 (high releasing frequency) and 5 time units (low releasing frequency).

Dispatching rules used for prioritizing the orders on the shop floor, were tested at two levels, namely: FCFS (First-Come-First-Served) and OSD (planned Operation Starting Date). FCFS is a flow conserving rule that selects the orders by the sequence they arrive at a machine or capacity group, i.e. the sequence of the orders in the outgoing flow at a capacity group is the same as the sequence in the incoming flow. OSD is a due date oriented rule that selects orders according to its lowest value, which is determined for each machine or capacity group by backward scheduling from the due date using the lead times of the operations still to be done.

Performance measures

The percentage of tardy orders, i.e. the percentage of orders delivered after the due date, is one of the key indicators for delivery performance.

The standard deviation of the lateness was also used as a measure of timing performance and due date adherence. It indicates how close the completion times of orders are to their due dates.

The system flow time is another performance measure used in the study that accounts for the time an order spends waiting in the pre-shop pool plus on shop floor. The latter is the time elapsed between job release and job completion and is referred as shop flow time.

SIMULATION REULTS AND DISCUSSION

The findings of this study are summarized and discussed below.
An overview of the system performance under the two dispatching rules (FCFS and OSD) for the two release period lengths (1 and 5 time units) is presented in Figures 1 to 3. Figure 1 shows the percentage of tardy orders for different values of the shop flow time, while Figures 2 and 3 show, for the same shop flow time values, the system flow time and the standard deviation of the lateness, respectively.

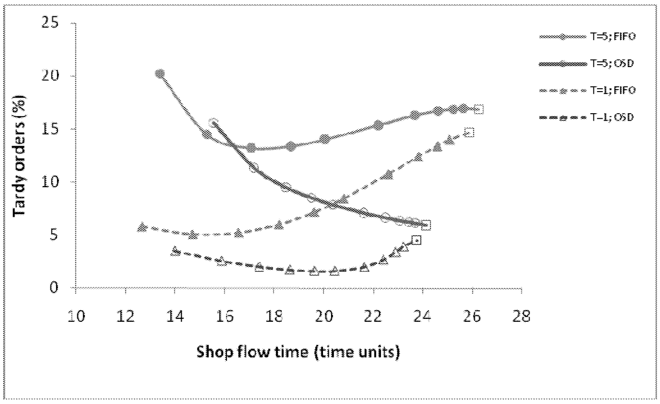


Figure 1: Percentage of tardy orders.

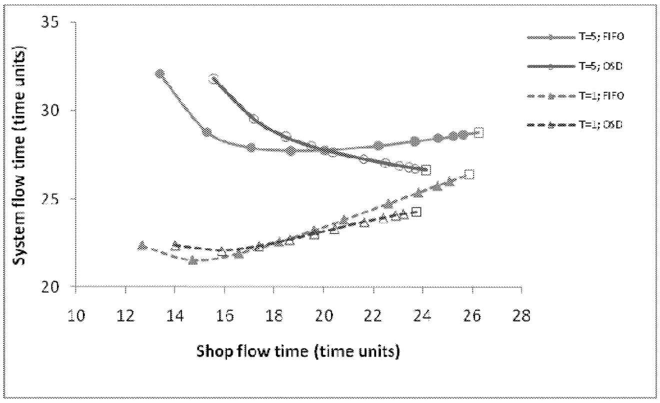


Figure 2: System flow time.

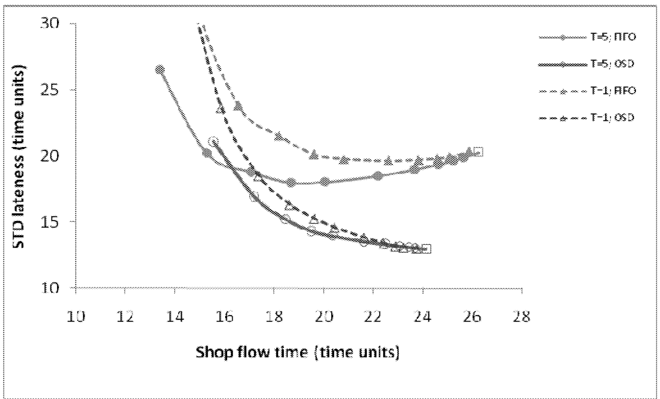


Figure 3: Std. deviation of the lateness.

Points on each logistic performance curve represent the eleven workload norm levels simulated. The utmost right point of each curve, represented by a square mark, results from an unrestricted norm level – usually called unrestricted periodic release. This means that each time order release is activated all the orders in the pre-shop pool are released to the shop floor.
Figure 1 clearly indicates that the best strategy in terms of the percentage of tardy orders depends on the release frequency, i.e. on the release period length. If the release period is short (1 time unit) then the lowest percentage of tardy orders is achieved under due date oriented dispatching and controlled order release. However, if the release period length is long (5 time units) then unrestricted release combined with due date oriented dispatching results in the best performance. In both situations the lowest percentage of tardy orders is obtained for due date oriented dispatching. The robustness of this rule towards due date adherence is reinforced by the fact that it also performs better than the flow conserving FCFS rule for the standard deviation of lateness, particularly for high levels of workload on the shop floor.
As explained in the introductory section of the paper a decrease on the percentage of tardy orders can result from both, a decrease on the average flow times and a decrease on the variance of lateness across orders. According to Figures 2 and 3, average system flow time decreases when frequency increases. With standard deviation of lateness the behaviour is opposite, i.e. increases with frequency. This means that the observed lower percentage of tardy orders for a release period length of one time unit results from the decrease of the average system flow times.
An important conclusion is that under the circumstances of low release frequency the best values for the percentage of tardy jobs and standard deviation of lateness is obtained for unrestricted order release, provided that OSD is adopted on the shop floor. This results may explain why some researchers, in their simulation studies, have found limited support to the use of workload controlled order release (Sabuncuoglu and Karapnnar 1999). This is usully known as the WLC paradox, referring to differing performance between WLC in theory and in practice.

CONCLUSIONS

In this paper an investigation has been made into the impact of the order release frequency on the due date adherence, measured as a function of the percentage of tardy orders and standard deviation of lateness, in a make-to-order job shop. Two dispatching rules (flow conserving and due date oriented) and two release strategies (unrestricted release and workload controlled) were studied in this investigation. Results show that due date oriented dispatching (OSD) performs better than flow conserving dispatching (FIFO) in terms of the percentage of tardy orders. However, the best strategy for order release clearly depends on the release frequency. When the release frequency is high, controlled order release gives the lowest percentage of tardy orders. On the other hand, if the release frequency is low, unrestricted order release performs best.

These findings can be seen as an important contribution for the design of production planning and control systems in practice. Future research work should address new issues of practical relevance, such as batch overlapping and batch splitting that have been overlooked in the literature of workload control. These are mecanisms for concurrent manufacturing that can have a geat influence on customer service related performance, such as the ones here addressed.

REFERENCES

Bechte, W. 1994. "Load oriented manufacturing control just in time production for job shops." *Production Planning and Control* 5(3), 292-307.

Fernandes, N.O. and Carmo-Silva, S. 2011. "Order release in a workload controlled flow-shop with sequence-dependent set-up times." *International Journal of Production Research 49,* No.8 (April), 2443-2454.

Hendry, L., Huang Y. and Stevenson, M. 2010. "Workload control: successful implementationtaking a contingency-based view of production planning & control ." *LUMS working paper 2010/50*.

Land, M. 2006. "Parameters and sensitivity in workload control." *International Journal of Production Economics* 104, No. 2, 625-338.

Land, M. and Gaalman, G. 1998. "The performance of workload control concepts in job shops: Improving the release method." *International Journal of Production Economics* 56-57, No. 20, 347-364.

Missbauer, H. 2009. "Models of the transient behavior of production units to optimize the aggregate material flow." *International Journal of Production Economics* 118, 387-397.

Oosterman, B. J., Land, M. J. and Gaalman, G. J. C. 2000. "The influence of shop characteristics on workload control." *International Journal of Production Economics* 60, 107-119.

Plossl, G.W. 1988. "Throughput time control." *International Journal of Production Research 26,* No.3, 493-499.

Sabuncuoglu, I. and Karapmnar H.Y. 1999. "Analysis of order review/release problems in production systems.", *International Journal of Production Economics* No.62, 259-279.

Thurer, M., Silva C. and Stevenson, M. 2011. "Optimising workload norms: the influence of shop floor characteristics on setting workload norms for the workload control concept." *International Journal of Production Research* 49, No.4 (Feb), 1151-1171.

LOCATION OF INSPECTION STATIONS IN MULTISTAGE MANUFACTURING PROCESSES

Ali G. Shetwan
Valentin I. Vitanov
Design Manufacture and Management,
School of Engineering, Durham University
South Road, DH1 3LE, UK.
E-mail: a.g.shetwan@durham.ac.uk v.i.vitanov@durham.ac.uk

KEYWORDS
Manufacturing, Quality, Resources Allocation, Inspection Optimisation.

ABSTRACT

In multistage manufacturing processes it is common to locate inspection stations after some or all of the processing workstations. The purpose of the inspection is to reduce the total cost of the manufacturing which is defined as the sum of the costs of production, inspection and failures (during production and after shipment). The objective of this paper is to find optimum allocation of inspection stations in a multistage manufacturing process such that the total manufacturing and quality cost can be reduced. The objective will be achieved under the assumption of limited number of inspection stations are available. A mathematical model for solving the inspection allocation and assignment problem in a multistage manufacturing system is developed. In the allocation problem when the number of workstations increases the processing time to solve the problem grows exponentially and the problem becomes NP hard problem. Max-Min Ant system (MMAS) algorithm is proposed to solve the formulated problem. The MMAS augmented with simple but very effective local search method. The performance of the proposed MMAS algorithm is compared to optimization method based on complete enumeration method (CEM). The experimental results show that the proposed MMAS algorithm produces solution near to the optimal with less time particularly when the number of workstations increased significance.

INTRODUCTION

The quality management policies in the majority of companies evolve continuously over a number of years by focusing on quality issues that are critical at any given instant of time. This approach usually focuses on particular critical operations and does not take into account the need for global analysis of a manufacturing system quality problem. As a result of that quality policies do not utilise fully the available financial, human and equipment resources. In the same time, present economic environment, the reduction of waste becomes of paramount importance because the increase in the product cost affects the overall competitiveness of the manufactured products. This paper targets specifically wastes resulting from unidentified defective items being processed unnecessarily during following manufacturing operations. The solution of inspection effort allocation issues needs to adopt the corresponding utility strategies. Such strategies aim to allocate an economically appropriate level of inspection effort by striking a balance among the different cost components connected with inspection, scrap, repair and replacement due to quality failure, and/or the warranty penalty in the case where a nonconforming product has been shipped to customers. The objective of this paper is to propose a methodology that tries to allocate limited inspection stations in a multistage manufacturing process in order to minimise the total manufacturing cost of a product without affecting with quality of the product.

INSPECTION ALLOCATION PROBLEM

Problem Background

The presented paper attempts to cover the major sources for the last fifty years that were available to the authors. The procedure of making decisions of whether or not to inspect a final or semi-finished product at every processing workstation is shown schematically in Figure 1. It is assumed that if inspection is performed after every workstation, the scrap and rework costs will stay at a minimum level. These savings have to be considered against the inspection costs which include equipment, staff, time, shop floor space and increase the number of works in process. Therefore, if these in process inspections are performed too often unnecessary costs will occur. In a practical situation usually the expectation is that manufacturing processes at each work station are capable of achieving the required quality tolerances for 99.73% of the items which is the 3σ level when the process is on target. It is possible that a shift of up to 1.5σ from the target quality level may occur undetected subject to the quality procedure which may lead to drop in quality level down to 93.3% and when such situation occurs our inspection stations have to be positioned in a way that will minimise the overall lost of profit. If the shift is larger than that we consider that the problem is not any longer a quality monitoring but a manufacturing issue that has to be fixed by different means. The purpose of inspection allocation strategy is to allocate an economically appropriate level of inspection activity by determining the correct balance among different cost components indicated above.

LIETRATURE REVIEW

Models of the Inspection Allocation Problem

The allocation of inspection effort in a serial manufacturing systems line has been extensively studied. The aim of almost all the proposed inspection models is to determine whether inspection operations should be performed immediately following each manufacturing stage to minimize the expected total cost. One of the first models to determine the fraction of production to be inspected at the various stages of a serial production system was proposed by Lindsay and Bishop (1964). In their model, the levels of inspection are determined by reducing the sum inspection and scrap unit costs. Hurst (1973) proposed the first model which accounted for two types of inspection errors: acceptance of nonconforming units and rejection of conforming units. The production scheme is assumed to be serial with only one inspection operation possible after each processing stage, with units perceived to be nonconforming separated from the production flow. Ballou and Pazer (1985) developed a computer program to perform a *what-if* simulation analysis of the serial systems with inspection errors. In a sequence of experiments, the authors found that inspection error rates have a major impact on cost. Type I errors (rejection of conforming units) were found to have a greater impact than type II errors (acceptance of nonconforming units). Multiple or repeated inspection provides one way to reduce the effects of inspection errors. A review of the literature on models for optimal allocation of inspection effort in multi-stage systems appears in Raz (1986). Lee and Chen (1998) examined the performance of inspection station allocation through heuristics rules on the basis of the parameters considered in serial production systems under different operating conditions. They proposed simulation scenarios examine the allocation and sequencing of the inspection operations with regard to the incurred costs of quality assurance and the requirements of efficient system operation in flexible manufacturing system FMS. Genetic algorithm was applied by Shiau et al. (2007) with realistic unit cost model to solve the manufacturing resource allocation problem by concurrently performing process planning and inspection planning in an advanced manufacturing system AMS. Large proportion of the studied papers focuses on several work stations representing part of a manufacturing line without attempting to solve the global optimisation problem which lead to solutions based on techniques such as dynamic programming and heuristic methods that are known to be computationally ineffective when the number of workstations increase which is considered as a computationally hard problem. However, in this paper, an attempt has been made to explore yet another application of MMAS to the AIE problem.

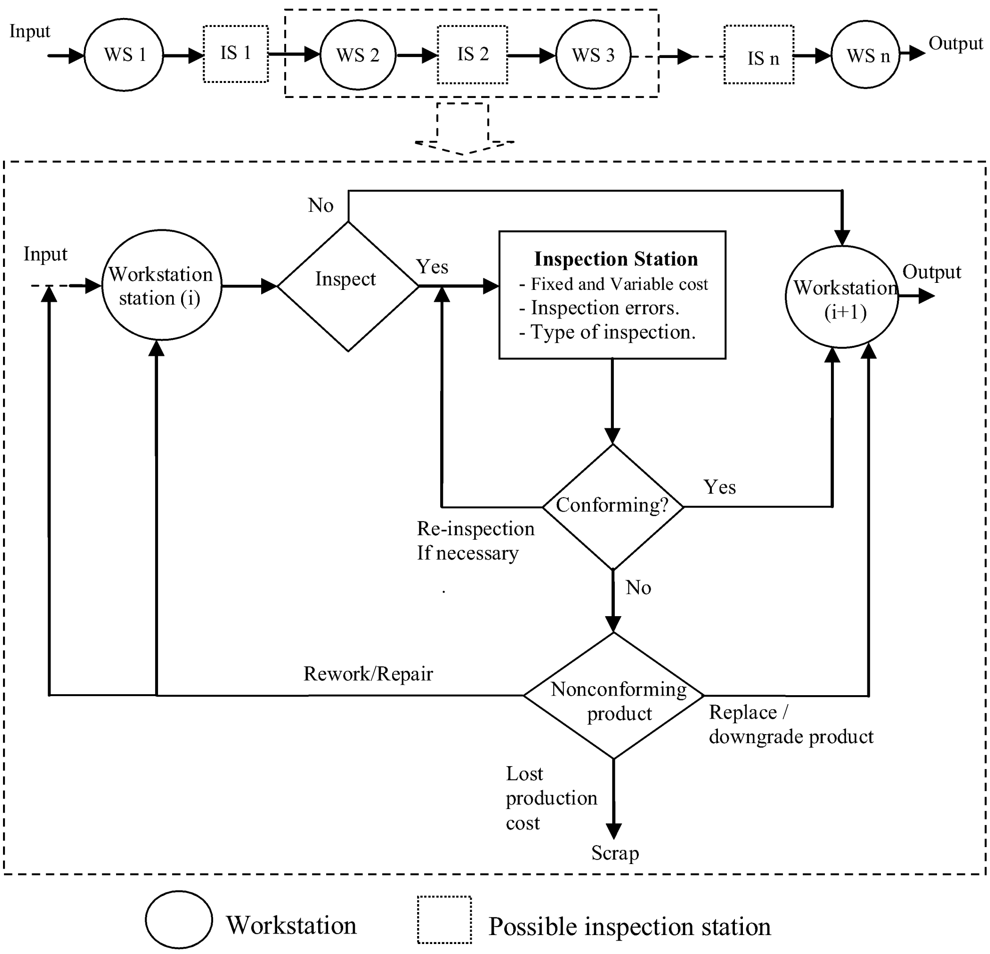


Figure 1: Inspection Allocation Problem in Multistage Manufacturing Processes

MODEL DEVELOPMENT

Model Description

To examine the inspection allocation problem, serial multistage manufacturing system has been studied under consideration as follows:

1. The system is considered to be made up of different (15 to 19) workstations arranged serially and parts are entering the system in batches.
2. Each workstation has a specific probability of producing defective parts.
3. A 100% inspection screen is applied to all parts processed in workstation if an inspection station is performed after it in the sequence.
4. Only one final product is considered in the system.
5. Two types of inspection errors are considered in the system. A type I error involves the classification of a conforming unit (CU) as a nonconforming unit (NCU), and a type II error means the classification of an NCU as a CU.
6. The system has limited number of inspection stations (e.g. five stations). Each inspection station can be assigned to perform inspection operation for one or more workstations.
7. Nonconforming items can either be scrapped or sent for rework as it is shown in Figure 1.
8. A probability of generating nonconforming items is assigned to each inspection station of selecting items for rework.
9. Each workstation has a specific probability of producing defective parts.

Description of Indices

m	Refers to the inspection station assigned.
k	Refers to workstation.

Description of Inputs

n:	Number of workstations in the system.
Q:	Number of parts entering the system.
q:	Number of inspection stations in the manufacturing system.
Z:	Probability of a nonconforming of part processing at the k workstation.
β:	Probability that the m inspection station incorrectly classifies a NCU as a CU.
α:	Probability that the m inspection operation incorrectly classifies a conforming unit (CU) as a nonconforming unit (NCU).
IFC:	Internal failure cost.
EFC:	External failure cost
MC:	Manufacturing cost.
IC:	Inspection cost.

Description of Variables

NG_k:	Number of conforming parts leaving the *k* workstation.
ND_k:	Number of defective parts leaving the *k* workstation.
NR_{km}:	Number of parts sent back for rework to the previous *k* workstation.
MC_k:	Total manufacturing cost for parts at the *k* workstation.
TIC_{km}:	Total inspection cost at *m* inspection station.
TC:	Total cost of manufacturing parts.

Inspection Stations

In the multistage manufacturing system being considered the model is developed based on assumption that there is limited number of inspection stations available (limited budget) and an inspection policy can be described as the following:

$$V_{km} = \begin{cases} 1 & \text{if an inspection station } m \text{ is assigined} \\ & \text{to inspect parts leaving workstation } k \\ 0 & \text{Otherwise} \end{cases}$$

WORK FLOW ANALYSIS

Flow constraints consider the number of conforming parts departure a workstation or inspection location, and the number of defective parts entering a following workstation or inspection location. The following equation is represented the first stage

$$NG_1 = Q.(1-Z_1) \qquad (1)$$

For all other workstations the equation is defined recursively as follows:

$$NG_k = V_{km}[NG_{k-1}.(1-\alpha_{k-1}) + ND_{k-1}.\beta_{k-1} + NR_{k-1}](1-Z_k).\Sigma V_{km} + (1-\Sigma V_{km}).NG_{k-1} \qquad (2)$$

The number of defective parts produced at first a processing workstation is:

$$NG_1 = Q.Z_1 \qquad (3)$$

For all other stations, is given by:

$$ND_k = V_{km}[NG_{k-1}.\alpha_{k-1} + ND_{k-1}.(1-\beta_{k-1}) + NR_{k-1}]Z_k.\Sigma V_{km} + (1-\Sigma V_{km}).ND_{k-1} \qquad (4)$$

The number of parts classified as defective by inspection station *m* after workstation *k* that can be repaired is given by:

$$NR_k = [NG_k.\alpha_m + ND_k.(1-\beta_m)].V_{km} \qquad (5)$$

The total cost (*TC*) of processing and inspection of *Q* parts in an *n*-stage serial manufacturing system is given by the following equation:

$$TC = \sum_{k=1}^{n}(IFC_k + MC_k + IC_m + EFC_k) \qquad (6)$$

The objective function for the inspection station allocation problem for serial manufacturing workstations producing parts could be formulated as:

$$Minimise\ TC = \Sigma_{k=1}^{n} TC \qquad (k = 1,2,...,n) \qquad (7)$$

The objective function is subjected to the following constraints:

$$\sum_{k=1}^{n} V_{km} \leq NI \quad (8)$$

$$\sum_{m=1}^{r} V_{km} \leq 1 \quad (9)$$

Equations (8) and (9) demonstrate the inspection resource limitations. Equation (8) represents that there are limited inspection stations available for the manufacturing process. Equation (9) represents that none or only one inspection station can be assigned after each workstation.

COMPLETE ENUMERATION METHOD (CEM)

The allocation of quality inspection station problem grows exponentially with the number of workstations. For instance if a relatively small number of 25 processes are required with a maximum of 5 inspection points there will be 6,375,600 assignment combinations in which to install these, as calculated using standard factorials using equation.

$$p(n,k) = \frac{n!}{(n-k)!} \quad (10)$$

Complete enumeration method is optimization method that checks for all possible combinations of inspection station allocation and assignment. If workstation k should be screened by an inspection station m, then X_{km}=1, otherwise X_{km}= 0. For example, inspection plan has **0 0 0 1 1** means that the inspection is performed after the fourth and fifth workstations. The CEM will be used as comparison method. The following steps are described CEM:

Step1 Generate the set of assignment location combinations of inspection plans for a multistage manufacturing system. $C = [(X_{11},..X_{k1},...X_{n1}),...(X_{1m},...X_{km},...X_{nm}),...(X_{1M},...X_{km},...X_{nM})]$. Where m =1,…, M, C is all assignment location combinations, $M=2^n$ and n is the number of workstations. Set m = 0 and k =0. Go to Step 2.

Step2 Set $m=m+1$, k=1 and TC = 0.Selecting $(X_{1m},...X_{km},...X_{nm})$, from C. Go to Step 3.

Step3 According to the order of inspection plans, if inspection plan matches the number of inspection stations required go to Step 4, otherwise this inspection plan will be rejected. Go back to Step 2.

Step4 According to the order of workstations, a workstation whether should be monitored by an inspection station is sequentially checked. If X_{km}=1, then inspection is needed after workstation k then go to Step 5. If k_{km} = 0, go to Step 6.

Step5 Calculating the sum of TIC which includes IC_m, MC_k, EFC_k and IFC_k that result from applying every available inspection station. Assigning a feasible inspection station that causes the lowest sum to monitor workstation k. Go to Step7.

Step6 Calculate the processing cost plus the external failure cost based on the inspection plan selected. Go to Step 7.

Step7 Total cost calculation. Calculate the total cost of the inspection plan (TC) in the manufacturing system. Update total cost TC and Go to Step8.

Step 8 Checking whether all workstations in the manufacturing system have been completely considered. If $k = n$, then go to Step 9. Otherwise, $k = k + 1$ and go to Step 4.

Step 9 Checking whether all assignment location combinations in C have been completely considered. If $m = M$, then go to Step 10. Otherwise, go back to Step 2.

Step 10 Checking whether all assignment location combinations in C have been completely considered. If $m = M$, then go to Step 10. Otherwise, go back to Step 2.

Max-MIN ANT SYSTEM for AIE PROBLEM

Marco Dorigio created an ant algorithm called Ant System. Afterwards many new ant algorithms and their variants were invented, among them the most important and successful was Max-Min ant system (MMAS) proposed by Stützle and Hoos (2000) which is a particular variant of the ACO metaheuristics. Experimental results in (Stutzle, 2000) demonstrate that MMAS achieves a strongly improved performance compared to AS and to other improved versions of *AS* for the *TSP*.

ACO algorithms are inspired by the foraging behaviour of real life ant colonies in which individual ants drop a substance called pheromone on the path while moving from the nest to the food sources and back. They communicate information about food source using pheromone along the ground. When an ant finds a food source it returns to the nest. As ants on short (i.e. objective function) path will return to the nest faster, more pheromone will be deposited on the shorter paths. Moving ants accordingly choose their path with a probability that depend on the amount of pheromone detected and consequently, paths that are more frequently travelled become more attractive and, by means of that self-strengthen behaviour, will be used more often. Further, the pheromone ‘‘evaporates” over time, so that pheromone trails of infrequently travelled become weaker while attractive paths are reinforced. Unlike the real ants that are almost blind creatures, artificial ants can include some heuristic information while assigning inspection station at workstations. For more detailed description of this Phenomenon, the interested reader is referred to Dorigo et al. (1996).

The standard Max-Min Ant System presented by Stützle and Hoos (2000), was derived from the standard ACO and it incorporated three key features to achieve a better performance:

(i) To exploit the best solutions found during iteration or during the run of the algorithm, after each iteration only one single ant adds pheromone. This ant may be the one which found the best solution in the current iteration (*iteration-best* ant) or the one which found the best solution from the beginning of the trial (*global-best* ant).

(ii) To avoid stagnation of the search the range of possible pheromone trails on each solution component is limited to an interval $[\tau_{min}, \tau_{max}]$. Where τ_{min} is the minimal pheromone trail and τ_{max} is the maximal pheromone trail on any arc.

(iii) Additionally, initialize the pheromone trails to τ_{max}, achieving in this way a higher exploration of solutions at the start of the algorithm.

Generally any *MMAS* algorithm must specify the following elements:

(1) Construction of solutions.
(2) Heuristic information.
(3) Pheromone updating.
(4) Pheromone trail limits
(5) Selection probability.
(6) Pheromone trail initialization.

In the following subsections these elements are described.

Construction of Solutions

A feasible and complete solution of the formulated inspection allocation problem is considered as static connected graph *V= (i, j),* where *i* is the number of workstations and *j* is the set of arcs connected them. Each arc is weighted by a pair of numbers $\{\tau_{ij}, \eta_{ij}\}$, where τ_{ij} is the trail level and η_{ij} is the visibility, computed initially according to a desirability measure derived from a greedy problem-specific heuristic as will be shown in next section. Assignment of inspection station to workstation *i* is called a move. A move V^k takes ant *k* from workstation 1to workstation 2. In the proposed algorithm, it is assumed that each ant initially assigns limited number of inspection stations to corresponding workstations in a serial multistage manufacturing system until a complete solution is obtained.

Heuristic Information

Unlike the real ants that are almost blind creatures, artificial ants can include some heuristic information while assigning inspection station at different workstations. The heuristic information pertaining to move *V (i, j)* is denoted by η_{ij}. This information indicates the desirability of assigning *i* inspection station to workstation *j* and is calculated using a heuristic approach. In multistage manufacturing systems, the higher manufacturing cost MC_k is the more important to allocate an inspection station for screening previous workstation *k-1*. The unnecessary MC_k of manufacturing nonconforming product by workstation *k-1* can then be avoided. Consequently, the higher the defective rate (z_k) of workstation *k*, the more important is to assign an inspection station for screening workstation *k*. Based on the concerns of manufacturing cost and defective rate, respectively, the weights to locate an inspection station after workstation *k* are determined as $W_{c,k}$ and $W_{z,k}$. While there are *n* workstations in the serial multistage manufacturing system, either $W_{c,k}$ or $W_{z,k}$ are determined to be 1; 2; . . ., *n*. The higher MC_k, the larger priority ($W_{p-1,k}$). An earlier workstation will have a higher priority $W_{p-1,k}$ if two or more workstations have equal manufacturing cost.
Concern should be paid that, $W_{n,k}$ is always set at the largest *n*. It added guarantees the product quality sold to customers.

Otherwise, unnecessary costs will be incurred due to costs of repair, replacement and quality loss. The higher p_k, the larger $W_{z,k}$ is. An earlier workstation will have a higher $W_{z,k}$, if two or more workstations have the same defective rate. The total weight of positioning an inspection station for each workstation, $W_{a,k}= W_{c,k}+ W_{z,k}$ can then be determined by considering simultaneously the manufacturing cost and defective rate. The larger $W_{a,k}$ is for a workstation, the higher the priority for positioning an inspection station to it. An earlier workstation will have a higher priority, if two or more workstations have no difference in $W_{a,k}$. Based on this assumption described above by Shiau (2002) and Lee and Unnikrishnan (1998), the following formula is proposed to calculate the desirability of move *V = (i, j)*:

$$\eta_{ij} = \frac{1}{W_{a,k}} \tag{11}$$

The smaller value of $W_{a,k}$ the more attractive to allocate inspection station. Hence, visibility η_{ij} is a static greedy heuristic value and can be pre-computed before applying the algorithm.

Pheromone Updating Rule

In *MMAS* only one single ant is used to update the pheromone trails after each iteration. Consequently, the modified pheromone trail update rule is given by the following equation:

$$\tau_{ijk}^{new} = \rho.\tau_{ijk}^{old} + \Delta\tau_{ij}^{best} \tag{12}$$

Where $\Delta\tau_{ij}^{best} = \dfrac{1}{f(s^{best})}$ and $f(s^{best})$ denotes the solution cost of either the iteration- best (s^{ib}) or the global-best solution (s^{gb}), and ρ is the pheromone evaporation rate in order to avoid unlimited accumulation of the trail, the value of ρ should be $(0 < \rho < 1)$.

MMAS focuses on the use of the iteration-best solutions. The use of only one solution, either s^{ib} or s^{gb}, for the pheromone update is the most important means of search exploitation in *MMAS*. By this choice, solution elements which frequently occur in the best found solutions get a large reinforcement.

Pheromone Trail Limits

Stutzle and Hoos (2000) proposed the provision of dynamically evolving bounds on pheromone trail intensities such that the pheromone intensity on all paths is always within a specified range. As a result, all paths will have a probability of being selected and, thus, a wider exploration of the search space is encouraged. MMAS uses upper and lower bounds to ensure that pheromone intensities lie within a given range $\tau_{\min} < \tau_{ij}(t) < \tau_{\max}$, hence the name. By limiting the range of values for the pheromone trail, the probability distribution given by Equation 24 can be indirectly influenced. The upper and lower pheromone bounds are calculated as following:

$$\tau_{\max} = \frac{1}{1-\rho}\,\frac{1}{f(s_{opt})} \tag{13}$$

Where $\tau_{\max}$ is the maximal pheromone trail and s_{opt} is the optimal solution value for the problem.

$$\tau_{\min} = \frac{\tau_{\max} \cdot (1 - p^{dec})}{avg. p^{dec}}, \quad p^{dec} = (p^{best})^{\frac{1}{n}} \qquad (14)$$

Where $\tau_{\min}$ represents the lower limit for the pheromone trail strength; p^{dec} is the probability that an ant constructs each component of the best solution again; p^{best} is the probability that the best solution is constructed again and *avg* is the average number of options available at decision points of the problem.

Selection Probability

In all the implementations of ant algorithms, an ant chooses a move to go from current state, to the next adjacent state, based on a rationale combination of two factors, namely the desirability of that move and the quantity of pheromone on the edge which is to be traversed. Ants prefer workstations connected by arcs with a high pheromone trail and in *AS* an ant *k* currently located at workstation *i* chooses to go to workstation *j* with a probability:

$$p(i,j) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\Sigma [\tau_{ij}]^{\alpha} \cdot [\eta_{ij}]^{\beta}} & if \;\; j \in N_i^k \\ 0 & otherwise \end{cases} \qquad (15)$$

Where α and β are the parameters that determine the relative dependence on pheromone trail intensity and local information and N_i^k is the feasible neighbourhood of ant *k*, that is, the set of workstations which ant *k* has not visited yet.

Pheromone Trail Initialization

In the max-min Ant system trail strength is initialised as the maximum possible trail strength for all edges. Then the trail strength will be reduced due to evaporation. After the first iteration of MMAS, the trails will be forced to take values within the specified limits. As only the best ant is allowed to update its tour, only the trails of arcs participating in the best tours are allowed to increase its intensities or maintain them at the upper trial level. Hence, arcs that do not receive any reinforcement will continuously lower their trail strength and be selected more rarely by the ants. This type of trail initialization is chosen to increase the exploration of solutions during the first iterations of the algorithm.

LOCAL SEARCH

Recent research in ant colony optimization has shown that for applications to combinatorial optimization problems, best performance is obtained if the ants are enhanced by additional capabilities. The reason for adding local search algorithms to MMAS is to enhance performance and to yield high quality solution, such that near optimal solution can be found. The most widely known iterative improvement algorithms for the assignment problems is certainly 2-opt method Stutzle (1998). When all inspection stations have been allocated to workstations, that is a complete solution has been constructed by an ant. This solution is improved by using local search method. In the 2-Opt method, 2 inspection stations are swapped between workstations as shown in Fig. 3. Whenever an improvement of the objective function is detected, the new solution replaces the old one, and the process continues until no further improvement. The improved solutions are then used to update the pheromone trail. Note that when applying ACO to tackle a problem without considering local search, the best results are obtained by increasing the number of ants proportionally to the instance size. Yet, this choice need not be the best when additionally applying local search Stutzle (1998).

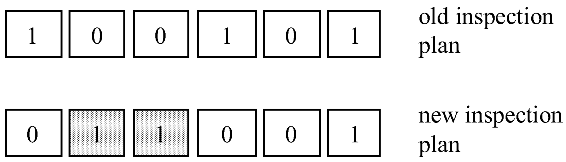


Figure 2-Opt exchange neighbourhood

MMAS ALGORITHM

The proposed algorithm is presented as follows:

Initialize
Begin
    Set parameters and initial pheromone trails
    Calculate the heuristic matrix information
    While *(termination condition not met)* ***do***
        For *k = 1 to A (number of ants)*
            For *i =1 to number of inspection stations*
            End for
        End for
            Calculate the objective function
            Apply local search to improve
            Update pheromone trails using
            Checking upper and lower pheromone trials
    End While
    Return *the best solution found*
End

EXPERIMENTAL RESULTS

In this section, the results obtained by our experimental tests will be described and analysed. The algorithms have been coded in Matlab and executed on a 2.8 GHz CPU and 2 GB RAM so as to remain comparable in terms of the computational effort of CEM and MMAS algorithm.

Experimental Framework

A multistage manufacturing system model with different workstations (15-19) arranged in a serial manner was used to allocate 5 inspection stations. The reason is to show the effective of the MMAS algorithm for tackling the AIE problem when number of workstation increases. The batch size used in the experiment was 100. The parameters of the developed model were randomly generated using a uniform random number generator to evaluate the processing time efficiency. The performance of the proposed MMAS algorithm is evaluated through CEM which produces optimal solution. The effective of employing local search methods on the MMAS algorithm is tested by using pure *MMAS* algorithm without local search which is denoted as

(*MMAS-wls*) and *MMAS* without local search which is denoted as (*MMAS-wols*).

Experiment Design

Experiments are critical in finding out the relative performance of the heuristic method as compared to the complete enumeration method. Two performance measures are considered in the experiments:

(1) Total system cost (*TC*). The total system cost is the sum of the total cost of processing and inspecting parts produced in the system.
(2) Processing time (*PT*). This is the *CPU* time taken to execute the computer programs for problem solving in computer system. An efficient *MMAS* algorithm should provide significant saving in processing time over the complete enumeration method.

MMAS Parameters Settings

In this section the influence of the parameter settings on Max-Min Ant system is studied. The amount of pheromone an ant deposits on an arc is always $Q = 1$, because the particular value of Q has not a significant influence on the final performance Stutzle (1998). The parameters of the MMAS algorithm are set to the following values: $\alpha = 1$, $\beta = 2$ and $\rho = 0.02$. This gives relatively more weight to the pheromone trail intensity than the problem-specific heuristic and greater emphasis on exploitation rather than exploration. The MMAS is not very sensitive to changes in these values, and tested well for quite a range of them. To avoid faster convergence of the algorithm iteration-best ant is allowed to update pheromone trails.

Results

Table 1 shows the results obtained by simulation experiments to evaluate the allocation problem through processing time efficiency and cost deviation. It was found that MMAS-wls algorithm has better saving processing time 99 % than CEM, particularly when the number of workstations increased significance. Table 1 also shows the expected total cost of the developed algorithms. The results show that MMAS-wls can produce solution near to the optimal, the maximum deviation was only 0.05. Table 2 shows inspection plans produced by CEM and $\text{MMAS-}_{\text{wls}}$. Clearly they are very similar because the cost deviation was very small the only difference is in the position of the second inspection station. For example in the case of 15 workstations and CEM the second inspection station was positioned after the seventh workstation whereas in the $\text{MMAS-}_{\text{wls}}$ it was positioned after the tenth workstation.

Table 3 shows the results obtained from experimental tests when $\text{MMAS-}_{\text{wols}}$ is used. It was found that the deviation of the total cost from the optimal solution in the case of MMAS-wols was slightly higher than $\text{MMAS-}_{\text{wls}}$. The maximum deviation was 0.1 and the minimum deviation was 0.07. It is clearly that adding a local search algorithm to MMAS enhanced the performance of and produced solution near to the optimal. Table 4 shows the results obtained by simulation experiments when number of Ants increased proportionally to the problem size. Also here the evaluation is through processing time efficiency and total cost. It was found that MMAS-wls algorithm has better saving processing time 98 % than CEM, particularly when the number of workstations increased significance.

Conclusions

Computational experiments have been carried out to validate the effectiveness of the algorithm in various problem sizes. The performance of the proposed MMAS algorithm is compared with the developed optimisation CEM. MMAS can produce solution near to optimal where the maximum deviation of the total cost of the algorithm from the optimal solution of approximately 6%.The reduction time of MMAS in comparison with CEM for the conducted case study was 99%. The reduction time is obviously when the number of workstations increased significance. The computational results of MMAS confirm that very high quality solutions on AIE can be obtained. To shorten the run-time for finding high quality solutions two possibilities are proposed.

The first is based on a gradual increase of the number of ants that apply local search. The second approach is used a constant number of ants and every ant applies local search to its tour.

The effectiveness of the algorithm lies in its considerably shorter execution time and its handiness and robustness. The present paper adds a new dimension to solve AIE problems by solving them through the ant colony optimization approach which emulates the real-life behaviour of ants.

The ACO approach has shown promising results that provide scope for further study.

Performance of the MMAS in comparison to CEM (15Ants are used)

Number of workstations	Processing time (minutes)		Time efficiency	Total cost		Cost deviation
	EM	MMAS-*wls*	$(1-\frac{MMAS-wls}{EM})\%$	CEM	MMAS-*wls*	$(1-\frac{MMAS-wls}{CEM})\%$
15	0.2	0.18	10	58622	61826	0.05
16	0.8	0.6	25	64692	68411	0.05
17	3	0.65	79	65933	68376	0.03
18	16	0.8	95	60725	62651	0.03
19	85	0.85	99	65951	67888	0.03

Table 2. Inspection plans of the *MMAS* in comparison to CEM

Number of workstations	No. of feasible solutions	Inspection plans	
		CEM	MMAS-$_{wls}$
15	360,360	010000100001011	010000000101011
16	524,160	0100000001000111	0100000000010111
17	742,560	01000000000010111	01000000001000111
18	1,028,160	010000000001100011	010000000000001111
19	1,395,360	0100000000000001111	0100000000000100111

Table 4. Performance of the MMAS in comparison to CEM

Number of workstations	No. of Ants	Processing time (minutes)	Time efficiency	Total cost	Cost deviation
		MMAS$_{-wls}$	$(1-\frac{MMAS\_wls}{CEM})\%$	MMAS$_{-wls}$	$(1-\frac{MMAS\_wls}{CEM})\%$
15	10	0.18	10	61188	0.04
16	15	0.5	62	64934	0.003
17	20	0.7	77	66505	0.008
18	25	0.9	94	64639	0.06
19	30	1.2	98	68273	0.03

References

Ballou, D.P. and Pazer, H. L. (1985). Process improvement versus enhanced inspection in optimised systems. *International Journal of Production Research*, 23, 1233-1245.

Bell, J.E. and McMullen, P. R. (2004). Ant colony optimization techniques for the vehicle routing problem. *Advanced Engineering Informatics,* 18, 41-48.

Dorigo, M. and Gambardella, L. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transaction on Evolutionary Computation,* 1, 53– 66.

Dorigo, M., Maniezzo, V. and Colorni, A. (1996). The ant systems: optimization by a colony of cooperative agents. *IEEE Transaction on systems, Man, Machine and Cybernetics, Part* B, 26(1), 29-41.

Dorigo, M. Maniezzo, V. and Colorni, A. (1991). Positive feedback as a search strategy. *Technical report*. Dipartimento di Elettronica, Politecnico di Milano, Milan.

Dorigo, M. and Stutzle, T. (2004). *Ant colony optimisation.* MIT Press, Cambridge, MA, USA, 2004.

Gambardella, L. Dorigo, M. (1995). Ant-Q: A reinforcement learning approach to the traveling salesman problem, in: Proceedings of the 11th International Conference on Machine Learning, Morgan Kaufmann, San Francisco, CA, pp. 252–260.

Hurst, E.G. (1973). Imperfect inspection in a multistage production process. *Management Science*, 20, 378-384.

Kumar, R. Tiwari, M.K. and Shankar, R. (2003). Scheduling of flexible manufacturing systems: Ant colony optimizations approach. Proceeding of the institution of mechanical engineers, Part B: J. Engineering Manufacture, 217(10), 1443-1453.

Liang, Y.C. and Smith, A.E. (2004). An ant colony optimization algorithm for the redundancy allocation problem. *IEEE Transaction on Reliability,* 53(3), 417-423.

Lindsay, G.F. and A.B. Bishop. 1964. ‘‘Allocation of Screening Inspection effort a dynamic programming approach’’10, No.2 (Jan), 342-352.

Lee, J. and Chen, F.F. (1996). Inspection sequencing and part scheduling for flexible manufacturing systems. *European Journal of Operational Research*, 95, 344-355.

Lee, J. and Unnikrishnan, S. (1998). Planning quality inspection operations in multi-stage manufacturing systems with inspection errors. *International Journal of Production Research*, 36, (1), 141-155.

Montgomery, D.C. (1997). *Introduction to statistical quality control*. 3rd ed. New York: John Wiley &Sons.

BIOGRAPHY

ALI SHETWAN received the B.Sc. from Alfateh University, Tripoli, Libya in 1986 in Industrial Engineering and M.Sc. degrees from the Higher Institute of Industry Misurata, Libya in Production and Quality Engineering in 1999. Since 2008 he is PhD student in school of engineering at the group of Prof. Vitanov, at Durham University. He is mainly interested in quality and production.

Valentin VITANOV is a Professor of Design Manufacture and Management in the School of Engineering and Computer Sciences at the University of Durham UK. He received his PhD in Industrial Automation and Robotics from the St. Petersburg Electrotechnical University (Russia). He is a Fellow of the Institute of Engineering and Technology, UK. His research interests include applied statistics and robust design, manufacturing systems simulation and optimisation, operational research and autonomous robotics.

BUILDING A STOCHASTIC INVENTORY MODEL FOR DETERIORATING ITEMS WITH SERVICE LEVEL CONSTRAINT USING SIMULATION MODELING

Hamidreza Eskandari
Mohammadmahdi Alizadeh
Tarbiat Modares University
School of Industrial Engineering
Tehran, 14117-13116, IRAN
eskandari@modares.ac.ir
mm.alizadeh@modares.ac.ir

Seyed M. Sajadifar
University of Science and Culture
Department of Industrial Engineering
Tehran, 13145-871, IRAN
sajadifar@usc.ac.ir

Afrouz Ghaemi
Brunel University
School of Engineering and Design
London, UB8 3PH, UK
im09aag1@brunel.ac.uk

KEYWORDS
Inventory, Deteriorating Items, Service Level, Simulation.

ABSTRACT

We consider an inventory system with one warehouse which faces independent Poisson demand. Inventory items are deteriorating at a constant rate and lead times are stochastic. Shortage is allowed and all the unsatisfied demands are backlogged. In addition replenishment is one for one. Our objective is to minimize long-run expected costs. To guarantee a specific service level, a constraint has been considered. We develop the analytical model for deterministic lead time. Since the stochastic lead time model is complicated especially when lead time has non-exponential distribution and it is difficult to prove convexity of the model, we have established simulation model which has no limitation for lead time or any other parameters. We validate our simulation model by comparing simulation and analytical models' results when lead time is deterministic. There is not significant differences between simulation model and mathematical model results. Furthermore, we try to find near optimal solutions for a number of examples with stochastic lead time by applying optimization module of the applied software.

INTRODUCTION

Most of inventory models assume that products can be stored for an indefinite time to satisfy future demands. However, there are many products for which this assumption is not true. Items which obsolete or decay are the examples of this kind of products. According to Goyal and Giri (2001) "Deterioration refers to the damage, spoilage, dryness, vaporization, etc of products". These two researchers classified deteriorating items into two groups. The first group is commodities like foodstuffs, green vegetables and human blood which have a maximum usable lifetime and known as perishable products. The second group contains products such as alcohol, gasoline, radioactive substances, etc. which have no shelf-life and known as decaying products. According to Dye et al. (2007a), there are many products such as medicine, volatile liquids, blood bank and many others which face different kinds of deterioration such as vaporization, damage, spoilage, dryness and so on during their storage period. In such cases the deterioration of products cannot be neglected for obtaining optimal inventory policy. Raafat (1991) states that "the analysis of decaying inventory problems began with Ghare and Schrader (1963)". During these years a lot of researches have been conducted in this area which consider deterministic demand. Since 1963 a large number of studies have been carried out in the field of deteriorating inventory models with deterministic demand especially constant demand rate. Although deteriorating inventory models with constant demand rate is almost saturated, in recent years some researches were conducted in this area in which new assumptions were applied. Li et al. (2007) and Hsieh et al. (2008) are two examples of those new studies. Li et al. (2007) considered postponement strategy in his EOQ based model. Hsieh et al. (2008) established a model for deteriorating items with constant demand rate and two warehouses (own and rented) to consider capacity limitation. In addition, the number of studies on decaying products inventory models with other types of deterministic demand is remarkable. In the field of deteriorating inventory models with time-varying demand considerable researches are conducted. Papachristos and Skouri (2000), Teng et al. (2002), Goyal and Giri (2003), Sana et al. (2004) and Chern et al. (2008) are such studies. Furthermore, there are noticeable studies with price-dependent demand Dye et al. (2007a) and Dye et al. (2007b) are such models. Moreover, a large number of papers have been published recently for deteriorating items inventory model with stock-dependent demand such as Teng and Chang (2005), Wu et al. (2006) and Chang et al. (2009).

Although in real world cases demand is usually random, there are a small number of studies on decaying products inventory models with stochastic demand. Lodree and Uzochukwu (2008) states that stochastic models of deteriorating items received less attention after 2001. Lodree and Uzochukwu (2008) developed a production planning model for a deteriorating item with stochastic demand, positive lead time and consumer choice. "The complexity of decay models having random demand depends strongly on the lead time assumptions. When lead times are zero, determining optimal order policies is relatively straightforward" (Nahmias, 1982). As a result of this fact, most of studies on deteriorating items inventory models considered zero lead times. "The assumption of positive lead times further increases the complexity of the analysis of these models and hence there are only a limited number of papers dealing with positive lead times" (Sivakumar, 2009). Since analyzing continuous decaying products inventory models with stochastic demand and positive or stochastic lead time is considerably difficult, there are a few papers in this area. It is obvious that exponential decaying problems with random demand become more complicated when lead time is assumed to be stochastic too. There are some stochastic inventory models for deteriorating items in which positive or stochastic lead time are considered. Kalpakam and Sapna (1994) in their paper analyze an (s, S) exponential decaying system with poission demand and exponentially distributed lead time. They assume demands during stock-out periods are lost. Sivakumar (2009) develops an inventory model for exponential decaying items. In that model a finite number of homogenous sources of demand are considered, and (s,S) inventory policy and exponentially distributed lead time are assumed. In addition, it was considered that demands occurring during stock-out periods enter into an orbit. The orbiting demands send out signal according to exponential distribution to compete for their demand.

To the best of our knowledge, there is no research in the field of deteriorating items inventory models with stochastic demand and lead time which considers service level constraint. Analyzing models with stochastic demand and lead time is necessary for real world problems. In this paper an inventory model for continuous decaying products with stochastic demand and lead time is developed. We also consider a constraint to guarantee a specific service level. Since stochastic lead times and demands make inventory models of deteriorating items complex, mathematical analysis of this kind of models is extremely difficult. Therefore, for solving this problem we apply simulation modeling. Before using the simulation model we should be sure that it is valid. Thus, we build the mathematical model of the system with deterministic lead time. We show the simulation model's solutions are near to mathematical model's solutions when lead time is deterministic as a validation for the simulation model. After determining model's validation, we start finding near optimal solution for stochastic lead times.

ASSUMPTIONS AND NOTATIONS

In this model a warehouse is assumed where customers enter the system on the basis of Poisson distribution. Each customer orders just one unit of the product. Since the items deteriorate according to exponential distribution, each time a customer orders one unit of the product, the warehouse orders an outside supplier (1+α) unit. In other words, since products are deteriorating, the warehouse orders α unit more than customers' demands to consider deterioration between ordering time and times of delivery. The inventory system is demonstrated in Figure 1:

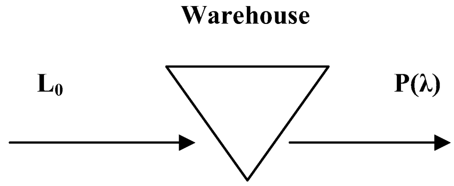


Figure 1: The Inventory System

The analytical model of the problem is developed on the basis of the following assumptions:

- Customers are served on the basis of first come, first served rule.
- Replenishment is one for one.
- Holding cost is calculated on the basis of amount of the product which is delivered to the warehouse.
- Shortage is allowed and all the unsatisfied demands are backlogged.
- A constant fraction of on hand inventory deteriorates per unit time.

We introduce the following notations for the model's parameters:

- L_0: lead time
- λ: demand intensity at the warehouse
- h_0: holding cost per unit per time unit at the warehouse
- β: shortage cost per unit per time unit at the warehouse
- γ: deterioration cost per unit
- θ: deterioration rate which is constant and $0<\theta<1$
- H: expected holding cost to fill one unit of demand
- Π: expected shortage cost to fill one unit of demand
- D: expected deterioration cost to fill one unit of demand
- S_0: inventory position at the warehouse
- α: extra amount of the product ordered by the warehouse
- B: unit purchase cost for the warehouse

MATHEMATICAL MODEL WITH DETERMINISTIC LEAD TIME

When customers enter the warehouse according to Poisson distribution with a rate λ, distribution of the time elapsed between the placement of an order and occurrence of its assigned demand will have Erlang distribution with parameters (λ, S_0). We show the density function of the Erlang distribution by $g^{S0}(.)$:

$$g^{S_0}(t) = \frac{\lambda^{S_0} t^{S_0-1} e^{-\lambda t}}{(S_0-1)!} \qquad (1)$$

We show the corresponding cumulative distribution function by $G^{S0}(t, \lambda)$:

$$G^{S_0}(t,\lambda) = \sum_{k=S_0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad (2)$$

When a demand occurs at the warehouse, (1+α) new unit is immediately ordered from the warehouse to the outside supplier. We designate this time as time zero. Note that the warehouse orders more than one unit for deterioration of the items during the lead time. If customers order while the warehouse is out of stock, the demand is satisfied with delay. As soon as units are again available at the warehouse, customers are served according to first come, first served rule. Furthermore, when a customer orders during stock out period, in fact, the related unit is virtually assigned to the customer. The following lemma which is similar to what Axsäter (1990) states, is necessary for understanding the modeling approach:

Lemma. *Any unit ordered by a specific customer is used to fill the S_0th demand following this order, hereafter, referred to as its demand.*

An order placed by a specific customer arrives after L_0 time units. If the order arrives before its assigned demand, it is considered as inventory stocks and holding cost is incurred. Note that (1+α) units is originally ordered and this amount decrease during the period between ordering and delivering times. When lead time is deterministic $H(S_0)$ is given by

$$H(S_0) = \int_{L_0}^{\infty} h_0 g^{S_0}(s)(s - L_0)(1+\alpha)e^{-\theta L_0} \quad (3)$$

Let us simplify expression (3) for $H(S_0)$ since

$$\int_0^t g^{S_0}(u)u du = \int_0^t g^{S_0+1}(u)\frac{S_0}{\lambda}du = \frac{S_0}{\lambda}G^{S_0+1}(t,\lambda) \quad (4)$$

We obtain

$$H(S_0) = (1+\alpha)h_0 e^{-\theta L_0}\left[\frac{S_0}{\lambda}(1-G^{S_0+1}(L_0,\lambda)) - L_0(1-G^{S_0}(L_0,\lambda))\right] \quad (5)$$

Note that $(1+\alpha)e^{-\theta L_0}$ is the amount of warehouse order (1+α) which remains after the lead time has elapsed.

When the order arrives after its assigned demand, the associated customer's order is backlogged and shortage cost is incurred. In other words for $s \le L_0$ shortage is incurred. Using the expression (4) we calculate and simplify the expected shortage cost $\Pi(S_0)$ as

$$\Pi(S_0) = \beta\int_0^{L_0}(L_0 - s)g^{S_0}(s)ds = \beta\left[L_0 G^{S_0}(L_0,\lambda) - \frac{S_0}{\lambda}G^{S_0+1}(L_0,\lambda)\right] \quad (6)$$

The expected deterioration cost is calculated as

$$\begin{aligned} D(S_0) &= \gamma\int_0^{L_0}(1+\alpha)(1-e^{-\theta L_0})g^{S_0}(s)ds + \gamma\int_{L_0}^{\infty}(1+\alpha)(1-e^{-\theta s})g^{S_0}(s)ds = \\ &\gamma(1+\alpha)\left[1 - e^{-\theta L_0}G^{S_0}(L_0,\lambda) - (\frac{\lambda}{\lambda+\theta})^{S_0}(1-G^{S_0}(L_0,(\lambda+\theta)))\right] \end{aligned} \quad (7)$$

Note that in expression (7) the first integral shows expected deterioration cost when the customer arrives before his assigned demand. The second integral explains the expected deterioration cost when the customer arrives after his assigned demand.

The long-run total expected cost is the sum of expected holding cost, shortage cost, deterioration cost and purchase cost of the products. Total expected cost is given by

$$C(S_0) = \lambda\left[H(S_0) + \Pi(S_0) + D(S_0) + B(1+\alpha)\right] \quad (8)$$

MATHEMATICAL MODEL WITH STOCHASTIC LEAD TIME

When lead time is stochastic with a specific density function, which is shown by *w(x)*, holding cost, shortage cost and deterioration cost incurred to fill a customer's demand are given by (9), (10) and (11) respectively.

$$\begin{aligned} H(S_0) &= h_0\int_X w(x)\int_x^{\infty} g^{S_0}(s)(s-x)(1+\alpha)e^{-\theta x}dsdx = \\ &(1+\alpha)h_0\int_X w(x)e^{-\theta x}\left[\frac{S_0}{\lambda}(1-G^{S_0+1}(x,\lambda)) - x(1-G^{S_0}(L_0,\lambda))\right]dx \end{aligned} \quad (9)$$

It is obvious that holding cost is incurred when a specific customer's assigned demand is delivered to the warehouse before that customer enters the system.

$$\begin{aligned} \Pi(S_0) &= \beta\int_X w(x)\int_0^x (x-s)g^{S_0}(s)dsdx = \\ &\int_X w(x)\beta\left[L_0 G^{S_0}(x,\lambda) - \frac{S_0}{\lambda}G^{S_0+1}(x,\lambda)\right]dx \end{aligned} \quad (10)$$

Shortage cost is incurred when a specific customer enters the system before his assigned demand.

$$\begin{aligned} D(S_0) &= \gamma(1+\alpha)(\int_X w(x)\int_0^x (1-e^{-\theta x})g^{S_0}(s)dsdx \\ &+ \int_X w(x)\int_x^{\infty}(1-e^{-\theta s})g^{S_0}(s)dsdx) \\ &= \gamma\int_X w(x)(1+\alpha)(1-e^{-\theta x})G^{S_0}(x,\lambda)dx \\ &+ \gamma\int_X w(x)(1+\alpha)\left[(1-G^{S_0}(x,\lambda)) - (\frac{\lambda}{\lambda+\theta})^{S_0}(1-G^{S_0}(x,(\lambda+\theta)))\right]dx \end{aligned} \quad (11)$$

Note that the first integral shows deterioration cost when a specific customer enters the system before his assigned demand and the second integral calculates deterioration cost when this customer enters the system after his assigned demand. Again total expected cost is given by (8).

SIMULATION MODEL

As it is illustrated in previous sections, when lead time is deterministic we can obtain different elements of the total cost easily. But when lead time is stochastic, these three elements become more difficult to obtain. In addition, in this situation it is extremely difficult to show whether the total cost function is convex and therefore it will be hard to find the optimal solution. Thus, we build a simulation model of this problem and we show how easy we can find solutions in different situations without any limitation on the lead time distribution or any other parameters. Furthermore, we can optimize the problem by using optimizer module of the applied software. We built the simulation model according to the analytical model's approach which was explained in previous sections.

We use deterministic lead time to validate our simulation model. For two different experiments we illustrate there are not significant differences between our analytical model and simulation model's variable cost containing holding, shortage and deterioration costs. Note that the holding cost ($H(S_0)$), shortage cost ($\Pi(S_0)$) and deterioration cost ($D(S_0)$) in our analytical model are obtained according to expressions (5), (6) and (7) respectively. In addition, we run our simulation model for 20 replications where each replication takes 10,000 hours. Table 1 shows the output results of two models, simulation and analytical, for the two different configurations where the model parameters are randomly chosen. Note that the confidence level is assumed to be 95%.

Table 1 indicates that there are very small differences between results obtained from simulation model with those of the analytical model. As it can be seen in the table, because of running model for 20 long replications (10,000 hours) half width of the simulation model's costs in most of the cases are very small. In other words, the simulation estimators of the system's costs are almost point estimators.

Table 1: Validation Experiments

Experiment 1 Parameters	$h_0=3$ $\beta=2$ $\gamma=15$ $\alpha=0.1$ $\theta=0.01$ $L_0=27$ $S_0=10$ $\lambda=3$		
Costs	H	Π	D
Simulation Results	0±0	47.3348±0.01	3.9042±0
Analytical Results	0	47.3333	3.9042
Error	0%	0.0032%	0%
Experiment 2 Parameters	$h_0=2$ $\beta=5$ $\gamma=8$ $\alpha=0.1$ $\theta=0.01$ $L_0=12$ $S_0=17$ $\lambda=0.25$		
Costs	H	Π	D
Simulation Results	109.33±0.41	0±0	4.2837±0.01
Analytical Results	109.2686	8.4543e-008	4.2823
Error	0.0562%	N/A	0.0327%

We build a simulation model for the problem and we show that the simulation model is valid. Now we can run our model with stochastic lead time. As mentioned before in stochastic demand inventory models of deteriorating items, stochastic lead time makes these models complicated. We also have shown this fact in expression (9), (10) and (11). Furthermore, it is extremely difficult to prove convexity of total cost function. In this inventory model, we have three stochastic parameters: the life time of the inventory items, lead time and the Poisson demand. It will be difficult to analyze this problem with analytical models. On the other hand, we can create the simulation model and choose any distribution for lead time or for other parameters without any limitation. We found in the inventory system literature that the lead time is typically normally or exponentially distributed. In this section we will take advantage of the simulation model for two examples with two different lead time distributions. We will use Arena optimizer module, OptQuest, and try to find near optimal solutions for these examples. We use the optimizer module of the applied software for two different problems with stochastic lead time which is obtained from the literature. In this study we have applied Arena software and its optimizer, OptQuest. Since customers' entrance are random and products decrease thorough the time, there is no guarantee to deliver the amount of product which is ordered by customers (one unit of the product). Hence, a constraint is defined to control percentage of the customers to whom less than one unit of the product is delivered. This constraint is called the service level constraint. Our objective is to minimize the total cost. The problem can be presented as below:

$$\textit{Minimize} \quad C(S_0,\alpha) = \lambda\left[H(S_0,\alpha) + \Pi(S_0) + D(S_0,\alpha) + B(1+\alpha)\right]$$

s.t.

$$P((1+\alpha)e^{-\theta s} < 1) \le 0.05$$
$$\alpha \ge 0$$
$$S_0 \in N$$

Note that in this problem we have two decision variables which are α and S_0. α is assumed to be a continuous variable and S_0 assumed as a discrete and non-negative variable. Notice that defining the service level constraint mathematically is not easy since we should obtain probability distribution function of $(1+\alpha)e^{-\theta S}$ in which s has Erlang distribution with parameters (λ,S_0). While this constraint can be easily defined in Arena optimizer, OptQuest. In this section we will take advantage of the simulation model for two examples with two different lead time distributions. In this paper we take advantage of simulation benefits to optimize abovementioned problem. Two different problems with two different lead times are optimized by OptQuest. OptQuest combines the metaheuristics of Tabu search, Scatter search and Neural Networks into a single, composite search algorithm to provide maximum efficiency in identifying new scenarios (April et al. 2003). In both the examples we have used the parameters which have been used by Dye et al. (2007c). These parameters are provided in Table 2.

Table 2: Common Parameters of the Numerical Examples

Shortage cost per unit per time unit	$\beta=2$
Deterioration cost per unit	$\gamma=10$
Holding cost per unit per time unit	$h_0=0.2$
Deterioration rate per unit per day	$\theta=0.02$
unit purchase cost for the warehouse	B=10

In addition, we use customer arrival rate which was applied by Kalpakam and Sapna (1994). In their research customer arrival rate (λ) was considered to be 5 (λ=5). In the literature of inventory models we have found normal and exponential distributions for lead time. In the first example we assume normal distribution for lead time (w(x)) with parameters μ=0.03 day and σ=0.06 day and exponential distribution with parameter b=3.33 days for the second one which were used by Maiti et al. (2009). Note that for each example parameters of the lead time distribution are given in the associated table. We take advantage of Arena optimizer, OptQuest, to minimize the total cost. Our goal is to find near optimal solution for decision variables, extra amount of ordering (α) and the inventory position at the warehouse (S_0) which cause near optimal total cost in a way that the service level constraint is met. The results of optimizations are illustrated in Table 3 and 4 for the example number 1 and 2 respectively. These tables show the trend in which the best found solutions are obtained during specified replication number by OptQuest software. Graphical view of each table is illustrated by the graph below these tables (Figures 2 and 3). These figures also show how the objective function improves during simulation runs.

Table 3: Optimization Results Obtained from OptQuest for Normal Lead Time (Example 1)

Lead time: Normal (μ=0.03, σ=0.06)(day)		No. of replication: 100	
Parameters h_0=0.2 β=2 γ=10 θ=0.02 B=10 λ=5			
Simulation Run	S_0	α	Total Cost
1	20	0.300000	75.129565
22	13	0.135168	62.530785
33	13	0.111632	61.233990
37	3	0077420	55.390465
55	2	0.049193	53.622100
61	2	0.048399	53.581775
62	2	0.048393	53.581460
95	2	0.024196	52.353035

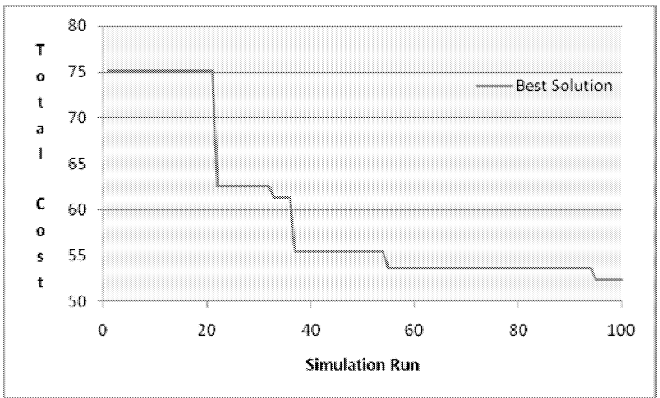


Figure 2: Optimization Results Obtained from OptQuest for Example 1

Table 4: Optimization Results Obtained from OptQuest for Exponential Lead Time (Example 2)

Lead time: Exponential (b=3.33)(day)		No. of replication: 100	
Parameters h_0=0.2 β=2 γ=10 θ=0.02 B=10 λ=5			
Simulation Run	S_0	α	Total Cost
1	20	0.300000	107.433605
24	1	0.300000	102.443665
61	1	0.258880	100.251510
62	1	0.238319	99.155265
63	1	0.228039	98.607145
64	1	0.222899	98.333250
65	1	0.220329	98.196135

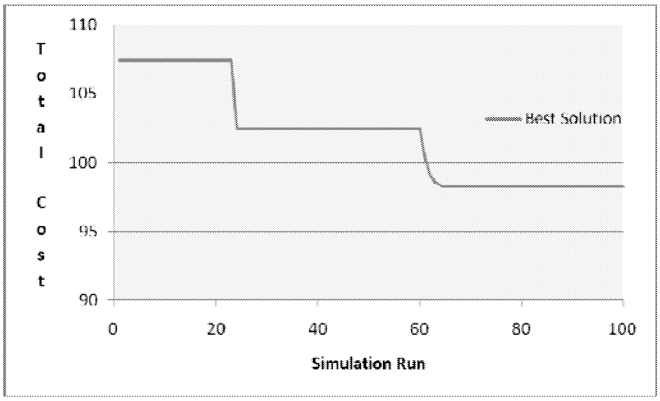


Figure 3: Optimization Results Obtained from OptQuest for Example 2

For the best solutions of each example we run simulation models for 20 replications where each replication takes 10000 hours to obtain average of H, Π and D separately. The results are demonstrated in Table 5.

Table 5: The Best Solutions of the Examples

	Example 1	**Example 2**
S_0	2	1
α	0.024196	0.220329
H	0.075890	0.047972
Π	0.072087	6.620500
D	0.081683	0.760200
Total Cost	52.358300	98.159950
Service Level	0.017353	0.049595

Note that there is no limitation in lead time or any other parameters in our simulation model which could be a great advantage of simulation approach. In addition, we can take advantage of simulation modeling to define service level constraint easily while it is extremely difficult to formulate this constraint mathematically. Furthermore, finding optimal or near optimal solutions for the stochastic constrained problem is another serious difficulty. While we can use simulation optimization as a decision support tool to obtain optimal or near optimal value of decision variables (S_0 and α).

CONCLUSION

Although, in most of real world cases inventory models' demands and lead times are stochastic, in the literature of deteriorating items inventory models there are a few research efforts in which these two parameters assume stochastic. In this paper a constrained inventory model of deteriorating items is presented. Since this model has three stochastic parameters (life time of products, demands and lead times), it is extremely difficult to build the mathematical model of the problem and prove the convexity of the model. Furthermore, the service level constraint is complicated to be defined since we should find the probability density function of $(1+\alpha)e^{-\theta S}$ in which s has Erlang distribution with parameters (λ,S_0). Thus, simulation modeling is applied for this problem. Firstly, it is shown that our simulation model is valid. To prove simulation model's validation we use relative error to show there are not significant differences between analytical and simulation models' output results when lead times are deterministic. After validating the simulation model we solve two different problems with stochastic lead times by the means of Arena optimizer module, OptQuest to find near optimal solutions. Although analytical model is too difficult to be established for stochastic lead time especially when the density function of lead time is complex, there is no limitation in our simulation model. In fact, difficult problems could be analyzed by this simulation model and simulation optimization can be used as a decision support tool to find optimal or near optimal solutions.

REFERENCES

April, J.; F. Glover; J. P. Kelly; and M. Laguna. 2003. "Practical introduction to simulation optimization." In *Proceeding of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin and D. J. Morrice (Eds). Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc., 71–78.

Axsäter, S. 1990. "Simple solution procedure for a class of two-echelon inventory problems." *Operations Research* 38:64-69.

Chang, C. T.; J. T. Teng; and S. K. Goyal. 2009. "Optimal replenishment policies for non-instantaneous deteriorating items with stock- dependent demand." *International Journal of Production Economics* 123:62–68.

Chern, M. S.; H. L. Yang; J. T. Teng; and S. Papachristos. 2008. "Partial backlogging inventory lot-size models for deteriorating items with fluctuating demand under inflation." *European Journal of Operational Research* 191:127–141.

Dye, C. Y.; T. P. Hsieh; and L. Y. Ouyang. 2007. "Determining optimal selling price and lot size with a varying rate of deterioration and exponential partial backlogging." *European Journal of Operational Research* 181:668–678.

Dye, C. Y.; L. Y. Ouyang; and T. P. Hsieh. 2007. "Inventory and pricing strategies for deteriorating items with shortages: A discounted cash flow approach." *Computers & Industrial Engineering* 52:29–40.

Dye, C. Y.; L. Y. Ouyang; and T. P. Hsieh. 2007. "Deterministic inventory model for deteriorating items with capacity constraint and time-proportional backlogging rate." *European Journal of Operational Research* 178:789-807

Goyal, S. K. and B. C. Giri. 2001. "Recent trends in modeling of deteriorating inventory." *European Journal of Operational Research* 134:1–16.

Goyal, S. K., and B. C. Giri. 2003. "The production-inventory problem of a product with time varying demand, production and deterioration rates." *European Journal of Operational Research* 147:549–557.

Hsieh, T. P.; C. Y. Dye; and L. Y. Ouyang. 2008. "Determining optimal lot size for a two-warehouse system with deterioration and shortages using net present value." *European Journal of Operational Research* 191:182–192.

Kalpakam, S. and K. P. Sapna. 1994. "Continuous review (s,S) inventory system with random lifetimes and positive leadtimes." *Operations Research Letters* 16:115–119.

Li, J.; T. C. E. Teng; and S. Wang. 2007. "Analysis of postponement strategy for perishable items by EOQ-based models." *International Journal of Production Economics* 107:31–38.

Lodree Jr.; J. Emmet; and B. M. Uzochukwu. 2008. "Production planning for a deteriorating item with stochastic demand and consumer choice." *International Journal of Production Economics* 116:219–232.

Maiti, A. K.; M. K. Maiti; and M. Maiti. 2009. "Inventory model with stochastic lead-time and price dependent demand incorporating advance payment." *Applied Mathematical Modelling* 33:2433-2443.

Nahmias, S. 2001. "Perishable inventory theory: A review." *Operations Research* 30:680–708.

Papachristos, S. and K. Skouri. 2000. "An optimal replenishment policy for deteriorating items with time-varying demand and partial-exponential type-backlogging." *Operations Research Letters* 27:175–184.

Raafat, F. 1991. "Survey of literature on continuously deteriorating inventory models." *Journal of the Operational Research Society* 42:27–37.

Sana, S.; S. K. Goyal; and K. S. Chaudhuri. 2004. "A production-inventory model for a deteriorating item with trended demand and shortages." *European Journal of Operational Research* 157:357–371.

Sivakumar, B. 2009. "A perishable Inventory system with retrial demands and a finite population." *Journal of Computational and Applied Mathematics* 224: 29-38.

Teng, J. T. and C. T. Chang. 2005. "Economic production quantity models for deteriorating items with price-and stock-dependent demand." *Computers & Operations Research* 32:297–308.

Wu, K.S.; L.Y. Ouyang; and C.T. Yang. 2006. "An optimal replenishment policy for non-instantaneous deteriorating items with stock-dependent demand and partial backlogging." *International Journal of Production Economics* 101:369-384.

AUTHOR LISTING

AUTHOR LISTING